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

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

Abstract

Do innovation disclosure decisions affect knowledge spillovers independently of R&D investment? We reinterpret patents as *disclosed* innovation, not just innovative output: a firm keeping trade secrets contributes less to publicly available knowledge regardless of R&D investment. We show this distinction matters fundamentally for process-based environmental technologies, which are harder to reverse-engineer than product innovation, making patent disclosure a first-order spillover channel. Within the public-firm valuation margin we observe, we find a 2.04% wedge between spillovers and competitive losses. The effect is driven entirely by process-based clean patents; product-based clean patents and other patents show no incremental spillovers beyond R&D investment. This is a lower-bound estimate, because total social benefits, including 5–9% lower peer-facility emissions over five years and spillovers to private firms exceed the market-value margin we measure. Current climate policy addresses underinvestment but ignores under-disclosure, leaving substantial social value untapped.

Keywords: Knowledge Spillovers; Process Innovation; Reverse-engineering; Environmental Innovation; Disclosure; Social 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, Emirhan Ilhan, Hao Liang, Kai Li, Stefano Lovo, Ralf Martin, Pedro Matos, Lilian Ng, Gordon Phillips, Kristina Marie Rennekamp, 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).

Introduction

Knowledge spillovers are a central source of social returns to innovation. A large literature, from Arrow (1962) and Griliches (1979) to Bloom, Schankerman, and Van Reenen (2013), documents that firms benefit from other firms' R&D investments, generating an investment wedge between social and private returns that motivates innovation policy. Yet, this literature treats spillovers as arising through reverse engineering, labor mobility, and informal channels, all of which operate regardless of firm public disclosure. Whether disclosure decisions independently shape knowledge diffusion has received little attention, despite its importance for measuring spillovers, understanding their sources, and designing policy.

We argue that disclosure is a first-order channel for a large and economically important class of innovations: process-based technologies. For product innovations, disclosure matters less. When a consumer electronics firm introduces a new device, competitors can disassemble its components and recover the underlying technology. Disclosure through patenting accelerates knowledge transfer that would occur anyway, while protecting the firm's intellectual property. Process innovations, typically embedded inside production facilities, are different. Without explicit disclosure, process-based knowledge remains hidden inside the innovating firm. This distinction is mirrored in firms' appropriation choices: secrecy is often viewed as more effective than patents for process innovations, reflecting their hidden nature and the informational costs of patenting (Arundel, 2001; Png, 2017; Mezzanotti and Simcoe, 2025). Process innovations are therefore both harder to infer and less likely to be patented, making patent disclosure a crucial channel for knowledge spillovers.

Environmental innovation provides an ideal empirical laboratory to test the disclosure channel of knowledge spillovers. Environmental technologies are disproportionately process-based: we document that 60% of environmental patents are process-based, compared to 43% of traditional patents. Moreover, environmental process innovation exhibits dual unobservability: not only is the technology embedded inside production facilities, but also the primary outcome (e.g., a reduction in emissions) does not reveal anything about the technology that achieved it. A faster production method developed by a steel manufacturer can be inferred by observable outputs that competitors can study. By contrast, consider selective catalytic reduction (SCR) for NO_X abatement, a process innovation used in coal power generation, cement production, and marine shipping. The emitted output (lower NO_X) does not reveal the underlying catalyst design or operating conditions; the technology stays hidden inside facilities absent explicit disclosure. The unobservability of both the technology and its primary outcome means that disclosure through patenting is especially consequential for environmental knowledge diffusion. We use this setting to identify a disclosure channel that

is independent of, and incremental to, the R&D investment channel in Bloom, Schankerman, and Van Reenen (2013) (BSV hereafter). BSV implicitly assumes knowledge is accessible once produced; we test whether accessibility itself is a margin. Our mechanism implies that disclosure should have no incremental effect when technologies are readily inferable; we test this prediction by comparing process-based and product-based innovations.

We reinterpret patents as a measure of *disclosed* innovation, not just a measure of innovative output. These are not the same thing. For process-based technologies, which are harder to infer from observed output, informal diffusion is weak and spillovers depend on disclosure. Disclosure without innovation generates no knowledge to transfer. Patents provide an observable proxy that jointly reflects inventive output and public disclosure. In exchange for a temporary monopoly, a patent must disclose the invention in sufficient detail that a person skilled in the relevant field can replicate it, a requirement enshrined in patent law's enablement doctrine (35 U.S.C. § 112(a)).[1] This makes patents fundamentally different from R&D expenditure as a measure of innovation: R&D stock captures investment irrespective of knowledge creation or diffusion, while patents capture both the knowledge disclosure decision and the knowledge itself. A firm that spends heavily on R&D but retains its innovations as trade secrets contributes substantially less to the public knowledge stock than one that patents. 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 augment BSV by introducing a disclosure-based spillover measure constructed from quality-weighted environmental patents, varying along both the extensive margin (patent versus trade secrecy) and the intensive margin (disclosure quality). If patents are a noisy proxy for R&D, disclosure quality would be irrelevant for generating knowledge spillovers, and there should be no observable differences in spillovers generated from clean and traditional patent disclosure. Neither is true in the data.

Our patent-based spillover coefficients allow us to quantify what we term the disclosure wedge: the gap between the spillover benefits that disclosed knowledge confers on technology peers and the competitive costs it imposes on product-market rivals. This wedge is conceptually distinct from the innovation investment wedge in the classic literature (Arrow,

[1]The enablement requirement reflects the fundamental quid pro quo of the patent system. As the US Supreme Court noted in *Kewanee Oil Co. v. Bicron Corp.* (1974; No. 73-187): "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."

1962): the innovation investment wedge refers to whether firms invest enough in R&D, while the disclosure wedge concerns whether firms, conditional on innovating, make that knowledge publicly available. This disclosure wedge captures the incremental effect of making innovations publicly available, above and beyond what R&D investment alone generates.

Our baseline estimates show 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 in the absence of public 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, consistent with an unobservability mechanism. Competitive effects from rivals' disclosure are economically negligible, generating a disclosure wedge of 2.04%. These effects translate beyond firm value. A one standard deviation increase in clean acquired knowledge is associated with a 5-9% reduction in peer facility-level toxic emissions over five years. In contrast, equivalent non-clean knowledge acquisition is associated with emission increases, consistent with traditional innovation enabling production expansion. The joint economic and environmental returns from clean knowledge spillovers underscore that the disclosure wedge has consequences beyond what firm value estimates alone can capture.

The central identification challenge is separating disclosure decisions from underlying innovation investment. We address this in three ways. First, by controlling for BSV's R&D spillovers, we ensure 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): 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. Enforcement actions target specific facilities for specific violations. Regulatory pressure increases firms' incentives to patent clean technologies as a compliance signal (Berrone et al., 2013), but is unlikely to directly change the spillover productivity of that technology for non-targeted firms. This generates exogenous variation in the stock of publicly available clean knowledge. Our identifying assumption is not that EPA actions have no private consequences for targeted firms, but that, conditional on the underlying innovation, they do not directly change the spillover value of that knowledge to non-targeted firms except through disclosure. The variation we exploit is in the availability of clean knowledge, not in its value.

We extensively validate the instrument's relevance and exclusion restrictions before turning to IV estimates. EPA actions immediately increase clean patent filings in the enforcement

year. In contrast, non-clean patent filings and R&D investments remain unchanged, while capital expenditures show a delayed response. These results reduce concerns that regulatory pressure reflects an increase in innovation investment rather than a disclosure decision. Targeted firms not only increase patent filings, but, conditional on patenting, also accelerate the *speed* of patent application disclosure following enforcement actions. We show that EPA-induced clean patents generate measurable signaling benefits for regulatory outcomes. We further show that downstream *emissions* fall for firms that cite EPA-induced clean patents.

Our IV estimates confirm and strengthen the baseline findings. A 10% increase in instrumented clean knowledge from technologically related firms increases focal firm value by 3.57%, nearly double the baseline. Competitive effects from rivals' instrumented patent disclosure are economically small. Traditional innovation disclosure generates no incremental spillovers beyond R&D investment. The disclosure wedge of 3.3% is driven by process-based clean patents, consistent with the unobservability mechanism.

These findings reveal a disclosure-specific externality distinct from the R&D under-investment problem at the center of the innovation policy literature. Firms that innovate in process-based environmental technologies generate large knowledge spillovers when they disclose. However, they cannot internalize these benefits, leading to under-disclosure independent of under-investment. The social return to disclosure substantially exceeds the private return. Even firms investing optimally in clean R&D may rationally choose secrecy, leaving substantial social value untapped. Existing policy instruments do not address this margin. More broadly, this disclosure channel is likely relevant for any innovation whose process or outcome is unobservable from the market, a feature that extends beyond environmental technology.

We extend a growing literature on knowledge spillovers and innovation (e.g., Moretti, 2021; Matray, 2022; Giroud, Liu, and Mueller, 2026) by documenting a distinct disclosure-based spillover channel that complements the R&D investment channel in BSV. For process-based environmental technologies, patent disclosure is not an accelerator of knowledge diffusion that would occur regardless, but a necessary condition for between-firm knowledge diffusion. This generates a disclosure-specific externality distinct from the R&D underinvestment problem. Firms can under-disclose even when they invest optimally in R&D. The two distortions require different policy instruments. Unlike BSV, which studies spillovers from innovation investment, we isolate a disclosure margin that matters only in settings where knowledge is not otherwise observable. For policy, R&D subsidies correct under-investment; for process-based innovations, under-disclosure is the binding constraint on whether that investment generates social returns.

We also contribute to the literature on sustainable finance and corporate environmental

action. Prior work studies how capital market frictions shape firms' green investment decisions (Krueger, Sautner, and Starks, 2020; Bolton and Kacperczyk, 2021; Flammer, 2021; Ilhan et al., 2023; Heeb et al., 2023; Sautner et al., 2023; Garel et al., 2024; Giglio et al., 2026). We study a complementary margin: the supply of environmental knowledge, specifically, which clean technologies become publicly available for adoption by other firms and what competitive incentives govern this decision. Market incentives substantially undersupply publicly available clean knowledge relative to the social optimum, indicating that the pace of the green transition depends not only on how much firms invest in clean innovation but on how much they choose to disclose.

1 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 using 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), who decompose spillovers into technology spillovers (knowledge externalities) and competitive (product market rivalry) effects. However, our focus differs fundamentally from BSV: whereas BSV examine spillovers from R&D investment (measured by R&D expenditure), we examine spillovers from disclosure decisions, i.e., the choice to make innovations publicly available through patenting as opposed to trade secrecy, and the informativeness of that disclosure. This distinction is critical for environmental process innovation, where 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.

1.1 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 R&D investment (as in BSV) and publicly disclosed innovations (patents). We can formalize the role of knowledge spillovers in shaping firm output as:

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

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

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

Empirically, we capture A_i through firm i's own innovation stock (R&D and patents), and A_j through the knowledge stock of other firms j. PROXIM_{ij} is the proximity between firm i and firm j in the technology and product market space, a weighting factor for spillovers. This formulation explicitly recognizes that knowledge spillovers occur through accessibility of knowledge: a firm that conducts R&D but maintains trade secrecy contributes substantially less to the public knowledge stock than a firm that publicly documents its innovation, limiting spillovers. Patents capture this disclosure margin while R&D investment does not.

We assume that firm i's output is proportional to its market value MV_i. Specifically, market value reflects the present discounted value of future output, assuming Y_i grows at a constant rate equal to the discount rate r such that $MV_i = \frac{Y_i}{r}$.

1.2 Empirical Specification

Our empirical strategy builds on BSV. 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. We assume r is constant across firms and time, so that $ln(r)$ is absorbed by the constant in our baseline specification:

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

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 firms j get a higher weight, as firm i may benefit more from their

knowledge creation. $SPILLTECH^{C}_{Clean}$ is the aggregate disclosed clean knowledge capital of all direct product market competitors j, weighted by their product market similarity to firm i (PRODSIM$_{ij}$).

Our specification controls for all R&D-based spillovers including: (i) technology spillovers from technology peers' R&D investments (SPILLTECH$^{S}_{R\&D}$), (ii) product market rivalry effects from competitors' R&D investments (SPILLTECH$^{C}_{R\&D}$), and (iii) firm i's own R&D stock ($R\&DStock/TA$), as well ass firm i's own clean patent stock (G). Thus, we ensure that our patent-based spillover coefficients (β_1, β_2) capture disclosure effects beyond BSV's innovation investment channel. That is, conditional on all firms' R&D investments and the spillovers they generate, our coefficients measure the incremental effect of making innovations publicly available through patent disclosure and the informativeness of that disclosure.

1.2.1 Construction of SPILLTECH Measures

To calculate $SPILLTECH^{S}_{Clean}$ and $SPILLTECH^{C}_{Clean}$, we measure disclosed clean knowledge capital ($G^{q}_{j,t}$) using the quality-weighted stock of clean patents filed by firms j.[2] 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 stock 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 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,

[2]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), see Section 2 for details.

but differ substantially in how much actionable knowledge they reveal through the quality of disclosed information (Wang, 2024). 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 patents), 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 $G^q_{j,t}$ reflects differences in disclosure decisions (whether to patent) and disclosure informativeness (how much actionable knowledge is revealed in patents). Conditioning on own R&D stock and all R&D-based spillovers absorbs variation in innovation investment, isolating the component of patent stock that reflects disclosure choices conditional on investment level.

To calculate $SPILLTECH^S_{R\&D}$ and $SPILLTECH^C_{R\&D}$, we follow BSV and use cumulative R&D stock as the innovation investment measure. The R&D stock is constructed using the perpetual inventory method: $R\&DStock_{i,t} = (1-\delta)R\&DStock_{i,t-1} + R\&D_{i,t}$, 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 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 $T_i = (T_{i1}, T_{i2}, \ldots, T_{ik})$ where T_{ik} is the share of firm i's ENV-TECH patents in 3-digit IPC class k for 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). The measure is constructed analogously for all patents and non-clean patents.

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 a more granular and objective competitive overlap than Standard Industry Classification (SIC). We exclude all firms with missing product similarity data.

[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.

1.2.2 Other Controls and Fixed Effects

$G_{i,t}$ is defined as firm i's own clean patent stock to capture direct effects from firms' own clean knowledge capital, and $R\&DStock/TA$ captures firm i's own innovation investment, scaled by total assets. $X_{i,t}$ 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. We assume the combination of fixed effects and control variables captures the elasticity of firm i's market value with respect to the other inputs in the production function (i.e., capital and labor).

1.2.3 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.

Competitive Effects (β_2). β_2 captures the effect of product market rivals' clean disclosure on firm value, capturing the private costs and benefits that disclosing firms j take into account when deciding whether to patent. These private returns reflect disclosing 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). Importantly, competitive effects need not arise from immediate product market competition. Disclosure of clean technologies may affect rivals' valuations through regulatory benchmarking, expectations about future compliance costs, or investor reassessment of technological positioning.

In Equation (2), a loss to disclosing (patenting) firms j mirrors gain to other firms i. If disclosed innovation by j can be easily appropriated by firms i, j suffers the classic

proprietary information loss while i gains, generating $\beta_2 > 0$. If instead j's private benefits from disclosure dominate j's proprietary loss, j captures positive private returns and i loses out, generating $\beta_2 < 0$. The sign β_2 reveals which side of j's disclosure trade-off dominates on average across the sample.

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 ($\beta_1 - \beta_2 > 0$) 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 close the gap.[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 returns to own innovation investment ($R\&DStock$), scaled by total assets, allowing us to distinguish the impact of innovation investment from disclosure. β_6 captures the elasticity of firm i's market value to its own clean patent stock ($G_{i,t}$), reflecting the net value of own clean patenting, including IP protection, signaling, and direct returns to clean innovation.

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. Our complementary analysis using facility-level emissions data provides additional evidence on direct environmental benefits, which constitute part of the total social return. Together, these analyses provide evidence on both the appropriable economic returns to other firms and the broader environmental externalities that motivate disclosure policy.

[4]For firms to fully internalize spillover benefits to other firms, their private competitive effects need to equal peer spillover effects ($\beta_1 = \beta_2$). Mechanisms that move private returns toward this benchmark include stronger licensing markets, enhanced IP protection, or disclosure-contingent policy instruments such as patent box tax incentives or regulatory approval tied to public registration of clean technologies.

1.3 Comparing Clean and Traditional Innovation

Our core analysis quantifies the disclosure wedge for environmental innovation. However, we also want to compare it to the disclosure 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 in ENV-TECH) as non-clean innovation. Our baseline specification therefore becomes:

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

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 $[(\beta_1^C - \beta_2^C) > (\beta_1^{NC} - \beta_2^{NC})]$.

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 predict $\beta_1^C > \beta_1^{NC}$. For competitive effects, we expect β_2^C to be less negative (closer to zero) than β_2^{NC} for two reasons. Traditional patents may 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. Second, environmental innovation may carry less competitive and strategic value for product-market rivals, particularly when clean technologies generate broad spillovers that offset narrow competitive advantages. We predict $\beta_2^C > \beta_2^{NC}$.

We expect the disclosure wedge to be larger for environmental innovation than for traditional innovation $(\beta_1^C - \beta_2^C) > (\beta_1^{NC} - \beta_2^{NC})$. Environmental innovation generates two distinct social benefits when disclosed: knowledge spillovers to other firms, and downstream pollution reductions as the technology diffuses. Both depend on disclosure for process-based environmental technologies, since neither knowledge nor pollution-reducing know-how diffuse efficiently without explicit documentation. Traditional technologies, in contrast, can be reverse-engineered and diffuse 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 clean patents into process- and

product-based to test our mechanism: if disclosure matters because process-based clean technologies are hidden inside production facilities, the spillover effect should concentrate in process-based patents. Third, we address potential endogeneity using EPA enforcement actions as instrumental variables for disclosure decisions.

2 Data and Summary Statistics

In this section, we provide an overview of the key data sources used in the paper. Our analysis focuses on US publicly listed firms. All variable descriptions are provided in Table B1. Accounting and financial variables are from the CRSP/Compustat Merged (CCM) database. We only retain firm-year observations with positive total assets.

2.1 Patent Data

We obtain 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

[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`.

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.

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.

2.2 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.

2.3 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 those from disclosure by product market rivals ($SPILLTECH^C$), reflecting that more firms operate in similar technological fields than directly compete in product markets. We take the ln(1+SPILLTECH) for all spillover variables.

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 (Koh et al., 2022; Mezzanotti and Simcoe, 2025). 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, 121,495 pounds 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.

3 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.

3.1 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 1, we separately estimate the aggregate economic impact from disclosed knowledge spillovers by technologically related firms (SPILLTECH^S) and competitive effects from spillovers by product market rivals (SPILLTECH^C). We ensure our estimates isolate the impact of disclosure, independent from innovation input, by controlling for R&D-investment based spillovers using 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., disclosure quality), we separate the decision to patent from the patent's ability to generate knowledge spillovers.

3.1.1 Baseline Results: Clean Innovation Disclosure Wedge

We first confirm BSV's findings using R&D spillovers in column (1) Table B3. 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, echoing BSV's findings. In column (2), we include patent-related spillovers. R&D spillover coefficients remain qualitatively unchanged, and patent technology spillovers are not significantly related to market value. We do find a product market competition effect for patents, which decreases the R&D product market effect. Further focusing on traditional (non-clean) patents, within all patents, in column (3) does not qualitative change the results. Patent stock overall does not substitute for R&D investment in these regressions.

Table 2 reports our baseline estimates for equation (3). 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 patent (publicly disclose) clean innovation, other firms in the same technology space can build on this knowledge to develop improved products, adopt cleaner production processes, or reduce regulatory compliance costs.

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 ($\beta_2^C = -0.010$, $t = -1.84$), 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. Coefficients of interest remain qualitatively similar. These patterns reveal a large clean disclosure wedge of approximately 2.04%: 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 ($\beta_1^{NC} = -0.03$, $t = -1.44$, column 3), reflecting the limited incremental social benefits from traditional innovation disclosure beyond R&D investment. The competitive effects of traditional innovation disclosure are modest at an 0.07% decrease in value for a 10% increase in rivals' non-clean innovation disclosure ($\beta_2^{NC} = -0.007$, $t = -2.5$, column 3), resulting in a near-zero 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's social spillovers, total R&D investment by technology peers is associated with a 1.42% increase in social returns, confirming BSV (column 2). However, jointly estimating R&D-based and disclosure-based spillovers in column (3) shows that disclosure absorbs most of the R&D spillover effect ($\beta_3^{R\&D} = 0.039$, $t = 0.79$), where columns (1) and (3) have identical R&D coefficients, implying that the absorption is mainly attributable to clean patent spillovers.

This result is consistent with disclosure mediating the R&D-spillover relationship. We use total R&D, bundling clean and non-clean innovation investment, so its spillover coefficient reflects a weighted average across both components. Clean patent disclosure captures the specific downstream output of clean R&D that generates spillovers, extracting the clean-relevant variation from the total R&D measure. For environmental technologies, disclosure is the channel through which clean R&D generates spillovers; without public disclosure, knowledge produced by clean R&D is much less likely to diffuse beyond the innovating firm.

The mediation interpretation is supported by a cross-innovation asymmetry that rules out a pure collinearity explanation. R&D spillovers are unaffected by the inclusion of non-clean patents (Table B3 column 2), consistent with non-clean knowledge diffusing through reverse engineering regardless of formal disclosure. If the R&D absorption were a statistical artifact, where any patent measure mechanically absorbs any R&D measure regardless of channel, it would appear symmetrically across both innovation types. The asymmetry is specific to the clean channel, where disclosure determines whether knowledge diffuses. This pattern supports the mediation interpretation: disclosure is an economically important margin in converting clean R&D into spillovers, not a competing explanation. Next we investigate the mechanism driving this asymmetry: the unobservability of environmental technologies and their outcomes, which makes patent disclosure a necessary condition for knowledge transfer.

3.1.2 Mechanism: Process- versus Product-Based Patents

Our baseline results show 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, are unobservable to external parties. 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. This mechanism is also

consistent with the broader appropriation literature, e.g., Arundel (2001), which shows that firms patent a smaller share of process than product innovations and often regard secrecy as relatively more effective for process innovations, precisely because they can remain hidden within production operations.

We classify patents as process- or product-based following the hypernym-based methodology developed in Tham et al. (2026). The method exploits the linguistic structure of patent titles and distinguishes 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 has 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 competitors' observable outputs. Process-based patents may differ from product patents along multiple dimensions, but their concentration in hidden operational changes is precisely why they provide a demanding test of our disclosure mechanism.

We investigate this hypothesis in Table 3 by decomposing clean patents into process-based and product-based patents. We find that our results are entirely driven by clean process-based knowledge spillovers, while clean product-based spillovers generate insignificant returns. A 10% increase in process-based patent disclosure generates social returns of 1.96% ($\beta_1^{C,Process} = 0.196$, $t = 1.86$), a similar coefficient to β_1^C in Table 2. In contrast, product-based clean patents generate near zero returns of 0.3% ($\beta_1^{C,Product} = 0.003$, $t = 0.02$).

The disclosure of clean process-based innovation generates larger social spillovers because competitors have fewer opportunities to infer the underlying technology; patent disclosure is therefore 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 disclosure.

3.1.3 Baseline Results: Additional Tests

We conduct extensive robustness analyses to validate our baseline findings in Table 2, presented in the Appendix. In our baseline analysis, we weight patents by disclosure specificity, as we are interested in spillovers to innovators in technology peers or product market rivals (who can interpret specific technical details), rather than the general public or other stakeholders (who may value readability and volume of information over technicality). In Table B4, we present results using alternative disclosure quality metrics (Dyer et al., 2024). Columns (1)-(3) use patent fullness (volume of information), columns (4)-(6) use patent readability, and columns (7)-(9) use a principal component analysis (PCA) combination of specificity, fullness, and readability. Across all measures, our results remain qualitatively unchanged.[6]

We conduct additional tests in which we identify competitors 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 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. 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 lower 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 that our results capture public disclosure effects rather than formal knowledge-sharing agreements. Ensuring knowledge spillovers materialize by restricting the sample to only 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 classifi-

[6]In Table C1, we present robustness for the process-based mechanism results in Table 3 using alternative disclosure quality metrics. Results remain qualitatively similar.

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

3.2 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 returns, e.g., pollution reductions and subsequent health and mortality improvements. Next establish whether disclosed clean innovation generates incremental real environmental outcomes.

3.2.1 Environmental Benefits

First, we estimate the relation between a firm's knowledge acquired from other firms' decisions to make clean technology publicly available, and its pollution outcomes. We plot estimates for facility-level regressions relating toxic emissions to knowledge acquired through patent citations in Figure 1, 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*).[7] Both variables are standardized by their respective standard deviation. 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.[8]

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. These effects persist through year five reaching 7.7% reduction, indicating that disclosed clean technologies enable lasting operational improvements in pollution abatement. In contrast, non-clean acquired

[7]We do not interpret citations as a direct measure of facility-level technology adoption. Rather, we consider them as a proxy for exposure to and use of disclosed clean knowledge, reflected empirically in downstream emissions reductions. By measuring acquired knowledge through backward patent citations, our estimates reflect a lower bound on environmental outcomes, as knowledge from observing patents can also transfer through informal channels.

[8]In Table B13, we also document the direct relation between a firm's clean innovation and its own emissions. Perhaps unsurprisingly, we find that clean patent filings are negatively related to a firm's facility emissions in the one- to four-year period following the filing.

knowledge is associated with emissions increases of up to 6.8% by year five, consistent with traditional innovation enabling production expansion, which increases pollution. 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. Results hold both within-facility over time (Table B11) and within-industry-year across facilities (Table B12). The former rules out time-invariant facility selection; the latter rules out industry-level selection.

The market value and emissions results are complementary and point to a real-effects interpretation of the spillover channel. The market value estimates establish that clean disclosure spillovers generate measurable economic value for technology peers but leave open whether this reflects genuine operational improvement or purely financial revaluation. Restricting the spillover measure to patents the focal firm cites (Table B9) retains roughly 70% of the baseline effect, indicating that observed citations, not just proximity-based exposure, drive a substantial portion of the market-value result. The emissions analysis uses the same citation-based measure of knowledge acquisition and shows that cited clean patents predict persistent pollution reductions. Taken together, the same citation-based exposure that delivers market-value gains also delivers emissions reductions, consistent with the market-value gains reflecting real adoption of cleaner production methods. On the other hand, the emissions results alone could reflect unobserved trends in regulatory pressure or abatement investment, but the market-value results, identified through cross-sectional variation in technological proximity and instrumented with EPA enforcement actions, point to disclosed knowledge as a first-order spillover channel.

3.2.2 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 then estimate that doubling disclosed clean innovation results in up to 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 $\mu g/m^3$

[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.

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 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 × $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. 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.

4 IV Estimation: EPA Enforcement as Shock to Disclosure Incentives

A central concern with OLS estimates is that peer firms' patent-based spillovers combine two distinct margins: peer firms' decisions to innovate in clean technologies and their decisions to disclose those innovations through patenting, so OLS estimates may conflate the disclosure channel with the investment channel. While our R&D-spillover controls (Section 3.1.1) and disclosure intensive margin (disclosure-quality weighting, Section 1.2.1) partially address this concern, we now implement an instrumental variable (IV) strategy using EPA enforcement actions against peer firms as a shock to their disclosure incentives. Enforcement exposes targeted firms to heightened scrutiny, legal risk, and sanction exposure, prompting them to patent pollution-abatement technologies as credible signals of remediation effort to regulators.

Our specification places focal firm i's market value on the left-hand side and spillover measures constructed from peer firms j's disclosure on the right-hand side. The instrument targets peer firms j, not firm i. This distinction is central to the identification strategy. EPA enforcement of firm j affects the targeted firm j's value through multiple direct channels: compliance investment, operational restructuring, regulatory risk premia, reputational effects. These effects are real but operate on firm j's value, which is not our dependent variable. For EPA enforcement to violate the exclusion restriction, it would need to affect firm i's value through channels other than the disclosed knowledge entering SPILLTECH^S.

For this IV to isolate the disclosure margin, two conditions must hold. First, EPA

enforcement must shift peer firms' disclosure decisions without changing their underlying innovation activity. Second, EPA enforcement of peer firm j must not affect focal firm i's value through channels other than the disclosed knowledge itself. The first condition is empirical: Table B14 shows that peer firms' R&D investment and non-clean patent filings are unchanged in the enforcement year, while clean patent filings rise sharply. The second is an identifying assumption supported by the cross-industry structure of SPILLTECH^S. Consider the SCR catalyst example, an EPA action against a utility induces disclosure of a catalyst design that enters the SPILLTECH^S of a cement producer with whom the utility shares IPC classes but no product-market overlap. For the cement producer's value to respond to the utility's enforcement through non-disclosure channels, the enforcement would need to propagate through input-output linkages or industry-wide regulatory anticipation, channels that are narrower than the direct effects EPA enforcement has on targeted firms. We investigate both conditions empirically in Section 4.2.

Table 4 presents two-stage least squares IV results. The endogenous variables are the clean disclosure spillovers SPILLTECH$^S_{Clean}$ and SPILLTECH$^C_{Clean}$. We instrument these with analogous spillover-weighted EPA enforcement measures: SPILLENF^S aggregates EPA enforcement actions faced by all firms $j \neq i$ up to year t, weighted by their technology proximity to firm i; SPILLENF^C aggregates EPA enforcement actions faced by all product-market rivals of firm i. In this specification we continue to control for R&D investment, so that variation in the instrumented disclosure spillovers is identified holding peer firms' innovation investment approximately constant.[10]

The first-stage regressions (columns 1 and 2) show that EPA enforcement strongly predicts clean patent disclosure by both technology peers and product-market 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 product-market competitors (SPILLENF^C) exhibit strong predictive power for product-market 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 through patent filings, EPA actions provide a positive shock to the available pool of clean knowledge.

[10]Table B15 provides facility- (Panel A), firm- (Panel B), and patent-level (Panel C) summary statistics within the EPA sample. Facilities receive 0.03 formal EPA actions per year on average, with frequency increasing over time (Figure B1). The average facility-level penalty amounts to \$150,203 USD and the maximum is \$20 million USD. However, direct monetary fines understate the true cost of enforcement: firms also face potential license withdrawals, escalating fines for repeat violators, regulatory scrutiny, and litigation risk (Blundell, Gowrisankaran, and Langer, 2020).

Column (3) presents second-stage estimates. The IV coefficient on technology spillovers nearly doubles relative to the baseline: a 10% increase in EPA-induced clean knowledge from technologically related firms raises firm value by 3.57% ($\beta_1^C = 0.357$, $t = 2.77$). This larger IV estimate is consistent with two non-exclusive interpretations. First, EPA-induced disclosure may be concentrated in technologies where disclosure is especially consequential for diffusion, particularly process-based clean technologies. Second, the IV identifies a local average treatment effect for firms whose disclosure responds to regulatory pressure, a subset that may differ from the population of disclosers captured by OLS. The competitive effect from rivals' EPA-induced disclosure is positive but not statistically distinguishable from zero ($\beta_2^C = 0.030$, $t = 1.63$). Non-clean innovation continues to show no meaningful disclosure wedge.

The IV identifies variation in the availability of disclosed clean knowledge driven by peer regulatory pressure, not variation in the productivity or quality of that knowledge, which is tied to the underlying innovation and absorbed by our controls. The disclosure wedge in our IV specification is 3.27% ($0.357 - 0.030 = 0.327$) per 10% increase in instrumented spillovers, exceeding the baseline wedge of 2.04%. This reflects a substantial misalignment between private incentives and social benefits.

In Table 5, we show that our results reflect process-based rather than product-based clean patents, in the IV setting. Column (3) shows that a 10% increase in instrumented process-based clean disclosure increases focal firm value by 3.46% ($t = 2.75$). This estimate is nearly identical to the aggregate clean patent IV effect of 3.57% in Table 4, indicating that almost all EPA-induced spillovers operate through process-based innovations. This is consistent with regulatory enforcement targeting operational pollution abatement rather than individual products. Competitive effects remain insignificant. The resulting disclosure wedge is concentrated in the technologies most critical to pollution reduction.[11]

To contextualize the IV estimates, consider the median U.S. publicly-listed firm in our sample with market value of $2.1 billion. A 10% increase in instrumented clean knowledge from technology peers generates approximately $75 million in focal firm value. The Table 5 decomposition shows this effect is driven almost entirely by process-based clean patents ($73 million per 10% increase, from the 3.46% coefficient), with product-based clean patents not contributing. For the technologies most critical to pollution reduction, process innovations embedded inside production facilities, the spillover magnitudes are both economically large and uncontaminated by offsetting competitive effects.

Critically, these estimates represent only the portion of spillover benefits captured by

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

publicly listed focal firms' shareholders through changes in market valuation. Total social benefits are substantially larger. Our emissions analysis documents 5-9% reductions in facility-level pollution from clean knowledge spillovers over a five-year post-citation window, translating to health benefits, regulatory cost savings, and environmental externalities that extend beyond the private returns captured in market valuations. Additionally, we only observe spillovers to public firms, excluding knowledge transfer to private firms, startups, universities, and consumers. The appropriable market-value benefit thus understates the full social value of environmental disclosure, potentially substantially so.

4.1 Environmental Benefits from EPA-Induced Disclosure

In order for regulatory pressure to trigger environmental spillovers, EPA-induced patents should reduce emissions beyond the innovating firms. Next we assess 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 citations of the subset of 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, with regression estimates reported in Table B16. 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.

4.2 EPA Actions as Shocks to Disclosure Incentives: Additional Tests

As discussed above, our identification strategy relies on two conditions. First, relevance requires that EPA actions generate variation in peer firms' disclosure decisions without inducing changes in their underlying innovation activity. Second, exogeneity (the exclusion restriction) requires that EPA actions on peer firm j affect focal firm i's value solely through disclosure, not through other channels. While the former can be assessed empirically, the latter remains an identifying assumption. In this section we provide evidence consistent with EPA actions that primarily shift disclosure behavior, rather than altering innovation activity or the economic value of the underlying technology.

4.2.1 EPA-Induced Clean Innovation Disclosure: Relevance

We investigate whether an increase in environmental regulatory pressure from EPA actions is associated with an increase in clean innovation disclosure at the extensive (number of patents filed) and intensive (disclosure speed) 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, 2026). 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 estimate Poisson specifications at both the firm-year and facility-year level, since EPA enforcement is observed at facility level while patent filings are attributed to firms. The dependent variable is the number of clean patents filed in year t, as every patent filing represents a disclosure decision. We control for facility- and firm-level characteristics including historical emissions, and include facility and industry $\times$ year fixed effects.

Our focus on patent responses in the enforcement year addresses a key concern: that EPA enforcement changes not only disclosure decisions but also firms' underlying innovation activity. If enforcement triggers new R&D, resulting patents would reflect innovation responses rather than disclosure responses. However, the lag from R&D expenditure to patent filing typically exceeds one year (Pakes and Griliches, 1980; Wang and Hagedoorn, 2014), so clean patents filed in the enforcement year reflect R&D initiated before enforcement. Table B14 shows that R&D spending is unchanged in the enforcement year, CapEx rises slightly at $t+1$, and clean patent filings rise sharply at $t=0$. The pattern is consistent with EPA enforcement shifting the disclosure margin rather than triggering new innovation in the short run.[12]

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).[13] Economically, a one-unit increase in formal EPA enforcement actions in

[12]In addition, CapEx and R&D budgets are typically planned in advance and adjust only gradually at annual frequencies, making sharp contemporaneous changes in total investment unlikely.

[13]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 patents including the year after EPA

facility f is associated with a 5.66% ($=(e^{0.055}-1)\times 100\%$) increase in clean patents filed by the facility's parent firm in the year of the enforcement action.[14] 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.

If EPA enforcement triggers a general increase in patenting activity rather than a targeted increase in clean disclosure, we would expect traditional (non-clean) patent filings to increase as well. In Panel B, we do not find evidence for an increase in traditional patent filings following EPA enforcement actions. The EPA-induced response is clean-specific patents, consistent with firms disclosing pollution-abatement capabilities rather than adjusting their overall patenting strategy.

4.2.2 EPA-Induced Disclosure Incentives: Patent Disclosure Timing

Having established that EPA enforcement shifts the extensive margin of disclosure, we now examine the intensive margin: the timing of disclosure. Evidence that EPA enforcement accelerates disclosure specifically, rather than shifting the underlying innovation or the general pace of patenting, strengthens the interpretation that the instrument operates through the disclosure channel, not through broader effects on innovation activity. 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 technological peer firm value.

One concern with the exclusion restriction is that EPA enforcement may cluster by industry or emission intensity, affecting focal firm value through industry-wide regulatory anticipation rather than through disclosed knowledge. This mechanism, if operative, would bias our estimates toward zero: if focal firms anticipate future scrutiny after observing EPA actions against their technology peers, their market value would decline, working against our finding of positive spillovers. In addition, our IV controls for spillovers from product-market rivals, which tend to operate in the same industry as the focal firm, absorbing within-industry regulatory anticipation channels directly.

Specifically, we exploit the *timing* of firms' decisions to publicly disclose clean innovation earlier than required by the USPTO. Since the American Inventor Protection Act (AIPA) of 2000, USPTO applications are published 18 months after filing by default, though firms can voluntarily publish earlier or, under certain conditions, delay publication until the patent grant decision (see details Appendix A). We measure disclosure speed using *Disclosure*

actions (t+1), and the results remain qualitatively similar.

[14] 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).

Latency and *% Latency* (Glaeser and Landsman, 2021; Hegde, Jovanovic, and Liu, 2023). *Disclosure Latency* is defined as 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 clean patent disclosure timing at the patent application level, comparing applications by the same firm that differ in disclosure speed. 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, 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. Non-clean patents show no timing response (columns 3-4), ruling out the possibility that EPA enforcement generically accelerates patent disclosure. The acceleration is specific to clean innovation, consistent with firms using clean disclosure as a targeted signal of compliance capability.

4.2.3 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 change 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 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 28.6% reduction in HPV duration compared to those that do

[15]Panel C of Table B15 provides the summary statistics for these variables. The average *Disclosure Latency* is 462 days, *% Latency* averages 68% indicating clean patents are published 32% earlier than mandated. 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.

not file clean patents. Given that the mean HPV duration is 600 days, this represents a 200-day quicker settlement with the EPA.

Given the potential costs and negative consequences of repeated violations, firms should also have incentives to disclose clean technology to investors. For clean patent filings to signal to investors, investors must be aware of the filings before USPTO publication, which occurs 18 months after filing under AIPA. Firms can disclose in parallel through SEC filings: the EPA retrieves corporate SEC filings during enforcement investigations (Li and Wang, 2024), so 10-K/Q disclosures reach both investors and regulators simultaneously.

We examine whether firms reference clean technology terminology in 10-K/Q filings around patent filing events. 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). We define *10K/Q* as an indicator variable equal to one if these terms co-occur in the same sentence, and zero otherwise.

Table B17 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 presented 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 10K/Q in the patent filing year, a 3.5% increase relative to the unconditional co-occurrence probability. 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. Clean-technology references continue rising up to $t + 2$, consistent with firms continuing to disseminate information through the USPTO-mandated disclosure period. The pattern indicates firms use SEC filings to signal clean innovation to investors in parallel with the patent filing, rather than waiting for USPTO publication.

5 Conclusion

Knowledge spillovers are typically assumed to materialize automatically, regardless of whether a firm chooses to publicly disclose its innovation. However, firms that innovate but retain trade secrecy contribute substantially less to the public knowledge stock, limiting spillovers to other firms. We examine this overlooked disclosure dimension of knowledge diffusion in the context of clean innovation, which is disproportionately difficult to reverse-engineer given the unobservability of process-based technologies and emissions outcomes. We show that

the social benefits of disclosing clean innovation, measured through spillovers and emissions, significantly exceed the private returns to disclosing firms. This disclosure wedge indicates a potential under-estimation of clean technologies' full societal contribution and a systemic under-disclosure of clean innovation from a social perspective. Crucially, the disclosure channel is relevant for any innovation whose process or outcome is unobservable to external participants, extending beyond environmental technology.

In addition, our documented effects represent a lower bound for the impact of clean innovation disclosure on social benefits. They 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 findings suggest that policymakers should not only incentivize R&D but also develop frameworks that promote the disclosure of innovation - with the appropriate intellectual property protections. Such a comprehensive strategy may maximize the impact of clean and process-based 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. Our study thereby underscores the central importance of disclosure requirements and knowledge spillovers in achieving global climate objectives and highlights 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.

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.

Arundel, A. 2001. The relative effectiveness of patents and secrecy for appropriation. Research Policy 30:611–24.

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? Review of Finance 28:1151–86.

Giglio, S., T. Kuchler, J. Stroebel, and X. Zeng. 2026. Biodiversity risk. Review of Finance 30:131–61.

Giroud, X., E. Liu, and H. Mueller. 2026. Innovation spillovers across U.S. tech clusters. Journal of Financial Economics 179:104264–.

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. McGraw-Hill, New York.

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

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.

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.

Klasa, S., H. Ortiz-Molina, M. Serfling, and S. Srinivasan. 2018. Protection of trade secrets and capital structure decisions. Journal of Financial Economics 128:266–86.

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

Koh, P.-S., D. Reeb, E. Sojli, W. W. Tham, and W. Wang. 2022. Deleting unreported innovation. Journal of Financial and Quantitative Analysis 57:2324 – 2354.

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.

Mezzanotti, F., and T. Simcoe. 2025. Innovation, patenting and appropriability: Survey evidence from a nationally representative sample of U.S. firms. Review of Economics and Statistics Forthcoming.

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

Pakes, A., and Z. Griliches. 1980. Patents and R&D at the firm level: A first report. Economics Letters 5:377–81.

Png, I. P. 2017. Law and innovation: Evidence from state trade secrets laws. Review of Economics and Statistics 99:167–79.

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. 2026. Green patenting and voluntary innovation disclosure. Review of Accounting Studies .

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

Tham, W. W., S. Baslandze, E. Sojli, and L. Liu. 2026. Foundational processes and growth. Journal of Financial Economics 180.

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

Wang, N., and J. Hagedoorn. 2014. The lag structure of the relationship between patenting and internal R&D revisited. Research Policy 43:1275–85.

Wang, X. 2024. Patenting and information disclosure. Working paper, Northwestern University.

Tables and Figures

Figure 1
Clean Innovation Knowledge Spillovers and Toxic Emissions

This figure 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.

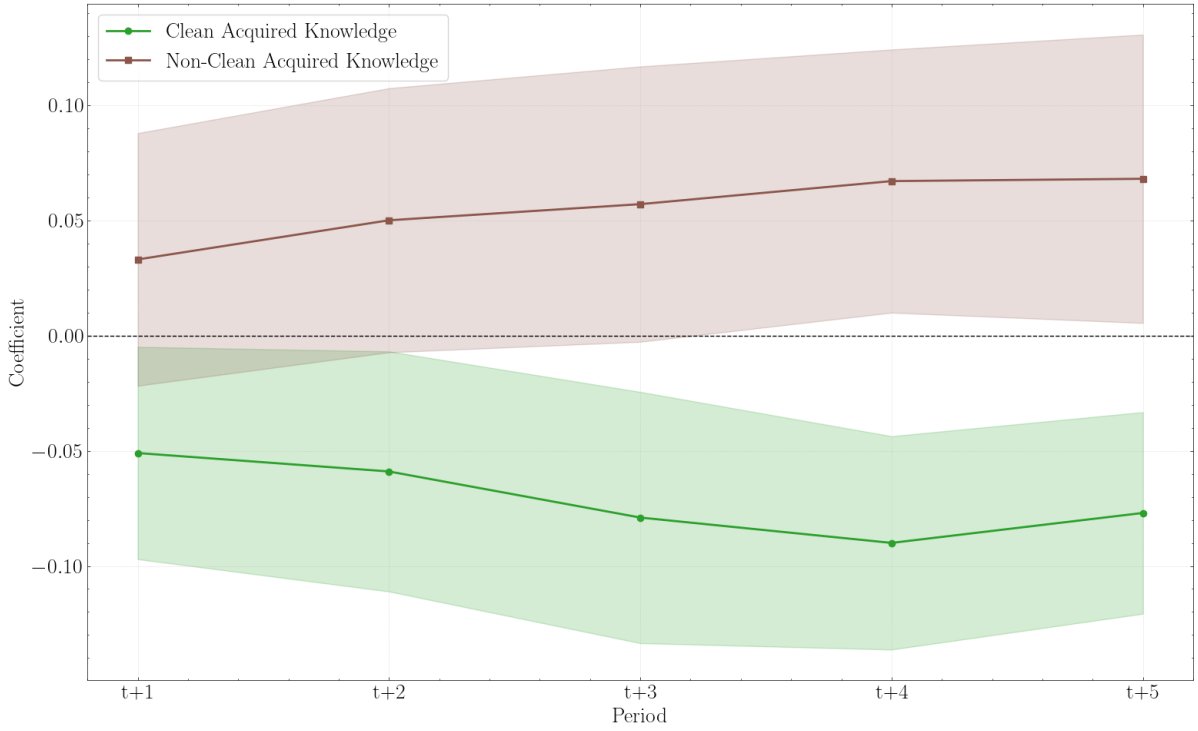


Figure 2
Clean Innovation Knowledge Spillovers from EPA-Induced Patents and Toxic Emissions

This figure presents regression coefficients for the relation between emissions and knowledge spillovers from clean patents filed in the year of or the year following EPA actions. 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 in the year of or the year following EPA actions. *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.

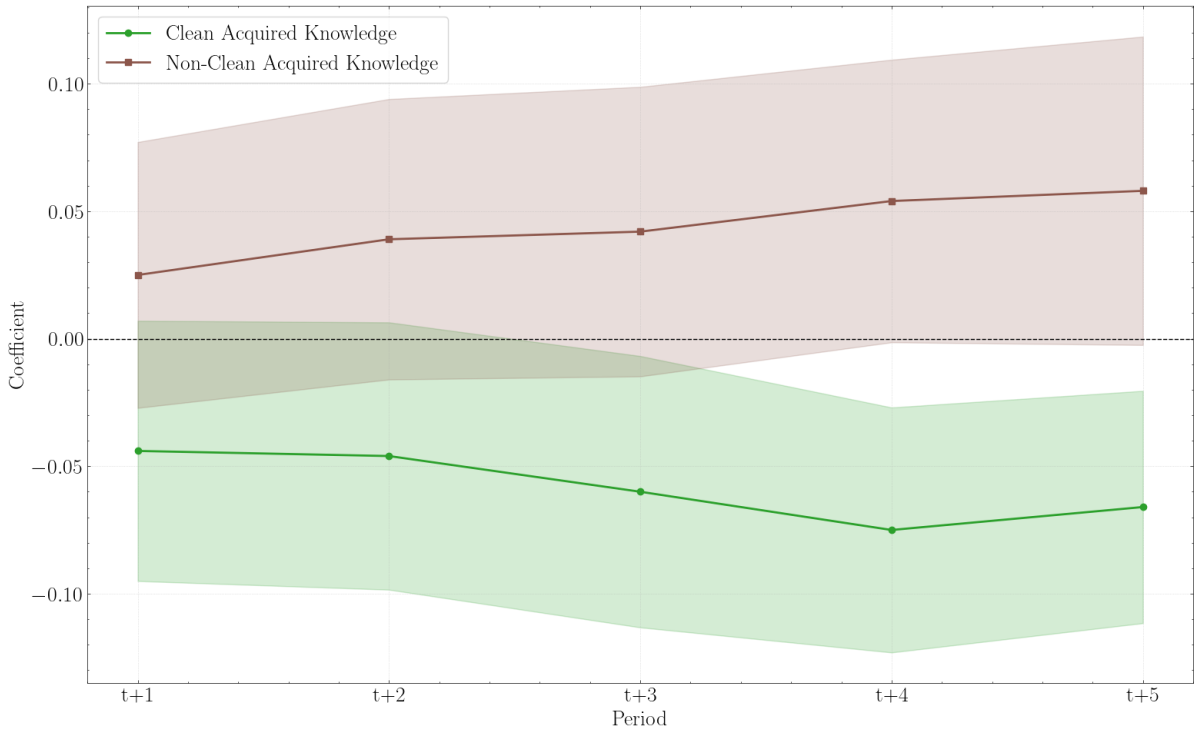


Figure 3
Relation between 10-K/Q Reporting and Clean Patents

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 B17. 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.

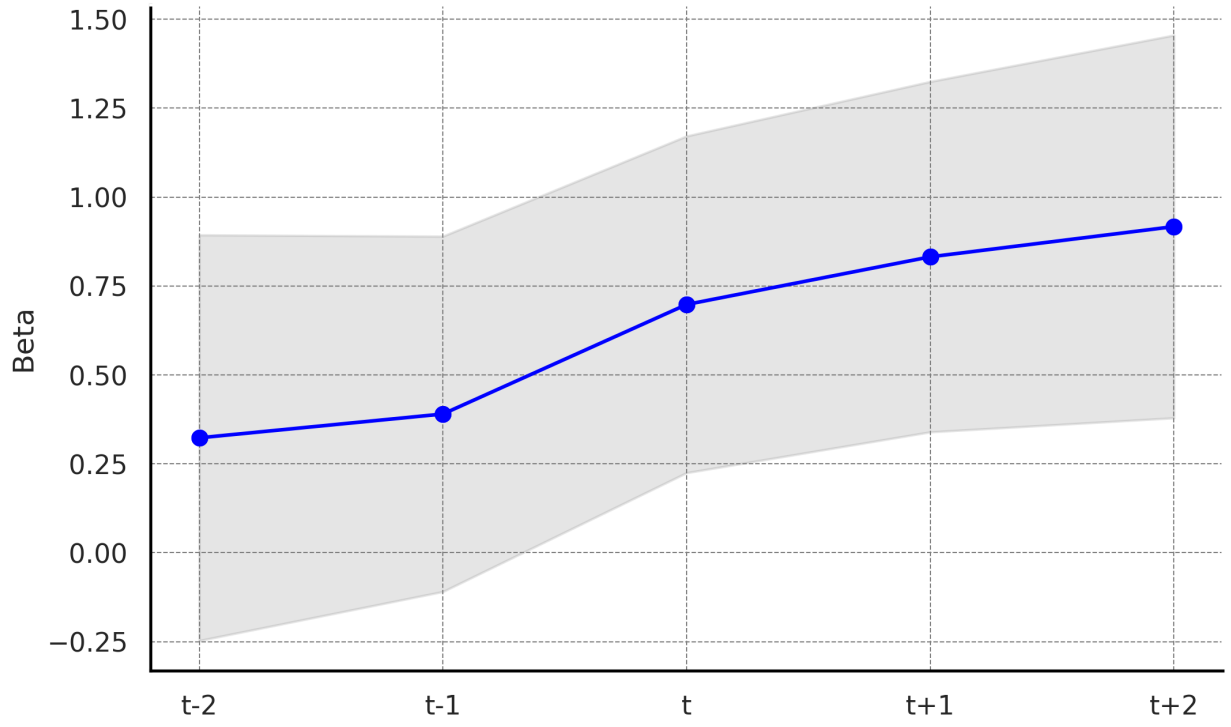


Table 1
Summary Statistics

This table presents summary statistics for the main variables used in our baseline analysis. Panel A covers all firms in the CRSP-Compustat (CCM) data with positive total assets. Panel B covers all facilities of patenting firms. 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 B15.

Panel A. Firm level

	N	Mean	Median	Std. Dev.
$\text{SPILLTECH}^{S,\text{Clean}}$	196,080	198.31	0.00	829.89
$\text{SPILLTECH}^{C,\text{Clean}}$	196,080	4.57	0.09	13.18
$\text{SPILLTECH}^{S,\text{Process}}$	196,080	124.34	0.00	510.73
$\text{SPILLTECH}^{S,\text{Product}}$	196,080	77.43	0.00	342.35
$\text{SPILLTECH}^{S,\text{Non-Clean}}$	196,080	40,035.27	0.00	76,415.12
$\text{SPILLTECH}^{C,\text{Non-Clean}}$	196,080	741.69	22.43	1,925.73
$\text{SPILLTECH}^{S,\text{R\&D}}$	196,080	11,132.74	0.00	45,222.12
$\text{SPILLTECH}^{C,\text{R\&D}}$	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:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). G_{jt} is weighted by quantile-transformed IPC3-year-adjusted specificity (number of specific pieces of information in the patent filing). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(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:

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

$\text{SPILLTECH}^S_{Clean,Process/Product}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^{q,Process/Product}_{jt}$, where Process/Product-Based patents in Tham et al. (2026) and TECH_{ij} measures technology proximity (Jaffe, 1986). 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). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$
$\text{Ln(SPILLTECH}^S_{\text{Clean, Process}}\text{)}$	0.196*
	(1.86)
$\text{Ln(SPILLTECH}^S_{\text{Clean, Product}}\text{)}$	0.003
	(0.02)
$\text{Ln(SPILLTECH}^C_{\text{Clean}}\text{)}$	-0.010*
	(-1.94)
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$	-0.031
	(-1.49)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$	-0.007**
	(-2.49)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.032
	(0.64)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.002
	(-1.03)
Ln(R&D Stock/TA)	0.231***
	(7.87)
Ln(G)	-0.007
	(-0.28)
Controls	✓
Year FE	✓
Firm FE	✓
Obs.	168,575
Adj. R^2	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. We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}}^S_{Clean})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \boldsymbol{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.$$

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean patent stock G^q_{jt} with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	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. We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\begin{aligned}\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = {} & \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}^S_{Clean,Process}})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} \\ & + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ & + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ & + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \mathbf{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G_{jt}^{q,Process}$, where $G_{jt}^{q,Process}$ is the disclosure-quality-weighted stock of clean process-based patents filed by firm j, with process-based patents defined as in Tham et al. (2026), and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G_{jt}^{q}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean process-based patent stock $G_{jt}^{q,Process}$ with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	First Stage	First Stage	Second Stage
	$\ln(\text{SPILLTECH}^{S,\text{ Clean, Process}})$	$\ln(\text{SPILLTECH}^{C,\text{Clean}})$	$\ln(\text{MV/TA})_{i,t+1}$
	(1)	(2)	(3)
$\ln(\text{SPILLENF}^S)$	0.196***	-0.015	
	(11.63)	(-0.72)	
$\ln(\text{SPILLENF}^C)$	-0.001	0.324***	
	(-0.73)	(39.54)	
$\ln(\text{SPILLTECH}^{S,\text{ Clean, Process}})$			0.346***
			(2.75)
$\ln(\text{SPILLTECH}^{C,\text{Clean}})$			0.030
			(1.62)
$\ln(\text{SPILLTECH}^{S,\text{Non-Clean}})$	0.016***	0.043***	-0.037*
	(6.08)	(2.87)	(-1.76)
$\ln(\text{SPILLTECH}^{C,\text{Non-Clean}})$	-0.000	0.156***	-0.013***
	(-0.21)	(52.18)	(-3.24)
$\ln(\text{SPILLTECH}^{S,\text{R\&D}})$	0.265***	0.038	-0.054
	(5.91)	(1.07)	(-0.72)
$\ln(\text{SPILLTECH}^{C,\text{R\&D}})$	-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 2.2. $\lambda_{ind,t}$ 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.

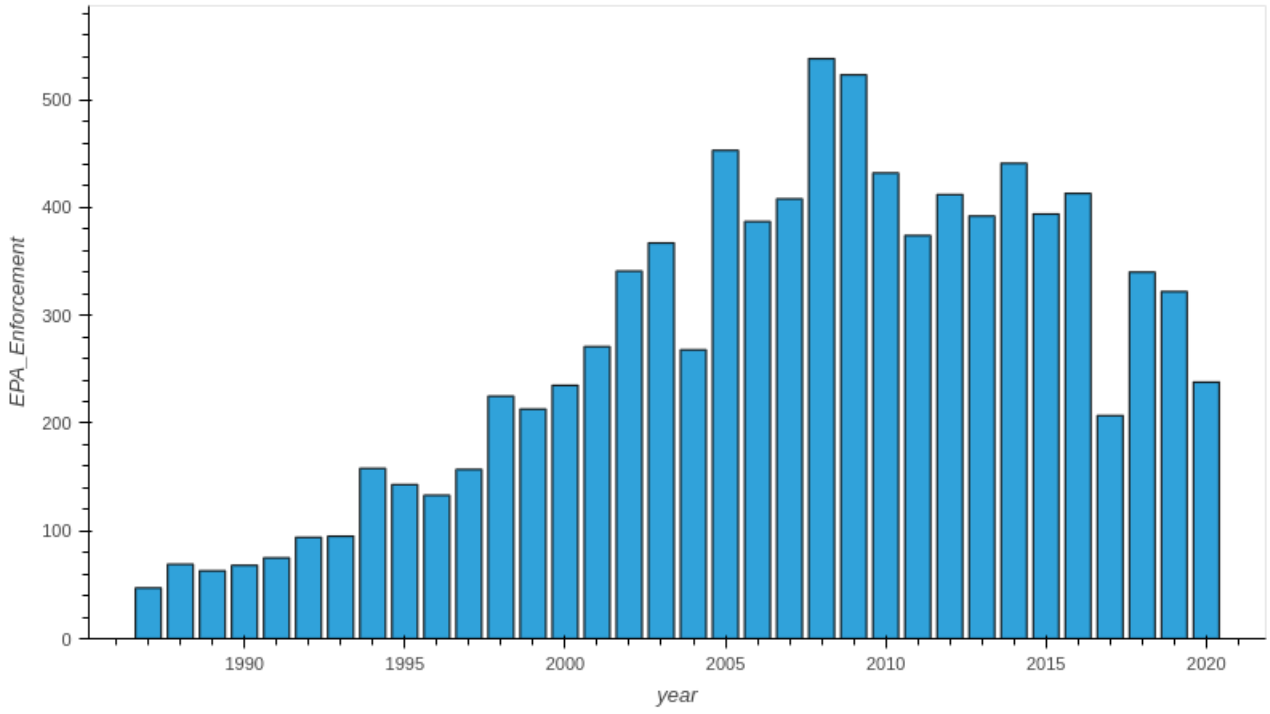


Table B1
Variable Definitions

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).
Disclosure Latency		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, since filing date.
Percentage Latency		Disclosure Latency divided by Latest 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
All-Patent, Non-Clean, and R&D-Based Knowledge Spillovers: BSV Approach

This table replicates and extends the BSV approach to knowledge spillovers. We estimate firm-year panel regressions of the form:

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

$\text{SPILLTECH}^S_{\cdot}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of the relevant patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{\cdot}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). $\text{SPILLTECH}^S_{R\&D}$ and $\text{SPILLTECH}^C_{R\&D}$ are constructed analogously using R&D stock in place of patent stock. We also control for the focal firm'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. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	$\text{Ln(MV/TA)}_{i,t+1}$	$\text{Ln(MV/TA)}_{i,t+1}$	$\text{Ln(MV/TA)}_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln(SPILLTECH}^S_{\text{All Pat.}}\text{)}$		-0.021	
		(-1.09)	
$\text{Ln(SPILLTECH}^C_{\text{All Pat.}}\text{)}$		-0.007***	
		(-3.12)	
$\text{Ln(SPILLTECH}^S_{\text{Non-Clean}}\text{)}$			-0.025
			(-1.23)
$\text{Ln(SPILLTECH}^C_{\text{Non-Clean}}\text{)}$			-0.008***
			(-3.42)
$\text{Ln(SPILLTECH}^S_{\text{R\&D}}\text{)}$	0.136***	0.141***	0.141***
	(4.22)	(4.28)	(4.30)
$\text{Ln(SPILLTECH}^C_{\text{R\&D}}\text{)}$	-0.009***	-0.004**	-0.004**
	(-6.55)	(-2.38)	(-1.97)
Ln(R&D Stock/TA)	0.226***	0.229***	0.230***
	(7.74)	(7.80)	(7.82)
Ln(G)	0.013	0.014	0.014
	(0.57)	(0.62)	(0.63)
Controls	✓	✓	✓
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, and a principal component combination (PCA) of fullness, readability, and specificity. We estimate firm-year panel regressions of the form:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). 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 (column 1), readability (Gunning (1952) FOG index multiplied by negative one (column 2), and a principal component combination of specificity, fullness, and readability (column 3). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	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:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). Industry competitors IND are defined as firms within the same SIC-3 or SIC-4 digit classification, with weights assigned either equally ($\text{PRODSIM} = 1/N$, columns 1 and 2) or based on product similarity as in Hoberg and Phillips (2016) (column 3). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(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:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(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 from public and private firms. We estimate firm-year panel regressions of the form:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). G^q_{jt} includes patents filed by private firms, in addition to the U.S. publicly-listed firms. $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum^{j \neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(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:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(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 cited by firm i. We estimate firm-year panel regressions of the form:

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

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the disclosure-quality-weighted stock of clean patents filed by firms j that are cited by firm i and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$	
	(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:

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

$\text{SPILLTECH}^S_{Clean/Dirty}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^{q,clean/dirty}_{jt}$, 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. $\text{SPILLTECH}^C_{Clean/Dirty}$ is defined as $\sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^{q,clean/dirty}_{jt}$, where PRODSIM_{ij} is between-firm product similarity from Hoberg and Phillips (2016). We also control for the focal firm'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.

	$\text{Ln(MV/TA)}_{i,t+1}$		
	(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), 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
EPA Enforcement Actions and Firm Innovation and Investment Responses

This table presents regression results for the relation between EPA enforcement actions and firms' clean patenting, R&D expenditure, and capital expenditure. We estimate firm-year panel regressions of the form:

$$Y_{i,t} = \beta_0 + \beta_1 \text{Formal Actions}_{i,t+\tau} + \Gamma X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t},$$

where $Y_{i,t}$ represents the number of clean patents filed (column 1), R&D expenditure scaled by total assets (column 2), and CAPX scaled by total assets (column 3) by firm i in year t. *Formal Actions* is the number of formal EPA enforcement actions (fines, civil lawsuits) against firm i at time $t+\tau$ for $\tau = -1, 0, +1$. X is a vector of control variables which includes: ln(Sales), Tobin's Q, Cash, PPE, and HHI, which 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. All potentially unbounded variables are winsorised at the 1% level. The sample period is 1988 to 2020.

	Clean Pat.	R&D	CAPX
	(1)	(2)	(3)
Formal Actions$_{t+1}$	0.248	-0.001	0.003
	(1.07)	(-0.55)	(1.28)
Formal Actions$_t$	0.218**	0.002	0.002
	(2.34)	(0.93)	(0.72)
Formal Actions$_{t-1}$	0.391***	0.001	0.008**
	(3.07)	(0.49)	(2.55)
Controls	✓	✓	✓
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
Obs.	6,598	6,598	6,589
R^2		0.89	0.65

Table B15
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 Table 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 B16
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 filed in the year of or the year following EPA actions. *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 B17
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.

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

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:

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

$\text{SPILLTECH}^S_{Clean,Process/Product}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G^{q,Process/Product}_{jt}$, where Process/Product-Based patents are defined following Tham et al. (2026); TECH_{ij} measures technology proximity (Jaffe, 1986). 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 (column 1), readability (Gunning (1952) FOG index multiplied by negative one (column 2), and a principal component combination of specificity, fullness, and readability (column 3). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum^{j\neq i}_{i_{TNIC3}=j_{TNIC3}} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). We also control for the focal firm'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.

	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 IV: Fullness-Weighted Patents

This table presents results for the disclosure wedge from clean knowledge spillovers, using EPA enforcement actions as an instrument for firms' clean patenting decisions, with patent stocks weighted by fullness (number of words). We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}}^S_{Clean})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \mathbf{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.$$

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the fullness (number of words)-weighted stock of clean patents filed by firm j and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean patent stock G^q_{jt} with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	First Stage	First Stage	Second Stage
	$\ln(\text{SPILLTECH}^S_{\text{Clean}})$	$\ln(\text{SPILLTECH}^C_{\text{Clean}})$	$\ln(\text{MV/TA})_{i,t+1}$
	(1)	(2)	(3)
$\ln(\text{SPILLENF}^S)$	0.192***	-0.015	
	(11.94)	(-0.72)	
$\ln(\text{SPILLENF}^C)$	-0.001	0.325***	
	(-1.09)	(39.98)	
$\ln(\text{SPILLTECH}^S_{\text{Clean}})$			0.351***
			(2.77)
$\ln(\text{SPILLTECH}^C_{\text{Clean}})$			0.030
			(1.63)
$\ln(\text{SPILLTECH}^S_{\text{Non-Clean}})$	0.013***	0.041***	-0.034
	(5.29)	(2.71)	(-1.60)
$\ln(\text{SPILLTECH}^C_{\text{Non-Clean}})$	-0.000	0.158***	-0.013***
	(-0.80)	(52.97)	(-3.22)
$\ln(\text{SPILLTECH}^S_{\text{R\&D}})$	0.252***	0.038	-0.051
	(5.88)	(1.08)	(-0.70)
$\ln(\text{SPILLTECH}^C_{\text{R\&D}})$	-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 IV: Fullness-Weighted Patents

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, with patent stocks weighted by fullness (number of words). We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\begin{aligned}\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = {} & \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}^S_{Clean,Process}})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} \\ & + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ & + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ & + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \mathbf{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^{q,Process}_{jt}$, where $G^{q,Process}_{jt}$ is the fullness (number of words)-weighted stock of clean process-based patents filed by firm j, with process-based patents defined as in Tham et al. (2026), and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean process-based patent stock $G^{q,Process}_{jt}$ with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	First Stage	First Stage	Second Stage
	$\ln(\text{SPILLTECH}^S_{\text{Clean, Process}})$	$\ln(\text{SPILLTECH}^C_{\text{Clean}})$	$\ln(\text{MV/TA})_{i,t+1}$
	(1)	(2)	(3)
$\ln(\text{SPILLENF}^S)$	0.198***	-0.015	
	(11.56)	(-0.72)	
$\ln(\text{SPILLENF}^C)$	-0.001	0.325***	
	(-0.79)	(39.98)	
$\ln(\text{SPILLTECH}^S_{\text{Clean, Process}})$			0.340***
			(2.75)
$\ln(\text{SPILLTECH}^C_{\text{Clean}})$			0.030
			(1.62)
$\ln(\text{SPILLTECH}^S_{\text{Non-Clean}})$	0.017***	0.041***	-0.035*
	(6.08)	(2.71)	(-1.65)
$\ln(\text{SPILLTECH}^C_{\text{Non-Clean}})$	-0.000	0.158***	-0.013***
	(-0.33)	(52.97)	(-3.22)
$\ln(\text{SPILLTECH}^S_{\text{R\&D}})$	0.271***	0.038	-0.054
	(5.90)	(1.08)	(-0.72)
$\ln(\text{SPILLTECH}^C_{\text{R\&D}})$	-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 IV: Readability-Weighted Patents

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. We estimate the following 2SLS firm-year panel regression:

$$\begin{aligned}
\textbf{1st Stage:}\quad \ln(\text{SPILLTECH}^S_{Clean})_{i,t} =& \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} \\
&+ \boldsymbol{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S, \\
\ln(\text{SPILLTECH}^C_{Clean})_{i,t} =& \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} \\
&+ \boldsymbol{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.
\end{aligned}$$

$$\begin{aligned}
\textbf{2nd Stage:}\quad \ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}}^S_{Clean})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} \\
&+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\
&+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\
&+ \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \boldsymbol{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.
\end{aligned}$$

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the readability-weighted stock of clean patents filed by firm j, weighting G_{jt} by the Gunning (1952) FOG index multiplied by negative one, and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean patent stock G^q_{jt} with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	First Stage	First Stage	Second Stage
	$\ln(\text{SPILLTECH}^S_{\text{Clean}})$	$\ln(\text{SPILLTECH}^C_{\text{Clean}})$	$\ln(\text{MV/TA})_{i,t+1}$
	(1)	(2)	(3)
$\ln(\text{SPILLENF}^S)$	0.208***	-0.010	
	(12.54)	(-0.48)	
$\ln(\text{SPILLENF}^C)$	-0.001	0.313***	
	(-0.94)	(40.29)	
$\ln(\text{SPILLTECH}^S_{\text{Clean}})$			0.323***
			(2.76)
$\ln(\text{SPILLTECH}^C_{\text{Clean}})$			0.031
			(1.64)
$\ln(\text{SPILLTECH}^S_{\text{Non-Clean}})$	0.013***	0.048***	-0.030
	(5.71)	(3.52)	(-1.52)
$\ln(\text{SPILLTECH}^C_{\text{Non-Clean}})$	-0.000	0.143***	-0.012***
	(-0.88)	(49.75)	(-3.12)
$\ln(\text{SPILLTECH}^S_{\text{R\&D}})$	0.271***	0.043	-0.050
	(6.16)	(1.27)	(-0.69)
$\ln(\text{SPILLTECH}^C_{\text{R\&D}})$	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 IV: Readability-Weighted Patents

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. We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \mathbf{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}^S_{Clean,Process}})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \mathbf{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.$$

$\text{SPILLTECH}^S_{Clean,Process}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G_{jt}^{q,Process}$, where $G_{jt}^{q,Process}$ is the readability-weighted stock of clean process-based patents filed by firm j, weighting $G_{jt}^{Process}$ by the Gunning (1952) FOG index multiplied by negative one, with process-based patents defined as in Tham et al. (2026), and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G_{jt}^q$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean process-based patent stock $G_{jt}^{q,Process}$ with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	First Stage	First Stage	Second Stage
	$\text{Ln}(\text{SPILLTECH}^S_{\text{Clean, Process}})$	$\text{Ln}(\text{SPILLTECH}^C_{\text{Clean}})$	$\text{Ln}(\text{MV/TA})_{i,t+1}$
	(1)	(2)	(3)
$\text{Ln}(\text{SPILLENF}^S)$	0.219***	-0.010	
	(12.29)	(-0.48)	
$\text{Ln}(\text{SPILLENF}^C)$	-0.001	0.313***	
	(-0.66)	(40.29)	
$\text{Ln}(\text{SPILLTECH}^S_{\text{Clean, Process}})$			0.308***
			(2.74)
$\text{Ln}(\text{SPILLTECH}^C_{\text{Clean}})$			0.031
			(1.63)
$\text{Ln}(\text{SPILLTECH}^S_{\text{Non-Clean}})$	0.017***	0.048***	-0.031
	(6.39)	(3.52)	(-1.57)
$\text{Ln}(\text{SPILLTECH}^C_{\text{Non-Clean}})$	-0.000	0.143***	-0.012***
	(-0.29)	(49.75)	(-3.12)
$\text{Ln}(\text{SPILLTECH}^S_{\text{R\&D}})$	0.288***	0.043	-0.051
	(6.09)	(1.27)	(-0.69)
$\text{Ln}(\text{SPILLTECH}^C_{\text{R\&D}})$	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 IV: PCA-Weighted Patents

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. We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:} \quad \ln(\text{SPILLTECH}^S_{Clean})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\textbf{2nd Stage:} \quad \ln(\text{MV/TA})_{i,t+1} = \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}}^S_{Clean})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} + \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} + \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} + \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \boldsymbol{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.$$

$\text{SPILLTECH}^S_{Clean}$ is defined as $\sum_{j \neq i} \text{TECH}_{ij} G^q_{jt}$, where G^q_{jt} is the stock of clean patents filed by firm j weighted by a principal component combination of specificity, fullness, and readability, and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j \neq i} \text{PRODSIM}_{ij} G^q_{jt}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean patent stock G^q_{jt} with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	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 IV: PCA-Weighted Patents

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. We estimate the following 2SLS firm-year panel regression:

$$\textbf{1st Stage:}\quad \ln(\text{SPILLTECH}^S_{Clean,Process})_{i,t} = \gamma_0^S + \gamma_1^S \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^S \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^S \mathbf{W}_{i,t} + \lambda_i^S + \lambda_t^S + u_{i,t}^S,$$

$$\ln(\text{SPILLTECH}^C_{Clean})_{i,t} = \gamma_0^C + \gamma_1^C \ln(\text{SPILLENF}^S)_{i,t} + \gamma_2^C \ln(\text{SPILLENF}^C)_{i,t} + \boldsymbol{\Theta}^C \mathbf{W}_{i,t} + \lambda_i^C + \lambda_t^C + u_{i,t}^C.$$

$$\begin{aligned}\textbf{2nd Stage:}\quad \ln(\text{MV/TA})_{i,t+1} =& \beta_0 + \beta_1^C \ln(\widehat{\text{SPILLTECH}^S_{Clean,Process}})_{i,t} + \beta_2^C \ln(\widehat{\text{SPILLTECH}}^C_{Clean})_{i,t} \\ &+ \beta_1^{NC} \ln(\text{SPILLTECH}^S_{Non\text{-}Clean})_{i,t} + \beta_2^{NC} \ln(\text{SPILLTECH}^C_{Non\text{-}Clean})_{i,t} \\ &+ \beta_3 \ln(\text{SPILLTECH}^S_{R\&D})_{i,t} + \beta_4 \ln(\text{SPILLTECH}^C_{R\&D})_{i,t} \\ &+ \beta_5 \ln(R\&D\ Stock/TA)_{i,t} + \beta_6 \ln(G)_{i,t} + \boldsymbol{\Gamma} X_{i,t} + \lambda_i + \lambda_t + \epsilon_{i,t+1}.\end{aligned}$$

$\text{SPILLTECH}^S_{Clean,Process}$ is defined as $\sum_{j\neq i} \text{TECH}_{ij} G_{jt}^{q,Process}$, where $G_{jt}^{q,Process}$ is the stock of clean process-based patents filed by firm j weighted by a principal component combination of specificity, fullness, and readability, with process-based patents defined as in Tham et al. (2026), and TECH_{ij} measures technology proximity (Jaffe, 1986). $\text{SPILLTECH}^C_{Clean}$ is defined as $\sum_{i_{TNIC3}=j_{TNIC3}}^{j\neq i} \text{PRODSIM}_{ij} G_{jt}^{q}$, where PRODSIM_{ij} is between-firm product similarity (Hoberg and Phillips, 2016). The instruments are $\ln(\text{SPILLENF}^S)$ and $\ln(\text{SPILLENF}^C)$, constructed analogously to SPILLTECH by replacing firm j's clean process-based patent stock $G_{jt}^{q,Process}$ with firm j's cumulative number of EPA enforcement actions up to year t. $\mathbf{W}_{i,t}$ denotes the vector of exogenous controls from the second stage (non-clean spillovers, R&D spillovers, own R&D stock, own clean patent stock G, and $X_{i,t}$). 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.

	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