The Value of Corporate Patent Utilization*

Jarrad Harford, Qiyang He, and Buhui Qiu

This Version: April 2026

Abstract

We use textual analysis of firm patent and product filings to construct a novel measure of patent utilization rate, which reflects the extent to which a firm's patents are applied in its new products. We validate the measure by showing that patents utilized in new products are more likely to belong to firms' core technology fields, receive more self-citations, and are less likely to be sold, and that new products supported by more patents receive higher announcement returns and are more likely to be breakthrough innovations. Firms with higher patent utilization rate experience more active new product development, stronger market share growth, and higher profitability improvement and valuation. The effects are predominantly driven by the utilization of high-value patents, and are more pronounced for firms in competitive product markets. We address endogeneity concerns using a shift-share instrumental-variable approach and show robust findings. Our results highlight the costs of patent underutilization and the strategic importance of innovation commercialization.

Keywords: Machine Learning, Textual Analysis, Patent Utilization, New Product Development, Market Share, Firm Valuation

JEL classification: G30, L25, O34

---

*Jarrad Harford is affiliated with The University of Washington; Address: 571 Paccar Hall, Foster School of Business, University of Washington, Seattle, WA 98195-3226, USA; Phone: +1-206-543-4796; Email: jarrad@uw.edu. Qiyang He is affiliated with Deakin University; Email: qiyang.he@deakin.edu.au. Buhui Qiu is affiliated with The University of Sydney; Email: buhui.qiu@sydney.edu.au. We thank Thomas Chemmanur (discussant), Baojun Gao, Alex He (discussant), Bin Li, Xing Liu (discussant), Jane Luo (discussant), Kaitang Zhou, conference participants of the 2026 American Finance Association (AFA) Annual Meeting, the 2025 CUHK-RAPS-RCFS Conference, the 2025 European Finance Association (EFA) Annual Meeting, the 2025 Monash Winter Finance Conference and the 2025 China Accounting and Finance Conference, and seminar participants at Wuhan University and Huazhong Normal University for helpful comments and suggestions. All errors are our own.

"*...There is an incredible amount of technology that's packed into the product. There are 5,000 patents in the product (Vision Pro) and it's, of course, built on many innovations that Apple has spent multiple years on, from silicon to displays and significant AI and machine learning.*"

— Tim Cook, *February 01, 2024, Apple Inc. 2024Q1 Earnings Call*

1 Introduction

Technological innovation is a fundamental driver of economic growth (Schumpeter, 1943; Aghion and Howitt, 1992). Among various forms of intellectual property, patents, legal instruments that grant exclusive rights to inventors and protect inventions from infringement, have received particular attention. Recent empirical work provides valuable evidence on links between patented innovations and firm growth (e.g., Kogan et al., 2017; Bowen et al., 2023; Ma, 2025). Nevertheless, patents represent only an intermediate stage of a firm's innovation process, while the ultimate goal is to commercialize these inventions for profits through incorporating the patented technologies into new products. In this study, we focus on the commercialization side of firms' patented innovations. We aim to deepen our understanding of the extent to which a firm's patent portfolio contributes to its new product development and the implications for future firm performance.

Our motivation stems from the observation that while U.S. patents per capita have more than doubled since 1976 (see Figure A1), many of them fail to achieve commercialization. The successful utilization of patented technologies in new products depends not only on scientific merit and economic viability, but also on firm-level capabilities and strategy. Some firms accumulate patents primarily to defend positions and deter rivals rather than to bring technologies to market,[1] whereas others may lack the resources, complementary assets, or organizational processes required to reach the final stage of commercialization. This distinction between patent

[1] Prior work suggests that U.S. firms are increasingly engaging in strategic patenting (Gilbert and Newbery, 1982; Argente et al., 2020). Such "patent portfolio races" coincide with a dramatic increase in the number of patent production accompanied by a noticeable decline in patent quality and stagnating productivity growth (Hall and Ziedonis, 2001; Choi and Gerlach, 2017; Bloom et al., 2020; Kalyani, 2022), as shown in Figure A1.

production and commercialization is economically first-order, yet it remains underexplored because standard patent-based metrics capture inventive output but not the extent to which patents are actually embedded in products. This paper fills this void by quantifying corporate patent utilization rate in new product development (patent utilization hereafter) using machine learning.

Underutilization of patent portfolios in new product development can erode competitive position and operating performance. The collapse of Eastman Kodak Company represents a prominent example: In the 1980s, Kodak was a leading camera film producer and the fifth largest patent inventor (Moretti, 2021). However, with the rise of digital photography, Kodak's market share began to decline substantially. It is noteworthy that Kodak engineers Gareth Lloyd and Steven Sasson had developed and patented the first digital camera as early as 1977.[2] Yet, Kodak refused to incorporate this breakthrough innovation into the product pipeline, as the management team feared it would cannibalize the firm's film-based business. This strategic decision ultimately proved to be a critical mistake. Digital photography soon emerged as a dominant technology and Kodak encountered a substantial decline in market share. The company eventually filed for bankruptcy in 2012. Therefore, understanding the extent to which a firm's patent portfolio contributes to new product development offers significant implications for future firm performance.

We measure corporate patent utilization based on the premise that a patent is utilized in a product if there is a high textual similarity between the patent filing text (obtained from PatentsView) and the new product launch text description (obtained from Capital IQ Key Development database).[3,4] A potential concern is that the language used in patent filings can differ significantly from that in product text descriptions. Hence, the traditional "bag-of-words"

[2] The patent is titled "Electronic still camera" (US4131919A), filed in 1977 and granted in 1978. For technical details, see https://patents.google.com/patent/US4131919A/en.

[3] The assumption is similar in spirit to the innovation literature investigating knowledge diffusion across firms. Prior studies typically use patent citation to proxy for knowledge diffusion (see., e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Singh and Marx, 2013; Arora et al., 2021; Fadeev, 2023; Cohen et al., 2023). Our criteria for patent utilization in products are more stringent in that we require high text similarity between patents and products.

[4] We focus on non-process (i.e., product) patents, as process patents primarily enhance production efficiency, which is not the focus of this study (Bena et al., 2022).

approach, which requires exact overlap in terms, may inaccurately measure patent-product pair text similarity.

To address this challenge, we employ the pretrained machine learning language model, *FastText* (Bojanowski et al., 2017), which builds on the architecture of the *Word2vec* model (Mikolov et al., 2013a). While both models represent words as semantic vectors, *FastText* takes a step further by accounting for rare or out-of-corpus words, providing a more nuanced understanding when comparing texts from different sources.[5] Leveraging the *FastText* model, we calculate the textual similarity score for each within-firm patent-product pair. Figure 1 illustrates the process of how we determine whether a patent is utilized in a new product. Suppose firm A launches two new products (NP1, NP2) in 2015. We collect the firm's patents applied for (that were later granted) in the prior five years (e.g., PAT1, PAT2, and PAT3). We compute the textual similarity for each patent-product pair and consider a patent utilized if its similarity score is above the 80th percentile of our sample patent-product pair scores.[6]

We conduct various validation tests at both the patent and product levels to verify our patent utilization measure. At the patent level, we investigate whether patents utilized in new products are more likely to belong to firms' core technology fields than unutilized ones, as firms tend to commercialize technologies in areas where they have established strong scientific foundations and accumulated deep market insights. The results confirm this expectation. We further hypothesize that commercialized patents will receive more self-citations (Hall et al., 2001), because firms are more likely to build upon these utilized patents when developing subsequent generations of products. Consistent with this conjecture, specifications including $Firm \times Class \times Year$ fixed effects show that these utilized patents are cited more by the same firm's later patents, exhibit higher self-citation ratio, and are less likely to be sold to other entities, relative to unutilized patents filed by the same firm in the same year and within the

[5] The *Word2vec* model fails to provide semantic vector representations for words that are rarely seen or out of the training corpus. As our text data originates from patent filings and new product launch text descriptions, they likely contain extensive technological words that are rarely seen or entirely absent in conventional training corpora, which can lead to absence of vector representations for those words. Details are provided in section B3.3 in Appendix B.

[6] Results are robust to alternative cutoffs (e.g., 70th percentile, 90th percentile) and different product/patent portfolio windows.

same technology class.

Next, at the product level, we examine whether patent-backed new products are of higher quality. We measure product quality using two metrics: i) economic value, defined as the cumulative abnormal stock return in a three-day window surrounding the new product announcement (Kogan et al., 2017), and ii) breakthrough index, a text-based measure capturing a product's impact and novelty (Kelly et al., 2021). The results reveal that new products incorporating more patents exhibit higher quality: they generate significantly higher announcement returns and are more likely to be classified as breakthrough products.

Having validated the measure, we turn to the main research question: does corporate patent utilization impact future firm performance? To answer this question, we first generate a firm-year-level patent utilization rate measure, calculated as the number of granted patents applied for in the past five years by a firm that have been utilized in the new products launched by the same firm in the current year, scaled by the total number of granted patents applied for in the past five years for that firm. This measure represents the proportion of patents from the past five years that a firm has utilized in new product development in the current year (our results are robust to varying the patenting and patent-utilization windows). Employing this measure of corporate patent utilization rate from 2002 to 2022, we investigate the future performance implications along the following dimensions: new product development, product market performance, profit improvement, and firm valuation.

We start by examining the relationship between a firm's patent utilization rate and its future new product development. Since patent utilization reflects a firm's propensity to commercialize its intellectual assets, we expect it to positively influence future product development of the firm. Consistent with this expectation, we find that a higher patent utilization rate is associated with an increase in both the number of new products and product announcement returns. Moreover, firms with greater patent utilization are more likely to develop breakthrough products in the future. These relationships remain robust after controlling for a comprehensive set of firm-level characteristics, as well as industry or firm fixed effects and year fixed effects. In terms of economic significance, a one-standard-deviation increase in a firm's patent utilization rate corresponds, on average, to a 30% (22%) increase in the number of new (breakthrough) products and a 1.1-percentage-point increase in cumulative abnormal returns around new product announcements in the subsequent year.

Building on the finding that firms develop more and higher-quality new products when patent utilization is high, we further investigate whether higher patent utilization rates lead to improvements in firms' future product market performance. Our analyses confirm this hypothesis: corporate patent utilization rate is positively associated with firms' sales growth and market share growth over the following three years. On average, a one-standard-deviation increase in patent utilization corresponds to a 0.86 to 1.29 percentage-point increase in sales growth and a 0.94 to 1.29 percentage-point increase in market share growth.

In addition, we find that firms with higher patent utilization rate experience significant improvements in profitability and market valuation. A one-standard-deviation increase in patent utilization rate is associated with an increase in gross profit margin by 0.70 percentage points, return on assets by 0.27 percentage points, and market-to-book equity ratio by 1.33% in the subsequent year. These results are consistent with prior literature documenting positive relationships between innovation inputs/outputs and future firm performance and valuation (e.g., Lev and Sougiannis, 1996; Hall et al., 2001). Taken together, the baseline findings suggest positive indications between a firm's patent utilization rate and its future product development, product market performance, profitability improvement, and valuation.

With the positive implications of patent utilization on various dimensions of firm performance, a natural question arises: are these effects primarily driven by the utilization of high-value patents? To answer this question, we construct two measures to capture the utilization rates of high-value and low-value patents, based on the economic value of patents estimated by Kogan et al. (2017). Our findings suggest that the utilization of high-value patents is the primary driver of positive firm outcomes. Additionally, we explore the heterogeneous effects of corporate patent utilization based on firms' product market competition. The results indicate that the benefits of patent utilization are more pronounced in competitive product markets.

The documented positive relationships between patent utilization rate and future firm performance may be subject to endogeneity concerns. For instance, high-performing firms may have greater incentives to utilize their patents to secure product market shares and maintain their leading positions. Omitted variables may also be correlated with both patent utilization rate and firm performance. To alleviate these endogeneity concerns and strengthen the causal interpretation of our findings, we construct a Bartik shift-share instrument to isolate plausibly exogenous variation in corporate patent utilization rate (Bartik, 1991; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2025). Specifically, our shift-share research design relies on two components: the preexisting technology-class share of a firm's patent portfolio (*share*), and the differential economy-wide growth in patent utilization rate (exclude patents belonging to the focal firm and its product market rivals) across technology classes (*shift*) over time. Our instrument is then calculated as the inner product of the shift and share components.

The rationale behind this instrument is that firms differ in the technological composition of their patent portfolios. The preexisting patent class shares reflect a firm's ex-ante exposure to specific technological fields. When certain technologies experience commercialization advances, firms with larger initial shares in those fields are more positively affected. This approach ensures that the variation in the shift-share instrument is driven by the growth in utilization rate across patent technology classes that are isolated from the firm's (and its rivals') product markets as well as from potential unobserved firm-specific factors. As expected, the first-stage regression results show a positive and significant relationship between the instrument and corporate patent utilization rate. The second-stage results indicate that the instrumented patent utilization rate continues to positively impact future firm performance across various dimensions.

Finally, we conduct a comprehensive set of robustness checks for our findings by: i) controlling for a firm's past patent outputs or product similarity score (Hoberg and Phillips, 2016); ii) conducting intensive margin analyses by limiting the firm-year observations to those with at least one new product launch; iii) constructing alternative patent utilization measures using different percentile cutoffs, varying the patent portfolio window as well as the patent-product incorporation window, or applying a 3-year moving average of the original measure, iv) measuring firm performance using alternative outcome variables, and iv) further including industry-by-time fixed effects to account for the sectoral shocks on firms' commercialization tendency.

Our study contributes to the literature on innovation and the commercialization of intangible assets. Endogenous growth theories highlight the creative destruction process, emphasizing technological innovation as a key driver of long-term economic growth (e.g., Schumpeter, 1943; Grossman and Helpman, 1991; Aghion and Howitt, 1992). A growing body of empirical research has provided valuable insights into whether creative innovations shape future firm growth and economic prosperity, with most studies relying on patent-based measures due to their broad data availability (e.g., Bloom and Van Reenen, 2002; Kogan et al., 2017; Kelly et al., 2021; Ma, 2025). While patents grant exclusive legal rights to protect inventors' ideas, they represent only an intermediate stage of the innovation process. The ultimate goal for a firm is to commercialize patented technologies through new product offerings that generate economic returns. This paper bridges the gap between patent production and product commercialization, offering new insights into how effectively firms translate their patent portfolios into new products.[7] This measure also speaks to growing concerns about "patent portfolio races", where firms accumulate patents primarily to deter competitors rather than to pursue commercialization (Choi and Gerlach, 2017; Argente et al., 2020).

Second, our paper extends the literature that investigates the implications of corporate innovation strength on firm performance. Prior studies document positive relationships between traditional R&D expenditures and firm profitability and valuation (e.g., Sougiannis, 1994; Lev and Sougiannis, 1996; Chan et al., 2001). Recent studies have explored alternative measures of firms' intangible capabilities, such as innovation efficiency (Hirshleifer et al., 2013), innovation originality (Hirshleifer et al., 2018), research quotient (Cooper et al., 2022), technology differentiation (Arts et al., 2023), and technological obsolescence (Ma, 2025). We complement the literature by measuring a firm's innovation strength in technology commercialization. The successful utilization of patented technologies in new products hinge on factors such as patent quality, commercial viability of the technologies, and the resources or skills a firm required to reach the final commercialization stage. We investigate how such patent utilization capability

[7] A similar study by Masclans et al. (2025) measures the commercial potential of scientific findings.

influences a firm's new product development, market share, profitability, and firm valuation. Our approach and analyses provide robust empirical evidence on the broader benefits of effectively utilizing patents, highlighting the importance of not just obtaining patents but actively integrating them into the firm's product pipeline.

Third, this study contributes to the literature on textual analysis in economics and finance (e.g., Loughran and McDonald, 2011; Garcia and Norli, 2012; Gentzkow et al., 2019). Prior studies generally use "bag-of-words" approach to measure textual similarity (Hoberg and Phillips, 2016; Kelly et al., 2021; Argente et al., 2020; Bowen et al., 2023). An emerging literature starts to adopt machine learning techniques to account for word semantics (Mikolov et al., 2013a; Pennington et al., 2014; Bojanowski et al., 2017).[8] For instance, Li et al. (2021) apply the *Word2vec* model to measure corporate culture. Hoberg and Phillips (2025) use the *Doc2vec* model to compute firm product market scope based on a firm's exposure to different industries. Similarly, Kogan et al. (2022) employ a machine learning model, *Glove*, to capture workers' technology exposure by calculating textual similarity between occupation descriptions and patent filings. This study leverages the *FastText* model to link products with patents within each firm, and develop a novel measure of corporate patent utilization.

The remainder of the paper is organized as follows. Section 2 describes our approach to measuring the patent utilization rate. Section 3 reports the results from various validation tests on the patent utilization measure. Section 4 discusses the implications of corporate patent utilization on new product development, product market share, profitability improvement, and valuation. Section 5 explores the heterogeneity of the documented effects. Section 6 presents the results from the shift-share instrumental-variable analysis and various robustness tests. Section 7 concludes. Appendix A provides variable definitions and additional empirical results. Appendix B provides technical details of our approach in measuring patent utilization rate.

[8] Seegmiller et al. (2023) show that these machine learning approaches significantly outperform the conventional "bag-of-words" approach.

2 Patent Utilization: Data and Measurement

This section describes how we construct the measure of a firm's patent utilization rate. In Section 2.1, we describe the sources of data used in the study. In Section 2.2, we compute the textual similarity score between a patent filing and a product description text. We regard a patent as utilized in a new product in a firm if the text description of the patent filing is *abnormally similar* to that of the new product description. Finally, we aggregate the patent-level utilization to firm level. Appendix B contains more technical details on the measurement of patent utilization.

2.1 Data

We obtain patent filing text from PatentsView, which provides title, abstract, description, and claims sections for each patent granted since 1976. We then aggregate title, abstract, and description sections of each patent document into a patent-level corpus for textual analysis.[9] To match patents with the U.S. publicly listed firms, we rely on the linking table developed by Kogan et al. (2017), which matches each patent assignee with a PERMNO ID from CRSP if available. Hence, our final patent text sample consists of 2,544,432 patents generated by U.S. public firms from 1976 to 2022. Figure A2 illustrates an example of patent text filing from the Google Patents website.

We further collect product-related text description data from the Capital IQ Key Development database. After restricting the product-related text descriptions to the sample of U.S. publicly listed firms, we obtain 269,472 product-related announcements from 2002 to 2022. As suggested by Cao et al. (2018), there are generally four types of product-related announcements: R&D progress, new product introduction, product improvement, and product retirement. In line with prior literature, we focus specifically on the category of new product introduction. To select new product introduction-related announcements, we build upon Cao et al. (2018) by

[9] We exclude the patent claim section because prior studies suggest that it is largely shaped by legal professionals and contains highly stylized legal language (Bena et al., 2022; Bowen et al., 2023). Our results are qualitatively similar if we exploit the full text of patent filings.

using new-product-launch keywords and employing an advanced natural language processing technique, *FinBert*, to help us determine whether a product-related announcement is related to new product introduction. Please see Section B1 for detailed descriptions of the training sample construction, the *FinBert* fine-tuning process, and the model classification performance.[10] In Figure A3, we demonstrate an example of Apple Inc. announcing a new product in 2020.

After requiring firms to have at least one patent granted throughout their histories, our final sample consists of 125,329 announcements related to new product launches.[11] We follow standard text cleaning procedures (e.g., Kelly et al., 2021; Kogan et al., 2022) to preprocess the patent documents and new product announcement text description, which are discussed step-by-step in Section B2. Finally, we obtain stock return data from the Center for Research in Security Prices (CRSP), financial data from Compustat, and corporate patent quantity and quality data from Kogan et al. (2017). Table 1 reports the summary statistics of the variables used in this study. Table A1 in Appendix A provides detailed variable definitions and data sources.

[Please insert Table 1 about here]

2.2 Measuring Patent-Product Pair Textual Similarity

We assume that a patent is utilized in a new product if the patent-product pair textual similarity is abnormally high.[12] This critical assumption is similar in spirit to the innovation literature that leverages pair-wise patent citations to investigate knowledge diffusion across firms (see., e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Singh and Marx, 2013; Arora et al., 2021; Fadeev, 2023). In essence, the literature hypothesizes that if a patent of Firm **A** cites a

[10] Panel A of Table A2 lists the new product launches keywords. Panel B further tabulates the classification performance in the testing sample. Our fine-tuned *FinBert* model can accurately classify 93% of the headlines. Panel C illustrates some (randomly) selected examples of new-product-introduction-related and non-related headlines predicted by our *FinBert* model.

[11] We also require our sample firms to have at least one new product launch in the key development database. Thus, our final sample contains 3,102 unique firms that have produced patents and launched products.

[12] Verifying whether a patented technology is utilized in a product poses a significant challenge as it requires consultations with technical experts. We acknowledge that high similarity may not indicate definite patent utilization in the new product. However, it does suggest that the new product is very likely to have been heavily influenced by or derived from the patented technology.

patent of Firm **B**, knowledge is diffused from Firm **B** to Firm **A**.[13] While the current data on new products does not specify information on patent utilization, we can infer the relationship between patents and products through their textual similarity.

On this basis, our first step is to measure textual similarity between patents and products. A conventional way to measure textual similarity in economics and finance literature is the "bag-of-words" approach (Hoberg and Phillips, 2016; Gentzkow et al., 2019; Kelly et al., 2021; Chen and Srinivasan, 2023). However, it does not account for semantic similarities between words. That is, words could possess similar meanings even if they are in different forms. For instance, the word "big" is semantically similar to the word "large," but the "bag-of-words" approach will count as a zero match.[14]

Importantly, the underestimation bias could be even more pronounced when comparing two documents from different text sources that exhibit diverse language styles (Seegmiller et al., 2023). In this study, we aim to compare the formal, standardized, and legalistic language used in patent filing text descriptions with the more informal and less structured tone in product announcement text descriptions. If we adopt the "bag-of-words" approach, the contrasting language styles of the two corpora could lead to sparse one-hot vectors, with many elements equal to zero and cosine similarity scores close to zero.

To overcome the issue, we exploit a machine learning technique, *Word2vec* (Mikolov et al., 2013a), that can transform words into semantic, low-dimension, and dense vectors (embeddings) via neural network. Hence, words with similar semantic meanings can have close spatial distance even if they are not exactly overlapped. We obtain pre-trained word embeddings from *FastText*, an extension of the *Word2vec* model developed by Bojanowski et al. (2017). In the following paragraphs, we briefly discuss how we use the *FastText* model to measure cosine similarities

[13] In a similar vein, Cohen et al. (2023) regard a firm as a user of an external patent if the firm has cited the patent previously.

[14] Consider an extreme case: document i contains the phrase "one beautiful house", while document j contains the phrase "a lovely dwelling". As humans, we can discern the closeness of the two documents. However, when using the "bag-of-words" approach, we transform the two documents into two one-hot vectors, $\mathrm{V_i}$ =[1, 1, 1, 0, 0, 0] and $\mathrm{V_j}$ = [0, 0, 0, 1, 1, 1]. We then compute the cosine similarity between the two vectors. In this example, we obtain a cosine similarity score of zero, which indicates that the two documents are unrelated. Please refer to Section B3.1 for more details on the challenges in "bag-of-words" approach.

between patents and product texts. Sections B3.2 to B3.4 contain more details.[15]

First, we aggregate *FastText* word vectors to document (i.e., patent/product text) level using the following equation:

$$D_i = \sum_{X_j \in Z_i} w_{i,j} x_j \tag{1}$$

where D is a vector for document i, measured as the weighted average of the word vectors x for each word j in the set of words Z in document i. Following prior textual analysis literature (see, e.g., Loughran and McDonald, 2011; Hoberg and Phillips, 2016; Li et al., 2021; Kelly et al., 2021), we use the term-frequency-inverse-document-frequency (TFIDF) as our weighting scheme to give different weights w on word vectors based on the importance of the words in our corpus.

After obtaining a dense semantic vector for each document, we use the following equation to measure the cosine similarity between a patent document vector D_p and a product description text D_t within a firm f:

$$Sim_{p,t,f} = \frac{D_{p,f}}{||D_{p,f}||} \cdot \frac{D_{t,f}}{||D_{t,f}||} \tag{2}$$

Equation 2 emphasizes within-firm patent-product pair similarity because we want to measure a firm's self-invented patent utilization in its new product development. It is worth noting that we solely focus on non-process (product) patents (Bena and Simintzi, 2025), as our goal is to understand whether product patents are utilized in new product development.[16] Moreover, for a firm's self-invented patents, we only focus on the firm's five-year patent application (later granted) portfolio before the launching date of a new product, since patents may become obsolete as other technologies evolve (Ma, 2025).[17] The calculation process of within-firm

[15] Please refer to Section B3.2 for technical details and advancement on the *Word2vec* model, Section B3.3 for information on the *FastText* model, and Section B3.4 for thorough description on the measurement of patent-product pair textual similarity using *FastText*.

[16] Process-related innovations are of less interest in our study, as these patents primarily focus on improving production processes. For technical details on how to distinguish product innovations from process innovation, please see Section B3.5

[17] The United States Patent and Trademark Office (USPTO) requires that for patent applications filed after June 8, 1995, the terms of patents will end 20 years after the patent application date. In robustness tests, we also consider the 10-year patent application (later granted) portfolio of a firm and obtain qualitatively similar

patent-product pair similarity is illustrated in Figure 1. Suppose that Firm **A** launched two products in 2015, NP1 and NP2. We then source the three product patents (PAT1, PAT2, and PAT3) that Firm **A** applied (and later granted) in the five years before 2015. For each patent-product pair, we compute its text similarity score using Equation 2.[18]

Next, since a majority of patent–product pairs within a firm have low textual similarity scores and are considered unrelated to one another, we follow the prior literature (e.g., Kogan et al., 2022; Hoberg and Phillips, 2016) to impose a stringent criteria: we only regard a patent as being utilized in a product if the textual similarity score is above 80th percentile of our sample patent-product pair scores.[19] In other words, for each within-firm patent-product pair, we replace the pair score with one if the raw similarity score is above 80th percentile, and otherwise replace it with zero.

In Panel A of Table A3, we demonstrate some examples of within-firm patent-product pair linkage. For each of the three randomly selected products, we show the 5 most (least) similar patents based on the patent-product similarity score. In Panel B of Table A3, we further provide excerpts from the text descriptions of the three new products, along with excerpts from the most and least similar patents for each product. These matching examples illustrate the effectiveness of the *FastText* model. For instance, the patent titled "Multi-functional hand-held device" filed in 2006 by Apple Inc. is most closely associated with the product "Apple IPhone 4," as their texts are semantically similar. In contrast, the patent titled "Transaction ID filtering for buffered programmed input/output (PIO) write acknowledgments" filed in 2009 by Apple Inc. is deemed as the least similar as its technical terms differ fundamentally from the iPhone 4 product description.

[Please insert Figure 2 about here]

Finally, having identified whether a patent is utilized by a firm, we can then measure a

results.

[18] For robustness check, we randomly select 250 patent-product pairs from our sample and use OpenAI's text embedding model (*text-embedding-3-small*) and Glove's word embedding model (Pennington et al., 2014) to compute their cosine similarity scores. Figure A4 illustrates the correlations between the similarity scores generated by *FastText* and *OpenAI* (correlation = 0.70), and by *FastText* and *Glove* (correlation = 0.71). The results suggest that the three embedding models produce comparable similarity measures.

[19] We also consider alternative percentile cutoffs such as 70th and 90th, and obtain qualitatively similar results.

firm's patent utilization rate, *Pat. Utilization Rate*, as the number of granted patents applied for by a firm in the past five years and utilized in the new products launched by the same firm in the current year, scaled by the total number of granted patents applied for by that firm in the past five years. We replace *Pat. Utilization Rate* with zero if a firm does not launch any new product for a firm-year, but has applied for (and was later granted) at least one patent in the past five years. Alternatively, if a firm does not have any successful patent applications in the past five years, we set *Pat. Utilization Rate* as missing.

[Please insert Figure 3 about here]

Table 1 shows that the average corporate patent utilization rate is 31.3%, which is analogous to prior literature that surveys inventors to analyze the commercialization outcomes of inventive activity. For instance, using survey data on 3,162 patented inventions, Webster and Jensen (2011) find that around 40% are advanced to subsequent new product launches and production. Similarly, Amesse et al. (1991) document that around 43% of patents are commercialized in Canada. Figure 2 further illustrates the variation of the average corporate patent utilization rate over time, which is fluctuated at around 30% over time and a slight decrease is observed since 2013, indicating a potential increase in defensive patenting (which would decrease utilization). Figure 3 further illustrates the top 10 industries (2-digit SIC) with the highest rates of corporate patent utilization. It shows that five out of the 10 industries are related to the manufacturing sector, with *36: Electronic & Other Electric Equipment* ranked the highest.

In Table A4, we further examine the determinants of firms' patent utilization rate. We regress *Pat. Utilization Rate* in the next year, the next two years, and the next three years on recent patent litigation activity and various firm characteristics,[20] controlling for firm and year fixed effects. We find that the number of patent cases in which the firm appears as a *plaintiff* over the prior three years is positively related to subsequent patent utilization, whereas the number of cases in which the firm is a *defendant* is negatively related. These findings are potentially consistent with two mechanisms. First, plaintiff status typically reflects active

[20] Patent litigation data is obtained from the USPTO Patent Litigation Docket Reports dataset (see https://www.uspto.gov/ip-policy/economic-research/research-datasets/patent-litigation-docket-reports-data).

enforcement of commercially relevant inventions: firms litigate to protect technologies that are already deployed or about to be embedded in products, which aligns with higher follow-on utilization.[21] Second, defendant status raises legal risk and organizational caution: infringement exposure imposes direct costs on firms and creates uncertainty around freedom-to-operate, encouraging defensive patent portfolio building (e.g., insulating claims) rather than the embedding of patents into product development, which is associated with lower future utilization. Moreover, firms facing greater product market competition (lower *TNIC HHI*), larger firms, younger firms, and firms with lower leverage, tend to exhibit higher future patent utilization. Greater product market competition and younger firm age heighten incentives to differentiate and grow, increasing the payoff to converting patents into products. By contrast, larger size and lower leverage proxy for commercialization capacity (i.e., more resources and complementary assets, and fewer debt-overhang or financing constraints), enabling higher subsequent utilization.

3 Validation

In this section, we validate the patent-utilization metric in two steps. First, at the patent level, we test whether patents flagged as utilized in new products align more closely with a firm's core technology fields, attract more self-citations, exhibit a higher self-citation ratio, and are less likely to be sold. Second, at the product level, we aggregate utilization to the focal product and examine whether patent-backed product launches are associated with higher quality, as indicated by stronger market reactions and a greater likelihood of being breakthrough products.

[21] Our sample firms are unlikely to be patent trolls described in Cohen et al. (2019) because we require that each sample firm must have launched at least one new product during the sample period. This requirement ensures that our firms are practicing entities engaged in commercialization rather than entities whose primary business model is patent assertion and litigation.

3.1 Utilized Patents, Core Technology fields, Self Citations, and Patent Sales

We first examine whether utilized patents fall within a firm's core technology field. Similar to scientific researchers who specialize in particular domains, firms also develop technological expertise within particular fields. When introducing new products, they are more likely to commercialize technologies in areas where they have established strong scientific foundations and accumulated deep market insights. Accordingly, we hypothesize that patents utilized in new products are more likely to be related to firms' core technology fields.

To identify a firm's core technology field, we use its historical distribution of patent classes.[22] We regard a patent class as the firm's core technology field if the firm has produced the largest number of patents in that class over the past 10 years. We then formally investigate whether utilized patents are more likely to belong to firms' core patent classes with the following patent-level regression specification:

$$Y_{p,f,c,t} = \beta_1 I\ (Utilized)_{p,f,c,t} + \theta_f + \delta_c + \mu_t + \epsilon_{p,f,c,t}. \tag{3}$$

In this regression, $Y_{p,f,c,t}$ is an indicator that equals one if firm f's patent p in technology class c filed in year t belongs to the firm's core technology field, otherwise equals zero. $I\ (Utilized_{p,f,c,t})$ is also an indicator that equals one if patent p is utilized in firm f's new products, based on our patent utilization measure described in Section 2.2. We further include firm fixed effects θ_f, patent class fixed effects δ_c, and year fixed effects μ_t. Panel A of Table 2 reports the results.

[Please insert Table 2 about here]

Column 1 reports estimates from the specification outlined in Equation 3, which includes firm, technology class, and year fixed effects. In Column 2, we add class-by-year fixed effects to account for national technology advancements. In Column 3, we further include firm-by-year fixed effects to absorb time-varying firm characteristics.[23] Across all specifications, the

[22] Patent class is defined as the 3-digit Cooperative Patent Classification (CPC) code.

[23] Unlike the analyses in Panel B of Table 2, we are not able to include firm-by-class fixed effects in this

coefficient estimates on *I (Utilized)* are positive and statistically significant at the 1% level. In terms of economic magnitude, our preferred specification (Column 3) suggests that, for patents applied by the same firm in the same year, those utilized in the firm's new products are on average 3.2% more likely to belong to the firm's core technology field than the unutilized patents.

Hall et al. (2001) suggest that self-citations (instances where a firm cites its own prior patents) reflect the cumulative nature of innovation and the increasing returns associated with knowledge accumulation. We therefore expect commercialized (utilized) patents to receive more self-citations and to exhibit a higher self-citation share, as subsequent generations of products build on these foundational inventions. Consistent with this logic, utilized patents should also be less likely to be sold, since they represent core technological building blocks for the firm.

To formally test this hypothesis, we reestimate Equation 3 using alternative dependent variables. Specifically, we replace the outcome with *# Self Citations*, denoted as the number of self citations received by a patent, and *Self Citation Ratio*, which is calculated as the number of self citations divided by the total citations received by a patent. To identify patent transactions, we leverage the raw patent assignment database from the USPTO and follow the methodology proposed by Brav et al. (2018). We then generate an indicator variable, *I (Sold)*, which equals one if a patent is sold to another entity by the original assignor within fours year after issuance, following Figueroa and Serrano (2019).[24] The results are reported in Panel B of Table 2.

In Columns 1, 4, and 7, we include class-by-year fixed effects to absorb technology shocks at the CPC–year level and firm-by-class fixed effects to account for the possibility that some

test because a firm's core technology field is relative stable over time. Including firm-by-class fixed effects will essentially absorb most of the variation.

[24] We capture patent transaction activities within four years after patent issuance due to two considerations. First, consistent with Figueroa and Serrano (2019), the majority (over 65%) of patent sales in our sample occur within the first few years after issuance and prior to the first patent maintenance date, which is four years after grant date. Thus, focusing on the 4-year window captures the majority of first-time patent reassignments. Second, commercialized patents are unlikely to be sold by firms in the short run, as they are strategically important to firms' ongoing product development. However, because these technologies are commercially viable, they are likely to be treated more favorably by potential buyers in the market. Therefore, in the longer run, firms may find them financially attractive to sell these patents once the underlying technologies become obsolete to their future innovation activities. Nonetheless, we conduct robustness checks of patent sales activities without limiting to the 4-year window. We find qualitatively similar, albeit weaker, results that utilized patents are less likely to be sold by firms.

firms are particular productive in specific technology fields. Columns 2, 5, and 8 further add firm-by-year fixed effects to control for time-varying firm characteristics. Columns 3, 6, and 9 include firm-by-class-by-year fixed effects, our most stringent specification. Consistent with expectations, utilized patents are positively related to the number of self-citations and to the self-citation ratio, and negatively related to the likelihood of sale. The results from the most stringent model suggests that, for patents filed in the same year, within the same technology class, and by the same firm, those that are utilized in the firm's new products on average receive 8.55% (i.e., exp(0.082)-1) more self-citations, exhibit 0.6% higher self-citation ratio, and are 0.1% less likely to be sold to other entities within four years after issuance.

3.2 New Product Quality

Next, we turn to the product-level analyses and investigate whether those new products supported by a greater number of patents would exhibit higher quality.

As discussed in Section 2.2, we use *FastText* model to calculate textual similarity scores for each patent-product pair within a firm, and we regard a patent as utilized in a new product if the patent-product pair text similarity score is above 80th percentile. Hence, to measure the number of patents utilized in each new product, we aggregate the within-firm patent-product pair scores (either one or zero) to product announcement event level.[25] To capture product quality, we follow the prior literature (Kogan et al., 2017; Mukherjee et al., 2017) and use a product's economic value, measured by the cumulative abnormal stock return (*CAR (-1, 1)*) over the three-day window surrounding the new product announcement. Specifically, we estimate the following regression equation:

$$Y_{i,f,t} = \beta_1 Log\ (1 + \#Patents\ Utilized_{i,f,t}) + \beta_2 Controls_{i,f,q-1} + \theta_f + \mu_t + \epsilon_{i,f,t} \qquad (4)$$

In Equation 3, Y represents the three-day CAR of product i of firm f in the product announcement event date t, and *Log (1+#Patents Utilized)* is the natural logarithm of one

[25] For each product announcement event, a firm may launch a single product or a bundle of products.

plus the number patents utilized by the product i of firm f in event date t. Because a firm may launch multiple products on single event date, we generate two variables to account for the situation of product bundle launching: *Log (1+#Patents Utilized*^{Sum}*)*, which is the natural logarithm of one plus the sum of the number of patents that are utilized in new product(s) for an event date, and *Log (1+#Patents Utilized*^{Average}*)*, which is the natural logarithm of one plus the average number of patents that are utilized in new product(s) for an event date. We further control for a variety of firm characteristics, such as firm size (*Firm Size*), firm age (*Log(Firm Age)*), leverage ratio (*Leverage*), research and development expenses (*R&D*), return on assets (*ROA*), cash holdings (*Cash*), Tobin's Q (*Log(Tobin's Q)*), sales growth (*Sales growth*), and past stock return (*Stock Return*), all measured one quarter before the product announcement quarter. Moreover, we control for the length of the product announcement text (*Log(Product Text Length)*) and the number of new products (*Log(1+#New Products Launched)*) that have already been launched by firm f in the same year before the event date. Finally, we include firm fixed effects θ and event-year-week fixed effects μ. The results are reported in Table 3.

[Please insert Table 3 about here]

In Columns 1 and 3, we include industry fixed effects to account for time-invariant industry characteristics, and event-year-week fixed effects to control for time-varying economic conditions. In Columns 2 and 4, we replace industry fixed effects with firm fixed effects to absorb time-invariant firm heterogeneity. We find that, across all specifications, the coefficient estimates of *Log(1+#Patents Utilized*^{Sum}*)* (*Log(1+#Patents Utilized*^{Average}*)*) are positive and statistically significant at the 5% level or better, indicating that new products integrating a larger number of patents are valued more by the stock market. The economic magnitude is also meaningful. For instance, the estimate in Column 2 implies that a 1-percentage-point increase in *Log(1+#Patents Utilized*^{Sum}*)* is associated with a 0.020% (= 0.01*0.020) increase in the product announcement return.[26]

Overall, these results are consistent with the expectation that more innovative products (those supported by a richer base of utilized patents) command higher market valuations and

[26] Note that the average product announcement return is 0.190%.

also provide an objective measure for a product's degree of innovation.

3.3 New Product Breakthrough Index

In addition to examining the economic value of new products, we investigate whether products that incorporate a greater number of patents are more likely to be breakthrough products. Following Kelly et al. (2021), we define breakthrough products as those that not only introduce novel features but also shape the direction of subsequent product developments. Building on this idea, we construct a text-based breakthrough index that captures a product's novelty and impact, providing a fresh perspective on what sets truly transformative products apart.

As stated above, a novel product is defined as one that is distinct from prior products. We follow Kelly et al. (2021) to measure a product's novelty as the inverse of its textual similarity with the prior products, where textual similarity is as follows:

$$BS^5{}_{\mathrm{j}} = \sum_{i \in \beta^5_{j,m}} \rho_{j,i} \tag{5}$$

where BS denotes the backward similarity of product j. $\rho_{j,i}$ is the pairwise similarity between product j and i. $\beta^5_{j,m}$ denotes the set of previous products that are launched in the 5 years before product j's offering and that are in the same product market m as product j.[27] Intuitively, novel products should have low backward similarity (BS) with the prior products.

On the other hand, an impactful product should shape future innovations, exhibiting high similarity with subsequent products. Thus, we measure a product's impact as follows:

$$FS^5{}_{\mathrm{j}} = \sum_{i \in \alpha^5_{j,m}} \rho_{j,i} \tag{6}$$

Similarly, FS denotes the forward similarity of product j. $\rho_{j,i}$ is the pairwise similarity between product j and i, and $\alpha^5_{j,m}$ denotes the set of future products that are launched in the 5 years after product j's offering and that are in the same product market m as product j. Thus, an

[27] The product market is based on parent firms' TNIC3 classification (Hoberg and Phillips, 2016).

influential product will have high similarity (FS) with future innovations.

Finally, the product breakthrough index, *BreakthroughIndex*, which reflects the novelty (inverse backward similarity BS) and impact (forward similarity FS) of new products, is measured as:

$$BreakthroughIndex^{5}{}_{\mathrm{j}} = \frac{FS^{5}{}_{\mathrm{j}}}{BS^{5}{}_{\mathrm{j}}} \tag{7}$$

The breakthrough index of a product tends to be higher if it exhibits low backward similarity with prior products (which is novel) but high forward similarity with subsequent products (which is impactful). To account for potential time-varying factors (such as fluctuations in the number of new products launched every year and changes in language over time), we follow Kelly et al. (2021) to adjust the breakthrough index by removing year fixed effects. We present the results on the relation between the number of patents utilized in new products and the breakthrough index in Table 4.

[Please insert Table 4 about here]

In Columns 1 and 2, we observe that the coefficients of *Log(1+#Patents Utilized)* are positively and significantly related to the *Breakthrough Index* at least at the 5% significance level. These results consistently show that new products supported by a greater number of patents tend to have higher breakthrough indices, suggesting that they are more novel and impactful. Furthermore, we examine whether these patent-embedded new innovations are more likely to be breakthrough products, defined by an indicator variable, *1 (Breakthrough Product)*, that equals one if the *Breakthrough Index* is above 95th percentile and zero otherwise (Kelly et al., 2021). The results presented in Columns 3 and 4 of Table 4 align with our expectation: New products with more patents are more likely to be breakthroughs.

In sum, the patent-level evidence in this section shows that utilized patents are more likely to lie in a firm's core technology fields, attract more self-citations, exhibit a higher self-citation share, and are less likely to be sold. These findings suggest that commercialized patents serve as foundational assets for subsequent innovation. At the product level, new product launches supported by more utilized patents earn higher announcement returns and are more likely to be breakthrough products. Taken together, these validations indicate that our text-based measure captures economically meaningful patent utilization in new products.

4 The Implications of Corporate Patent Utilization

Since Schumpeter introduced the concept of creative destruction, economists have developed various endogenous growth models demonstrating that technological innovation is a central driver of economic growth and firm success (e.g., Aghion and Howitt, 1992; Lentz and Mortensen, 2008; Akcigit and Kerr, 2018). Empirical research consistently shows that innovation capabilities are a key determinant of future firm performance (e.g., Hall et al., 2001; Hirshleifer et al., 2013; Kogan et al., 2017; Cooper et al., 2022). However, recent trends reveal that many firms increasingly file patents as a strategic tool to block competitors rather than to drive genuine innovation. This practice risks stifling research productivity and impeding technological advancement, with potentially adverse effects on long-term economic growth (Bloom et al., 2020; Kalyani, 2022). Over time, unused patents may become obsolete, offering little practical application in new product development (Ma, 2025).

In contrast to simply tracking patent filings, out study introduces a new metric of innovation strength: the proportion of patents that are incorporated into new products. By focusing on patent utilization, we offer a metric of how well a firm translates its innovation efforts into tangible product market outcomes. This approach captures the extent to which a firm's patents contribute to new product development, offering fresh insights into the role of patents in sustaining competitive advantage and driving growth. In this section, we aggregate patent-product pair scores at the firm level to generate a measure of corporate patent utilization rate. This measure effectively captures the extent to which a firm's past patent portfolio is incorporated into new products. We then explore the implications of patent utilization for a firm's future performance, focusing on four key dimensions: new product development, product market performance, profit improvement, and firm value.

4.1 Corporate Patent Utilization and New Product Development

First, we shed light on the association between a firm's patent utilization rate and its future new product development. Since patent utilization rate indicates a firm's proficiency in commercializing patents, we anticipate that firms with higher rates of patent utilization will produce more (and higher-quality) new products. To investigate this research question, we employ the fixed-effects Poisson regression model for count-dependent and/or highly-skewed variables following the literature (Cohn et al., 2022; Chen and Roth, 2024). The model is as follows:

$$Y_{f,j,t+1} = \beta_1 Pat.\ Utilization\ Rate_{f,j,t} + \beta_2 Controls_{f,j,t} + \theta_j + \mu_t + \epsilon_{f,j,t} \tag{8}$$

The dependent variable Y represents the new products development of firm f in industry j in year *t+1*. To evaluate a firm's new product development, we follow Mukherjee et al. (2017) to concentrate on a firm's number of new products launched, which is measured as the raw number of new product announcements (*#New Products*) a firm releases in a year.[28] To capture the quality of new products, we follow Kogan et al. (2017) and Mukherjee et al. (2017) to estimate the economic value of a new product using the cumulative abnormal returns around the product announcement. We then generate the variable, *Sum CARs*, which is calculated as the sum of all positive three-day cumulative abnormal stock returns of the new products a firm launches in a year. Moreover, we create another outcome variable, *#Breakthrough Products*, that measures the number of breakthrough products a firm develops in a year.[29]

The independent variable, *Pat.Utilization Rate*, represents the patent utilization rate of firm f in industry j in year t. We also include a variety of standard firm-level controls: firm

[28] Consistent with Mukherjee et al. (2017), we primarily focus on major new product introductions by firms. Specifically, we count only new products with cumulative abnormal returns (CARs) above the 80th percentile in a given calendar year. To test the robustness of our results, Table A4 examines highly valued new products, defined as those with CARs above the 95th percentile. We construct two outcome variables: *#Highly-Valued New Products*, which represents the number of highly valued new products a firm develops in a year, and *1(Highly-Valued New Products)*, a binary indicator that equals one if a firm launches at least one highly valued new product in a year and zero otherwise. Our findings remain consistent, showing that firms with higher patent utilization rates tend to develop more highly valued new products and are more likely to introduce such products in the following year. Additionally, we confirm that our results remain qualitatively similar even when considering all new product launches.

[29] Breakthrough products are defined as those products with breakthrough indices above the 95th percentile.

size (*Firm Size*), firm age (*Log(Firm Age)*), leverage ratio (*Leverage*), stocks of research and development expenses(*R&D Stock*), cash holdings (*Cash*), and past stock return (*Past Stock Return*), all measured in year *t*. Additionally, we control for a firm's new product intensity (*#New Products/Sales*), as the patent utilization rate may be positively correlated with the number of new products launched in the same year.[30] Lastly, we include industry or firm fixed effects (θ) and year fixed effects (μ) to account for time-invariant industry or firm heterogeneity and time-varying economic factors. The results are presented in Table 5.

[Please insert Table 5 about here]

In Columns 1 and 2, we investigate the relationship between a firm's patent utilization rate and its one-year-ahead raw number of new products (*#New Products*); in Columns 3-4, we further shed light on how a firm's patent utilization rate is related to its one-year ahead new product CARs (*Sum CARs*); in Columns 5-6, we explore whether firms with higher patent utilization rate will develop more breakthrough products (*#Breakthrough Products*). Columns 1, 3, and 5 include industry and year fixed effects, while Columns 2, 4, and 6 replace industry fixed effects with firm fixed effects.

Consistent with expectation, the fixed effects Poisson regression results in Table 5 show that *Pat.Utilization Rate* is positive and significantly (all at the 1% level) associated with the firm's new product quantity and quality in the subsequent year. The economic magnitude is meaningful. Take our preferred specifications (Columns 2, 4, and 6) as examples: A one-standard-deviation increase in patent utilization rate is on average related to a 30.17% (i.e., exp(0.390*0.676)-1) increase in its one-year-ahead number of new products, a 21.34% (i.e., exp(0.390*0.496)-1) increase in subsequent year's cumulative abnormal returns of new products, and a 21.82% (i.e., exp(0.390*0.506)-1) increase in the number of breakthrough products in the subsequent year.[31]

[30] To further address the concern that the relation between patent utilization and future firm performance may depend on whether firms launch new products in a given year, we also conduct intensive margin analyses as a robustness check by restricting the firm-year observations to those with at least one new product launch. Panel B of Table 11 shows that the results remain qualitatively similar.

[31] Given that the sample mean of *Sum CARs* is 5.291 percentage points, a one-standard-deviation increase in patent utilization rate is on average related to an increase in cumulative abnormal returns of new products by 1.13 percentage points.

Overall, the results suggest that a higher patent utilization rate is linked to both a greater quantity and higher quality of future new products.

4.2 Corporate Patent Utilization and Product Market Performance

Building on the previous subsection, which shows that a higher corporate patent utilization rate predicts more numerous and higher-quality product launches, we now examine firm performance implications. Endogenous growth theory posits that performance improves when firms introduce more innovative products. Guided by this logic and our prior evidence, this subsection investigates how short- to medium-term product-market performance evolves with a firm's patent utilization rate.

We estimate variants of Equation 7 using Ordinary Least Squares (OLS) instead of Poisson to study product-market outcomes. To capture firms' product market performance, we follow prior literature (e.g., Campello, 2006; Fresard, 2010; Billett et al., 2017) and focus on two measures: sales growth (*Sales Growth*), defined as the natural logarithm of total sales of a firm in the current year minus that of the previous year; market share growth (*MSG(SIC4)*), defined as the sales growth of a firm in a year minus the industry (4-digit SIC) median sales growth in the same year. The results are reported in Table 6.

[Please insert Table 6 about here]

Panel A reports sales-growth results; Panel B reports market-share growth. In both panels, we find that the coefficient estimates on *Pat.Utilization Rate* are positive and statistically significant at least at the 5% level, indicating that patent utilization rate is associated with better product market performance in the next three years. The results are robust to controlling for industry and year fixed effects (Columns 1, 3, and 5 in each panel) or firm and year fixed effects (Columns 2, 4, and 6 in each panel). In terms of economic magnitude, the specifications with firm and year fixed-effects (even-numbered columns) show that a one-standard-deviation increase in *Pat.Utilization Rate*, on average, corresponds to a 0.858 to 1.287 percentage-point increase (i.e., 0.390*0.022 to 0.390*0.033) in sales growth, and a 0.936 to 1.287 percentage-point increase (i.e., 0.390*0.024 to 0.390*0.033) in market share growth in the subsequent three years.

We also control for a firm's new-product intensity (*#New Products/Sales*) in all specifications. As expected, new-product intensity is positively related to both future sales growth and market-share growth, though the relationship is statistically significant only in year *t+2*. These findings suggest that the *Pat.Utilization Rate* captures a capability distinct from mere launch intensity and exhibits more persistent predictive power for product-market performance.

Overall, the evidence indicates that firms with higher patent utilization experience stronger subsequent product-market outcomes, consistent with endogenous-growth theory in which innovation (here, the effective embedding of patents into new products) drives firm growth.

4.3 Corporate Patent Utilization, Profit Improvement, and Firm Value

Having documented that a higher patent utilization rate predicts richer new-product development and stronger product-market performance, we next examine its implications for firm profitability and valuation. A large strand of literature has investigated the implications of innovation inputs and/or outputs on the financial performance and market valuations of firms. For example, regarding the innovation inputs, prior studies show that R&D expenditures and R&D intensity can positively predict future firm values and stock returns (see, e.g., Sougiannis, 1994; Lev and Sougiannis, 1996; Chan et al., 2001). On the innovation output side, several studies find that firms that generate more patent citations and patent economic values are associated with higher market valuations (see, e.g., Hall et al., 2001; Kogan et al., 2017). Recent studies also explore alternative measures of firms' intangible capabilities, such as innovation efficiency (Hirshleifer et al., 2013) and research quotient (Cooper et al., 2022). Both measures imply a positive relationship between firms' innovation strengths and future firm valuations.

Our measure of corporate patent utilization can be regarded as a firm's ability to transform its patent portfolio into new product development. Therefore, a higher patent utilization rate may be favorably valued by investors as it signals a stronger capacity for patent commercialization, thereby increasing the economic value of the firm's intangible assets. In this regard, we hypothesize that patent utilization rate of a firm is positively associated with its future firm valuations. Moreover, based on our previous findings that firms with higher patent utilization rate tend to introduce a greater number of higher-quality new products and achieve stronger product market performance, we expect these firms to experience greater subsequent profitability improvements.

We capture profitability changes using the changes in gross profit margin, ΔGPM, and in return on assets, ΔROA. Following prior literature, we measure firm value based on the log market-to-book ratio (*Log(MTB)*). As in the prior subsection, we estimate OLS versions of Equation 7 and report results in Table 7.

[Please insert Table 7 about here]

We find that a higher *Pat.Utilization Rate* predicts positive and statistically significant one-year-ahead profitability changes and firm valuation. The results are robust to industry and year fixed effects or firm and year fixed effects. A one-standard-deviation increase in *Pat.Utilization Rate* is on average associated with an increase in gross profit margin by 0.702 percentage point (= 0.390*0.018), return on assets by 0.273 percentage point (= 0.390*0.007), and 1.33% (i.e., exp(0.390*0.034)-1) increase in market-to-book ratio. Interestingly, we do not find consistent and significant relationships between a firm's new product launch intensity and its profitability improvement and market valuation in the subsequent year.

To summarize, the results in this section show that higher patent utilization rate of a firm is associated with more and higher-quality new products, stronger product market performance, greater profitability improvement, and higher firm value in the future. Together, these findings suggest that corporate patent utilization has positive value implications on future firm performance.

5 Heterogeneity of Patent Utilization Rate

In this section, we further explore the heterogeneous effects of corporate patent utilization on future firm performance. We first shed light on the economic value of patents utilized in firms' new products. We next investigate the role of product market competition.

5.1 High-value versus Low-value Patent Utilization

Patent quality varies in both scientific and economic dimensions (Hall et al., 2001; Kogan et al., 2017). While scientifically advanced patents can attract future forward citations and generate positive knowledge externalities, economically significant patents should be the ones that most impact future firm performance. Therefore, we hypothesize that the documented positive implications of patent utilization on firms' future new product development, product market performance, profit improvement, and valuation would be primarily driven by the use of patents that possess significant economic value.

We conduct the analyses by acquiring patent economic value data from Kogan et al. (2017), which assesses the economic significance of each innovation by analyzing the stock market reaction around the patent grant date. To distinguish whether patents are economically meaningful, we regard a patent as a high-value (low-value) patent if its economic value is above (below) the median across the sample patent economic value. We then construct two separate measures: *High-Value (Low-Value) Pat.Utilization Rate*, which are defined as the number of granted high-value (low-value) patents that are applied for in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, scaled by the total number of granted patents applied for in the past five years for that firm. We standardize both variables to unit standard deviation for ease of interpretation. We then estimate Poisson or OLS versions of Equation 7 to compare the predictive power of high- versus low-value utilization for future firm outcomes. Results are reported in Table 8.

[Please insert Table 8 about here]

Consistent with expectations, the utilization of economically high-value patents primarily drives the results. Although utilizing low-value patents is associated with some subsequent product development (Panel A), its effect is economically and statistically smaller than that of high-value utilization. More importantly, the effects of low-value utilization on product-market performance (Panel B) and on profitability and valuation (Panel C) are negative or statistically indistinguishable from zero. By contrast, high-value utilization exhibits pronounced and robust positive effects across all outcomes. Taken together, these findings indicate that investors effectively differentiate between high- and low-value patents at the point of commercialization, and that the performance gains we document are concentrated in the utilization of economically valuable patents.

5.2 Product Market Competition

Next, we investigate the role of product market competition. In competitive product markets, firms often struggle to differentiate their products from those of competitors. Therefore, corporate innovation and patent utilization are crucial for these firms to survive intense competition and gain market share. Thus, we hypothesize that the positive effects of patent utilization on future firm performance are stronger for firms operating in more competitive product markets.

We capture a firm's product market competition using two alternative proxies: Lerner index, which is measured as the median gross profit margin in the firm's two-digit SIC industry following Aghion et al. (2013); and TNIC HHI (Hoberg and Phillips, 2016), a sales-based Herfindahl-Hirschman index for the firm's industry that is defined by text-based network industry classifications. We then split the full sample into the above-median (High) and below-median (Low) subsamples based on the median value of Lerner index or TNIC HHI. Next, we re-estimate Equation 7 for each subsample and report the results in Table 9.

[Please insert Table 9 about here]

We find that firms utilizing more patents in highly competitive product markets (proxied by lower Lerner index or lower TNIC HHI) tend to launch a greater number of higher-quality new products (Panel A), achieve stronger product-market performance (Panel B), and realize higher profitability improvement and valuations (Panel C) than firms in less competitive markets (higher Lerner index / TNIC HHI). These results are consistent with the hypothesis that the benefits of patent utilization are amplified under stronger product-market competition.

Overall, the findings in this section suggest that the positive effects of patent utilization on firms' future performance are driven by the utilization of high-value patents and are more pronounced for firms in competitive product markets.

6 Additional Analyses

In this section, we address potential endogeneity concerns around corporate patent utilization. We implement a Bartik shift–share instrument to strengthen causal interpretation and conduct a series of additional checks to verify robustness.

6.1 Addressing Endogeneity Concerns

In the previous sections, we document positive associations between a firm's patent utilization and its subsequent new-product development, product-market performance, profitability, and valuation. These relationships may, however, be affected by endogeneity. First, better-performing firms may be more inclined to commercialize their patents to defend market share and sustain their leadership position. Second, unobserved factors, such as managerial quality, organizational culture, or demand shocks, may jointly drive both patent utilization and future firm performance, generating spurious correlations.

To help alleviate these endogeneity concerns and strengthen the causal interpretation of our findings, we construct a Bartik shift-share instrument for the corporate patent utilization rate (Bartik, 1991; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2025). The Bartik instrument aims to identify the treatment effect by measuring the differential impact of common shocks on units with distinct predetermined exposures. For instance, Bartik (1991) instruments a county's employment rate by interacting the nationwide growth of industry employment with the county's preexisting industry shares. More recently, Jaravel (2019) applies a similar shift-share framework to causally show that increases in relative demand leads to greater product variety and lower inflation for continuing products. In his setting, the endogenous variable, product demand, is instrumented using the interaction between predetermined spending shares across product categories for different sociodemographic groups and the changes in the number of households in each group.

In line with the prior literature, our shift-share research design combines two components: the predetermined technology-class share of a firm's patent portfolio (*share*), and the differential growth in patent utilization rate across technology classes (*shift*). Specifically, the share component, denoted as $s_{c,i,t}$, is the share of patents in class c held by firm i , computed over a 10-year period preceding the construction window of *Pat. Utilization Rate*.[32] The shift component, denoted as $g_{c,t}$, captures the economy-wide growth in the patent utilization rate of technology class c from year *t-3* to t. Importantly, when measuring the utilization rate of patent class c, we exclude patents belonging to the focal firm and its product market rivals (TNIC2). This ensures that the patent class-level growth in utilization rate captures class-specific advances in technology commercialization which is plausibly exogenous to the focal firm's competitive environment. We then instrument the endogenous variable, *Pat. Utilization Rate*$_{i,t}$, by *Bartik IV*$_{i,t}$, which is calculated as the sum of the products of $s_{c,i,t}$ and $g_{c,t}$ (i.e., $\sum_c s_{c,i,t} \cdot g_{c,t}$), in two-stage least squares (2SLS) regressions.

The rationale behind this instrument is that firms differ in the technological composition of their patent portfolios.[33] The preexisting patent class shares reflect a firm's ex-ante exposure to specific technological fields. When certain technology classes experience commercialization advances, firms with larger preexisting shares in those fields are more positively affected. Our approach ensures that the variation in the shift-share instrument is driven by the heterogeneous, economy-wide growth in utilization rates across patent technology classes that is isolated from the firm's product markets as well as potential unobserved firm-specific factors. Table 10 reports the 2SLS regression results.

[Please insert Table 10 about here]

Column 1 reports the first-stage regression of *Pat. Utilization Rate* on *Bartik IV*. As ex-

[32] Recall that our patent utilization rate measure is constructed based on a firm's past 5 years (from year *t-4* to t) patent portfolio. Our findings remain robust if we change from the 10-year historical window to all the years for which the focal firm's patent data are available.

[33] According to the USPTO, there are 132 unique three-digit patent classes.

pected, the coefficient on the instrument is positive and statistically significant at the 1% level. The first-stage F-statistic (Cragg–Donald Wald) is 17.84, which rejects the null of a weak instrument. Columns 2 to 9 present the second-stage estimates using the instrumented utilization rate, $\widehat{Pat.Utilization}\ Rate$. Consistent with our earlier findings, the results suggest that the instrumented patent utilization rate continues to be positively related to future firm performance: firms with higher patent utilization rates tend to launch more and higher-quality products, achieve stronger product market performance, and yield higher profitability improvement in the future.

In sum, the 2SLS estimates based on a Bartik shift–share design support a causal link from corporate patent utilization to future firm performance. The findings indicate that the capitalization of a firm's patent stock into commercial products is an economically meaningful driver of future growth.

6.2 Robustness Checks

6.2.1 Controlling for Firms' Patent Outputs

A natural concern is that our patent–utilization rate, constructed partly with a firm's patent application portfolio in the denominator, may simply proxy patenting output. If so, our results could be driven by the scale or value of the patent stock rather than by utilization per se.

To address this, we re-estimate Equation 7 with firm and year fixed effects and include, on the right-hand side, (i) the citation-weighted patent stock scaled by total assets (*#CW Patents/AT*) and (ii) the economic value of patent stock scaled by total assets (*Patent Values/AT*), each constructed over the prior five years. If the baseline findings were primarily driven by patenting outputs, the coefficient on *Pat.Utilization Rate* should attenuate to zero once these patent stock measures are controlled for. Conversely, if *Pat.Utilization Rate* captures commercialization capability beyond standard output metrics, it should remain positive and significant. Results are reported in Panel A of Table 11.

[Please insert Table 11 about here]

After controlling for a firm's patent output, we continue to find a positive and statistically significant relationship between patent utilization rate and firm performance in the subsequent year. The economic magnitudes of *Pat.Utilization Rate* are similar to the baseline results. Consistent with Kogan et al. (2017) and Ma (2025), we also find that patent economic value is positively and significantly associated with future firm performance, whereas the relationship between citation-weighted patent counts and firm performance is insignificant.[34] These findings imply that the measure of corporate patent utilization provides information beyond traditional innovation output metrics. While a firm's patent portfolio reflects the stock of its innovation output, the patent utilization rate captures the extent to which the firm leverages these patents in new product development.

6.2.2 Restrict to Firm-year Observations with at least One New Product Launch

Another concern is that our findings could be driven by the extensive margin, that is, whether a firm launches any new product in a given year, rather than by variation in utilization conditional on launching. Although our baseline specifications already control for new-product intensity, we further address this issue by restricting the sample to firm–year observations with at least one new product launch. This design conditions on crossing the extensive margin and isolates the intensive-margin relation between patent utilization and subsequent firm outcomes.

The results in Panel B of Table 11 show that *Pat.Utilization Rate* remains positively and significantly associated with future new-product development, product-market performance, and profitability changes. Therefore, our findings are not driven by the mere incidence of product launching; rather, the measure of patent utilization rate captures economically meaningful variation beyond the extensive margin.

[34] This could be because citation-weighted measures capture the scientific value of inventions, which may generate positive externalities for society without necessarily benefiting the firm.

6.2.3 Alternative Measures of Firm Performance

In Table A5, we further examine the robustness of our baseline findings by exploiting alternative measures of firm performance.

Specifically, Panel A sheds light on highly valued new products, which is defined as those with CARs above the 95th percentile. We construct two outcome variables: *#Highly-Valued New Products*, which represents the number of highly valued new products a firm launches in a year, and *1(Highly-Valued New Products)*, an indicator that equals one if a firm launches at least one highly valued new product in a year and zero otherwise. Panel B provides an alternative measure of a firm's market share growth, calculated as the difference between a firm's sales growth and the median sales growth of its Fama-French 49 industry peers in the same year. In Panel C, we focus on a firm's profit improvement measured by the changes in its operating cash flow between year t and $t+1$. Finally, in Panel D, we replace the market-to-book ratio with Tobin's Q. Across all these alternative performance measures, we find that the results from Table A5 are both qualitatively and quantitatively similar to our baseline findings.

6.2.4 Controlling for Product Similarity Score

One might be concerned that our patent utilization measure overlaps with the product similarity score constructed by Hoberg and Phillips (2016), which is based on the business description section of firms' 10-K filings. However, while their measure primarily captures product market competition, our measure differs both conceptually and methodologically. Specifically, we aim to quantify a firm's patent portfolio utilization rate rather than its competitive positioning. To achieve this, we compare the textual content of the firm's new products with its patent portfolio from the past five years, generating a *within-firm* patent-product similarity score. In contrast, Hoberg and Phillips (2016)'s measure emphasizes *firm-to-firm* product market similarity.

Nevertheless, to formally address this concern, we control for the product similarity score in addition to our baseline control variables. The regression results, presented in Table A6, remain qualitatively and quantitatively consistent with our baseline findings.

6.2.5 Alternative Measures of Patent Utilization

In the baseline regressions, we consider the past 5 years' patent portfolio of a firm and classify a patent as being utilized in a product if the textual similarity score is above 80th percentile across our sample's patent-product pair scores. In this subsection, we further construct alternative measures of patent utilization rate using either 10-year patent portfolio window of a firm or alternative patent-product pair score cutoffs (70th or 90th). In addition, we further employ a 3-year moving average approach to generate an alternative measure of firm-level patent utilization rate, accounting for the possibility that a firm's patent portfolio utilization in new product development may be more stable over the medium term. Similarly, instead of focusing on a 1-year (i.e., current year t) patent-product incorporation rate over the past five-year patent portfolio window, we extend the patent usage window to three years. That is, we count the number of unique patents that have been incorporated into new products launched over the past three years from year t-2 to year t, scaled by the total number of unique patents applied for and later granted by the firm from year t-6 to year t.

The robustness results in Table A7 show that our findings are qualitatively similar across alternative definitions of the patent–utilization rate. In the patent-level analyses (Panels A–B), utilized patents remain more likely to lie in a firm's core technology domains, receive more self-citations, exhibit a higher self-citation ratio, and be less likely to be sold. In the product-level analyses (Panels C–D), products embedding a larger number of utilized patents continue to earn significantly higher announcement returns (CARs) and display higher quality (breakthrough index), regardless of the utilization definition. Finally, in the firm-year analyses (Panel E), alternative utilization measures remain positively and significantly associated with subsequent new-product development, product-market performance, and profitability improvement. These findings indicate that the results are not an artifact of any single construction choice (e.g., similarity cutoff or look-back window).

6.2.6 Control for Industry-by-Time Fixed Effects

In Table A8, we augment the specifications with industry-by-time fixed effects to absorb time-varying industry shocks that could jointly affect firms' utilization of patented technologies and their performance. The results are unchanged: patent-backed products exhibit higher quality (Panels A–B); firms with higher patent-utilization rates develop more and higher-quality new products (Panel C), achieve stronger product-market performance (Panels D–E), and show greater profitability improvements and higher market valuation (Panel F).

7 Conclusion

This study introduces a machine learning approach to capture the extent to which a firm's patent portfolio contributes to new product development, based on the textual analysis of firm patent and product texts. Leveraging the pre-trained *FastText* model, we construct a product- and firm-level patent utilization rate for 3,102 firms from 2002 to 2022. We exploit this novel measure to deepen our understanding of firms' patent utilization rate in new product development, and its implications for future firm performance.

We first validate the measure based on both the patent- and product-level analyses. Our findings show that, at the patent level, patents utilized in new products are more likely to belong to firms' core technology fields, receive more self citations, exhibit higher self-citation ratio, and are less likely to be sold to other entities than unutilized ones. At the product level, new products incorporating more patents tend to be of higher quality, as measured by cumulative abnormal stock returns in the three-day window surrounding the announcement, and by the text-based breakthrough index, which captures a product's novelty and impact.

Aggregating the patent utilization measure to firm level, we find that corporate patent utilization is positively and significantly associated with future new product development, product market performance, profitability improvement, and firm valuation. Heterogeneity tests further reveal that these positive effects are primarily driven by the utilization of high-value patents. Moreover, firms operating in competitive product markets benefit more from effective patent utilization, highlighting the role of product market dynamics in shaping the value derived from intellectual assets.

To address potential endogeneity concerns, we implement a Bartik shift-share instrument that isolates plausibly exogenous variation in corporate patent utilization. The instrument interacts each firm's predetermined technology-class shares with economy-wide growth in patent-utilization rates across classes (excluding the focal firm and its product market rivals). Instrumental-variable estimates confirm the robustness of the findings, reinforcing the potential causal relationship between patent utilization and future firm performance.

Our study bridges the gap between patent production and product commercialization, offering new insights into how effectively firms translate their patent portfolios into new products. Our patent–utilization measure also speaks to concerns about "patent portfolio races," in which firms accumulate patents primarily to block rivals rather than to commercialize technologies. Our findings highlight the value of integrating patents into the product-development pipeline. The results offers practical insights for corporate managers to reconsider the use of their patent portfolios, and for policymakers to design initiatives that promote the effective adoption of intellectual assets.

References

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Aghion, P., Van Reenen, J., and Zingales, L. (2013). Innovation and institutional ownership. *American economic review*, 103(1):277–304.

Akcigit, U. and Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, 126(4):1374–1443.

Amesse, F., Desranleau, C., Etemad, H., Fortier, Y., and Seguin-Dulude, L. (1991). The individual inventor and the role of entrepreneurship: A survey of the canadian evidence. *Research policy*, 20(1):13–27.

Argente, D., Baslandze, S., Hanley, D., and Moreira, S. (2020). Patents to products: Product innovation and firm dynamics.

Arora, A., Belenzon, S., and Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3):871–898.

Arts, S., Cassiman, B., and Hou, J. (2023). Position and differentiation of firms in technology space. *Management Science*, 69(12):7253–7265.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Bena, J., Ortiz-Molina, H., and Simintzi, E. (2022). Shielding firm value: Employment protection and process innovation. *Journal of Financial Economics*, 146(2):637–664.

Bena, J. and Simintzi, E. (2025). Machines could not compete with chinese labor: Evidence from us firms' innovation. *Review of Finance, forthcoming*.

Billett, M. T., Garfinkel, J. A., and Yu, M. (2017). The effect of asymmetric information on product market outcomes. *Journal of Financial Economics*, 123(2):357–376.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144.

Bloom, N. and Van Reenen, J. (2002). Patents, real options and firm performance. *The Economic Journal*, 112(478):C97–C116.

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146.

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204.

Bowen, D. E., Frésard, L., and Hoberg, G. (2023). Rapidly evolving technologies and startup exits. *Management Science*, 69(2):940–967.

Brav, A., Jiang, W., Ma, S., and Tian, X. (2018). How does hedge fund activism reshape corporate innovation? *Journal of Financial Economics*, 130(2):237–264.

Campello, M. (2006). Debt financing: Does it boost or hurt firm performance in product markets? *Journal of Financial Economics*, 82(1):135–172.

Cao, S. S., Ma, G., Tucker, J. W., and Wan, C. (2018). Technological peer pressure and product disclosure. *The Accounting Review*, 93(6):95–126.

Chan, L. K., Lakonishok, J., and Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *The Journal of finance*, 56(6):2431–2456.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Chen, W. and Srinivasan, S. (2023). Going digital: Implications for firm value and performance. *Review of Accounting Studies*, pages 1–47.

Choi, J. P. and Gerlach, H. (2017). A theory of patent portfolios. *American Economic Journal: Microeconomics*, 9(1):315–351.

Cohen, L., Gurun, U. G., and Kominers, S. D. (2019). Patent trolls: Evidence from targeted firms. *Management Science*, 65(12):5461–5486.

Cohen, L., Gurun, U. G., Moon, S. K., and Suh, P. (2023). Patent hunters. *Available at SSRN 4635609*.

Cohn, J. B., Liu, Z., and Wardlaw, M. I. (2022). Count (and count-like) data in finance. *Journal of Financial Economics*, 146(2):529–551.

Cooper, M., Knott, A. M., and Yang, W. (2022). Rq innovative efficiency and firm value.

Journal of Financial and Quantitative Analysis, 57(5):1649–1694.

Fadeev, E. (2023). Creative construction: Knowledge sharing and cooperation between firms.

Figueroa, N. and Serrano, C. J. (2019). Patent trading flows of small and large firms. *Research Policy*, 48(7):1601–1616.

Firth, J. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, pages 10–32.

Fresard, L. (2010). Financial strength and product market behavior: The real effects of corporate cash holdings. *The Journal of finance*, 65(3):1097–1122.

Garcia, D. and Norli, Ø. (2012). Geographic dispersion and stock returns. *Journal of Financial Economics*, 106(3):547–565.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Gilbert, R. J. and Newbery, D. M. (1982). Preemptive patenting and the persistence of monopoly. *The American Economic Review*, pages 514–526.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The review of economic studies*, 58(1):43–61.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). Market value and patent citations: A first look.

Hall, B. H. and Ziedonis, R. H. (2001). The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995. *rand Journal of Economics*, pages 101–128.

Hirshleifer, D., Hsu, P.-H., and Li, D. (2013). Innovative efficiency and stock returns. *Journal of financial economics*, 107(3):632–654.

Hirshleifer, D., Hsu, P.-H., and Li, D. (2018). Innovative originality, profitability, and stock returns. *The Review of Financial Studies*, 31(7):2553–2605.

Hoberg, G. and Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Hoberg, G. and Phillips, G. M. (2025). Scope, scale, and concentration: The 21st-century firm. *The Journal of Finance*, 80(1):415–466.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, 108(3):577–598.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics*, 134(2):715–783.

Kalyani, A. (2022). The creativity decline: Evidence from us patents. *Available at SSRN 4318158*.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring technological innovation over the long run. *American Economic Review: Insights*, 3(3):303–320.

Kogan, L., Papanikolaou, D., Schmidt, L., and Seegmiller, B. (2022). Technology, vintage-specific human capital, and labor displacement: Evidence from linking patents with occupations. *Available at SSRN 3983906*.

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

Lentz, R. and Mortensen, D. T. (2008). An empirical model of growth through product innovation. *Econometrica*, 76(6):1317–1373.

Lev, B. and Sougiannis, T. (1996). The capitalization, amortization, and value-relevance of r&d. *Journal of accounting and economics*, 21(1):107–138.

Li, K., Mai, F., Shen, R., and Yan, X. (2021). Measuring corporate culture using machine learning. *The Review of Financial Studies*, 34(7):3265–3315.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65.

Ma, S. (2025). Technological obsolescence. *The Review of Financial Studies*, page hhaf059.

Masclans, R., Hasan, S., and Cohen, W. M. (2025). Measuring the commercial potential of science. *Strategic Management Journal*, 46(9):2199–2236.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., and Joulin, A. (2017). Advances in pre-training distributed word representations. *arXiv preprint arXiv:1712.09405*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10):3328–3375.

Mukherjee, A., Singh, M., and Žaldokas, A. (2017). Do corporate taxes hinder innovation? *Journal of Financial Economics*, 124(1):195–221.

Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Schumpeter, J. A. (1943). *Capitalism, socialism and democracy.* routledge.

Seegmiller, B., Papanikolaou, D., and Schmidt, L. D. (2023). Measuring document similarity with weighted averages of word embeddings. *Explorations in Economic History*, 87:101494.

Singh, J. and Marx, M. (2013). Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, 59(9):2056–2078.

Sougiannis, T. (1994). The accounting based valuation of corporate r&d. *Accounting review*, pages 44–68.

Thompson, P. and Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, 95(1):450–460.

Webster, E. and Jensen, P. H. (2011). Do patents matter for commercialization? *The Journal of Law and Economics*, 54(2):431–453.

Figure 1. Patent-Product Pair Matching Process Illustration

This figure illustrates the patent-product pair matching process. The patent filing text data is obtained from PatentsView, while the new product launch text data is obtained from Capital IQ Key Development Database. The similarity score of patent-product pair is generated via the *Word2vec* model which compares the text similarities between patent filings and new product launch text description.

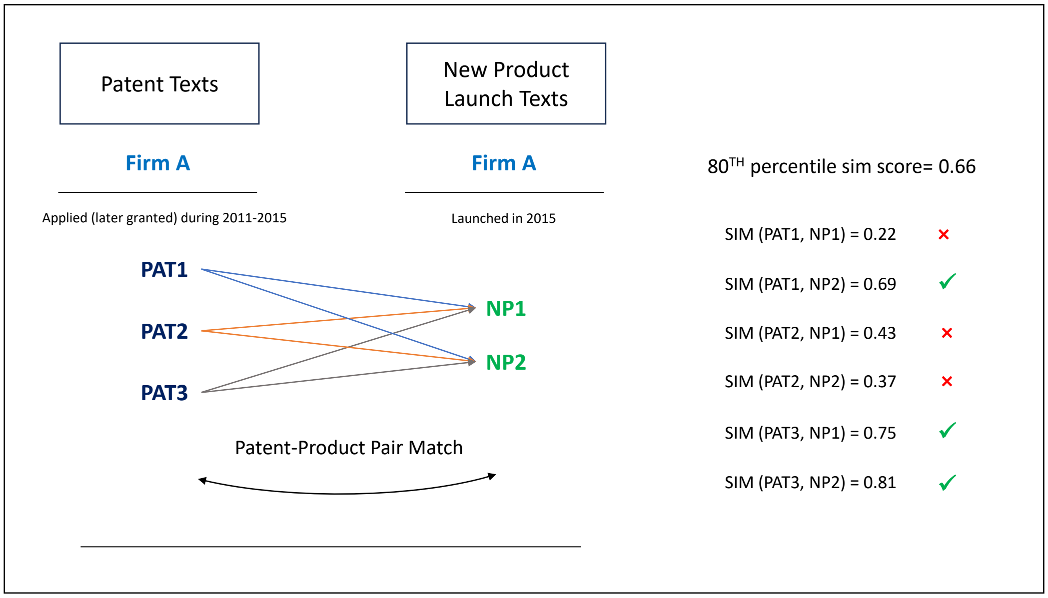


Figure 2. Annual Variation of Patent Utilization Rate

This figure illustrates the annual variation of corporate patent utilization rate by year from 2002 to 2022.

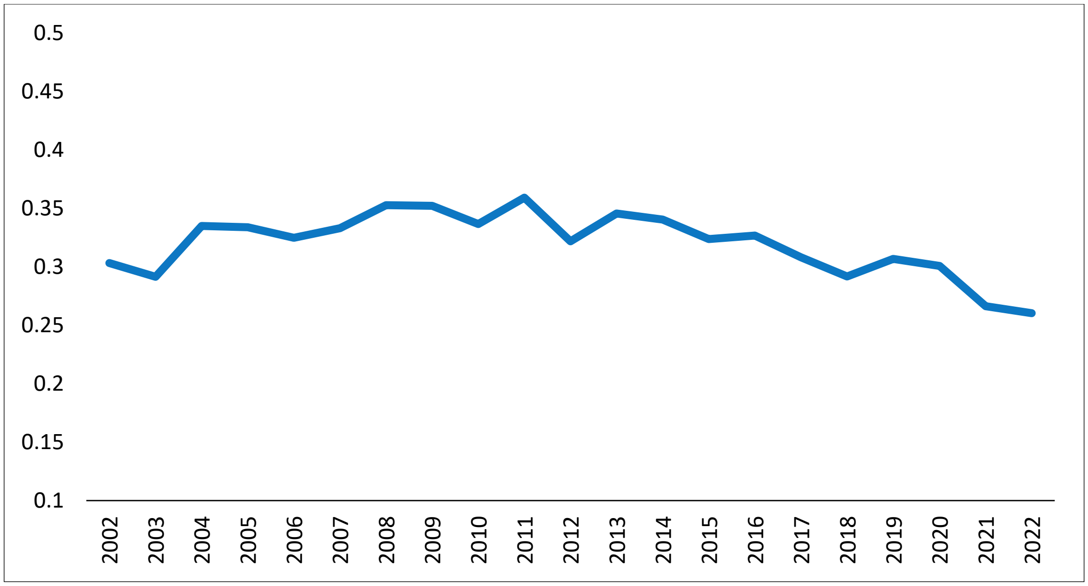


Figure 3. Industry Variation of Patent Utilization Rate

The figure illustrates the top 10 industries (2-digit SIC) with the highest rates in patent utilization. The y-axis denotes the 2-digit SIC and the related industry classification, and the x-axis reports the rate of patent utilization.

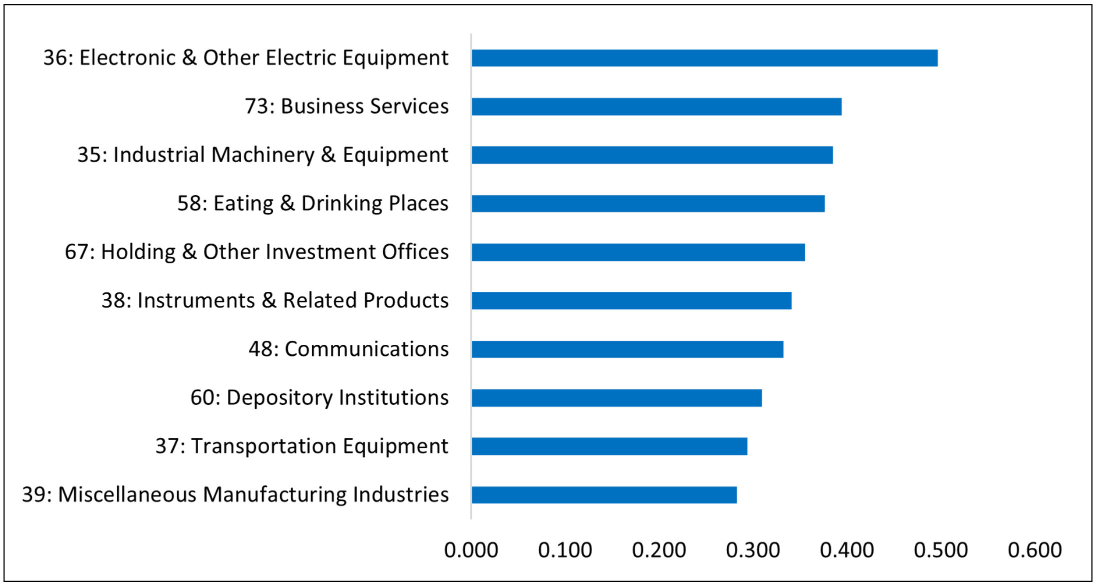


Table 1. Summary Statistics

This table reports the summary statistics for the new product announcement event sample (Panel A) and firm-year regression sample (Panel B) The sample period starts from 2002 to 2022. We report the number of observations, standard deviation, mean, 25th percentile, median, and 75th percentile for each of the variables used in the study. All continuous variables are winsorized at the 1st and 99th percentiles. Table A1 in Appendix A provides detailed variable definitions.

Panel A. Patent Sample

Variables	Obs.	Std.	Mean	P25	Median	P75
I (Core Technology Class)	1,380,261	0.496	0.435	0.000	0.000	1.000
#Self Citations	1,380,261	4.837	1.457	0.000	0.000	1.000
Self Citation Ratio	1,380,261	0.277	0.136	0.000	0.000	0.100
I (Sold)	1,380,261	0.325	0.120	0.000	0.000	0.000
I (Utilized)	1,380,261	0.459	0.698	0.000	1.000	1.000

Panel B. New Porduct Announcemente Event Sample

Variables	Obs.	Std.	Mean	P25	Median	P75
Dependent Variables						
CAR (-1,1)	94,239	4.079	0.118	-1.722	-0.020	1.774
Breakthrough Index	105,196	0.797	0.000	-0.211	-0.106	0.077
1 (Breakthrough Product)	105,196	0.155	0.025	0.000	0.000	0.000
Independent Variables						
Log(1+#Patents Used^{Sum})	94,239	2.367	2.795	0.693	2.565	4.635
Log(1+#Patents Used^{Average})	94,239	2.305	2.732	0.405	2.565	4.543
Log(Product Text Length)	94,239	0.587	5.025	4.635	5.011	5.425
Log(1+#New Products Launched)	94,239	1.263	1.777	0.693	1.609	2.708
Firm Size	94,239	2.324	6.206	4.567	6.207	8.093
Log(Firm Age)	94,239	0.936	2.769	2.197	2.890	3.497
Leverage	94,239	0.161	0.163	0.002	0.133	0.267
R&D	94,239	0.023	0.023	0.007	0.019	0.032
ROA	94,239	0.038	0.026	0.014	0.029	0.044
Cash	94,239	0.192	0.266	0.112	0.225	0.383
Log(Tobin's Q)	94,239	0.530	0.672	0.273	0.614	1.002
Sales Growth	94,239	0.185	0.014	-0.042	0.022	0.082
Past Stock Return	94,239	0.273	0.034	-0.100	0.019	0.137

Panel C. Firm-Year Sample

Variables	Obs.	Std.	Mean	P25	Median	P75
Dependent Variables						
#New Products	27,843	2.863	1.088	0.000	0.000	1.000
Sum CARs	27,843	10.910	5.291	0.000	0.000	5.769
#Breakthrough Products	27,843	0.877	0.098	0.000	0.000	0.000
Sales Growth	27,795	0.308	0.066	-0.033	0.066	0.171
MSG (SIC4)	27,795	0.283	-0.003	-0.079	0.000	0.077
△GPM	27,844	0.135	0.033	-0.015	0.022	0.070
△ROA	27,812	0.115	0.015	-0.017	0.010	0.041
Log(MTB)	26,456	0.901	1.012	0.439	0.936	1.477
Independent Variables						
Pat.Utilization Rate	27,844	0.390	0.313	0.000	0.000	0.696
#New Products/Sales	27,844	0.574	0.019	0.000	0.000	0.001
Firm Size	27,844	2.463	6.397	4.574	6.356	8.211
Log(Firm Age)	27,844	0.815	2.834	2.303	2.890	3.434
Leverage	27,844	0.215	0.198	0.005	0.156	0.312
R&D Stock	27,844	0.396	0.197	0.007	0.073	0.221
Cash	27,844	0.376	0.312	0.077	0.199	0.421
Past Stock Return	27,844	0.656	0.148	-0.204	0.063	0.344

Table 2. Validation: Utilized Patents, Firms' Core Technology Fields, Self Citations, and Patent Sales

This table reports the regression results that validate the patent utilization measure at the patent level. Panel A investigates whether patents that are utilized in new product development by firms are more likely to belong to firms' core technology fields. Panel B examines whether those utilized patents receive more self citations, have higher self citation ratio, and are less likely to be sold. The dependent variable *I (Core Technology Class)* is an indicator that equal one if a patent belongs to class *c* in which the patent-owning firm has produced the most patents over the past 10 years, otherwise equals zero. *# Self Citations* is the number of self citations received by a patent. *Self Citation Ratio* is measured as the number of self citations divided by the total number of citations received by a patent. *I (Sold)* is an indicator that equals one if the patent is sold by the owning firm to another entity prior to the first patent maintenance date (four years after patent issuance), otherwise equals zero. The independent variable *I (Utilized)* is an indicator that equals one if a patent is utilized in new products by the owning firm, otherwise equals zero. We define a patent is utilized in a new product if the patent-product pair text similarity score is above 80th percentile. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the patent level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Utilized Patents and the Likelihood of Belonging to the Firm's Core Technology Class

	(1)	(2)	(3)
VARIABLES		I (Core Technology Class)	
I (Utilized)	0.034***	0.033***	0.032***
	(0.001)	(0.001)	(0.001)
Model	OLS	OLS	OLS
Firm FE	✓	✓	
Class FE	✓		
Year FE	✓		
Class-Year FE		✓	✓
Firm-Year FE			✓
Obs.	1,380,261	1,380,130	1,375,022
Adj. R2	0.366	0.382	0.406

Panel B. Utilized Patents, Self Citations, and Patent Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES		# Self Citations			Self Citation Ratio			I (Sold)	
I (Utilized)	0.049***	0.083***	0.082***	0.005***	0.006***	0.006***	-0.004***	-0.001**	-0.001**
	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS	OLS
Firm-Class FE	✓	✓		✓	✓		✓	✓	
Class-Year FE	✓	✓		✓	✓		✓	✓	
Firm-Year FE		✓			✓			✓	
Firm-Class-Year FE			✓			✓			✓
Obs.	1,349,952	1,318,303	1,218,706	1,375,093	1,369,142	1,328,513	1,375,093	1,369,142	1,328,513
Pseudo/Adj. R2	0.362	0.394	0.402	0.121	0.139	0.144	0.336	0.480	0.509

Table 3. Validation: Number of Patents Utilized in New Products and New Product Announcement Return

This table reports the regression results that investigate the association between the number of unique patents utilized in a new product and the product's announcement return. The dependent variable *CAR (-1, 1)* is the cumulative abnormal stock returns during a three-day event window of (-1, 1) following the new product announcement event. Because a firm may launch multiple products on one event date, we then generate two independent variables to account for the situation of product bundle launching: *Log(1+#Patents Utilized$^{Sum(Average)}$)*, which is the natural logarithm of one plus the sum (average) of the number of patents that are utilized in new product(s) for an event-date. We define a patent is utilized in a new product if the patent-product pair text similarity score is above 80th percentile. All regression specifications include product and firm level control variables. Columns 1 and 3 include industry and event-year-week fixed effects, while Columns 2 and 4 include firm and event-year-week fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm and event-year-week level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR (-1,1)			
Log (1+#Patents Used^{Sum})	**0.024***	**0.020**		
	(0.008)	**(0.008)**		
Log (1+#Patents Used^{Average})			**0.023***	**0.018**
			(0.008)	**(0.009)**
Log(Product Text Length)	-0.018	0.009	-0.017	0.010
	(0.026)	(0.027)	(0.026)	(0.027)
Log(1+# New Products Launched)	-0.019	-0.008	-0.018	-0.007
	(0.018)	(0.026)	(0.018)	(0.026)
Firm Size $_{q-1}$	-0.073***	-0.256***	-0.072***	-0.255***
	(0.014)	(0.055)	(0.014)	(0.055)
Log(Firm Age) $_{q-1}$	-0.028	-0.039	-0.028	-0.039
	(0.023)	(0.083)	(0.023)	(0.083)
Leverage $_{q-1}$	0.302**	0.299	0.302**	0.298
	(0.122)	(0.186)	(0.122)	(0.186)
R&D $_{q-1}$	4.257***	6.115***	4.271***	6.122***
	(1.059)	(1.879)	(1.060)	(1.879)
ROA $_{q-1}$	1.489**	5.511***	1.484**	5.508***
	(0.743)	(1.107)	(0.742)	(1.107)
Cash $_{q-1}$	0.203*	0.139	0.204*	0.139
	(0.112)	(0.183)	(0.112)	(0.183)
Log(Tobin's Q) $_{q-1}$	-0.139***	-0.484***	-0.138***	-0.484***
	(0.048)	(0.083)	(0.048)	(0.083)
Sales Growth $_{q-1}$	0.437***	0.319***	0.437***	0.319***
	(0.108)	(0.112)	(0.108)	(0.112)
Stock Return $_{q-1}$	0.058	0.003	0.057	0.003
	(0.092)	(0.094)	(0.092)	(0.094)
Industry FE	✓		✓	
Event-Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	92,614	92,235	92,614	92,235
Adj. R2	0.036	0.054	0.036	0.054

Table 4. Validation: Number of Patents Utilized in New Products and Product Breakthrough

This table reports the regression results that investigate the association between the number of unique patents utilized in a new product and the product's breakthrough index and the likelihood of being a breakthrough product. The dependent variable *Breakthrough Index* is a text-based measurement that captures product significance. Following Kelly et al. (2021), the breakthrough index considers a product's novelty and impact, which is constructed as: $BreakthroughIndex^5{}_\mathrm{j} = \frac{FS^5{}_\mathrm{j}}{BS^5{}_\mathrm{j}}$, where $BS^5{}_j$ measures the backward similarity (novelty dimension) and $FS^5{}_j$ measures the forward similarity (impact dimension). Specifically, $BS^5{}_\mathrm{j} = \sum_{i \in \beta^5_{j,m}} \rho_{j,i}$, where $\rho_{j,i}$ is the pairwise similarity between product j and i, and $\beta^5_{j,m}$ denotes the set of previous products that are launched in the 5 years before product j's offering and that are in the same product market m (based on parent firms' TNIC3 classification (Hoberg and Phillips, 2016) as product j). Similarly, the forward similarity $FS^5{}_\mathrm{j} = \sum_{i \in \alpha^5_{j,m}} \rho_{j,i}$. $\rho_{j,i}$ is the pairwise similarity between product j and i, and $\alpha^5_{j,m}$ denotes the set of future products that are launched in the 5 years after product j's offering and that are in the same product market m (based on parent firms' TNIC3 classification) as product j. Thus, a product with low backward similarity with the prior products (which is novel) but high forward similarity with the subsequent products (which is impactful) has a high breakthrough index. The other dependent variable, *1 (Breakthrough Product)*, is an indicator that equals 1 if the product's breakthrough index is above the 95th percentile, otherwise equals 0. The independent variable *Log(1+#Patents Utilized)* is the natural logarithm of one plus the number of patents that are utilized in the new product. We define a patent is utilized in a new product if the patent-product pair text similarity score is above 80th percentile. All regression specifications include product and firm level control variables. Columns 1 and 3 include industry and event-year-week fixed effects, while Columns 2 and 4 include firm and event-year-week fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm and event-year-week level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	Breakthrough Index		1 (Breakthrough Product)	
Log (1+#Patents Utilized)	**0.019*** **	**0.006** **	**0.003** **	**0.002** **
	(0.006)	**(0.002)**	**(0.001)**	**(0.001)**
Log(Product Text Length)	-0.007	-0.011*	-0.003	-0.001
	(0.015)	(0.006)	(0.005)	(0.002)
Log(1+#New Products Launched)	0.015	-0.003	-0.009**	-0.008
	(0.017)	(0.014)	(0.004)	(0.005)
Firm Size $_{q-1}$	-0.033***	-0.055	-0.002	-0.004
	(0.013)	(0.039)	(0.003)	(0.010)
Log(Firm Age) $_{q-1}$	0.034	0.126*	0.006	0.036***
	(0.023)	(0.067)	(0.005)	(0.012)
Leverage $_{q-1}$	0.101	-0.078	0.002	0.017
	(0.098)	(0.099)	(0.021)	(0.031)
R&D $_{q-1}$	-1.149	-0.772	-0.327	-0.040
	(0.711)	(0.759)	(0.204)	(0.149)
ROA $_{q-1}$	-0.194	-0.422	-0.016	0.018
	(0.472)	(0.357)	(0.100)	(0.086)
Cash $_{q-1}$	-0.173**	-0.138	-0.078***	-0.029
	(0.084)	(0.097)	(0.018)	(0.021)
Log(Tobin's Q) $_{q-1}$	0.026	-0.037	-0.001	0.002
	(0.032)	(0.030)	(0.008)	(0.007)
Sales Growth $_{q-1}$	0.007	0.034	-0.010	-0.002

	(0.029)	(0.034)	(0.007)	(0.007)
Stock Return $_{\text{q-1}}$	-0.012	0.005	0.004	0.008
	(0.023)	(0.022)	(0.006)	(0.006)
Industry FE	✓		✓	
Event-Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	103,509	103,133	103,509	103,133
Adj. R2	0.052	0.354	0.125	0.374

Table 5. Corporate Patent Utilization Rate and New Product Development

This table reports the regression results that investigate the association between a firm's patent utilization rate and one-year-ahead new product development. The dependent variable *#New Products* is measured as the number of new products (with three-day CARs above 80th percentile) a firm launches in a year. *Sum CARs* is measured as the sum of all positive three-day cumulative abnormal stock returns of the new products that a firm launches in a year. *#Breakthrough Products* is measured as the number of breakthrough products (new products with breakthrough indexes above 95th percentile) a firm launches in a year. The independent variable *Pat. Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We regard a patent as utilized in a product if the patent-product pair similarity score is above 80th percentile. All specifications include firm controls. Columns 1, 3, and 5 include industry and year fixed effects, while Columns 2, 4, and 6 include firm and year fixed effects. The results are estimated with Poisson regressions. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	#New Products $_{t+1}$		Sum CARs $_{t+1}$		#Breakthrough Products $_{t+1}$	
Pat. Utilization Rate	**1.738***** **(0.063)**	**0.676***** **(0.053)**	**1.574***** **(0.057)**	**0.496***** **(0.046)**	**1.411***** **(0.123)**	**0.506***** **(0.155)**
#New Products/Sales	0.050*** (0.012)	0.167*** (0.032)	0.038*** (0.013)	0.048 (0.046)	0.036 (0.048)	-0.259*** (0.087)
Firm Size	0.318*** (0.019)	0.222*** (0.056)	0.218*** (0.018)	0.107** (0.046)	0.122*** (0.026)	-0.109 (0.123)
Log(Firm Age)	-0.074 (0.046)	-0.032 (0.098)	-0.102** (0.042)	-0.058 (0.087)	0.159** (0.073)	0.145 (0.351)
Leverage	-0.337** (0.142)	-0.060 (0.123)	-0.078 (0.131)	0.197* (0.113)	0.154 (0.282)	-0.564 (0.470)
R&D Stock	0.533*** (0.051)	-0.054 (0.064)	0.462*** (0.044)	-0.131*** (0.051)	-1.437** (0.651)	-2.301** (1.077)
Cash	0.409*** (0.047)	0.083* (0.050)	0.344*** (0.048)	0.085* (0.048)	-0.134 (0.181)	-0.198 (0.229)
Past Stock Return	-0.096*** (0.019)	-0.075*** (0.015)	-0.033* (0.019)	-0.023 (0.015)	0.013 (0.080)	-0.018 (0.059)
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Industry FE	✓		✓		✓	
Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,838	23,397	27,839	24,915	24,411	6,192
Pseudo R2	0.495	0.495	0.595	0.595	0.381	0.381

Table 6. Corporate Patent Utilization Rate and Product Market Performance

This table reports the regression results that investigate the association between a firm's patent utilization and one-year-ahead product market performance. In Panel A, we report the relationship between a firm's patent utilization rate and sales growth. In Panel B, we shed light on the association between a firm's patent utilization rate and market share growth. The dependent variable *Sales Growth* is measured as the natural logarithm of total sales for a firm in a year minus the natural logarithm of total sales for that firm in the previous year. *MSG(SIC4)* is measured as the sales growth of a firm in a year minus the industry (4-digit SIC) median sales growth in the same year. The independent variable *Pat. Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm at the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We regard a patent as utilized in a product if the patent-product pair similarity score is above 80th percentile. All specifications include firm controls. Columns 1, 3 and 5 include industry and year fixed effects. Columns 2, 4 and 6 include firm and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Patent Utilization Rate and Firm Future Sales Growth

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Sales Growth $_{t+1}$		Sales Growth $_{t+2}$		Sales Growth $_{t+3}$	
Pat. Utilization Rate	**0.016*****	**0.022*****	**0.030*****	**0.033*****	**0.034****	**0.031****
	(0.005)	**(0.007)**	**(0.009)**	**(0.010)**	**(0.013)**	**(0.013)**
#New Products/Sales	0.009	0.040	0.059	0.082**	0.047	0.062
	(0.007)	(0.032)	(0.041)	(0.038)	(0.050)	(0.042)
Firm Size	-0.004***	-0.185***	-0.006**	-0.334***	-0.009**	-0.456***
	(0.001)	(0.010)	(0.003)	(0.016)	(0.004)	(0.019)
Log(Firm Age)	-0.024***	0.028**	-0.042***	0.076***	-0.063***	0.092***
	(0.003)	(0.013)	(0.006)	(0.023)	(0.008)	(0.031)
Leverage	0.068***	0.085***	0.112***	0.093**	0.161***	0.128***
	(0.013)	(0.024)	(0.025)	(0.037)	(0.036)	(0.047)
R&D Stock	-0.069***	-0.093***	-0.126***	-0.143***	-0.147***	-0.165***
	(0.011)	(0.014)	(0.019)	(0.024)	(0.028)	(0.036)
Cash	0.077***	0.012	0.152***	0.042**	0.211***	0.050**
	(0.010)	(0.014)	(0.016)	(0.021)	(0.022)	(0.022)
Past Stock Return	0.058***	0.042***	0.081***	0.044***	0.103***	0.052***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.007)	(0.005)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Industry FE	✓		✓		✓	
Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,796	27,621	25,700	25,479	23,489	23,260
Adj. R2	0.072	0.205	0.090	0.362	0.099	0.478

Panel B. Patent Utilization Rate and Future Market Share Growth

VARIABLES	(1) MSG (SIC4) $_{t+1}$	(2)	(3) MSG (SIC4) $_{t+2}$	(4)	(5) MSG (SIC4) $_{t+3}$	(6)
Pat. Utilization Rate	**0.014*****	**0.024*****	**0.024*****	**0.032*****	**0.025****	**0.033*****
	(0.005)	**(0.006)**	**(0.009)**	**(0.010)**	**(0.012)**	**(0.012)**
#New Products/Sales	0.010	0.039	0.055	0.082**	0.041	0.062
	(0.006)	(0.028)	(0.039)	(0.036)	(0.045)	(0.043)
Firm Size	-0.002	-0.164***	-0.003	-0.300***	-0.004	-0.414***
	(0.001)	(0.009)	(0.003)	(0.015)	(0.004)	(0.018)
Log(Firm Age)	-0.020***	0.020*	-0.034***	0.058***	-0.052***	0.057**
	(0.003)	(0.012)	(0.006)	(0.021)	(0.008)	(0.029)
Leverage	0.065***	0.078***	0.106***	0.076**	0.151***	0.108**
	(0.013)	(0.022)	(0.024)	(0.035)	(0.035)	(0.046)
R&D Stock	-0.067***	-0.079***	-0.130***	-0.129***	-0.155***	-0.150***
	(0.010)	(0.013)	(0.019)	(0.024)	(0.028)	(0.035)
Cash	0.071***	0.018	0.135***	0.045**	0.181***	0.048**
	(0.010)	(0.013)	(0.015)	(0.020)	(0.021)	(0.022)
Past Stock Return	0.041***	0.025***	0.067***	0.033***	0.085***	0.040***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.007)	(0.005)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Industry FE	✓		✓		✓	
Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,796	27,621	25,700	25,479	23,489	23,260
Adj. R2	0.030	0.166	0.044	0.328	0.048	0.445

Table 7. Corporate Patent Utilization Rate, Profit Improvements, and Firm Value

This table reports the regression results that investigate the association between a firm's patent utilization rate and one-year-ahead firm profit improvements and valuation. The dependent variable $\triangle GPM$ is measured as the change in the gross profit margin of a firm between year *t+1* and *t*; gross profit margin is defined as a firm's sales minus cost of goods sold, further divided by the firm's book value of total assets at the beginning of the year. $\triangle ROA$ is measured as the change in the return on assets (ROA) of a firm between year *t+1* and *t*; ROA is defined as a firm's operating income before depreciation divided by the firm's book value of total assets at the beginning of the year. *Log(MTB)* is measured as the natural logarithm of a firm's market value of assets divided by the book value of total assets. The independent variable *Pat.Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We regard a patent as utilized in a product if the patent-product pair similarity score is above 80th percentile. All specifications include firm controls. Columns 1, 3 and 5 include industry and year fixed effects. Columns 2, 4 and 6 include firm and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	$\triangle$GPM $_{t+1}$		$\triangle$ROA $_{t+1}$		Log(MTB) $_{t+1}$	
Pat. Utilization Rate	**0.016*** **	**0.018*** **	**0.008*** **	**0.007** **	**0.120*** **	**0.034** **
	(0.003)	**(0.004)**	**(0.002)**	**(0.003)**	**(0.024)**	**(0.016)**
#New Products/Sales	0.004	0.018	-0.002	-0.018	-0.008	-0.043*
	(0.004)	(0.020)	(0.003)	(0.013)	(0.006)	(0.024)
Firm Size	-0.011***	-0.089***	-0.000	-0.044***	0.032***	-0.085***
	(0.001)	(0.005)	(0.001)	(0.004)	(0.007)	(0.017)
Log(Firm Age)	0.002	0.040***	-0.010***	-0.011	-0.024	-0.146***
	(0.002)	(0.008)	(0.002)	(0.007)	(0.017)	(0.036)
Leverage	0.010	0.008	0.021***	0.077***	1.075***	0.895***
	(0.008)	(0.015)	(0.005)	(0.011)	(0.077)	(0.081)
R&D Stock	0.038***	-0.010	0.016***	0.010	0.315***	0.105***
	(0.008)	(0.011)	(0.004)	(0.007)	(0.036)	(0.034)
Cash	-0.147***	-0.220***	-0.073***	-0.104***	0.424***	0.018
	(0.010)	(0.014)	(0.009)	(0.012)	(0.027)	(0.022)
Past Stock Return	-0.000	-0.002	0.008***	0.003	0.166***	0.127***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.010)	(0.007)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Industry FE	✓		✓		✓	
Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,844	27,667	27,813	27,638	26,461	26,277
Adj. R2/Pseudo R2	0.091	0.130	0.047	0.120	0.101	0.533

Table 8. High-Value versus Low-Value Patent Utilization Rate

This table compares the effects of high-value versus low-value patent utilization rate on future firms' new product development (panel A), product market performance (panel B), and profit improvements and firm value (panel C). We define a patent as high value if its economic value is above median across all patents' value. Similarly, a patent is regarded as low value if the economic value is below median. The independent variable *High-Value (Low-Value) Pat.Utilization Rate* is measured as the number of high-value (low-value) patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We further standardize these two variables to unit standard deviation in order to facilitate magnitude interpretations. We regard a patent as utilized in a product if the patent-product pair similarity score is above 80th percentile. All specifications include firm controls, industry (firm) fixed effects, and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. New Product Development

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	#New Products $_{t+1}$				Sum CARs $_{t+1}$				#Breakthrough Products $_{t+1}$			
Patent Values	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
High-Value (Low-Value) Pat.Utilization Rate	0.326***	0.298***	0.216***	0.069***	0.301***	0.302***	0.162***	0.057***	0.335***	0.274***	0.190***	0.056
	(0.024)	(0.023)	(0.021)	(0.019)	(0.019)	(0.017)	(0.015)	(0.015)	(0.037)	(0.053)	(0.063)	(0.063)
P Value of Diff.	0.478		0.000		0.995		0.000		0.705		0.189	
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	27,838	27,838	23,397	23,397	27,839	27,839	24,915	24,915	24,411	24,411	6,192	6,192
Pseudo R2	0.298	0.283	0.488	0.483	0.279	0.274	0.552	0.548	0.213	0.203	0.335	0.333

Panel B. Product Market Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Sales Growth $_{t+1}$				MSG (SIC4) $_{t+1}$			
Patent Values	High	Low	High	Low	High	Low	High	Low
High-Value (Low-Value) Pat.Utilization Rate	0.020***	-0.009***	0.014***	-0.000	0.017***	-0.007***	0.014***	0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
P Value of Diff.	0.000		0.000		0.000		0.000	
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	27,796	27,796	27,621	27,621	27,796	27,796	27,621	27,621
Adj. R2	0.075	0.072	0.206	0.205	0.033	0.030	0.167	0.166

Panel C. Profit Improvements and Firm Value

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	$\triangle$GPM $_{t+1}$				$\triangle$ROA $_{t+1}$				Log(MTB) $_{t+1}$			
Patent Value	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
High-Value (Low-Value)	0.008***	-0.003***	0.004***	0.001	0.004***	-0.001	0.003***	0.001	0.126***	-0.051***	0.033***	-0.014**
Pat.Utilization Rate	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.010)	(0.009)	(0.007)	(0.007)
P Value of Diff.	0.000		0.072		0.000		0.156		0.000		0.003	
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	27,844	27,844	27,667	27,667	27,813	27,813	27,638	27,638	26,461	26,461	26,277	26,277
Adj. R2	0.054	0.052	0.161	0.161	0.045	0.044	0.095	0.094	0.228	0.214	0.587	0.586

Table 9. Heterogeneous Tests: Product Market Competition and Patent Utilization Rate

This table investigates the heterogeneous effects of patent utilization rate on future firms' new product development (panel A), product market performance (panel B), and profit improvements and firm value (panel C) based on firms' product market competition, which is proxied by Lerner index and the Herfindahl-Hirschman index based on text-based network industry classification. The independent variable *Pat.Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We regard a patent as utilized in a product if the patent-product pair similarity score is above the 80th percentile. All specifications include firm controls, industry (firm) fixed effects, and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. New Product Development

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	#New Products $_{t+1}$				Sum CARs $_{t+1}$				#Breakthrough Products $_{t+1}$			
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
					Lerner Index							
Pat.Utilization Rate	1.615***	1.859***	0.543***	0.726***	1.454***	1.538***	0.366***	0.436***	0.965***	1.641***	0.177	0.844***
	(0.080)	(0.077)	(0.068)	(0.070)	(0.062)	(0.062)	(0.055)	(0.060)	(0.199)	(0.156)	(0.204)	(0.246)
P Value of Diff.	0.001		0.008		0.273		0.415		0.006		0.037	
Obs.	14,087	13,736	11,469	10,534	14,088	13,738	12,370	11,408	12,437	11,812	2,251	2,702
Pseudo R2	0.318	0.389	0.467	0.520	0.320	0.364	0.548	0.582	0.150	0.270	0.221	0.403
					TNIC HHI							
Pat.Utilization Rate	1.522***	1.588***	0.531***	0.741***	1.352***	1.369***	0.416***	0.470***	1.226***	1.635***	0.231	0.658
	(0.076)	(0.095)	(0.071)	(0.091)	(0.067)	(0.068)	(0.057)	(0.068)	(0.140)	(0.211)	(0.182)	(0.473)
P Value of Diff.	0.558		0.083		0.846		0.537		0.112		0.402	
Obs.	12,725	11,755	10,192	9,237	12,762	11,808	11,009	9,839	11,038	8,894	3,161	1,481
Pseudo R2	0.235	0.424	0.391	0.529	0.233	0.428	0.489	0.595	0.188	0.447	0.263	0.598
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Panel B. Product Market Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Sales Growth $_{t+1}$				MSG (SIC4) $_{t+1}$			
	High	Low	High	Low	High	Low	High	Low
				Lerner Index				
Pat.Utilization Rate	0.007	0.026***	0.007	0.032***	0.004	0.023***	0.007	0.036***
	(0.007)	(0.008)	(0.007)	(0.012)	(0.006)	(0.008)	(0.007)	(0.011)
P Value of Diff.	0.060		0.070		0.051		0.026	
Obs.	14,094	13,700	13,916	13,380	14,094	13,700	13,916	13,380
Adj. R2	0.134	0.054	0.308	0.150	0.059	0.022	0.256	0.119
				TNIC HHI				
Pat.Utilization Rate	0.003	0.030***	0.009	0.042***	0.005	0.027***	0.011	0.043***
	(0.007)	(0.008)	(0.008)	(0.012)	(0.007)	(0.008)	(0.008)	(0.011)
P Value of Diff.	0.011		0.023		0.031		0.021	
Obs.	12,761	11,769	12,453	11,458	12,761	11,769	12,453	11,458
Adj. R2	0.104	0.059	0.270	0.167	0.044	0.027	0.215	0.141
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Panel C. Profit Improvements and Firm Value

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	△GPM $_{t+1}$				△ROA $_{t+1}$				Log(MTB) $_{t+1}$			
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
						Lerner Index						
Pat.Utilization Rate	0.002	0.012***	0.004	0.010*	0.001	0.009***	0.003	0.008*	0.107***	0.125***	-0.006	0.071***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)	(0.003)	(0.003)	(0.005)	(0.031)	(0.030)	(0.020)	(0.024)
P Value of Diff.	0.073		0.308		0.045		0.455		0.648		0.013	
Obs.	14,105	13,737	13,927	13,417	14,096	13,715	13,921	13,396	13,419	13,040	13,243	12,722
Adj. R2	0.117	0.032	0.273	0.102	0.035	0.054	0.100	0.089	0.092	0.111	0.570	0.527
						TNIC HHI						
Pat.Utilization Rate	0.002	0.011**	0.002	0.016***	0.003	0.007**	0.003	0.010**	0.118***	0.080**	0.022	0.036
	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)	(0.005)	(0.031)	(0.033)	(0.024)	(0.024)
P Value of Diff.	0.095		0.044		0.297		0.222		0.537		0.686	
Obs.	12,771	11,804	12,462	11,491	12,760	11,786	12,454	11,471	12,119	11,151	11,813	10,837
Adj. R2	0.071	0.039	0.202	0.138	0.037	0.053	0.116	0.077	0.119	0.121	0.568	0.511
Model	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓			✓	✓			✓	✓		
Firm FE			✓	✓			✓	✓			✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 10. Addressing Endogeneity Concerns: Shift-Share Instrumental-Variable Analyses

This table report the results of shift-share instrumental-variable analyses. We construct an instrument *Bartik IV*, by summing the products of $s_{c,i}$ and $g_{c,t}$ for each patent class c of firm i in year t. Specifically, $s_{c,i}$ is the historical share of patents in class c held by firm i. $g_{c,t}$ is the growth (shift) in the patent utilization rate of class c from year *t-3* to t. When measuring the utilization rate in patent class c, we exclude those patents from the focal firm and its product market rivals (TNIC2). All specifications include firm controls, firm fixed effects, and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Pat.Util.Rate $_{t}$	IHS(#New Products) $_{t+1}$	IHS(Sum CARs) $_{t+1}$	IHS(# Break-through Products) $_{t+1}$	Sales Growth $_{t+1}$	MSG (SIC4) $_{t+1}$	$\triangle$GPM $_{t+1}$	$\triangle$ROA $_{t+1}$	Log(MTB) $_{t+1}$
Bartik IV	0.162***								
	(0.047)								
$\widehat{\textit{Pat. Utilization Rate}}$		1.238**	1.991**	0.007	0.614**	0.143	0.300**	0.366**	0.676
		(0.595)	(0.982)	(0.173)	(0.303)	(0.234)	(0.140)	(0.146)	(0.574)
Cragg-Donald Wald F statistic	17.84								
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	22,106	22,106	22,106	22,106	22,047	22,047	22,089	22,072	20,916

Table 11. Robustness Checks

This table conducts robustness checks for the baseline results. In panel A, we further control for *#CW Patents/AT*, which is the citation-weighted number of granted patents that are applied for by a firm in a year scaled by the firm's book value of total assets at the beginning of the year, and *Patent Values/AT*, which is the economic values of granted patents that are applied for by a firm in a year scaled by the firm's book value of total assets at the beginning of the year. In panel B, we restrict firm-year observations to have at least one new product launch. All specifications include firm controls, firm fixed effects, and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

A. Controlling Patent Count and Patent Values

VARIABLES	(1) #New Products $_{t+1}$	(2) Sum CARs $_{t+1}$	(3) #Breakthrough Products $_{t+1}$	(4) Sales Growth $_{t+1}$	(5) MSG (SIC4) $_{t+1}$	(6) $\triangle$GPM $_{t+1}$	(7) $\triangle$ROA $_{t+1}$	(8) Log(MTB) $_{t+1}$
Pat.Utilization Rate	**0.674*** **	**0.462*** **	**0.452*** **	**0.022*** **	**0.024*** **	**0.007** **	**0.006** **	**0.033** **
	(0.053)	**(0.041)**	**(0.151)**	**(0.007)**	**(0.006)**	**(0.003)**	**(0.003)**	**(0.016)**
#CW Patents/AT	-0.015	0.111	0.079	-0.025	-0.021	0.014	-0.015	0.022
	(0.108)	(0.094)	(0.601)	(0.029)	(0.028)	(0.015)	(0.013)	(0.059)
Patent Values/AT	0.108**	0.128***	1.088***	0.040*	0.037*	0.036***	0.021***	0.247***
	(0.046)	(0.025)	(0.305)	(0.021)	(0.021)	(0.010)	(0.007)	(0.040)
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	23,397	24,915	6,192	27,621	27,621	27,667	27,638	26,277
Pseudo/Adj. R2	0.488	0.553	0.341	0.206	0.167	0.165	0.096	0.537

B. At least one new product launch

VARIABLES	(1) #New Products $_{t+1}$	(2) Sum CARs $_{t+1}$	(3) #Breakthrough Products $_{t+1}$	(4) Sales Growth $_{t+1}$	(5) MSG (SIC4) $_{t+1}$	(6) $\triangle$GPM $_{t+1}$	(7) $\triangle$ROA $_{t+1}$	(8) Log(MTB) $_{t+1}$
Pat.Utilization Rate	**0.576*** **	**0.383*** **	**0.386* **	**0.019** **	**0.022*** **	**0.007** **	**0.004**	**0.029**
	(0.061)	**(0.046)**	**(0.217)**	**(0.007)**	**(0.007)**	**(0.004)**	**(0.003)**	**(0.020)**
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	15,490	15,916	4,327	17,067	17,067	17,082	17,068	16,240
Pseudo/Adj. R2	0.440	0.502	0.381	0.291	0.236	0.235	0.125	0.569

Appendix A

Table A1. Variable Definition

Variables	Definition
Dependent Variables	
I (Core Technology Class)	An indicator that equal one if a patent belongs to a class in which the patent-owning firm has produced the most patents over the past 10 years, otherwise equals zero. **Source:** PatentsView
#Self Citations	The number of self citations received by a patent. **Source:** PatentsView
Self Citation Ratio	The number of self citations divided by the total number of citations received by a patent. **Source:** PatentsView
I (Sold)	An indicator that equals one if the patent is sold by the owning firm to another entity prior to the first patent maintenance date (four years after patent issuance), otherwise equals zero. **Source:** PatentsView
CAR(-1, 1)	Cumulative abnormal stock return within a three-day event window of (-1, 1) following the new product announcement event. **Source:** CRSP
Breakthrough Index	A text-based measurement that captures the novelty and impact of a product, following Kelly et al. (2021). Please refer to Section 3.2 of the paper for detailed technical construction of the variable. **Source:** PatentsView and Capital IQ Key Development Database.
1(Breakthrough Product)	An indicator that equals 1 if a product's breakthrough index (*BreakthroughIndex*) is above 95th percentile and equals zero otherwise. Please refer to Section 3.2 of the paper for detailed technical construction of the *BreakthroughIndex*. **Source:** PatentsView and Capital IQ Key Development Database.
#New Products	The number of new products (with three-day CARs above 80th percentile) a firm launches in a year. **Source:** CRSP and Capital IQ Key Development Database.
#Breakthrough Products	The number of breakthrough products (with the breakthrough index (*BreakthroughIndex*) above 95th percentile) a firm launches in a year. Please refer to Section 3.2 of the paper for detailed technical construction of the *BreakthroughIndex*. **Source:** CRSP, PatentsView and Capital IQ Key Development Database.
Sum CARs	The sum of all positive three-day cumulative abnormal stock returns of the new products that a firm launches in a year. **Source:** CRSP and Capital IQ Key Development Database.
Sales Growth	Natural logarithm of total sales for a firm in a year minus the natural logarithm of total sales for that firm in the previous year. **Source:** Compustat.
MSG(SIC4)	The sales growth of a firm in a year minus the industry (4-digit SIC) median sales growth in the same year. **Source:** Compustat.
△GPM	The change in the gross profit margin of a firm between year $t{+}1$ and t; the gross profit margin is defined as a firm's sales minus cost of goods sold, further divided by the firm's book value of total assets at the beginning of the year. **Source:** Compustat.
△ROA	The change in the return on assets (ROA) of a firm between year $t{+}1$ and t; the ROA is defined as a firm's operating income before depreciation divided by the firm's book value of total assets at the beginning of the year. **Source:** Compustat.

Log(MTB)	Natural logarithm of a firm's market value of assets divided by the book value of total assets. **Source:** Compustat.
Independent Variables	
I (Utilized)	An indicator that equals one if a patent is utilized in new products by the owning firm, otherwise equals zero. We define a patent is utilized in a new product if the patent-product pair text similarity score is above 80th percentile. **Source:** PatentsView
Log(1+#Patents Utilized$^{\text{Sum}}$)	Natural logarithm of one plus the total number of unique patents utilized in a new product (or a series of new products launched on the same date) by a firm. **Source:** PatentsView and Capital IQ Key Development Database.
Log(1+#Patents Utilized$^{\text{Average}}$)	Natural logarithm of one plus the average number of unique patents utilized in a new product (or a series of new products launched on the same date) by a firm. **Source:** PatentsView and Capital IQ Key Development Database.
Log(1+#Patents Utilized)	Natural logarithm of one plus the total number of patents utilized in a new product by a firm. **Source:** PatentsView and Capital IQ Key Development Database.
Pat. Utilization Rate	The number of patents that are utilized in products scaled by the total number of patents applied (and later granted) in the past five years for a firm in a year. We regard a patent as utilized in a product if the patent-product pair similarity is above 80th percentile. **Source:** PatentsView and Capital IQ Key Development Database.
Log(Product Text Length)	Natural logarithm of the new product announcement text description length. **Source:** Capital IQ Key Development Database.
Log(1+#New Products Launched)	Natural logarithm of one plus the number of new products that have been launched in the year. **Source:** Capital IQ Key Development Database.
#New Products/Sales	The number of new products (with three-day CARs above 80th percentile) a firm launches in a year divided by the sales of that firm in year *t-1*. **Source:** CRSP, Capital IQ Key Development Database and Compustat.
Firm Size	Natural logarithm of the sales of a firm in a year. **Source:** Compustat.
Log(Firm Age)	Natural logarithm of one plus the current year of observation minus the first year a firm appears in Compustat. **Source:** Compustat.
Leverage	The sum of a firm's current liabilities and long-term debt divided by the book value of total assets of the firm. **Source:** Compustat.
R&D Stock	The accumulated R&D expenses (XRD) of the firm over the five-year period ending in year t-2, assuming an annual depreciation rate of 20% following Hirshleifer et al. (2013). **Source:** Compustat.
Cash	A firm's cash holdings divided by the book value of assets. **Source:** Compustat.
Past Stock Return	Buy-and-hold stock return of a firm. **Source:** CRSP and Compustat.

Table A2. Selecting New Product Launch Announcement Texts

This table reports the keywords that are used to select the training sample of new product launch announcement texts (panel A), the FinBert's classification performance (panel B), and a randomly selected sample of new product launch headlines predicted by the fine-tuned FinBert model (panel C).

Panel A. Keywords about New Product Launches

launch, product, introduce, begin, unveil, release, debut and their variants.

Panel B. FinBert Classification Performance

	Precision	**Recall**	**F1-score**	**# Headlines**
Negative	0.94	0.90	0.92	139
Positive	0.92	0.95	0.93	161
Overall Accuracy			**0.93**	300
Macro Average	0.93	0.92	0.93	300
Weighted Average	0.93	0.93	0.93	300

Panel C. Randomly Selected New-Product-Launch Headlines that are Predicted by the Fine-tuned FinBert Model

Headline	**Company Name**	**Date**	**New Product Launch**
Sally Beauty Holdings, Inc. Announces Nationwide Launch of the Vernon Franois Collection	Sally Beauty Holdings, Inc.	2019-07-25	YES
Quest Diagnostics, Inc. Announces the Availability of OVA1 Blood Test to Aid Pre-Surgical Evaluation of Women for Ovarian Cancer	Quest Diagnostics Inc.	2010-03-09	YES
Thermo Fisher Scientific Launches 300mm FT-IR Metrology Tool	Thermo Fisher Scientific Inc.	2007-07-30	YES
Zoom(R) Modems Ship With ENERGY STAR(R) Qualified Adapters	ZOOM Technologies, Inc.	2007-11-13	YES
Lockheed Martin Offers Advanced Electro-Optical Targeting System for the F-35 Lightning II	Lockheed Martin Co.	2015-09-10	YES
Anthera Pharmaceuticals, Inc. Provides Clinical Program Updates for Blisibimod and Sollpura	Anthera Pharmaceuticals, Inc.	2016-06-28	NO
Acura Pharmaceuticals Provides Update on FDA Discussions Surrounding Development of Aversion Hydrocodone with Acetaminophen Tablet	Acura Pharmaceuticals, Inc.	2014-08-15	NO
MAIA Biotechnology, Inc. Announces HREC Approval in Australia for its THIO-101 Phase 2 Trial for NSCLC	MAIA Biotechnology, Inc.	2022-03-15	NO
Northern Vertex Mining Corp. Announces to Report Recent Results from Its Multi-Phase Infill and Resource Expansion Drilling Program At the Moss Mine in Nw Arizona	Elevation Gold Mining Co.	2021-06-10	NO
Delta Expands Trans-Pacific Service with Nonstop Shanghai-Atlanta Flight	Delta Air Lines, Inc.	2017-07-20	NO

Table A3. Overview of Within-Firm Patent-Product Linkage

This table demonstrates some randomly selected samples of within-firm patent-product linkages. Panel A shows three products and their 5 most (least) similar patents based on the patent-product similarity scores. For each patent-product linkage, we report the patent title, patent number, patent filing year, product name, product launching year, and the similarity score. In Panel B, we further report the excerpts of the three new product descriptions. For each product, we also report the most (least) similar patent text excerpt. The similarity score is estimated based on the *Word2vec* model which computes text similarities between patent text filings and product descriptions.

Panel A. Top-5 (green) and bottom-5 (red) patent-product linkages based on the pair similarity score

Patent Title	Patent Num. (Filing Year)	Product Name (Launching Year)	Similarity
Content analytics system configured to support multiple tenants	9183230 (2012)	IBM Watson Analytics (2014)	0.880
Concurrent execution of request processing and analytics of requests	8819183 (2009)		0.843
Automatic log sensor tuning	9507847 (2013)		0.821
Analytics platform spanning unified subnet	9342345 (2014)		0.791
Analytic solution integration	9098821 (2013)		0.786
Dynamic scan	8516318 (2010)		0.171
Immersion-cooled and conduction-cooled electronic system	8947873 (2012)		0.171
Dynamically reconfiguring time zones in real-time using plural time zone libraries	9740176 (2014)		0.207
Non-uniformity evaluation apparatus, non-uniformity evaluation method, and display inspection apparatus and program	8368750 (2009)		0.208
Land grid array interposer with compressible conductors	8672688 (2012)		0.211
Techniques to transfer data among hardware devices	11132326 (2020)	Nvidia GeForce RTX 3060 (2021)	0.770
Technique for sharing context among multiple threads	11080111 (2020)		0.741
Asynchronous data movement pipeline	11294713 (2020)		0.728
Graphics processing unit systems for performing data analytics operations in data science	11307863 (2019)		0.717
Real-time hardware-assisted GPU tuning using machine learning	10909738 (2018)		0.699
Cross talk reduction differential cross over routing systems and methods	10600730 (2018)		0.345
System and method for procedurally synthesizing datasets of objects of interest for training machine-learning models	10643106 (2018)		0.350
Three state latch	10009027 (2017)		0.385
System and method for cooperative game control	10252171 (2016)		0.387
Resistance and capacitance balancing systems and methods	10685925 (2018)		0.389
Multi-functional hand-held device	11275405 (2006)	Apple IPhone 4 (2010)	0.814
Establishing a video conference during a phone call	8744420 (2010)		0.814
Integrated touch screen	8390582 (2009)		0.813
In conference display adjustments	8502856 (2008)		0.813
Multipoint touchscreen	8125463 (2008)		0.812
Technique for reducing wasted material on a printed circuit board panel	8650744 (2010)		0.395
Low power peer detection circuit	8291241 (2009)		0.393
System and method for internet connected service providing heterogeneous mobile systems with situational location relevant content	8538685 (2007)		0.375
Methods and apparatus for shielding circuitry from interference	8071893 (2009)		0.372
Transaction ID filtering for buffered programmed input/output (PIO) write acknowledgements	8032673 (2009)		0.317

Panel B. Excerpts of patent and product Text

Patent Name & Excerpt	**Product Name & Excerpt**
1. Patent Name: Content analytics system configured to support multiple tenants (**Similarity Score:** 0.880) **Excerpt:** Techniques are disclosed for a software as a ***service (SaaS) provider*** to host a ***content analytics*** tool used to ***evaluate data collections*** for multiple customers (referred to as tenants) using one dedicated and expandable computing infrastructure, without requiring that the service provider obtain, install, license, and manage a separate copy of the content analytics tools for each tenant. ***Customers are provided access to resources dedicated to their enterprise***, but do not have access, or even awareness, of data collections or analytics resources hosted for other customers. That is, embodiments presented herein allow a provider to host content analytics tools used by customers to evaluate their enterprise data in a secure and timely manner. **2. Patent Name:** Land grid array interposer with compressible conductors (**Similarity Score:** 0.211) **Excerpt:** An electrical interconnect is provided for use within, for example, a land grid array (LGA) interposer such as a module-to-board connector. The electrical interconnect includes an electrically-conductive, compressible conductor which has a first conductor end portion and a second conductor end portion. The first and second conductor end portions physically contact in slidable relation each other with compression of the compressible conductor to facilitate inhibiting rotation of the compressible conductor. In one embodiment, the first end portion includes at least one first leg and the second end portion includes at least two second legs, and the at least one first leg and at least two second legs are interdigitated. Further, in one embodiment, the first end portion and the second end portion are each in slidable contact with an inner-facing surface of the compressible conductor.	**Product Name:** IBM Watson Analytics. **Excerpt:** IBM announced Watson Analytics, a natural language-based ***cognitive service*** that can provide instant access to powerful predictive and visual analytic tools for businesses. Watson Analytics is designed to make ***advanced and predictive analytics*** easy to acquire and use for anyone. The first release of Watson Analytics will include a freemium version of its cloud-based service designed to run on desktop and mobile devices. Watson Analytics offers a full range of ***self-service*** analytics, including access to easy to use ***data refinement and data warehousing services*** that make it easier for business users to acquire and prepare data - beyond simple spreadsheets - for analysis and visualization that can be acted upon and interacted with.
1. Patent Name: Techniques to transfer data among hardware devices (**Similarity Score:** 0.770) **Excerpt:** Apparatuses, systems, and techniques to route data transfers between hardware devices. In at least one embodiment, a path over which to ***transfer data from a first hardware component of a computer system to a second hardware component of a computer system*** is determined based, at least in part, on one or more characteristics of different paths usable to transfer the data. In at least one embodiment, first ***CPU*** is communicatively coupled with a first ***peripheral component interconnect (PCI) express (PCIe)*** switch, and second CPU is communicatively coupled with a second PCIe switch. A first ***graphics processing unit (GPU)***, designated at GPU 0, is coupled with third PCIe switch, and a second GPU, designated as GPU 1, is coupled with fourth PCIe switch. In at least one embodiment, memory can include various types of memory devices including ***graphics double data rate ("GDDR") memory***. **2. Patent Name:** Resistance and capacitance balancing systems and methods (**Similarity Score:** 0.389) **Excerpt:** Systems and methods that facilitate resistance and capacitance balancing are presented. In one embodiment, a system comprises: a plurality of ground lines configured to ground components; and a plurality of signal bus lines interleaved with the plurality of ground lines, wherein the interleaving is configured so that plurality of signal bus lines and plurality of ground lines are substantially evenly spaced and the plurality of signal bus lines convey a respective plurality of signals have similar resistance and capacitance constants that are balanced. The plurality of signals can see a substantially equal amount ground surface and have similar amounts of capacitance. The plurality of signal bus lines can have similar cross sections and lengths with similar resistances. The plurality of signal bus lines interleaved with the plurality of ground lines can be included in a two copper layer interposer design with one redistribution layer (RDL).	**Product Name:** Nvidia GeForce RTX 3060. **Excerpt:** NVIDIA Corporation announced that it is bringing the NVIDIA Ampere architecture to millions more PC gamers with the new GeForce RTX 3060 ***GPU***. When combined with a compatible motherboard, this advanced ***PCI Express*** technology enables ***all of the GPU memory to be accessed by the CPU at once***, providing a performance boost in many games. The RTX 3060's key specifications include: 13 shader-TFLOPs; 25 RT-TFLOPs for ray tracing; 101 tensor-TFLOPs to power NVIDIA DLSS (Deep Learning Super Sampling); 192-bit memory interface; 12GB of ***GDDR6 memory***.

1. Patent Name: Multi-functional hand-held device (**Similarity Score:** 0.814) **Excerpt:** The term "multi-functional" is used to define a device that has the capabilities of two or more traditional devices in a single device. The multi-functional device may, for example, include two or more of the following device functionalities: ***cell phone, music player, video player, game player, digital camera, handtop, Internet terminal, GPS or remote control***. The multi-functional ***hand-held device*** also incorporates a variety of input mechanisms, including ***touch sensitive screens***, touch sensitive housings, display actuators, audio input, etc. The device also incorporates a user-configurable GUI for each of the multiple functions of the devices. **2. Patent Name:** Transaction ID filtering for buffered programmed input/output (PIO) write acknowledgments (**Similarity Score:** 0.317) **Excerpt:** A PIO transaction unit includes an input buffer, a response buffer, and a control unit. The input buffer may receive and store PIO write operations sent by one or more transactons sources. Each PIO write operation may include a source identifier that identifies the transaction source. The response buffer may store response operations corresponding to respective PIO write operations that are to be transmitted to the transaction source identified by the identifier. The control unit may store a particular response operation corresponding to the given PIO write operation in the response buffer prior to the given PIO write operation being sent from the input buffer. The control unit may store the particular response operation within the response buffer if the given PIO write operation is bufferable and there is no non-bufferable PIO write operation having a same source identifier stored in the input buffer.	**Product Name:** Apple IPhone 4. **Excerpt:** Apple Inc. presented the new iPhone 4 featuring FaceTime, which makes the dream of ***video calling*** a reality, and Apple's stunning new ***Retina display***, the highest resolution display ever built into a phone, resulting in super crisp ***text, images and video***. In addition, iPhone 4 features a 5 megapixel ***camera*** with LED flash, ***HD video recording***, Apple's A4 processor, a 3-axis gyro and up to 40% longer talk time in a beautiful all-new design of glass and stainless steel that is the thinnest smartphone in the world. It can shoot high-definition video, catching up to some other smart phones. It has a gyroscope in addition to other sensors, to enable more advanced motion-sensing applications, such as ***games and mapping services***. The 3.5-inch screen is the same size as on previous models but features 326 pixels per inch, four times more pixels than the earlier iPhones.

Table A4. Drivers of Corporate Patent Utilization Rate

This table reports the regression results that investigate the drivers of corporate patent utilization rate. labor-shortage exposure. The dependent variable *Pat.Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. The independent variable *#Plaintiff Cases* (*#Defendant Cases*) is the number of patent litigation cases that a firm has involved as a plaintiff (defendant) in a year. *TNIC HHI* is Herfindahl-Hirschman index based on text-based network industry classification Hoberg and Phillips (2016). All specifications include firm and year fixed effects. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

VARIABLES	(1) Pat.Utilization Rate $_{t+1}$	(2) Pat.Utilization Rate $_{t+1,t+2}$	(3) Pat.Utilization Rate $_{t+1,t+3}$
#Plaintiff Cases $_{t,t-2}$	0.011***	0.012***	0.012***
	(0.003)	(0.003)	(0.003)
#Defendant Cases $_{t,t-2}$	-0.006***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)
TNIC HHI	-0.025	-0.026*	-0.025*
	(0.017)	(0.015)	(0.014)
#New Products/Sales	0.034*	0.020	0.008
	(0.018)	(0.023)	(0.019)
Firm Size	0.030***	0.025***	0.023***
	(0.007)	(0.006)	(0.006)
Log(Firm Age)	-0.078***	-0.067***	-0.057***
	(0.016)	(0.015)	(0.015)
Leverage	-0.062**	-0.065***	-0.067***
	(0.026)	(0.023)	(0.022)
R&D Stock	-0.021*	-0.015	-0.009
	(0.011)	(0.010)	(0.010)
Cash	0.013	0.012*	0.015**
	(0.009)	(0.007)	(0.006)
Past Stock Return	-0.004	-0.006**	-0.004*
	(0.004)	(0.003)	(0.002)
ROA	-0.047**	-0.022	-0.021
	(0.021)	(0.018)	(0.017)
Log(MTB)	0.003	0.005	0.006
	(0.004)	(0.004)	(0.004)
Sales Growth	-0.008	-0.004	-0.006
	(0.006)	(0.005)	(0.004)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Obs.	22,345	23,212	23,672
Adj. R2	0.513	0.654	0.710

Table A5. Alternative Measures of Firm Performance

This table reports the regression results that investigate the association between a firm's patent utilization rate and alternative measures of firm performance. The dependent variable *#Highly-Valued New Products* is measured as the number of highly-valued new products (with three-day CARs above 95th percentile) a firm launches in a year. *1(Highly-Valued New Products)* is an indicator that equals 1 if a firm launches at least one highly-valued new product (with three-day CARs above 95th percentile) in a year. *MSG (FF49)* is the difference between a firm's sales growth and the median sales growth of its Fama-French 49 industry peers in the same year. $\triangle OCF$ is measured as the change in the operating cash flow of a firm between year t and $t{+}1$; operating cash flow is defined as a firm's operating cash flow divided by the firm's book value of total assets at the beginning of the year. *Log(Tobin's Q)* is measured as the natural logarithm of a firm's book value of assets minus book value of equity plus market value of equity further divided by the book value of total assets. The independent variable *Pat.Utilization Rate* is measured as the number of patents that are applied (and later granted) in the past five years by a firm and that are utilized in the new products launched by the same firm in the current year, further scaled by the total number of patents applied (and later granted) in the past five years for that firm. We regard a patent as utilized in a product if the patent-product pair similarity score is above 80th percentile. All specifications include firm controls. In each panel, Columns 1 and 3 include industry and year fixed effects, while Columns 2 and 4 include firm and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Alternative Measures of New Products

	(1)	(2)	(3)	(4)
VARIABLES	#Highly-Valued New Products $_{t+1}$		1 (Highly-Valued New Product) $_{t+1}$	
Pat.Utilization Rate	1.587***	0.597***	0.226***	0.068***
	(0.072)	(0.090)	(0.009)	(0.008)
Model	Poisson	Poisson	OLS	OLS
Firm Controls	✓	✓	✓	✓
Industry FE	✓		✓	
Year FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	27,672	17,668	27,870	27,691
Pseudo/Adj. R2	0.243	0.350	0.132	0.254

Panel B. Alternative Measures of Market Share Growth

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	MSG (FF49) $_{t+1}$		MSG (FF49) $_{t+2}$		MSG (FF49) $_{t+3}$	
Pat.Utilization Rate	0.017***	0.023***	0.032***	0.036***	0.036***	0.034***
	(0.005)	(0.006)	(0.009)	(0.010)	(0.013)	(0.012)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓
Industry FE	✓		✓		✓	
Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,796	27,621	25,700	25,479	23,489	23,260
Adj. R2	0.037	0.175	0.053	0.337	0.061	0.457

Panel C. Alternative Measures of Firm Profitability Improvement and Valuation

	(1)	(2)	(3)	(4)
VARIABLES	$\triangle$OCF $_{t+1}$		Log(Tobin's Q) $_{t+1}$	
Pat.Utilization Rate	0.008***	0.005	0.071***	0.022**
	(0.002)	(0.003)	(0.016)	(0.010)
Model	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓
Industry FE	✓		✓	
Year FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	27,828	27,651	26,461	26,277
Adj. R2	0.042	0.074	0.112	0.592

Table A6. Control for Product Similarity

This table conducts robustness checks for the baseline results where we control for the average product similarity (*TNIC Product Sim.*) of a firm in a year. All specifications include firm controls, firm fixed effects, and year fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

VARIABLES	(1) #New Products $_{t+1}$	(2) Sum CARs $_{t+1}$	(3) #Breakthrough Products $_{t+1}$	(4) Sales Growth $_{t+1}$	(5) MSG (SIC4) $_{t+1}$	(6) $\triangle$GPM $_{t+1}$	(7) $\triangle$ROA $_{t+1}$	(8) Log(MTB) $_{t+1}$
Pat. Utilization Rate	**0.645*** **	**0.454*** **	**0.442** **	**0.023*** **	**0.025*** **	**0.007** **	**0.006** **	**0.035** **
	(0.056)	**(0.044)**	**(0.176)**	**(0.007)**	**(0.007)**	**(0.003)**	**(0.003)**	**(0.017)**
TNIC Product Sim.	-0.091	-0.245**	0.612*	-0.029**	-0.028**	0.001	0.008**	0.039
	(0.145)	(0.110)	(0.335)	(0.013)	(0.011)	(0.005)	(0.004)	(0.043)
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	20,558	21,869	5,664	24,368	24,368	24,411	24,383	23,100
Pseudo/Adj. R2	0.479	0.546	0.377	0.201	0.161	0.159	0.093	0.528

Table A7. Alternative Measures of Patent Utilization Rate

This table examines the robustness of the main results using alternative measures of patent utilization rate. We construct alternative measures of patent utilization rate using either 10-year patent portfolio of a firm or alternative patent-product pair score cutoffs (70th or 90th). In addition, we employ a 3-year moving average approach to generate an alternative measure of firm-level patent utilization rate. Moroever, instead of focusing on a 1-year (i.e., current year) patent-product incorporation rate over the past five-year patent portfolio window, we extend the patent usage window to three years. That is, we count the number of unique patents that have been incorporated into new products launched over the past three years (including the current year) from year *t-2* to year *t*, scaled by the total number of unique patents applied for and later granted by the firm from year *t-6* to year *t*. In Panel A, we report the results that investigate whether patents that are utilized in new product development by firms are more likely to belong to firms' core technology class. In Panel B, we examine whether those utilized patents receive more self citations, have higher self citation ratio, and are less likely to be sold. In Panel C (D), we report the results that investigate the relationship between the number of unique patents utilized in a new product and the product's announcement return (breakthrough index). In Panel E, we further investigate the relationship between a firm's patent utilization and the firm's new product development, product market performance, profit improvement, and firm value. All specifications include firm/product controls but are omitted for succinctness. Table A1 in Appendix A provides detailed variable definitions. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Utilized Patents and the Likelihood of Being within the Firm's Core Technology Class

	(1)	(2)	(3)
VARIABLES		I (Core Technology Class)	
		90th Percentile Cutoff	
I (Utilized)	0.039***	0.038***	0.036***
	(0.001)	(0.001)	(0.001)
Obs.	1,380,261	1,380,130	1,375,022
Adj. R2	0.367	0.383	0.407
		70th Percentile Cutoff	
I (Utilized)	0.032***	0.030***	0.030***
	(0.001)	(0.001)	(0.001)
Obs.	1,380,261	1,380,130	1,375,022
Adj. R2	0.366	0.382	0.406
		10-Year Patent Portfolio Window	
I (Utilized)	0.036***	0.035***	0.033***
	(0.001)	(0.001)	(0.001)
Obs.	1,380,261	1,380,130	1,375,022
Adj. R2	0.366	0.382	0.406
Model	OLS	OLS	OLS
Firm FE	✓	✓	
Class FE	✓		
Class-Year FE		✓	✓
Firm-Year FE			✓

Panel B. Utilized Patents, Self Citations, and the Likelihood of Being Sold

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	# Self Citations			Self Citation Ratio			I (Sold)		
				90 Percentile Cutoff					
I (Utilized)	0.066***	0.087***	0.086***	0.005***	0.005***	0.005***	-0.003***	-0.001	-0.001
	(0.005)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Obs.	1,349,952	1,318,303	1,218,706	1,375,093	1,369,142	1,328,513	1,375,093	1,369,142	1,328,513
Pseudo/Adj. R2	0.362	0.394	0.402	0.121	0.139	0.144	0.336	0.480	0.509
				70 Percentile Cutoff					
I (Utilized)	0.048***	0.088***	0.085***	0.005***	0.006***	0.006***	-0.004***	-0.002***	-0.002***
	(0.007)	(0.007)	(0.007)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	1,349,952	1,318,303	1,218,706	1,375,093	1,369,142	1,328,513	1,375,093	1,369,142	1,328,513
Pseudo/Adj. R2	0.362	0.394	0.402	0.121	0.139	0.144	0.336	0.480	0.509
				10-Year Patent Portfolio Window					
I (Utilized)	0.062***	0.079***	0.078***	0.006***	0.006***	0.006***	-0.004***	-0.001*	-0.001
	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs.	1,349,952	1,318,303	1,218,706	1,375,093	1,369,142	1,328,513	1,375,093	1,369,142	1,328,513
Pseudo/Adj. R2	0.362	0.394	0.402	0.121	0.139	0.144	0.336	0.480	0.509
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS	OLS
Firm-Class FE	✓	✓		✓	✓		✓	✓	
Class-Year FE	✓	✓		✓	✓		✓	✓	
Firm-Year FE		✓			✓			✓	
Firm-Class-Year FE			✓			✓			✓

Panel C. Number of Patents Utilized and New Product Announcement Return

	(1)	(2)	(3)	(4)
VARIABLES		CAR (-1,1)		
		90th Percentile Cutoff		
Log(1+#Patents Utilized$^{\mathrm{Sum}}$)	0.023***	0.022***		
	(0.007)	(0.008)		
Log(1+#Patents Utilized$^{\mathrm{Average}}$)			0.023***	0.021***
			(0.007)	(0.008)
Obs.	92,614	92,235	92,614	92,235
Adj. R2	0.036	0.054	0.036	0.054
		70th Percentile Cutoff		
Log(1+#Patents Utilized$^{\mathrm{Sum}}$)	0.029***	0.024**		
	(0.008)	(0.010)		
Log(1+#Patents Utilized$^{\mathrm{Average}}$)			0.028***	0.022**
			(0.008)	(0.010)
Obs.	92,614	92,235	92,614	92,235
Adj. R2	0.036	0.054	0.036	0.054
		10-Year Patent Portfolio Window		
Log(1+#Patents Utilized$^{\mathrm{Sum}}$)	0.030***	0.041***		
	(0.010)	(0.013)		
Log(1+#Patents Utilized$^{\mathrm{Average}}$)			0.029***	0.039***
			(0.010)	(0.014)
Obs.	92,614	92,235	92,614	92,235
Adj. R2	0.036	0.054	0.036	0.054
Controls	✓	✓	✓	✓
Industry FE	✓		✓	
Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓

Panel D. Number of Patents Utilized and Breakthrough Product

	(1)	(2)	(3)	(4)
VARIABLES	Breakthrough Index		1 (Breakthrough Product)	
		90th Percentile Cutoff		
Log(1+#Patents Utilized)	0.020***	0.005**	0.003**	0.001*
	(0.006)	(0.002)	(0.001)	(0.001)
Obs.	103,509	103,133	103,509	103,133
Adj. R2	0.053	0.354	0.125	0.374
		70th Percentile Cutoff		
Log(1+#Patents Utilized)	0.019***	0.007**	0.003**	0.002**
	(0.007)	(0.003)	(0.002)	(0.001)
Obs.	103,509	103,133	103,509	103,133
Adj. R2	0.052	0.354	0.125	0.374
		10-Year Patent Portfolio Window		
Log(1+#Patents Utilized)	0.026**	0.016***	0.005**	0.004***
	(0.010)	(0.006)	(0.002)	(0.002)
Obs.	103,509	103,133	103,509	103,133
Adj. R2	0.053	0.354	0.126	0.374
Controls	✓	✓	✓	✓
Industry FE	✓		✓	
Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓

Panel E. Patent Utilization Rate and firms' new product development, product market performance, profit improvement, and firm value

VARIABLES	(1) #New Products $_{t+1}$	(2) Sum CARs $_{t+1}$	(3) #Breakthrough Products $_{t+1}$	(4) Sales Growth $_{t+1}$	(5) MSG (SIC4) $_{t+1}$	(6) $\triangle$GPM $_{t+1}$	(7) $\triangle$ROA $_{t+1}$	(8) Log(MTB) $_{t+1}$
90th Percentile Cutoff								
Pat.Utilization Rate	0.713***	0.463***	0.565***	0.028***	0.028***	0.010**	0.008**	0.030
	(0.058)	(0.044)	(0.174)	(0.008)	(0.008)	(0.004)	(0.003)	(0.019)
Obs.	23,397	24,915	6,192	27,621	27,621	27,667	27,638	26,277
Pseudo/Adj. R2	0.488	0.551	0.335	0.205	0.166	0.161	0.095	0.533
70th Percentile Cutoff								
Pat.Utilization Rate	0.635***	0.447***	0.537***	0.022***	0.023***	0.006**	0.006**	0.023
	(0.050)	(0.040)	(0.143)	(0.006)	(0.006)	(0.003)	(0.003)	(0.015)
Obs.	23,397	24,915	6,192	27,621	27,621	27,667	27,638	26,277
Pseudo/Adj. R2	0.487	0.552	0.336	0.205	0.166	0.161	0.095	0.533
10-Year Patent Portfolio Window								
Pat.Utilization Rate	0.774***	0.565***	0.780***	0.022**	0.023***	0.007*	0.008**	0.038*
	(0.071)	(0.053)	(0.185)	(0.009)	(0.008)	(0.004)	(0.003)	(0.020)
Obs.	23,666	25,520	6,207	28,308	28,308	28,359	28,330	26,937
Pseudo/Adj. R2	0.495	0.562	0.341	0.201	0.163	0.159	0.092	0.530
3-Year Moving Average								
Pat.Utilization Rate	0.969***	0.601***	0.783*	0.037***	0.033***	0.008*	0.011***	0.022
	(0.089)	(0.068)	(0.430)	(0.011)	(0.010)	(0.005)	(0.004)	(0.031)
Obs.	24,665	26,439	7,332	29,363	29,363	29,417	29,388	27,899
Pseudo/Adj. R2	0.491	0.554	0.385	0.202	0.164	0.157	0.092	0.532
3-Year Patent-Product Incorporation over 5-Year Patent Portfolio								
Pat.Utilization Rate	0.650***	0.391***	1.273***	0.016**	0.014*	0.006	0.011***	0.018
	(0.067)	(0.054)	(0.289)	(0.008)	(0.007)	(0.003)	(0.003)	(0.020)
Obs.	23,344	25,062	6,904	27,840	27,840	27,891	27,863	26,492
Pseudo/Adj. R2	0.491	0.557	0.395	0.202	0.163	0.159	0.092	0.533
Model	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Table A8. Control for Industry-by-Time Fixed Effects

This table conducts robustness checks for the baseline results by replacing industry fixed effects with industry-by-time fixed effects to further account for time-varying industrial shocks. All specifications include product/firm controls. The odd-numbered columns control for industry-year (or industry-year-week) fixed effects, while the even-numbered columns control for both firm and industry-year (or industry-year-week) fixed effects. Table A1 in Appendix A provides detailed variable definitions. Robust standard errors clustered at the firm level (or firm and event-week level) are provided in parentheses. All financial variables are winsorized at the 1st and 99th percentiles. ***, **, and * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Number of Patents Utilized in New Products and NewProduct Announcement Return

	(1)	(2)	(3)	(4)
VARIABLES	CAR (-1,1)			
$\text{Log}(1+\#\text{Patents Utilized}^{\text{Sum}})$	0.020**	0.013		
	(0.008)	(0.009)		
$\text{Log}(1+\#\text{Patents Utilized}^{\text{Average}})$			0.019**	0.013
			(0.009)	(0.009)
Controls	✓	✓	✓	✓
Industry-Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	85,209	84,728	85,209	84,728
Adj. R2	0.050	0.066	0.050	0.066

Panel B. Number of Patents Utilized in New Products and Product Breakthrough

	(1)	(2)	(3)	(4)
VARIABLES	Breakthrough Index		1 (Breakthrough Product)	
Log (1+#Patents Utilized)	0.018***	0.005**	0.002	0.002**
	(0.006)	(0.002)	(0.002)	(0.001)
Controls	✓	✓	✓	✓
Industry-Year-Week FE	✓	✓	✓	✓
Firm FE		✓		✓
Obs.	96,333	95,868	96,333	95,868
Adj. R2	0.164	0.459	0.250	0.486

Panel C. Corporate Patent Utilization Rate and New Product Development

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	$\#\text{New Products}_{t+1}$		Sum CARs_{t+1}		$\#\text{Breakthrough Products}_{t+1}$	
Pat.Utilization Rate	1.738***	0.637***	1.568***	0.462***	1.378***	0.516***
	(0.065)	(0.048)	(0.057)	(0.045)	(0.122)	(0.176)
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Firm Controls	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	26,626	22,557	27,077	24,314	20,355	5,079
Pseudo R2	0.495	0.495	0.595	0.595	0.381	0.381

Panel D. Corporate Patent Utilization Rate and Firm Future Sales Growth

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Sales Growth $_{t+1}$		Sales Growth $_{t+2}$		Sales Growth $_{t+3}$	
Pat.Utilization Rate	0.017***	0.023***	0.032***	0.035***	0.035**	0.031**
	(0.005)	(0.007)	(0.010)	(0.010)	(0.014)	(0.013)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,596	27,412	25,495	25,264	23,287	23,047
Adj. R2	0.077	0.211	0.094	0.372	0.097	0.487

Panel E. Corporate Patent Utilization Rate and Firm Market Share Growth

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	MSG (SIC4) $_{t+1}$		MSG (SIC4) $_{t+2}$		MSG (SIC4) $_{t+3}$	
Pat.Utilization Rate	0.014***	0.024***	0.024***	0.032***	0.024*	0.029**
	(0.005)	(0.007)	(0.009)	(0.010)	(0.013)	(0.013)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,596	27,412	25,495	25,264	23,287	23,047
Adj. R2	0.012	0.152	0.022	0.318	0.025	0.440

Panel F. Corporate Patent Utilization Rate and Firm Future Profits and Value

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	$\triangle$GPM $_{t+1}$		$\triangle$ROA $_{t+1}$		Log(MTB) $_{t+1}$	
Pat.Utilization Rate	0.016***	0.018***	0.008***	0.006*	0.121***	0.029*
	(0.003)	(0.005)	(0.002)	(0.003)	(0.024)	(0.017)
Model	OLS	OLS	OLS	OLS	OLS	OLS
Firm Controls	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
Obs.	27,643	27,457	27,611	27,429	26,239	26,047
Adj. R2	0.086	0.120	0.041	0.114	0.217	0.595

Figure A1. (High-Quality) Patent Count per capita in the United States

This figure illustrates the (high-quality) patents per capita in the United States from 1976 to 2016. The blue line represents the number of total patents granted per capita; the purple line is the number of highly-cited patents (with citations above 95th percentile) granted per capita; the grey line denotes the number of breakthrough patents granted per capita, where breakthrough patents are defined by Kelly et al. (2021); the green line shows the number of creative patents granted per capita, where creative patents are measured by Kalyani (2022).

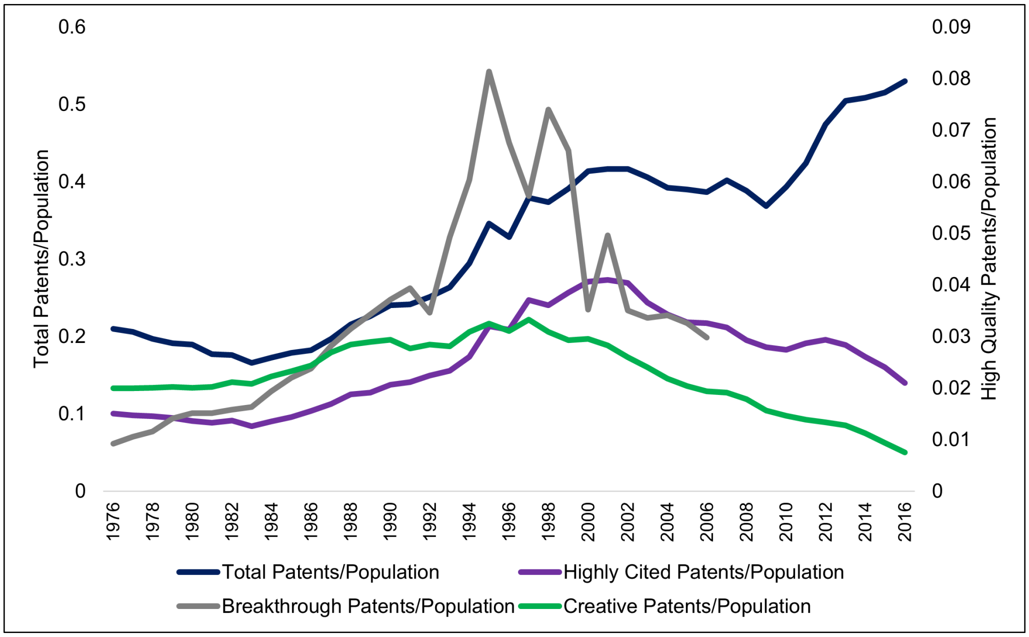


Figure A2. An Example of Patent Text Description

This figure illustrates the text description of the patent "Consistently-tight watch band" applied by Apple Inc. in 2016. The patent text description web page is sourced from Google Patent.

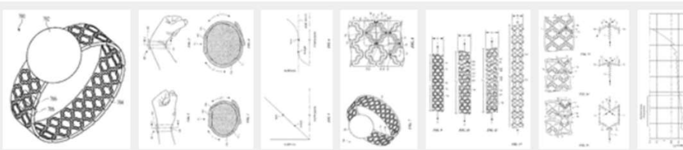
Google Patents   10455906

Consistently-tight watch band

Abstract

A watch band is disclosed. The watch band maintains a substantially constant tension throughout changes in its length while worn by a user. Such changes in length may occur automatically to accommodate changes in the size and circumference of a user's wrist as they move their wrist normally. By maintaining a constant tension, the watch band also maintains a constant force on the user's wrist, and causes a watch body attached to the watch band to also maintain a constant force on the user's wrist. This can increase a user's comfort, since the watch will not get tighter or constrict their wrist as they straighten and bend their wrist. It can also help optimize operation of any sensors in the watch band or watch body that benefit from being held against the user's wrist with a constant force, such as some physiological sensors.

Images (12)



Classifications

A44C5/0069 Flexible straps extensible

View 5 more classifications

Landscapes

Life Sciences & Earth Sciences

Health & Medical Sciences

Show more

US10455906B1
United States

Download PDF Find Prior Art Similar

Inventor: Yiwen Wu, Erik G. DE JONG, Colin M. Ely, Sameer Pandya

Current Assignee : Apple Inc

Worldwide applications

2016 US 2019 US

Application US15/269,710 events

2016-09-19	Application filed by Apple Inc
2016-09-19	Priority to US15/269,710
2016-11-03	Assigned to APPLE INC.
2019-10-02	Priority to US16/591,511
2019-10-29	Application granted
2019-10-29	Publication of US10455906B1
Status	Active
2037-09-14	Adjusted expiration

Info: Patent citations (14), Cited by (7), Legal events, Similar documents, Priority and Related Applications

External links: USPTO, USPTO PatentCenter, USPTO Assignment, Espacenet, Global Dossier, Discuss

Description

FIELD

The described embodiments relate generally to watch bands. More particularly, the present embodiments relate to watch bands that maintain a substantially constant tension when worn.

BACKGROUND

Watch bands may become tight around a user's wrist as the user moves their wrist. Such tightening can be uncomfortable.

SUMMARY

The present disclosure describes watch bands that maintain a substantially constant tension despite changes in their lengths while worn by a user. Such changes in length may occur automatically to accommodate changes in the size and circumference of a user's wrist as they move their wrist normally (e.g., moving

Claims (18)   Show Dependent

What is claimed is:

1. A watch band, comprising:

a first end for connecting to a watch body;

a second end for connecting to the watch body,

repetitive compliant mechanisms along the watch band between the first end and the second end, each of the repetitive compliant mechanisms being movable between a non-extended position and an extended position, each of the repetitive compliant mechanisms comprising two opposing spring segments connected at a pivot point, wherein the opposing spring segments form an angle less than 180 degrees in the non-extended position, wherein the opposing spring segments form an angle greater than 180 degrees in the extended position.

Figure A3. An Example of Product Text Description

This figure illustrates the text description of Apple Watch SE by Apple Inc. in 2020. The product text description is obtained from Capital IQ key development database.

S&P Capital IQ PRO

Apple Inc. | Key Development Details

NASDAQGS: AAPL (MI KEY: 4004205; SPCIQ KEY: 24937)

Apple Inc. Announces Apple Watch® SE

Apple Inc. announced Apple Watch® SE, packing the essential features of Apple Watch into a modern design — all at a more affordable price. The most advanced Retina® display allows customers to easily see more details and the information that matters most, right on their wrist. Apple Watch SE features the same accelerometer, gyroscope, and always-on altimeter as AppleWatch Series 6, and with the latest motion sensors and microphone, it offers robust health and safety capabilities including fall detection, Emergency SOS, international emergency calling, and the Noise app. With watchOS® 7, users can take advantage of powerful new features including Family Setup, which allows kids or older family members without an iPhone® to enjoy Apple Watch, plus sleep tracking, automatic handwashing detection, and new workouts. Apple Watch SE is available in three beautiful case finishes made of 100% recycled aluminum, and compatible with all Apple Watch bands including the new Solo Loop and Braided Solo Loop. Apple Watch SE features a Retina display, with thin borders and curved corners, that is 30% larger than Series 3. The interface allows for large and easy-to-read app icons and fonts, while complications are precise and informative. A variety of new watch faces are optimized for the display, so customers can easily view notifications, text messages, workout metrics, and more. With the S5 System in Package (SiP) and dual-core processor, Apple Watch SE delivers incredibly fast performance, up to two times faster than AppleWatch Series 3. The Digital Crown® with haptic feedback generates incremental clicks with an extraordinary mechanical feel as it is rotated. Apple Watch SE features the latest speaker and microphone, which are optimized for better sound quality for phone calls, Siri®, and Walkie-Talkie, along with Bluetooth® 5.0. The next-generation always-on altimeter provides real-time elevation all day long by using a new, more power-efficient barometric altimeter, along with GPS and nearby Wi-Fi networks. This feature allows for the detection of small elevation changes above ground level, up and down to the measurement of 1 foot, and can be shown as a new watch face complication or workout metric. The built-in compass provides users with better directions and compass headings, in addition to incline, elevation, and latitude and longitude. Users can add one of three new Compass complications to their watch face, and developers can take advantage of compass information in their apps to create rich experiences. With Emergency SOS on Apple Watch, customers can quickly and easily call for help and alert emergency services with just a push of a button. For added personal safety while traveling, users with cellular models of Apple Watch SE can complete international calls to emergency services, regardless of where the device was originally purchased or if the cellular plan has been activated. Fall detection uses a custom algorithm and the latest accelerometer and gyroscope in Apple Watch SE to detect when a user falls. By analyzing wrist trajectory and impact acceleration, Apple Watch sends the user an alert after a hard fall, which can be dismissed or used to initiate a call to emergency services. If the watch senses immobility for approximately 60 seconds after the fall, it will automatically call emergency services and play an audio message that provides the user's location as latitude and longitude coordinates, in addition to sending a message to emergency contacts. To provide enhanced insights into hearing health, Apple Watch SE utilizes the latest-generation microphone to measure ambient sound levels in a user's environment. Apple Watch SE sends a notification if the decibel level of surrounding sound has risen to a point that it could cause damage, and users can check noise levels at any time through the Noise app or Noise watch face complication.

Company Name	Apple Inc.
Event Date	16/09/2020
Announcement Date	15/09/2020
Development Type	Product-related Announcement
Source	Business Wire
Advisors	NA

Figure A4. Correlations of Similarity Scores by Different Word Embedding Models

This figure compares the cosine similarity scores for 250 randomly selected patent product pairs using the *FastText*, *OpenAI* (*text-embedding-3-small*), and *Glove* word embedding models.

Figure A4a. FastText & OpenAI Word Embeddings

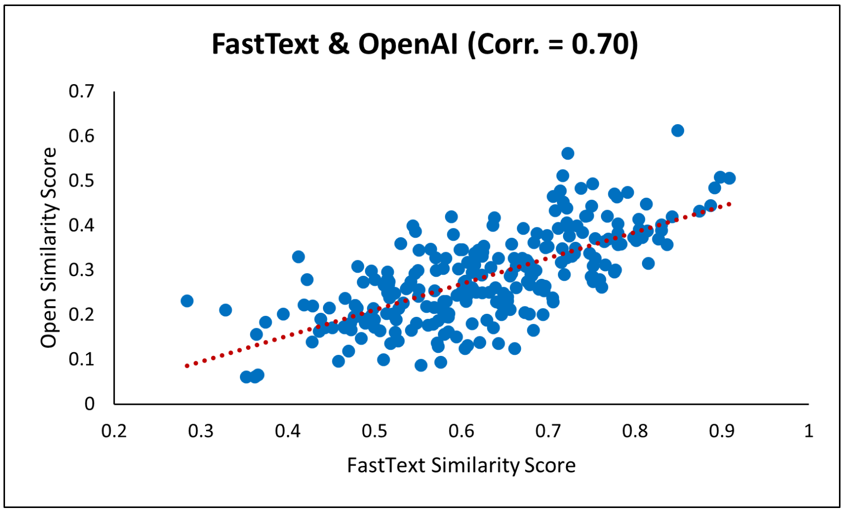
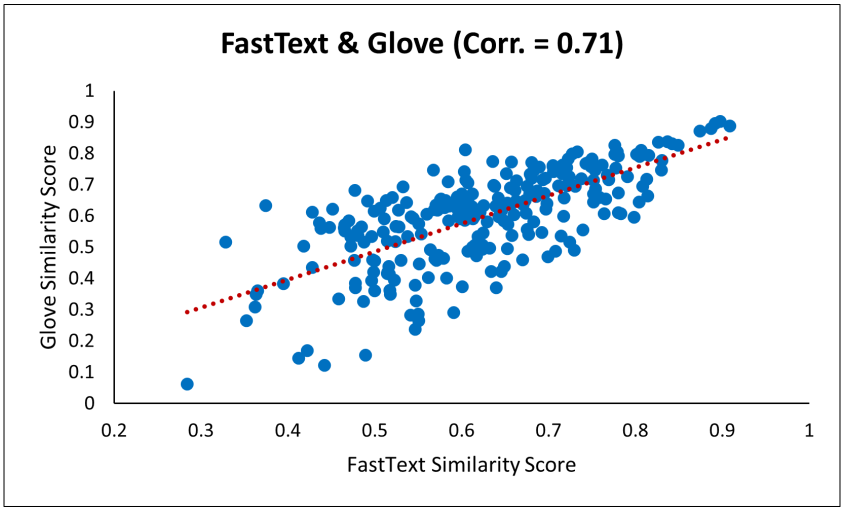


Figure A4b.FastText & Glove Word Embeddings



Appendix B Technical Details

In Appendix B, we discuss in detail the preprocessing steps for patent and product texts, the advantages of the Word2vec model, the measurement of patent-product pair similarity, the decision of whether a patent is utilized in a product, and the measurement of corporate patent utilization rate.

B1 Text Data Sources and Sample Construction

We first obtain patent filing text data from PatentsView, which text data for each patent granted since 1976. We then aggregate title, abstract, and description sections of each patent document into a patent-level corpus for textual analysis. To match patents with the U.S. publicly listed firms, we rely on the linking table developed by Kogan et al. (2017), which matches each patent assignee with a PERMNO ID from CRSP if available. Hence, our final patent text sample consists of 2,544,432 patents generated by the U.S. public firms from 1926 to 2022. Figure A2 illustrates an example of patent text filing from Google Patents website.

We further collect product-related text description data from the Capital IQ Key Development database. After merging each product-related text description with the U.S. publicly listed firms, we obtain 269,472 product-related announcements from 2002 to 2022. As suggested by Cao et al. (2018), there are generally four types of product-related announcements: R&D progress, new product introduction, product improvement, and product retirement. We follow prior studies to focus on the category of new product introduction. To select new product introduction-related announcements, we construct a list of keywords that are related to new product launches following Cao et al. (2018) and Mukherjee et al. (2017).[1] However, this keyword-discovery approach potentially suffers from two issues. First, if the keywords are of a narrow scope, we may not be able to fully capture all announcements that are related to new product launches (false negative). Second, it is also possible that some product-related announcements that we regard as new product launches may actually belong to other types of product announcements (false positive).

To improve the accuracy rate of our new product launch classification, we further employ an advanced natural language processing technique, *FinBert*, to help us automatically determine whether a product-related announcement is about new product introduction or not. Specifically, based on the new product launch keywords, we first construct a training sample that covers 3,000 randomly selected product announcement headlines, of which 1,500 headlines contain at least one of those keywords, and the other 1,500 headlines do not. We then manually read each of the 3,000 headlines to decide whether it is related to new product launch or not. With this training sample, we then fine-tune the *FinBert* model. Panel B of Table A2 tabulates the classification performance in the testing sample. Our fine-tuned *FinBert* model can accurately classify 93% of the headlines. Panel C further illustrates some (randomly) selected examples of new-product-introduction-related and non-related headlines predicted by our *FinBert* model. We then use this fine-tuned *FinBert* model to help us classify all the 269,472 product-related announcements. After requiring firms to have at least one patent granted throughout their histories, our final sample consists of 125,329 announcements related to new product launches. We also require our sample firms to have at least one new product launch in the key development

[1] Panel A of Table A2 lists the new product launches keywords.

database. Thus, our final sample contains 3,102 unique firms that have both produced patents and launched products. In Figure A3, we demonstrate an example of Apple Inc. announcing a new product in 2020.

B2 Preprocessing Text Data

We first remove all non-alphabetic characters, including numbers and punctuation marks, from both the patent filing text and the product announcement text. Next, we split the full text into lists of word tokens. Consistent with the natural language processing (NLP) literature, we further remove stop words from both the patent and product text documents. Stop words are those widely used in a language but contain little significant information. For example, some common stop words include articles (e.g., "the," "an"), prepositions (e.g., "in," "on") and conjunctions (e.g., "and," "but"). To construct the stop word list, we combine multiple sources that are commonly used in NLP: NLTK[2], Spacy[3], Scikit-learn[4], Bill Mcdonald Software Repository for Accounting and Finance[5], WebConfs[6], and MySQL[7]. The final list contains 938 unique stop words.[8]

After removing the stop words, we expect that a considerable proportion of the remaining words in the patent (product) text may provide little information for understanding the functions and characteristics of the patent (product). Thus, we follow Kogan et al. (2022) and Seegmiller et al. (2023) and retain only nouns and verbs, as these two syntactic terms likely contain more informative content. To identify the syntax of each word, we use the part-of-speech tagger package from NLTK (Natural Language Toolkit) in Python. Finally, we convert all the remaining words to lowercase and use the NLTK Lemmatizer package to lemmatize them. Lemmatization is a natural language processing technique that aims to reduce inflected forms of a word to one single form. For example, "running" and "ran" will be lemmatized to "run." After completing the preprocessing steps, we construct a cleaned list of word tokens for each patent and product text.

B3 Measuring Patent Utilization Rate

In this study, we assume that a patent is utilized in a new product within a firm if the text description of the patent filing is abnormally similar to that of the product announcement. Therefore, we first need to compute the textual similarity score between a patent filing and a product description text.

B3.1 Challenges in "Bag-of-Words" Approach

To compute patent-product pair textual similarity score, a conventional approach in economics and finance literature is the "bag-of-words" approach (see Gentzkow et al., 2019). Con-

[2] https://www.nltk.org/book/ch02.html
[3] https://github.com/explosion/spaCy/blob/master/spacy/lang/en/stop_words.py
[4] https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html
[5] https://sraf.nd.edu/textual-analysis/stopwords/
[6] https://www.webconfs.com/stop-words.php
[7] https://dev.mysql.com/doc/refman/8.0/en/fulltext-stopwords.html
[8] When cleaning the patent documents, we further filter out the following words that are commonly used in patent description: claim, present, invention, united, states, patent, description, background, and their variants.

sider two separate document D_i and D_j. In the "bag-of-words" approach, we convert each document into a one-hot vector V_i and V_j, with dimension equal to 1×N (N represents the number of unique words in these two documents). Each element of the vector, corresponding to a word, is set to zero if the word does not occur in the respective document, otherwise it is set to one. The two documents can thus be represented in two vectors, respectively. Next, we can compute the distance between the two documents using cosine similarity as follows:

$$Sim_{i,j} = \frac{V_i}{||V_i||} \cdot \frac{V_j}{||V_j||} \tag{1}$$

This traditional approach has been frequently employed in prior studies. For example, Hoberg and Phillips (2016) compute product similarity scores for the U.S. public firms by comparing the pairwise distance between their product description sections in 10-K filings. Chen and Srinivasan (2023) construct an industry-level AI technology exposure by comparing text similarity between AI patent abstract and the industry description from NAICS. Kelly et al. (2021) identify breakthrough patents by calculating their patent text similarities. However, the "bag-of-words" approach does not account for semantic similarities between words. That is, words could possess similar meanings even if they are in different forms. For instance, the word "big" is semantically similar to the word "large," but the "bag-of-words" approach will count as a zero match. Consider another extreme case: document i contains the phrase "one beautiful house," while document j contains the phrase "a lovely dwelling." As humans, we can discern the closeness of the two documents. However, when using the "bag-of-words" approach, we transform the two documents into two one-hot vectors, $V_i = [1, 1, 1, 0, 0, 0]$ and $V_j = [0, 0, 0, 1, 1, 1]$. Using Equation 1, we obtain a cosine similarity score of zero, indicating that the two documents are unrelated.

The underestimation bias could be even more pronounced when comparing two documents from different text sources that exhibit diverse language styles (Seegmiller et al., 2023). In this study, we aim to compare the formal, standardized, and legalistic language used in patent filing text descriptions with the more informal and less structured tone typically found in product announcement text descriptions. If we follow the "bag-of-words" approach, the contrasting language styles of the two corpora could lead to sparse one-hot vectors with many elements equal to zero. Consequently, this can result in an underestimated cosine similarity score that is close to zero. Moreover, as the corpus size increases (i.e., the number of unique words), the dimension of the one-hot vector also increases, significantly slowing down computational efficiency.

To summarize, the "bag-of-words" approach has two limitations: i) it fails to capture the semantics of words, and ii) it generates high-dimensional but sparse vectors that are computationally inefficient. To address these issues, we leverage on an advanced machine learning technique, *Word2vec* (Mikolov et al., 2013a), which can produce semantic, low-dimension, and dense word vectors via neural network. Seegmiller et al. (2023) have thoroughly discussed the advantage of *Word2vec* over the "bag-of-words" approach. They also replicate prior text-based measures, such as linking occupations with patents (Kogan et al., 2022), and find that *Word2vec* indeed outperforms the "bag-of-words" approach. We discuss more on the *Word2vec* model in the following subsection.

B3.2 *Word2vec* Model

The essence of the *Word2vec* model is based on the distributional hypothesis that "You should know a word by the company it keeps" (Firth, 1957), which suggests that the meaning of a word can be inferred from its neighboring words. For example, by comparing "I am majoring in Mathematics" and "I am majoring in Finance," we can easily understand that "Mathematics" and "Finance" both refer to specific subjects because they are surrounded by "I am majoring in." Recently, this linguistic concept has been incorporated into neural networks by Mikolov et al. (2013a), where a focal word is used to predict its neighboring words. The final product of *Word2vec* is a $N \times V$ parameter matrix: N denotes the dimension of a vector and V denotes the number of unique words in a corpus. This parameter matrix records the semantic vector representation of each word. Thus, Mikolov et al. (2013a) quantify words into dense and low-dimension vectors that also contain semantic information.[9]

Figure B1 illustrates a simple neural network framework for the *Word2vec* model. Specifically, a focal word X_c is first initialized as a $1 \times V$ one-hot vector in the input layer of the neural network, where V represents the number of unique words in the corpus (e.g., patent text). Next, we multiply X_c by W, a $V \times N$ parameter matrix where N denotes the dimension of the final word vectors, which generally varies from 50 to 1,000 depending on research interest.[10] In this step, the initial one-hot vector X_c is projected, from the input layer, into a $1 \times N$ vector V_c in the hidden layer.

[Please insert Figure B1 about here]

Next, V_c is further multiplied by the other $N \times V$ parameter matrix W', which produces the final $1 \times V$ vector, $Y_{c\text{-}m+j}$, where m denotes the window length of neighboring words.[11] We then use the Softmax function to transform $Y_{c\text{-}m+j}$, which is a vector of raw numbers, to a vector of probabilities that predicts the most likely neighboring word of the focal word X_c.[12] Let us call this vector of probabilities $S_{c\text{-}m+j}$.

To maximize the probability of predicting the correct neighboring word, we use the following likelihood function:

$$\begin{aligned} L(W, W') &= P(S_{c+j} | X_c) \\ &= \prod_{c=1}^{V} \prod_{\substack{-m<=j<=m \\ j \neq 0}} P(S_{c+j} | X_c; W, W') \end{aligned} \tag{2}$$

Note that W and W' are the two randomly initialized parameter matrices before the start of the model training process. When the training begins, each focal word in the corpus will be fed forward (i.e., from the input layer to the output layer) in the neural network, predicting

---

[9] *Word2vec* has two different model architectures to produce semantic word vectors. The first one is Continuous Skip-gram (SG), which uses the focal (center) word to predict its neighboring words. The other is Continuous Bag of Words (CBOW), which instead uses neighboring words to predict the focal (center) words. Please see Mikolov et al. (2013b) for more details.

[10] Intuitively, this $V \times N$ parameter matrix W, after model training, records the $1 \times N$ word vector for each of the V unique words. It hence reduces the sparse $1 \times V$ one-hot vector to a dense $1 \times N$ vector for each word.

[11] For example, when m equals to five, it means that this neural network will predict five words before and after the focal word X_c.

[12] The Softmax function restricts the vector of numbers to range from zero to one. The probability of each value in an element is proportional to the relative proportion of each value in the vector.

its neighboring words. It is common that the model will make prediction errors, that is, the forecasted neighboring words are not the ground truth. To reduce the errors, the model will then feed backward (also called backpropagation in machine learning domain) to fine tune the parameter matrices W and W'. After rounds of iterations, the prediction errors converge and the two parameter matrices become stable. The best parameters should maximize the probability in Equation B2. When the training process is completed, the *Word2vec* model will regard V_c as the focal word X_c's numeric vector (also called word embeddings). Intuitively, V_c is one of the V embeddings in the parameter matrix W. Each embedding has a $1 \times N$ dimension, in which the numeric values indicate the semantic information of the word.

As the *Word2vec* model can produce semantic word embeddings, it significantly alleviates the underestimation issue inherent in the "bag-of-words" approach. The dense and low-dimension word vectors also allow for more computationally efficient comparison between documents.

B3.3 *FastText*: An Improved Version of *Word2vec*

Despite the significant progress made by the *Word2vec* model in producing semantic vector representations for words in the vocabulary, there are still limitations: i) it does not provide vectors for words that are rare or out of the training corpus, and ii) it ignores the internal structure of words.[13] Since our text data originate from patent filings and product announcement texts, they likely contain extensive technological descriptions. However, many technical words are rarely seen or entirely absent in conventional training corpora. This can lead to no vector representations for those words when we use the pre-trained language model in later stages. The ignorance of technical words could potentially bias the patent-product pair similarity.

To overcome these challenges, we leverage *FastText* (Bojanowski et al., 2017), an extension of *Word2vec* model that takes into account subword information and also computes word vector representations for words that do not appear in the training corpus. Bojanowski et al. (2017) adopts a similar neural network structure and continuous skip-gram model as *Word2vec* to train *FastText*. But, instead of using a one-hot vector to represent each word in the input layer as outlined in Figure B1, *FastText* splits each word into n-grams (subword).[14] For example, the word "apple" can be split into 3-grams: "app," "ppl," and "ple." After neural network training, we obtain word embeddings not for the simple word "apple," but for each 3-gram "app," "ppl," and "ple." The final word embedding of "apple" will be represented as the sum of all these 3-gram word embeddings. Therefore, the advantage of *FastText* is that rare words or words that are out of the corpus can now be properly represented in semantic vectors by n-grams, as some of their n-grams are likely to appear in other words.

Bojanowski et al. (2017) and Mikolov et al. (2017) empirically examine the performance of the *FastText* model in different language tasks. They find that *FastText* outperforms other models such as the original *Word2vec* (Mikolov et al., 2013b) and *Glove* (Pennington et al., 2014). Thus, we leverage *FastText* to measure document similarity between patent filing text and product announcement text.

[13] Many English word formations follow rules, so morphologically similar words could share similar meanings.For example, the adjective "happy" and the noun "happiness," which are close in meaning, share the same root "happ" and differ only in their suffix. In English, the suffix "ness" generally indicates a noun.

[14] N-grams are all the combinations of adjacent letters with length n in a word.

B3.4 Measuring Patent-Product Pair Textual Similarity using *FastText*

We download the pre-trained English word vectors using the *FastText* model.[15] These 300-dimensional vectors are estimated using skip-gram model with default parameters as introduced in Bojanowski et al. (2017), where the training corpus is sourced from Wikipedia. These pre-trained word embeddings are well recognized, publicly available, and frequently adopted in the computer science domain. Using the publicly available word embeddings also increase the replicability of our results in this paper.

There are alternative word vectors trained on general corpus such as Common Crawl.[16] We choose to use word embeddings that are pre-trained using Wikipedia text as the training corpus because our patent and product text data are more related to scientific fields. General training corpora may not work well in capturing the meanings of technical words. Additionally, Wikipedia generally includes substantial parts of technical descriptions. Therefore, pre-trained word vectors derived from Wikipedia are more likely to capture semantic information closely aligned with the technical context.

After obtaining the word vectors for each word in our corpus (i.e., all unique words in patent and product text), we next aggregate these word vectors to document level using the following equation:

$$D_i = \sum_{X_j \in Z_i} w_{i,j} x_j \tag{3}$$

where D is a vector for document i, measured as the weighted average of the word vectors x for each word j in the set of words Z in document i. Following prior textual analysis literature (see, e.g., Loughran and McDonald, 2011; Li et al., 2021; Hoberg and Phillips, 2016; Kelly et al., 2021), we give different weights w on word vectors based on the importance of the words in our corpus. Consistent with the "bag-of-words" approach, we use the term-frequency-inverse-document-frequency (TFIDF) as our weighting scheme. Specifically, the TFIDF is calculated as:

$$TFIDF_{i,j} \equiv w_{i,j} \equiv TF_{i,j} \times IDF_k \tag{4}$$

The first component of the weight, term frequency (TF), is defined as follows:

$$TF_{i,j} = \frac{c_{i,j}}{\sum_j c_{i,j}} \tag{5}$$

where it counts the number of times word j appears in the document i, further divided by the total number of words in document i. *TF* thus captures the relative importance of a word in a document. Similar to Loughran and McDonald (2011), the second component of the weight, inverse-document-frequency (IDF), is measured as:

$$IDF_j = Log\left(\frac{\#\text{ Documents in the sample}}{\#\text{ Documents that include the word } c}\right) \tag{6}$$

Thus, if a word appears frequently across the set of documents, *IDF* will attenuate its impact using a log transformation. The product of *TF* and *IDF* in Equation B4 can then capture the

[15] The pre-trained word embeddings can be downloaded here: https://fasttext.cc/docs/en/pretrained-vectors.html.

[16] See https://commoncrawl.org/.

importance of a given word (vector) j in a given document i. In addition, as we use two distinct sets of text data (patent and product), the corpus of each text source could be significantly different. In this case, we follow the suggestion of Kogan et al. (2022) and Seegmiller et al. (2023) to compute the *IDF* for the words in patent and product text separately. This approach assures that, for instance, the word "patent" will be assigned a much lower weight if it appears in patent documents due to its common occurrence.

After obtaining a dense semantic vector for each document, we use the following equation to measure the cosine similarity between a patent document vector D_p and a product description text D_t within a firm f:

$$Sim_{p,t,f} = \frac{D_{p,f}}{||D_{p,f}||} \cdot \frac{D_{t,f}}{||D_{t,f}||} \tag{7}$$

Unlike Equation B1, Equation B7 emphasizes within-firm patent-product pair similarity because we want to measure a firm's self-invented patent utilization in its new product development. It is worth noting that for a firm's self-invented patents, we only focus on the firm's five-year patent application (later granted) portfolio before the launching date of a new product, since patents may become obsolescent as other technologies evolve (Ma, 2025).[17] The calculation process of within-firm patent-product pair similarity is illustrated in Figure 1. Suppose that firm A launches two products in 2015, NP1 and NP2. We then source the three patents (PAT1, PAT2, and PAT3) that firm A applied (and later granted) in the five years before 2015. For each patent-product pair, we compute its text similarity score using Equation B7.

B3.5 Determining Whether a Patent is Utilized in a Product

Critical to our study is the assumption that a patent is utilized in a new product if the patent-product pair similarity is abnormally high. We acknowledge that this assumption is strong, as high similarity may not definitively indicate utilization in the product. However, it does suggest that the product is very likely to have been heavily influenced by or derived from the patented technology.[18] In this regard, we are similar in spirit to the innovation literature that investigates knowledge diffusion across firms. Prior studies typically use patent citations to determine whether knowledge is diffused across firms (see., e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Singh and Marx, 2013; Arora et al., 2021; Fadeev, 2023).[19]

Moreover, the recent literature on innovation has categorized patents into process patents, which are inventions of new methods or processes that could improve firms' production efficiency, and non-process (product) patents, which generally refer to inventions of new or improved products (see, e.g., Bena et al., 2022; Bena and Simintzi, 2025). Process-related innovations are of less interest in our study, as these patents primarily focus on improving production

[17] The USPTO requires that for patent applications filed after June 8, 1995, the terms of patents will end 20 years after the patent application date. In robustness tests, we also consider the 10-year patent application (later granted) portfolio of a firm and obtain qualitatively similar results.

[18] Confirming whether a patented technology is indeed being utilized in a product poses a challenge, as it requires seeking advice from technical experts to confirm the usage of the patented technology.

[19] The literature assumes that if a patent of firm A cites a patent of firm B, knowledge is then spillovered from firm B to firm A. In a similar vein, Cohen et al. (2023) assume the utilization of a patent by a firm if the firm has cited the patent previously. Our criteria for patent utilization in products are more stringent, as we require extremely high text similarity between patents and products.

processes, while our focus is on whether product patents are utilized in new product development. Therefore, we follow the classification algorithm by Bena et al. (2022) to differentiate process and non-process patents.

Specifically, we define a patent as a process patent if the first claim and over 50% of patent claims are process claims (Bena et al., 2022; Bena and Simintzi, 2025). A patent claim is defined as a process claim if it contains words such as "*A method for . . .*" or "*A process for . . .*", followed by a verb.[20] For example, General Motor's patent "Method for automatic wireless replenishment using DTMF" (US7313382B2) is a process patent as its first claim "*A method for replenishing call-use authorization to a mobile vehicle ...*" is a process claim. After the classification, we retain patent-product pairs where the patents are non-process patents.

Furthermore, since a majority of patent–product pairs within a firm have low textual similarity scores and are considered unrelated to one another, we follow the prior literature (e.g., Kogan et al., 2022; Hoberg and Phillips, 2016) to impose a stringent criteria: we only regard a patent as being utilized in a product if the textual similarity score is above 80th percentile of our sample patent-product pair scores.[21] Table A3 demonstrates some (randomly) selected examples of within-firm patent-product pair linkage.

Finally, we measure a firm's patent utilization rate as the number of granted patents applied forby a firm in the past five years and utilized in the new products launched by the same firm in the current year, scaled by the total number of granted patents applied forby that firm in the past five years.

[20] We use the following keywords to identify process patent claims: "a method,""the method," "method for," "method of," "method in," "method, comprising, "method comprising," "method to," "method applicable to," "a process," "the process," "process for," "process according," "process in," and "process of."

[21] We also consider alternative percentile cutoffs such as 70th and 90th, and obtain qualitatively similar results.

Figure B1. Neural Network Framework for *Word2vec*

This figure presents a simple neural network for the *Word2vec* model. X_c is the focal word which is initialized as a $1\times V$ one-hot vector in the input layer. V represents the number of unique words in the corpus. $W_{V\times N}$ is a $V\times N$ parameter matrix, where N is a dimension of interest which generally varies from 50 to 1000. V_c is a projected vector with the size of $1\times N$ in the hidden layer. $W'_{N\times V}$ is the other parameter matrix with the size of $N\times V$. $Y_{c\text{-}m}$, $Y_{c\text{-}m+1}$, and $Y_{c\text{-}m+2}$ are $1\times V$ prediction vector in the output layer.