Patent Disclosure and Reliance on External Financing

Donald E. Bowen III
Lehigh University

Kathleen Weiss Hanley
Lehigh University

Sungjoung Kwon
Wayne State University

William Mann
Emory University

Qianqian Yu
Lehigh University

This version: July 22, 2026

Patent disclosures involve a fundamental tradeoff between concealing information from competitors and revealing it to capital providers. Using large language models, we decompose patent disclosure complexity into fundamental complexity and strategic obfuscation for 2.6 million U.S. patents. Exploiting quasi-random examiner assignment, we show that obfuscation deters competitors by reducing citations and litigation but lowers patents' marketability and usefulness as collateral. Three additional disclosure and financing shocks demonstrate that firms reliant on external finance produce clearer patent disclosures. Our findings identify a previously unexplored channel through which financing shapes innovation and the diffusion of technological knowledge.

Keywords: Patent disclosure; Obfuscation; Large language models; Innovation; Collateral; Information asymmetry; Proprietary costs

JEL classification: G32; O31; O34; G14; D82

Corresponding author: Donald E. Bowen III, Perella Department of Finance, College of Business, Lehigh University, 621 Taylor Street, Bethlehem PA 18015; deb219@lehigh.edu. For helpful comments, we thank Tony Cookson, Stephen Glaeser, Scott Guernsey, Winston Xu, Tengfei Zhang and seminar participants at the University of Colorado and the University of Pennsylvania, and conference discussants and attendees at 2026 Western Finance Association, 2026 FMA Napa, 2026 Eastern Finance Association, 2026 Lake District Workshop, and 2026 Clemson Finance Research Conference. First version: November 11, 2025.

AI Disclosure: During the preparation and revision of this manuscript the authors used ChatGPT (OpenAI GPT-5.5, https://chatgpt.com), Refine.ink and Claude Sonnet (Anthropic, https://claude.ai) to assist with language editing and improving clarity. All use of AI tools to generate data are disclosed in the paper and these tools were not used to conduct analyses, generate results, or draw scientific conclusions. All AI-assisted text was reviewed and revised by the authors, who take full responsibility for the manuscript's content. Neither tool was modified or trained on proprietary data for this research.

1 Introduction

The patent system is built on a bargain: inventors receive exclusivity in exchange for the public disclosure of their inventions.[1] The stakes are substantial: according to the U.S. Patent and Trademark Office, patents generate $8 trillion of economic activity each year.[2] Yet disclosure is privately costly. The more clearly a patent explains its invention, the easier it is for competitors to design around the patent and enter related markets without infringing. On the other hand, with deliberately vague language, a firm may be able to obfuscate the boundaries of the patent's protection, honoring the letter of the disclosure requirement but not its substance. Indeed, there is evidence that obfuscation in patent documents is widespread, undermining the very bargain on which the patent system is based (Bhattacharya and Ritter, 1983; Figueroa and Lemus, 2024; Dyer et al., 2024; Ireland et al., 2024).

However, the value of a patent lies not only in its ability to deter competitors, but also in its ability to support external financing (e.g., Mann, 2018). Outside investors also care about what is protected by the firm's patents, and what competitors *think* is protected, since these competitors are the most likely buyers of the firm or its patents. Obfuscated language that creates uncertainty about the patent's boundaries can therefore impair its value and perceived liquidity, and hence the firm's financing capacity. A firm that relies on external financing thus faces a tradeoff: obscure the patent's contribution and boundaries in order to insulate it from competition, or make the patent transparent in order to attract greater outside capital.[3] In this paper, we examine whether firms' reliance on external financing mitigates their incentives to strategically obfuscate patent disclosures.

One empirical challenge is to construct an appropriate measure of obfuscation. Prior work studies obfuscation by measuring the general complexity of a given text, applying tools from computational linguistics to quantify readability (Li, 2008). The major limitation of these methods is the lack of a counterfactual. A patent may be difficult to read simply because the technology being described is inherently complex, without necessarily implying that the firm could have described

[1] As the Supreme Court describes it, "the patent system represents a carefully crafted bargain that encourages both the creation and the public disclosure of new and useful advances in technology, in return for an exclusive monopoly for a limited period of time" (*Pfaff v. Wells Electronics, Inc.*, 525 U.S. 55, 63 (1998)).

[2] https://www.uspto.gov/blog/quality-us-patents-drive-our-economy-and-solve-world-problems

[3] A large literature emphasizes this tradeoff at the firm level: increased disclosure transparency can reduce a firm's cost of capital but reveal costly information to competitors. See Healy and Palepu (2001), Leuz and Wysocki (2016), and Goldstein and Yang (2017) for reviews.

the patent using simpler language.

We propose a novel LLM-based approach that separates the fundamental and strategic components of patent complexity by constructing the missing counterfactual. We instruct a leading LLM, Claude Sonnet 3.5, to provide a clearer restatement of each patent's abstract.[4] We then measure the linguistic complexity of both the original abstract (*Total Complexity*) and the LLM-simplified abstract (*Fundamental Complexity*) using standard linguistic tools.[5] We interpret the difference between the two as a proxy for strategic obfuscation (*Obfuscation*) in the writing of the patent. Our methodology is new, but it has strong parallels with prior work on disclosure and obfuscation, for example in conference calls (e.g. Bushee et al., 2018; Bai et al., 2025), as we discuss in detail below.

We validate our decomposition in several ways. First, we confirm that the LLM maintains the information content of the original patents both by manually inspecting the output for 1,000 randomly chosen patents and by calculating the cosine similarity of the embeddings of the original and LLM-rewritten abstracts. We find that more than 80% of the rewritten abstracts have a cosine similarity above 70% with the original patent, and 99% exceed a similarity of 55%. Second, we show that obfuscation is positively related to information asymmetry and substitutes for trade secrets. Third, we confirm that obfuscation is more common among firms where strategic deterrence or concealment is more valuable, such as firms with larger market share and firms in more concentrated industries. Fourth, we find that more obfuscated patents are used less in follow-on innovation (consistent with deterrence or poorer disclosure reducing knowledge spillovers) and that a reduction in the value of preemptive deterrence (due to competitors receiving valuable patents) is associated with a decline in obfuscation. Importantly, these patterns are not present for fundamental complexity or abstract length, and are weaker for *Total Complexity*.

We next propose a counterpoint to the existing literature on patent complexity and obfuscation: while that literature has primarily highlighted the strategic *benefits* of obfuscation, we argue that obfuscation also carries a *cost* in the form of reduced financing capacity. To motivate this tradeoff, we examine how firms change their disclosure behavior around the time they first pledge one

[4]This model is now deprecated, but we show that its outputs on our task are highly similar to those of more recent generations of models.

[5]We focus on the abstract because it is the first text that algorithmic searches use to identify relevant technologies, because it is required to be a concise summary of the invention rather than a detailed legal description, and for computational efficiency. We discuss this choice in detail in Section 2.

or more patents as loan collateral, an event that indicates the firm is more likely to rely on the financing capacity of subsequent patents. We find that the level of obfuscation in subsequent patent applications drops significantly, consistent with firms reducing obfuscation to enhance the collateral value of their patent portfolio.

The rest of our analysis focuses on further establishing and studying this link between the financing capacity of patents and the tradeoffs inherent in patent disclosure. The identification challenge facing this analysis is to isolate the effects of patent disclosure from firms' broader corporate disclosure practices. Changes in firms' overall transparency or disclosure policies could simultaneously influence both financing outcomes and patent-writing behavior, making it difficult to identify the financing consequences of patent-specific disclosure choices.

Our empirical strategy addresses this identification challenge by exploiting two classes of shocks. The first consists of patent-level shocks that affect firms' ability or incentives to obfuscate patent disclosures while leaving broader corporate disclosure practices unchanged. The second consists of financing shocks that alter firms' reliance on external capital without directly affecting the technological content of their patents. Together, these complementary approaches isolate the financing consequences of patent-specific disclosure choices. Across both settings, we find that firms more reliant on external financing engage in less strategic obfuscation, suggesting that financing needs promote more transparent patent disclosures.

The first patent-level shock we exploit is an instrumental variable for obfuscation based on the quasi-random assignment of patent examiners (Dyer et al., 2024). To validate the instrument, we show that examiner strictness meaningfully affects changes in patent language between application and grant. Using this approach, we find that obfuscated patents receive fewer citations from outside firms and are less likely to be involved in infringement litigation, consistent with a strategic deterrence mechanism (Freilich, 2017; Ashtor, 2022). At the same time, obfuscated patents are less likely to be sold or transferred and less likely to be pledged as collateral, demonstrating that strategic opacity reduces the financing value of patents.

The second patent-level shock is the Supreme Court's decision in *Nautilus v. Biosig* (2014), which strengthened the clarity requirements for patent disclosures under §112 of the Patent Act. Firms with the greatest pre-Nautilus obfuscation significantly reduced strategic obfuscation following the decision, consistent with the ruling limiting firms' ability to obscure their inventions.

However, this response is substantially weaker for firms that have previously pledged patents as collateral or rely more heavily on external debt. Because Nautilus directly affects patent disclosure rather than firms' broader disclosure policies, these results provide additional evidence that financing needs independently discipline patent obfuscation, leaving the Court's mandate less binding for firms already facing strong market incentives to produce transparent patent disclosures.

The final patent-level shock is the enactment of the American Inventors Protection Act (AIPA) in 1999. The Act required US patent applications to be made public eighteen months after filing, rather than upon grant, regardless of whether the application was ultimately successful. Firms that did not seek foreign patent protection could opt out of this accelerated disclosure schedule. Building on Ireland et al. (2024), we treat patents as more exposed if they were issued to firms that sought foreign protection more often pre-AIPA. We find that these patents exhibit increased obfuscation after the Act took effect. However, this response is substantially weaker for patents of firms that relied on patent collateral for financing or were highly leveraged, consistent with the notion that dependence on external finance dampens incentives to obfuscate patent filings.

Our final test exploits an exogenous shock to firms' financing constraints: Moody's 2006 revision to its leverage-adjustment methodology, following Kraft (2015) and Kisgen (2019). A central feature of the revision was the reduced weight placed on underfunded pension liabilities, which lowered adjusted leverage ratios and effectively eased financing constraints for firms with significant pension underfunding. Using a difference-in-differences design, we find that firms with large, underfunded pensions, the treated group, subsequently increase their level of patent obfuscation. In other words, once these firms experience a more favorable reassessment of their creditworthiness and thus become less financially constrained, the benefits of obfuscation begin to outweigh the costs of reduced transparency, leading them to produce more obfuscated disclosures.

Collectively, these findings underscore the inherent tradeoff in patent disclosure strategies. While obfuscation can expand a firm's uncontested technological space, it also reduces the salability and pledgeability of patents. Accordingly, firms that rely on external financing may find obfuscation costly, so financing needs can mitigate the incentives to obfuscate, incentivize transparency, and facilitate the use of intellectual property as collateral. Because obfuscation degrades the public disclosure that the patent system is designed to elicit, our results also imply that capital-market frictions shape not only firm outcomes, but also the quality of the technological knowledge

that enters the public domain.

Our paper makes three important contributions to the literature. First, we introduce a new way to separate the strategic and fundamental components of disclosure complexity that can be applied across a variety of settings. Our approach leverages the widespread availability of AI tools to contribute to the literature on financial disclosure (e.g., Gunning, 1969; Li, 2008; Loughran and McDonald, 2014; Bushee et al., 2018). Rather than relying solely on the linguistic complexity of the original patent abstract, we use AI tools to construct a counterfactual abstract that improves clarity relative to the original text. By applying traditional readability metrics to both the original and counterfactual versions, we decompose textual complexity into two components, a strategic component and a fundamental component. We believe that this decomposition will allow researchers to better identify strategic disclosure by firms in many different contexts. This approach is complementary to Bai et al. (2025), which uses LLMs to generate counterfactual managerial answers in Q&A sessions. While they focus on differences in semantic content and information in their research setting, we focus on differences in clarity and presentation, holding content fixed.

Second, we show that disclosure choices carry financing costs as well as competitive benefits, contributing to a large literature on the tradeoffs firms face in reducing disclosure to hide proprietary information versus increasing transparency to reduce information asymmetry. Disclosure theory proposes that companies, particularly those with developing technologies, may wish to shield proprietary information from competitors and other market participants (Bhattacharya and Ritter, 1983; Horstmann et al., 1985; Maksimovic and Pichler, 2001) as they seek to obtain a competitive advantage over incumbent firms. However, greater opacity can increase information asymmetry between firms and external investors, reduce market liquidity, and raise the firm's cost of capital, particularly when firms rely on external financing (Diamond and Verrecchia, 1991; Botosan, 1997; Leuz and Verrecchia, 2000). The result is an inherent tradeoff between the proprietary costs of revealing sensitive information and the capital-market benefits of transparency.

Prior work predicts that firms optimally adjust both the level and the precision of disclosure in response to this tradeoff, disclosing more when financing needs and information frictions are high, and withholding or obfuscating information when proprietary costs from competition or entry threats are more salient (Verrecchia, 1983; Darrough and Stoughton, 1990; Healy and Palepu, 2001). Our approach advances this literature by measuring disclosure choices at the asset level within

firms. Firm-level disclosure proxies conflate heterogeneous incentives across assets that differ in proprietary value and financing relevance; asset-level measurement lets us test directly how firms tailor transparency to the proprietary and financing considerations attached to individual assets.

Third, we show that financing considerations shape the choices firms make when drafting patent disclosures, extending research on the determinants of disclosure in patent applications. For instance, Dyer et al. (2024) show that examiner leniency with respect to §112 is associated with disclosures that contain less substantive content, and Ashtor (2022) demonstrate more broadly that examination procedures and policies can significantly affect claim clarity. Relatedly, Kong et al. (2023) find that university patents, which are more frequently licensed, are less complex than corporate patents. Because obfuscation reduces knowledge spillovers, our results identify an additional channel through which financing can shape the trajectory of innovation (Brown et al., 2009; Acharya and Xu, 2017).

2 Measuring Patent Obfuscation

2.1 Measurement

A central challenge for our analysis is to determine whether the linguistic complexity of a patent reflects the underlying technical complexity of the invention or represents a deliberate attempt to obscure the text. The key is to identify a counterfactual version of the text that is unlikely to be subject to obfuscation.

Our approach is motivated by Bushee et al. (2018), who focus on the possibility of obfuscation in quarterly earnings calls. To separate the linguistic complexity of corporate disclosure in quarterly earnings conference calls into informative versus obfuscatory components, the authors use a two-step approach. First, they measure informational complexity using analysts' language (measured by the Gunning Fog Index), under the assumption that analysts have no incentive to obfuscate. Second, they measure obfuscation as the residual from regressing managers' linguistic complexity on analysts' linguistic complexity after controlling for business complexity. This residual captures the portion of managers' complexity that exceeds what would be expected given the underlying informational complexity of the topic. Thus, the counterfactual text that they employ is the analyst's summary of the call, under the assumption that analysts do not have the same incentives

to obfuscate as the firm's management.

Our approach constructs counterfactual text for patent filings by exploiting recently-developed large language models (LLMs). We instruct the LLM to rewrite patent text to enhance clarity, without changing the substance of the text. The difference between the linguistic complexity of the original text and the counterfactual version is our measure of obfuscation.[6]

Our analysis focuses on the patent abstract, for several reasons. First, the abstract is the logical starting point for parties who need to understand a patent's contribution: patent searches are often conducted using algorithmic textual analysis with a focus on the abstract to identify relevant technologies, allowing the patent writer to influence what is learned about the patent through the wording of this section.[7] Second, USPTO rules require the abstract to summarize the technical disclosure and prohibit the legal phraseology that characterizes the claims, such as "means" or "said" (MPEP §608.01(b); 37 CFR 1.72(b)). Complexity in the abstract is therefore not a response to legal requirements, but rather is discouraged by the examination process. This feature is important for our measurement strategy: it means that when the LLM simplifies an abstract, it is not stripping out legally required language, and the residual complexity we attribute to obfuscation is not an artifact of drafting conventions imposed by patent law. By contrast, the claims define the subject matter and scope of the patent in technical legal language that is notoriously difficult even for specialists (Bena et al., 2022; Mille and Wanner, 2008). Third, limiting our analysis to the abstract substantially reduces token usage and thus cost.

We begin by downloading the abstracts and identifiers from PatentsView for the full sample of 2.6 million patents granted from 1976 through 2023 that have been matched with public firm identifiers using the link tables of Kogan et al. (2017) (henceforth, KPSS) and the CRSP/Compustat Merged (CCM) database.[8] We then remove patents granted to firms in utility and financial industries (SIC 49 and 60-69).

[6]Other papers also use LLMs to generate counterfactual text for patents and conference calls, though with slightly different goals. Yang (2023) proposes a measure of application quality based on AI-generated patent abstracts. He finds that AI-generated text significantly improves the chances of patent success particularly for smaller firms and less experienced lawyers. Bai et al. (2025) construct a measure of new information provided by managers by comparing the similarity between the manager's answers to questions during conference calls to those generated by ChatGPT for the same questions. A larger discrepancy generates more active trading and larger price responses.

[7]Yang (2023) highlights that download rates for USPTO bulk textual data have increased rapidly over time, and argues that this indicates widespread commercial application of textual analysis to patents.

[8]Available at https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data. We use the October 2024 data release.

We provide each patent abstract to Claude Sonnet 3.5 (with the temperature parameter set to zero) and instruct it to generate, if possible, an LLM-simplified version that represents the abstract's fundamental content using the following prompt:[9]

> *Give me a simpler and clearer version of this patent abstract. Reply with JUST your version, not any label, title, header, or comments.*

We intentionally refrain from instructing the LLM to preserve technical or legal detail, as these are not critical in the abstract, and we keep the prompt deliberately brief to minimize token count.

After obtaining the LLM rewrites, we quantify the complexity of both the original abstract and the LLM-simplified version using tools from the linguistics literature. Specifically, we measure linguistic complexity using the Gunning Fog Index, Flesch-Kincaid Grade Level, Flesch Reading Ease, Automated Readability Index (ARI), Coleman-Liau Index, LIX, RIX, and entropy scores. These metrics use various approaches to estimate the difficulty or readability of a text, often expressed as a grade level.

Using these metrics, we define the amount of obfuscation in patent text as the difference between the linguistic complexity of the original abstract, which is representative of the total complexity, and the linguistic complexity of the LLM-simplified abstract, which represents the fundamental complexity:

$$\begin{aligned} \text{Obfuscation} &= f(\text{Original abstract}) - f(\text{LLM-simplified abstract}) \\ &= \textit{Total Complexity} - \textit{Fundamental Complexity} \end{aligned} \tag{1}$$

where $f(\cdot)$ denotes any of the readability measures, all of which are signed so that they are increasing in the complexity of the text. The greater the difference between the complexity of the original abstract and the LLM-simplified version, the greater the scope for obfuscation in the patent text.

In the paper, we present results using two primary measures of linguistic complexity: the Gunning Fog Index, and a composite measure based on the principal component (PCA1) that combines and exploits information from all the available readability measures (as in Kong et al.,

[9]We access the model through AWS Bedrock (model ID: `anthropic.claude-3-5-sonnet-20240620-v1:0`), which pins a specific model build. At the end of this section, we show that the resulting measure is highly stable across repeated runs with this model and across alternative LLMs.

2023).[10] As can be seen in Appendix Table A.2, PCA1 loads approximately evenly across all the indices except for entropy (the difficulty of predicting the next word in a sequence), while the second principal component (PCA2) loads mainly on entropy.

2.2 Examples and Validation of Information Preservation

Table 1 presents three examples from this process, comparing the original abstract in the first column with the LLM-simplified version in the second column. The rightmost column reports the Gunning Fog Index and PCA1 values for each example, as well as the resulting measure of *Obfuscation*.

The first patent in the table is a potential case of obfuscation: Patent No. 9,547,842, IBM's patent for an email autoreply feature. The invention appears relatively straightforward to describe, yet the abstract is surprisingly complex and contains unnecessary information and ambiguous wording.[11] Indeed, the LLM explains the core idea much more clearly in the second column of Table 1. The original abstract has a Gunning Fog score of 19.71, consistent with postgraduate-level text, whereas the AI-generated abstract has a Gunning Fog score of 12.81, corresponding to a high-school-graduate reading level. Thus, the obfuscation score for this patent is 6.90 grade levels, indicating an economically meaningful reduction in textual complexity.

[Insert Table 1 about here]

The next two examples illustrate that the LLM does not always find opportunities to simplify patent abstracts, even when the patents exhibit different levels of complexity. Patent No. 8,786,025 describes a method of manufacturing resistors that was granted to TSMC for use in semiconductor production. The abstract is written in a direct manner accessible to undergraduate students, leaving relatively little scope for simplification. In fact, the LLM revision is slightly *more* complex. The final example is Patent No. 8,786,825, which was granted to ASML, whose high-end lithography machines have been described as "among the most complex devices humans have ever created" (Mackrael, 2026). Unsurprisingly, ASML's patent on lithographic techniques is the example with

[10] We calculate the principal components for the measures across the combined samples of original and LLM-simplified abstracts.

[11] The Electronic Frontier Foundation called Patent No. 9,547,842 the "stupid patent of the month." https://www.eff.org/deeplinks/2017/02/stupid-patent-month-ibm-patents-out-office-email

the highest level of *Fundamental Complexity* (18.88). Nevertheless, the structure and terms of the granted *'825* abstract are similar to the LLM rewrite, and thus the measured obfuscation is only 1.69, well below the median patent.

Together, the three examples suggest that the LLM preserves the semantic meaning of the patent abstracts, and we find strong systematic evidence that it does so. Specifically, we obtain embeddings of the original and rewritten abstracts and compute the cosine similarity between the two embedding vectors for each patent. The distribution of these similarity scores is displayed in Figure A.1 of the Appendix. More than 80% of patents have a similarity above 70% and virtually all rewrites are above 55%, supporting the conclusion that the rewrites preserve the information contained in the original abstracts. This finding is consistent with recent literature showing that LLM-generated summaries exhibit high fidelity to their source texts (e.g., Huang et al., 2023). As such, we conclude that this approach succeeds at holding content fixed and allows us to isolate differences in clarity and presentation.

2.3 Discussion

As with any application of LLMs to historical data, concerns may be raised about the possibility of look-ahead bias. In our setting, one may wonder whether Claude simplifies patent abstracts by using technical information and vocabulary that are accepted as simple today but did not yet exist at the time the patent was filed. In such cases, look-ahead bias would inflate our measure of obfuscation. We believe that this concern is minor for a few reasons. First, the necessary technical lexicon for the vast majority of patents is well developed by the time the patent is filed, even if the invention is in a new area and especially when it is incremental. Second, even when this concern is present, it would manifest as a strong *downward* trend in obfuscation over time (because the older patents experience more look-ahead bias), a pattern we do not observe in our results below. Third, the variation in this bias is mainly at the level of time and industry or technology class, not the individual patent, meaning it will be absorbed by the fixed effects included in our empirical design.

A related concern is that our measure could reflect the LLM's own notion of "good writing" rather than strategic choices by the patent drafter. Three features of our setting and results argue against this interpretation. First, because USPTO rules make complexity in the abstract discretionary rather than mandated (as discussed above), the simplification opportunities the LLM

identifies are not artifacts of required drafting conventions. Second, as we show next, models from five different LLM developers produce highly correlated versions of the measure, indicating that the rewrites reflect a broadly shared standard of clarity rather than the stylistic preferences of any one model. Third, the results in Section 3.4 indicate that our measure varies systematically with strategic incentives, information asymmetry, and the use of alternative protection mechanisms, patterns that would be difficult to explain if the measure captured an arbitrary stylistic benchmark.

Additionally, a common external-validity question for research that uses LLMs is whether results are sensitive to the model used. The costs of computing our measure preclude a full re-estimation using other models, but we do so for a random subsample of 1,000 patents in the Appendix. Table A.3 shows the correlations between obfuscation estimated with Claude Sonnet 3.5 and six additional LLMs from Anthropic, Google, Meta, OpenAI, and xAI. The lowest correlation is 92% (Llama 3.3), and the average across the other models is 96%, including the two Claude generations released after 3.5. We conclude that the patterns produced by our approach are robust across different LLMs.

Finally, we assess whether our measure is sensitive to the inherent non-determinism of LLM-generated text. Using the same subsample of 1,000 patents, we generate five independent simplified versions of each patent abstract with the same LLM and prompt. We then recompute *Obfuscation* for each version. The pairwise correlations across the five resulting measures all exceed 0.99, indicating that our measure is highly stable across repeated LLM runs.

3 Data and Descriptive Statistics

This section describes our data sources and summary statistics, validates the decomposition, and presents first evidence that obfuscation carries a financing cost.

3.1 Data Sources

We link our patent sample with data on citations within 10 years of the grant date and Cooperative Patent Classification (CPC) technology classes (PatentsView), patent sales and use as collateral (USPTO Assignments database), infringement allegations (USPTO and the Stanford Non-Practicing Entity (NPE) Litigation Database), art-unit assignments (the USPTO Patent Ex-

amination Research Dataset, PatEx), competitor patenting (Hoberg and Phillips, 2016), trade secrets (Glaeser, 2018), and firm-year financials (Compustat). We exclude firms in SIC codes 49 and 60-69 (industrials and finance). Firm-year data are merged with patents using the application year to measure the firm's condition at the time of the patent's invention. Trading data from CRSP is used to estimate a proxy for information asymmetry. Finally, data used to estimate examiner strictness come from the USPTO Office Action Research Dataset and PatEx, and we describe the construction of these variables in detail in Section 4.1.2.

Key variables are defined as follows, and formal variable definitions are listed in Table A.1 in the Appendix. We follow Bushee et al. (2018) and measure information asymmetry with the illiquidity measure of Amihud (2002). Illiquidity is measured for each day in the 25 trading-day period following patent disclosures, winsorized at the event-day level, and then averaged across the patent disclosure window to produce a patent-level measure. We use the percentile rank of this measure in tests. Patent value estimates come from KPSS, and we obtain a measure of whether the patent pertains to a technological area that is rapidly evolving (i.e., following breakthroughs) or stable (*RETech*) from Bowen III et al. (2023). We decompose patent citations into self-cites and non-self cites using the firm identifiers in the KPSS link tables. We determine which patents have been pledged and sold following Mann (2018) and Bowen III (2016), respectively. Pledging data are available for patents granted from 1991 through 2020. Examiner strictness variables used in our IV tests are described in more detail in Section 4.1.2.

Additional tests and descriptive statistics examine variables that capture dimensions of competition. *HHI (SIC3)* is industry sales concentration measured within SIC 3-digit industries. *Peer Patent Count* is the number of patents granted to TNIC3 industry competitors (Hoberg and Phillips, 2016).

3.2 Summary Statistics

Table 2 presents summary statistics for our patent-level dataset of 2,607,015 patents granted from 1976 to 2023, constructed as described in Section 2.

[Insert Table 2 about here]

The average patent abstract has a *Total Complexity* Gunning Fog Index (GF) score of 22.13,

corresponding to a reading level requiring substantial postgraduate education. In contrast, the LLM-simplified abstract, which captures *Fundamental Complexity*, has an average GF score of 12.07, approximately the reading level of a high school graduate. The average difference of 10.06 grade levels forms our measure of *Obfuscation*, indicating considerable scope to simplify patent disclosures without altering their underlying technical content.

A few additional summary statistics are worth noting. Patents in our sample have an average estimated value of approximately $13 million (in 1982 dollars) and receive 3.84 citations within ten years of grant, although the median patent receives none. Fifteen percent of patents are pledged as collateral and 26% are sold, highlighting the importance of patents as financing assets in addition to their role in innovation. Patent litigation is relatively rare, with just over 11,000 patents involved in infringement disputes, although these patents are disproportionately valuable.

3.3 Time Series

Figure 1 shows how patent complexity, decomposed into its fundamental and obfuscation components, has evolved since 1976.[12] The left subpanel of Panel A shows the average Gunning Fog Index score of patents across the sample period.

[Insert Figure 1 about here]

The solid red lines show the average level of *Obfuscation* across nearly five decades and reveal a marked rise since 2000. After a fall from 11 in 1990 to 9.5 in 2000, there is a notable upward trend to almost 12 in 2020. Focusing on the level, the figure shows that patent abstracts are written in a much more complex manner than is necessary to summarize the embedded idea, corresponding to an extra 10 grade levels. This pattern is consistent with concerns in both academic research and the popular press about growing patent obfuscation. Figure A.2 of the Appendix further shows that this rise is a pervasive phenomenon across technology classes.[13]

While *Obfuscation* contains significant amounts of time series variation, Figure 1 shows that the average level of fundamental complexity of patent abstracts, displayed in dashed black lines, is

[12] All tests and figures in the paper use the application date of the patent to align the data with the timing of the inventive act.

[13] It also shows that there is substantial heterogeneity across technology areas at points in time and across time. For example, patents in human necessities are near the bottom of the obfuscation distribution in most years and transport patents are frequently at or near the top, while other categories change rank order over time.

relatively stable over time. Once strategic verbosity is stripped from the text, readability consistently hovers around a Gunning Fog score of approximately 12, a level consistent with high-school comprehension. This stability implies that trends in the *Total Complexity* of patent disclosures are largely driven by the strategic component of patent language.

This stability does not imply that all LLM-simplified abstracts are rewritten to the same readability level. Figure A.3 of the Appendix shows systematic differences in the readability of simplified abstracts across technology classes. This verifies that the LLM is not simply reducing all patents to a similar degree of complexity, but rather encounters significant variation in the fundamental complexity of the technology being described. However, these cross-sectional differences are largely stable over time when comparing one technology class with another.

The right subpanel in Panel A of Figure 1 removes CPC technology-class fixed effects. The resulting time-series patterns directly address another natural concern, namely, that patents may appear more obfuscated simply because technologies become more complex over time. The stability of the fundamental complexity makes it difficult to attribute the rising obfuscation documented in the left subpanel to this mechanical effect or to shifts in the types of technologies patented over time. Panel B replaces the Gunning Fog Index with a principal-component-based measure of linguistic complexity and yields qualitatively identical patterns, highlighting the robustness of the time-series patterns to the choice of complexity measure.

3.4 Correlations

Having described the measure's construction and its aggregate behavior, we now validate the decomposition in the cross-section. Table 3 reports correlations between our text-based measures, *Obfuscation*, *Fundamental Complexity*, *Total Complexity*, and the length of the abstract, and variables capturing information asymmetry, alternative protection mechanisms, strategic incentives, and patent impact. All correlations control for application year, and variables are defined in Section 3.1 and Table A.1.

[Insert Table 3 about here]

The first column shows that *Obfuscation* behaves as strategic disclosure theory predicts. First, obfuscation is positively correlated with information asymmetry, as proxied by illiquidity (Bushee

et al., 2018) following patent disclosures (0.024), consistent with obfuscated writing increasing the difficulty of extracting informative signal from the patent document. Obfuscation is also negatively correlated with the use of trade secrets (−0.020). Because obfuscation is itself a form of intellectual property protection, it can substitute, on the margin, for trade secrecy; indeed, Glaeser (2018) finds that shocks to the use of trade secrets reduce proprietary disclosure in the form of less patenting and more redacted filings.

Second, obfuscation is more common where strategic deterrence or concealment is more valuable. At the firm level, obfuscation is positively correlated with industry concentration (0.068), where the threat of entry into a profitable market is most salient. At the patent level, obfuscation is negatively correlated with RETech (−0.037): it is concentrated in slower-moving technological domains, which are characteristic of mature or stagnant industries where patents tend to serve strategic purposes rather than reflect substantive innovation. Moreover, obfuscation responds to changes in the value of deterrence. After competing firms receive additional patents, the value of preemptive deterrence falls; consistent with this mechanism, obfuscation is strongly negatively correlated with realized peer patenting (−0.201).

Third, more obfuscated patents are used less in follow-on innovation: they receive fewer citations (−0.050), consistent with deterrence.

Importantly, the remaining columns show that these patterns are unique to the strategic component of complexity. *Fundamental Complexity* and *Log(Claims)* flip the sign of every correlation: such patents are more valuable and more cited, are situated in rapidly evolving technology areas, and are relatively prevalent in competitive industries. Abstract length also fails to reproduce the patterns above: longer abstracts are associated with *lower* information asymmetry, *more* trade secret use, and *more* citations, and where its signs do align with *Obfuscation*, the magnitudes are a fraction of the size.[14] Finally, *Total Complexity* carries the same signs as *Obfuscation* but with attenuated magnitudes, because it mixes the strategic and fundamental components.

These correlations are descriptive rather than causal. Nevertheless, they are difficult to reconcile with the concern that our measure merely captures arbitrary writing style. Instead, each component of the decomposition is systematically related to information frictions, competition, and innovation

[14]Claim counts, which Dyer et al. (2024) reports are the strongest predictor of non-final rejections due to insufficient disclosure, behave similarly; their correlation with *Obfuscation* is only −0.07.

outcomes in ways consistent with its proposed economic interpretation. Having established the construct validity of the measure, we next examine the central question of the paper: whether a firm's need for financing affects their incentives to obfuscate their patents.

3.5 *Obfuscation* Falls After Firms Begin Using Patents as Collateral

Table 3 contains one further pattern, and it motivates the remainder of the paper: obfuscated patents have markedly lower estimated market values (-0.219 with the KPSS value measure), the largest correlation in the table. The most natural buyers of a patent are the firm's competitors, and obfuscated language that creates uncertainty about what the patent protects should reduce what they are willing to pay. A lower and less certain resale value matters beyond the patent market itself: it reduces the patent's value in liquidation, which is precisely what lenders care about when they accept patents as collateral.

When firms rely on the collateral value of their patents, they should therefore have incentives to avoid obfuscation and use the clearest possible language. To investigate this prediction within-firm, we examine how obfuscation changes around the first date that a firm pledges any patents as collateral. Typically, this is not an isolated event but rather the beginning of an ongoing lending relationship in which the firm also commits to pledge its future patents as collateral (Mann, 2018). The firm should have a permanently stronger incentive to avoid obfuscation after this date, knowing that its future patents will be added to the collateral pool.

To test this prediction, we estimate a patent-level event time regression specified by

$$y_{p,i,t} = \sum_{k=-5}^{5} \Big(\beta_k \Phi(t - P(i) = k)\Big) + \phi_{c(p),t} + \gamma_i + \epsilon_{p,i,t} \tag{2}$$

where $\Phi(t - P(i) = k)$ is an indicator function for whether the patent p from application year t is k years after the year of firm i's first patent pledge $P(i)$. We omit $k = -1$, so all coefficients are relative to the year before the first pledge. We include patent technology category-by-year ($\phi_{c(p),t}$) and firm (γ_i) fixed effects. The sample includes patents within the [-5,+5] window and patents of firms that never pledge as controls. We estimate Equation 2 for $y \in \{$*Fundamental Complexity*, *Obfuscation*, *Total Complexity*$\}$. We report the event-time coefficients (β_k) in Figure 2 with 95% confidence intervals based on standard errors clustered by firm-year.

[Insert Figure 2 about here]

Figure 2 shows a strong shift in the complexity of firm patent disclosures after the firm starts using patents as collateral in its debt contracts. In the years before a firm's first pledge, the event-time coefficients for the overall complexity of the patent text (reported as blue triangles) are close to zero and statistically insignificant: eventual pledgers show no differential trend relative to firms that never pledge, supporting the parallel-trends assumption underlying this design.

In contrast, patent applications become less complex in the years following a firm's first patent pledge. This decline in complexity is statistically significant at the 5% level in years 3 and 4 and at the 10% level in years 2 and 5. While the estimates for the fundamental complexity of patents (reported as black circles) show a slight decline after pledging starts, the overall decline in complexity is almost fully attributable to a reduction in obfuscation (reported as red diamonds). For example, three years after firms pledge a patent, total complexity is 0.35 grade levels lower than that of control firms, and 81% of this is due to a 0.29 decline in obfuscation (roughly 3% of the sample mean of *Obfuscation*).

This finding strongly suggests that obfuscation becomes less attractive when firms begin to rely on patents as collateral. It is consistent with a more general interpretation: strategic patent obfuscation carries not only the benefits described in prior literature but also a novel cost in the form of weaker access to external financing. The remaining analysis in the paper focuses on establishing that interpretation more rigorously.

Additionally, we acknowledge that our measure of obfuscation may reflect not only the strategic incentives that are our focus, but also non-strategic factors that affect the patent drafting process, such as legal requirements. The findings above include year or technology-by-year fixed effects, which will absorb the level component of non-strategic factors at a point in time. In the remaining sections of the paper, we improve on this by exploiting multiple identification strategies to isolate exogenous variation in patent obfuscation, disclosure incentives, and financing constraints.

4 Shocks to Patent Disclosure

In this section, we examine the tradeoffs associated with patent obfuscation by exploiting several shocks to firms' ability or incentives to obfuscate patents. Our central hypothesis is that firms

balance the strategic benefits of obfuscation against its financing costs. If clearer patent disclosures facilitate external financing by making patents easier for investors and lenders to value, then firms that rely more heavily on external finance should have weaker incentives to obfuscate their patents. Consequently, when exogenous shocks alter firms' ability or incentives to use obfuscated disclosures, financing-dependent firms should respond less because they already face stronger incentives to maintain transparent patent disclosures.

We examine this prediction using three complementary research designs. First, we employ an instrumental variable (IV) approach based on patent examiner §112 strictness following Dyer et al. (2024). Second, we exploit the Supreme Court's decision in Nautilus v. Biosig, which raised the legal standard for patent disclosure. Third, we analyze the passage of the American Inventors Protection Act (AIPA), which accelerated the publication of patent applications. Each shock directly affects the patent disclosure process without directly affecting firms' financial reporting obligations or other voluntary corporate disclosures. As a result, they provide a cleaner setting to identify the consequences of changes in patent disclosure, rather than shifts in a firm's overall disclosure strategy.

4.1 IV Analysis: Patent Examiner §112 Strictness

Our first strategy uses an IV approach to identify the causal effects of patent obfuscation. Following Dyer et al. (2024), we exploit plausibly exogenous variation in patent obfuscation arising from the quasi-random assignment of patent examiners with different propensities to enforce disclosure standards under §112. This approach allows us to determine whether the benefits of obfuscation in deterring competitors, as well as its costs in reducing the financing value of patents, reflect causal effects rather than endogenous disclosure choices by firms.

4.1.1 Institutional Details on Patent Examination

Our instrument exploits two features of the patent examination process: patent applications are assigned quasi-randomly to examiners within art units, and examiners differ in the stringency with which they enforce the disclosure requirements of §112. We briefly summarize the institutional details underlying this identification strategy below. Upon arriving at the USPTO, patent applications are received by a central office that performs an initial review to ensure applications are ready

for examination. Applications are then assigned a filing date, patent class, and subclass codes, and allocated to an art unit specializing in the relevant technology area (Lemley and Sampat, 2012; Gaulé, 2018). After an application is delegated to an art unit, the supervisory patent examiner (SPE) assigns it to a patent examiner, who maintains responsibility for examining the application throughout the process until it is either issued or abandoned (Cockburn et al., 2002; Carley et al., 2015). After receiving the assignment, the examiner evaluates the application and issues a preliminary determination of its patentability, the "first-action decision", which is communicated to the applicant through an official letter.

The most frequent grounds for patent rejection are noncompliance with the following sections of the code: §101 (subject-matter eligibility), whereby the claimed invention does not fall within statutory categories or is directed to a judicial exception without additional inventive elements; §102 (novelty), where the invention is already disclosed in prior art; §103 (non-obviousness), where the invention would have been obvious to a person having ordinary skill in the art in light of existing knowledge; and §112 (written description and claim clarity), where the specification fails to adequately describe the invention or the claims are indefinite or ambiguous. For our identification strategy, we focus on examiners' rejection rates based solely on §112, excluding applications with rejections on other grounds. This approach generates exogenous variation in the degree of patent obfuscation that is independent of the underlying quality of the invention or applicant.

Our instrument for patent obfuscation, examiner strictness with respect to disclosure-related reasons (§112), is motivated by the following. First, both the institutional details and the existing literature (see, e.g., Lemley and Sampat, 2012 and Sampat and Williams, 2019) suggest that the assignment of patent applications to examiners is quasi-random within an art unit and year. For example, interviews with patent examiners conducted by Lemley and Sampat (2012) document that applications are assigned based on the last four digits of the application number, docket flow management, or familiarity with certain technologies, none of which is related to the application or applicant (firm) quality.[15]

Second, patent examiners exercise considerable discretion in the review process, and they differ significantly in their strictness regarding compliance with disclosure requirements. Therefore, a

[15]In October 2020, the USPTO replaced its quasi-random assignment process with automated examiner routing (Armstrong et al. (2024); Pairolero and de Grazia (2026)). Therefore, we restrict the IV analysis to applications assigned before the policy change.

patent application reviewed by a stricter examiner is more likely to receive a first-action rejection on disclosure-related grounds and to be revised with less obfuscation to pass the review process.

Importantly for our abstract-based measure, examiner strictness also applies to the abstract of the document. Patent examiners are required to ensure that patent abstracts comply with the USPTO guidelines. According to the Manual of Patent Examining Procedure (MPEP) 608.01(b), if an application lacks an abstract or the abstract does not meet the guidelines, the examiner must point out the defect and require correction.[16] This implies that examiners are likely to actively enforce clarity and compliance in abstracts as part of the examination process.

4.1.2 Data on Patent Examiners and Construction of IV

Our instrument for patent obfuscation, *Examiner §112 Strictness*, is computed as the fraction of an examiner's total applications that receive non-final rejections in the first action solely due to failure to meet disclosure clarity requirements under §112, excluding the focal application.

To construct our instrument, we first utilize the USPTO Office Action Research Dataset, which provides information on the first-action decisions for all patent applications filed between 2008 and 2023, irrespective of their ultimate disposition (granted, abandoned, or pending). Specifically, the Office Action data list whether a patent application has been issued a §101, §102, §103, or §112 rejection. We then merge these records with PatEx to compile the track record of rejection rates by reason for each examiner over the same period. Finally, we restrict the sample to end in 2020 to accommodate the change in how patents are assigned to examiners.

For a patent application a reviewed by examiner k, we compute the instrument as:

$$\text{Examiner §112 Strictness}_{ka} = \frac{\text{Total §112 Rejections}_k - 1(\text{§112 Rejection}_a = 1)}{\text{Total Applications}_k - 1}. \tag{3}$$

In the above, Total §112 Rejections$_k$ and Total Applications$_k$ are the total number of §112 rejections that examiner k issued at the first action and the total number of applications she reviewed. 1(§112

[16] https://www.uspto.gov/web/offices/pac/mpep/s608.html: "The Office of Patent Application Processing (OPAP) will review all applications filed under 35 U.S.C. 111(a) for compliance with 37 CFR 1.72 and will require an abstract, if one has not been filed. In all other applications which lack an abstract, the examiner in the first Office action should require the submission of an abstract directed to the technical disclosure in the specification."... "If the abstract contained in the application does not comply with the guidelines, the examiner should point out the defect to the applicant in the first Office action, or at the earliest point in the prosecution that the defect is noted, and require compliance with the guidelines."

Rejection$_a$=1) is equal to one if patent application a was issued only a §112 rejection (and not §101, §102, or §103) by examiner k who reviewed it and zero otherwise. To isolate an examiner's focus on clarity and ensure the instrument is not driven by low quality patents flagged on multiple dimensions, Total §112 Rejections$_k$ counts patents that *only* cite §112 in the first action decision (and not §101, §102, or §103). In doing this, we exclude the focal application, following the existing literature (e.g., Gaulé, 2018; Melero et al., 2020).[17]

In our IV regressions, we include art unit–by–year fixed effects, as both institutional details and prior research indicate that the assignment of examiners to patent applications is quasi-random within an art unit and year. Our findings remain robust when we instead employ an art unit–adjusted measure of examiner §112 strictness, defined as an examiner's individual §112 strictness relative to the average §112 strictness within the same art unit, in place of art unit–by–year fixed effects.

To validate our use of the IV, Table 4 examines a subset of patent applications for which we observe the evolution of the patent text throughout the examination process, from initial review and rejection to final allowance. We relate changes in patent readability to examiner rejection rates under §112, while separately controlling for rejection rates under §101, §102, and §103. Our estimates show that patents assigned to examiners with higher §112 rejection propensities become significantly more readable during the examination process. This finding is consistent with Ashtor (2022), who shows that patent claims rejected under §112 become more simplified during prosecution. We extend this evidence by showing that the same pattern also holds for patent abstracts, which form the basis of our measure of patent obfuscation.

[Insert Table 4 about here]

4.1.3 IV Results

We report the results of our IV analysis in Table 5. Patent obfuscation and fundamental complexity are measured based on readability levels, using either the Gunning Fog Index in Panel A or the first principal component (PCA1) in Panel B. Because the results across panels are similar both

[17] Following Dyer et al. (2024), we calculate examiners' §112 strictness using the full sample, as this approach leverages more data and thus provides greater variation. This is also supported by existing evidence suggesting that examiners' strictness in issuing rejections is at least partly driven by examiner-specific, time-invariant attributes (e.g., Lemley and Sampat, 2012).

statistically and economically, we focus our discussion on Panel A.

In column 1, we report the first-stage result. As expected, examiner §112 strictness is negatively associated with obfuscation, and the coefficient is significant at the 1% level. The first-stage F-stat is 27.71, which is significantly greater than the critical value suggested in Stock and Yogo (2005). This demonstrates that examiner §112 strictness is a strong determinant of patent obfuscation, confirming that the relevance condition required for a valid instrument is satisfied.

[Insert Table 5 about here]

The remainder of Table 5 presents the second-stage results, examining the causal impact of patent obfuscation on forward citations (columns 2-4), patent litigation (column 5), and financing outcomes (columns 6 and 7).[18] Column 2 shows that obfuscated patents receive fewer citations overall, with the decline coming both from outside firms (column 3) and the inventing firm itself (column 4).[19]

Decomposing the drop in citations supports the notion that obfuscation carries strategic benefits. The lower level of citations from other firms is consistent with obfuscation deterring other firms from operating in related technological areas or pursuing related patents, because obfuscation makes it costlier for outsiders to learn what the patent covers and to assess their exposure to infringement. Moreover, obfuscated patents receive fewer citations from the inventing firm. This is consistent with the deterrent effect reducing their incentives to build defensive thickets of related patents, although it could also reflect a cost of obfuscation if opaque disclosure makes it harder for the firm's own researchers to build on the invention. In column 5, we identify a novel benefit of patent obfuscation that is also consistent with this deterrence-oriented view: obfuscated patents are less likely to be infringed upon, and less likely to be involved in patent litigation. Because litigation is rare, the magnitude is meaningful: a one-standard-deviation increase in obfuscation lowers the probability of litigation by 0.046 percentage points, roughly 11% of the litigation rate in the IV sample.

[18] To be precise, the IV captures the effect of obfuscation for the subset of "complier" patents, that is, patents that actually exhibited lower (higher) levels of obfuscation due to the assignment to a strict (lax) examiner.

[19] The implied reduced-form effect is remarkably similar to that reported by Dyer et al. (2024). A one-standard-deviation increase in examiner §112 strictness reduces patent obfuscation by 0.01 standard deviations in the first stage, which, combined with the second-stage coefficient of -2.162, implies an increase of approximately 0.022 log points in patent citations. This closely matches the 0.017 log point effect implied by Dyer et al. (2024).

Based only on these perceived benefits, one might expect that firms tend to obfuscate as much as possible, and that strong enforcement is required to achieve the public good of disclosure that is core to the patent system. However, the remaining results in Table 5 demonstrate offsetting costs of obfuscation. In column 6, we find that obfuscated patents are significantly less likely to be resold, which is consistent with the notion that obfuscation reduces the liquidity of patents. In column 7, we find that obfuscated patents are less likely to be pledged as collateral for loans. In economic terms, a one-standard-deviation increase in obfuscation reduces the probability that a patent is sold by 0.38 percentage points, approximately 6% of the IV-sample mean, and reduces the probability that it is pledged as collateral by 0.20 percentage points, or 1.3% of the mean. This draws a direct connection between patent obfuscation and external financing.

A natural question is whether firms can undo these financing costs by communicating privately with capital providers, since lenders in particular often acquire borrower-specific information outside of public disclosures (e.g., Plumlee et al., 2015). Two features of our evidence suggest that this alternative channel is incomplete. First, to the extent that private communication substitutes for public clarity, it should attenuate the relationship between obfuscation and financing outcomes, implying that our estimates understate the underlying tradeoff. The negative effect on pledging that we nevertheless observe indicates that private disclosure does not fully offset opaque public filings. Second, obfuscation also reduces the probability that a patent is sold, pointing to a loss of redeployability and liquidation value, components of collateral value that sharing private signals with a current lender cannot restore. Consistent with this interpretation, Section 5 shows that firms increase obfuscation after a positive shock to their credit profile, which would be surprising if private channels fully neutralized the financing costs of opacity.

Note that our IV analyses control for examiner strictness based on §101-103 grounds (usefulness, novelty, non-obviousness), which captures the extent to which examiners scrutinize substantive aspects of the patented invention in a patent application. This ensures that our main coefficient is identified from variation in how the invention is written. For five of the six outcome variables, the coefficient on *Ex. Other Strictness* is not statistically significant.

We also include *Fundamental Complexity* to control for the inherent complexity of the ideas embedded in the patent. We find that patents that are fundamentally less complex and more readable receive more citations, are more likely to be resold, and are more frequently pledged as

collateral. Each of these estimates is statistically significant at conventional levels and is about one-fifth of the size of the obfuscation coefficients.

Overall, our findings in this section support a causal interpretation of the costs and benefits of patent obfuscation: while obfuscation deters competition and reduces the likelihood of patent infringement, it also hinders firms' access to external financing.

4.2 Nautilus v. Biosig

The next shock to patent disclosure that we analyze is *Nautilus v. Biosig* (2014), a U.S. Supreme Court ruling that tightened the standard for evaluating patent claim definiteness under 35 U.S.C. §112. The case involved Biosig Instruments, a medical equipment company that held a patent on a heart-rate monitor, and Nautilus, a fitness equipment manufacturer that challenged the patent on the grounds that its claims were ambiguous and too indefinite to enforce. Prior to this ruling, patent claims had to be clear enough so they were not "insolubly ambiguous," establishing a fairly permissive threshold that tolerated substantial ambiguity in claim language. This lenient standard effectively allowed patent applicants considerable latitude in drafting vague or imprecise claims, creating strategic opportunities for obfuscation while maintaining patent enforceability. The Federal Circuit upheld the patent under the prevailing "insolubly ambiguous" standard, finding that while the claims might lack precision, they were not so ambiguous as to be incapable of any reasonable interpretation.

In a unanimous decision, the Supreme Court rejected the "insolubly ambiguous" standard as insufficiently rigorous and inconsistent with §112(b)'s requirement that patent specifications "conclude with one or more claims particularly pointing out and distinctly claiming the subject matter." The Court established a new standard requiring that "a patent is invalid for indefiniteness if its claims, read in light of the specification and prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention."[20]

The "reasonable certainty" standard represented a middle ground: more demanding than the Federal Circuit's permissive approach, yet mindful of what the Court called "the inherent limitations of language." Consistent with the ruling's objective of preventing unclear claim boundaries,

[20] *Nautilus, Inc. v. Biosig Instruments, Inc.*, 2014, p. 910, https://supreme.justia.com/cases/federal/us/572/898/

Ashtor (2022) documents a significant increase in §112(b) rejections post-*Nautilus*. The decision thus heightened the standard of clarity in patent applications and increased the expected costs of obfuscation by raising the likelihood that strategically vague patents would be rendered unenforceable. Although the ruling directly concerned claim language, its practical effect was to raise the premium on clarity throughout the application, and examiners enforce clarity guidelines for the abstract as well (see Section 4.1). We therefore expect the heightened standard to be reflected in our abstract-based measure.

The *Nautilus* decision provides a plausibly exogenous shock to the feasibility of patent obfuscation, generating time-series as well as cross-sectional variation in disclosure practices. If financing considerations already discipline disclosure choices, the ruling should bind least where those considerations are strongest: firms reliant on external financing obfuscate less to begin with, and thus have less obfuscation to unwind when the ruling arrives. We therefore predict that the post-ruling decline in obfuscation is smaller among firms that are financially constrained or reliant on external financing.

To examine how the ruling influences firms' patent obfuscation, we implement a difference-in-differences design in Table 6 using 503,093 patent applications filed three years before and after the decision date (June 2, 2014). We require firms to have filed patent applications both before and after the ruling.

[Insert Table 6 about here]

In Panel A, the dependent variable is our obfuscation measure constructed using the Gunning Fog index. Our treatment variable, *Exposure*, is defined as the firm's median obfuscation level across patent applications filed before the ruling, and *Post* is an indicator variable equal to one if a patent application is filed after June 2, 2014 (the decision date), and zero otherwise. The motivation for *Exposure* is straightforward: while the decision imposed more stringent disclosure requirements on all applicants, firms with higher pre-ruling obfuscation levels face greater pressure to modify their disclosure practices than firms with lower obfuscation levels.

As shown in column 1, the interaction term *Exposure* × *Post* is -0.158 and statistically significant at the 1% level. In terms of economic magnitude, a one standard deviation increase in pre-period

exposure corresponds to a 5.4 percent reduction in obfuscation.[21] This finding shows that the ruling reduced obfuscation most among firms with more obfuscated patents and is in line with Ashtor (2022), who reports a significant improvement in patent claim clarity following the ruling.

Next, we examine whether the *Nautilus* ruling had differential effects on firms hypothesized to rely more heavily on external financing. We consider firms that pledge their patents as collateral and those with high leverage as more reliant on external financing. In column 2, we interact our treatment variables with an indicator *Pledged* (the *Z* variable in the table), equal to one if the firm pledged patents as collateral during the 2011-2013 pre-ruling period, and zero otherwise. The coefficient on *Exposure* × *Post* is negative and significant, indicating that among non-pledging firms, those with higher pre-ruling obfuscation reduced obfuscation more relative to firms with lower pre-ruling obfuscation. However, this reduction is significantly attenuated for firms that had previously pledged patents as collateral: the coefficient on *Pledged* × *Exposure* × *Post* is positive (0.163) and significant at the 1% level. Thus, for pledging firms, a one standard deviation increase in pre-period exposure implies a decline in obfuscation that is 5.5 percent smaller than that of non-pledging peers.

To examine how the effect of *Nautilus* varies with a firm's reliance on debt capital, in column 3 we interact our treatment variables with an indicator *HighLev* (the *Z* variable), equal to one if the firm's average leverage during 2011-2013 exceeds the median of its Fama-French 12 industry, and zero otherwise. We find that the reduction in obfuscation is similarly attenuated for highly levered firms: while the coefficient on *Exposure* × *Post* remains negative and significant (-0.271), the coefficient on *HighLev* × *Exposure* × *Post* is positive (0.292) and significant at the 1% level. These coefficients almost completely offset each other, indicating that the decrease in obfuscation was concentrated among firms with relatively low leverage.

Columns 4-6 indicate that the results are robust to the inclusion of patent application year fixed effects and firm fixed effects.[22] In addition, as shown in Panel B, the conclusions are similar when obfuscation is measured using the first principal component of the readability measures (PCA1).

These patterns are consistent with the hypothesis that external financing creates incentives for

[21] The standard deviation of pre-period *Exposure* is 3.3, and the mean of obfuscation is 9.7. It follows that -0.158*3.3/9.7 = -0.054, or -5.4%.

[22] Note that *Post* dummy is not dropped from the specification because it is not perfectly collinear with year fixed effects. In 2014, some patent applications fall in the pre-period, and others in the post-period.

transparent patent disclosure that substitute for the legal discipline imposed by the Court. Firms that depend more heavily on external finance already face stronger incentives to avoid strategically vague patent disclosures because doing so reduces the financing value of their patents. Consequently, *Nautilus* requires fewer changes to their disclosures, leading to smaller reductions in patent obfuscation.

4.3 The Enactment of AIPA

The next shock to patent disclosure that we study is the American Inventors Protection Act (AIPA), which was enacted on November 29, 1999, and applied to patents filed on or after November 29, 2000. Prior to its enactment, the United States maintained a system under which patent applications remained confidential throughout the examination process. Applications were only published upon patent grant, and applications that were abandoned or rejected were never publicly disclosed. This regime stood in contrast to the publication practices of most other major patent jurisdictions, including Europe and Japan, which had adopted early publication systems whereby applications were automatically published 18 months after filing, regardless of examination outcome.

The AIPA altered U.S. patent disclosure practices by requiring publication of most patent applications 18 months after their earliest effective filing date, including any earlier foreign priority filing, harmonizing the U.S. patent system with the approach of most other countries. This change weakened the information asymmetry present in the patenting system by forcing disclosure of U.S. applications that otherwise could have remained confidential until they were granted (or forever if they were never granted). For applicants who also filed abroad, much of the underlying technical content was already published by foreign patent offices 18 months after the first filing; for these applications, the AIPA's main effect was to accelerate and centralize dissemination through the USPTO, making the disclosure more accessible and salient to U.S. audiences (Armstrong et al., 2024). However, as demonstrated by Ireland et al. (2024), firms responded in part by using greater rates of complex language in the patents that they were no longer able to conceal, suggesting that the new regime created incentives for obfuscation. We show that these incentives are not uniform across firms. While opacity brings the strategic benefit of concealing information from competitors, it simultaneously conceals that information from providers of external capital, impairing their ability to value the firm's patent portfolio. Thus, we predict that while AIPA should lead to an increase

in obfuscation for the average firm, this effect should be significantly weaker among firms that rely more heavily on external financing.

We adopt the identification strategy of Ireland et al. (2024). To identify patents exposed to AIPA, they note that the Act included an important exception: applicants could opt out of publication by certifying that the invention had not been and would not be the subject of a patent application filed in another country requiring publication 18 months after filing. This opt-out provision created a natural experiment by generating cross-sectional variation in publication requirements based on firms' international patenting strategies. Based on this insight, Ireland et al. (2024) classify patents as treated or control based on whether they seek foreign protection. They find that after AIPA, linguistic complexity increases significantly for patents that seek foreign protection relative to those that do not. Their interpretation is that firms strategically respond to mandatory disclosure by increasing obfuscation in patents that can no longer remain confidential.[23]

The decision to seek foreign protection post-AIPA, however, is itself a choice and may be endogenously affected by the change in regime. We mitigate this potential endogeneity by constructing a firm-level exposure measure based on pre-AIPA international patenting behavior. Specifically, we construct *Exposure* as the percentage of the firm's patents that sought foreign protection in 1996.[24] This pre-treatment measure captures firms' established international patenting strategies while remaining unaffected by AIPA itself.

We use this *Exposure* measure to confirm that the passage of AIPA led to an increase in obfuscation among more exposed firms. We predict that this effect should be weaker among firms that are more reliant on external financing. To do this, we follow the same logic as in Table 6 by interacting the overall effect of AIPA with indicators for firms that previously pledged patents as collateral or had above-median leverage ratios in our sample. The sample consists of all patents

[23]We confirm the principal finding of Ireland et al. (2024) using our measure of patent obfuscation. In untabulated analyses, we find that patents seeking foreign protection after the passage of the AIPA exhibit a relative increase in obfuscation compared to patents that did not. Whereas Ireland et al. (2024) measure the overall linguistic complexity of patent disclosures, our measure isolates strategic obfuscation from the underlying technical complexity of the invention. Thus, our results indicate that the AIPA increased strategic obfuscation rather than merely increasing the overall complexity of patent disclosures.

[24]We use 1996 as the reference year because it was the last year before legislative efforts to enact AIPA-style rules began. The initial failed attempt (HR 400) was introduced in 1997, and the debate culminated with AIPA's enactment in November 1999. Importantly, the mandatory publication provisions remained substantively unchanged from the initial draft bill through final enactment, while legislative debate centered on other aspects of patent reform such as prior user rights and patent term adjustments. This stability in the publication requirement supports our use of 1996 filing patterns as a pre-determined measure of exposure to the eventual mandatory disclosure regime.

with application years from 1990 to 2010 that can be matched to publicly traded firms, excluding utilities and financial firms (SIC codes 49 and 60-69). We exclude applications filed in 2000 from the analysis, as this transition year falls after AIPA's enactment (November 1999) but before the law's effective date for new applications (November 29, 2000). However, our results are not sensitive to this choice. Our results are presented in Table 7.

[Insert Table 7 about here]

In column 1 of Table 7, Panel A, the outcome variable is our obfuscation measure based on the Gunning Fog index. We regress this on an indicator variable if a patent is filed after 2000 (*Post*), the fraction of the filing company's 1996 patents that sought foreign protection (*Exposure*), and an interaction between the two. We find that patents that were more likely affected by the new AIPA disclosure requirements (due to seeking foreign protection) were more obfuscated post-AIPA. For a firm that filed only patents seeking foreign protection, as opposed to a firm that filed none (a 0 to 1 change in *Exposure*), the average patent pre-AIPA was written at a level of 5.3 grade levels higher according to Gunning Fog, but this gap grew by 1.3 grade levels to 6.6 post-AIPA. (For reference, the pre-period standard deviation of *Exposure* was 0.3, so a one standard deviation increase in exposure corresponds to a 0.38 increase in obfuscation, which is a 3.9% increase relative to its mean.) The finding is consistent with the perspective that greater disclosure is costly to patenting firms, and they counteract this effect through greater obfuscation, making it harder for outsiders to learn about the invention from the now-public filing.

Next, we examine whether the effect of the Act on disclosure is weaker for firms with high reliance on external financing. In column 2, we include an indicator variable for whether the filing firm pledged patents as collateral during the pre-AIPA period (*Pledged*, the Z variable in the table) and interact it with all the explanatory variables from column 1. Among firms that had not pledged, the treatment effect ($Exposure \times Post$) rises to 4.2. By contrast, the triple interaction term ($Z \times Exposure \times Post$) is -2.9, indicating that patent pledging moderates the effect of AIPA. The results indicate that firms were much more likely to obfuscate their patents in response to AIPA if they did not rely on the collateral value of those patents.

In column 3, we directly measure the firm's reliance on external financing based on its leverage ratio ($Z = HighLev$) to see if this moderates the effect of AIPA on patent obfuscation. We find

that among low-debt firms, the effect of AIPA (*Exposure* × *Post*) is more than twice as large as the average effect in column 1 (2.9 compared to 1.3). However, for high-debt firms this effect is absent: the triple-interaction term ($Z \times$ *Exposure* × *Post*) is -3.842, more than offsetting the effect among low-debt firms, such that the effect of AIPA on obfuscation for high-debt firms is slightly negative (-0.938) and not significantly different from zero. As in column 2, the results indicate that firms did not obfuscate their patents in response to AIPA when they were more reliant on external financing.

In Panel B, we use the first principal component as the measure of linguistic complexity and the results are qualitatively similar. Both panels also display results with firm and year fixed effects added to the specification in columns 4-6. The interaction effect with patent pledging becomes insignificant in both panels, but otherwise the results are similar in each specification.

Overall, this analysis indicates that firms have an incentive to increase the obfuscation of their patents in order to mitigate the new disclosure requirements imposed by AIPA. Our results are also consistent with Saidi and Žaldokas (2021), who show that firms that are more affected by the new disclosure requirements of AIPA find it easier to establish new banking relationships. Enhancing transparency in patent applications, therefore, may facilitate access to capital markets.

5 Shock to Financial Constraints

The prior section exploits disclosure shocks that alter firms' ability or incentives to obfuscate patents. To explore the other side of the relationship and examine how financing shocks impact patent obfuscation, this section employs an identification strategy that uses Moody's 2006 change in its leverage adjustment methodology as a plausibly exogenous shock to affected firms' financial constraints. Our approach builds on Kraft (2015) and Kisgen (2019), which demonstrate that credit rating agency methodology changes can generate exogenous variation in firms' financing conditions.

Credit ratings play a critical role in shaping firms' access to debt financing: rated firms have significantly greater access to public debt markets, face lower borrowing costs, and adopt higher leverage ratios than comparable unrated firms (Faulkender and Petersen, 2006). Rating downgrades can trigger restrictive covenants, increase the cost of existing debt, and impair firms' ability to refinance maturing obligations. Consequently, firms actively manage their financial policies to

maintain favorable ratings, treating rating thresholds as binding constraints on their financing decisions (Kisgen, 2019).

Leverage is among the most important quantitative inputs to a rating, but when calculating leverage, rating agencies do not rely solely on GAAP-reported debt figures. Instead, they routinely adjust firms' reported financial statements to account for obligations that create debt-like claims but are not classified as debt under GAAP, such as defined benefit pension liabilities, operating leases, securitization transactions, and hybrid securities. The rationale for these adjustments is to reflect the underlying economics of firms' obligations and to improve comparability across firms with different accounting treatments (Moody's, 2006).

In February 2006, Moody's announced a significant revision to its methodology for adjusting several off-balance-sheet obligations, most notably underfunded defined benefit pension liabilities. This change had substantial implications for how Moody's assessed the creditworthiness of firms with large pension obligations. Under the pre-2006 methodology, Moody's treated the full amount of underfunded pension liabilities (the difference between pension benefit obligations and plan assets) as debt-equivalent when computing adjusted leverage. The logic was straightforward: underfunded pensions represent a contractual obligation to employees that must ultimately be satisfied, either through future contributions or lump-sum payments, creating a claim on the firm's cash flows similar to debt.

The 2006 revision fundamentally changed this treatment. Rather than adding 100% of the underfunded amount to debt, Moody's instead modeled pension funding as occurring gradually over time through contributions that would preserve the firm's existing capital structure. Specifically, Moody's assumed that firms would fund their pension shortfalls using the same debt-to-equity mix currently employed in their operations. Under this "going concern" approach, only a fraction of the underfunded pension liability, corresponding to the debt portion of the assumed funding mix, would be added to adjusted debt.

This methodological change mechanically reduced the adjusted leverage of firms with substantial underfunded pensions, even though their underlying economic fundamentals, operating performance, cash flows, actual GAAP-reported debt levels, and pension funding status remained unchanged.

The key insight for our identification strategy is that this methodology change generated hetero-

geneous exposure to an improved credit profile based on firms' pre-existing pension funding status, a characteristic unlikely to be directly related to patent disclosure strategies. Firms with larger underfunded pension obligations experienced a more favorable reassessment of their creditworthiness following the methodology change, effectively relaxing their financial constraints.

We exploit this plausibly exogenous shock to test whether financial constraints causally affect patent obfuscation decisions. If access to financing shapes disclosure strategies when filing a patent, we predict that firms experiencing a relaxation of financial constraints should subsequently increase patent obfuscation.

[Insert Table 8 about here]

To test this conjecture, we employ a difference-in-differences framework, regressing our measures of obfuscation and fundamental complexity on the interaction between an indicator for firms with large, underfunded pensions (*Underfunded*) and an indicator for the post-2006 period (*Post*). The sample for this analysis includes patents applied for between 2005 and 2008 by firms rated by Moody's Investors Service and covered in the Moody's Financial Metrics database. The results are reported in Table 8.

In column 1, we find that firms with large, underfunded pensions (treated firms) exhibit greater patent obfuscation in 2007 and 2008 following Moody's revision to its adjustment methodology. Specifically, their patents increase 0.584 grade levels more than those of control firms, which corresponds to 6.0% of the mean for the pre-period dependent variable within this test's sample. However, as shown in column 2, these treated firms do not appear to alter the fundamental content of their patent abstracts after the revision. Columns 3 and 4 report results using measures of obfuscation and fundamental complexity based on the first component of PCA, and again show an increase in obfuscation.

We interpret these findings to mean that when firms become less financially constrained and therefore potentially less reliant on pledging patents for financing, they are more likely to increase their obfuscation. In other words, under certain circumstances, the benefits to obfuscation outweigh the potential cost of reducing its value as collateral.

Taken together, the four designs in Sections 4 and 5, quasi-random examiner assignment, a judicial clarity mandate, a statutory acceleration of disclosure, and a credit-rating methodology

change, rely on different samples, different sources of variation, and different identifying assumptions. Yet they paint a consistent picture: obfuscation carries real financing costs; firms whose reliance on external finance already disciplines their drafting respond least when the returns to obfuscation change; and when financing constraints loosen, obfuscation rises.

6 Conclusion

The patent system is built on a basic bargain: inventors receive legal protection in return for disclosing knowledge that supports further innovation. This paper shows that the nature of that exchange can be shaped by firms' financing needs. We introduce an LLM-based method that distinguishes the fundamental clarity of a patent abstract from the additional complexity associated with strategic obfuscation, providing a scalable and generalizable measure of disclosure quality.

Using this measure, we document that obfuscation is higher when a firm faces more competitive threats, decreases after firms begin using patents as collateral, and has increased substantially over the past 15 years across all major technology sectors. This rise in strategic complexity carries meaningful economic consequences.

We exploit several exogenous shocks to firms' incentives to obfuscate patent disclosures. An instrumental variables analysis shows that patent obfuscation reduces follow-on citations and litigation risk, consistent with deterring competitive entry, but simultaneously lowers a patent's marketability and usefulness as collateral. These findings reveal a fundamental tradeoff between protecting proprietary information and preserving financing capacity. Evidence from the *Nautilus* decision, the AIPA disclosure reforms, and Moody's 2006 revision to its leverage-adjustment methodology points to the same conclusion.

Our findings highlight a channel through which financial considerations influence the clarity of information that enters the public domain. Because firms that rely more heavily on external financing face stronger incentives to produce transparent patent disclosures, external capital creates a positive informational externality by improving the quality of publicly disclosed technological information. Our results therefore identify a new mechanism through which financial markets shape the diffusion of technological knowledge.

More broadly, our LLM-based framework provides a general approach for separating funda-

mental complexity from strategic obfuscation in textual disclosures. By generating counterfactual text that preserves informational content while improving clarity, the method identifies the portion of complexity attributable to the underlying information and the portion arising from strategic disclosure choices. As large language models continue to improve, this framework can be applied well beyond patents to earnings announcements, SEC and other regulatory filings, loan contracts, scientific writing, and other settings where necessary complexity and strategic ambiguity coexist.

References

Acharya, V., Xu, Z., 2017. Financial dependence and innovation: The case of public versus private firms. Journal of financial economics 124, 223–243.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets 5, 31–56.

Armstrong, C., Glaeser, S., Park, S.Y., Timmermans, O., 2024. The assignment of intellectual property rights and innovation. Journal of Accounting Research .

Ashtor, J.H., 2022. Modeling patent clarity. Research Policy 51, 1–15.

Bai, J.J., Boyson, N.M., Cao, Y., Liu, M., Wan, C., 2025. Executives vs. chatbots: Unmasking insights through human-ai differences in earnings conference q&a. Northeastern U. D'Amore-McKim School of Business Research Paper .

Bekamiri, H., Hain, D.S., Jurowetzki, R., 2024. Patentsberta: A deep nlp based hybrid model for patent distance and classification using augmented sbert. Technological Forecasting and Social Change 206, 123536.

Bena, J., Ortiz-Molina, H., Simintzi, E., 2022. Shielding firm value: Employment protection and process innovation. Journal of Financial Economics 146, 637–664.

Bhattacharya, S., Ritter, J.R., 1983. Innovation and communication, signalling with partial disclosure. Review of Economic Studies 50, 331–346.

Botosan, C.A., 1997. Disclosure level and the cost of equity capital. The Accounting Review 72, 323–349.

Bowen III, D.E., 2016. Patent acquisition, investment, and contracting. Robert H. Smith School Research Paper No. RHS 2870112.

Bowen III, D.E., Frésard, L., Hoberg, G., 2023. Rapidly evolving technologies and startup exits. Management Science 69, 940–967.

Brown, J.R., Fazzari, S.M., Petersen, B.C., 2009. Financing innovation and growth: Cash flow, external equity, and the 1990s r&d boom. The journal of finance 64, 151–185.

Bushee, B.J., Gow, I.D., Taylor, D.J., 2018. Linguistic complexity in firm disclosures: Obfuscation or information? Journal of accounting research 56, 85–121.

Carley, M., Hegde, D., Marco, A.C., 2015. What is the probability of receiving a u.s. patent? Yale Journal of Law & Technology 17, 203–217.

Cockburn, I.M., Kortum, S., Stern, S., 2002. Are All Patent Examiners Equal? The Impact of Examiner Characteristics. Working Paper 8980. National Bureau of Economic Research.

Darrough, M.N., Stoughton, N.M., 1990. Financial disclosure policy in an entry game. Journal of Accounting and Economics 12, 219–243.

Diamond, D.W., Verrecchia, R.E., 1991. Disclosure, liquidity, and the cost of capital. Journal of Finance 46, 1325–1359.

Dyer, T., Glaeser, S., Lang, M.H., Sprecher, C., 2024. The effect of patent disclosure quality on innovation. Journal of Accounting and Economics 77, 1–28.

Faulkender, M., Petersen, M.A., 2006. Does the source of capital affect capital structure? The Review of Financial Studies 19, 45–79.

Figueroa, N., Lemus, J., 2024. The benefit of the doubt: Patent examination under strategic obfuscation. Management Science 70, 5853–5870.

Freilich, J., 2017. Patent clutter. Iowa L. Rev. 103, 925.

Gaulé, P., 2018. Patents and the success of venture-capital backed startups: Using examiner assignment to estimate causal effects. The Journal of Industrial Economics 66, 350–376.

Glaeser, S., 2018. The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. Journal of Accounting and Economics 66, 163–193.

Goldstein, I., Yang, L., 2017. Information disclosure in financial markets. Annual Review of Financial Economics 9, 101–125.

Gunning, R., 1969. The fog index after twenty years. Journal of Business Communication 6, 3–13.

Healy, P.M., Palepu, K.G., 2001. Information asymmetry, corporate disclosure, and the capital markets, a review of the empirical disclosure literature. Journal of Accounting and Economics 31, 405–440.

Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. Journal of political economy 124, 1423–1465.

Horstmann, I., MacDonald, G.M., Slivinski, A., 1985. Patents as information transfer mechanisms: To patent or (maybe) not to patent. Journal of Political Economy 93, 837–858.

Huang, A.H., Wang, H., Yang, Y., 2023. Finbert: A large language model for extracting information from financial text. Contemporary Accounting Research 40, 806–841.

Ireland, K., Schuster, W.M., Samples, T.R., 2024. Disclose by law, obfuscate by strategy: A corpus-based analysis of patent filings. BYU Law Review 50, 123–154.

Kisgen, D.J., 2019. The impact of credit ratings on corporate behavior: Evidence from moody's adjustments. Journal of Corporate Finance 58, 567–582.

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

Kong, N., Dulleck, U., Jaffe, A.B., Sun, S., Vajjala, S., 2023. Linguistic metrics for patent disclosure: Evidence from university versus corporate patents. Research policy 52, 104670.

Kraft, P., 2015. Rating agency adjustments to gaap financial statements and their effect on ratings and credit spreads. The Accounting Review 90, 641–674.

Lemley, M.A., Sampat, B., 2012. Examiner characteristics and patent office outcomes. The Review of Economics and Statistics 94, 817–827.

Leuz, C., Verrecchia, R.E., 2000. The economic consequences of increased disclosure. Journal of Accounting Research 38, 91–124.

Leuz, C., Wysocki, P.D., 2016. The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. Journal of accounting research 54, 525–622.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and economics 45, 221–247.

Loughran, T., McDonald, B., 2014. Measuring readability in financial disclosures. the Journal of Finance 69, 1643–1671.

Mackrael, K., 2026. The race to make the world's most in-demand machine. *The Wall Street Journal*. URL: https://www.wsj.com/business/the-race-to-make-the-worlds-most-in-demand-machine-092e8cea. accessed July 20, 2026.

Maksimovic, V., Pichler, P., 2001. Technological innovation and initial public offerings. Review of Financial Studies 14, 459–494.

Mann, W., 2018. Creditor rights and innovation: Evidence from patent collateral. Journal of Financial Economics 130, 25–47.

Melero, E., Palomeras, N., Wehrheim, D., 2020. The effect of patent protection on inventor mobility. Management Science 66, 5485–5504.

Mille, S., Wanner, L., 2008. Making text resources accessible to the reader: the case of patent claims, in: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Tapias, D. (Eds.), Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), European Language Resources Association (ELRA), Marrakech, Morocco.

Moody's, 2006. Moody's Approach to Global Standard Adjustments in the Analysis of Financial Statements for Non-Financial Corporations. Technical Report 96760. Moody's Investors Service.

Pairolero, N.A., de Grazia, C.A.W., 2026. Automated routing, patent examiner specialization, and quasi-random assignment at the uspto. Research Policy 55, 105460. doi:10.1016/j.respol.2026.105460.

Plumlee, M., Brown, D., Hayes, R.M., Marshall, R.S., 2015. Voluntary environmental disclosure quality and firm value: Further evidence. Journal of accounting and public policy 34, 336–361.

Saidi, F., Žaldokas, A., 2021. How does firms' innovation disclosure affect their banking relationships? Management Science 67, 742–768.

Sampat, B., Williams, H.L., 2019. How do patents affect follow-on innovation? evidence from the human genome. American Economic Review 109, 203–36.

Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear iv regression, in: Andrews, D.W.K., Stock, J.H. (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, New York, pp. 80–108.

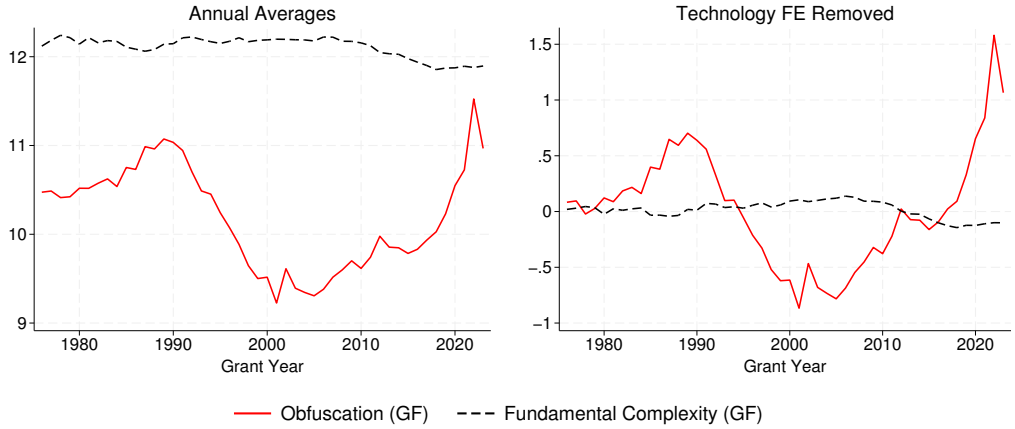
Verrecchia, R.E., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194.

Yang, S., 2023. Understanding innovation quality and success with large-language models. Available at SSRN 4482536 .

Figure 1: Time Series of Obfuscation

This figure plots the aggregate evolution of our measures of obfuscation from 1976-2023. Panel A shows the average level of our main measure of *Obfuscation* (solid red line) and the average complexity of LLM-simplified abstract (dotted line), with both variables using Gunning Fog to compute textual complexity. Panel B repeats this graph, using the first principal component of the readability measures listed in Table A.2 as the measure of patent text complexity. Left panels report raw annual averages and right panels show technology class-adjusted averages. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog



Panel B: Patent obfuscation measured via PCA1

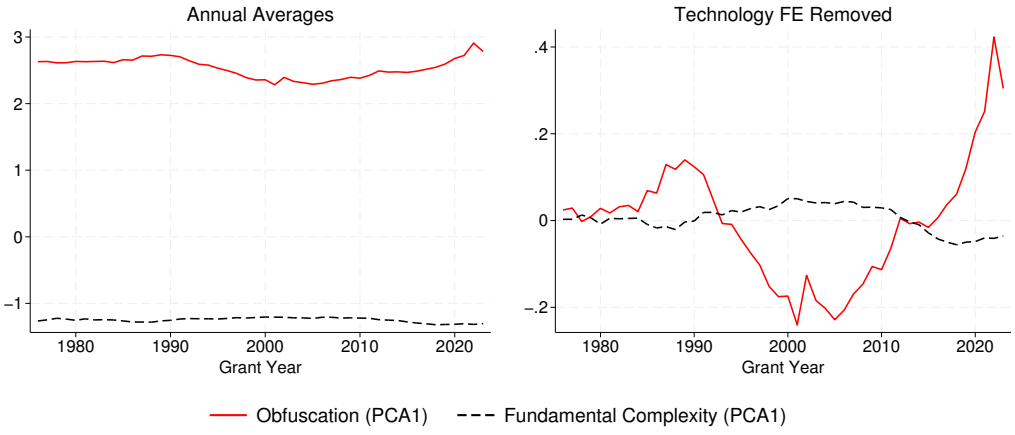


Figure 2: Obfuscation Around Firm's First Patent Pledge

This figure shows how patent complexity evolves around the year of a firm's first pledging event. Event-time coefficients (β_k) are estimated by

$$y_{p,i,t} = \sum_{k=-5}^{5} \Big(\beta_k \Phi(t - P(i) = k)\Big) + \phi_{c(p),t} + \gamma_i + u_{p,i,t}$$

where $\Phi(t-P(i)=k)$ is an indicator function for whether the patent p from application year t is k years after the year of the firm's first patent pledge $P(i)$. We omit $k=-1$, so all coefficients are relative to the year before the first pledge. Fixed effects are included for patent technology category-by-year fixed effects ($\phi_{c(p),t}$) and firm (γ_i). We estimate this equation for $y_{p,i,t} \in \{\text{Total Complexity}, \text{Fundamental Complexity}, \text{Obfuscation}\}$ measured via Gunning Fog. The 95% confidence interval is based on standard errors clustered by firm-year. (Back to text.)

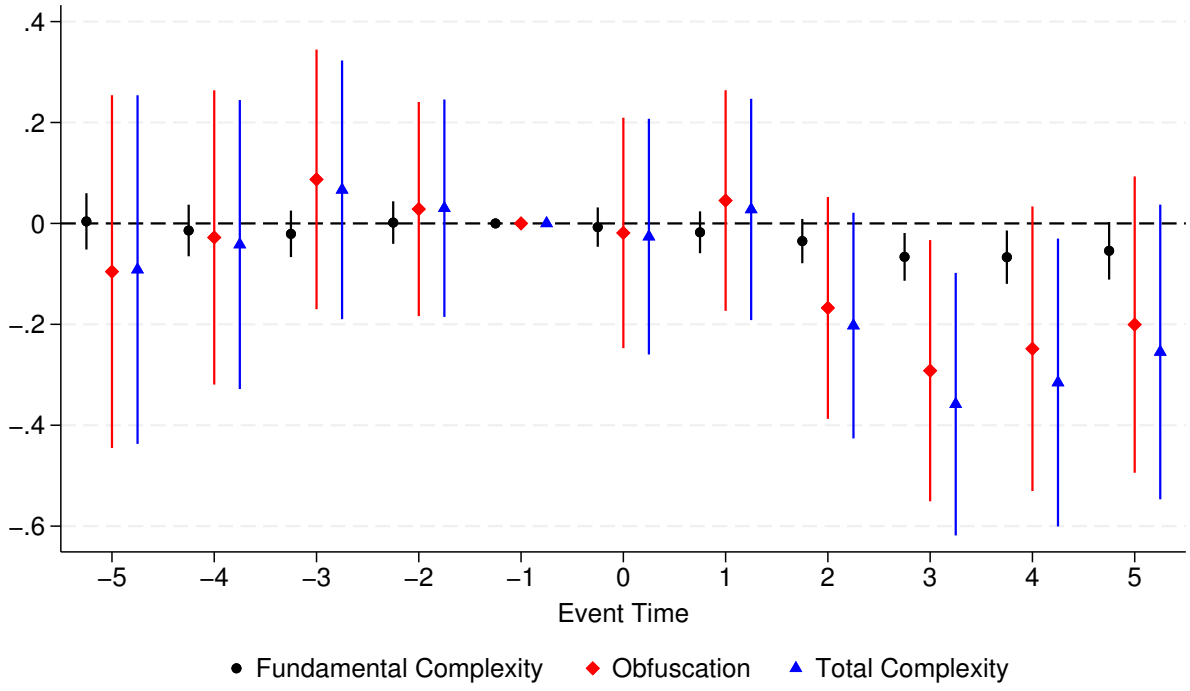


Table 1: Examples of Patents with Obscured Language

This table presents three patent abstracts, their LLM-simplified versions, and the corresponding measures of total complexity, fundamental complexity, and obfuscation. Patent 9,547,842 is IBM's "out of office" auto-reply patent. Patent 8,786,025 is TSMC's "resistor" patent. Patent 8,786,825 is ASML's "lithographic technique" patent. The complexity of a patent is measured via Gunning Fog (*GF*) or the first principal component of the readability measures listed in Table A.2 (*PCA1*). (Back to text.)

Patent	Abstract *Source for Total Complexity*	LLM-Simplified Version *Source for Fundamental Complexity*	Complexity Measures
9,547,842	A method, a system and a computer program product of transmitting availability information in an electronic mail system are disclosed. The system includes a first system including a sender, the first system is adapted to activate an availability indicator having a start date, wherein the activation is performed by the sender and extract availability indicator metadata from the availability indicator. The availability indicator metadata includes the start date, an end date, and at least one availability indicator message. If an e-mail has not been sent to a receiver since activation and if a current date is prior to the start date, the first system is further adapted to attach the availability indicator metadata to the e-mail, and send the e-mail to the receiver. The electronic mail system further includes a scheduling system functionally coupled to the first system, and is adapted to define the availability indicator.	This invention describes a way to share availability information through email. It involves a system where a sender can set up an availability indicator with start and end dates, along with a message. When the sender emails someone for the first time after setting this up, and before the start date, the system automatically adds the availability information to the email. The system also includes a scheduling feature that helps create these availability indicators.	*Total Complexity* GF: 19.71 PCA1: 0.62 *Fundamental Complexity* GF: 12.81 PCA1: -1.17 *Obfuscation* GF: 6.90 PCA1: 1.79
8,786,025	A system and method for forming a resistor system is provided. An embodiment comprises a resistor formed in a U-shape. The resistor may comprise multiple layers of conductive materials, with a dielectric layer filling the remainder of the U-shape. The resistor may be integrated with a dual metal gate manufacturing process or may be integrated with multiple types of resistors.	A new way to make resistors is described. It involves creating a U-shaped resistor using layers of conductive materials. The empty space in the U-shape is filled with a non-conductive material. This design can be used in various manufacturing processes, including those that make different types of resistors or use two metal gates.	*Total Complexity* GF: 13.29 PCA1: -1.21 *Fundamental Complexity* GF: 14.77 PCA1: -0.94 *Obfuscation* GF: -1.48 PCA1: -0.26
8,786,825	The present invention makes the use of measurement of a diffraction spectrum in or near an image plane in order to determine a property of an exposed substrate. In particular, the positive and negative first diffraction orders are separated or diverged, detected and their intensity measured to determine overlay (or other properties) of exposed layers on the substrate.	This invention uses diffraction patterns near an image plane to analyze properties of a substrate with exposed layers. It specifically measures the intensity of separated positive and negative first-order diffractions to determine how well the layers align or to assess other characteristics of the substrate.	*Total Complexity* GF: 20.56 PCA1: 0.63 *Fundamental Complexity* GF: 18.88 PCA1: 0.48 *Obfuscation* GF: 1.69 PCA1: 0.15

Table 2: Summary Statistics

This table presents summary statistics at the patent level for patents granted from 1976-2023 to publicly traded firms. *Obfuscation* is the difference between the complexity of the abstract in a given patent grant and the LLM-simplified version of that abstract. The complexity of a patent is measured via Gunning Fog (*GF*) or the first principal component of the readability measures listed in Table A.2 (*PCA1*). *Pledged*, *PatSold*, and *Infringed* are booleans equal to one if the patent is used as collateral, sold to another entity, or involved in litigation by year-end 2024. Other variables are defined in Section 3 and Table A.1. (Back to text.)

	N	Mean	SD	p10	p50	p90
Main variables						
Total Complexity (GF)	2,607,015	22.13	11.84	12.87	18.65	35.46
Fundamental Complexity (GF)	2,607,015	12.07	2.91	8.56	11.88	15.74
Obfuscation (GF)	2,607,015	10.06	11.96	0.86	6.57	23.53
Total Complexity (PCA1)	2,607,015	1.26	2.87	-1.05	0.46	4.49
Fundamental Complexity (PCA1)	2,607,015	-1.26	0.83	-2.28	-1.29	-0.20
Obfuscation (PCA1)	2,607,015	2.52	2.92	0.20	1.70	5.81
IV variables						
Examiner 112 Strictness	915,443	0.04	0.06	0.00	0.02	0.11
Examiner Other Strictness	915,443	0.54	0.19	0.26	0.57	0.78
Patent-level variables						
Illiquidity	1,805,160	0.50	0.29	0.10	0.49	0.90
RETech	2,586,165	1.48	2.13	-0.06	1.10	3.31
Log(1+Cite10)	2,607,015	0.52	0.95	0.00	0.00	1.95
Log(1+NonSelf10)	2,607,015	0.44	0.88	0.00	0.00	1.61
Log(1+Self10)	2,607,015	0.15	0.52	0.00	0.00	0.69
KPSS	2,607,015	12.54	34.05	0.07	3.91	29.30
Infringed	2,607,015	0.00	0.06	0.00	0.00	0.00
PatSold	2,607,015	0.26	0.44	0.00	0.00	1.00
Pledged	2,006,173	0.15	0.36	0.00	0.00	1.00
Firm-year variables						
TradeSecret	1,457,848	0.75	0.43	0.00	1.00	1.00
HHI (SIC3)	2,600,729	0.15	0.14	0.04	0.10	0.34
Log(Peer Patent Count)	2,355,338	1.97	1.74	0.00	1.95	4.43

Table 3: Pairwise Correlations between Complexity Measures and Economic Outcomes

This table reports correlations between our text-based complexity measures and proxies for information asymmetry, trade secrecy, strategic incentives, and patent impact. The sample is patents granted from 1976-2023 to publicly traded firms; each cell uses the subsample for which the row variable is available. All correlations control for patent application year. *Total Complexity (GF)* is the Gunning Fog score of the patent abstract. *Fundamental Complexity (GF)* is the Gunning Fog score of the LLM-simplified abstract. *Obfuscation (GF)* is the difference between total and fundamental complexity. *Abstract Length* is the number of words in the patent abstract. Other variables are defined in Section 3 and Table A.1. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. (Back to text.)

	Obfuscation (GF)	Fundamental Complexity (GF)	Total Complexity (GF)	Abstract Length	Log(Claims)
Illiquidity	0.024***	-0.021***	0.018***	-0.013***	-0.008***
TradeSecret	-0.020***	0.022***	-0.013***	0.023***	0.028***
HHI (SIC3)	0.068***	-0.039***	0.059***	0.008***	-0.048***
Log(Peer Patent Count)	-0.201***	0.058***	-0.189***	-0.010***	0.183***
RETech	-0.037***	0.084***	-0.017***	-0.018***	0.050***
Log(1+Cite10)	-0.050***	0.006***	-0.049***	0.018***	0.112***
Log(KPSS)	-0.219***	0.031***	-0.213***	-0.053***	0.179***

Table 4: Abstract Simplification During Patent Examination Process

This table examines the changes in patent abstract readability before and after the examination process. The sample is based on a subset of patent applications for which abstracts are observable both at filing and at grant. Across columns, the dependent variable is an indicator variable that equals 100 if the readability of the abstract improved during the examination process, and zero otherwise. *Examiner 112 Strictness* is the fraction of applications considered by an examiner that do not meet disclosure clarity requirements under §112 in the first review action but pass other requirements, excluding the focal application. Variables representing strictness for other sections are defined in a similar manner. *Fundamental Complexity (GF)* is the Gunning Fog score of the LLM-simplified abstract. When constructing strictness measures, we use examiners' raw rejection rates in columns 1-2 and unit-adjusted rejection rates in columns 3-4. In columns 1 and 3 (2 and 4), rejection rates are calculated based on all applications (only on eventual grants). To facilitate interpretation, all independent variables are standardized. Standard errors are clustered by firm-year. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. All models include technology-by-application year and art-unit-by-application year fixed effects. (Back to text.)

Rej Rates Are:	Examiner Only		Unit-Adjusted	
	(1)	(2)	(3)	(4)
Rej Rates For:	Apps.	Grants	Apps.	Grants
Examiner 112 Strictness (z)	0.647***	0.639***	0.470***	0.468***
	(0.06)	(0.06)	(0.04)	(0.04)
Examiner 101 Strictness (z)	0.764***	0.763***	0.359***	0.360***
	(0.07)	(0.07)	(0.03)	(0.03)
Examiner 102 Strictness (z)	0.202***	0.201***	0.175***	0.175***
	(0.04)	(0.04)	(0.04)	(0.04)
Examiner 103 Strictness (z)	0.157***	0.157***	0.152***	0.152***
	(0.04)	(0.04)	(0.04)	(0.04)
Fundamental Complexity (GF) (z)	0.082**	0.082**	0.082**	0.082**
	(0.04)	(0.04)	(0.04)	(0.04)
Fixed effects:				
CPC-year	Y	Y	Y	Y
Art-unit-year	Y	Y	Y	Y
Obs.	395,304	395,303	395,273	395,272
R^2	0.06	0.06	0.06	0.06

Table 5: Patent Obfuscation and Patent Examiner Strictness

This table reports the effects of patent-level obfuscation estimated using an IV framework. The sample is patents granted from 2008-2020 to publicly traded firms where examiner strictness can be measured. Column 1 shows the first-stage. Subsequent columns show second-stage IV regression results. *Ex. 112 Strictness* is the fraction of applications considered by an examiner that do not meet disclosure clarity requirements under §112 in the first review action but pass other requirements, excluding the focal application. *Ex. Other Strictness* is the fraction of applications considered by an examiner where the first action decision flags §101, §102, or §103, excluding the focal application. *Fund. Complexity* is the readability level (using the Gunning Fog index in Panel A and PCA in Panel B) of the LLM-simplified patent abstract. Citation variables are the natural logarithm of one plus the number of forward citations received within 10 years of the grant date. We separately report results using all cites, cites from the inventing firm (*Self*), and non-self citations. *Infringed*, *PatSold*, and *Pledged* are indicator variables equal to 100 if the patent is involved in litigation, sold to another entity, or used as collateral by year-end 2024, and zero otherwise, respectively, so their coefficients are in percentage points. Other variables are defined in Section 3 and Table A.1. To facilitate interpretation, all independent variables are standardized. Standard errors are clustered by firm-year. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. All models include technology-by-application year and art-unit-by-application year fixed effects. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog

	1st Stg	Log of 1 plus cites received from			Other outcomes		
	(1) Obf. (z)	(2) All	(3) Others	(4) Self	(5) Infringed	(6) PatSold	(7) Pledged
Ex. 112 Strictness (z)	-0.010***						
	(0.00)						
$\widehat{Obfuscation}(z)$		-2.162***	-1.273***	-1.515***	-0.046***	-0.381***	-0.202**
		(0.47)	(0.32)	(0.33)	(0.02)	(0.12)	(0.08)
Ex. Other Strictness (z)	-0.009***	-0.006	-0.005*	-0.002	-0.000	-0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Fund. Complexity (z)	-0.200***	-0.409***	-0.236***	-0.292***	-0.009***	-0.074***	-0.039**
	(0.00)	(0.09)	(0.06)	(0.07)	(0.00)	(0.02)	(0.02)
Fixed effects:							
CPC-year	Y	Y	Y	Y	Y	Y	Y
Art-unit-year	Y	Y	Y	Y	Y	Y	Y
Obs.	733,013	733,013	733,013	733,013	733,013	733,013	733,013
Cragg-Donald F	27.71						

Panel B: Patent obfuscation measured via PCA1

	1st Stg	Log of 1 plus cites received from			Other outcomes		
	(1) Obf. (z)	(2) All	(3) Others	(4) Self	(5) Infringed	(6) PatSold	(7) Pledged
Ex. 112 Strictness (z)	-0.009***						
	(0.00)						
$\widehat{Obfuscation}(z)$		-2.266***	-1.334***	-1.588***	-0.048***	-0.400***	-0.212**
		(0.51)	(0.34)	(0.35)	(0.02)	(0.12)	(0.09)
Ex. Other Strictness (z)	-0.009***	-0.006	-0.005*	-0.002	-0.000	-0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Fund. Complexity (z)	-0.235***	-0.508***	-0.295***	-0.361***	-0.011***	-0.091***	-0.048**
	(0.00)	(0.12)	(0.08)	(0.08)	(0.00)	(0.03)	(0.02)
Fixed effects:							
CPC-year	Y	Y	Y	Y	Y	Y	Y
Art-unit-year	Y	Y	Y	Y	Y	Y	Y
Obs.	733,013	733,013	733,013	733,013	733,013	733,013	733,013
Cragg-Donald F	26.01						

Table 6: Nautilus v. Biosig and Patent Obfuscation

This table examines the impact of *Nautilus v. Biosig* (2014) on patent obfuscation. The sample is based on patent applications by public firms three years before and after the decision date (6/2/2014). In Panel A, the dependent variable is *Obfuscation (Gunning Fog)*. In Panel B, the dependent variable is *Obfuscation (PCA1)*. *Exposure* is the median obfuscation level for a given firm across patent applications filed during the pre-period. *Post* is an indicator for patent applications after the ruling. *Pledged* is an indicator for firms that pledged patents as collateral during 2011-2013. *HighLev* is an indicator for firms with above-median leverage, where medians are defined at the Fama-French 12-industry level based on firms' average leverage during 2011-2013. The variable used as the triple interaction term *Z* for a given column is listed in the header of each panel. To be included in the sample, firms should have filed patents in both the pre- and post-periods. Standard errors are clustered by the intersection of firm and grant year. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog

	(1)	(2)	(3)	(4)	(5)	(6)
Z variable:	None	Pledged	HighLev	None	Pledged	HighLev
Exposure	1.214***	1.072***	1.242***			
	(0.03)	(0.06)	(0.03)			
Post	1.478***	3.668***	2.109***	0.420**	1.488***	0.830***
	(0.33)	(0.72)	(0.46)	(0.21)	(0.45)	(0.19)
Exposure x Post	-0.158***	-0.512***	-0.271***	-0.071**	-0.254***	-0.138***
	(0.05)	(0.12)	(0.08)	(0.03)	(0.07)	(0.03)
Z		-0.226	0.967***			-0.509
		(0.17)	(0.36)			(0.41)
Z x Exposure		0.061**	-0.081			
		(0.03)	(0.05)			
Z x Post		-1.016***	-1.767***		-0.455**	-1.105***
		(0.28)	(0.62)		(0.18)	(0.36)
Z x Exposure x Post		0.163***	0.292***		0.078***	0.174***
		(0.05)	(0.10)		(0.03)	(0.06)
Fixed effects:						
Year	N	N	N	Y	Y	Y
Firm	N	N	N	Y	Y	Y
Obs.	503,093	503,093	503,086	503,093	503,093	503,086
R^2	0.12	0.12	0.12	0.14	0.14	0.14

Panel B: Patent obfuscation measured via PCA1

	(1)	(2)	(3)	(4)	(5)	(6)
Z variable:	None	Pledged	HighLev	None	Pledged	HighLev
Exposure	1.211***	1.071***	1.243***			
	(0.03)	(0.06)	(0.03)			
Post	0.415***	0.992***	0.579***	0.126**	0.411***	0.229***
	(0.09)	(0.19)	(0.12)	(0.05)	(0.12)	(0.05)
Exposure x Post	-0.171***	-0.529***	-0.280***	-0.079***	-0.266***	-0.141***
	(0.05)	(0.12)	(0.08)	(0.03)	(0.07)	(0.03)
Z		-0.064	0.254***			-0.122
		(0.05)	(0.09)			(0.10)
Z x Exposure		0.061**	-0.091*			
		(0.03)	(0.05)			
Z x Post		-0.269***	-0.468***		-0.121***	-0.282***
		(0.07)	(0.16)		(0.05)	(0.10)
Z x Exposure x Post		0.166***	0.289***		0.080***	0.164***
		(0.04)	(0.10)		(0.03)	(0.06)
Fixed effects:						
Year	N	N	N	Y	Y	Y
Firm	N	N	N	Y	Y	Y
Obs.	503,093	503,093	503,086	503,093	503,093	503,086
R^2	0.11	0.11	0.11	0.13	0.13	0.13

Table 7: AIPA and Patent Obfuscation

This table examines how the enactment of the American Inventors Protection Act (AIPA) affected patent complexity, and how that effect was mediated by reliance on patent collateral and external financing. The identification strategy is a difference-in-differences approach following Ireland et al. (2024); see text for discussion. The regressions study all patents with application years from 1990 to 2010 that can be matched with public firm identifiers, excluding year 2000. In Panel A, the dependent variable is *Obfuscation (Gunning Fog)*. In Panel B, the dependent variable is *Obfuscation (PCA1)*. *Post* is an indicator for the patent application year being later than 2000. *Exposure* is the percent of a filing firm's patents that sought foreign protection as of 1996 (the year before the introduction of HR 400, the precursor to AIPA). *Pledged* is an indicator for whether the firm ever pledged patents as collateral in the pre-AIPA era. *HighLev* is an indicator for firms with above-median leverage, where the median is measured pre-AIPA. The variable used as the triple interaction term Z for a given column is listed in the header of each panel. Standard errors are clustered by the intersection of firm and grant year. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog

	(1)	(2)	(3)	(4)	(5)	(6)
Z variable:	None	Pledged	HighLev	None	Pledged	HighLev
Exposure	5.312***	1.220	4.260***			
	(0.381)	(0.825)	(0.483)			
Post	-0.541*	-1.361***	-0.801**			
	(0.311)	(0.439)	(0.374)			
Exposure x Post	1.266**	4.200***	2.904***	1.192***	0.755	1.621***
	(0.543)	(1.168)	(0.700)	(0.238)	(0.579)	(0.281)
Z		-0.856*	0.479			-0.607*
		(0.449)	(0.481)			(0.313)
Z x Exposure		4.497***	1.753***			0.750*
		(0.919)	(0.648)			(0.416)
Z x Post		0.861	1.045*		0.529	0.740**
		(0.553)	(0.597)		(0.386)	(0.325)
Z x Exposure x Post		-2.930**	-3.842***		0.478	-1.111**
		(1.308)	(0.963)		(0.629)	(0.441)
Fixed effects:						
Year	N	N	N	Y	Y	Y
Firm	N	N	N	Y	Y	Y
Obs.	982,497	982,497	976,986	982,496	982,496	976,985
R^2	0.02	0.02	0.02	0.10	0.10	0.10

Panel B: Patent obfuscation measured via PCA1

	(1)	(2)	(3)	(4)	(5)	(6)
Z variable:	None	Pledged	HighLev	None	Pledged	HighLev
Exposure	1.297***	0.0606	1.004***			
	(0.0918)	(0.153)	(0.115)			
Post	-0.102	-0.198*	-0.194**			
	(0.0771)	(0.102)	(0.0926)			
Exposure x Post	0.283**	1.137***	0.700***	0.300***	0.521***	0.403***
	(0.130)	(0.239)	(0.169)	(0.0624)	(0.138)	(0.0712)
Z		-0.194*	0.173			-0.135*
		(0.106)	(0.118)			(0.0780)
Z x Exposure		1.365***	0.326**			0.166
		(0.182)	(0.155)			(0.104)
Z x Post		0.106	0.256*		-0.0553	0.180**
		(0.132)	(0.145)		(0.0941)	(0.0815)
Z x Exposure x Post		-0.864***	-0.913***		-0.242	-0.279**
		(0.279)	(0.228)		(0.152)	(0.110)
Fixed effects:						
Year	N	N	N	Y	Y	Y
Firm	N	N	N	Y	Y	Y
Obs.	992,264	992,264	976,986	992,264	992,264	976,985
R^2	0.02	0.02	0.02	0.10	0.10	0.09

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Financing Constraints and Patent Obfuscation: Moody's Change in Adjustment Methodology in 2006

This table examines how a firm's financing constraints affect patent obfuscation using Moody's 2006 revision of its adjusted leverage methodology as a plausibly exogenous shock. The sample includes patents filed by firms that were rated by Moody's and covered in the Moody's Financial Metrics Database from 2005 to 2008. The identification strategy is a difference-in-differences approach. *Post* is an indicator variable equal to one for patents that were filed after 2006. *Underfunded* is an indicator variable equal to one for firms whose underfunded pension-to-total-assets ratio exceeds the pre-period median. Standard errors are clustered by the intersection of firm and grant year. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. All models include firm and technology-by-application year fixed effects. (Back to text.)

Complexity Measure:	Gunning Fog		PCA1	
	(1) Obfuscation	(2) Fundamental Complexity	(3) Obfuscation	(4) Fundamental Complexity
Underfunded × Post	0.584**	-0.043	0.148**	-0.021*
	(0.25)	(0.04)	(0.06)	(0.01)
Fixed effects:				
CPC-year	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Obs.	151,618	151,618	151,618	151,618
R^2	0.05	0.08	0.05	0.11

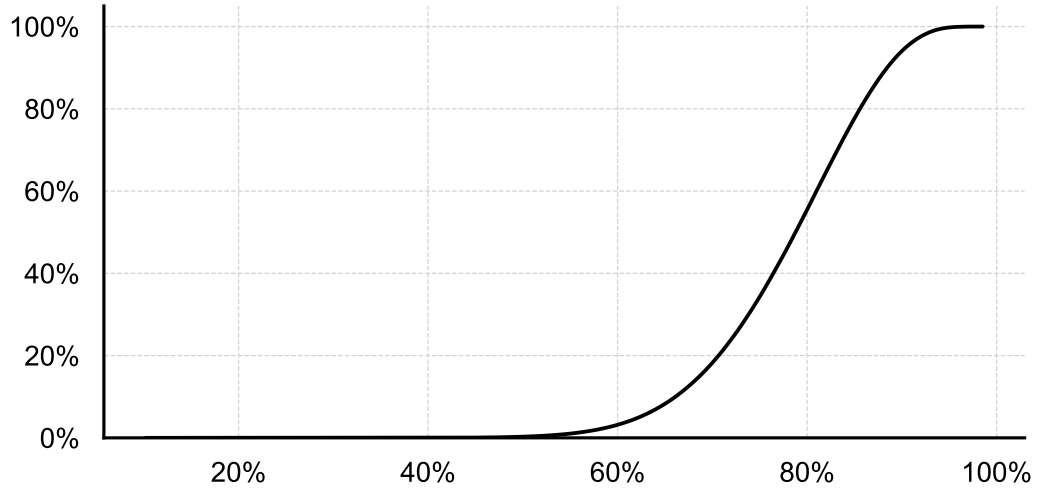
Appendix

- Figure A.1 shows the semantic similarity of patent abstracts and the LLM written versions.
- Figure A.2 examines how *Obfuscation* evolves over time across the six largest technology sections in the CPC classification scheme from 1976 to 2023.
- Figure A.3 examines how *Fundamental Complexity* evolves over time across the six largest technology sections in the CPC classification scheme from 1976 to 2023.
- Table A.1 defines variables at the patent level.
- Table A.2 shows the loadings from a principal components analysis of readability scores of patent abstracts. The model is estimated on the real and LLM-summarized patents simultaneously.
- Table A.3 shows the correlation between Obfuscation measures computed using six alternative LLMs.

Figure A.1: Distribution of the Cosine Similarity between Granted and LLM-Rewritten Abstracts

This figure compares the semantic similarity of patent abstracts to the rewritten versions we obtain from Claude Sonnet 3.5. For each abstract and its rewritten version, we obtain embeddings using the *PatentSBERTa* embedding model from Bekamiri et al. (2024). From these embeddings, we compute cosine similarities for each patent-rewrite pair. We show the empirical CDF of the cosine similarities in Panel A and the kernel density estimate (KDE) in Panel B.

Panel A: Empirical CDF of Cosine Similarity Distribution



Panel B: KDE of Cosine Similarity Distribution

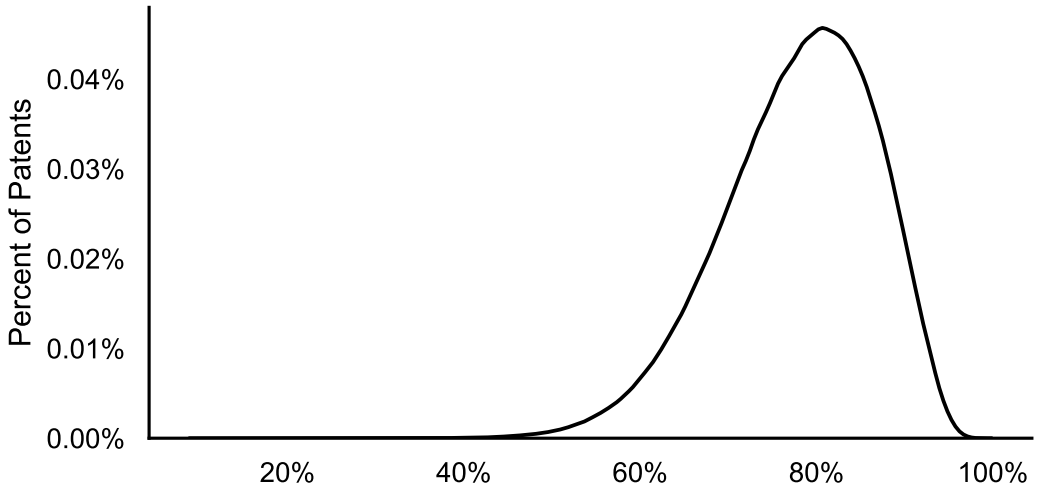


Figure A.2: Time Series of Obfuscation, by Technology Category
This figure plots the evolution of *Obfuscation (GF)* across the six largest technology sections in the CPC classification scheme from 1976 to 2023. Each subpanel reports in parentheses the fraction of patents over the entire sample belonging to that category.

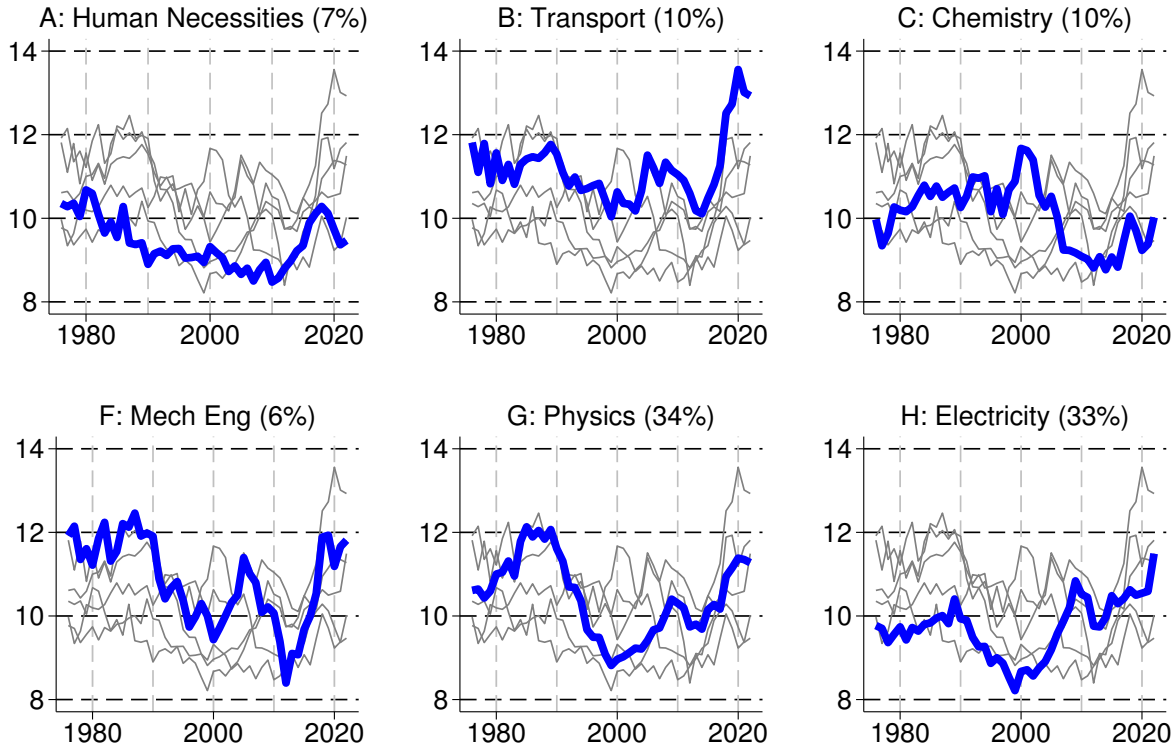


Figure A.3: Time Series of *Fundamental Complexity*, by Technology Category
This figure plots the evolution of *Fundamental Complexity*, measured by applying Gunning Fog to LLM rewritten patent abstracts. The figure separately reports averages across the six largest technology sections in the CPC classification scheme from 1976 to 2023. Each subpanel reports in parentheses the fraction of patents over the entire sample belonging to that category.

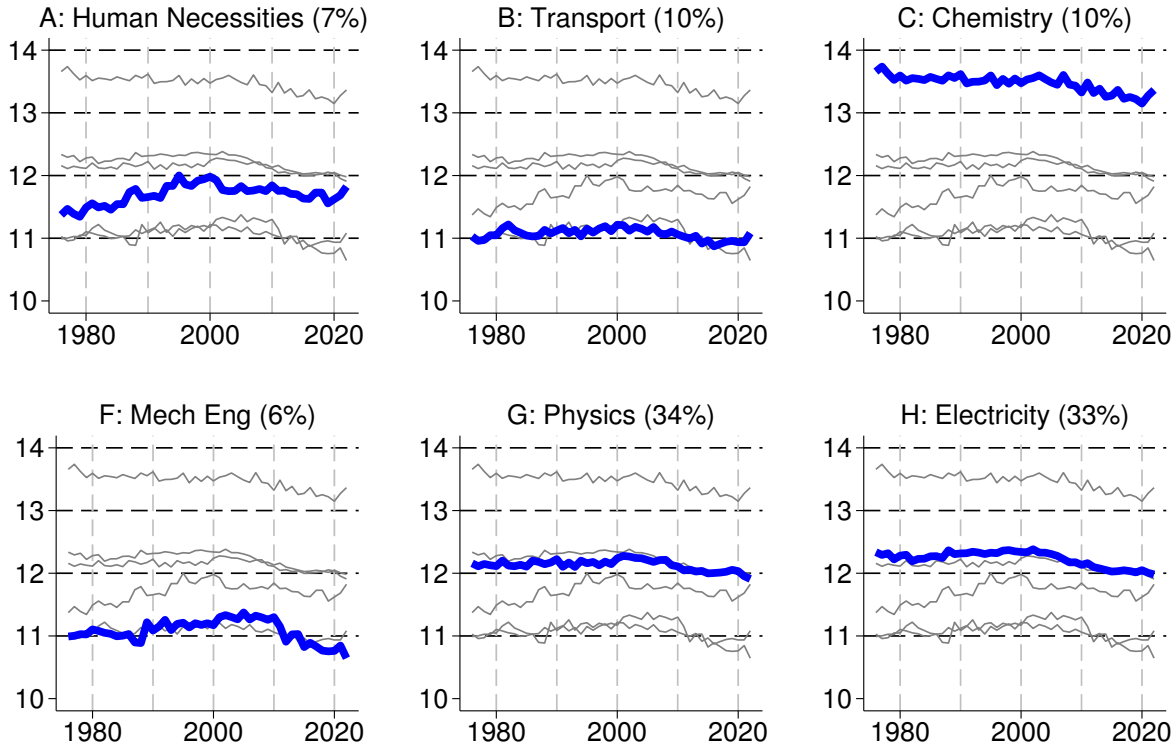


Table A.1: Variable Definitions

Data sources are described in Section 3.

Variable	Definition
	Main variables
Obfuscation (GF)	*Total Complexity (GF)* minus *Fundamental Complexity (GF)*
Fundamental Complexity (GF)	Gunning Fog grade level of the LLM-rewritten patent abstract
Total Complexity (GF)	Gunning Fog grade level of a patent abstract
Obfuscation (PCA1)	*Total Complexity (PCA1)* minus *Fundamental Complexity (PCA1)*
Fundamental Complexity (PCA1)	PCA1 complexity measure for the rewritten abstract
Total Complexity (PCA1)	PCA1 complexity measure for the patent abstract
	IV variables
Examiner 112 Strictness	For a given patent, the fraction of other applications considered by the examiner whose first-action rejection cites §112 (disclosure clarity) and no other rejection grounds (§101, §102, or §103)
Examiner Other Strictness	For a given patent, the fraction of other applications handled by the examiner whose first-action rejection cites §101, §102, or §103 and does not cite §112
	Patent-level variables
Illiquidity	Follows Bushee et al. (2018), adapted for patent disclosures. For each trading day $t \in [T+0, T+25]$ around patent disclosures (where T is grant date pre-AIPA, min(grant date, filing date + 18months) post-AIPA), compute $\lvert r_t \rvert / DailyVolInMill_t$, and winsorize this daily ratio. Then average this ratio for each event, and convert to percentile (0-100) ranks.
RETech	Exposure to rapidly evolving technology areas (i.e., following breakthroughs) (Bowen III et al., 2023)
Cite10	Number of citations within 10 years of grant
NonSelf10	Citations within 10 years from non-assignee firms
Self10	Citations within 10 years from the focal assignee
KPSS	Inflation adjusted patent value estimate from Kogan et al. (2017)
Infringed	Indicator: patent involved in litigation by 2024
PatSold	Indicator: patent sold to another entity by 2024
Pledged	Indicator: patent used as collateral by 2024
	Firm-year variables
TradeSecret	Indicator: the firm's 10-K in the patent application year discusses trade secrets. Data from Glaeser (2018) available at https://www.stephenglaeser.com/data
HHI (SIC3)	Industry sales concentration measured within SIC 3 digit industries
Log(Peer Patent Count)	Log of one plus the count of patents of product market competitors. Competitors are defined by TNIC3 industries. This variable is measured using the *grant* date of competitor patents to reflect public disclosure.

Table A.2: Principal Components Analysis of Readability Factors

This table shows the loadings on the first three principal components of the readability scores. The loadings are estimated using the readability scores measured on both the granted abstracts and the LLM rewritten abstracts. The sample of patents is those granted from 1976 to 2023 that are linked to public firm identifiers by Kogan et al. (2017).

	Comp1	Comp2	Comp3
Component Loadings			
Flesch Reading Ease	-0.40	0.12	-0.07
Flesch Kincaid Grade	0.40	0.07	-0.13
Gunning Fog	0.40	0.06	-0.12
Automated Readability Index	0.40	0.05	-0.13
Coleman Liau Index	0.16	-0.68	0.68
LIX	0.40	-0.02	-0.06
RIX	0.40	0.02	-0.10
Entropy	0.12	0.72	0.69
Principal Components			
Eigenvalues	6.04	1.20	0.63
CumulExplained	0.76	0.91	0.98

Table A.3: Correlation of Obfuscation (GF) From Different LLMs

This table shows the correlations between versions of *Obfuscation (GF)* created by using various leading LLMs. The sample consists of 1,000 random patents. Claude Sonnet 3.5 is the baseline model for the paper. We obtained simplifications for other models in February 2026 via the *Expected Parrot* platform.

Model	Legend	(1)	(2)	(3)	(4)	(5)	(6)	(7)
anthropic/claude-3-5	(1)							
anthropic/claude-3-7-sonnet-latest	(2)	0.97						
anthropic/claude-4-opus	(3)	0.96	0.94					
google/gemini-2.5-flash-lite	(4)	0.94	0.95	0.91				
meta/Llama-3.3-70B-Instruct-Turbo	(5)	0.92	0.93	0.90	0.91			
openai/gpt-4o	(6)	0.98	0.98	0.94	0.95	0.93		
xai/grok-4-1-fast-non-reasoning	(7)	0.96	0.97	0.93	0.94	0.92	0.97	