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: May 31, 2026

We use LLMs to decompose patent disclosure complexity into fundamental and strategic obfuscation components. Patent obfuscation has increased sharply across technology sectors over the past 15 years. Obscure disclosure deters competitor innovation and reduces infringement litigation but decreases patents' financing capacity. Using examiner strictness as an instrument, we show obfuscation reduces both competitor innovation and the probability of litigation while lowering the rate at which patents are sold or pledged as collateral. Following policy or financing shocks, firms' obfuscation responses depend on their reliance on external finance. Our findings reveal a tradeoff: firms balance opaque disclosure in patents against their need to attract external capital, a new channel through which financing can shape the trajectory of innovation.

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, Tengfei Zhang and seminar participants at the University of Colorado and the 2025 Philly Five Conference. Refine.ink was used to check the paper for consistency and clarity. All errors and omissions are our own. First version: November 11, 2025.

1 Introduction

According to the U.S. Patent and Trademark Office, patents generate $8 trillion of economic activity each year,[1] giving firms an incentive to protect the value of their patents. One way to do so is to deliberately obfuscate the language in patent filings. By making it difficult to judge the exact scope of a patent, firms create uncertainty for competitors about the probability of infringement, thereby deterring entry into related markets and follow-on innovation (Bhattacharya, 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 (Mann, 2018). If the patent's contribution or innovation is unclear to a company's competitors, it will be even more obscure to outside investors and lenders, who are likely to have less subject matter expertise. Thus, a firm that relies on external financing faces a tradeoff: obscure the patent's contribution to insulate it from competitors and patent trolls, or make its potential value more transparent to facilitate external financing.[2] In this paper, we examine whether firms' reliance on external financing mitigates their incentives to strategically obfuscate patent disclosures.

Our first 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 the patent using simpler language.

We propose a novel LLM-based approach to separate the fundamental and strategic components of patent complexity. Our methodology evaluates each patent by comparing the

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

[2] A large literature shows that increased disclosure transparency can reduce a firm's cost of capital but reveal costly information to competitors. See Healy and Palepu (2001a), Leuz and Wysocki (2016), and Goldstein and Yang (2017) for reviews.

complexity of its original abstract to an AI-generated version that is explicitly instructed to state the text in the simplest possible manner. We first instruct a leading LLM, Claude Sonnet 3.5, to provide the clearest possible restatement of the patent's abstract.[3] We then measure the linguistic complexity of both the original abstract (*Total Complexity*) and the LLM-simplified abstract (*Fundamental Complexity*) using standard linguistic tools.[4] We interpret the difference between the two as a proxy for strategic obfuscation (*Obfuscation*) in the writing of the patent. Our methodology is novel, but has strong parallels with prior work on disclosure and obfuscation, for example in conference calls (e.g. Bushee et al., 2018a; 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 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 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. Third, we also 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.[5]

We next propose a novel counterpoint to the existing literature on patent complexity and obfuscation: While the disclosure literature has primarily highlighted strategic *benefits* of obfuscation, we argue that there is also a potential *cost* through reduced financing capacity. To empirically explore this tradeoff, we examine how firms change their disclosure behavior

[3]This model is now deprecated and outdated on frontier research tasks but we show that its outputs on our task is highly similar to more recent generations of models.

[4]We focus on the abstract because it serves as the first line of defense for searches using algorithmic textual analysis to identify relevant technologies and for computational efficiency. We discuss this choice in more detail below.

[5]Importantly, these patterns are not present for fundamental complexity or abstract length.

around the time they first pledge one 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. This result is consistent with the hypothesis that firms reduce obfuscation to enhance the collateral value of their patent portfolio. The rest of our analysis focuses on establishing and studying this novel linkage between the financing capacity of patents and the tradeoffs inherent in patent disclosure.

Our empirical strategy links a firm's financing needs to its patent obfuscation behavior by exploiting two types of shocks: (i) patent-level shocks that affect a firm's ability or incentives to obfuscate, and (ii) financing-constraint shocks that influence its access to external capital. Across both settings, we find that firms more reliant on external financing are less likely to engage in strategic obfuscation, suggesting that dependence on external capital is associated with more transparent disclosure.

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), as detailed in Section 4.A. We document that examiner strictness has a meaningful impact on the change in the patent language from application to grant date. Using this approach, we find that obfuscated patents receive fewer citations from outside firms and are less likely to be involved in infringement litigation. Both findings are consistent with the strategic deterrent effect of obfuscation described in prior work (Freilich, 2017; Ashtor, 2022). However, obfuscated patents are significantly less likely to be sold or transferred after being granted and are less likely to be pledged as collateral for loans, thus reducing their value in financing.

The second patent-level shock to patent obfuscation is the Supreme Court's decision in *Nautilus v. Biosig* (2014), which heightened the clarity requirements for patent disclosures under §112 of Title 35 of the U.S. Code. We find that the firms most likely to be affected by the ruling, those that engaged in greater obfuscation prior to *Nautilus*, significantly reduce their obfuscation afterward, consistent with the decision's intent to limit firms' ability to obscure their patents. However, this effect is substantially smaller for firms that have previously pledged patents as collateral or rely heavily on leverage, and these firms already employed clearer disclosures prior to the ruling. These patterns support the hypothesis that firms reliant on external financing avoid obfuscating their patents, in order to enhance the value of their intellectual property to capital providers.

The third patent-level shock we examine is the enactment of the American Inventors Protection Act (AIPA) in 1999 and its effects on patent disclosure. 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. This suggests that financing needs can mitigate the incentives to obfuscate, illustrating how financial constraints can incentivize transparency and facilitate the use of intellectual property as collateral.

Our paper makes three important contributions to the literature. First, we introduce an LLM-based approach to measuring obfuscation 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., 2018b). 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 contribute 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, 1983; Bhattacharya and Ritter, 1995; 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). As a result, firms face an inherent tradeoff in their disclosure choices, balancing the proprietary costs of revealing sensitive information against 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, 2001b). A benefit of our approach is that we measure disclosure choices at the asset level within firms, rather than relying on firm-level disclosure proxies that may conflate heterogeneous disclosure incentives across assets with different proprietary value and financing relevance. This granularity allows us to isolate within-firm variation in disclosure behavior and directly test how firms tailor transparency to the specific proprietary and financing considerations associated with individual assets.

Third, we extend 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. We contribute to this literature by showing that financing considerations can influence the strategic choices made during the patent filing process. Because obfuscation reduces knowledge spillovers, our results identify an additional channel through which financing can shape the trajectory of innovation (Brown et al., 2009; Bernstein, 2015; Acharya and Xu, 2017).

2 Data

2.A Measuring Patent Obfuscation

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 not subject to obfuscation. It is useful to motivate our approach by comparing it with Bushee et al. (2018a), who focus on the possibility of obfuscation in quarterly earnings calls. 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. The authors use the Gunning Fog index to measure the complexity of both the management discussion and the analyst's summary, and the difference between the two serves as their measure of obfuscation.

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 and summary are the logical starting point for parties who need to understand a patent's contribution. Second, 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 in this section.[7] Third,

[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

the abstract and summary narrative allows the patent writer more room for discretion in comparison to the claims section. By contrast, the claims section defines 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). Fourth, limiting our analysis to the abstract substantially reduces its token usage and thus its 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 in the KPSS+CCM link tables.[8] We then remove patents granted to firms in utility and financial industries (SIC 49 and 60-69).

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:

> *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.[9]

Table I displays three examples from this process. The left column of the table presents the abstracts of the patents. The first two patents are often cited as potential examples of obfuscation. The first is Patent No. 8,786,366, Intel Mobile Communications' patent for an amplifier circuit. The second is Patent No. 9,547,842, IBM's patent for an email autoreply feature. Both inventions appear relatively straightforward to describe, yet the abstracts in both cases are surprisingly complex, containing unnecessary information and ambiguous

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.

[9]In ongoing work, we check robustness of our results to more token-intensive approaches applied to smaller subsamples of patents, including longer prompts and examining the patent summary in addition to the abstract.

wording.[10] However, as can be seen in the second column of Table I, the LLM is able to explain the core idea of each patent much more clearly than the original.

Moreover, we find strong evidence that the LLM succeeds at preserving the semantic meaning of the patent abstracts. In Figure A1 of the Appendix, we validate the information-preserving nature of these summaries by computing the cosine similarity of the embeddings of the original and rewritten abstracts. More than 80% of patents have a similarity above 70% and virtually all rewrites are above 55%, indicating that the rewritten abstracts contain the same content. This is consistent with recent literature that finds strong evidence that the fidelity of the summarizations produced by LLMs is high (e.g., Huang et al., 2023).

It is worth noting that the LLM does not always find opportunities to simplify a patent. As an example, the last patent listed in the table is Patent No. 8,786,025, a new method of making resistors that was granted to TSMC. In contrast to the prior two abstracts, the original and LLM texts are very similarly structured and the core idea is clear in each.

[Insert Table I about here]

Next, 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

[10] Indeed, 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 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 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., 2023).[11] As can be seen in Appendix Table A2, 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.

The table below shows the Gunning Fog Index and PCA values for the examples in Table I as well as the calculation for *Obfuscation*:

Patent Number	Complexity Measure	Total Complexity	Fundamental Complexity	Obfuscation
8,786,366	Gunning Fog	94.17	11.97	82.20
	PCA	17.86	-1.39	19.25
9,547,842	Gunning Fog	19.71	12.81	6.90
	PCA	0.62	-1.17	1.79
8,786,025	Gunning Fog	13.29	14.77	-1.48
	PCA	-1.21	-0.94	-0.26

As shown in the table above, our obfuscation measure captures economically meaningful reductions in textual complexity. For Patent No. 8,786,366, the Gunning Fog index declines from 94.17 to 11.97. For Patent No. 9,547,842, the readability of the original abstract is 19.71, consistent with post-graduate level text, while the AI-generated abstract has a Gun-

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

ning Fog score of 12.81, corresponding to a high-school graduate reading level. Importantly, the results also validate that the third patent listed in the table, Patent No. 8,786,025, was not obfuscated: The original is written in a direct manner accessible to undergraduate students, leaving relatively limited scope for simplification. In fact, the LLM revision exhibits slightly *more* complexity.

As with any application of LLMs to historical data, one may be concerned about the possibility of lookahead 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, lookahead 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 lookahead bias), which we do not observe in our results below. Third, the degree of this bias will vary over time and should be absorbed by the time-varying fixed effects included in all of our empirical designs, which focus on cross-sectional variation.

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 A3 shows the correlations between obfuscation estimated with Claude Sonnet 3.5 and six additional LLMs from 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.

2.B Other Data Sources

We link our patent sample with data on citations within 10 years of the grant date and CPC technology classes (PatentsView), patent sales and use as collateral (USPTO Assignments database), infringement allegations (USPTO and the Stanford NPE Litigation Database), art-unit assignments (PatEx), competitor patenting, and firm-level financial data (Compustat). Firm-level data are merged with patents using the application year to measure the firm's condition at the time of the patent's invention. 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.A.2.

Formal variable definitions are listed in Table A1 in the Appendix. Outcome variables are as follows. Patent value estimates come from Kogan et al. (2017) (henceforth, 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.

Additional tests and stylized facts examine variables that capture dimensions of competition. *HHI (SIC3)* and *HHI (TNIC3)* are industry sales concentration measured within SIC 3 digit industries and TNIC3 industries from Hoberg and Phillips (2016), respectively. *Market Share* is based on the market share of sales within SIC3 industries. *Peer Patent Count* is the number of patents granted to TNIC3 industry competitors. *Peer Patent Value* is the total KPSS value of patents granted to competitors in the same TNIC3 industry.

2.C Summary Statistics

Table II presents the summary statistics of our patent-level dataset. Our full sample consists of 2,607,015 patents spanning grants from 1976 to 2023. We exclude firms in SIC codes 49 and 60-69 (industrials and finance).

[Insert Table II about here]

The average patent abstract has a Gunning Fog Index (GF) score of 22.13, requiring substantial post-graduate education to understand. LLM-simplified abstracts score 12.07, accessible to high school graduates. The difference of 10.06 represents our measure of *Obfuscation (GF)*, suggesting substantial potential for strategic disclosure in typical patents.

Examiner strictness data are available only after 2008. The average examiner issues a first-action rejection citing only §112 for 4% of patents. The average patent receives 3.8 citations (median is zero) and has a KPSS value of approximately $13 million (in 1982 dollars). In our sample, 15% of patents are pledged as collateral and 26% are sold. Patent infringement is somewhat rare, with slightly over 11,000 patents litigated, though these are disproportionately valuable. On average, patents receive 3.84 citations over a ten-year period (and the average of the log of citation count is 0.52).

3 Stylized Facts about Patent Obfuscation

We next document several stylized facts about our measure of patent obfuscation that provide an empirical foundation for understanding its behavior across time, industries, and firms. Importantly, we validate that obfuscation rises when strategic benefits of poorer disclosure are higher. We then report new evidence on how obfuscation changes after firms begin using patents as lending collateral.

Fact 1: *Obfuscation* fell during the 1990s but has steadily risen since.

[Insert Figure I about here]

Figure I 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.

The solid red lines show that obfuscation exhibits substantial time variation across the nearly five decades of data. 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.[13]

Fact 2: *Fundamental Complexity* has less time series variation. Thus, *Total Complexity* is largely driven by the strategic component of patent language.

The dashed black line in Figure I displays the average fundamental complexity of patent abstracts, measured by the readability of the LLM-simplified abstracts. In contrast to the substantial time-series variation in obfuscation, this fundamental component of patent complexity remains 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 does not imply that all LLM-simplified abstracts are rewritten to the same readability level. Figure A2 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

[12] Figure I and Figure II are presented based on grant year. The patterns are very similar by application year. All remaining tests in the paper are based on application year.

[13] https://ipwatchdog.com/2017/12/06/changes-patent-language-ensure-eligibility-alice/

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 I 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, nor 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, reinforcing this interpretation.

Fact 3: There is large variation in *Obfuscation* across technology type and time.

Figure II plots the time series of our obfuscation measure across six broad technology classes: human necessities, transport, chemistry, mechanical engineering, physics, and electricity. Obfuscation does not follow identical trajectories across categories, revealing substantial cross-sectional heterogeneity. While 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, other categories change rank order over time. For example, physics patents exhibit relatively high obfuscation around 1990 but move toward the bottom of the distribution by 2000, while obfuscation for chemistry patents is relatively low at the beginning and end of the sample but elevated between 1990 and 2005.

Finally, Figure II shows that the recent rise in obfuscation is a pervasive phenomenon across technology classes rather than an artifact of any single domain, reinforcing the interpretation that the upward trend reflects economy-wide disclosure incentives rather than sector-specific technological change.

[Insert Figure II about here]

Fact 4: *Obfuscation (Fundamental Complexity)* is negatively (positively) related to patent impact measures.

Figure III presents binscatter plots relating *Obfuscation* (Panel A) and *Fundamental Complexity* (Panel B) to patent value from Kogan et al. (2017), citations, and RETech (a measure of exposure to rapidly evolving technology areas, drawn from Bowen III et al., 2023).

The subpanels in Panel A show that highly obfuscated patents have significantly lower patent valuations and receive fewer citations. The negative correlation with RETech shows that firms obfuscate more in slower-moving technological domains. Such domains are characteristic of mature or stagnant industries, where patents tend to serve strategic purposes rather than reflect substantive innovation. This correlation thus demonstrates that obfuscation captures a strategic dimension of patenting.

[Insert Figure III about here]

The subpanels in Panel B relate the same outcomes to *Fundamental Complexity*, which captures the inherent technological sophistication of the invention, distinct from strategic disclosure choices. The relationships between *Fundamental Complexity* and the outcome metrics are positive, indicating that fundamentally more complex inventions are more valuable. This pattern aligns with several theoretical mechanisms: technologically complex inventions represent larger departures from existing knowledge and thus command higher monopoly rents; they require broader claim scope to adequately protect the innovation, enhancing their protection from infringement; they impose higher costs on competitors attempting to design around the patent; and they contain more codified technical knowledge that generates valuable spillovers to follow-on inventors.

As measured by stock market reactions to patent grants, fundamentally complex patents create greater shareholder value. They receive higher citation counts, as subsequent inventors must build upon rather than circumvent these foundational innovations. They are also situated in more dynamic technology areas where the pace of cumulative innovation is faster. Importantly, these positive relationships contrast sharply with the negative effects of strategic complexity, validating that our LLM-based measure successfully separates the technological signal (fundamental complexity) from the strategic component (obfuscation).

Fact 5: *Obfuscation* increases as competitive threats rise, and falls as the value of deterrence drops.

In the previous section, the negative correlation between *Obfuscation* and RETech already suggested a strategic aspect to obfuscation, in that it becomes more common in mature or stagnant industries. We now connect obfuscation with strategy more explicitly by directly examining how it varies with exposure to competitive forces.

[Insert Figure IV about here]

Figure IV sorts patents by obfuscation (Panel A) and fundamental complexity (Panel B), again using binned scatterplots that control for application year fixed effects. Panel (1) examines the relationship with HHI measured using SIC 3-digit industries, and panel (2) uses HHI measured using the text-based industries of Hoberg and Phillips (2016). Both measures of obfuscation rise as the firm faces threats from more dominant industry leaders. Additionally, because deterring competitor entry is more valuable for market leaders, panel (3) shows the relationship with market share. Across all three columns, obfuscation is positively related to industry concentration and market share, consistent with the strategic value of deterrence.

We also find evidence in panels (4) and (5) that obfuscation is lower for patents after competing firms develop more patents and more valuable patents. After competitors receive additional patents, the value of deterrence drops. Consistent with this mechanism, the relationship between realized competitor innovation and deterrence is strongly negative and linear.

Panel B compares HHI, market share, and peer patent outcomes with fundamental complexity and reports correlations in the opposite direction of those in row 1. We find that patents are relatively simpler when the inventing firm is in a more concentrated industry, has a larger market share, and when peers have less innovation success. The flipped relationships highlight the value of our decomposition of total patent complexity into its two components. If we focus on total complexity instead of the two components separately, we will attenuate our estimates of how competitive pressures relate to patent disclosure choices.

To highlight the value of our new measure in contrast to existing measures that might proxy for patent disclosure clarity, Figure A3 repeats the analysis in Figures III and IV, but focuses on the length of the patent abstract and the number of claims.[14] The correlation between these variables and *Obfuscation* are not strong: 0.18 and -0.07, respectively. In contrast to the results above regarding *Obfuscation*, Figure A3 shows that both abstract length and claim counts are positively related to citations and realized peer patent success, and negatively related to competition and market share. This analysis highlights the value of our approach to measuring strategic disclosure in patents.

Fact 6: *Obfuscation* falls after firms begin using patents as collateral.

For a patent to be valuable as collateral, competitors and lenders need to understand the language of the patent clearly. Thus, when firms rely on the collateral value of their patents, they should 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

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

[14] Dyer et al. (2024) reports that the number of claims is the strongest predictor of non-final rejections due to insufficient disclosure.

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 $y \in \{\textit{Fundamental Complexity}, \textit{Obfuscation}, \textit{Total Complexity}\}$. We report the event-time coefficients (β_k) in Figure V with 95% confidence intervals based on standard errors clustered by firm-year.

[Insert Figure V about here]

Figure V 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 pledged patent, the overall complexity of the patent text (reported as blue triangles) is not statistically different from that of control firms that do not pledge patents.

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.

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.

4 Shocks to Patent Disclosure

In this section, we present evidence on the tradeoffs associated with patent obfuscation by exploiting several shocks to firms' ability or incentives to obfuscate patents. First, we employ an instrumental variable (IV) approach using patent examiner §112 strictness following Dyer et al. (2024). Second, we exploit a Supreme Court ruling, Nautilus v. Biosig, that raised the standard for patent disclosure. Third, we analyze the passage of the American Inventors Protection Act (AIPA), which affected the disclosure choices of certain firms. If financing considerations are important to firms' disclosure choices, we expect more-exposed firms to exhibit greater sensitivity to each of these shocks.

4.A IV Analysis: Patent Examiner §112 Strictness

Our first strategy is an instrumental-variables approach to identify the causal effect of patent obfuscation on subsequent financing choices and other outcomes. Following Dyer et al. (2024), we exploit plausibly exogenous variation in patent obfuscation due to the quasi-random assignment of patent examiners. As demonstrated in our IV results below, we identify reduced access to external financing as a novel cost of patent obfuscation, even after accounting for potential endogeneity concerns.

4.A.1 Institutional Details on Patent Examination

Similar to Dyer et al. (2024), our instrumental variable strategy leverages exogenous variation in patent obfuscation arising from the random assignment of patent examiners, who differ in their stringency when rejecting applications for disclosure-related reasons under Section 112. As described in detail in Section 4.A.2, we measure examiner strictness as the fraction of an examiner's applications that receive non-final rejections due solely to inadequate disclosure under §112, and use this measure as an instrument for patent obfuscation. In doing so, we build upon a growing literature that exploits the patent examination process at the USPTO

(Gaulé, 2018; Sampat and Williams, 2019; Farre-Mensa et al., 2020; Melero et al., 2020; Ertugrul et al., 2024; Dyer et al., 2024). Below we briefly describe the institutional details of the patent examination process, focusing on the features most relevant to our instrumental variable strategy.

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.

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

It is worth noting that patent 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.[15] This implies that examiners are likely to actively enforce clarity and compliance in abstracts as part of the examination process.

4.A.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

[15]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."

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 the Patent Examination Research Dataset to compile the track record of rejection rates by reason for each examiner over the same period.

We compute the instrument for obfuscation of a patent application a that is reviewed by examiner k as the §112 strictness of examiner k faced by a patent application a using the following:

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

In the above, Total §112 Rejections_k and 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 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).[16]

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

[16]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).

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.

As a way of validating our instrument, in Table IV we focus on a subset of patent applications for which we can observe the text of the application evolving through the review process, from initial review and rejection to final acceptance. We separately include measures for examiner rejection rates with respect to §112 along with §101, §102, and §103. Our estimates confirm that patents assigned to examiners that flag clarity issues exhibit improved readability by the end of the review process. This is consistent with Ashtor (2022), who shows that claims within patents rejected on §112 grounds become more simplified during the review process. Our findings in Table IV confirm that this pattern holds as well for the abstract of the patent, which is the focus of our analysis.

[Insert Table IV about here]

4.A.3 IV Results

We report the results of our IV analysis in Table V. 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 statistically and economically, we focus our discussion on Panel A.

In column 1, we report the first-stage result. As expected, the coefficient of examiner §112 strictness is negative and 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 the examiner §112 strictness is a strong determinant of patent obfuscation, confirming that the relevance condition required for a valid instrument is satisfied.

[Insert Table V about here]

The remainder of Table V 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).[17] 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). The lower level of citations from other firms is consistent with the notion that obfuscation deters other firms from operating in related technological areas or pursuing related patents by creating ambiguity around the exact scope of patent protection. 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. 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.

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 V 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. This draws a direct connection between patent obfuscation and external financing.

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

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

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.B Nautilus v. Biosig

In this section, the shock 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."[18]

Critically, the Court emphasized that this standard must account for "the inherent limitations of language" while still providing meaningful boundaries around patent rights. The "reasonable certainty" standard represented a middle ground: more demanding than the Federal Circuit's permissive approach, yet acknowledging that absolute precision in claim drafting is neither achievable nor required. Consistent with the ruling's objective of preventing unclear claim boundaries that could create legal uncertainty, Ashtor (2022) documents a significant increase in §112(b) rejections post-*Nautilus*. The Court's decision explicitly 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.

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 constrain firms' obfuscation choices, we predict heterogeneous treatment effects based on firms' reliance on external financing. Specifically, if obfuscation incentives are weaker for firms reliant on external financing, then such firms should already obfuscate less even in the absence of the *Nautilus* decision, such that the decision itself does not force any change in their behavior. Altogether, the decision should induce a decrease in obfuscation that is relatively 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 VI 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 VI about here]

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

In Panel A, the dependent variable is our obfuscation measure constructed using the Gunning Fog index. Our treatment variable, *Exposure*, captures firms that are more likely to be affected by the Nautilus ruling and is defined as the firm's median obfuscation level across patent applications filed before the ruling. The intuition 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. *Post* is an indicator variable equal to one if a patent application is filed after June 2, 2014 (the decision date), and zero otherwise.

As shown in column 1, the interaction term of *Exposure* x *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.[19] 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* as the *Z* variable, equal to one if the firm pledged patents as collateral during 2011-2013 in the period before the ruling, and zero otherwise. The coefficient on *Exposure* $\times$ *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* x *Exposure* x *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.

[19] 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%.

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* as the Z variable, equal to one if the firm's leverage exceeds the sample median (0.215) during 2011-2013, and zero otherwise. We find that the reduction in obfuscation is similarly attenuated for highly levered firms: while the coefficient on *Exposure* $\times$ *Post* remains negative and significant, the coefficient on *HighLev* $\times$ *Exposure* $\times$ *Post* is positive (0.292) and significant at the 1% level. Indeed, 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.[20] 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 firms relying more heavily on external financing maintain more transparent disclosure practices even before the ruling to optimize their patents' value to capital providers. Consequently, these firms face less pressure to adjust their disclosure practices in response to the Court's stricter clarity requirements, as their pre-existing disclosure strategies already emphasize transparency over strategic obfuscation.

4.C The Enactment of AIPA

The next disclosure shock 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

[20]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.

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, 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 patent applications that otherwise could have remained confidential until they were granted (or forever if they were never granted). 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 build on this finding by pointing out that these incentives for obfuscation vary 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 activity for the average firm, this effect should be significantly weaker among firms that are highly reliant on external financing.

Our methodology builds on Ireland et al. (2024). To sort patents based on exposure to AIPA, they highlight 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 that requires publication 18 months after filing. This opt-out provision created a natural experiment, as it generated 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.

We begin by confirming their core insight using our obfuscation measure. In untabulated results, we find that patents that sought foreign protection post-AIPA saw a relative increase in obfuscation compared to patents that did not. Compared to Ireland et al. (2024), the key difference is to use our obfuscation measure that separates obfuscation from fundamental complexity, rather than looking at the total complexity of the language in the patent filing as they do. The result thus confirms that AIPA led to an increase in obfuscation, not just an increase in patents that are fundamentally difficult to describe.

We next elaborate slightly on their strategy. The decision to seek foreign protection post-AIPA 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.[21] 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 VI 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. 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

[21]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.

presented in Table VII.

[Insert Table VII about here]

In column 1 of Table VII, 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 1.2 grade levels to 6.5 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, creating ambiguity about the exact scope of the patent being filed.

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* as the Z variable) 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 ($\mathrm{Z} \times Exposure \times Post$) 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 relative 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 ($Exposure \times Post$) 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×*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 tables, 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 are a key indicator of a firm's default risk and play a critical role in shaping firms' access to debt financing. For example, Faulkender and Petersen (2006) documents that rated firms have significantly greater access to public debt markets, face lower borrowing costs, and adopt higher leverage ratios than comparable unrated firms. The importance of credit ratings extends beyond mere access: rating downgrades can trigger restrictive covenants, increase the cost of existing debt, and impair firms' ability to refinance maturing obligations. Consequently, as Kisgen (2019) demonstrates, firms actively manage their financial policies to maintain favorable credit ratings and avoid costly downgrades, treating rating thresholds as binding constraints on their financing decisions.

Credit rating agencies assign ratings based on a combination of quantitative and qualitative factors, among which leverage is one of the most important quantitative indicators. However, when calculating a firm's leverage, CRAs 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. These adjustments include items such as defined benefit pension liabilities, operating lease obligations, securitization transactions, and hybrid securities with both debt and equity characteristics. The economic rationale for these adjustments is to better reflect the underlying economics of transactions and events and to improve the comparability of financial statements across firms with different accounting treatments (Moody's, 2006). For instance, two otherwise identical firms might report different GAAP leverage if one leases its assets while the other purchases them with debt. Credit rating agencies aim to see through such accounting differences and assess firms' true economic leverage and debt servicing capacity.

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.

Beginning in February 2006, Moody's fundamentally revised 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 heterogeneous 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 VIII 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 VIII.

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.

6 Conclusion

The patent system is built on a basic exchange: inventors receive legal protection in return for disclosing knowledge that supports further innovation. This paper shows that the nature of that exchange is 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.

An instrumental variables approach finds that obfuscation deters competitive entry and reduces litigation risk, but simultaneously lowers a patent's marketability and its usefulness as collateral. This establishes a fundamental tradeoff between strategic ambiguity and financing capacity, and evidence from three additional identification strategies supports this interpretation.

Our findings highlight a channel through which financial considerations influence the clarity of information that enters the public domain. Because financial constraints can incentivize inventors to reduce obfuscation and improve disclosure choices at the level of individual patents, our results establish a positive externality of financial constraints and demonstrate a new way that financial forces can shape the diffusion of technological knowledge.

The LLM-based framework developed here offers a general tool for analyzing similar disclosure tradeoffs in other contexts. By constructing counterfactual text that maintains information content while improving clarity, our approach allows researchers to decompose textual complexity into fundamental and strategic components. This decomposition can be applied to corporate disclosures beyond patents, including earnings announcements, regulatory filings, and loan contracts, wherever clarity and strategic ambiguity may coexist. More broadly, our findings suggest that policies affecting either patent disclosure requirements or firms' access to financing can have spillover effects on the transparency of technological innovation, with implications for knowledge diffusion, follow-on innovation, and the efficiency of intellectual property markets.

References

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

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.

Bernstein, S., 2015. Does going public affect innovation? The Journal of finance 70, 1365–1403.

Bhattacharya, S., 1983. Imperfect information, dividend policy, and "The Bird in the Hand" fallacy. Bell Journal of Economics 14, 259–270.

Bhattacharya, S., Ritter, J.R., 1995. Innovation and communication, signalling with partial disclosure. Review of Economic Studies 62, 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., Gow, I., Taylor, D., 2018a. Liguistic complexity in firm disclosures: Obfuscation or information? Journal of Accounting Research 56, 85–121.

Bushee, B.J., Gow, I.D., Taylor, D.J., 2018b. 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.

Ertugrul, M., Krishnan, K., Xu, B., Yu, Q., 2024. Corporate patenting, customer capital, and financial market outcomes. Journal of Financial and Quantitative Analysis 59, 2425–2458.

Farre-Mensa, J., Hegde, D., Ljungqvist, A., 2020. What is a patent worth? evidence from the u.s. patent "lottery". The Journal of Finance 75, 639–682.

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., 2001a. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. Journal of accounting and economics 31, 405–440.

Healy, P.M., Palepu, K.G., 2001b. 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.

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.

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.

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.

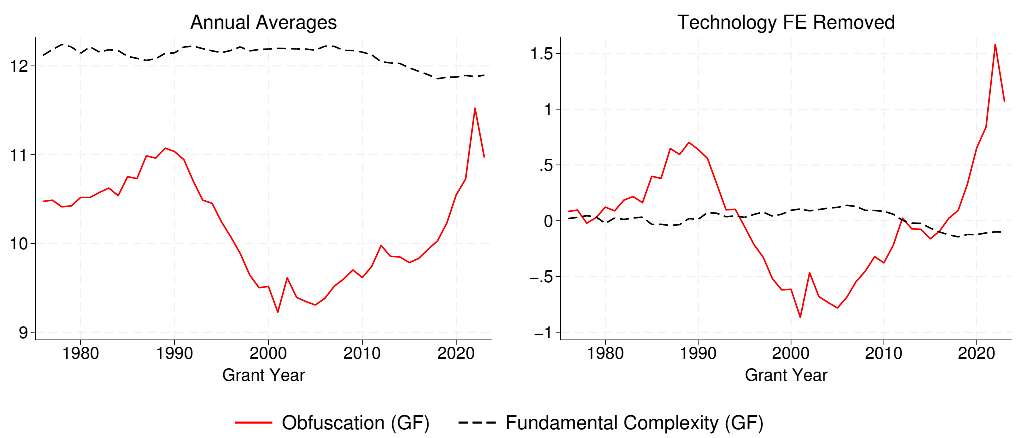
Verrecchia, R.E., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194.

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

Figure I: 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 principle component of the readability measures listed in Table A2 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

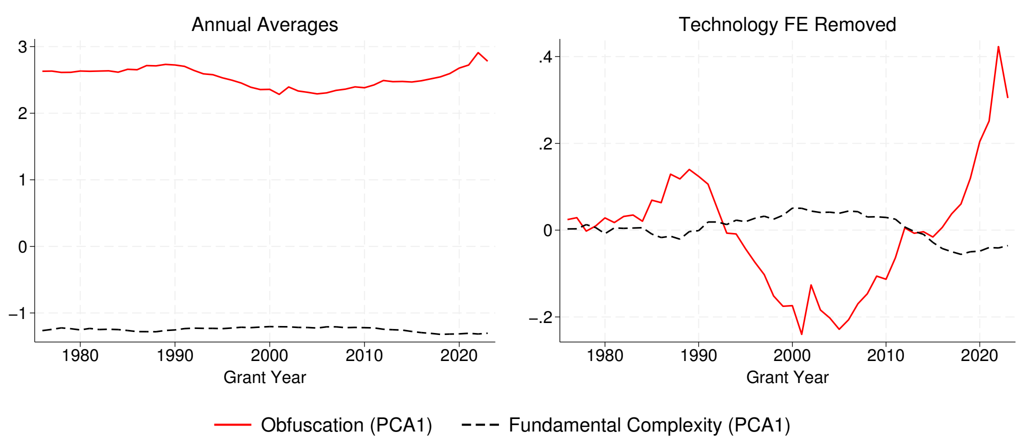


Figure II: Time Series of Obfuscation, by Technology Category

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. (Back to text.)

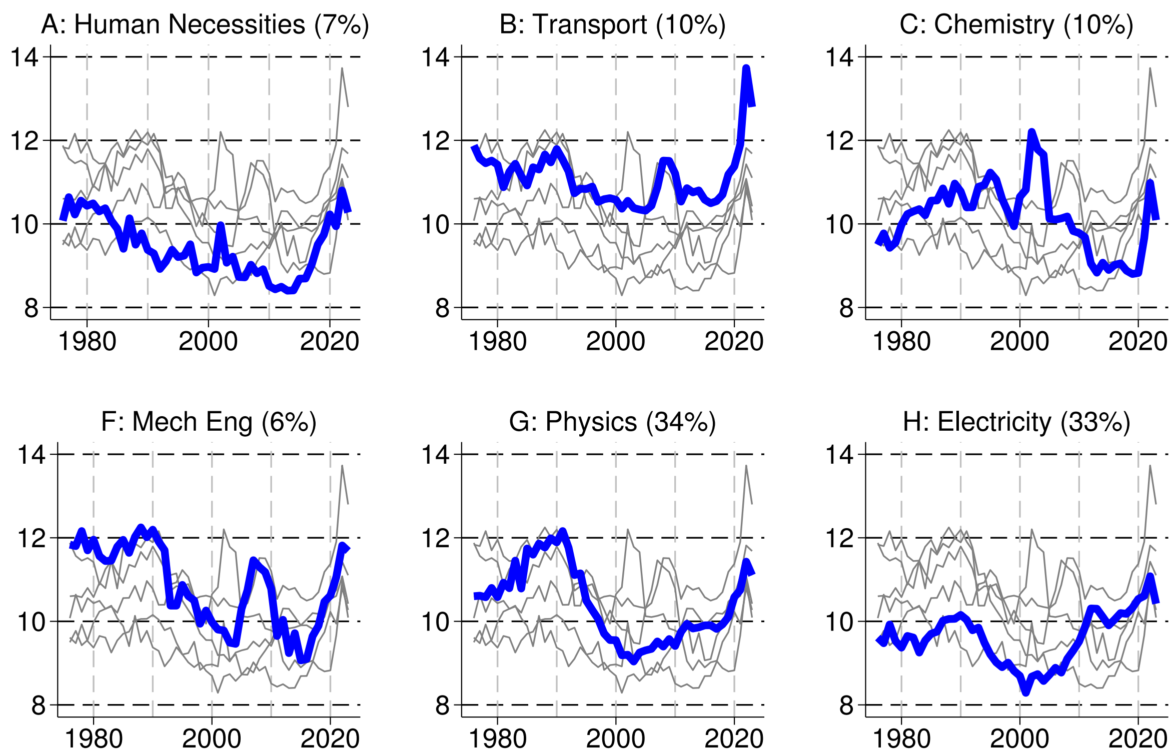
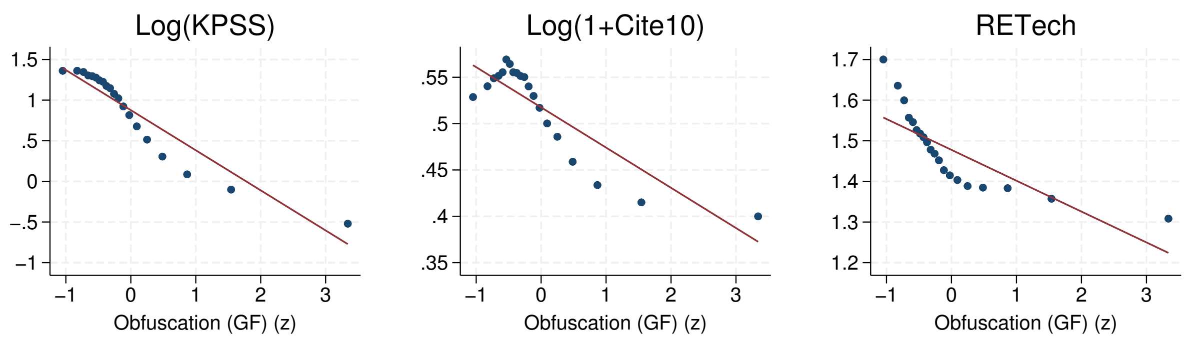


Figure III: Relation with Key Patent Metrics

This figure shows binned scatter plots illustrating the relation between *Obfuscation (GF)* and key patent metrics (Panel A). As a benchmark, we also plot the relation between *Fundamental Complexity (GF)* and key patent metrics (Panel B). *Log(KPSS)* is the natural log of the patent value estimate from Kogan et al. (2017). *Log(1+Cite10)* represents the natural logarithm of one plus the number of forward citations received within 10 years of the grant date. *RETech* measures whether a patent pertains to a technological area that is rapidly evolving (i.e., following breakthroughs) or stable (Bowen III et al., 2023). To facilitate interpretation, both *Obfuscation* and *Fundamental Complexity* are standardized. All plots control for application year fixed effects. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog



Panel B: Fundamental complexity measured via Gunning Fog

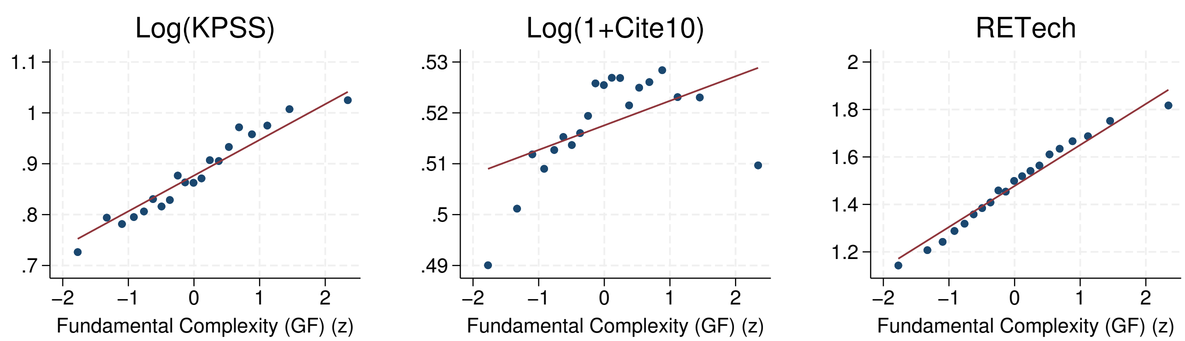
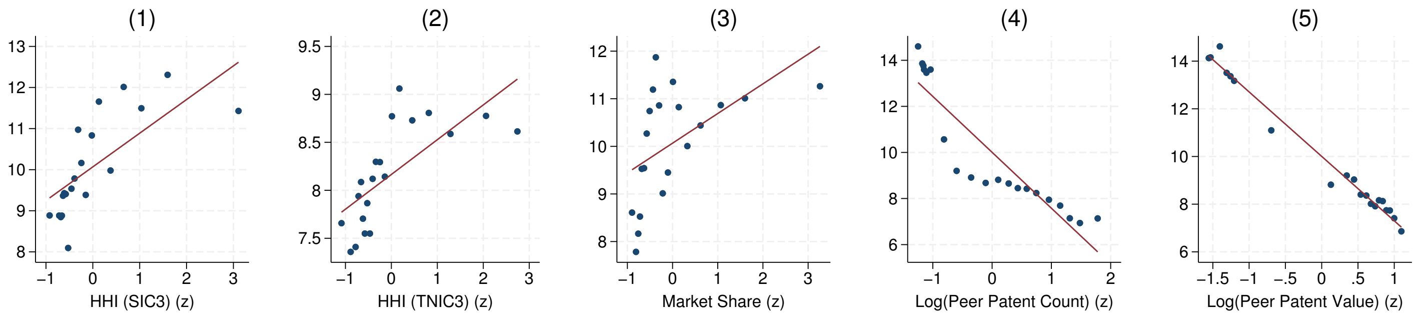


Figure IV: Relation with Competition Measures

This figure shows binned scatter plots illustrating the relation between *Obfuscation (GF)* and competition metrics in Panel A. As a benchmark, we also plot the relation between *Fundamental Complexity (GF)* and competition metrics in Panel B. *HHI (SIC3* and *HHI (TNIC3)* are industry sales concentration measured within SIC 3 digit industries and TNIC3 industries from Hoberg and Phillips (2016), respectively. *Market Share* is based on the market share of sales within SIC3 industries. *Peer Patent Count* is the number of patents granted to TNIC3 industry competitors. *Peer Patent Value* is the KPSS value of patents granted to TNIC3 industry competitors. To facilitate interpretation, independent variables in each figure are standardized. All plots control for application year fixed effects. (Back to text.)

Panel A: Patent obfuscation measured via Gunning Fog



Panel B: Fundamental complexity measured via Gunning Fog

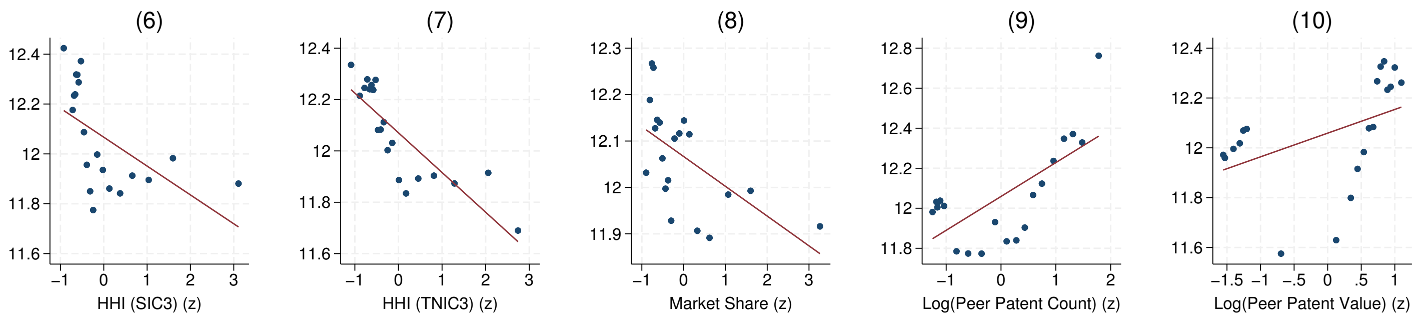


Figure V: Obfuscation Around Firm's First Patent Pledge
This figure shows how patent complexity evolves around the year of a firm's first pledging event. Event-time coefficients (β_k) are estimated by

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

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 $y_{p,i,t} \in \{\textit{Total Complexity}, \textit{Fundamental Complexity}, \textit{Obfuscation}\}$ measured via Gunning Fog. The 95% confidence interval is based on standard errors clustered by firm-year.
(Back to text.)

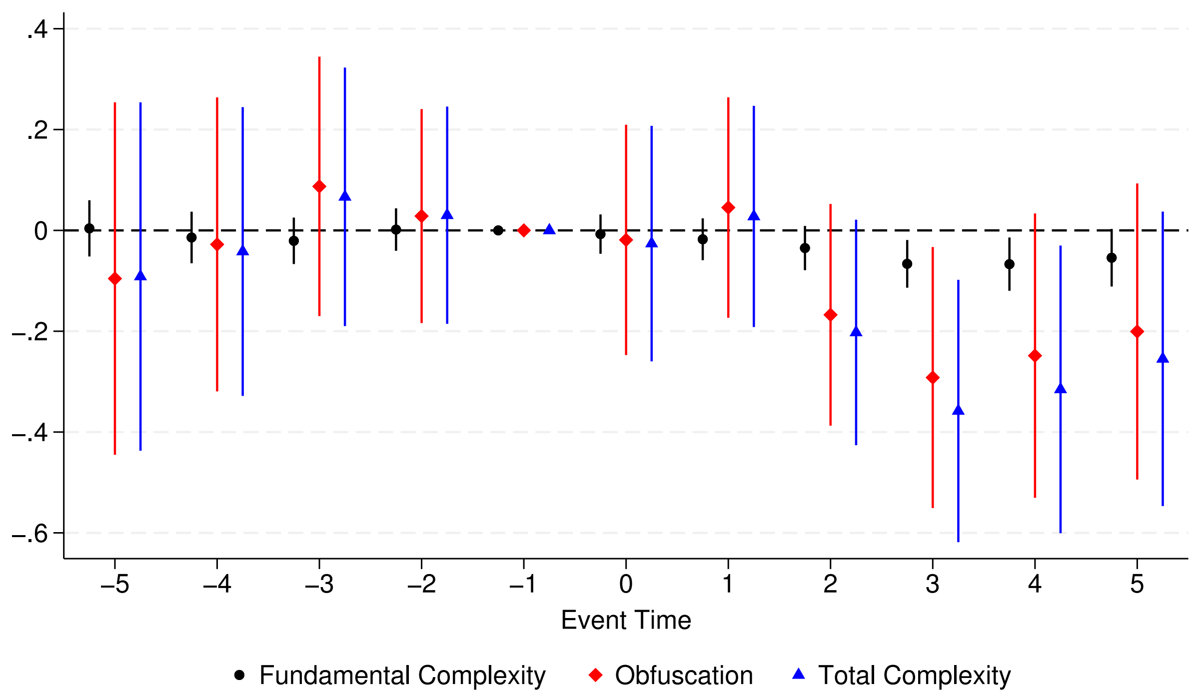


Table I: Examples of Patents with Obscured Language

This table shows three patent abstracts and the LLM-simplified version. Patent 8786366 is Intel Mobile Communications' "Amplifier circuit" patent. Patent 9547842 is IBM's "out of office" auto reply patent. Patent 8786025 is TSMC's "resistor" patent. (Back to text.)

Patent	Abstract *Source for Total Complexity*	LLM-Simplified Version *Source for Fundamental Complexity*
8786366	An amplifier circuit is described comprising a first field effect transistor comprising a first source/drain terminal coupled to a first supply terminal, a second source/drain terminal coupled to an output of the amplifier circuit and a gate terminal; a second field effect transistor comprising a first source/drain terminal coupled to an input of the amplifier circuit, a second source/drain terminal coupled to the gate terminal of the first field effect transistor and a gate terminal; a third field effect transistor comprising a first source/drain terminal coupled to a first bias current source of the amplifier circuit, a second source/drain terminal and a gate terminal coupled to its first source/drain terminal and the gate terminal of the second field effect transistor; a fourth field effect transistor comprising a first source/drain terminal coupled to a second bias current source, a second source/drain terminal coupled to a second supply terminal and a gate terminal coupled to the second source/drain terminal of the third field effect transistor; and a control circuit coupled to the gate of the fourth field effect transistor configured to control the source drain voltage of the fourth field effect transistor by means of the gate of the fourth field effect transistor to be equal to a reference voltage.	This patent describes an amplifier circuit with four transistors and a control circuit. The first transistor connects to the power supply and the output. The second transistor links the input to the first transistor's gate. The third transistor connects to a bias current source and the second transistor's gate. The fourth transistor connects to another bias current source and the power supply. A control circuit regulates the fourth transistor's voltage to match a reference voltage.
9547842	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.
8786025	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.

Table II: 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 principle component of the readability measures listed in Table A2 (*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 2 and Table A1. (Back to text.)

	N	Mean	SD	p10	p50	p90
Main variables						
Obfuscation (GF)	2,607,015	10.06	11.96	0.86	6.57	23.53
Fundamental Complexity (GF)	2,607,015	12.07	2.91	8.56	11.88	15.74
Total Complexity (GF)	2,607,015	22.13	11.84	12.87	18.65	35.46
Obfuscation (PCA1)	2,607,015	2.52	2.92	0.20	1.70	5.81
Fundamental Complexity (PCA1)	2,607,015	-1.26	0.83	-2.28	-1.29	-0.20
Total Complexity (PCA1)	2,607,015	1.26	2.87	-1.05	0.46	4.49
Outcome variables						
KPSS	2,607,015	12.54	34.05	0.07	3.91	29.30
Log(KPSS)	2,607,015	0.88	2.28	-2.59	1.36	3.38
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
RETech	2,586,165	1.48	2.13	-0.06	1.10	3.31
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
HHI (SIC3)	2,600,729	0.15	0.14	0.04	0.10	0.34
HHI (TNIC3)	2,607,015	0.17	0.25	0.00	0.09	0.51
Market Share	2,602,010	0.15	0.19	0.01	0.09	0.40
Log(Peer Patent Count)	2,355,338	1.97	1.74	0.00	1.95	4.43
Log(Peer Patent Value)	2,607,015	6.68	5.56	0.00	9.40	12.64
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

Table III: Information Asymmetry, Trade Secrets, and Obfuscation

This table examines how our measures of patent complexity are related to information asymmetry and the use of trade secrets. The sample is patents granted from 1976-2023 to publicly traded firms where the dependent variables can be measured. The dependent variable in Panel A is the percentile rank (0-100) of firm information asymmetry for trading days [t+0, t+25] around the patent disclosure. The dependent variable in Panel B is equals 100 if the firm's 10-K in the patent application year discusses trade secrets, and 0 otherwise. *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 difference between total and fundamental complexity. Other variables are defined in Section 2 and Table A1. 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: Information Asymmetry

	(1)	(2)	(3)	(4)	(5)
Obfuscation (GF) (z)	0.423***		0.658***	0.410***	0.398***
	(0.09)		(0.17)	(0.09)	(0.05)
Total Complexity (GF) (z)		0.390***	-0.245		
		(0.09)	(0.18)		
Fundamental Complexity (GF) (z)				-0.060	0.039*
				(0.04)	(0.02)
Log(AT) (z)					-18.220***
					(0.15)
Fixed effects:					
CPC-year	Yes	Yes	Yes	Yes	Yes
Obs.	1,802,286	1,802,286	1,802,286	1,802,286	1,802,281
R^2	0.49	0.49	0.49	0.49	0.84

Panel B: Trade Secrets

	(1)	(2)	(3)	(4)	(5)
Obfuscation (GF) (z)	-0.338**		-0.671**	-0.320**	-0.277*
	(0.15)		(0.33)	(0.16)	(0.16)
Total Complexity (GF) (z)		-0.300*	0.348		
		(0.16)	(0.36)		
Fundamental Complexity (GF) (z)				0.085	0.066
				(0.09)	(0.09)
Log(AT) (z)					-7.232***
					(0.75)
Fixed effects:					
CPC-year	Yes	Yes	Yes	Yes	Yes
Obs.	1,455,104	1,455,104	1,455,104	1,455,104	1,455,103
R^2	0.27	0.27	0.27	0.27	0.29

Table IV: 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 one 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 V: Effects of Patent-Level Obfuscation

This table reports the effects of patent-level obfuscation estimated using an IV framework. The sample is patents granted from 2008-2023 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 one if the patent is involved in litigation, sold to another entity, or used as collateral by year-end 2024, and zero otherwise, respectively. Other variables are defined in Section 2 and Table A1. 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)	(2)	(3)	(4)	(5)	(6)	(7)
	Obf. (z)	All	Others	Self	Infringed	PatSold	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)	(2)	(3)	(4)	(5)	(6)	(7)
	Obf. (z)	All	Others	Self	Infringed	PatSold	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 VI: 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 VII: 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-difference 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 VIII: Financing Constraints and Patent Obfuscation: Moody's Change in Adjustment Methodology in 2006

This table examines how a firm's financing constraints affects 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-difference 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

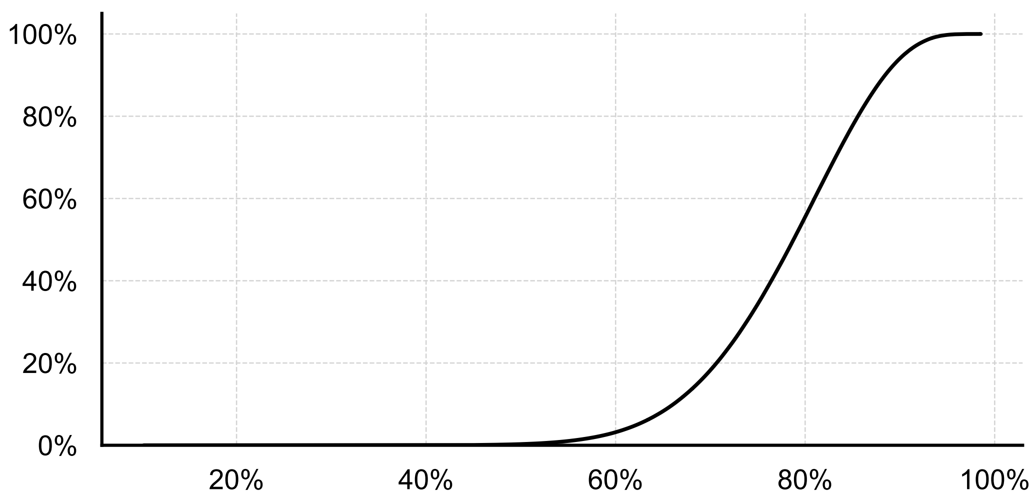
Appendix

- Figure A1 shows the semantic similarity of patent abstracts and the LLM written versions.
- Figure A2 examines how *Fundamental Complexity* evolves over time across the six largest technology sections in the CPC classification scheme from 1976 to 2023.
- Table A1 defines variables at the patent level.
- Table A2 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 A3 shows the correlation between Obfuscation measures computed using six alternative LLMs.

Figure A1: 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

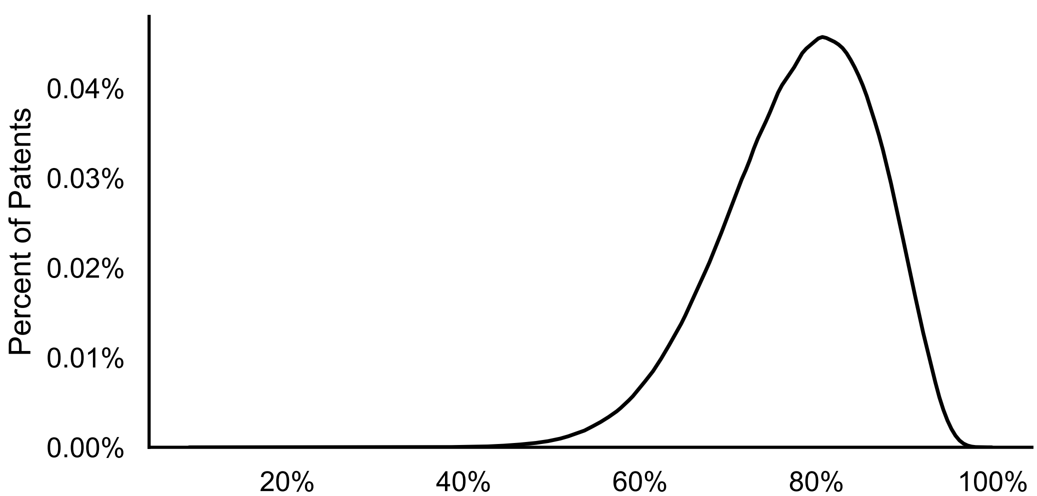


Figure A2: Time Series of *Fundamental Complexity*, by Technology Category
This figure plots the evolution of *Fundamental Complexity*, measured by applying Gunning Fog to LLM rewritten patent abstracts. The figure separately reports averages across the six largest technology sections in the CPC classification scheme from 1976 to 2023. Each subpanel reports in parentheses the fraction of patents over the entire sample belonging to that category.

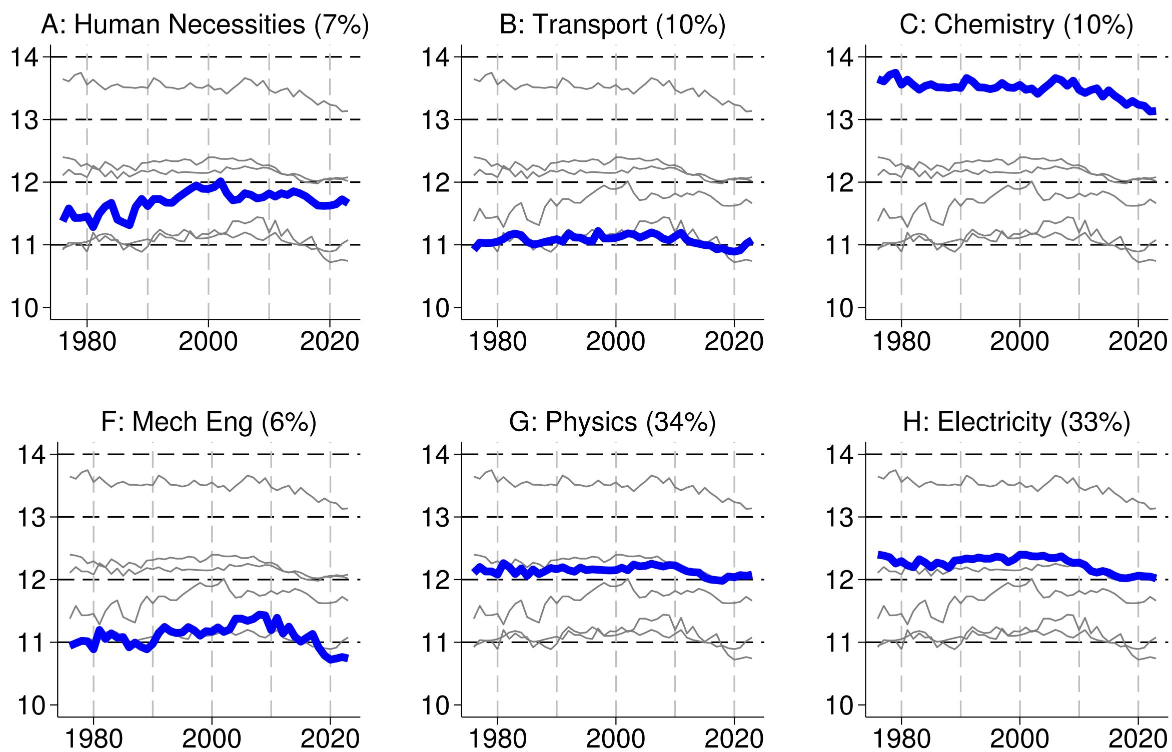
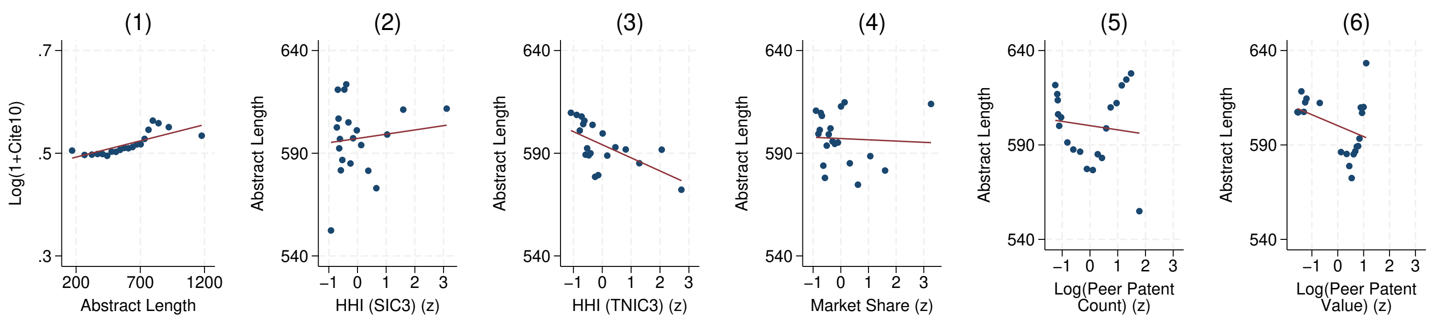


Figure A3: Other Possible Proxies for Complexity

This figure shows binned scatter plots relating possible proxies for complexity to strategic outcomes and as a function of strategic determinants. Panel A examines patent length, measured by the number of characters in the abstract, and Panel B examines the number of claims in the patent. The left column of each panel reports citations as the dependent variable, as in Figure III, where a *decrease* in citations can be interpreted as effective deterrence as in Dyer et al. (2024). The remaining columns report competition metrics as in Figure IV. *HHI (SIC3* and *HHI (TNIC3)* are industry sales concentration measured within SIC 3 digit industries and TNIC3 industries from Hoberg and Phillips (2016), respectively. *Market Share* is based on the market share of sales within SIC3 industries. After peers receive patents, the value of deterrence falls. *Peer Patent Count* is the number of patents granted to TNIC3 industry competitors. *Peer Patent Value* is the KPSS value of patents granted to TNIC3 industry competitors. To facilitate interpretation, independent variables denoted with "(z)" are standardized. All plots control for application year fixed effects. (Back to text.)

Panel A: Abstract Character Length



Panel B: Number of Patent Claims

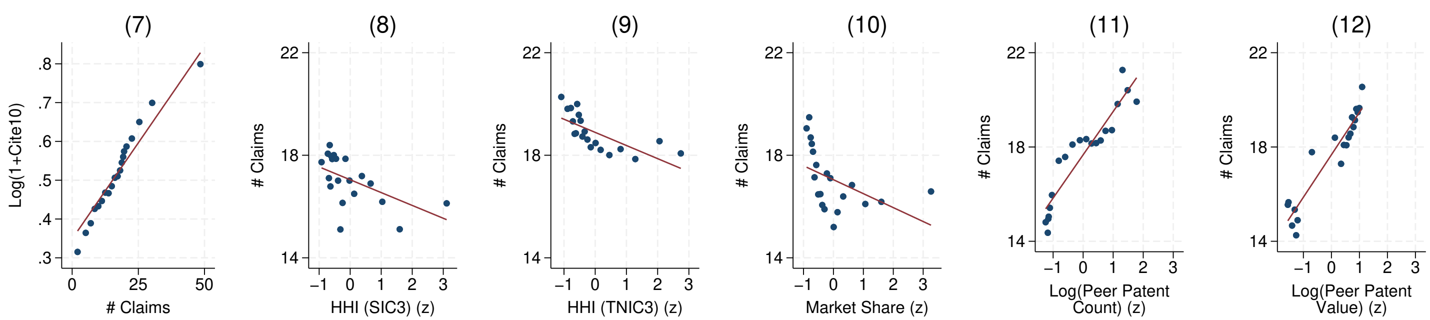


Table A1: Variable Definitions

Data sources are described in Section 2.

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 that do not meet disclosure clarity requirements under §112 in the first review action but pass other requirements
Examiner Other Strictness	For a given patent, the fraction of other applications handled by the examiner that do not meet requirements under §101, §102, or §103 in the first review action but pass other requirements
	Outcome variables
Information Asymmetry	Follows Bushee et al. (2018b), 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.
Trade Secrets	100 if the firm's 10-K in the patent application year discusses trade secrets, and 0 otherwise. Data from Glaeser (2018) available at `https://www.stephenglaeser.com/data`
KPSS	Patent value estimate from Kogan et al. (2017)
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
RETech	Exposure to rapidly evolving technology areas (i.e., following breakthroughs) (Bowen III et al., 2023)
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
HHI (SIC3)	Industry sales concentration measured within SIC 3 digit industries
HHI (TNIC3)	Industry sales concentration measured within TNIC3 industries (Hoberg and Phillips, 2016)
Market Share	Market share of sales within the firm's SIC3 industry
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.

Variable	Definition
Log(Peer Patent Value)	Log of one plus the total KPSS value for 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 A2: 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
Principle Components			
Eigenvalues	6.04	1.20	0.63
CumulExplained	0.76	0.91	0.98

Table A3: 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	