The Disclosure Channel of Knowledge Spillovers: Evidence from Environmental Innovation*†

Leo Liu, Elvira Sojli, Wing Wah Tham, Cara Vansteenkiste

Abstract

Do disclosure decisions affect knowledge spillovers independently of R&D investment? We reinterpret patents not as innovative output but as *disclosed* innovation: a firm retaining trade secrets contributes nothing to the public knowledge stock regardless of R&D investment. We show this distinction matters fundamentally for process-based environmental technologies, which cannot be reverse-engineered, making patent disclosure the primary spillover channel. Clean technology disclosure generates nearly $11 in peer firm value for every $1 captured by rivals and reduces facility-level emissions by 5%. Traditional innovation shows no disclosure effects beyond R&D investment. Current climate policy addresses underinvestment but ignores under-disclosure, leaving substantial social value uncaptured.

Keywords: Knowledge Spillovers; Process Innovation; Reverse-engineering; Environmental Innovation Disclosure; Societal Economic Returns; Private Returns; Emissions; Environmental Technology

*We thank Ilona Babenko, Salome Baslandze, Jan Bena, Catherine Casamatta, Eugenie Dugoua, Nickolay Gantchev, Janet Gao, Xavier Giroud, Elise Gourier, Swarnodeep Homroy, Po-Hsuan Hsu, Kai Li, Stefano Lovo, Ralf Martin, Pedro Matos, Lilian Ng, Gordon Phillips, Rik Sen, Chander Shekhar, Jan Starmans, Noah Stoffman, Johan Sulaeman, Alison Taylor, Margarita Tsoutsoura, Hannes Wagner, Mindy Xiaolan, Alminas Zaldokas, Le Zhang, Yifei Zhang, and participants at the 2025 AEA, AFRIC, GRASFI Asia, CEIBS Finance & Accounting Symposium, GEIRC, 2024 PRI Academic Network Conference, SGFIN Sustainability Research Conference, Luxembourg Conference in Sustainable Finance, HEC-HKUST Sustainable Finance Webinar, NFA, Corporate Finance Days, FIRN Corporate Finance Meeting, FIRN Annual Meeting, NTHU-UNSW Sustainability Symposium and seminars at University of Sydney, University of Technology Sydney, and UNSW Sydney for very helpful suggestions.

†Liu is at University Technology Sydney, Sojli and Tham are at UNSW Sydney, and Vansteenkiste is at University of Sydney. Contact: leo.liu@uts.edu.au (Liu), e.sojli@unsw.edu.au (Sojli), w.tham@unsw.edu.au (Tham), cara.vansteenkiste@sydney.edu.au (Vansteenkiste).

1 Introduction

Knowledge spillovers are a central source of social returns to innovation. A large literature beginning with Arrow (1962), Griliches (1979), and Bloom, Schankerman, and Van Reenen (2013) documents that firms benefit from other firms' R&D investments, generating a wedge between social and private returns that motivates innovation policy. Yet this literature treats spillovers as arising primarily through reverse engineering, labor mobility, and informal channels, all of which operate regardless of whether innovating firms formally disclose their knowledge. Whether disclosure decisions themselves independently shape knowledge diffusion has received little attention, despite having direct implications for the design of innovation policy.

We argue that disclosure matters fundamentally for a large and economically important class of innovations: process-based technologies embedded inside production facilities. When a pharmaceutical firm develops a new drug, competitors can reverse-engineer the compound from the product itself once it reaches the market. Disclosure through patenting merely accelerates knowledge transfer that would occur anyway. Process innovations are different. When a chemical firm develops a novel emissions capture method, the only observable output is a reduction in emissions. Competitors cannot recover the underlying technology from observing outcomes alone. Without explicit disclosure, the knowledge remains locked inside the innovating firm. For this class of innovations, patent disclosure is not merely informative at the margin but it is the primary channel through which knowledge spillovers materialize.

Environmental innovation provides an ideal empirical laboratory to test this mechanism, and not merely a convenient one. Unlike most process innovations where competitors can at least partially recover technical knowledge by observing product outcomes, environmental innovation exhibits dual unobservability: both the underlying process and its primary outcome, an emissions reduction, are hidden from market observation. A steel manufacturer developing a faster production method leaves observable outputs that competitors can study; a chemical firm developing a novel emissions capture method leaves only a reduction in reported pollution, which reveals nothing about the technology that achieved it. This makes environmental innovation the strongest possible test of whether disclosure independently determines whether knowledge spillovers materialize. Environmental technologies are disproportionately process-based: 60% of environmental patents in our sample are process-based compared to 43% of traditional patents. The unobservability of both the technology and its outcomes, given that a firm can reduce emissions without revealing how, means that disclosure through patenting is especially consequential for knowledge diffusion. We use this setting to identify a disclosure channel in knowledge spillovers that is independent of,

and incremental to, the R&D investment channel documented by Bloom, Schankerman, and Van Reenen (2013) (BSV hereafter).

We reinterpret patents not as a measure of innovative output but as a measure of *disclosed* innovation. These are not the same thing. Innovation that is not disclosed generates no knowledge spillovers regardless of its quality or quantity. Disclosure without innovation generates no knowledge to transfer. Patents uniquely capture both simultaneously: to be granted a patent, a firm must publicly reveal the technical details of its innovation in sufficient depth that a skilled practitioner could replicate it, in exchange for a temporary monopoly on the technology.[1] This makes patents fundamentally different from R&D expenditure as a measure of innovation: R&D stock captures how much a firm invests regardless of whether the resulting knowledge ever leaves the firm, while patents capture both the decision to make knowledge public and the knowledge itself. A firm that spends heavily on R&D but retains its innovations as trade secrets contributes nothing to the public knowledge stock. A firm that patents contributes directly and verifiably. This distinction is the foundation of our empirical strategy.

Our empirical framework builds directly on BSV, who decompose innovation spillovers into technology spillovers from technologically related firms and competitive effects from product market rivals, using R&D expenditure as the measure of innovation. We depart from BSV by introducing a parallel disclosure-based spillover measure constructed from quality-weighted environmental patents. Conditional on the full set of R&D-based spillovers identified by BSV, our patent-based spillover coefficients isolate the disclosure wedge: the incremental effect of making innovations publicly available through patenting, above and beyond what R&D investment alone generates. If patents were merely a noisy proxy for R&D, disclosure quality would be irrelevant and the clean and traditional innovation disclosure channels would be symmetric. Neither is true in the data.

The central identification challenge is separating disclosure decisions from underlying innovation investment. A firm that patents more may also be better positioned to absorb spillovers for reasons unrelated to disclosure, for instance because it operates in a more productive technology space or faces greater competitive pressure to innovate. We address this in three ways. First, controlling for BSV's R&D spillovers ensures variation in patent stocks reflects disclosure choices rather than innovation input. Second, we exploit the intensive margin of disclosure through patent specificity, fullness, and readability (Dyer et al., 2024):

[1] "When a patent is granted and the information contained in it is circulated to the general public and those especially skilled in the trade, such additions to the general store of knowledge are of such importance to the public weal that the Federal Government is willing to pay the high price of 17 years of exclusive use for its disclosure, which disclosure, it is assumed, will stimulate ideas and the eventual development of further significant advances in the art." US Supreme Court, *Kewanee Oil Co. v. Bicron Corp.* (1974; No. 73-187).

if spillovers operate through investment rather than knowledge dissemination, disclosure quality should be irrelevant. Third, we instrument disclosure using EPA enforcement actions against technology peers and product market rivals. Regulatory pressure increases firms' incentives to patent clean technologies as a compliance signal (Berrone et al., 2013; He and Qiu, 2025), generating exogenous variation in the stock of publicly available environmental knowledge. Critically, enforcement actions target specific facilities for specific violations, and we validate that they affect clean patent filings but not non-clean patent filings, that they accelerate disclosure timing at the patent level, and that downstream emissions fall for firms that cite EPA-induced patents, all of which is consistent with the instrument operating through a disclosure channel rather than through correlated productivity shocks.

Our baseline estimates reveal a stark contrast between environmental and traditional innovation. For traditional innovation, patent disclosure generates no incremental spillovers beyond R&D investment, consistent with knowledge diffusing through reverse engineering regardless of formal disclosure. For environmental innovation, a 10% increase in quality-weighted clean knowledge from technology peers is associated with a 1.94% increase in firm value. This effect is entirely driven by process-based clean patents; product-based clean patents generate no incremental spillovers, directly confirming the unobservability mechanism. Competitive effects from rivals' disclosure are economically negligible, generating a disclosure wedge, defined as the gap between spillover benefits to technology peers and competitive costs to disclosing firms, of approximately 1.8%. These value effects translate to social returns beyond firm value: a one standard deviation increase in clean acquired knowledge is associated with a 5% reduction in facility-level toxic emissions over five years, while equivalent non-clean knowledge acquisition is associated with emissions increases, consistent with traditional innovation enabling production expansion rather than pollution abatement. That clean knowledge spillovers generate measurable economic and environmental returns jointly underscores that the disclosure wedge has consequences extending well beyond what firm value estimates alone can capture.

To address potential endogeneity, we instrument disclosure-driven knowledge spillovers with EPA enforcement actions in technology and product market peers. EPA actions increase firm-specific regulatory risk and prompt affected firms to publicly register clean technologies as a reputational and compliance signal (Berrone et al., 2013; He and Qiu, 2025), but do not directly affect the knowledge spillovers generated by these technologies beyond a disclosure channel. We extensively validate the instrument's relevance and exclusion restriction: EPA enforcement actions increase clean patent filings but leave non-clean patent filings unchanged, ruling out the interpretation that regulatory pressure broadly stimulates innovation activity; affected firms accelerate the *timing* of clean technology disclosure fol-

lowing enforcement; clean *patenting* generates measurable *signaling* benefits for regulatory outcomes; and downstream *emissions* fall for firms that cite EPA-induced clean patents. Our IV estimates confirm and strengthen our baseline findings. A 10% increase in instrumented clean knowledge from technologically related firms raises focal firm value by 3.9%, nearly double the baseline estimate. Notably, competitive effects from rivals' instrumented patent disclosure are economically small and positive (0.35%), indicating that learning benefits from accessing technical knowledge revealed following regulatory pressure outweigh proprietary information costs. In contrast, traditional innovation disclosure generates no incremental spillovers beyond R&D investment. The disclosure wedge of 3.5% is driven primarily by process-based clean patents, highlighting the unobservability mechanism inherent to clean innovation through which disclosure triggers knowledge spillovers.

These findings reveal a disclosure-specific externality distinct from the R&D underinvestment problem at the center of the innovation policy literature. Firms that innovate in process-based environmental technologies generate large knowledge spillovers when they disclose, but cannot internalize these benefits, leading to under-disclosure independent of under-investment. The social return to disclosure, measured through peer firm value and downstream emissions reductions, substantially exceeds the private return: clean innovation disclosure generates nearly $11 in peer firm value for every $1 captured by rivals, a ratio that reflects the scale of under-disclosure relative to the social optimum. This gap is not addressed by existing policy instruments. More broadly, this disclosure channel is likely relevant for any innovation whose process or outcome is unobservable from the market, a feature that extends well beyond environmental technology.

Our paper contributes to several literatures. First, we extend research on knowledge spillovers and innovation by documenting a distinct disclosure-based spillover channel that complements the R&D investment channel in BSV.[2] The prior literature treats disclosure and investment as inseparable: more R&D produces more patents, and both generate spillovers through the same informal channels. We show these are distinct margins with different consequences. For traditional innovation, the two are indeed inseparable: patents add nothing once R&D investment is controlled for. For environmental innovation, disclosure is the operative channel: R&D spillovers are fully absorbed by patent-based spillovers once both are included, indicating that clean knowledge reaches other firms through explicit documentation rather than informal diffusion. This result is consistent with environmental technologies being disproportionately process-based (Bena, Ortiz-Molina, and Simintzi, 2022;

[2] A large body of work studies how these spillovers propagate through public R&D (Matray, 2022), local inventor networks (Moretti, 2021), and firms' plant-level innovation networks (Giroud, Liu, and Mueller, 2024).

Tham et al., 2025): embedded inside production facilities, they cannot be reverse-engineered from market outputs, making explicit patent documentation the binding constraint on knowledge diffusion. This distinction matters for policy: the spillover literature has focused almost entirely on correcting under-investment, but our results suggest that for process-based innovations, correcting under-disclosure is the binding constraint.

Second, we contribute to the literature on sustainable finance and corporate environmental action. We show that disclosure is a real corporate environmental action with climate consequences. Prior work examines how capital markets shape the green transition: carbon pricing and asset valuation (Bolton and Kacperczyk, 2021; Hong, Karolyi, and Scheinkman, 2020), green investments and firm value (Flammer, 2021; Garel et al., 2024), climate risk disclosure (Sautner et al., 2023; Giglio et al., 2023), and investor sustainability preferences (Ilhan et al., 2023; Krueger, Sautner, and Starks, 2020; Heeb et al., 2023). We demonstrate that disclosure choices, shaped by competitive forces in the market, also affect which environmental technologies become publicly accessible and can be adopted by other firms. In doing so, we highlight disclosure as an overlooked but economically consequential margin that shapes the diffusion of environmental technologies and ultimately the pace of the green transition.

2 Empirical Framework and Estimation Strategy

In this section, we outline our empirical framework for estimating the private and social economic returns of environmental innovation disclosure. Although non-economic environmental outcomes can be estimated by evaluating toxic emissions, estimating aggregate economic returns is less straightforward. Our approach builds on Griliches (1979), which assumes knowledge capital is a key driver of firm performance, and adapts the empirical strategy of Bloom, Schankerman, and Van Reenen (2013, BSV hereafter), which decomposes innovation spillovers into technology spillovers (knowledge externalities) and competitive (product market rivalry) effects. However, our focus differs fundamentally from BSV: whereas BSV examines spillovers from R&D investment (measured through R&D expenditure), we examine spillovers from disclosure decisions, i.e. the choice to make innovations publicly available through patenting versus maintaining trade secrecy, and the informativeness of that disclosure. This distinction is critical in the context of environmental innovation, for which social knowledge spillover benefits depend not only on whether firms innovate, but also on whether they choose to disclose that innovation and how much actionable knowledge they reveal.

2.1 Production Function with Disclosed Knowledge

Our econometric specification incorporates the Griliches (1979) knowledge capital approach into a standard Cobb-Douglas production function, augmenting the traditional R&D-based measures with disclosed knowledge capital. Specifically, we measure knowledge capital through both R&D investment (as in BSV) and publicly disclosed innovations (patents), allowing us to separately identify spillovers from innovation investment versus disclosure decisions. This allows us to formalize the role of disclosed clean knowledge spillovers in shaping firm output:

$$Y_i = K_i^{\alpha} \left(L_i \cdot A_i^{\text{total}} \right)^{1-\alpha}, \tag{1}$$

where Y_i represents firm i's output, K_i is capital, L_i is labor, and A_i^{total} denotes the total disclosed clean knowledge stock available to the firm. Critically, we assume that A_i^{total} consists of disclosed knowledge from internally generated innovation and spillovers from external sources' public disclosures:

$$A_i^{\text{total}} = A_i + \sum_{j \neq i} \text{PROXIM}_{ij} A_j.$$

Empirically, we capture A_i through firm i's own clean innovation stock (patents), and A_j through the disclosed clean knowledge stock of other firms j. PROXIM_{ij} is the proximity between firm i and firm j in the technology and product market space, serving as a weighting factor for spillovers. This formulation explicitly recognizes that clean knowledge spillovers occur through accessibility of knowledge: a firm that conducts environmental R&D but maintains trade secrecy contributes zero to the public knowledge stock, generating no spillovers regardless of innovation quality. By including both R&D-based measures (following BSV) and patent-based measures in our empirical specification, we can test whether disclosure decisions matter for spillovers independently of innovation investment.

2.1.1 From Production Function to Market Value Estimation

We assume that firm i's output is proportional to its market performance, captured through market value. Specifically, market value MV_i reflects the present discounted value of future output, assuming Y_i grows at a constant rate equal to the discount rate r such that $MV_i = \frac{Y_i}{r}$. In our empirical specification, we therefore substitute Y_i for MV_i.

Our market value specification captures the portion of knowledge spillovers that is appropriated by publicly listed firms through changes in their market valuation. This represents a lower bound on total social benefits. First, we observe only spillovers to other publicly traded

firms, excluding spillovers to private firms, startups, and non-profit research institutions. Second, market value reflects only the financial returns captured by firms, not broader social benefits including consumer surplus from improved products, health benefits from pollution reduction, and other climate benefits that are not fully capitalized into any firm's market value. We complement our market value analysis with measures of direct environmental outcomes (facility-level emissions) to capture these non-market social benefits. Together, these analyses provide evidence on both the appropriable economic returns to other firms and the broader environmental externalities that motivate disclosure policy.

2.2 Empirical Specification

Our empirical strategy builds on BSV by measuring spillovers through disclosed patent stocks in addition to R&D investment. Econometrically, we log-linearize Equation (1) and separate knowledge spillovers into two components: technology spillovers from technologically related firms and competitive spillovers from product market rivals. Our baseline specification follows the following structure:

$$\begin{aligned}\ln\left(\frac{MV}{TA}\right)_{i,t+1} = \beta_0 &+ \beta_1 \underbrace{\sum_{j\neq i} \text{TECH}_{ij} \ln(G^q_{j,t})}_{\ln(\text{SPILLTECH}^S_{Clean})_{i,t}} + \beta_2 \underbrace{\sum_{\substack{j\neq i \\ j\in\text{prod. rivals}}} \text{PRODSIM}_{ij} \ln(G^q_{j,t})}_{\ln(\text{SPILLTECH}^C_{Clean})_{i,t}} \\ &+ \beta_3 \underbrace{\sum_{j\neq i} \text{TECH}_{ij} \ln(R\&DStock)_{j,t}}_{\ln(\text{SPILLTECH}^S_{R\&D})_{i,t}} + \beta_4 \underbrace{\sum_{\substack{j\neq i \\ j\in\text{prod. rivals}}} \text{PRODSIM}_{ij} \ln(R\&DStock)_{j,t}}_{\ln(\text{SPILLTECH}^C_{R\&D})_{i,t}} \\ &+ \beta_5 \ln(G)_{i,t} + \beta_6 \ln\left(\frac{R\&DStock}{TA}\right)_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned} \tag{2}$$

In Equation (2), $SPILLTECH^S_{Clean}$ is the aggregate disclosed clean knowledge capital ($G^q_{j,t}$) of all firms j , weighted by their technological proximity to firm i (TECH_{ij}). Firms that are technologically closer to firm i therefore get a higher weight, as firm i may benefit more from their knowledge diffusion. $SPILLTECH^C_{Clean}$ is the aggregate disclosed clean knowledge capital of all direct product market competitors j, weighted by their product market similarity to firm i (PRODSIM_{ij}).

Critically, our specification controls for firm i's own clean patent stock (G) as well as the complete set of R&D-based spillovers identified by BSV, including (i) firm i's own R&D stock ($R\&DStock/TA$), (ii) technology spillovers from technology peers' R&D investments ($\text{SPILLTECH}^S_{R\&D}$), and (iii) product market rivalry effects from competitors'

R&D investments ($\text{SPILLTECH}^C_{R\&D}$). We thereby ensure that our patent-based spillover coefficients (β_1, β_2) capture disclosure effects beyond the innovation investment channel that BSV identify. That is, conditional on all firms' R&D investments and the spillovers they generate through BSV's mechanism, our coefficients measure the incremental effect of making innovations publicly available through patent disclosure and the informativeness of that disclosure.

2.2.1 Construction of SPILLTECH Measures

To calculate $SPILLTECH^S_{Clean}$ and $SPILLTECH^C_{Clean}$, we measure disclosed clean knowledge capital ($G^q_{j,t}$) through the quality-weighted stock of clean patents filed by firms j. We classify patents as clean if at least 20% of a patent's IPC classifications are in the OECD ENV-TECH classification criteria for Environmental Management technologies (pollution abatement). Clean patents explicitly codify technical environmental knowledge through detailed public disclosure as a condition for patent protection. Unlike R&D expenditures, which capture innovation investment regardless of disclosure decisions, patents capture both the disclosure decision and the disclosed knowledge itself, making them a direct empirical proxy for innovation diffusion and knowledge spillovers accessible to other firms.

Moreover, patents enable us to capture the intensive margin of disclosure through the quality of information in patent filings, which we use as a weighting factor in firms' patent stocks. Following Dyer et al. (2024), we measure disclosure quality through three patent characteristics: specificity (the number of specific technical parameters such as percentages, quantities, and measurements), fullness (the volume of technical information disclosed), and readability (the clarity and accessibility of information). Because we are primarily interested in knowledge spillovers to technology peers and product market rivals, rather than the general public or non-specialists, we use specificity in our baseline estimates and show that our results remain virtually unchanged when using patent fullness and readability. We weight each patent in firm j's clean patent stock $G_{j,t}$ by the patent's quantile-transformed IPC3-year-adjusted quality (ranging from 0 to 1) when calculating SPILLTECH^S and SPILLTECH^C.

This quality-weighting is critical for isolating disclosure effects from innovation output effects. Two firms may file the same number of patents representing similar innovations, but differ substantially in how much actionable knowledge they reveal through the quality of disclosed information. If spillovers operate purely through innovation investment, disclosure quality should not matter. A patent's knowledge diffusion should then depend only on the existence of the underlying technology, not on the informativeness of disclosed technical information. In contrast, if spillovers operate through knowledge dissemination (the transfer of actionable information via patent documentation), more informative patents should

generate larger spillovers holding patent counts constant. This prediction uniquely identifies the disclosure channel and provides a sharp test of our mechanism.[3]

Variation in $G^q_{j,t}$ therefore reflects differences in disclosure decisions (whether to patent versus maintain trade secrecy) and disclosure informativeness (how much actionable knowledge is revealed in patent documentation). Because our empirical specification conditions on own R&D stock and all R&D-based spillovers, we thereby ensure patents capture disclosure decisions rather than R&D quality (i.e., more productive research generating both more spending and more patents).

To calculate $SPILLTECH^S_{R\&D}$ and $SPILLTECH^C_{R\&D}$, we following BSV and measure innovation investment through cumulative R&D stock constructed using the perpetual inventory method with a 15% depreciation rate. For firm i at time t, the R&D stock is:

$$R\&DStock_{i,t} = (1-\delta)R\&DStock_{i,t-1} + R\&D_{i,t},$$

where $R\&D_{i,t}$ is firm i's R&D expenditure in year t and $\delta = 0.15$ is the depreciation rate. We use total R&D stock rather than separating clean and traditional R&D because firms do not report R&D expenditures by technology type.

In all $SPILLTECH$ measures, technological proximity between firms i and j (TECH_{ij}) is based on technology (IPC) class overlap in firms' patent portfolios, following Jaffe (1986). For each firm i, we construct a patent portfolio vector $T_i = (T_{i1}, T_{i2}, \ldots, T_{iN})$ where T_{ik} is the share of firm i's ENV-TECH patents in 3-digit IPC class k, calculated over the period 1988-2020. Technological proximity is then calculated as the uncentered correlation:

$$TECH_{ij} = \frac{T_i \cdot T_j'}{\sqrt{(T_i \cdot T_i')(T_j \cdot T_j')}},$$

which ranges from 0 (no technological overlap) to 1 (identical technology portfolios).

Product similarity between firms i and j (PRODSIM_{ij}) is based on text-based analysis of firms' product descriptions in their 10-K filings, following Hoberg and Phillips (2016). This measure provides more granular and objective competitive overlap than traditional industry classifications.

2.2.2 Other Controls and Fixed Effects

$G_{i,t}$ is defined as firm i's own clean patent stock to capture direct value effects from firms' own clean knowledge capital, and $R\&DStock/TA$ captures firm i's own innovation investment,

[3]Dyer et al. (2024) show that higher quality patent disclosure leads to more follow-on innovation, supporting the view that disclosure informativeness facilitates knowledge transfer.

scaled by total assets. We assume that r is negligible and constant across all firms and time, and is therefore captured in β_0. $X_{i,t}$ is a vector of firm-specific characteristics that may affect firm i's market value, including: log total sales, cash holdings, capital expenditures (CAPX), property plant and equipment (PPE), and industry competitiveness (proxied by the Herfindahl-Hirschman Index, HHI). λ_i are firm fixed effects, which capture firm-invariant unobserved heterogeneity, λ_t are time (year) fixed effects, which capture time-invariant unobserved heterogeneity, and $\epsilon_{i,t+1}$ is the error term. The combination of fixed effects and control variables should capture the elasticity of firm i's market value with respect to the other inputs in the production function (i.e., capital and labor).

2.2.3 Interpretation of Coefficients

Technology Spillovers (β_1). In Equation (2), β_1 quantifies the portion of clean knowledge spillovers that firm i appropriates in its market valuation. If spillovers operate purely through R&D investment (as in BSV), then conditional on R&D-based spillovers, patent-based spillovers should have no additional effect ($\beta_1 = 0$). If disclosure independently affects knowledge diffusion, we should observe $\beta_1 > 0$ even after controlling for all R&D spillovers. A positive β_1 indicates that an increase in environmental technology disclosure by technologically related firms increases firm i's market value, reflecting firm i's ability to generate follow-on innovation, adopt disclosed process improvements, or otherwise benefit from freely available technical knowledge to reduce emissions and regulatory costs.

Importantly, β_1 measures only the financial spillover benefits captured by other publicly traded firms, and does not measure other spillover benefits captured by non-public firms, consumer surplus from improved or cheaper products, or direct and indirect environmental and health benefits. It is therefore a conservative lower bound for aggregate social returns from clean innovation disclosure. Our complementary analysis using facility-level emissions data provides direct evidence on environmental benefits, which we combine with market value effects to characterize the full social return.

Competitive Effects (β_2). β_2 captures the effect of product market rivals' clean disclosure on firm value, reflecting the private costs and benefits that firms internalize when deciding whether to patent. These private returns reflect firms' profit-maximizing patenting decisions and the trade-off between expected proprietary information costs (e.g., the loss of intellectual property control from imitation by competitors, reduced bargaining power in licensing negotiations, or increased litigation risk) and private disclosure benefits (e.g., licensing opportunities and stronger legal protection) (Anton and Yao, 1994, 2004; Verrecchia, 1983).

In Equation (2), a loss to disclosing (patenting) firms reflects a gain to non-disclosing

(non-patenting) firms. Thus, proprietary information costs imply an average positive effect of rivals' patenting on firm i's valuation ($\beta_2 > 0$), whereas private benefits imply an average negative effect ($\beta_2 < 0$). β_2's sign therefore depends on whether firm j's proprietary information costs outweigh legal protection benefits. If clean process-based innovations are difficult to imitate in the absence of formal patent disclosure, learning benefits from accessing codified technical knowledge may disclosing firms' private benefits ($\beta_2 > 0$).

The Disclosure Wedge ($\beta_1 - \beta_2$). The disclosure wedge ($\beta_1 - \beta_2$) quantifies the difference in valuation impact between technology spillovers and competitive effects. When firm j decides whether to disclose an innovation, it internalizes how the disclosure affects rival firms (β_2). However, it does not internalize the social spillover benefits generated for technologically related firms (β_1). A positive wedge ($\beta_1 - \beta_2 > 0$) therefore indicates that society gains more from disclosure than firms capture privately, leading to under-disclosure, with a large wedge providing support for disclosure-specific policy interventions to create social welfare gains.[4]

R&D-Based Spillovers (β_3 and β_4). β_3 and β_4 capture spillover effects from R&D investment as identified in BSV. β_3 measures technology spillovers from technology peers' R&D investments ($\beta_3 > 0$), while β_4 measures competitive effects from rivals' R&D investments ($\beta_4 < 0$). These coefficients control for innovation investment, allowing us to isolate the incremental effect of disclosure decisions in the wedge between β_1 and β_2.

Clean Patent and R&D Stock (β_5 and β_6). β_5 captures the elasticity of firm i's market value to its own clean patent stock ($G_{i,t}$), quantifying the direct value of IP protection. β_6 captures returns to own innovation investment ($R\&DStock$), scaled by total assets, allowing us to distinguish the impact of innovation investment from disclosure.

2.3 Comparing Clean and Traditional Innovation

Our core analysis quantifies the disclosure wedge for environmental innovation. However, we also want to compare it to the wedge for traditional innovation. As a final step, we augment Equation (2) by also including traditional (non-clean) spillovers, where non-clean spillovers are defined analogously to Equation (2) for the subset of non-clean patents. We classify all patents that are not defined as clean (i.e., 20% of a patent's IPC classifications

[4]For firms to fully internalize spillover benefits to other firms, their private competitive effects need to equal peer spillover effects ($\beta_1 = \beta_2$). This could happen through stronger licensing markets, enhanced IP protection, or government policies that reward disclosure (e.g., disclosure-contingent subsidies, regulatory approval tied to public registration of clean technologies, patent box tax incentives for disclosed innovations).

in ENV-TECH) as non-clean innovation. Our baseline specification therefore becomes:

$$\begin{aligned}\ln\left(\frac{MV}{TA}\right)_{i,t+1} &= \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_1^{NC} \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_2^C \ln(\text{SPILLTECH}^S_{Non-Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non-Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(G_{i,t}) + \beta_6 \ln\left(\frac{R\&DStock}{TA}\right)_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned} \tag{3}$$

Equation (3) allows us to directly test whether disclosure decisions matter for clean knowledge spillovers independently of R&D investment ($\beta_1^C - \beta_2^C$), and whether the disclosure wedge differs between environmental and traditional innovation $[(\beta_1^C - \beta_1^{NC}) > (\beta_2^C - \beta_2^{NC})]$. We expect that environmental innovation disclosure generates larger social spillover benefits because environmental technologies, which are more often process-based, require explicit disclosure, whereas traditional technologies can often be reverse-engineered regardless of disclosure strategy. We therefore predict $\beta_1^C > \beta_1^{NC}$.

Traditional patents may however provide larger competitive benefits (i.e., costs to product market rivals) because firms' internalization of competitive effects implies they choose to disclose only when the benefits of doing so outweigh the cost. In contrast, environmental patents may exhibit smaller competitive effects if knowledge spillover effects offset disclosing firms' private benefits or if environmental innovation carries less competitive and strategic value for product market rivals. We therefore predict $\beta_2^C > \beta_2^{NC}$.

We then hypothesize that $(\beta_1^C - \beta_1^{NC}) > (\beta_2^C - \beta_2^{NC})$, i.e. the disclosure wedge is fundamentally larger for environmental innovation than for traditional innovation because environmental technologies generate dual externalities (knowledge spillovers to other innovators and environmental benefits to society) that require public disclosure to be realized, whereas traditional technologies can be reverse-engineered regardless of disclosure strategy.

Our empirical analysis proceeds in three stages. First, we estimate Equation (3) using OLS to establish baseline patterns. Second, we decompose our results in process- and product-based patents to validate our mechanism in which formal patent disclosure matters specifically for technologies hidden inside production facilities that cannot be reverse-engineered from market observations. Third, we address potential endogeneity using EPA enforcement actions as instrumental variables for disclosure decisions.

3 Data and Summary Statistics

In this section, we provide an overview of the key data sources used to estimate our baseline specification in Equation (3) and the variables used in our main identification strategy based on EPA enforcement actions.

3.1 Patent Data

To measure firms' patent stock, we obtain data on U.S. firms' patent filings from Patentsview provided by the USPTO. Patentsview holds the full history of patents granted by the USPTO from 1976 and patent applications from 2001. We use patents on environment-related technologies (ENV-TECH) as defined by the OECD. The structured retrieval and identification methods for mapping patent documents to environment-related technologies have been developed by experienced patent examiners working in relevant fields in cooperation with external experts. These examiners are technical experts who retrieve pertinent patent information, perform prior art searches, and classify patents into Cooperative Patent Classification (CPC) and International Patent Classification (IPC) classifications based on their relevance as an environment-related technology.

The ENV-TECH classification features eight environmental families, separated into three areas: Environmental Management, Water-Related Adaptation Technologies, and Climate Change Mitigation Technologies (CCMT). We focus on patents in Environmental Management (i.e., pollution abatement technologies), as these are most closely aligned with regulatory requirements and EPA enforcement actions and, therefore, firms' private disclosure incentives.[5] To identify clean patents, we first identify all USPTO-granted patents from 1976 to 2020 with the ENV-TECH IPC classification. On average, USPTO patents in our sample have a mean (median) of 8.45 (6) IPC classifications. Patents with an ENV-TECH classification have a mean (median) of 12.6 (10) IPC classifications, indicating that these patents are more general technology-based. To ensure a patent has a substantial contribution to environmental technologies, we require at least 20% of its classifications to be in environmental management classes (i.e., 1 in 5). The 20% cutoff is arbitrary, but the results are qualitatively similar for higher cutoffs and the full sample.

[5]ENV-TECH classification is based on the European Patent Office's classification scheme for environment-related technologies, where relevant patent publications are classified into a separate scheme within the CPC, see `https://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y.html`.

3.2 Pollution and EPA Enforcement Action Data

We obtain facility-level data on toxic emissions from the Toxic Release Inventory (TRI) database, administered by the U.S. EPA. TRI data for each toxic chemical (level of pollutants and name of the chemical) are reported by individual industrial and federal facilities. The data covers a broad spectrum of industries, including manufacturing, mining, and electric utilities, and provides detailed information on more than 650 specified toxic chemicals. Facilities employing 10 or more full-time staff must submit annual reports detailing the quantities of each released TRI-listed chemical. We use data for total on-site releases for all facilities owned by publicly-listed U.S. firms for the sample period 1988 to 2020, as TRI data coverage starts in 1987.

In our IV analysis, we leverage within-firm and within-facility variation in environmental enforcement actions by the EPA to capture changes in firms' incentives to disclose clean innovation through patenting. We obtain facility-level data on formal (e.g., fines and civil lawsuits) and informal (e.g., violation notices) actions by the EPA and local environmental regulators reported through Enforcement and Compliance History Online (ECHO). Figure B1 reports the total number of EPA actions faced by U.S. public firms over time, showing an increasing trend in environment-related regulatory pressure.

Accounting and financial variables are from the CRSP/Compustat Merged (CCM) database using parent company names. All variable descriptions are provided in Table B1. We only retain firm-year observations with positive total assets. We link patents to U.S. publicly listed firms using the Permanent Stock Identifier (PERMNO) provided by Kogan et al. (2017). We identify 41,190 clean patents granted to publicly listed firms between 1988 and 2020. We report some of the most cited clean patents in Table B2 in the Appendix.

3.3 Knowledge Spillover Summary Statistics

In Panel A of Table 1, we provide summary statistics for the main firm-level variables used in Equation (3), for the full sample of U.S. publicly listed firms. The clean and non-clean knowledge stocks from disclosure by technology peers ($SPILLTECH^S$) are larger than that from disclosure by product market rivals ($SPILLTECH^C$), reflecting that more firms are technology peers than direct product market rivals. Similar patterns hold for R&D stock, where technology peers' R&D stock ($SPILLTECH^S_{R\&D}$) is larger than product market rivals' R&D stock ($SPILLTECH^C_{R\&D}$). Process-based clean knowledge stock from technology peers' disclosure is larger than the product-based clean knowledge stock, reflecting that clean innovation tends to be process-, rather than product-based.

The average firm in our sample has a mean (median) MV/TA ratio of 1.4 (0.79). Firms

in the sample invest on average 16% of their total assets in R&D. Firms' mean and median annual clean patent stock G is 3.45 and 0 patents, respectively, indicating that many firms do not disclose their technology or do not engage in R&D. Other firm-level variables align with the literature: firms achieve $2,677 million in annual sales, have an average cash ratio of 18%, CAPX ratio of 5%, PPE ratio of 24%, and average HHI is 0.23.

Panel B shows facility-level summary statistics for the sample used in our tests for facility-level emissions, based on the sample of patent-filing firms. Facilities emit, on average, 1,754 pounds ($=e^{7.47}$) in toxic emissions per year. *Clean Acquired Knowledge* and *Non-Clean Acquired Knowledge*, measured as the firm's number of backward citations of clean and non-clean patents, respectively, capture the effect of other firms' innovation disclosure on facility-level emissions. *Clean Acquired Knowledge* is lower than *Non-Clean Acquired Knowledge* given it is based on a smaller subset of patents.

4 Results

In this section, we present analyses estimating the disclosure wedge for environmental innovation and compare it to that for non-clean innovation. We then document the real environmental outcomes from clean knowledge spillovers by examining facility-level emissions, and address endogeneity using EPA enforcement actions as instruments for disclosure decisions and validate their relevance and exclusion assumptions.

4.1 Disclosure Wedge for Clean Knowledge Spillovers

We begin by quantifying the wedge between private (competitive) and social (technology) returns from disclosed clean knowledge spillovers following Equation (3). As detailed in Section 2, we separately estimate the aggregate economic impact from disclosed knowledge spillovers by technologically related firms (SPILLTECH^S) versus competitive effects from spillovers by product market rivals (SPILLTECH^C). We ensure our estimates isolate the impact of disclosure, independent from innovation input (R&D investment), by explicitly controlling for R&D-based spillovers through firms' own R&D stock, spillovers from technology peers' R&D, and product market effects from competitors' R&D, following BSV. Moreover, by weighting our spillover measures with the informativeness of disclosed information (i.e., patent filing disclosure quality), we separate the decision to patent from the patent's ability to generate knowledge spillovers.

4.1.1 Baseline Results: Clean Innovation Disclosure Wedge

Table 2 reports our baseline estimates. Column (1) shows that clean innovation disclosure generates large social returns through technologically peers' knowledge spillovers, incremental to spillovers from R&D investment. A 10% increase in quality-weighted clean knowledge from technology peers is associated with a 1.93% increase in firm i's market value ($\beta_1^C = 0.193$, $t = 2.46$). When firms publicly disclose clean innovation through patenting, other firms in the same technology space can benefit and build on this knowledge to develop improved products, adopt cleaner production processes, reduce regulatory compliance costs, and otherwise benefit from freely available technical information.

In contrast, spillovers from rivals' clean innovation disclosure generate negative and economically negligible incremental returns. In column (2), a 10% increase in rivals' clean knowledge corresponds to a 0.10% decrease in firm value ($\beta_2^C = -0.010$, $t = -1.84$), indicating limited knowledge spillover benefits and reflecting innovating firms' decision to disclose only when private benefits marginally exceed competitive costs.

We estimate the social and private returns of environmental innovation disclosure jointly in column (3), thereby quantifying the disclosure wedge of disclosure. Social returns from environmental innovation disclosure spillovers remain high at 1.94% ($\beta_1^C = 0.194$, $t = 2.48$), while its competitive effects remains economically small and marginally significant ($\beta_2^C = -0.010$, $t = -1.90$). These patterns reveal a large clean disclosure wedge of approximately 1.8%: clean innovation disclosure generates large social spillover benefits that firms cannot internalize, in contrast to competitive effects, which can be internalized and which incentivize firms to disclose only when private benefits exceed competitive costs.

This disclosure wedge indicates a potential misalignment between private disclosure incentives and social benefits, an issue that matters significantly more for environmental than traditional innovation. Social returns from traditional (non-clean) innovation disclosed by technology peers are not significantly different from zero ($\beta_1^{NC} = -0.03$, $t = 1.46$, column 1), reflecting the limited incremental social benefits from traditional innovation disclosure beyond R&D investment. The competitive effects of traditional innovation disclosure also remain economically negligible at an 0.08% decrease in value for a 10% increase in rivals' non-clean innovation disclosure ($\beta_2^{NC} = -0.008$, $t = 3.35$, column 1), resulting in an economically negligible disclosure wedge for traditional innovation.

The importance of disclosure as a distinct channel is reinforced by comparing disclosure-based and R&D-based spillover coefficients. Without accounting for clean disclosure spillovers, R&D investment by technology peers is associated with a 1.42% increase in social returns, confirming BSV ($\beta_3^{R\&D} = 0.142$, $t = 4.32$, column 2). However, jointly estimating R&D-based and disclosure-based spillovers in column (3) reveals that disclosure absorbs most

of the R&D spillover effect ($\beta_3^{R\&D} = 0.039$, $t = 0.80$). This pattern is consistent with our hypothesis: for environmental technologies, R&D investment generates social value primarily *through* disclosure. Without public disclosure, knowledge produced by clean R&D remains locked inside the innovating firm. Disclosure is the mechanism through which R&D spillovers materialize, rather than a competing explanation. Importantly, this absorption occurs only for clean innovation; traditional R&D spillovers are unaffected by the inclusion of traditional disclosure measures, consistent with traditional knowledge diffusing through reverse engineering regardless of formal disclosure decisions. In the next section, we investigate a potential explanatory mechanism: the unobservability of environmental outcomes (emissions) and technologies (process-based), which require explicit patent disclosure for knowledge transfers to materialize.

4.1.2 Mechanism: Process- versus Product-Based Patents

Our baseline results reveal that environmental and traditional innovation differ significantly in the social returns generated by disclosure. We next investigate the mechanism driving this difference. Clean innovation differs from traditional innovation because both the technologies, which generally involve cleaner production methods hidden inside production facilities - and the outcomes - emissions reductions that cannot be reverse-engineered - are unobservable to other market participants. We therefore expect the disclosure wedge to be particularly pronounced for clean innovations that reduce emissions by changing firms' production processes, where patent disclosure is the primary channel through which technical knowledge can diffuse.

Measuring process and product patents. We classify patents as process- or product-based following the hypernym-based methodology developed in Tham et al. (2025). The method exploits the linguistic structure of patent titles and claims to distinguish innovations describing activities (processes) from those describing physical entities (products). Specifically, an expanding set of activity hypernyms (e.g., "method," "treatment," "fabrication") and physical entity hypernyms (e.g., "apparatus," "device," "compound") is constructed using the WordNet lexical database. Each patent title is partitioned using conjunctions, and the key words within each partition are mapped to the activity and physical entity sets: patents whose titles map exclusively to activities are classified as process patents, while those mapping exclusively to physical entities are classified as product patents. The classification achieves 93.3% concordance with independent classifications by intellectual property experts at Maxval Group Inc. and patent examiners at IP Australia. In our sample, 60% of environmental patents are classified as process-based, compared to 43%

for traditional patents (difference significant at the 1% level), consistent with environmental innovation disproportionately involving changes to production methods, such as emissions capture techniques, waste treatment processes, and filtration methods, that are embedded inside production facilities and cannot be reverse-engineered from market observations.

We investigate this hypothesis in Table 3 by decomposing clean patents into process-based versus product-based. In column (1) and (2), we estimate the disclosure wedge for process- and product-based clean patents, respectively. A 10% increase in process-based patent disclosure generates social returns of 2.44% ($\beta_1^{C,Process} = 0.198$, $t = 2.77$). In contrast, product-based clean patents generate returns of 1.64% ($\beta_1^{C,Product} = 0.164$, $t = 2.07$). Estimating both types of patents jointly in column (3), we find that our results are entirely driven by clean process-based knowledge spillovers, while clean product-based spillovers generate insignificant returns.

The disclosure of clean process-based innovation generates large social spillovers because competitors cannot reverse-engineer them from market observation; patent disclosure is a first-order channel that enables knowledge transfer. In contrast, the disclosure of technologies such as clean product innovation generates limited incremental spillovers beyond those for R&D investment, because they can be analyzed through market transactions regardless of formal disclosure.

4.1.3 Baseline Results: Additional Tests

We conduct extensive robustness analyses to validate our baseline findings; full results appear in the Appendix. Our baseline specification augments the specification in BSV by including quality-weighted disclosure spillovers for clean and non-clean innovation, in addition to R&D-based spillovers. In Table B3, we confirm BSV's findings by including only R&D-based spillovers (column 1). A 10% increase in spillovers from technology peers' R&D investment is associated with a 1.36% increase in value, whereas spillovers from product market rivals are associated with a 0.9% decrease in value. We then show that the social returns from R&D-based spillovers are not absorbed by general disclosure-based spillovers (column 2): general disclosure-based knowledge capital by technology peers is not significantly related to market value. Crucially however, clean disclosure-based knowledge capital is strongly positively related to market value, with R&D-based spillovers providing no further incremental value (column 3), highlighting the unique importance of disclosure for clean relative to traditional innovation.

In our baseline method, we weight patents by disclosure specificity, as we are interested in spillovers to innovators in technology peers or product market rivals (which can interpret specific technical details), rather than the general public or other stakeholders (which may

value readability and volume of information over technicality). In Table B4, we present results using alternative quality metrics identified in Dyer et al. (2024). Columns (1) to (3) use patent fullness (volume of information), columns (4) to (6) use patent readability, and columns (7) to (9) use a principal component analysis (PCA) combination of specificity, fullness, and readability. Across all measures, our results remain qualitatively unchanged.[6]

We conduct an additional tests in which we identify rivals using the more conventional SIC industry codes in Table B5 and show that the private returns from disclosure are not driven by the choice of TNIC-3 based product market rivals. We find that competitive effects become insignificant when using the 3- or 4-digit Standard Industry Classification (SIC) codes (columns 1 and 2), likely reflecting the coarseness of SIC codes in identifying product market rivals (Hoberg and Phillips, 2016). We find similar results when using technological proximity ($TECH_{ij}$) to weight SIC-4 rivals' clean patent stock in the calculation of SPILLTECH^{C,Clean} (column 3).

Our empirical framework includes labor as an input in the production function, which we omit from the baseline estimation because employment is a less populated variable in Compustat. However, we include firm-level Ln(Employment) as an additional control variable in Table B6 and continue to find that social returns from clean innovation disclosure outweigh the private return effects of disclosure.

Alternative sample compositions and comparative tests further validate our findings. Including clean patents filed by private firms does not diminish our estimates (Table B7), indicating our focus on public firms does not overstate social returns. A licensing placebo test finds no significant spillover effects from licensed clean patents (Table B8), confirming our results capture public disclosure effects rather than formal knowledge-sharing agreements. Ensuring knowledge spillovers materialize by restricting the sample to actually cited patents also yields similar, although quantitatively smaller, results (Table B9). Across all tests, clean innovation disclosure generates substantial social returns with minimal competitive costs, while traditional innovation disclosure generates no incremental returns beyond R&D investment.

Finally, comparing clean to dirty innovation - proxied through patents with IPC classification F02: internal combustion engine-related technologies following Aghion et al. (2016) - shows returns from clean patent knowledge spillovers exceed those for dirty patent spillovers by more than 2:1 (Table B10). Dirty innovation generates no significant social returns beyond those for R&D investment, underscoring the unique social value of clean innovation disclosure.

[6]In Table C1, we also confirm the process-based mechanism using alternative disclosure quality metrics.

4.2 Non-Economic Social Benefits from Clean Knowledge Spillovers

In addition to the significantly larger disclosure wedge for clean innovation relative to traditional innovation, clean innovation also differs from traditional innovation through its sizeable social benefits beyond economic, in terms of pollution reductions, and health and mortality improvements. We therefore next establish whether disclosed clean innovation generates incremental real environmental outcomes.

4.2.1 Environmental Benefits

First, we plot pollution benefits in terms of reduced toxic emissions in Figure 1. We plot estimates for facility-level regressions relating emissions to knowledge acquired through patent citations, with the underlying coefficients reported numerically in Table B11. We measure acquired knowledge using backward citations, distinguishing between citations to clean patents (*Clean Acquired Knowledge*) and citations to non-clean patents (*Non-Clean Acquired Knowledge*). Both variables are standardized by their respective standard deviation. By measuring acquired knowledge through backward patent citations, our estimates reflect a lower bound, as knowledge from observing patents can also transfer through informal channels. We include facility fixed effects and industry $\times$ year fixed effects, absorbing time-invariant facility characteristics and common industry trends, as well as firm-level controls including sales, Tobin's Q, cash, CAPX, PPE, R&D, and market concentration.[7] The full regression specification is reported in the caption to Figure 1.

We find that clean knowledge acquisition is associated with substantial and persistent emissions reductions. A one standard deviation increase in clean acquired knowledge corresponds to a 5.1% reduction in total emissions one year after citation, increasing to 9% by year four. These effects persist through year five, indicating that disclosed clean technologies enable lasting operational improvements in pollution abatement. In contrast, non-clean acquired knowledge is associated with emissions increases of up to 6.8%, consistent with traditional innovation enabling production expansion rather than pollution reduction. This stark contrast validates our hypothesis that clean innovation generates environmental externalities requiring explicit disclosure for knowledge diffusion, whereas traditional innovation operates through different channels.[8]

[7] In Table B12, we confirm that our results also hold cross-sectionally, when excluding facility fixed effects.

[8] In Table B13, we document the direct relation between a firm's own clean patent filings and its emissions, finding negative associations consistent with He and Qiu (2025).

4.2.2 Health and Mortality Benefits

Second, we estimate two back-of-the-envelope effects of broader societal benefits of clean innovation disclosure: infant mortality (Currie and Schmieder, 2009) and social mortality costs (lives lost and medical expenses) from PM2.5 emissions (Deryugina et al., 2019). We assume that clean acquired knowledge increases with the amount of available clean innovation information at an equal rate, such that doubling the available clean innovation spillover pool results in an approximately 10% reduction in emissions, equivalent to 12,149 pounds fewer emissions. Following Currie and Schmieder (2009), we can then estimate that doubling disclosed clean innovation may result in 1.3 fewer infant deaths per 1,000 live births.[9]

To estimate social mortality costs, we assume industrial emissions contribute 11% of fine particulate matter (PM2.5) (Alves et al., 2023). Deryugina et al. (2019) show that a 1 $\mu g/m^3$ decrease in PM2.5 reduces social mortality costs by \$299,000 per million beneficiaries. In our setting, an average reduction of 12,149 pounds of emissions within a two-mile cubic kilometer radius of a polluting facility is expected to lower PM2.5 concentrations by 1.18 $\mu g/m^3$. Thus, the reduction in emissions leads to an estimated \$352,820 reduction in social mortality costs ($= 1.18 \times$ \$299,000). This means that for every ten million Medicare recipients, the estimated savings in lives lost and medical expenses due to reduced air pollution exceeds \$3.5 million.

Taken together, our results provide three important insights. First, they demonstrate that patent-mediated knowledge spillovers generate measurable real effects, not merely paper gains. Second, they establish that our market value estimates represent only a lower-bound economic component of total social benefits, excluding direct environmental benefits. Third, they confirm that clean innovation fundamentally differs from traditional innovation in its societal impact, justifying our focus on disclosure-specific policy interventions for environmental technologies.

5 IV Estimation: EPA Enforcement as Shocks to Disclosure Incentives

A potential concern is that our baseline estimates may suffer from endogeneity if firms more likely to patent are also better positioned to benefit from knowledge spillovers, or are those facing contemporaneous opportunities in clean innovation. To address this concern, we implement an instrumental variable (IV) strategy using EPA enforcement actions as a shock to disclosure incentives that should not affect the clean disclosure wedge through channels

[9] Currie and Schmieder (2009) estimate 0.106 fewer infant deaths per 1,000-pound reduction in emissions, resulting in 1.3 (= 0.106 x 12.149) fewer infant deaths in our scenario.

other than the increase in available clean knowledge. Enforcement shocks expose targeted firms to heightened scrutiny, legal risk, and the possibility of future sanctions, which in turn motivate firms to patent pollution-abatement innovations to credibly signal compliance and reputational responsiveness.

Table 4 presents two-stage least squares IV results. We instrument clean innovation disclosure ($\text{SPILLTECH}^S_{Clean}$ and $\text{SPILLTECH}^C_{Clean}$) with EPA enforcement actions faced by technologically related firms and product market rivals, respectively. We aggregate enforcement actions faced by all firms $j \neq i$ in the economy up to year t, weighted by their technological proximity to firm i (SPILLENF^S), or by all firms j that are product market rivals of firm i, weighted by their product market similarity with firm i (SPILLENF^C). As before, we control for returns generated from R&D investment to isolate the impact of disclosure from innovation input.[10].

The first-stage regressions (columns 1 and 2) show that EPA enforcement strongly predicts clean patent disclosure by both technology peers and competitors. EPA enforcement actions in technology peers (SPILLENF^S) are positively related to technology peers' disclosure ($\gamma = 0.189$, $t = 11.81$), while EPA enforcement actions in competitors (SPILLENF^C) exhibit strong predictive power for rivals' disclosure ($\gamma = 0.324$, $t = 39.54$). The Kleibergen-Paap Wald F-statistic of 781 substantially exceeds conventional thresholds, confirming instrument strength and ruling out weak instrument concerns. Consistent with environmental regulatory pressure increasing firms' incentives to disclose clean innovation via patent filings, EPA actions provide a positive shock to the available pool of clean knowledge.

Column (3) presents second-stage estimates. The IV coefficient on technology spillovers nearly doubles relative to baseline: a 10% increase in EPA-induced clean knowledge from technologically related firms raises firm value by 3.57% ($\beta^C_1 = 0.357$, $t = 2.77$). This larger magnitude likely reflects that EPA enforcement selects for particularly high-spillover process-based innovations, or corrects for attenuation bias due to measurement error in baseline patent quality measures. In contrast to our baseline results, the competitive effect from rivals' EPA-induced disclosure becomes insignificant ($\beta^C_2 = 0.030$, $t = 1.63$), indicating a small benefit - offsetting gains to disclosing firms at baseline - from EPA-induced disclosure in product market competitors. We also continue to find that non-clean innovation does not

[10]Table B14 provides facility- (Panel A), firm- (Panel B), and patent-level (Panel C) summary statistics within the EPA sample. Facilities receive an average of 0.03 formal EPA actions per year, indicating regulatory interventions are infrequent - albeit increasing over time, as indicated in Figure B1. The average facility-level penalty amounts to $150,203 USD and can be as high as $20 million USD. Monetary penalties from environmental violations may be relatively small, but they often result in much larger direct and indirect costs, such as the withdrawal of emission and operating licenses, exponential increases in fines for repeat violators, regulatory scrutiny, reputation damage, and litigation risk (Blundell, Gowrisankaran, and Langer, 2020)

generate a meaningful disclosure wedge after accounting for R&D investment spillovers; if anything, non-clean spillovers primarily generate competitive effects, reflecting firms' decision to disclose only when private gains exceed competitive costs.

The disclosure wedge in our IV specification is 3.27% ($0.357 - 0.030 = 0.327$, i.e. 3.27% for a 10% increase in disclosure-induced spillovers), exceeding the baseline wedge of 2.3%. This substantial gap between spillover benefits to technology peers (3.57%) and competitive effects on disclosing firms 0.30%) demonstrates a substantial misalignment between private incentives and social benefits, with firms generating nearly 11 times as much in peer firm value for every 1 they might lose from rivals' learning.

In Table 5, we confirm that our results reflect process-based clean patents, rather than product-based clean patents, in the IV setting. The second-stage regressions in column (3) show that instrumented process-based clean patents generate 3.46% social returns per 10% increase in disclosure ($\beta_1^{S,Process} = 0.346$, $t = 2.75$). This estimate is larger than our baseline and nearly identical to the aggregate clean patent IV effect of 3.57% in Table 4, indicating that virtually all EPA-induced spillovers operate through process-based innovations, consistent with regulatory enforcement targeting operational pollution abatement methods rather than product reformulations. As before, competitive effects remain small and positive ($\beta_2^C = 0.030$), reinforcing that environmental disclosure generates no meaningful competitive costs even for process innovations that might be considered more proprietary. The combination of large process-based spillovers and negligible competitive effects creates an especially pronounced disclosure wedge for the technologies most critical to pollution reduction.[11]

To contextualize our IV estimates, consider the median publicly listed firm in our sample with market value of $2.1 billion. A 10% increase in available clean knowledge arising from patent disclosure by technology peers generates approximately $75 million in market value for the focal firm (based on the 3.57% IV coefficient on SPILLTECH^S). However, when the focal firm's competitors disclose clean innovation, the focal firm does not significantly lose. This asymmetry reveals the disclosure wedge: clean innovation disclosure generates nearly $11 in peer firm value for every $1 captured by competitors, reflecting private disincentives to disclose despite large social benefits.

Critically, these estimates represent only the portion of the spillover benefits appropriated by other firms' shareholders. Total social benefits are substantially larger. Our emissions analysis documents up to 6% reductions in facility-level pollution from clean knowledge spillovers, translating to health benefits, regulatory cost savings, and climate externalities

[11]These results also remain robust to alternative disclosure quality weighting measures in Appendix C using the number of words (Tables C2-C3) and readability (Tables C4-C5).

not reflected in market valuations. Additionally, we observe only spillovers to public firms, excluding knowledge transfer to private firms, startups, universities, and consumers. The $75 million appropriable benefit thus understates the full social value of environmental disclosure by an order of magnitude or more.

5.1 Environmental Benefits from EPA-Induced Disclosure

In order for regulatory pressure to trigger environmentally meaningful spillovers, EPA-induced patents should reduce emissions beyond innovating firms' boundaries. We therefore next investigate whether these patents reflect true clean innovation by investigating environmental outcomes for firms that acquired clean knowledge from EPA-induced spillovers. Specifically, we focus on the subset of cited clean patents filed by firms immediately following EPA actions.

In Figure 2, we plot estimates for the relation between emissions and EPA-induced clean knowledge spillovers. The underlying coefficients are reported in Table B15. EPA-induced knowledge spillovers are associated with equally large emissions reductions as in our baseline results: a one standard-deviation increase in EPA-induced clean acquired knowledge is associated with a 4.4% to 7.5% reduction in total emissions in years $t+1$ to $t+5$ following the citation of EPA-induced clean innovation. These results imply that clean patents filed after EPA actions include important knowledge for the reduction of emissions, which would have remained untapped without the EPA intervention. Regulation-induced disclosure, therefore, provides large social benefits in terms of reducing emissions beyond the filing firm and does not merely represent an alternative form of greenwashing.

5.2 EPA Actions as Shocks to Disclosure Incentives: Additional Tests

The use of EPA actions as a shock to firms' incentives to publicly disclose clean innovation hinges on two key assumptions: EPA actions increase firms' incentives to disseminate their clean innovation, and EPA actions are exogenous to social returns. While the EPA does not mandate that firms patent pollution abatement technologies, patenting can demonstrate a firm's capability to meet future requirements for advanced pollution control technologies, which can facilitate quicker settlement. In addition, EPA actions should not affect aggregate economic returns in non-targeted firms through channels other than shifts in clean knowledge spillovers. Next, we provide evidence consistent with these two assumptions.

5.2.1 EPA-Induced Clean Innovation Disclosure: Relevance

We investigate whether an increase in environmental regulatory risk exposure from EPA actions is associated with an increase in clean innovation disclosure at the extensive (number of patents filed) and intensive (disclosure timing) margin. Because regulatory actions increase uncertainty regarding a firm's ability to address environmental risk and reduce emissions, firms should have increased incentives to signal information regarding their emission abatement capabilities to key stakeholders, such as regulators and investors (Skinner and Valentine, 2023). Patents provide highly detailed information subject to a rigorous third-party verification process during which a USPTO patent examiner verifies the novelty and utility of the filing. As implied by signaling theory (Spence, 1973), patent filings thereby provide one of the most highly accurate, credible, and third-party verified signals about a firm's capabilities to reduce emissions. In the absence of increased regulatory pressure, firms may generally choose to keep their technology as a trade secret.

We first estimate the impact of EPA enforcement in year t on the number of clean patents filed by affected firms in year t in Poisson specifications at the firm and facility level, as regulatory pressure is observed at the facility level, whereas patent outcomes are observed at the firm level. We consider number of clean patents filed, as every patent filing represents a disclosure decision. We control for facility- and firm-level control variables, including the facility's total emissions as firms' incentives to develop emission-abatement technologies may be correlated with their emissions output, in addition to facility fixed effects and industry $\times$ year fixed effects.

In Table 6, we find that the number of formal actions taken against a firm's facilities is positively related to clean patents filed at both the firm-(column 1) and facility-(column 2) level (Panel A).[12] Economically, a one-unit increase in formal EPA enforcement actions in facility f is associated with a 5.65% ($=(e^{0.055}-1)\times 100\%$) increase in clean patents filed by the facility's parent firm in the year of the enforcement action.[13] In absolute numbers, this percentage change is equivalent to approximately 1.10 more clean patents filed by a parent firm when moving from 0 to 1 enforcement action in year t.

EPA pressure should only affect firms' incentives to disclose clean innovation, and not those of traditional (non-clean) innovation. In Panel B, we do not find evidence for an

[12]We assume the speed and timing of firms' patent filings in response to EPA pressure is fully captured in the enforcement action year. However, there may be calendar effects (e.g., EPA action in November and patent filing the following February) or delays in preparing patent filings, resulting in patent filings occurring in the year after the EPA action. In unreported results, we analyze filings in the year after EPA actions (t+1), and the results remain qualitatively similar.

[13]In a Poisson specification, the magnitude of a one-unit change in an independent variable (e.g., EPA enforcement actions) is calculated as $e^{\beta}-1$, which represents the percentage change in the dependent variable (e.g., clean patent count) when multiplied by 100% (Sautner et al., 2023).

increase in non-clean patent filings following EPA enforcement actions. Firms therefore only make clean technology publicly available, indicating that the increase in patent filings is not due to a change in firms' patenting strategy.

5.2.2 EPA-Induced Disclosure Incentives: Patent Disclosure Timing

A second key assumption is that EPA enforcement shocks are exogenous to aggregate social returns.[14] We provide evidence consistent with firms' use of clean patent filings as a short-term, credible signaling mechanism towards key stakeholders, an incentive that should be exogenous to the social returns from such public disclosure.

Specifically, we exploit the *timing* of firms' decisions to publicly disclose clean innovation earlier than required by the USPTO. The American Inventor Protection Act (AIPA), implemented in November 2000, changed the disclosure requirements for filed patents, requiring firms to publish their patent application after 18 months from the application date. The AIPA and USPTO changes to patent publication are summarized in Appendix A. We capture the timing and speed of patent publication by measuring *Disclosure Latency* and *% Latency* (Glaeser and Landsman, 2021; Hegde, Jovanovic, and Liu, 2023). *Disclosure Latency* measures the number of days between the clean patent filing date and the earliest date of its publication by the USPTO, thereby capturing not only the filer's voluntary publication before the 18-month mark but also its opt-out decision to delay publication beyond the 18-month mark. *% Latency* normalizes disclosure latency by the number of days from filing to the USPTO-mandated 18-month publication mark or the patent application decision date.[15]

Table 7 estimates the effect of EPA actions on the timing of firms' clean patent disclosure decisions, conducted at the patent application level, which enables us to compare successful clean patent applications by the same firm that differ in how fast they were publicly disclosed. Column (1) shows that a one-unit increase in formal actions by the EPA is associated with a 3.3% decrease in the time to disclosure ($e^{-0.033} - 1) \times 100\%$, equivalent to roughly 15 days faster publication for EPA-induced patents. Column (2) shows that an additional EPA enforcement action is associated with an 11.5% decrease in disclosure latency, equivalent to 87 days (approximately 3 months) earlier publication. In contrast, there is no reduction in

[14] It is also worth nothing that, if EPA actions cluster by industry or emission output and thereby affect social returns by signaling higher potential regulatory scrutiny in related firms, such a mechanism would bias against finding positive social returns and/or divergence in social and private returns.

[15] Panel C of Table B14 shows that the average *Disclosure Latency* is 462 days, i.e. clean patent filings are made public 15 months after the filing date (3 months before the USPTO's mandatory 18-month publication requirement). *% Latency* averages 68%, indicating clean patents are published 32% earlier than mandated. The majority of firms publish their patents at the mandated deadline, as median *Latency* is 547 days. Under certain conditions, filers can delay publication until the patent grant date: the latest possible publication date (*Days to Mandatory Disclosure*) averages 759 days after filing.

the time to publication for non-clean patents (columns 3 and 4), indicating that these effects are unique to firms' clean innovation disclosure decisions.

5.2.3 EPA-Induced Disclosure Incentives: Signaling to Regulators and Investors

Given the large direct and indirect costs associated with repeated violations, firms may patent clean innovation as a signal to the EPA (Berrone et al., 2013). A natural question is then whether clean patent filings effectively affect EPA enforcement actions. In Table 8, we focus on a facility's High Priority-Violator (HPV) status, which indicates serious noncompliance with environmental regulations and is designed to be a substantial burden to firms (Blundell, Gowrisankaran, and Langer, 2020). Due to the heightened scrutiny, potential operational disruptions, and additional compliance costs, firms must allocate more resources to address their environmental violations the longer they stay in HPV status.

We investigate all HPV cases in our sample and estimate the relation between HPV status duration (number of days) and firms' decision to disclose clean innovation through patent filings. Column (1) shows that the number of clean patents filed is negatively related to HPV status duration. Column (2) similarly shows that firms subject to HPV status filing at least one clean patent experience a 33% reduction in HPV duration compared to those that do not file clean patents. Given that the mean HPV duration is 600 days, this represents a 200-day reduction in HPV status and a quicker settlement with the EPA. These results support anecdotal evidence that the EPA may view a firm's clean patent filing as evidence of its commitment to addressing pollution-related regulatory issues.

Given the potential costs and negative consequences of repeated violations, firms should also have incentives to disclose clean technology to investors. Moreover, for clean patents to provide a credible signal to investors and regulators, these stakeholders must be aware of their existence. Patent filings can be observed through the USPTO's central repository, but firms should also have incentives to disseminate information about their clean patents more widely: Li and Wang (2024) document that the EPA consistently retrieves corporate SEC filings, particularly during enforcement investigations.

We consider the extent to which firms disseminate clean innovation information by using terminology related to clean technology in their annual and quarterly reports (10-K and 10-Q filings). We search 10-K/Q filings for the co-occurrence of terms related to "clean" (e.g., variations of clean, green, environmental, sustainable, pollution, waste, toxic, emission, contamination, eco-friendly, energy-efficient, recyclable, and biodegradable) and terms related to "technology" (e.g., variations of technology, innovation, patent, research, and solution), and define *10K/Q* as an indicator variable equal to one if these terms co-occur in the same sentence, and zero otherwise.

Table B16 reports the main specification estimating changes in clean innovation references around patent filings. Coefficient estimates for $\beta_{1,\tau}$ over a five-year period around a patent filing are graphically depicted in Figure 3. Firms are generally unlikely to refer to clean technology in 10K/Q reports in years prior to a clean patent filing. However, they are 0.7 percentage points more likely to include clean technology references in the patent filing year. Notably, at this time the patent filing is not publicly available and not visible to stakeholders, indicating the firm's urgency to inform market participants about their clean technology. Moreover, this effect represents a 3.5% increase relative to the unconditional co-occurrence probability. The inclusion of clean technology references increases over time, as firms get closer to the USPTO-mandated 18 months disclosure requirement. The highest use of clean technology terminology is two years after the patent filing, which captures the time from filing a patent to the average publication date.

6 Conclusion

We examine the disclosure wedge between social and private returns for clean innovation in a panel of U.S. firms filing clean patents. We show that the social benefits of disclosing clean innovation, as measured through environmental outcomes and aggregate economic returns, exceed the private returns to disclosing firms. This wedge indicates a potential under-estimation of clean technologies' full societal contribution and a potential under-disclosure of clean innovation from the social perspective that remains unaddressed by existing climate policy.

Crucially, our documented effects represent a lower boundary for the impact of clean innovation disclosure on social benefits. Our results capture effects on domestic U.S. firms from observing clean technology made available by other U.S. firms. However, clean knowledge spillovers do not stop at the U.S. border: the development of groundbreaking clean technologies by highly skilled U.S. firms is likely to have significant value for polluting firms in other countries, particularly those in developing and emerging markets, which we do not account for.

Overall, our study underscores the central importance of disclosure requirements and knowledge spillovers in achieving global climate objectives. These findings highlight a critical policy gap: while environmental tax incentives and R&D subsidies encourage investment in clean innovation, the broader social benefits of such investment may remain untapped if firms lack sufficient incentives to make their advancements available to others. Our findings suggest that policymakers should not only incentivize R&D but also develop frameworks that promote the disclosure of clean technologies - with the appropriate intellectual property

protections. Such a comprehensive strategy may maximize the impact of clean innovation on environmental and economic progress in addressing the urgent challenges of the climate crisis. Future research might explore how different regulatory frameworks, industry structures, or financial constraints further shape firms' disclosure behaviors and affect the broader uptake of clean technology.

REFERENCES

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen. 2016. Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. Journal of Political Economy 124:1–51.

Alves, C., M. Evtyugina, E. Vicente, A. Vicente, I. C. Rienda, A. S. de la Campa, M. Tomé, and I. Duarte. 2023. Pm2.5 chemical composition and health risks by inhalation near a chemical complex. Journal of Environmental Sciences 124:860–74.

Anton, J. J., and D. A. Yao. 1994. Expropriation and inventions: Appropriable rents in the absence of property rights. The American Economic Review 84:190–209.

———. 2004. Little patents and big secrets: Managing intellectual property. The RAND Journal of Economics 35:1–22.

Arrow, K. J. 1962. Economic welfare and the allocation of resources for invention. In R. R. Nelson, ed., The rate and direction of inventive activity: Economic and social factors, 609–26. Princeton: Princeton University Press.

Bena, J., H. Ortiz-Molina, and E. Simintzi. 2022. Shielding firm value: Employment protection and process innovation. Journal of Financial Economics 146:637–44.

Berrone, P., A. Fosfuri, L. Gelabert, and L. R. Gomez-Mejia. 2013. Necessity as the mother of 'green'inventions: Institutional pressures and environmental innovations. Strategic Management Journal 34:891–909.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. Identifying technology spillovers and product market rivalry. Econometrica 81:1347–93.

Blundell, W., G. Gowrisankaran, and A. Langer. 2020. Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations. American Economic Review 110:2558–85.

Bolton, P., and M. Kacperczyk. 2021. Do investors care about carbon risk? Journal of Financial Economics 142:517–49.

Currie, J., and J. F. Schmieder. 2009. Fetal exposures to toxic releases and infant health. American Economic Review 99:177–83.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif. 2019. The mortality and medical costs of air pollution: Evidence from changes in wind direction. American Economic Review 109:4178–219.

Dyer, T. A., S. Glaeser, M. H. Lang, and C. Sprecher. 2024. The effect of patent disclosure quality on innovation. Journal of Accounting and Economics 77:101647–.

Flammer, C. 2021. Corporate green bonds. Journal of Financial Economics 142:499–516.

Garel, A., A. Romec, Z. Sautner, and A. F. Wagner. 2024. Do investors care about biodiversity? Journal of Financial Economics Forthcoming.

Giglio, S., T. Kuchler, J. Stroebel, and X. Zeng. 2023. Biodiversity risk. Working Paper 31137.

Giroud, X., E. Liu, and H. Mueller. 2024. Innovation spillovers across U.S. tech clusters. Working paper, NBER.

Glaeser, S. A., and W. R. Landsman. 2021. Deterrent disclosure. The Accounting Review 96:291–315.

Griliches, Z. 1979. Issues in assessing the contribution of research and development to productivity growth. The Bell Journal of Economics 10:92–116.

Gunning, R. 1952. The technique of clear writing. (No Title) .

Haščič, I., and M. Migotto. 2015. Measuring environmental innovation using patent data. Working Paper, OECD.

He, Q., and B. Qiu. 2025. Environmental enforcement actions and corporate green innovation. Journal of Corporate Finance 91:102711–.

Heeb, F., J. F. Kölbel, F. Paetzold, and S. Zeisberger. 2023. Do investors care about impact? The Review of Financial Studies 36:1737–87.

Hegde, D., B. Jovanovic, and J. Liu. 2023. Patent signaling. Working paper, New York University.

Hoberg, G., and G. Phillips. 2016. Text-based network industries and endogenous product differentiation. Journal of Political Economy 124:1423–65.

Hong, H., G. A. Karolyi, and J. A. Scheinkman. 2020. Climate finance. The Review of Financial Studies 33:1011–23.

Ilhan, E., P. Krueger, Z. Sautner, and L. T. Starks. 2023. Climate risk disclosure and institutional investors. Review of Financial Studies 36:2617–50.

Jaffe, A. B. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. The American Economic Review 76:984–1001.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. Technological innovation, resource allocation, and growth. Quarterly Journal of Economics 132:665–712.

Krueger, P., Z. Sautner, and L. T. Starks. 2020. The importance of climate risks for institutional investors. The Review of Financial Studies 33:1067–111.

Li, B., and A. Y. Wang. 2024. On the epa's radar: The role of financial reports in environmental regulatory oversight. Journal of Accounting Research 62:1849–900.

Matray, A. 2022. The local innovation spillovers of listed firms. Journal of Financial Economics 141:395–412.

Moretti, E. 2021. The effect of high-tech clusters on the productivity of top inventors. American Economic Review 111:3328–75.

Sautner, Z., L. van Lent, G. Vilkov, and R. Zhang. 2023. Firm-level climate change exposure. The Journal of Finance 78:1449–98.

Skinner, N., and K. Valentine. 2023. Green patenting and voluntary innovation disclosure. Working paper, SSRN.

Spence, M. 1973. Job market signaling. The Quarterly Journal of Economics 87:355–74.

Tham, W. W., S. Baslandze, E. Sojli, and L. Liu. 2025. Foundational processes and growth. Journal of Financial Economics Forthcoming.

Verrecchia, R. E. 1983. Discretionary disclosure. Journal of Accounting and Economics 5:179–94.

Tables and Figures

Figure 1
Clean Innovation Knowledge Spillovers and Toxic Emissions

This table presents regression coefficients for the relation between emissions and knowledge spillovers from clean innovation. The analysis is at facility-year level.

$ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Acq. Know.}_{i,t} + \beta_2 \text{Non-Clean Acq. Know.}_{i,t} + \Gamma X_{i,f,t} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}$.

Total Emissions are a facility's total onsite emissions in years $t + \tau$ for $\tau = 1$ to $\tau = 5$. *Clean Acquired Knowledge* is the number of backward citations by the focal firm i's patents to other firms' clean patents and *Non-Clean Acquired Knowledge* is the number of backward citations to other firms' non-clean patents. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI, which are omitted for brevity. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level. The sample period is 1988 to 2020.

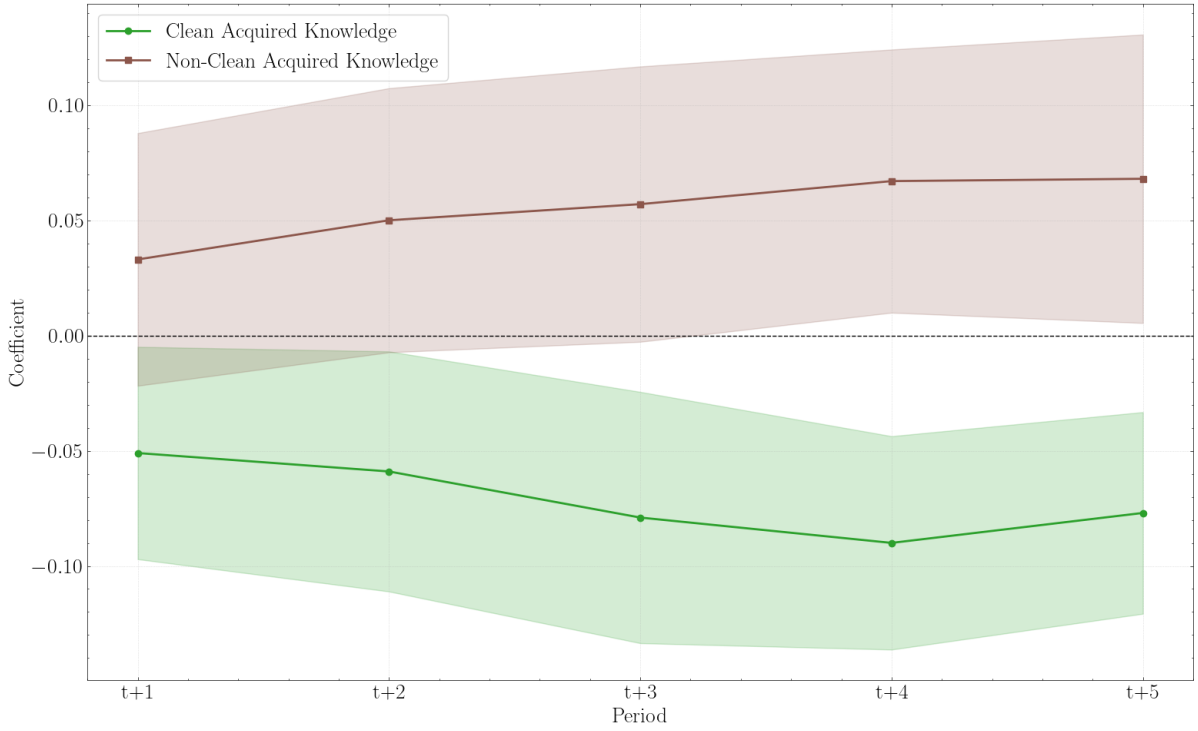


Figure 2
Clean Innovation Knowledge Spillovers from EPA-Induced Patents and Toxic Emissions

This table presents regression coefficients for the relation between emissions and knowledge spillovers from clean innovation. The analysis is at facility-year level.

$ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Acq. Know.}_{i,t} + \beta_2 \text{Non-Clean Acq. Know.}_{i,t} + \Gamma X_{i,f,t} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}$.

Total Emissions are a facility's total onsite emissions in years $t + \tau$ for $\tau = 1$ to $\tau = 5$. *Clean Acquired Knowledge* is the number of backward citations by the focal firm i's patents to other firms' clean patents filed during EPA enforcement years and *Non-Clean Acquired Knowledge* is the number of backward citations to other firms' non-clean patents. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI, which are omitted for brevity. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level. The sample period is 1988 to 2020.

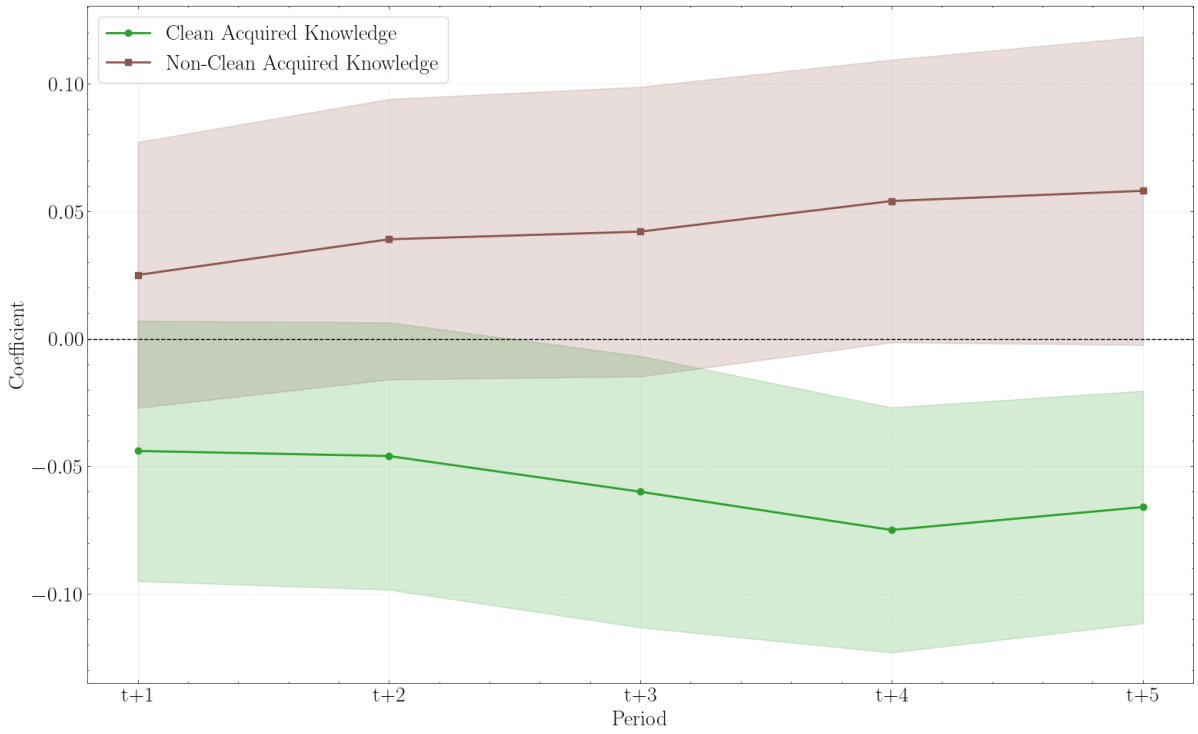


Figure 3
Relation between 10-K/Q Reporting and Clean Patents

This figure presents coefficient estimates of the relation between firms' references to clean technologies in their 10-K/Q filings and their clean patent filings. We estimate the following regression at the firm level:

$$10K/Q_{i,t+\tau} = \beta_0 + \beta_{1,\tau}\text{Clean Patents}_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+\tau}.$$

Full regression results are presented in Table B16. We search for the co-occurrence of terminology relating to 'clean' and 'technology' within the same sentence in the 10-K/Q text. Clean terms include clean, green, environmental, sustainable, pollution, waste, toxic, emission, contamination, eco-friendly, energy-efficient, recyclable, and biodegradable, and their variations. Technology terms include technology, technologies, innovation, patent, research, and solution, and their variations. $10K/Q_{i,t+\tau}$ is an indicator variable equal to one if the firm mentions co-occurrences of clean and technology terms measured at time $t+\tau$, and zero otherwise. We report β_1 for $\tau = -2, -1, 0, 1, 2$. *Clean Patents* is the number of clean patents filed in year t, λ_i are firm fixed effects and λ_t are year fixed effects. X is a vector of control variables which including ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI. The shaded area is the 95% confidence interval calculated using clustered standard errors at the firm level. The sample period is 1993 to 2020, as per SEC-EDGAR filings availability.

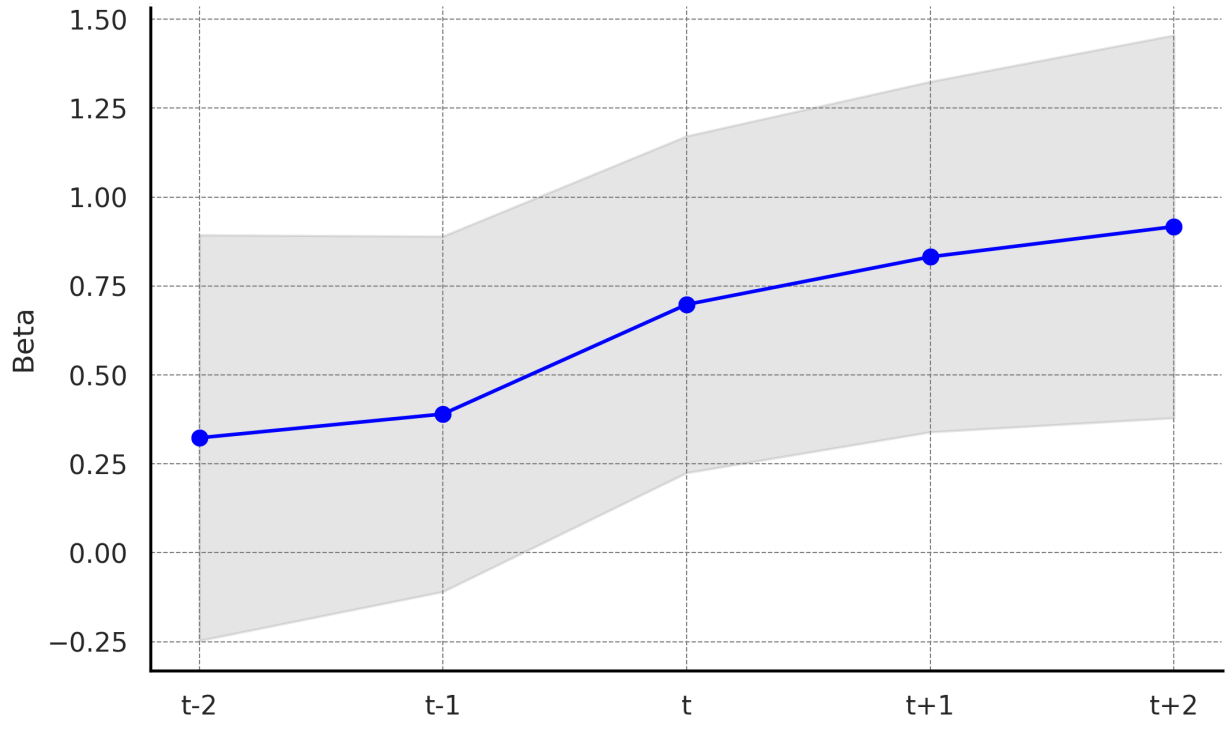


Table 1
Summary Statistics

This table presents summary statistics for the main variables used in our baseline estimations in Tables 2 to 5 (Panel A) and Figure 1 (Panel B). All variable definitions are provided in Table B1 in the Appendix. The sample period is 1988 to 2020. Summary statistics for other variables used in additional tests are presented in Table B14.

Panel A. Firm level

	N	Mean	Median	Std. Dev.
$SPILLTECH^{S,Clean}$	196,080	198.31	0.00	829.89
$SPILLTECH^{C,Clean}$	196,080	4.57	0.09	13.18
$SPILLTECH^{S,Process}$	196,080	124.34	0.00	510.73
$SPILLTECH^{S,Product}$	196,080	77.43	0.00	342.35
$SPILLTECH^{S,Non\text{-}Clean}$	196,080	40,035.27	0.00	76,415.12
$SPILLTECH^{C,Non\text{-}Clean}$	196,080	741.69	22.43	1,925.73
$SPILLTECH^{S,R\&D}$	196,080	11,132.74	0.00	45,222.12
$SPILLTECH^{C,R\&D}$	196,080	1,330.38	0.00	4,708.01
MV/TA	196,080	1.40	0.79	1.84
R&D Stock/TA	196,080	0.16	0.00	0.86
Clean Patent Stock	196,080	3.45	0.00	68.72
Sales	195,372	2,677.49	172.25	13,576.80
Cash	196,080	0.18	0.08	0.23
CAPX	196,080	0.05	0.03	0.08
PPE	196,080	0.24	0.14	0.25
HHI	195,804	0.23	0.17	0.34

Panel B. Facility level

	N	Mean	Median	Std. Dev.
Ln(Total Emission)	133,465	7.47	8.32	4.13
Clean Acquired Knowledge	133,465	13.63	0.00	81.06
Non-Clean Acquired Knowledge	133,465	1,470.01	51.00	4,167.76

Table 2
Disclosure Wedge from Clean Knowledge Spillovers

This table presents results for the disclosure wedge from clean knowledge spillovers. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =&\beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$	0.193**		0.194**
	(2.46)		(2.48)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$		-0.010*	-0.010*
		(-1.84)	(-1.90)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	-0.030	-0.025	-0.030
	(-1.46)	(-1.21)	(-1.44)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.008***	-0.007***	-0.007**
	(-3.35)	(-2.59)	(-2.50)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.039	0.142***	0.039
	(0.80)	(4.32)	(0.79)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.004**	-0.002	-0.002
	(-2.00)	(-1.04)	(-1.04)
Ln(R&D Stock/TA)	0.231***	0.229***	0.231***
	(7.87)	(7.82)	(7.87)
Ln(G)	-0.004	0.015	-0.003
	(-0.16)	(0.67)	(-0.12)
Ln(Sales)	-0.091***	-0.091***	-0.091***
	(-16.85)	(-16.86)	(-16.83)
Cash	0.700***	0.701***	0.700***
	(24.35)	(24.37)	(24.35)
CAPX	0.826***	0.827***	0.826***
	(10.65)	(10.64)	(10.64)
PPE	-0.241***	-0.244***	-0.241***
	(-4.93)	(-4.98)	(-4.93)
HHI	0.009*	0.009*	0.009*
	(1.90)	(1.91)	(1.89)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75

Table 3
Disclosure Wedge from Clean Knowledge Spillovers: Product- and Process-Based Patents

This table presents results for the disclosure wedge from clean product-based and process-based knowledge spillovers. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process/Product})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process/Product}$ captures spillovers from clean innovation disclosed by other firms defined as $\text{SPILLTECH}^S_{Clean,Process/Product} = \sum_{j\neq i} \text{TECH}_{ij} G^{q,Process/Product}_{jt}$, where Process/Product-Based patents are defined following Tham et al. (2025); TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$	0.198***		0.196*
	(2.77)		(1.86)
$\text{Ln(SPILLTECH}^S_{\text{Clean, Product}}\text{)}$		0.164**	0.003
		(2.07)	(0.02)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	-0.010*	-0.010*	-0.010*
	(-1.94)	(-1.87)	(-1.94)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	-0.031	-0.029	-0.031
	(-1.49)	(-1.38)	(-1.49)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.007**	-0.007**	-0.007**
	(-2.49)	(-2.51)	(-2.49)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.032	0.065	0.032
	(0.66)	(1.37)	(0.64)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.002	-0.002	-0.002
	(-1.03)	(-1.05)	(-1.03)
Ln(R&D Stock/TA)	0.231***	0.231***	0.231***
	(7.88)	(7.87)	(7.87)
Ln(G)	-0.007	0.001	-0.007
	(-0.28)	(0.05)	(-0.28)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75

Table 4
Disclosure Wedge from Clean Knowledge Spillovers: IV

This table presents results for the disclosure wedge from clean knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$	$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.189***	-0.015	
	(11.81)	(-0.72)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.324***	
	(-0.89)	(39.54)	
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$			0.357***
			(2.77)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$			0.030
			(1.63)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	0.012***	0.043***	-0.036*
	(5.16)	(2.87)	(-1.70)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.000	0.156***	-0.013***
	(-0.88)	(52.18)	(-3.23)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.248***	0.038	-0.051
	(5.87)	(1.07)	(-0.70)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.000	0.132***	-0.008***
	(-0.27)	(55.87)	(-2.61)
Ln(R&D Stock/TA)	-0.002	-0.022	0.233***
	(-1.07)	(-1.24)	(7.90)
Ln(Clean Patent Stock)	0.015	0.079***	-0.023
	(1.45)	(3.36)	(-0.86)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			781

Table 5
Disclosure Wedge from Clean Process-Based Knowledge Spillovers: IV

This table presents results for the disclosure wedge from clean process-based knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean,Process} = \sum_{j\neq i} \text{TECH}_{ij} G^{q,Process}_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. $G^{q,Process}_{jt}$ is the disclosure-quality-weighted stock of clean process-based patents filed by firm j where process-based patents are defined as in Tham et al. (2025). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean,Process}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^{S,\text{ Clean, Process}}\text{)}$	$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.196***	-0.015	
	(11.63)	(-0.72)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.324***	
	(-0.73)	(39.54)	
$\text{Ln(SPILLTECH}^{S,\text{ Clean, Process}}\text{)}$			0.346***
			(2.75)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$			0.030
			(1.62)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	0.016***	0.043***	-0.037*
	(6.08)	(2.87)	(-1.76)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.000	0.156***	-0.013***
	(-0.21)	(52.18)	(-3.24)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.265***	0.038	-0.054
	(5.91)	(1.07)	(-0.72)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.000	0.132***	-0.008***
	(-0.18)	(55.87)	(-2.60)
Ln(R&D Stock/TA)	-0.002	-0.022	0.233***
	(-0.81)	(-1.24)	(7.90)
Ln(Clean Patent Stock)	0.029***	0.079***	-0.027
	(3.14)	(3.36)	(-1.01)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			782

Table 6
Clean Innovation Disclosure as a Response to EPA Pressure

This table examines whether firms file clean patents in response to EPA actions. Panels A and B focus on samples of clean and non-clean patents, respectively. We present firm-year (column 1) and facility-year (column 2) Poisson regressions of the form:

$$Patents_{i,f,t} = \beta_0 + \beta_1 Formal\ Action_{i,f,t} + \Gamma X_{i,f,t-1} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t},$$

where $Patents_{i,f,t}$ represents the number of clean (Panel A) and non-clean (Panel B) patents filed by firm i in year t. $Formal\ Action_{i,f,t}$ denotes the number of EPA actions recorded in the ECHO system, as described in Section 3.2. $\lambda_{ind,t}$ are industry fixed effects, defined using the facility's NAICS 6-digit classification. Firm-level estimations measure total EPA actions for firm i, scaled by the number of operating facilities. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, HHI, and ln(Total Emissions) that are suppressed for brevity. T-statistics adjusted for standard errors clustered at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020.

Panel A - Clean Patents

	Clean Patents$_t$	
	(1)	(2)
Formal Actions	0.422***	0.055***
	(2.83)	(2.68)
Controls	✓	✓
Firm FE	✓	
Year FE	✓	
Facility FE		✓
Industry × Year FE		✓
Obs.	7,120	62,366

Panel B - Non-Clean Patents

	Non-Clean Patents$_t$	
	(1)	(2)
Formal Actions	0.147	0.019
	(1.12)	(1.43)
Controls	✓	✓
Firm FE	✓	
Year FE	✓	
Facility FE		✓
Industry × Year FE		✓
Obs.	7,120	62,366

Table 7
Disclosure Timing and EPA Actions

This table presents a regression analysis of the relationship between patent publication times and EPA actions, conducted at the patent level.

$$\ln(\text{Disclosure Latency})/\%\ \text{Latency}_{p,i,t} = \beta_0 + \beta_1 \text{Formal Action}_{p,i,t} + \Gamma X_{p,i,t} + \lambda_i + \lambda_t + \epsilon_{p,i,t},$$

where *Disclosure Latency* is the natural logarithm of the number of days between the patent filing date and the earliest publication date, and *% Latency* is the percentage latency, defined as the fraction of disclosure latency days scaled by the number of days until the latest possible publication. *Formal Action* is an indicator variable equal to one if the patent filing occurs in the same year as a formal EPA action and zero otherwise. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI that are suppressed for brevity. T-statistics, adjusted for standard errors clustered at the firm level, are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 2001 to 2020, following the implementation of the American Inventor Protection Act (AIPA).

	Clean		Non-Clean	
	ln(Disclosure Latency)	% Latency	ln(Disclosure Latency)	% Latency
	(1)	(2)	(3)	(4)
Formal Action	-0.033**	-0.115***	0.002	0.003
	(-2.56)	(-2.65)	(0.19)	(0.14)
Ln(Sales)	0.028	0.044	0.044	0.032
	(0.37)	(0.36)	(0.64)	(0.51)
Tobin's Q	0.019	0.112	0.021	0.027
	(0.32)	(1.07)	(1.43)	(1.12)
Cash	-0.605	-1.044	0.256	0.255
	(-0.80)	(-0.83)	(1.56)	(0.84)
CAPX	-1.712	-3.092	0.370	-0.178
	(-1.63)	(-1.60)	(0.69)	(-0.25)
PPE	1.348***	1.859**	0.830***	0.997***
	(3.54)	(2.34)	(3.22)	(2.82)
R&D	-3.634**	-3.602	1.045	0.347
	(-2.45)	(-1.34)	(1.22)	(0.47)
Missing R&D	0.081	-0.208	-0.118	-0.104
	(0.47)	(-0.66)	(-0.85)	(-0.67)
HHI	0.412	0.694	0.133	0.243
	(1.03)	(0.90)	(0.72)	(1.37)
Ln(Days to Mandatory Disclosure)	0.402***	-0.674***	0.328***	-0.646***
	(8.73)	(-7.13)	(14.46)	(-13.22)
Year FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Obs.	23,392	23,392	450,767	450,770
Adj. R^2	0.31	0.48	0.23	0.45

Table 8
Clean Patent Filings as Signaling to Regulators

This table presents coefficient estimates for the effect of clean patent filings on the time to resolve a facility's High Priority Violator (HPV) status. The analysis is conducted at the individual HPV incident level indexed by z.

$$\ln(\text{HPV Duration})_{z,i,t} = \beta_0 + \beta_1 \text{Clean Patent}_{i,t} + \Gamma X_{i,t} + \epsilon_{z,i,t},$$

where *HPV Duration* is the number of days from the start of a facility's HPV status until its resolution. *Clean Patent* is the number of clean patents filed during the HPV designation, and *I(Clean Patent)* is an indicator variable denoting whether at least one clean patent has been filed during the HPV designation. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI that are suppressed for brevity. T-statistics, adjusted for clustered standard errors at the firm level, are reported in parentheses. ***, **, and * denote significance levels at 1%, 5%, and 10%, respectively. The sample period is 1988 to 2020.

	HPV Duration	
	(1)	(2)
Clean Patents	-0.001***	
	(-4.91)	
I(Clean Patents)		-0.336***
		(-4.89)
Controls	✓	✓
Year FE	✓	✓
Firm FE	✓	✓
Obs.	5,827	5,827
Adj. R^2	0.19	0.20

Appendix A. The American Inventor Protection Act

The American Inventor Protection Act (AIPA), implemented in November 2000, mandated patent applications to be publicly published by the USPTO at three possible times: at the grant decision date, 18 months after the USPTO filing date, or 18 months after the foreign filing date (i.e., priority date) for applications seeking international protection that have not yet received a decision. If a patent has a parent application through continuations, its filing date is adjusted to match that of the parent application, in accordance with AIPA rules, which mandate that patent publication occurs 18 months from the application date of the earliest parent application.

Before the implementation of AIPA, patent applications remained confidential and hidden from the public until a patent was granted. If a patent was never granted, the application remained inaccessible to the public. After the enactment of AIPA, the "18-month publication" rule required that patent applications were made public latest at 18 months after their filing date, with inventors having the option to publish before the 18-month mark at their request. Nevertheless, inventors can also "opt out" and keep their applications confidential until the patent is granted, similar to the pre-AIPA process. To opt-out, inventors must file an application for non-publication request before the decision date, see `https://www.uspto.gov/web/offices/pac/mpep/s1122.html`. Applicants with or who intend to apply for foreign filings cannot opt-out. Applicants choosing to opt-out forfeit the right to file foreign counterpart patent applications. This setting therefore enables us to identify firms that possess information on environment-related capabilities but opt to withhold it in a comparative analysis between applicants who choose to disclose their information earlier versus those who delay disclosure.

Appendix B. Additional Tables and Figures

Figure B1
Number of EPA Enforcement Actions

This figure presents the number of EPA enforcement actions faced by U.S. publicly listed firms, by year. EPA enforcement data is obtained from the EPA Enforcement and Compliance History Online (ECHO) database. The sample period is 1987 to 2020.

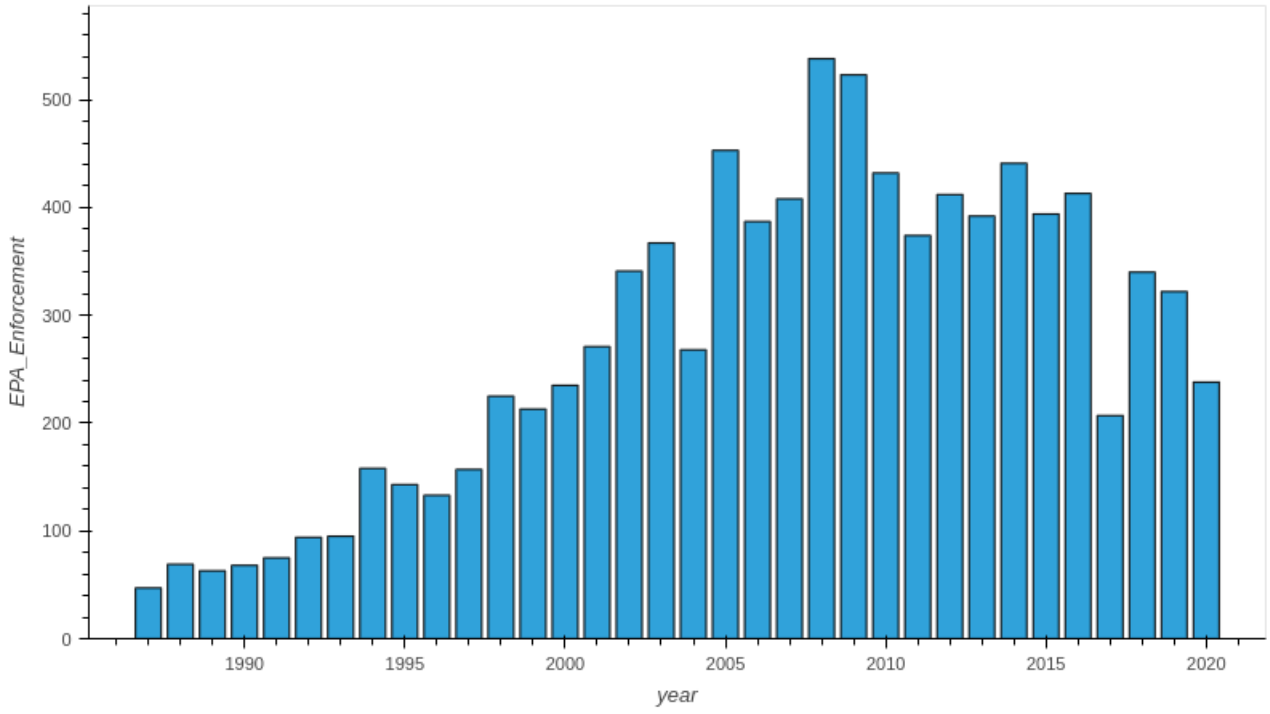


Table B1
Variable Definitions

Variable Name	Abbrev.	Description
Patent Measures		
Clean Patent		Patent defined as abatement technologies using technology classifications [OECD Haščič and Migotto (2015)]
Non-Clean Patent		Patent not defined as abatement patent using technology classifications [OECD Haščič and Migotto (2015)]
N Words		The number of words in the written patent description, in 000s
Readability		The Gunning (1952) FOG index multiplied by negative one of the written patent description
Specificity		The number of specific pieces of information in the written patent description, divided by the total number of words in the written description, multiplied by 1000
Latest Disclosure		The number of days until the patent application must be published (for applications seeking foreign protection, the earlier of 18 months after filing abroad and the patent decision (grant or not) date, and for all others, the application decision date).
Actual Disclosure		The number of days until the USPTO publishes the patent filing, either at the request of the applicant or because the disclosure deadline passes, less 14 weeks for publication delays
Disclosure Latency		The number of days until the disclosure of a patent filing, divided by the number of days until the latest possible disclosure
Chemical Release and EPA Variables		
Total Emission		Total On-Site Releases (pounds) [TRI BASIC PLUS file]
Formal Actions		Formal actions taken by the EPA or local government [EPA ECHO]
HPV		High Priority Violator Status set by the EPA [EPA ECHO]
Firm Characteristics		
G		The stock of clean patents filed by the firm
Market Value	MV	[COMPUSTAT: PRCC_F × CSHO])
Total Assets	TA	[COMPUSTAT: AT])
Sales		[COMPUSTAT: SALE])
MV/TA		Sum of total assets plus market value of equity minus book value of equity divided by total assets [COMPUSTAT (AT+CSHO× PRCC_F - CEQ) / AT)]
Tobin's Q		Sum of total assets plus market value of equity minus book value of equity divided by total assets [COMPUSTAT (AT+CSHO× PRCC_F - CEQ) / AT)]
Cash		[COMPUSTAT CHE/AT]
CAPX		Capital expenditures divided by total assets [COMPUSTAT CAPX/AT]
PPE		Property, plant, and equipment scaled by assets [COMPUSTAT PPENT/AT]
R&D		R&D Research and development expenses divided by total assets [COMPUSTAT XRD/AT]
Missing R&D		Indicator equal to one if R&D Research and development expenses is missing
Employment		Number of employees [COMPUSTAT Employees]

Table B2
Most Cited Clean Patents

The table presents the most-cited clean patents. We present the patent number (*Pat. No.*), title, number of forward citations until 2020 (Citations), and the patent grant date.

Pat. No.	Title	Citations	Grant date
5080556	Thermal seal for a gas turbine spacer disc	946	19920114
6270916	Complete discharge device for lithium battery	486	20010807
5651821	Battery disposal and collection apparatus	483	19970729
6825620	Inductively coupled ballast circuit	386	20041130
6436299	Water treatment system with an inductively coupled ballast	379	20020820
5820646	Inline filter apparatus	379	19981013
4984594	Vacuum method for removing soil contamination utilizing surface electrical heating	363	19910115
6023554	Electrical heater	356	20000208
5318116	Vacuum method for removing soil contaminants utilizing thermal conduction heating	338	19940607

Table B3
R&D-Based Knowledge Spillovers: BSV Approach

This table presents results for the disclosure wedge from clean knowledge spillovers, starting from the approach in BSV. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1 \ln(\text{SPILLTECH}^S)_{i,t} + \beta_2 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} + \beta_4 \ln(R\&DStock/TA)_{i,t} \\ &+ \beta_5 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

SPILLTECH^S is excluded in column (1) and captures spillovers from disclosed innovation (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{it} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity and where G^q_{jt} is the disclosure-quality-weighted stock of patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). G_{jt} includes all patents filed by firm j in column (2) and clean and non-clean patents separately in column (3). $\text{SPILLTECH}^S_{R\&D}$ captures spillovers from R&D investment by technology peers, and $\text{SPILLTECH}^C_{R\&D}$ captures spillovers from R&D investment by product market rivals defined as $\text{SPILLTECH}^C_{R\&D} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} R\&DStock_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^S_{\text{All Pat.}}\text{)}$		-0.021	
		(-1.07)	
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$			0.201***
			(2.58)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$			-0.030
			(-1.46)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.136***	0.142***	0.036
	(4.22)	(4.30)	(0.74)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.009***	-0.008***	-0.008***
	(-6.55)	(-6.39)	(-6.34)
Ln(R&D Stock/TA)	0.226***	0.229***	0.231***
	(7.74)	(7.80)	(7.87)
Ln(G)	0.013	0.015	-0.003
	(0.57)	(0.66)	(-0.13)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75

Table B4
Disclosure Wedge from Clean Knowledge Spillovers: Alternative Quality Weighting

This table presents results for the disclosure wedge from clean knowledge spillovers weighted by disclosure quality, where disclosure quality is captured through patent fullness, readability, or a principal component combination of fullness, readability, and specificity. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by the filing's quantile-transformed IPC3-year-adjusted number of words (columns 1 to 3), readability (Gunning (1952) FOG index multiplied by negative one, columns 4 to 6), and a principal component combination of specificity, fullness, and readability (columns 7 to 9). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	Fullness			Fog			PCA		
	$\text{Ln(MV/TA)}_{i,t+1}$			$\text{Ln(MV/TA)}_{i,t+1}$			$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$	0.190**		0.192**	0.195***		0.198***	0.195***		0.197***
	(2.48)		(2.50)	(2.62)		(2.66)	(2.61)		(2.63)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$		-0.010**	-0.011**		-0.012**	-0.013**		-0.008*	-0.009*
		(-2.00)	(-2.06)		(-2.31)	(-2.41)		(-1.84)	(-1.91)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	-0.029	-0.023	-0.028	-0.025	-0.019	-0.025	-0.028	-0.023	-0.028
	(-1.37)	(-1.12)	(-1.35)	(-1.28)	(-0.99)	(-1.25)	(-1.39)	(-1.13)	(-1.38)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.008***	-0.007**	-0.006**	-0.008***	-0.006**	-0.006**	-0.007***	-0.006**	-0.005**
	(-3.34)	(-2.51)	(-2.43)	(-3.07)	(-2.24)	(-2.14)	(-3.19)	(-2.36)	(-2.28)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.038	0.142***	0.038	0.027	0.142***	0.027	0.032	0.142***	0.032
	(0.79)	(4.30)	(0.79)	(0.54)	(4.30)	(0.53)	(0.66)	(4.31)	(0.66)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.004**	-0.002	-0.002	-0.004**	-0.002	-0.002	-0.004**	-0.002	-0.002
	(-2.05)	(-1.01)	(-1.02)	(-2.18)	(-0.97)	(-0.97)	(-2.36)	(-1.23)	(-1.22)
Ln(R&D Stock/TA)	0.231***	0.229***	0.231***	0.231***	0.229***	0.230***	0.231***	0.229***	0.231***
	(7.86)	(7.81)	(7.87)	(7.85)	(7.80)	(7.86)	(7.86)	(7.81)	(7.86)
Ln(G)	-0.004	0.015	-0.003	-0.006	0.015	-0.005	-0.004	0.015	-0.003
	(-0.16)	(0.67)	(-0.12)	(-0.26)	(0.67)	(-0.22)	(-0.16)	(0.67)	(-0.12)
Ln(Sales)	-0.091***	-0.091***	-0.090***	-0.091***	-0.091***	-0.090***	-0.091***	-0.091***	-0.091***
	(-16.85)	(-16.86)	(-16.82)	(-16.83)	(-16.86)	(-16.81)	(-16.85)	(-16.86)	(-16.83)
Cash	0.700***	0.701***	0.700***	0.700***	0.701***	0.701***	0.700***	0.701***	0.700***
	(24.36)	(24.38)	(24.36)	(24.37)	(24.39)	(24.37)	(24.35)	(24.38)	(24.36)
CAPX	0.826***	0.827***	0.826***	0.826***	0.827***	0.826***	0.826***	0.827***	0.826***
	(10.65)	(10.64)	(10.64)	(10.65)	(10.64)	(10.64)	(10.64)	(10.64)	(10.64)
PPE	-0.241***	-0.243***	-0.241***	-0.240***	-0.243***	-0.240***	-0.241***	-0.243***	-0.241***
	(-4.92)	(-4.97)	(-4.92)	(-4.90)	(-4.96)	(-4.90)	(-4.92)	(-4.97)	(-4.92)
HHI	0.009*	0.009*	0.009*	0.009*	0.009*	0.009*	0.009*	0.009*	0.009*
	(1.90)	(1.91)	(1.89)	(1.90)	(1.91)	(1.89)	(1.90)	(1.91)	(1.89)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	168,575	168,575	168,575	168,575	168,575	168,575	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75

Table B5
Disclosure Wedge from Clean Knowledge Spillovers: Industry Definitions

This table presents results for the disclosure wedge from clean knowledge spillovers using SIC-based industry definitions to identify product market rivals. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). Industry rivals are defined as firms within the same SIC-3 or SIC-4 digit classification, with weights assigned either equally (PRODSIM = $1/N$, columns 1 and 2) or based on product similarity as in Hoberg and Phillips (2016) (column 3). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$	0.192**	0.193**	0.178**
	(2.46)	(2.46)	(2.22)
$\text{Ln(SPILLTECH}^{C,\text{Clean, Equally-Weighted within SIC3}}\text{)}$	-0.007		
	(-1.15)		
$\text{Ln(SPILLTECH}^{C,\text{Clean, Equally-Weighted within SIC4}}\text{)}$		0.003	
		(0.34)	
$\text{Ln(SPILLTECH}^{C,\text{Clean, Weighted by Tech Similarity, within SIC4}}\text{)}$			0.031
			(0.81)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	-0.028	-0.031	-0.031
	(-1.37)	(-1.47)	(-1.49)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.008***	-0.008***	-0.008***
	(-3.35)	(-3.35)	(-3.36)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.040	0.039	0.037
	(0.81)	(0.79)	(0.74)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.004**	-0.004**	-0.004**
	(-1.99)	(-2.01)	(-2.01)
Ln(R&D Stock/TA)	0.231***	0.231***	0.231***
	(7.88)	(7.86)	(7.87)
Ln(Clean Patent Stock)	-0.002	-0.005	-0.008
	(-0.08)	(-0.20)	(-0.34)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75

Table B6
Disclosure Wedge from Clean Knowledge Spillovers: Including Employment

This table presents results for the disclosure wedge from clean knowledge spillovers when including employment in the underlying production function. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =&\beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, HHI, and ln(Employment). T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$	0.186**		0.187**
	(2.37)		(2.39)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$		-0.008	-0.009
		(-1.56)	(-1.62)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	-0.016	-0.012	-0.016
	(-0.75)	(-0.54)	(-0.75)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.006**	-0.005*	-0.005*
	(-2.38)	(-1.76)	(-1.68)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.033	0.133***	0.033
	(0.69)	(4.16)	(0.69)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.002	-0.000	-0.000
	(-0.86)	(-0.15)	(-0.15)
Ln(R&D Stock/TA)	0.215***	0.213***	0.215***
	(7.21)	(7.15)	(7.21)
Ln(Clean Patent Stock)	0.004	0.022	0.004
	(0.16)	(0.98)	(0.19)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	156,917	156,917	156,917
Adj. R^2	0.74	0.74	0.74

Table B7
Disclosure Wedge from Clean Knowledge Spillovers: Private Firms

This table presents results for the disclosure wedge from clean knowledge spillovers attributable to public and private firms. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by public and private technology peers defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(1)	(2)
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$ (Include Private)	0.198**	0.200**
	(2.55)	(2.57)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$		-0.010*
		(-1.90)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	-0.031	-0.030
	(-1.47)	(-1.45)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.008***	-0.007**
	(-3.35)	(-2.50)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.037	0.037
	(0.75)	(0.75)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.004**	-0.002
	(-2.00)	(-1.04)
Ln(R&D Stock/TA)	0.231***	0.231***
	(7.87)	(7.87)
Ln(Clean Patent Stock)	-0.004	-0.003
	(-0.16)	(-0.12)
Controls	✓	✓
Year FE	✓	✓
Firm FE	✓	✓
Obs.	168,575	168,575
Adj. R^2	0.75	0.75

Table B8
Disclosure Wedge from Clean Knowledge Spillovers: Licensed Patents

This table presents results for the disclosure wedge from clean knowledge spillovers attributable to licensed patents. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation obtained through licensing by non-focal licensee firms defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(1)	(2)
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$ (Only Licensed Firm)	-0.042	-0.042
	(-0.60)	(-0.60)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$		-0.010*
		(-1.84)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	-0.025	-0.025
	(-1.22)	(-1.20)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.008***	-0.007***
	(-3.43)	(-2.59)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.143***	0.144***
	(4.32)	(4.34)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.004**	-0.002
	(-1.96)	(-1.03)
Ln(R&D Stock/TA)	0.229***	0.229***
	(7.81)	(7.81)
Ln(Clean Patent Stock)	0.016	0.017
	(0.70)	(0.74)
Controls	✓	✓
Year FE	✓	✓
Firm FE	✓	✓
Obs.	168,575	168,575
Adj. R^2	0.75	0.75

Table B9
Disclosure Wedge from Clean Knowledge Spillovers: Cited Patents

This table presents results for the disclosure wedge from clean knowledge spillovers attributable to patents are cited by firm i. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =&\beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed through patents cited by firm i, defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j that are cited by firm i, weighting G_{jt} by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and total clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(1)	(2)
$\text{Ln(SPILLTECH}^{S,\text{Clean, Cited}}\text{)}$	0.135**	0.136**
	(1.99)	(2.00)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$		-0.010*
		(-1.93)
$\text{Ln(SPILLTECH}^{S,\text{Non-Clean}}\text{)}$	-0.029	-0.029
	(-1.42)	(-1.40)
$\text{Ln(SPILLTECH}^{C,\text{Non-Clean}}\text{)}$	-0.008***	-0.007**
	(-3.38)	(-2.52)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.084*	0.085*
	(1.77)	(1.78)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.004**	-0.002
	(-2.01)	(-1.04)
Ln(R&D Stock/TA)	0.230***	0.230***
	(7.84)	(7.85)
Ln(Clean Patent Stock)	-0.002	-0.001
	(-0.11)	(-0.06)
Controls	✓	✓
Year FE	✓	✓
Firm FE	✓	✓
Obs.	168,363	168,363
Adj. R^2	0.75	0.75

Table B10
Disclosure Wedge from Clean and Dirty Knowledge Spillovers

This table presents results for the disclosure wedge from clean and dirty knowledge spillovers. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} + \beta_1^D \ln(\text{SPILLTECH}^S_{Dirty})_{i,t} \\ &+ \beta_2^D \ln(\text{SPILLTECH}^C_{Dirty})_{i,t} + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \beta_7 \ln(\text{G}^{Dirty})_{i,t} \\ &+ \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean/Dirty}$ capture spillovers from clean/dirty innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean}/Dirty = \sum_{j\neq i} \text{TECH}_{ij} G^{q,clean/dirty}_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. $G^{q,clean/dirty}_{jt}$ is the disclosure-quality-weighted stock of clean/dirty patents filed by firm j. $\text{SPILLTECH}^C_{Clean/Dirty}$ captures spillovers from clean/dirty innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean/Dirty} = \sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^{q,clean/dirty}_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean (G) and dirty (G^{Dirty}) patent stock. X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^{S,\text{Clean}}\text{)}$	0.164**		0.165**
	(2.03)		(2.05)
$\text{Ln(SPILLTECH}^{C,\text{Clean}}\text{)}$		-0.016**	-0.016**
		(-2.40)	(-2.39)
$\text{Ln(SPILLTECH}^{S,\text{Dirty}}\text{)}$	0.096		0.096
	(1.62)		(1.64)
$\text{Ln(SPILLTECH}^{C,\text{Dirty}}\text{)}$		0.002	0.001
		(0.32)	(0.25)
$\text{Ln(SPILLTECH}^{S,\text{R\&D}}\text{)}$	0.033	0.135***	0.033
	(0.67)	(4.19)	(0.67)
$\text{Ln(SPILLTECH}^{C,\text{R\&D}}\text{)}$	-0.009***	-0.005***	-0.005***
	(-6.64)	(-2.88)	(-2.88)
Ln(R&D Stock/TA)	0.228***	0.226***	0.228***
	(7.81)	(7.76)	(7.81)
Ln(Clean Patent Stock)	-0.022	-0.004	-0.022
	(-0.85)	(-0.18)	(-0.84)
Ln(Dirty Patent Stock)	0.010	0.051*	0.013
	(0.28)	(1.66)	(0.34)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75

Table B11
Clean Innovation Knowledge Spillovers and Toxic Emissions

This table presents regression results for the relation between emissions and knowledge spillovers from other firms' disclosed clean patents. We estimate the following specification at the facility-year level:

$$ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Acquired Knowledge}_{i,t} + \beta_2 \text{Non-Clean Acquired Knowledge}_{i,t} + \Gamma X_{i,f,t} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}.$$

Total Emissions are the facility's total onsite emissions in years $t + \tau$ for $\tau = 1$ to $\tau = 5$. *Clean Acquired Knowledge* is the number of backward citations by the focal firm i's patents to other firms' clean patents and *Non-Clean Acquired Knowledge* is the number of backward citations to other firms' non-clean patents. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI that are suppressed for brevity. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level. The sample period is 1988 to 2020.

	(t+1)	(t+2)	(t+3)	(t+4)	(t+5)
	(1)	(2)	(3)	(4)	(5)
Clean Acquired Knowledge	-0.051**	-0.059**	-0.079***	-0.090***	-0.077***
	(-2.17)	(-2.22)	(-2.84)	(-3.81)	(-3.45)
Non-Clean Acquired Knowledge	0.033	0.050*	0.057*	0.067**	0.068**
	(1.18)	(1.71)	(1.87)	(2.30)	(2.13)
Facility FE	✓	✓	✓	✓	✓
Industry × Year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Obs.	110,204	98,362	88,100	79,185	71,475
Adj. R^2	0.86	0.86	0.86	0.87	0.87

Table B12
Clean Innovation Knowledge Spillovers and Toxic Emissions: Cross-Sectional Results

This table presents regression results for the relation between emissions and knowledge spillovers from other firms' disclosed clean patents. The analysis is at facility-year level.

$ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Acq. Know.}_{i,t} + \beta_2 \text{Non-Clean Acq. Know.}_{i,t} + \Gamma X_{i,f,t} + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}$.

Total Emissions are a facility's total onsite emissions in years $t + \tau$ for $\tau = 1$ to $\tau = 5$. *Clean Acquired Knowledge* is the number of backward citations by the focal firm i's patents to other firms' clean patents and *Non-Clean Acquired Knowledge* is the number of backward citations to other firms' non-clean patents. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI, which are omitted for brevity. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level. The sample period is 1988 to 2020.

	(t+1)	(t+2)	(t+3)	(t+4)	(t+5)
	(1)	(2)	(3)	(4)	(5)
Clean Acquired Knowledge	-0.083**	-0.091***	-0.099***	-0.101***	-0.100***
	(-2.54)	(-2.69)	(-2.89)	(-2.93)	(-2.94)
Non-Clean Acquired Knowledge	0.088	0.089	0.086	0.087	0.092
	(1.49)	(1.42)	(1.35)	(1.34)	(1.39)
Industry × Year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Obs.	111,693	99,618	89,400	80,409	72,484
Adj. R^2	0.48	0.48	0.48	0.47	0.47

Table B13
Clean Patent Filings and Toxic Emissions

This table presents regression results for the relation between firms' facility-level toxic emissions and their own clean innovation disclosure through patent filings. The analysis is conducted at the facility-year level.

$$\ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Patents}_{i,f,t} + \Gamma X_{i,t} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}.$$

Total Emissions represents the pounds of air, water, and land pollutants emitted by facility f of firm i in year $t+\tau$ for $\tau = 1$ to $\tau = 5$. *Clean Patents* denotes the number of clean patents filed by facility f's parent firm at time t. X is a vector of control variables which includes: ln(Sales), Cash, Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI that are suppressed for brevity. T-statistics, adjusted for clustered standard errors at the firm level, are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020.

	t+1	t+2	t+3	t+4	t+5
	(1)	(2)	(3)	(4)	(5)
Clean Patents	-0.117***	-0.106***	-0.114***	-0.094**	-0.074
	(-3.65)	(-3.41)	(-2.76)	(-2.06)	(-1.48)
Controls	✓	✓	✓	✓	✓
Facility FE	✓	✓	✓	✓	✓
Industry × Year FE	✓	✓	✓	✓	✓
Obs.	50,016	46,014	42,118	38,671	35,398
Adj. R^2	0.84	0.84	0.85	0.85	0.86

Table B14
Summary Statistics for EPA Sample

This table presents summary statistics for the variables employed in our analysis using EPA actions. We report variables at the facility-year, firm-year, and patent level. Each sample is limited to firms with at least one clean patent filing and no missing control variables. Variables are defined in Appendix B1. Penalty is conditional on the facility receiving formal actions. The sample period is 1988 to 2020 for all variables, except for latency measures, which start post-AIPA in January 2001.

	N	Mean	Median	Std. Dev
Facility-Year				
Formal Actions	163,147	0.03	0.00	0.17
Penalty	4,813	150,203.87	6,000.00	876,137.76
Total Emission	163,147	121,495.55	700.00	1,421,631.48
Clean Acquired Knowledge	163,147	60.67	0.00	273.70
Non-Clean Acquired Knowledge	163,147	1,424.11	50.00	4,087.23
Firm-Year				
Clean Patents	23,861	0.72	0.00	3.65
Non-Clean Patents	23,861	37.78	0.00	138.01
Formal Actions	23,861	0.04	0.00	0.14
Patent level				
Ln(Disclosure Latency)	23,410	5.95	6.30	0.64
Ln(Days to Mandatory Disclosure)	23,410	6.33	6.55	0.83
Percentage Latency	23,410	0.68	0.68	0.31
Disclosure Latency	23,410	461.97	547.00	272.62
Days to Mandatory Disclosure	23,410	759.50	700.00	542.55

Table B15
EPA-Induced Clean Innovation Knowledge Spillovers and Toxic Emissions

This table presents regression results for the relation between emissions and knowledge spillovers from clean patents filed in the year of or the year following EPA actions. We estimate the following specification at the facility-year level:

$$ln(\text{Total Emissions})_{i,f,t+\tau} = \beta_0 + \beta_1 \text{Clean Acquired Knowledge}_{i,t} + \beta_2 \text{Non-Clean Acquired Knowledge}_{i,t} + \Gamma X_{i,f,t} + \lambda_f + \lambda_{ind,t} + \epsilon_{i,f,t+\tau}.$$

Total Emissions are the facility's total onsite emissions in years $t + \tau$ for $\tau = 1$ to $\tau = 5$. *Clean Acquired Knowledge* is the number of backward citations by the focal firm i's patents to other firms' clean patents and *Non-Clean Acquired Knowledge* is the number of backward citations to other firms' non-clean patents. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI that are suppressed for brevity. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% level. The sample period is 1988 to 2020.

	(t+1)	(t+2)	(t+3)	(t+4)	(t+5)
	(1)	(2)	(3)	(4)	(5)
Clean Acquired Knowledge	-0.044*	-0.046*	-0.060**	-0.075***	-0.066***
	(-1.69)	(-1.72)	(-2.21)	(-3.06)	(-2.84)
Non-Clean Acquired Knowledge	0.025	0.039	0.042	0.054*	0.058*
	(0.94)	(1.39)	(1.45)	(1.91)	(1.88)
Facility FE	✓	✓	✓	✓	✓
Industry × Year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
Obs.	110,204	98,362	88,100	79,185	71,475
Adj. R^2	0.86	0.86	0.86	0.87	0.87

Table B16
Disclosure through 10-K/Q

This table reports the results of regressing firms' references to clean technology in their 10-K/Q filings. The analysis is conducted at the firm level.

$$10K/Q_{i,t+\tau} = \beta_0 + \beta_1 \text{Clean Patent}_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t}.$$

We search for the co-occurrence of terminology relating to 'clean' and 'technology' within the same sentence in the 10-K/Q text. Clean terms include clean, green, environmental, sustainable, pollution, waste, toxic, emission, contamination, eco-friendly, energy-efficient, recyclable, and biodegradable, and their variations. Technology terms include technology, technologies, innovation, patent, research, and solution, and their variations. $10K/Q_{i,t+\tau}$ is an indicator variable equal to one if the firm mentions co-occurrences of clean and technology terms measured at time τ, and zero otherwise. We report β_1 for $\tau = -2, -1, 0, 1, 2$. *Clean Patent* is the number of clean patents filed in year t, λ_i are firm fixed effects and λ_t are year fixed effects. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, CAPX, PPE, R&D, missing R&D, and HHI, which are omitted for brevity. Standard errors are clustered at the firm level. The sample period is 1993 to 2020.

	I(Disclosure of Clean Tech in 10-K/Q)				
	t-2	t-1	t	t+1	t+2
	(1)	(2)	(3)	(4)	(5)
Clean Patents	0.323	0.390	0.698***	0.832***	0.917***
	(1.11)	(1.53)	(2.89)	(3.31)	(3.34)
Controls	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Obs.	17,656	19,297	21,325	19,367	17,760
Adj. R^2	0.53	0.53	0.53	0.54	0.54

Appendix C. Online Appendix Alternative Disclosure Quality Measures

Table C1
Disclosure Wedge from Clean Process-Based Spillovers: Alternative Quality Weighting

This table presents results for the disclosure wedge from clean process-based knowledge spillovers weighted by disclosure quality in terms of patent fullness, readability, or a principal component combination of specificity, fullness, and readability. We estimate firm-year panel regressions of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process/Product})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process/Product}$ captures spillovers from clean innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean,Process/Product} = \sum_{j\neq i} \text{TECH}_{ij} G^{q,Process/Product}_{jt}$, where Process/Product-Based patents are defined following Tham et al. (2025); TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j, weighting G_{jt} by the filing's quantile-transformed IPC3-year-adjusted number of words (columns 1 to 3), readability (Gunning (1952) FOG index multiplied by negative one, columns 4 to 6), and a principal component combination of specificity, fullness, and readability (columns 7 to 9). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). X is a vector of control variables including ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, ** indicate significance at 1%, 5% and 10% level. The sample period is 1988 to 2020. The sample includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	Fullness			Fog			PCA		
	$\text{Ln(MV/TA)}_{i,t+1}$			$\text{Ln(MV/TA)}_{i,t+1}$			$\text{Ln(MV/TA)}_{i,t+1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}})$	0.190***		0.190*	0.186***		0.161*	0.194***		0.205**
	(2.68)		(1.82)	(2.84)		(1.69)	(2.83)		(2.05)
$\text{Ln(SPILLTECH}^S_{\text{Clean, Product}})$		0.157**	0.000		0.178**	0.037		0.152**	-0.016
		(2.02)	(0.00)		(2.24)	(0.32)		(1.99)	(-0.14)
$\text{Ln(SPILLTECH}^C_{\text{Clean}})$	-0.011**	-0.011**	-0.011**	-0.013**	-0.013**	-0.013**	-0.009*	-0.009*	-0.009*
	(-2.09)	(-2.03)	(-2.09)	(-2.44)	(-2.36)	(-2.44)	(-1.94)	(-1.87)	(-1.94)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}})$	-0.029	-0.027	-0.029	-0.025	-0.023	-0.025	-0.029	-0.026	-0.029
	(-1.39)	(-1.28)	(-1.39)	(-1.28)	(-1.19)	(-1.28)	(-1.42)	(-1.30)	(-1.42)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}})$	-0.006**	-0.006**	-0.006**	-0.006**	-0.006**	-0.006**	-0.005**	-0.005**	-0.005**
	(-2.42)	(-2.44)	(-2.42)	(-2.14)	(-2.15)	(-2.13)	(-2.27)	(-2.29)	(-2.27)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}})$	0.034	0.066	0.034	0.028	0.050	0.024	0.027	0.065	0.028
	(0.70)	(1.41)	(0.69)	(0.57)	(1.01)	(0.47)	(0.56)	(1.37)	(0.58)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}})$	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-1.01)	(-1.03)	(-1.01)	(-0.95)	(-0.98)	(-0.96)	(-1.21)	(-1.23)	(-1.21)
Ln(R&D Stock/TA)	0.231***	0.231***	0.231***	0.230***	0.231***	0.231***	0.231***	0.230***	0.231***
	(7.87)	(7.86)	(7.86)	(7.86)	(7.86)	(7.86)	(7.86)	(7.85)	(7.86)
Ln(G)	-0.006	0.001	-0.006	-0.008	-0.001	-0.009	-0.007	0.003	-0.007
	(-0.26)	(0.06)	(-0.26)	(-0.35)	(-0.03)	(-0.36)	(-0.28)	(0.11)	(-0.27)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	168,575	168,575	168,575	168,575	168,575	168,575	168,575	168,575	168,575
Adj. R^2	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75

Table C2
Disclosure Wedge from Clean Knowledge Spillovers: Fullness-Weighted Patents - IV

This table presents results for the disclosure wedge from clean fullness-weighted knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by technology peers defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the patent fullness (number of words)-weighted stock of clean patents filed by firm j. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firms's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.192***	-0.015	
	(11.94)	(-0.72)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.325***	
	(-1.09)	(39.98)	
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$			0.351***
			(2.77)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.030
			(1.63)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.013***	0.041***	-0.034
	(5.29)	(2.71)	(-1.60)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.158***	-0.013***
	(-0.80)	(52.97)	(-3.22)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.252***	0.038	-0.051
	(5.88)	(1.08)	(-0.70)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.000	0.132***	-0.008***
	(-0.18)	(56.60)	(-2.63)
Ln(R&D Stock/TA)	-0.002	-0.019	0.233***
	(-1.07)	(-1.07)	(7.89)
Ln(G)	0.014	0.082***	-0.022
	(1.50)	(3.48)	(-0.84)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			799

Table C3
Disclosure Wedge from Clean Process-Based Knowledge Spillovers: Fullness-Weighted Patents - IV

This table presents results for the disclosure wedge from clean fullness-weighted process-based knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean,Process} = \sum_{j\neq i} \text{TECH}_{ij} G^{q,process}_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. $G^{q,process}_{jt}$ is the patent fullness (number of words)-weighted stock of clean patents filed by firm j where *process* is defined as in Tham et al. (2025). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean,Process}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.198***	-0.015	
	(11.56)	(-0.72)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.325***	
	(-0.79)	(39.98)	
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$			0.340***
			(2.75)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.030
			(1.62)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.017***	0.041***	-0.035*
	(6.08)	(2.71)	(-1.65)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.158***	-0.013***
	(-0.33)	(52.97)	(-3.22)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.271***	0.038	-0.054
	(5.90)	(1.08)	(-0.72)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.000	0.132***	-0.008***
	(-0.11)	(56.60)	(-2.63)
Ln(R&D Stock/TA)	-0.002	-0.019	0.233***
	(-0.86)	(-1.07)	(7.89)
Ln(G)	0.031***	0.082***	-0.028
	(3.22)	(3.48)	(-1.02)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			799

Table C4
Disclosure Wedge from Clean Knowledge Spillovers: Readability-Weighted Patents - IV

This table presents results for the disclosure wedge from clean readability-weighted knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the patent readability (FOG Index)-weighted stock of clean patents filed by firm j. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.208***	-0.010	
	(12.54)	(-0.48)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.313***	
	(-0.94)	(40.29)	
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$			0.323***
			(2.76)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.031
			(1.64)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.013***	0.048***	-0.030
	(5.71)	(3.52)	(-1.52)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.143***	-0.012***
	(-0.88)	(49.75)	(-3.12)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.271***	0.043	-0.050
	(6.16)	(1.27)	(-0.69)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	0.000	0.134***	-0.009***
	(0.10)	(57.00)	(-2.69)
Ln(R&D Stock/TA)	-0.001	-0.012	0.232***
	(-0.65)	(-0.70)	(7.86)
Ln(G)	0.017**	0.080***	-0.023
	(2.01)	(3.42)	(-0.87)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			812

Table C5
Disclosure Wedge from Clean Process-Based Knowledge Spillovers: Readability-Weighted Patents - IV

This table presents results for the disclosure wedge from clean readability-weighted process-based knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean,Process} = \sum_{j \neq i} \text{TECH}_{ij} G^{q,process}_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. $G^{q,process}_{jt}$ is the patent readability (FOG Index)-weighted stock of clean patents filed by firm j where *process* is defined as in Tham et al. (2025). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean,Process}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.219***	-0.010	
	(12.29)	(-0.48)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.313***	
	(-0.66)	(40.29)	
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$			0.308***
			(2.74)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.031
			(1.63)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.017***	0.048***	-0.031
	(6.39)	(3.52)	(-1.57)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.143***	-0.012***
	(-0.29)	(49.75)	(-3.12)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.288***	0.043	-0.051
	(6.09)	(1.27)	(-0.69)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	0.000	0.134***	-0.009***
	(0.02)	(57.00)	(-2.69)
Ln(R&D Stock/TA)	-0.001	-0.012	0.232***
	(-0.54)	(-0.70)	(7.86)
Ln(G)	0.037***	0.080***	-0.029
	(3.78)	(3.42)	(-1.05)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			812

Table C6
Disclosure Wedge from Clean Knowledge Spillovers: PCA-Weighted Patents - IV

This table presents results for the disclosure wedge from clean PCA-weighted knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean} = \sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. G^q_{jt} is the patent PCA-weighted stock of clean patents filed by firm j. $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.190***	-0.012	
	(10.97)	(-0.52)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.356***	
	(-1.14)	(41.97)	
$\text{Ln(SPILLTECH}^S_{\text{Clean}}\text{)}$			0.356***
			(2.81)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.027
			(1.63)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.013***	0.042**	-0.034
	(5.28)	(2.51)	(-1.63)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.175***	-0.012***
	(-0.73)	(56.64)	(-3.05)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.274***	0.042	-0.060
	(5.94)	(0.98)	(-0.80)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.000	0.174***	-0.010***
	(-0.21)	(67.93)	(-2.68)
Ln(R&D Stock/TA)	-0.002	-0.030	0.233***
	(-0.99)	(-1.55)	(7.90)
Ln(G)	0.013	0.088***	-0.022
	(1.44)	(3.42)	(-0.84)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			881

Table C7
Disclosure Wedge from Clean Process-Based Knowledge Spillovers: PCA-Weighted Patents - IV

This table presents results for the disclosure wedge from clean PCA-weighted process-based knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions. The second-stage specification is a firm-year panel regression of the form:

$$\begin{aligned}\ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} + \beta_2^C \ln(\text{SPILLTECH}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&DStock/TA)_{i,t} + \beta_6 \ln(\text{G})_{i,t} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ captures spillovers from clean innovation disclosed (through patenting) by other firms defined as $\text{SPILLTECH}^S_{Clean,Process} = \sum_{j \neq i} \text{TECH}_{ij} G^{q,process}_{jt}$, where TECH_{ij} is Jaffe (1986)'s measure of technology proximity. $G^{q,process}_{jt}$ is the patent PCA-weighted stock of clean patents filed by firm j where *process* is defined as in Tham et al. (2025). $\text{SPILLTECH}^C_{Clean}$ captures spillovers from clean innovation disclosed by product market rivals defined as $\text{SPILLTECH}^C_{Clean} = \sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm's R&D stock, scaled by total assets, and clean patent stock (G). In the first stage, we instrument $\text{SPILLTECH}^S_{Clean,Process}$ and $\text{SPILLTECH}^C_{Clean}$ by replacing G^q_{jt} with G^{ENF}_{jt}. ENF is defined as the number of EPA enforcements the firms j received in the past. X is a vector of control variables including: ln(Sales), Cash, CAPX, PPE, and HHI. T-statistics adjusted for clustered standard errors at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 1988 to 2020 and includes all firms in the CRSP-COMPUSTAT merged database with non-missing product similarity data.

	First Stage	First Stage	Second Stage
	$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$	$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLENF}^S\text{)}$	0.196***	-0.012	
	(10.69)	(-0.52)	
$\text{Ln(SPILLENF}^C\text{)}$	-0.001	0.356***	
	(-0.88)	(41.97)	
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$			0.345***
			(2.78)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$			0.027
			(1.62)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	0.017***	0.042**	-0.035*
	(6.10)	(2.51)	(-1.69)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.000	0.175***	-0.012***
	(-0.26)	(56.64)	(-3.06)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.299***	0.042	-0.065
	(6.07)	(0.98)	(-0.84)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.000	0.174***	-0.010***
	(-0.14)	(67.93)	(-2.67)
Ln(R&D Stock/TA)	-0.002	-0.030	0.233***
	(-0.78)	(-1.55)	(7.90)
Ln(G)	0.031***	0.088***	-0.028
	(3.21)	(3.42)	(-1.02)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	168,575	168,575	168,575
R^2			0.04
Kleibergen-Paap Wald F			881