

Article

Linking Macro Policy Shifts to Technology Evolution: Integrating a Difference-in-Differences Model with Patent Keyword Networks

Taeyeoun Roh ¹, Hyejin Jang ^{1,2} and Byungun Yoon ^{1,*}

¹ Department of Industrial & Systems Engineering, Dongguk University, 3-26, Pil-dong 3-ga, Jung-gu, Seoul 100-715, Republic of Korea; arongi1320@naver.com (T.R.); hj@hufs.ac.kr (H.J.)

² Division of AI Data Convergence, Hankuk University of Foreign Studies, Yongin-si 17035, Republic of Korea

* Correspondence: postman3@dongguk.edu; Tel.: +82-2-2260-8659; Fax: +82-2-2260-8743

Highlights

Please indicate how your work links to systems science via your contributions to systems practice, theory, and/or methodology.

- Methodologically, the study couples a difference-in-differences design for policy exposure with a patent keyword network, yielding a technology change index (new elements, new links, changing link intensity) that makes structural change inside a technological system measurable rather than inferred from filing counts alone.
- For systems practice and theory, it moves policy evaluation from detecting whether an effect exists to characterizing what the change consists of, and folds external-environment analysis and R&D item selection into one procedure that can be reproduced for other policies and domains.

What are the main findings and/or the implications of the main findings?

- In the battery industry around the Paris Agreement, post-policy change is heterogeneous across battery sub-technologies, and it persists whether or not covariates are included—though the two specifications disagree on which sub-technology leads.
- Decomposing that change into new elements, new combinations, and intensified existing links points each field to a different action: collaborate on shared foundational elements, or secure proprietary performance-critical ones.

Abstract

Macro policies such as the Paris Agreement are key external drivers of research and development (R&D) planning. Measuring their impact and reflecting the resulting technological changes in the R&D process are essential for sustaining competitive advantage. This study proposes a methodology for measuring the association between macro-level policy shifts and micro-level technological change. The proposed methodology integrates a difference-in-differences (DID) model with a patent keyword network to construct a technology change index that captures the emergence of new technological elements, the formation of new relationships, and changes in the intensity of existing relationships. A case study examining the impact of the Paris Agreement on the battery industry reveals heterogeneous policy sensitivity across different battery technologies—a heterogeneity that holds under both a covariate-adjusted and a covariate-free specification, although which technology carries the largest differential depends on that choice—and identifies, for each technology, the technological elements associated with post-policy innovation. These elements are further classified into firm-specific proprietary components and shared



Academic Editor: Luna Leoni

Received: 8 August 2026

Revised: 18 September 2026

Accepted: 24 September 2026

Published: 4 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

foundational components, providing insights for technology development strategies. By linking macro-level policy shocks to micro-level technological evolution, the proposed approach provides strategic implications for firms' R&D planning and supports more targeted technology and innovation policies.

Keywords: R&D planning; macro policy; difference in differences; keyword network; policy sensitivity

1. Introduction

R&D strategy comprises stages that involve identifying the internal and external environment of a company and executing plans to achieve a competitive advantage. External environment analysis involves examining the relationships between various factors that can impact R&D, such as policies, competition, and society. Among these, technology policy is the most influential factor in the R&D planning stage as it has a multifaceted effect not only on technology but also on economic and social aspects [1]. Therefore, objectively evaluating the impact of policies is essential [2,3] as even the same policy may have different effects depending on the specific technology sector [4]. While there exist sectors where significant changes occur in the R&D environment due to macro policies such as the Paris Climate Agreement, there also exist sectors, such as semiconductors, artificial intelligence (AI), and software, where overcoming technical limitations and improving performance are the main goals. These technology sectors, where policies are not as significant, may be less sensitive to policy changes, which would allow for aggressive R&D strategies disregarding the influence of policies.

To analyze the relationship between policies and technology, the influencing factors are identified and examined using two main approaches. First, methodologies such as system dynamics and computable general equilibrium (CGE) are utilized to examine the relationships between these factors. System dynamics analyzes the relationships between various factors and suggests measures that either promote or hinder these relationships to achieve strategic objectives [5,6]. The CGE model is used to simulate assumptions and interactions related to environmental, social, and economic activities, enabling the assessment of policy impacts and the proposal of subsequent policies and government actions [7,8].

In studies that analyze the quantitative relationships among factors influenced by policies within specific technology sectors or industries, regression analysis based on difference-in-differences (DiD) [9] is the most widely used method. Various quantitative indicators are employed as dependent variables in DiD-based regression models to interpret the development of the selected technologies or industries. These include changes in the scale of related industries, increments in the sales revenue of relevant companies, growth in the number of patents filed, and reductions in carbon emissions. The independent variables include factors such as the policy itself, increments in R&D investment, growth in the number of relevant scientists, and government subsidies, as these factors can impact technology development. The results from various types of regression models are then analyzed to interpret whether the policy and independent variables have a statistically significant impact, which, in turn, informs policy recommendations or the establishment of R&D strategies.

In both system dynamics and CGE-based policy impact analysis, patents are utilized as indicators representing technology or industries. As patents are recognized by national patent offices for their novelty, industrial applicability, and inventiveness, they constitute

relatively objective and standardized measures. Patents contain quantitative information, such as the number of inventors, number of citations, and number of claims, as well as qualitative information in the form of text, including the patent's title, abstract, and claims. Patents explicitly highlight the key elements of technologies in the form of keywords within their qualitative information, which enables systematic identification and specification of policy-related technologies through well-structured patent queries tailored to capturing these critical elements.

Studies examining policy impacts have largely relied on aggregate quantitative indicators—such as carbon emissions, energy efficiency, or patent counts—to assess relationships between policies and technological sectors. While these measures capture macro-level outcomes, they provide limited insight into how policies shape the internal structure and evolution of technologies. Some studies further restrict their contribution to proposing broad directional recommendations, such as enhancing or reducing interactions among technological elements, without offering concrete mechanisms for implementation. As a result, when policy effects appear statistically weak or insignificant, such approaches fail to generate meaningful analytical or strategic implications.

In addition, patent-based analyses often focus solely on the number of applications as a proxy for policy influence, neglecting the qualitative dimensions embedded in patents, including technological principles, novelty, inventiveness, and industrial applicability. Moreover, prior research has more often focused on large-scale technological trends with broad industrial impact—such as the commercialization of artificial intelligence or the expansion of hydrogen fuel cell technologies—rather than examining how specific technological principles evolve or how new technological elements emerge to achieve concrete technical objectives.

This study interprets patent applications as indicators of technological advancement and quantifies policy–technology linkages by integrating a difference-in-differences (DID) design with a patent keyword network. The network maps keywords to technology elements and tracks structural, time-varying changes; the DID captures exogenous policy exposure. These measures enter a regression framework to estimate the effect of policy on technological change, yielding evidence to inform R&D strategy. The primary contribution of this study is methodological: it provides a unified framework linking policy exposure to structural change in technology elements, rather than a definitive causal estimate for a single policy or a prescriptive strategy recommendation.

The remainder of this paper is organized as follows. Section 2 discusses the DID approach for quantifying policy impact and the keyword-network methodology for analyzing changes in technology elements represented in patents. Section 3 presents a framework for analyzing and interpreting changes in technology elements and the impacts of policies. Section 4 describes the application of the proposed methodology in a case study to assess the impact of the Paris Climate Agreement on the battery industry. Finally, Section 5 discusses the results of the proposed methodology, its limitations, and directions for future research.

2. Background

2.1. Macro and Micro R&D Policy

R&D policy can be classified into macro policy and micro policy, depending on its scope and the extent of its impact. Macro policies are those that can impact entire nations or industries. Rather than targeting specific countries or industries, macro policies are designed to achieve international objectives such as reducing unemployment, stabilizing prices, and protecting the environment. The entities potentially impacted by policies include industries, the economy, and businesses. Depending on how each factor is defined,

methods such as system dynamics and CGE are employed to identify and simulate the interactions between multiple policy-affected factors.

You and Jung [6] employed system dynamics to emphasize the importance of an adequate evaluation period for properly assessing and refining macro policies. Liu and Xiao [10] demonstrated that in addition to economic factors, concrete macro policy support is essential for successful technology diffusion. Miyata et al. [11] simulated economic and environmental policies from a long-term perspective to evaluate resource allocation and interactions between industries.

By contrast, micro policies aim to address specific issues within individual industries or companies, focusing on a single element among the ones included within macro policies, such as industry, the economy, or society [12]. Micro policy studies primarily focus on utilizing clear, quantifiable indicators aligned with policy goals, such as changes in carbon emissions due to regulations [13] and variations in R&D investment resulting from tax incentives [14,15].

Research examining micro-level policy impacts often focuses on technology-specific instruments or statistical hypothesis testing, providing limited insight from a broader technology development perspective. Even when macro-level policy effects are identified, such analyses rarely indicate which specific technological elements should be prioritized for strategic investment. For example, in promoting a hydrogen fuel cell society, it remains unclear whether priority should be given to expanding refueling infrastructure, improving hydrogen refueling efficiency, or strengthening safety-related technologies. Without identifying the concrete technological elements requiring focused investment, policy impact studies offer limited guidance for R&D strategy formulation.

To address this limitation, this study conceptualizes diverse micro policies as components of a unified macro policy environment and employs a difference-in-differences (DID) approach to detect structural shifts in technological development. Importantly, the analysis goes beyond estimating aggregate policy effects by integrating changes in keyword-based patent networks to capture the evolution of specific technology elements. By jointly examining policy-level shifts and structural changes in technology elements, the study not only assesses the magnitude and direction of policy influence but also identifies concrete technological priorities for effective R&D planning.

2.2. Policy Impact Analysis Based on DID

Technology policy covers a wide range of factors that affect companies, society, and the economy. Research in this field often focuses on quantitatively interpreting the changes in inputs and outputs resulting from policy interventions. System dynamics and CGE models define the factors influenced by policies and highlight potential directions for R&D planning that either promote or hinder these relationships to achieve the intended competitive technological advantage. Specific action plans for the proposed directions are developed using qualitative methods, such as expert workshops, surveys, and analysis of international cases [16]. Additionally, the relationships before and after policy implementation are quantitatively analyzed to evaluate the effectiveness of the implemented policies [17].

The DID approach is actively utilized to analyze the relationship between policies and specific factors. DID distinguishes between groups affected by a particular event and those not affected by it, dividing the timeline into pre-event and post-event periods. This methodology separates the data into two sets, as illustrated in Figure 1, where δ represents the difference between the observed trend and the trend that would have been maintained without the implemented policy. This difference serves as a statistical measure to assess the policy's impact. The basic regression model based on this approach is presented in Formula (1) and is generalized for broader application in Formula (2).

$$y = \beta * \text{Dummy } G + \gamma * \text{Dummy } T + \delta * \text{Dummy } G * \text{Dummy } T + a_1 * x_1 + a_2 * x_2 + \dots + \varepsilon \text{ where } \text{Dummy} \in \{1, 0\} \quad (1)$$

Formula (1) Basic DID Regression

$$\text{Dependent Variable} = \beta * \text{Dummy } G + \gamma * \text{Dummy } T + \delta * \text{Dummy } DID + \text{Independent Variable} + \varepsilon \quad (2)$$

Formula (2) Generalized DID Regression.

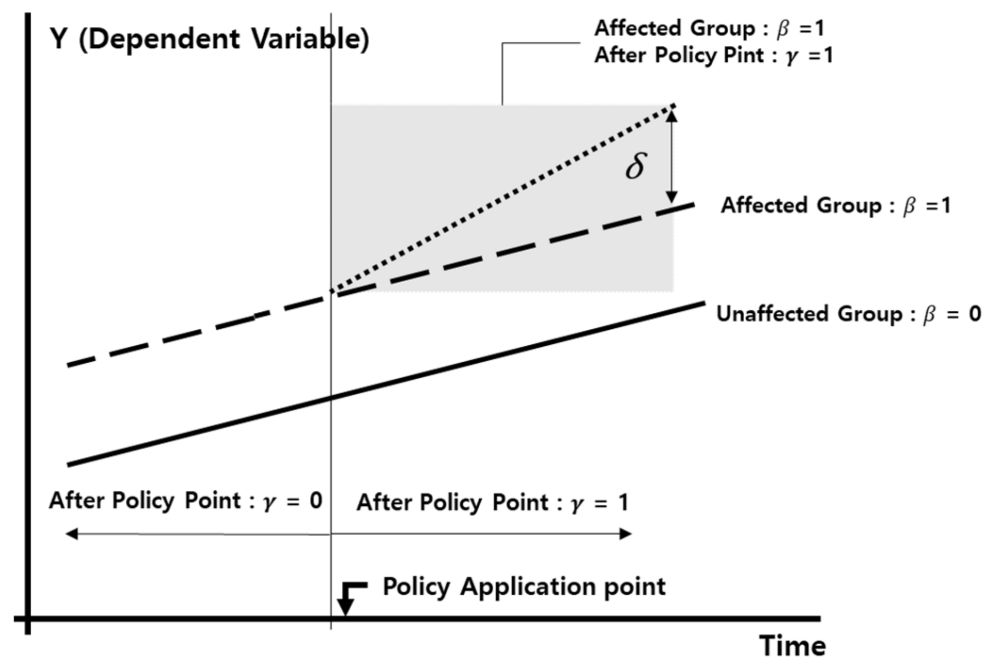


Figure 1. Conceptual diagram of DID.

Leveraging the ability of a dummy DID variable—a single quantitative value—to represent various changes resulting from a specific event, Büttner et al. [18] interpreted an increase in the number of cited patents as an outcome of knowledge diffusion and defined it as the dependent variable. They set the development of specialized patent translation tools, the number of International Patent Classification (IPC) classes, bibliographic information on domestic patent citations as independent variables to verify the impact of patent translation tool development on knowledge diffusion. Pippel and Seefeld [19] statistically examined the impact of R&D collaboration on product innovation with the number of innovative products as the dependent variable and the number of R&D collaborations between firms and investment indicators for each company as independent variables. To evaluate the effectiveness of patent incentive policies, Gu [20] used the number of patent applications as the dependent variable and the number of students in research institutes, the number of related companies, and the number of commercialized patents as independent variables. Gao and Yuan [21] investigated the role of the National Innovation City policy in economic growth deceleration and increased environmental pollution with the concentration of R&D personnel and urban development as independent variables.

Patents are widely employed as dependent variables associated with technology advancement as they provide objective and quantifiable information. For instance, de Rassenfosse and Palangkaraya [22] analyzed the impact of patent pledges on technology advancement based on the growth rate of pledged patents, the number of citations, and other related indicators. Li, Xue, and Qin [23] investigated the changes in environmental

technological innovation resulting from environmental information disclosure by defining related patents as independent variables and utilizing the Pollution Information Transparency Index as an indicator. Li et al. [23] analyzed the impact of patent disclosure on technology diffusion based on how citation counts, the proportion of early-published patents, and patent attorney involvement affect patent applications.

While studies have adopted patents as dependent variables, they have focused primarily on the number of patent applications. They have only categorized patents by their respective fields, without examining the specific information contained within the patents themselves. Moreover, although the quantitative aspects of patents—such as the number of applications, number of inventors, and number of cited patents—have been used as independent variables, the core textual information in patents has not been sufficiently considered in assessing the impact of patents on technology advancement.

Therefore, the growth in the number of patent applications is defined as an indicator of technology advancement, and the textual information within patents is considered a key technology element. Subsequently, macro policies are assumed to impact specific technological fields based on their implementation timing and are reflected in the dummy variable *T*. Consequently, a regression model based on the DID approach is employed, with changes in technology elements, changes in patents, the pre- and post-policy implementation period, whether the technological field is affected by the policy, and the degree of policy influence as independent variables, and technological advancement as the dependent variable. This model is used to statistically verify the influence of each independent variable.

2.3. Technology Policy and Patents

Technology policy plays a critical role in guiding industrial growth and technological advancement, as well as setting the direction for R&D efforts. For the successful formulation of technology policy, it is essential to secure objective supporting evidence and assess the outcomes of policy implementation. In this regard, patents serve as objective records granted by national patent offices based on novelty, inventive step, and industrial applicability. They represent formalized outputs from major R&D performers and are therefore widely used as reliable data sources for both the formulation and evaluation of technology policy.

Patents can be utilized in the policy-making process from two perspectives: as a basis for formulating technology policy and as an outcome indicator for evaluating policy effectiveness. Altuntas [24], Petralia [25], and Yuan and Li [26] each analyzed patent data in the fields of renewable energy, artificial intelligence, and hydrogen fuel cell vehicles, respectively, identifying slower technological progress in these areas compared to other industries. Based on their findings, they highlighted the necessity of policy support in these specific sectors. Additionally, Fabrizi et al. [27] conducted a comparative analysis of the impact of university–industry collaboration and policy instruments on patents, focusing on energy efficiency and green technologies. Their findings provided empirical evidence for establishing effective policies to strengthen collaborative R&D activities. Beyond cataloging these two roles, however, prior work rarely examines how the textual content of patents, rather than their counts, changes under policy; the present study addresses this gap, and the proposed NTE, NTF, and VTF indicators are intended to capture the emergence, recombination, and shifting emphasis of technology elements rather than the novelty, recombination, or centrality measures emphasized in earlier keyword-network studies.

Meanwhile, several studies have actively employed patents as indicators for evaluating the outcomes of technology policies. Long and Wang [28] found that China's patent incentive policies led to an increase in the number of patent applications, but had limited impact on the quality of those patents, thereby calling for a reconsideration of policies

overly focused on quantitative indicators. Donis et al. [29] analyzed the effects of policy types and institutional capacities on patent performance in the green technology sector across OECD countries, emphasizing the importance of designing policy instruments that are aligned with both technological characteristics and national contexts.

While previous studies have largely focused on quantitative aspects of patents, such as application and grant counts, they have made only limited use of the rich textual information embedded in patents, such as technical content and innovation principles. To overcome this limitation, recent research has introduced advanced text-mining techniques. In this study, we employ a methodology that integrates natural language processing (NLP) with network analysis to analyze patents as policy-influenced technological outputs.

2.4. Evaluation of Changes in Patent Contents Through Natural Language Processing and Network Analysis

A patent is certified technical information that has been approved by national patent offices based on three key factors: novelty, inventiveness, and industrial applicability [30]. These three elements are documented in the text of the title, abstract, and claims. Numerous studies have adopted text-mining techniques to analyze this technical information. Park, Yoon, Kim, and Seol [31] explored promising technology fields in the AI domain by applying keyword topic modeling to patent titles and abstracts. Roh, Jeong, Jang, and Yoon [32] sought to identify customized technology opportunities by analyzing keywords related to user requirements within patents. Yang and Yu [33] investigated technological risks arising from legal disputes by examining patent claims.

However, text information is not analyzed merely at the word level; instead, research is being conducted to capture more nuanced meanings by identifying structures or connections within the text. First, NLP techniques are employed to derive and utilize meaningful semantic connections by analyzing the grammatical functions of words within sentences through part of speech (POS) tagging [34]. A widely used approach involves employing the Subject–Action–Object (SAO) structure. The SAO structure consists of the subject, verb, and object of a sentence. In this context, the “S” (subject) can be interpreted as the Solution and the “AO” (action–object) as the problem, which forms a problem–solution structure. This approach allows for interpreting the SAO structure as a framework involving technology elements that represent “methods for solving problems.” Wang et al. [35] mapped the SAO structure onto the seven layers of a technology roadmap. Park, Yoon, and Kim [36] analyzed patent infringement by calculating similarities between patents from a problem–solution perspective.

Second, from a network perspective, complex meanings have been derived from connections among words by examining groups of words at the phrase, sentence, paragraph, and document levels. Figure 2 illustrates the process of connecting words while excluding articles (such as “the”) and prepositions (such as “to”) from two sentences describing a battery. The resulting network combines these two sentences to visualize their connections. The connected words in the keyword network can be interpreted as being semantically related. In the combined network, “battery” and “energy,” which are commonly connected in both sentences, are represented with thicker edges than in other connections. This suggests that the relationship between these two terms holds greater importance than the connections between other words. Additionally, while “rechargeable” is not directly connected to “supply” in the diagram, both are linked to “battery.” This indicates an indirect connection between these two words, indicating a relationship from a technology-element perspective.

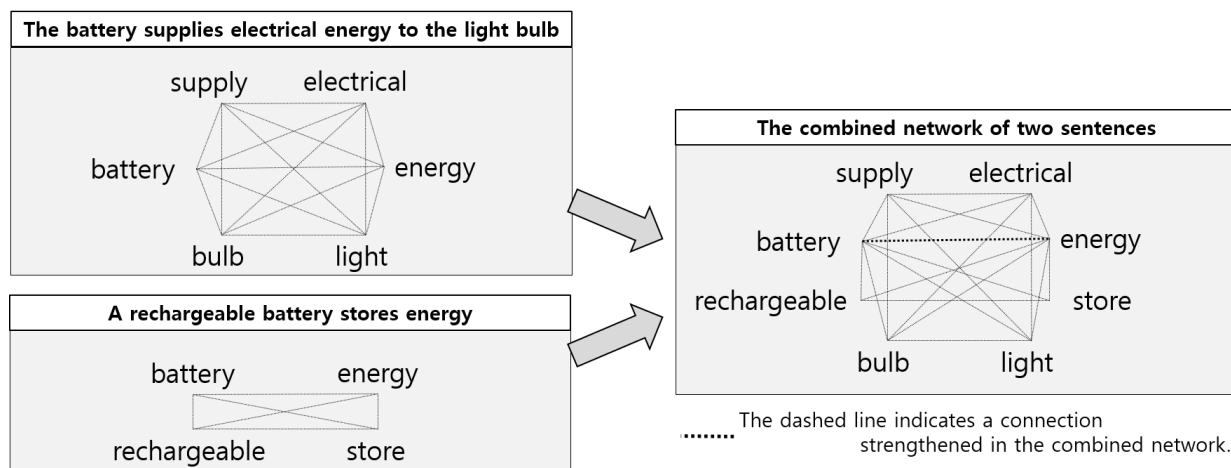


Figure 2. Examples of independent sentence networks and combined network.

Based on the example in Figure 2, Ree, Jeong, Park, and Kim [37] analyzed changes in areas of interest within the battery field by interpreting the relationships between technology elements based on a patent keyword network and constructing a context–problem network. Similarly, Choi and Hwang [38] analyzed changes and trends in research areas within the LED display field and identified keywords within the network as critical technology elements.

The keywords in patents are not interpreted independently but rather as terms representing technology elements that combine to form solutions to problems, technology convergence, or diverse perspectives within a specific technological field. However, as shown in Figure 2, when keywords appear within a single unit, such as a document or a sentence, connections are formed between all elements, potentially creating links between technology elements that are not technologically related. Additionally, previous studies have given limited attention to temporal changes in the relationships between technology elements as networks were not constructed and analyzed separately for different years. Therefore, this study employs NLP to extract and define meaningful keywords representing technology elements, which are then used to construct networks. The temporal changes in these networks are quantified as indicators of shifts in technology elements, enabling an analysis of their impact on technology advancement.

The reviewed approaches reveal two separate streams: policy evaluation tools that quantify macro effects (e.g., DID) and text/network analytics that capture micro-level technological structure. Few integrated frameworks treat macro policy shocks as exogenous exposure while explicitly modeling the network dynamics of technology-element evolution. This study occupies this intersection by (i) defining policy exposure in a DID setting and (ii) operationalizing dynamic technology change via a network-based index. This integrated setup directly serves the paper’s objective: to explain why policy effects differ across product classes and how such differences translate into actionable R&D choices.

2.5. Theoretical Channel: From Policy Exposure to Technology-Element Change

The link between macro policy and technology-element change assumed in this study draws on the induced innovation hypothesis in its patent-based empirical form [39] and its formalization as directed technical change [40,41]. In this view, a policy shock alters the relative returns to alternative technological inputs and thereby redirects firms’ research effort. Empirical work in the environmental-policy literature supports this redirection channel: carbon taxes have been shown to shift automakers’ patenting activity toward clean rather than dirty engine technologies [42], and well-designed environmental regulation can induce innovation offsets that firms would not otherwise pursue [43]. Demand-pull and

technology-push policy instruments have similarly been shown to differ in which stage of the innovation process they redirect [44].

Building on this literature, we posit a four-step micro-level channel connecting the Paris Agreement to the patent-text evidence examined in this study: (i) the policy changes firms' expected future market size and relative returns across battery sub-technologies; (ii) firms reallocate R&D budgets and search direction toward sub-technologies with improved expected returns; (iii) this reallocation shifts the locus of firms' local technological search—the set of existing and new technology elements firms recombine when developing new solutions; and (iv) the shift becomes observable in patent text as changes in the technology-element network, captured here by the NTE, NTF, and VTF indices. Figure 3 summarizes this chain.

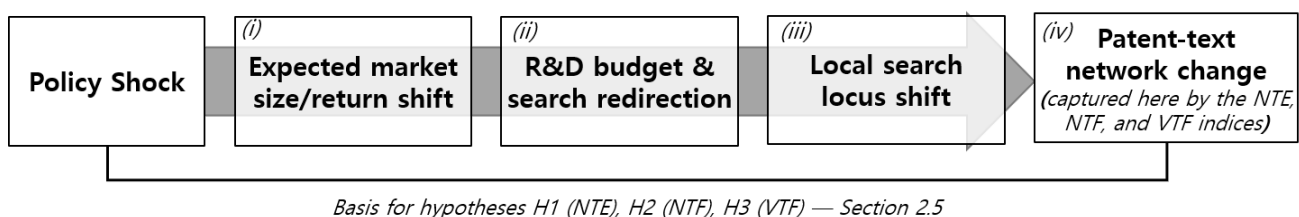


Figure 3. Micro-level channel from policy exposure to patent-text technology-element change.

This channel yields three hypotheses that structure the interpretation of the empirical results in Section 4. H1: sub-technologies with higher policy relevance exhibit a larger increase in NTE, reflecting entry of new technology elements as firms search beyond their existing vocabulary. H2: policy-relevant sub-technologies also exhibit a disproportionate increase in NTF, reflecting new combinations formed as firms recombine newly entered and existing elements. H3: sub-technologies with lower policy relevance are instead characterized by VTF-dominant change, reflecting exploitation of already-established technological combinations rather than exploratory search. Because all three indices scale with filing volume, the hypotheses are evaluated per filing in Section 4.2; and because the estimated policy differentials prove sensitive to the covariate specification (Section 4.3.1), they are evaluated against the composition of technological change within each sub-technology rather than against a ranking of sub-technologies.

3. Frameworks

This study examines the impact of various factors on technology advancement, including macro policies—which affect diverse R&D environments based on social and economic conditions—and changes in text-based technology elements. A DID design identifies the magnitude and direction of policy impact, while the network index pinpoints which technology elements evolve; the two insights are then combined to derive strategy rules. Statistically significant relationships are analyzed, and the evolving technology elements are identified. Based on the findings, an R&D strategy at the micro-level is proposed, which focuses on R&D themes shaped by policy and technology elements. The research methodology is divided into three stages, as shown in Figure 4.

In Module 1, patents in the target field for R&D strategy formulation are collected, and the top 20 applicants are filtered to reflect different technology strategies based on the applicant. Next, sub-technologies are defined as per IPC guidelines to establish detailed strategies within the technology field based on policy changes and shifts in technology elements.

In Module 2, patent indicators and the technology change index are measured to explore the factors influencing technology advancement. The patent indicators comprise elements identified in previous research as impacting technology development, while the technology change index is derived using NLP techniques and network dynamics.

In Module 3, the measured indicators and macro policies represented as dummy variables are used to perform a DID regression analysis to establish an R&D strategy. Statistically significant relationships among policies, characteristics of technological fields, temporal changes represented by dummy variables, technology-element change indicators, and patent indicators are identified. On this basis, an R&D strategy is formulated.

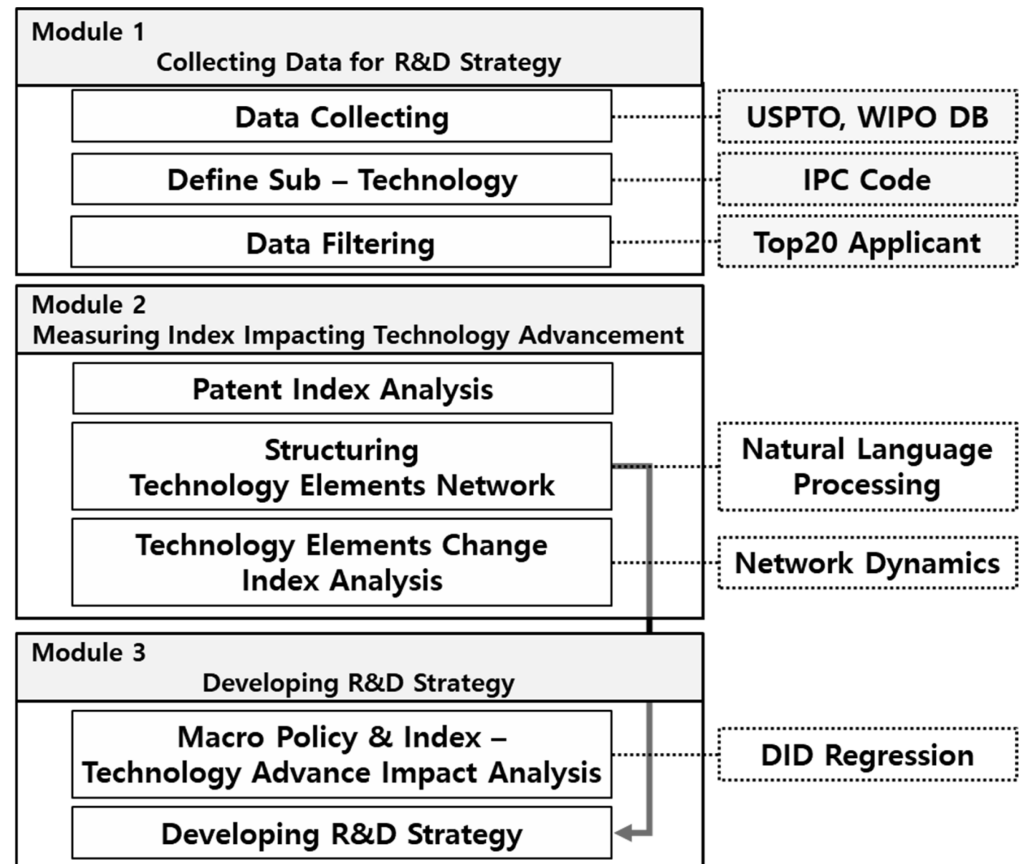


Figure 4. Research framework.

3.1. Data Collection

U.S. patents from the United States Patent and Trademark Office (USPTO) and Patent Cooperation Treaty (PCT) patents from the World Intellectual Property Organization (WIPO) were collected to analyze the impact of changes in policy and technology elements on technology advancement. U.S. patents are considered the most influential IP assets and represent a market essential for IP business, while PCT applications, governed by the Patent Cooperation Treaty, allow applicants to secure legal rights in over 150 member countries through a single application. Both U.S. and PCT patents are written in English and encompass key technologies in the U.S. as well as globally. Therefore, U.S. and PCT patents are considered herein as the most comprehensive, impactful, and meaningful datasets related to technological fields. The patents in the target technology fields were collected based on IPC codes, which are supported by both the WIPO and USPTO databases. IPC codes structurally categorize technologies into sections, classes, and subclasses, enabling the collection of patents from specific technology fields based on clear and standardized criteria. The collection of patents is conducted over an equal period before and after the impact of the macro policy to ensure a balanced dataset of patents influenced and not influenced by the policy.

Two features of the retrieval require comment. First, the IPC condition was specified at the subclass level, so that each target subclass was retrieved in its entirety; the retrieved

records were then assigned at the group level, according to whether they fall within the scope of the policy, to form the analysis datasets. For example, in the case study presented in Section 4, patents in the battery field were collected at the subclass level and subsequently divided into battery product types at the group level. Second, no restriction was placed on the publication stage of the record: both pre-grant publications and granted patents were retrieved.

Because the two sources overlap—a single invention may appear both as a U.S. national filing and as a PCT application—records were then consolidated at the patent-family level. Documents sharing an INPADOC family identifier were treated as one invention, and the member with the earliest application date was retained; where several members shared that date, one member was selected by a fixed country precedence of US, EP, CN, KR, and JP. Where the retained member was not in English, it was replaced by an English-language member of the same family; families containing no English-language member were excluded, since the technology-element extraction described in Section 3.2.2 operates on English text. Together these steps removed 127,864 of the 231,099 retrieved records, leaving 103,235 unique English-language inventions. The complete search strategy, field-code definitions, and the order in which the filters were applied are reported in Appendix B.

The collected data were filtered by applicant to control for firm-level heterogeneity in R&D strategies and technological positioning. Since technological development is largely driven by firms as primary innovation actors, differences in strategic orientation can substantially influence patent characteristics and technology-element dynamics [45,46]. Prior research further suggests that industry-level technological trajectories are typically shaped by leading incumbents whose patent portfolios structure the broader innovation landscape, particularly in energy and automotive-related sectors [26,27]. Concentrating on the top 20 applicants—who together account for 29,573 of the 103,235 family-consolidated H01M inventions filed between 2010 and 2022, or 28.6% of the analytic universe defined in Section 3.1—therefore enables the identification of structurally meaningful policy responses while minimizing noise from sporadic or marginal patenting activities. Because this ranking is computed over the full observation window, the selected sample may over-represent applicants whose patent output expanded after 2016. The selection rule, however, operates at the applicant level, whereas treatment and comparison groups are defined at the sub-technology level: the same 20 applicants contribute observations to every sub-technology, so any tilt in the applicant pool enters the treated and comparison groups alike and is differenced out of the quantity the DID specification identifies, and applicant fixed effects absorb time-invariant differences in scale across the selected firms. The concern can also be checked directly. Recomputing the ranking on pre-treatment filings alone (2010–2015) returns sixteen of the same twenty applicants, and the identical sixteen when the window is narrowed to 2011–2015; the four applicants that enter the pre-treatment list rank twenty-second, twenty-third, twenty-sixth and thirty-fourth over the full window, so the two rules do not select different populations, and the sixteen applicants common to both account for 83.3% of the filings the top twenty contribute over 2010–2022. Three of the four applicants that leave are consequences of identifying applicants at the level of the filing entity, as described below, rather than post-policy entrants: LG Energy Solution was demerged from LG Chemical in 2020 and therefore has no pre-treatment filings, while LG Chemical appears in both lists; Murata Manufacturing acquired Sony’s battery business in 2017, and Sony ranks tenth in the pre-treatment list; and Kia Motors ranks twenty-fifth in the pre-treatment window, with Hyundai Motor present in both. Contemporary Amperex Technology is the only applicant whose presence in the top twenty is attributable to post-2016 expansion. This is a check on the sampling frame; the estimates themselves have not been re-computed on the alternative sample (Section 5). Applicant identity was taken from the standardized

applicant-name field supplied by the database (Appendix B.4). Records naming more than one applicant were split, and each named applicant credited with one record. No consolidation was applied across distinct legal entities: applicants are identified at the level of the entity named in the record rather than at the level of the corporate group (Section 4.1). Assignments, mergers, and acquisitions occurring after the application date were not traced, so each patent remains attributed to the entity that filed it.

To establish diverse technology strategies within the target technology fields, key areas of interest at the subclass level were selected, and they were defined as sub-technologies at the group level. As per IPC guidelines, subclasses denote specific technologies or application fields, while groups divide technologies based on their functions, structures, methodologies, or product-related characteristics. Since the impact of policies may vary across different technology fields, sub-technologies were categorized according to their sensitivity to policy changes, with the goal of setting diverse directions for technological development and suggesting corresponding R&D strategies.

Because the IPC assigns to a patent every code that describes its technical content, a single patent may carry codes in more than one of the retained groups. Such patents are assigned to every group whose code they carry. This rule preserves technical content that a principal-code rule would discard, at the cost of a partial overlap between the treated and comparison groups in the DID design. Because an overlapping patent moves both series in the same direction, the overlap compresses the estimated differential (Section 3.3.1).

3.2. *Measuring Index for Technology Advancement*

The collected patents were categorized by applicant and sub-technology to create distinct datasets, and for each dataset, the patent index and technology change index were calculated. The patent index draws on patent indicators that reflect applicant strategies, while the technology change index is determined by constructing a keyword network and applying NLP techniques to identify significant relationships within the network. Finally, changes in the technology elements influencing advancement were quantified using the technology change index.

3.2.1. Patent Index Analysis

The patent indicators were calculated for each applicant based on the patents they had filed within each sub-technology during the specified year. These indicators include the number of inventors, the number of cited patents, the number of keywords appearing in the title and abstract, and the number of IPC classes. The number of inventors refers to the amount of individuals involved in the patent filing because a single patent can have multiple inventors. Patents represent technical documents that consolidate the specialized knowledge of each inventor involved. As shown in Figure 5, since a single inventor can contribute to multiple patents, the number of unique inventors within the analyzed patent cluster was calculated as a key indicator for the index.

The number of cited patents refers to the count of patent documents referenced or cited when a patent is filed. When the technology elements in a filed patent are not entirely novel, similar prior technology principles are cited to clarify the differences between the existing principles and the new ones being introduced. Thus, a higher number of cited patents can be interpreted as indicating a greater presence of distinguishable novel technology principles. Since the same patent can be cited multiple times within a patent cluster, the calculation focuses on the number of unique cited patents within the cluster, similar to the approach adopted for inventors (Figure 5).

Bibliographic by Patent						
Patent No	Technology Field	Year	IPC	Inventor	Citing Patent	Keyword in Title / Abstract
Pat_1	Tech_A	2015	A01M A02M	Dr. A Dr. B	Pat_10 Pat_11	KW1, KW2
Pat_2	Tech_A	2015	A02M A03M	Prf. A Prf. B	Pat_10	KW3, KW4
Pat_3	Tech_A	2016	A01M A02M	Prf. A Prf. B	Pat_13 Pat_14	KW1, KW5
Pat_4	Tech_A	2016	A03M A04M	Prf. A Dr. B	Pat_15 Pat_16	KW1, KW6
Pat_5	Tech_B	2015	B01M B02M	Dr. B Prf. C	Pat_17 Pat_18	KW7, KW8
Pat_6	Tech_B	2015	B02M	Dr. D	Pat_18	KW9, KW10
Pat_7	Tech_C	2015	C01M C02M	Dr. E	Pat_20 Pat_21	KW1, KW2
Pat_7	Tech_C	2015	C03M C04M	Dr. E	Pat_20 Pat_21	KW1, KW10

Patent Index by Technology Field & Year					
Technology Field	Year	IPC_#	Inventor_#	Citing_Patent #	Keyword in Title/Abstract #
Tech_A	2015	3 (A01M, A02M, A03M)	4(Dr. A, Dr. B, Prf. A, Prf. B)	2 (Pat_10, Pat_11)	4 (KW1, KW2, KW3, KW4)
Tech_A	2016	4 (A01M, A02M, A03M, A04M)	3(Prf. A, Prf. B, Dr. B)	4(Pat_13, Pat_14, Pat_15, Pat_16)	3(KW1, KW,5 KW6)
Tech_B	2015	2(B01M, B02M)	3(Dr. B, Prf. C, Dr. D)	2(Pat_17, Pat_18)	4(KW7, KW7, KW9, KW10)
Tech_C	2015	4(C01M,C02M, C03M,C04M)	1 (Dr. E)	2(Pat_20, Pat_21)	3(KW1, KW2, KW10)

Figure 5. Example of calculation of applicant's patent index.

The number of IPC classes indicates the number of technology fields to which a patent belongs when it is filed. Assigning multiple IPC classes to a patent suggests that it encompasses technology elements relevant to various fields. While a single patent may have multiple IPC classes, overlaps can exist within the patent cluster being analyzed. Therefore, similar to the approach used for inventors and cited patents, the calculation focuses on the unique number of IPC classes within the cluster (Figure 5).

Keywords in the title and abstract offer a straightforward method to identify the technology elements documented in patents. Changes in technology elements can be contrasted based on the number of unique keywords identified in the title and abstract. Since a patent may repeatedly use the same keyword to emphasize or describe a specific technology element, the count of unique keywords is utilized as another indicator. The same keyword can be identified multiple times in the analyzed patent cluster; hence, the approach focuses on the number of unique keywords, similar to the handling of inventors and cited patents (Figure 5). Figure 5 illustrates the patent index calculation for a

specific applicant; the dataset was constructed through the same process for all of the top 20 applicants.

3.2.2. Structuring Technology-Element Network

NLP and keyword-network analysis were performed not only to track the presence of technology elements but also to interpret changes in the relationships between them. From a text-mining perspective, connections between keywords can be defined in various manners. In this study, NLP was used to extract technically meaningful POSs as technology elements, and keywords with relations to these technology elements were structured into a network. The analysis of sentences via NLP, as illustrated in Figure 6, revealed the POSs for each word and the connections between them. The word-level analysis provided information on the lemmas and POSs. A lemma represents the root form of a word, while the POS indicates the grammatical role of each word within a sentence. Based on the obtained details, the structure of an entire sentence composed of multiple phrases can be identified. Each phrase comprised multiple words, with the key word in each phrase highlighted in bold.

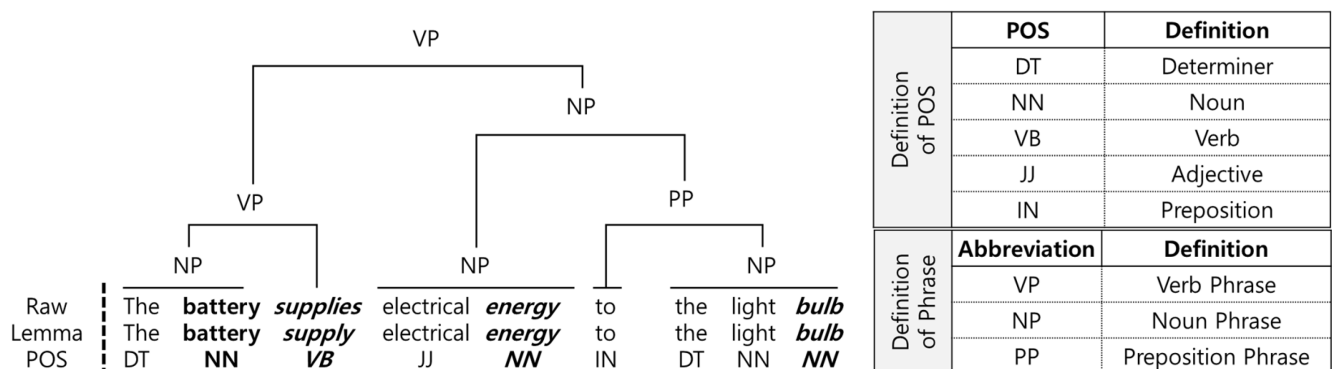


Figure 6. Example of NLP result: Phrase structure and POS–lemma information of words.

In this study, a network of technology elements was constructed based on three criteria:

- Only words that can be meaningfully interpreted as technology elements, such as nouns (NN), verbs (VB), and adjectives (JJ), can be used to form the network.
 - Elements such as determiners (DT) and prepositions (IN) are excluded as they do not represent technology elements.
 - Examples: DT: “the,” “a,” “an”; IN: “to,” “for,” “of”
- The lowest-level phrase forms a meaningful relationship network.
 - A phrase is defined as the smallest unit combining two or more words to convey meaning.
 - Examples: “light bulb,” “electrical energy”
- Connections are established among the key words within nested or linked phrases.
 - These connections link different POSs, such as nouns and prepositions.
 - Examples: “battery–energy,” “energy–bulb”

Figure 7 illustrates a technology-element network comprising two connected sentences, constructed based on the three rules outlined above. Compared with a combined network, which connects all elements within sentences, the NLP-based network is more semantically and technologically meaningful. For example, the “supply”–“light” connection appears in the network where all elements are linked but is absent in the NLP-based network as the two relevant sentences do not technically imply that light is supplied. Additionally,

“battery” is directly connected to “supply,” “store,” and “rechargeable” but is connected to “bulb” only indirectly through three steps: “store”–“energy”–“light.” This indicates that from a technological perspective, the “battery”–“bulb” connection has a lower degree of technical relevance.

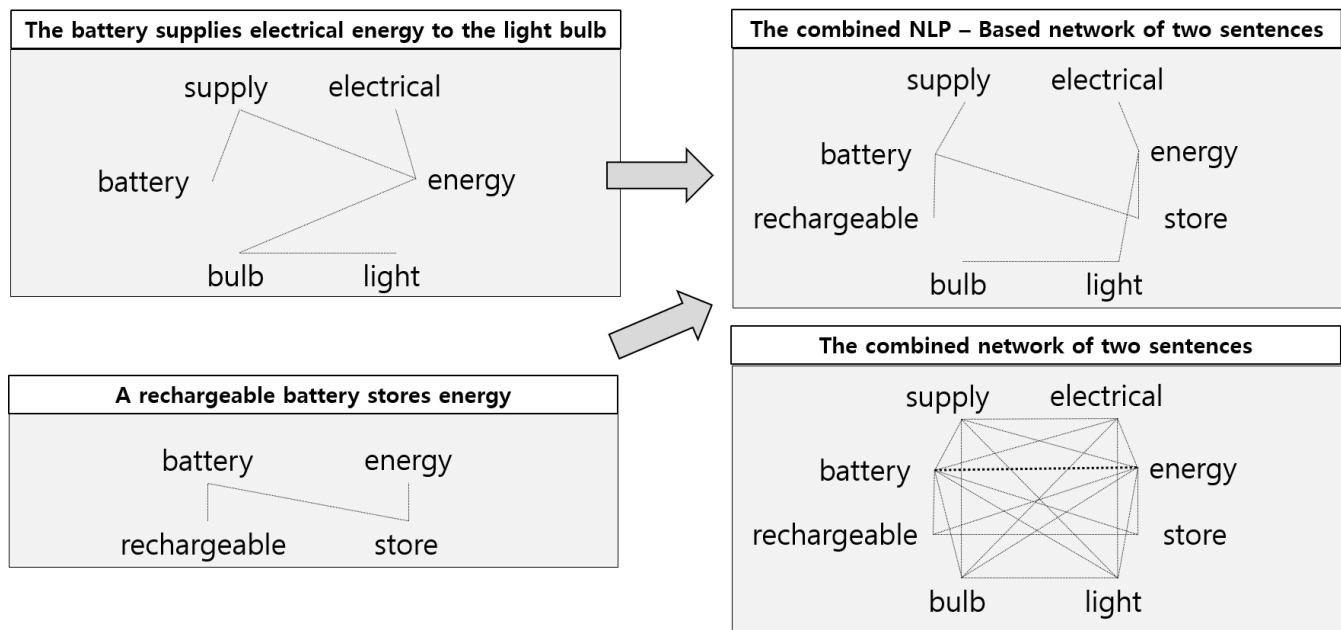


Figure 7. Comparison between Combined NLP-based network and combined fully connected network with two sentences.

3.2.3. Technology Change Index Analysis

The network of technology-elements evolves with changes in patent contents over time. For instance, if the sentence “A rechargeable battery stores long-lasting energy” is added, the “rechargeable”–“battery” and “battery”–“store” connections are strengthened. Additionally, the keyword “long-lasting” is introduced, which forms a new connection with “energy.” In this context, the addition of “long-lasting” is interpreted as the emergence of a new technology element (NTE), while its link to an existing term such as “energy” represents the formation of a new technology feature (NTF). Reinforced connections like “battery–store” and “battery–rechargeable” are interpreted as a change in the level of interest in existing technology principles (VTF). These three perspectives—NTE, NTF, and VTF—are quantified as shown in Formula (3) and summarized in Table 1. The corresponding example is visually illustrated in Figure 8, which demonstrates how these three types of changes are reflected in the evolution of a technology element network.

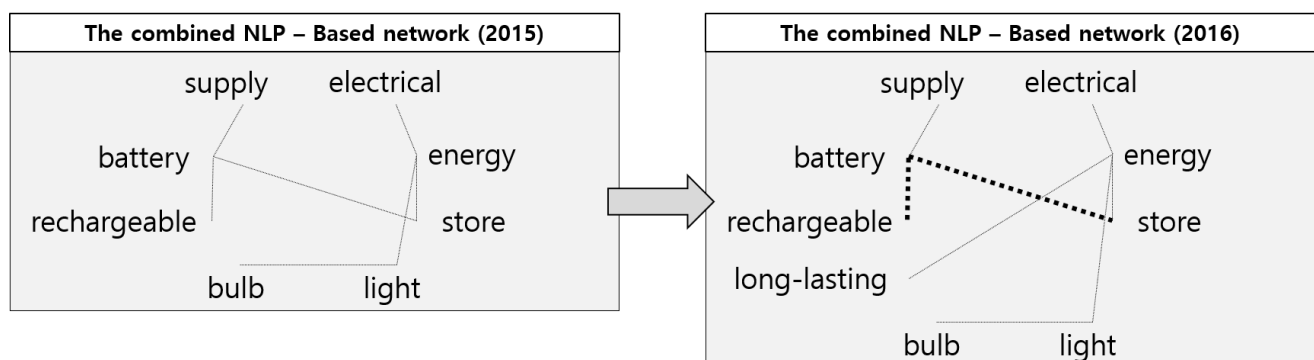


Figure 8. Example of changes in technology-element network.

Table 1. Examples of NTE, NTF, and VTF calculation based on network changes in Figure 8.

Technology Change Index	Calculation Process				Value
NTE	[battery, supply, electrical, energy, store, light, bulb, long-lasting, rechargeable] – [battery, supply, electrical, energy, store, light, bulb, rechargeable] = [long-lasting]				1
NTF	energy–long-lasting				1
VTF	Technology Element 1	Technology Element 2	Number of Connections		Variance
			2016	2015	
	Battery	Store	2	1	1
	Battery	Rechargeable	2	1	1
	Sum				2

Because NTE, NTF, and VTF are computed as deterministic, rule-based graph statistics from a fixed term-extraction and network-construction procedure (Sections 3.2.1 and 3.2.2) they are not subject to the small-sample estimation instability that motivates corpus-size diagnostics for such model-based approaches. The indices nonetheless scale mechanically with the number of patents filed in a given year. This scale effect is differenced out within the DID comparison; descriptive comparisons across sub-technologies are made per filing (Section 4.2).

These three components are grounded in the recombinant view of technological innovation [47–49], in which new technologies emerge through the entry, combination, and reuse of existing elements rather than through entirely novel invention. In this framing, NTE corresponds to the entry of a previously unused element into a firm’s technological vocabulary—an exploratory search for new building blocks [48]—NTF corresponds to the formation of a novel combination between existing elements [47], and VTF corresponds to the intensification of an already-established combination, reflecting exploitation of a known technological trajectory rather than exploration of a new one. Because network-based measurement tracks changes in the structural connections between lemmatized technology elements rather than surface-level word counts, a simple synonym substitution in patent drafting does not by itself alter NTF or VTF unless it also changes the underlying element-level network structure.

This study analyzes not only the presence of technology elements in patent texts, but also their structural connections and dynamic changes over time. By doing so, it complements traditional keyword-based approaches and enables a more structured and contextual interpretation of technological evolution.

Consequently, this framework offers deeper insights for identifying not only which technologies are emerging, but also how they are embedded and integrated within the broader technological landscape. It supports more strategic decision-making in R&D planning and enables the detection of policy impacts with greater precision. By moving beyond surface-level term frequency, this method captures the underlying dynamics of innovation and provides a stronger basis for actionable policy and strategic recommendations.

$$\begin{aligned}
V_n &= \text{Network composed with technology elements in } n \text{ year;} \\
|V_n| &= \text{The number of technology elements in network } V_n; \\
V_n \setminus V_{n-1} &: \text{Set of technology elements in } V_n \text{ that are not present in } V_{n-1}; \\
v_1^n, v_2^n, v_3^n \dots v_z^n \in V_n &: \text{Technology elements 1, 2, } \dots, \text{ and } z \text{ in belonging network } V_n; \\
N(v_{a,b}^n) &= \text{The number of connection between } v_a^n \text{ and } v_b^n; \\
\text{NTE} &: \text{The number of new technology elements in } n = |V_n \setminus V_{n-1}|; \\
\text{NTF} &: \text{The number of new technology feature in } n = \sum N(v_{a,b}^n); \\
&\text{if } v \in V_n \text{ and } v \notin V_{n-1} \\
\text{VTF} &: \text{The variance of interest in technology feature in } n = \sum (N(v_{a,b}^n) - N(v_{a,b}^{n-1})). \\
&\text{if } v \in V_n \text{ and } v \in V_{n-1}
\end{aligned} \tag{3}$$

Formula (3) Network Element-based Indicators: NTE, NTF, and VTF.

3.3. Developing R&D Strategy

This section describes how econometric (DID) and network outputs from Module 2 are translated into R&D choices. In this study, technological development is operationalized as the number of patent applications, and DID regressions are conducted using the DID variables, the proposed technology change index, and patent-related indicators to develop an R&D strategy. DID regressions are estimated to obtain policy-effect coefficients by sub-technology (direction and strength), and the network-based technology-element index is examined to identify specific R&D themes underlying significant changes. All empirical estimates and interpretations are reported in Section 4; here, we specify the rule set that maps these outputs to strategy choices at the macro and micro levels.

3.3.1. Macro Policy and Index—Technology Advance Impact Analysis

The factors influencing technological advancement are derived based on the interpretation of the DID regression results. The two differences used to construct the dataset for the DID regression, Difference T and Difference G, are defined as follows. Difference T represents the time dimension, distinguishing between the pre-policy and post-policy periods. Difference G represents whether a sub-technology is affected by the policy. Policy exposure is determined at the sub-technology level. If a target sub-technology is assumed to be affected by the policy, the remaining sub-technologies are assumed to be unaffected. Accordingly, this study repeatedly performs DID regressions to test the hypothesis for each sub-technology and derive the corresponding results.

This procedure is repeated for all four sub-technologies without presupposing which one is affected by the policy; policy exposure is therefore an empirically tested outcome for each sub-technology rather than an assumption fixed in advance. The resulting DID coefficients are accordingly interpreted as differential post-2016 changes relative to the remaining sub-technologies, not as causal effects of the Agreement on a predetermined treated group.

A further identification concern is that the stable unit treatment value assumption (SUTVA) underlying the DID design is unlikely to hold exactly across battery sub-technologies. For example, advances in electrode or electrolyte materials developed for one sub-technology are often adapted by firms working on another. Because a positive spillover would raise patenting activity in the comparison sub-technologies as well as the treated one, its effect is to narrow rather than widen the estimated differential. The DID coefficients reported in Section 4.3 should therefore be interpreted as conservative lower bounds on the

true divergence in policy responsiveness across sub-technologies. A quantitative diagnostic of the magnitude of such spillovers is left to future work.

In addition to the DID variables, the patent index and the technology change index, derived in Sections 3.1 and 3.2, are included as independent variables. Formula (4) is written with applicant fixed effects to absorb time-invariant heterogeneity across firms. Section 4.3.1 reports it, for comparability with the originally submitted results, both without and with them. Because both indices are computed from the same patent documents that generate the dependent variable, they may themselves respond to the policy rather than being predetermined with respect to it. Conditioning on variables that are realized after treatment can absorb part of the effect under study. Section 4.3.1 therefore reports Formula (4) alongside a covariate-free baseline that retains only the treatment dummies and applicant fixed effects, and the two are read as complementary.

$$\begin{aligned} \text{The Number of Patent}_{\gamma,s,t} &= \beta_0 + \beta_1 DID_{s,t} + \beta_2 G_s + \beta_3 T_t + \sum_{i=1}^n \beta_i PatentIndex_{i,\gamma,s,t} \\ &+ \sum_{j=1}^m \beta_j TechChangeIndex_{j,\gamma,s,t} + FE_{\gamma} + \varepsilon_{\gamma,s,t} \end{aligned}$$

The Number of Patent _{γ,s,t}

Number of patent applications by affiliation γ in sub-technology s at time t

DID: Dummy variable for treatment sub-technology s in the post-policy period time t

G_s: Dummy variable for treatment sub-technology s

(if hypothesis H_0 assumes that sub-technology s is affected by the macro policy, $G_s = 1$; otherwise, $G_s = 0$)

T_t: Dummy variable for post-policy period time t

(if period time t is post-policy, $T_t = 1$; otherwise, $T_t = 0$)

PatentIndex _{i,γ,s,t} : the i -th patent index of affiliation γ in sub-technology s at time t

TechChangeIndex _{j,γ,s,t}

the j -th sub-technology change index of affiliation γ in sub-technology s at time t

FE _{γ} : fixed effects of affiliation γ

(4)

Formula (4) DID regression model with patent index and technology change index.

For example, considering two sub-technologies, **G1** and **G2**, the DID regression was performed twice. When **G1** was assumed to be influenced by a policy implemented in 2016, the dataset was structured as shown in Table 2. Since each affiliation holds patents in each sub-technology, the data are organized in a structure defined by the product of the number of years, the number of sub-technologies, and the number of affiliations. Only patents filed after 2016 have a value of 1 for the dummy variable T ; moreover, since **G1** is assumed to be affected by the policy, only data corresponding to **G1** have a value of 1 for the dummy variable G . The dummy variable DID is set to 1 for instances where the time is post-2016 and the sub-technology is **G1**. The term FE_i denotes applicant fixed effects, one indicator per affiliation; as noted in Section 3.3.1, the estimates reported for this specification in Section 4.3.1 are obtained without them, while the covariate-free baseline reported alongside includes them. The process is then repeated with the assumption that **G2** is affected by the policy, whereby two additional regression analyses are conducted. Each analysis yields the DID coefficient along with the technology-element change index and the patent index.

Table 3 presents the results of the DID regression. Both Tech_A and Tech_B show an adjusted R^2 value of over 0.90, indicating an explanatory power of over 90% for the data. Regarding the dummy variable, Tech_A exhibits a DID coefficient of 13.662 with a p -value of 0.001; this suggests that Tech_A is statistically significant and has a positive effect on patent applications. Accordingly, the policy positively influenced Tech_A, leading to an increase in patent filings. In the case of Tech_B, the technology-specific variable G is statistically significant, indicating that the increase in patent filings is attributable to

the unique characteristics of the technology field rather than the policy. Regarding the technology-element change index, Tech_A shows a statistically significant impact from the increase in new technology principles. For Tech_B, however, the change in the level of interest in technology principles negatively impacted patent application.

Table 2. Sample DID data structure assuming sub-technology Tech_A to be affected by policy implemented in 2016.

Year	Sub Tech.	Affiliation	Dependent Var.	Independent Var.					Control Var.	
			# of Patents	G	T	DID	Patent Index	Technology Change Index	FE ₁	FE ₂
2014	G1	Aff.1	...	1	0	0	...	...	1	0
2014	G2	Aff.2	...	0	0	0	...	...	0	1
2014	G1	Aff.1	...	1	0	0	...	...	1	0
2014	G2	Aff.2	...	0	0	0	...	...	0	1
...										
2015	G1	Aff.1	...	1	1	1	...	...	1	0
2015	G2	Aff.2	...	0	1	0	...	...	0	1
2015	G1	Aff.1	...	1	1	1	...	...	1	0
2015	G2	Aff.2	...	0	1	0	...	...	0	1
...										

Table 3. Sample result of DID regression for sub-technology.

			Tech_A	Tech_B
Adjusted R ²			0.9425	0.9021
Variable Coef (P > t)	Constant		...	...
	G		...	12.33 (0.004)
	T		...	...
	DID		13.662 (0.001)	−2.33 (0.2333)
	Patent Index	IPC_#	...	...
		Inventor_#	...	...
		Citing Patent #	6.122 (0.004)	...
		Keyword #	...	...
	Technology Element Change Index	NTE	...	...
		NTF	7.222 (0.001)	...
		VTF	...	−5.223 (0.001)

3.3.2. Developing R&D Strategy

Based on the results of the DID regression, which highlight the impact of policies on each sub-technology, methods and directions for R&D are proposed, along with an R&D strategy informed by an analysis of keywords related to the technology-element change index.

When policy impact is considerable, identifying and prioritizing the development of technology elements supported or regulated by the government or selectively reducing investments in certain areas might be an effective strategy. Alternatively, collaboration with national institutions could help prioritize the development of government-supported items.

In cases where the impact is minimal or the sensitivity to policy is low, companies could pursue innovative, independent technology elements. Such cases can be interpreted as involving fewer regulatory constraints on technology advancement, enabling companies to select unique technology elements and formulate IP strategies accordingly. Notably, policy impact is not always positive; restrictive policies may have negative effects. In such cases, companies could establish R&D strategies that avoid or circumvent regulated technology elements by identifying areas affected by existing policies.

4. Illustration

4.1. Data Collection

The Paris Agreement (2015) is an international climate accord aimed at addressing global climate change, which has placed significant pressure on industries heavily dependent on fossil fuels. The battery industry, in particular, will play a crucial role in replacing fossil fuels or improving energy efficiency and is therefore central to achieving the goal of carbon neutrality by 2050. Consequently, patents from the top 20 applicants in the battery sector filed between 2010 and 2022—six years before (2010–2015) and seven years after (2016–2022) the adoption of the Paris Agreement—were collected and analyzed. Applicants affiliated with the same corporate group (for example, LG Chemical and LG Energy Solution, or Panasonic and Panasonic Intellectual Property Management) are treated as distinct R&D entities, on the assumption that each formulates and pursues its own R&D strategy. One qualification applies to this treatment. LG Energy Solution was established in December 2020 as a spin-off of the battery business of LG Chemical, and its record as an applicant of patents is therefore discontinuous by construction. The theme-matched counts reported for it in Section 4.4 (11 applications before 2016 against 922 after) reflect the emergence of a new legal entity rather than a policy-induced change in research activity within a continuously operating organization. Its rows in Section 4.4 are therefore interpreted at the group level. In the IPC code, the battery field corresponds to the subclass H01M, which was searched in its entirety; the four groups that denote a battery type, were then identified from the IPC codes of the retrieved records and form the units of analysis. Data were collected and filtered based on the top 20 applicants; the resulting dataset is outlined in Table 4, and the top 20 applicants and their respective patent counts are listed in Table 5.

Table 4. Data collection details.

Patent Query/Patent Filtering	Number of Patents
ICH = (H01M) AND (AD $\geq$ 2010 AND AD < 2023)	231,099
After INPADOC family consolidation and English-language correction	103,235
Limited to Top 20 Patent Applicants	29,573

Notes: ICH: IPC search field, matching a record if any IPC position, whether inventive or additional, falls under the specified subclass. AD: application year. The retrieval covered subclass H01M as a whole and was not restricted by publication stage; the four battery-type groups analyzed in this study were identified from the IPC codes of the retrieved records at the analysis stage. The search was executed in September 2024. Retrieval, family consolidation, and English-language correction are described in Section 3.1; the full search strategy and filtering sequence are reported in Appendix B.

The sub-groups of IPC H01M are presented in Table 6, with each group categorized as a product, technology principle, or key component. While multiple products may share technology principles and key components, battery product classifications can be distinctly defined based on technology principles. Therefore, primary batteries, fuel cells, secondary batteries, and hybrid cells are defined as sub-technologies when analyzing the impact of policies and changes in technology principles. The records retrieved for subclass H01M

were accordingly assigned to the four groups that denote a battery type. H01M-050/00, which corresponds to battery module or pack structures, represents a product-component category rather than a distinct battery type; it is therefore unsuitable for assessing which battery types respond to policy and is excluded from the analysis.

Table 5. Top 20 applicants and number of patents.

Applicant *	Number of Patents	Applicant *	Number of Patents
TOYOTA JIDOSHA	3586	SANYO ELECTRIC	1094
LG CHEMICAL	3471	SAMSUNG ELECTRONICS	1052
SAMSUNG SDI	3383	FORD GLOBAL TECHNOLOGIES	924
LG ENERGY SOLUTION	2469	NISSAN MOTOR	792
HYUNDAI MOTOR	1882	KIA MOTORS	789
PANASONIC INTELLECTUAL PROPERTY MANAGEMENT	1657	GS YUASA INTERNATIONAL	691
GM GLOBAL TECHNOLOGY OPERATIONS	1555	TOSHIBA	631
HONDA MOTOR	1510	SUMITOMO ELECTRIC INDUSTRIES	584
ROBERT BOSCH GMBH	1248	PANASONIC	561
CONTEMPORARY AMPEREX TECHNOLOGY	1138	MURATA MANUFACTURING	556

* Applicants are ranked by the number of patents filed over the full 2010–2022 observation window; the implications of this selection rule are discussed in Sections 3.1 and 5. Sixteen of these twenty applicants are also returned when the ranking is recomputed on pre-treatment filings alone (2010–2015); see Section 3.1.

Table 6. Subcategories of IPC H01M.

IPC Code	IPC Title	Type
H01M-004/00	Electrodes	Component of Product
H01M-006/00	Primary cells; Manufacture thereof	Product
H01M-008/00	Fuel cells; Manufacture thereof	Product
H01M-010/00	Secondary cells; Manufacture thereof	Product
H01M-012/00	Hybrid cells; Manufacture thereof	Product
H01M-014/00	Electrochemical current or voltage generators not provided for in groups H01M6/00–H01M12/00; Manufacture thereof	Core Technology
H01M-016/00	Structural combinations of different types of electrochemical generators	Core Technology
H01M-050/00	Constructional details or processes of manufacture of the non-active parts of electrochemical cells other than fuel cells, e.g., hybrid cells	Core Technology

is computed by aggregating the corresponding yearly values and using the summed totals for evaluation. The identified elements can serve as reference points for specifying concrete R&D directions after the DID regression analysis, and they can also be used independently to interpret technology-specific trends over time. For example, Fuel Cells (H01M-008) shows an early shift toward new cell-centered recombinations (NTF: cell, 2015), followed by a strengthening of gas-related linkages (VTF: gas, 2016). More recently, the emergence of flow as a new element (NTE: flow, 2020) suggests increasing attention to flow/transport-oriented themes, indicating that the technology focus has expanded from core cell configuration to gas handling and flow-field optimization.

Table 8. Elements with the highest value for NTE, NTF, and VTF.

Sub Technology	NTE		NTF		VTF	
	Top Element	Top Year	Top Element	Top Year	Top Element	Top Year
Primary Cells (H01M-006)	surface	2014	electrode	2022	surface	2019
Fuel Cells (H01M-008)	flow	2020	cell	2015	gas	2016
Secondary Cells (H01M-010)	voltage	2015	material	2022	compound	2020
Hybrid Cells (H01M-012)	electrolyte	2015	anode	2014	product	2019

Since the technology-element change index is based on network changes, it can be constructed only for an applicant that has filed at least twice in a given sub-technology, and it is computed between successive application years rather than between adjacent calendar years. For example, if an applicant filed in the hybrid cell field (H01M-012/00) in 2015, filed nothing in 2016, and filed again in 2017, the network change is calculated between 2015 and 2017. Applying this rule yields 533 applicant $\times$ sub-technology $\times$ year cells, a partial extract of which is shown in Table 9. Seven cells belong to the residual group H01M-014/00 and are retained in the comparison group; excluding them changes no result in of the DID regressions in Section 4.3.1 materially.

Because all three components rise mechanically with filing volume (Section 3.2.3), Table 9 reports the indices per filing. Normalization reverses part of the picture. In levels, NTE rises after 2016 in every sub-technology; per filing it falls in all three fields that grew, most steeply in secondary cells (0.99 to 0.71). Per filing, NTF and VTF rise everywhere. The post-2016 period is thus characterized not by a broader search for new technology elements but by denser recombination and intensified use of elements already in the vocabulary. Table 9 tests this on the normalized measures directly. One coefficient survives: secondary cells record a significant increase in new links per filing (0.492, $p = 0.028$). This supports H2; H1 is not supported, since normalized NTE declines in the growing fields, and H3 cannot be evaluated as stated because no sub-technology is VTF-dominant in absolute terms. The significant NTE coefficient for primary cells is not interpreted, for the reason given in Section 4.3.1: the field records 80 filings in the entire post-policy window.

Normalization is applied here and not in the DID regressions of Section 4.3.1. Dividing an index by the number of filings and then using the ratio to explain that same number of filings would place the dependent variable in the denominator of a regressor and induce a mechanical negative association.

Table 9. Calculated patent index and technology-element change index (Partial). (a) Technology-element change index per filing, before and after 2016. (b) Difference-in-differences on the normalized indices.

Sub-Technology	Applicant	Year	# of Patents	Patent Index				Technology-Element Change Index			
				# of Inventors	# of IPC Classes	# of Citing Patents	# of Keywords	NTE	NTF	VTF	
H01M-010	TOYOTA JIDOSHA	2013	85	158	98	203	86	95	107	14	
H01M-010	TOYOTA JIDOSHA	2014	97	141	102	219	88	65	120	14	
...											
H01M-010	TOYOTA JIDOSHA	2021	114	161	165	55	105	63	258	12	
H01M-010	TOYOTA JIDOSHA	2022	196	182	238	73	126	154	417	40	
H01M-010	LG CHEMICAL	2013	123	254	103	444	105	110	86	4	
H01M-010	LG CHEMICAL	2014	334	526	168	415	173	210	677	22	
H01M-010	LG CHEMICAL	2015	168	336	153	78	109	72	330	26	
H01M-010	LG CHEMICAL	2016	229	418	185	104	150	127	437	32	
H01M-010	LG CHEMICAL	2017	313	639	234	289	185	214	860	48	
...											
H01M-010	SAMSUNG SDI	2011	95	133	80	332	83	89	163	6	
H01M-010	SAMSUNG SDI	2012	90	151	95	462	80	57	149	8	
...											
a											
Sub-technology	Filings		NTE per filing			NTF per filing			VTF per filing		
	Pre	Post	Pre	Post	Δ	Pre	Post	Δ	Pre	Post	Δ
Primary Cells (H01M-006)	82	80	4.45	4.62	+4%	0.73	0.94	+28%	0	0.25	—
Fuel Cells (H01M-008)	1986	3118	1.45	1.2	−17%	1.61	1.81	+12%	0.1	0.16	+62%
Secondary Cells (H01M-010)	4014	12,013	0.99	0.71	−28%	1.29	1.68	+30%	0.13	0.32	+155%
Hybrid Cells (H01M-012)	141	188	3.96	3.34	−16%	2.11	2.71	+29%	0.1	0.22	+125%

Table 9. Cont.

Sub-technology (G = 1)	b								
	NTE per filing			NTF per filing			VTF per filing		
	DID	SE	p	DID	SE	p	DID	SE	p
Primary Cells (H01M-006)	1.915	0.782	0.024	−0.392	0.43	0.373	−0.022	0.102	0.83
Fuel Cells (H01M-008)	0.142	0.353	0.691	−0.126	0.369	0.736	−0.015	0.047	0.754
Secondary Cells (H01M-010)	−0.346	0.339	0.32	0.492	0.207	0.028	0.025	0.052	0.631
Hybrid Cells (H01M-012)	−0.549	0.992	0.587	−0.178	0.955	0.854	−0.020	0.08	0.807

NTE: Number of New Technology Elements. NTF: Number of New Technology Features. VTF: Variance of Interest in Technology Features. Notes: Each entry divides the sum of the index over the period by the sum of filings over the same period; the unit is index value per filing. The pre-policy window is 2011–2015 and the post-policy window 2016–2022. No growth rate is shown for VTF in primary cells because the pre-policy total is zero. Across the 533 cells the index correlates with the number of filings at 0.86 (NTE), 0.93 (NTF) and 0.73 (VTF), which is why levels are not compared across sub-technologies. Notes: Each entry is the interaction coefficient from a regression of the normalized index on the group, time and interaction dummies with applicant fixed effects; sample, rotation of the exposed group and clustered inference follow the note to Section 4.3.1. Normalization is applied here but not to the covariates of Section 4.3.1 because the normalized index is the dependent variable in this table and therefore carries no mechanical correlation with its denominator, whereas normalizing a regressor would place the dependent variable in the denominator of an explanatory variable.

4.3. Developing R&D Strategy

4.3.1. Macro Policy and Index—Technology Advance Impact Analysis

Prior to estimating the DID regressions, we tested whether the treated and comparison sub-technologies followed a common trajectory before the policy shift. The test is specified on the pooled pre-policy subsample ($T = 0$; application years 2011–2015; 243 observations on 19 applicants) and is run separately for each sub-technology taken in turn as the exposed group. The first enables interaction between group membership and a linear application-year trend and tests whether that interaction is zero. The second replaces the linear trend with a full set of year indicators interacted with group membership, using 2015 as the reference year, and jointly tests whether the four group-specific leads are zero [50]. Standard errors are clustered at the applicant level, and because the number of clusters is small, p -values are also obtained from a restricted wild cluster bootstrap (Rademacher weights, 9999 replications) [51]. The results are reported in Table 10.

Table 10. Tests of differential pre-treatment trends (2011–2015).

Panel A. Without covariates							
Sub-technology (G = 1)	Group $\times$ Trend	SE	p	p (WCB)	Joint F (4, 18)	p	p (WCB)
Primary Cells (H01M-006)	−5.028	1.862	0.015	0.025	3.393	0.031	0.07
Fuel Cells (H01M-008)	−6.894	2.975	0.032	0.028	1.88	0.158	0.207
Secondary Cells (H01M-010)	9.302	3.787	0.024	0.011	1.612	0.215	0.252
Hybrid Cells (H01M-012)	−4.957	1.93	0.019	0.034	2.227	0.107	0.213
Panel B. With the covariates of Formula (4)							
Sub-technology (G = 1)	Group $\times$ Trend	SE	p	p (WCB)	Joint F (4, 18)	p	p (WCB)
Primary Cells (H01M-006)	0.228	0.641	0.726	0.729	3.774	0.021	0.083
Fuel Cells (H01M-008)	−1.755	1.319	0.2	0.246	2.846	0.055	0.115
Secondary Cells (H01M-010)	1.851	1.558	0.25	0.291	2.628	0.069	0.122
Hybrid Cells (H01M-012)	−0.633	1.045	0.552	0.578	3.441	0.03	0.154

Notes: The sample is the pre-policy window only ($T = 0$; application years 2011–2015): $N = 243$ applicant–sub-technology–year cells on 19 applicants in every column, with the rotation and the clustered inference described in the note to Table 11. Group $\times$ Trend is the coefficient on G interacted with a linear application-year trend centered on 2015; the joint test is a Wald test that the four interactions between G and the 2011–2014 year indicators are zero, with 2015 as the reference year. Panel B adds the covariates of Formula (4). Adding applicant fixed effects leaves Panel B unchanged, with bootstrap p -values for Group $\times$ Trend between 0.26 and 0.77.

Panel A, which omits covariates, shows that the four sub-technologies were not on a common unconditional path: the group-by-trend interaction is significant at the 5% level in every case. Panel B adds the covariates of Formula (4). Conditional on these, no group-by-trend interaction is distinguishable from zero, with bootstrap p -values between 0.24 and 0.73, and the joint test of the group-specific leads is not rejected at the 5% level for any sub-technology, with bootstrap p -values between 0.08 and 0.15. It is thus the covariates of Formula (4) that render the pre-policy paths comparable. Because the joint test is nonetheless rejected at the 10% level for primary cells, conditional parallelism is supported but not established beyond doubt.

To verify whether each of the four sub-technologies (primary batteries, fuel cells, secondary batteries, and hybrid cells) had been influenced by the implemented policy, four

DID regressions were conducted based on the assumption that if one sub-technology is affected, the others are unaffected. The DID regression results are presented in Table 11. Multicollinearity was assessed using variance inflation factors evaluated against the conventional threshold of 10, and in the estimated specification the largest factor is 8.76. One measured variable, the number of keywords, was excluded before estimation because it is a second measure of the same lexical volume that NTE records.

Panel B reproduces the specification in Formula (4), with standard errors clustered on applicants in both panels. The post-2016 time indicator is negative in all four columns, indicating that filing levels decline over the second half of the observation window irrespective of sub-technology. In Panel B only primary cells depart from zero at a conventional threshold ($DID = 5.8901$, $p = 0.004$), but this estimate cannot be read as an economic effect: primary cells average 2.16 filings per applicant–year with a maximum of six, and the same sub-technology carries a large negative and equally significant coefficient in Panel A. Primary and hybrid cells are the two smallest fields in the sample, and their interaction coefficients are unstable in sign and implausible in magnitude across both panels; we therefore read them as an artifact of thin cells (Section 4.3.2). Neither fuel cells nor secondary cells carry a significant interaction in Panel B. The three components of the technology-element change index are significant in every column, NTE and NTF at the 5% level and VTF at the 10% level.

Panel A re-estimates the same four regressions with no covariates at all. The interaction coefficients are roughly five times larger in absolute value than in Panel B, while the explanatory power of the model falls from an R^2 of about 0.95 to between 0.31 and 0.43. In the covariate-free specification the only positive and statistically significant interaction is that of secondary cells (29.8859, bootstrap $p = 0.013$); primary cells (-29.4596 , $p < 0.001$) and hybrid cells (-20.5000 , $p = 0.033$) are significantly negative, and fuel cells remain negative but are significant only at the 10% level (-17.4853 , $p = 0.077$). The two panels thus disagree on which sub-technology carries the differential. We therefore do not claim that any single sub-technology can be identified as the policy beneficiary from these estimates. What both panels support is the weaker and more robust proposition that post-2016 filing growth was not common across battery sub-technologies. Because four interactions are tested, one per sub-technology, a Bonferroni adjustment multiplies each p -value by four; under it the primary-cell interaction remains significant in both panels (bootstrap $p = 0.002$ and 0.032), the secondary-cell interaction in Panel A becomes marginal ($p = 0.053$), and no other interaction is significant, so the adjustment changes no conclusion drawn here. The R^2 of about 0.95 in Panel B accordingly reflects the mechanical relation between the indices and the filing count rather than overfitting of the treatment effect: the interaction terms add almost nothing to the fit, and the Poisson pseudo- R^2 of the same specification is 0.79–0.83 (Table 12).

These magnitudes also record relative, not absolute, movements. Mean filings per applicant–year rose in every sub-technology after 2016—from 25.8 to 36.3 in fuel cells, from 44.1 to 89.7 in secondary cells, from 4.4 to 6.1 in hybrid cells, and remaining flat at 2.2 in primary cells—so a negative interaction indicates growth slower than the comparison group rather than contraction. Because each regression assigns the remaining sub-technologies to the comparison group, the rapid expansion of secondary cells enters the comparison group of the other three specifications and mechanically depresses their interaction coefficients (Section 3.3.1). Re-estimating Panel B under three successively wider covariate sets, each with and without applicant fixed effects, does not change the pattern: primary cells carry the only significant interaction in all eight estimates and the other three fields are insignificant throughout, and restoring the number of keywords leaves the interactions essentially unchanged.

The two panels report a single interaction for the whole post-2016 window, which leaves open when—and whether—any differential appears. We therefore re-estimated both specifications with the time and interaction dummies replaced by a full set of application-year indicators and their interactions with the group dummy, with 2015 as the reference. This yields four pre-policy leads (2011–2014) and seven post-policy lags (2016–2022) for each sub-technology, plotted with 95% confidence intervals in Figure 10.

The covariate-free specification does not describe a change that begins in 2016. Five of the sixteen pre-policy leads are individually significant at the 5% level, and the primary-cell and secondary-cell leads converge monotonically toward zero between 2011 and 2014. The post-2016 coefficients continue these paths rather than departing from them, so the large Panel A interactions in Table 11 measure the average gap between two halves of a pre-existing divergence. Panel A of Tables 10 and 12 records the same property in other forms.

Under the specification of Formula (4) the picture is different but no more supportive of a datable response. None of the twenty-eight post-2016 interactions is significant at the 5% level, and the joint test that the seven lags are zero is far from rejection in every column (Figure 10). The event study therefore locates no discrete break at 2016. The one significant interaction in Panel B, that of primary cells, is correspondingly the gap between a slightly negative pre-policy level and a modestly positive post-policy level (mean lag minus mean lead = 5.9031) rather than a shift at the onset. Because every unit shares the same onset year, the decomposition problems of staggered designs [52,53] do not arise.

Because neither panel can be validated against a known null inside the observation window, we added an in-time placebo test [54]. The sample is restricted to the pre-policy window (2011–2015), a counterfactual onset is imposed at 2013, and both specifications of Table 11 are re-estimated with each sub-technology in turn as the exposed group (see the note to Table 12). A specification that isolates a policy-associated differential should return an interaction indistinguishable from zero here. The test has little power against small differentials, and it does not address the post-treatment status of the covariates (Section 3.3.1); it asks only whether a specification manufactures differentials in the absence of a policy. The results are reported in Table 12.

Table 11. Results of DID regression under two specifications.

Panel A. Baseline: treatment dummies and applicant fixed effects only					
Variable	Statistic	Primary Cells (H01M-006)	Fuel Cells (H01M-008)	Secondary Cells (H01M-010)	Hybrid Cells (H01M-012)
G	Coefficient	−30.5628	3.2238	32.5331	−31.8642
	SE	7.3022	8.1252	11.4806	6.9
	<i>p</i>	0.001	0.696	0.011	<0.001
T	Coefficient	25.7441	28.6237	6.2923	24.4486
	SE	5.6358	5.3681	4.3871	5.1919
	<i>p</i>	<0.001	<0.001	0.168	<0.001
DID	Coefficient	−29.4596	−17.4853	29.8859	−20.5000
	SE	6.47	9.4762	10.6476	7.6931
	<i>p</i>	<0.001	0.081	0.011	0.015
	<i>p</i> (WCB)	<0.001	0.077	0.013	0.033
R ²		0.3587	0.306	0.4293	0.3405
Within-R ²		0.1181	0.0457	0.2153	0.0931

Table 11. Cont.

Panel B. With the Patent Index and the Technology-Element Change Index (Formula (4))					
Variable	Statistic	Primary Cells (H01M-006)	Fuel Cells (H01M-008)	Secondary Cells (H01M-010)	Hybrid Cells (H01M-012)
R ² -Patent Index Only		0.8353	0.8311	0.8456	0.8328
R ² -Patent Index and Technology-Element Change Index		0.9400	0.9398	0.9395	0.9394
Constant	Coefficient	−2.0489	−1.2451	−0.0429	−1.1897
	SE	4.0315	3.8085	3.6286	4.126
	<i>p</i>	0.617	0.747	0.991	0.776
G	Coefficient	1.6715	1.328	−2.8808	−0.3739
	SE	2.3991	2.0168	3.6718	1.921
	<i>p</i>	0.494	0.518	0.442	0.848
T	Coefficient	−6.6467	−3.4820	−6.1420	−5.6503
	SE	1.7138	1.9167	1.6278	1.4733
	<i>p</i>	0.001	0.085	0.001	0.001
DID	Coefficient	5.8901	−5.5528	1.8393	2.1253
	SE	1.7978	4.5174	4.4326	2.3398
	<i>p</i>	0.004	0.234	0.683	0.375
	<i>p</i> (WCB)	0.008	0.308	0.696	0.41
Number of Inventors	Coefficient	−4.2081	−4.6546	−4.6892	−4.2959
	SE	4.2717	4.2791	4.4112	4.3416
	<i>p</i>	0.337	0.29	0.301	0.335
Number of IPC Classes	Coefficient	0.4251	0.402	0.4216	0.4108
	SE	0.0567	0.0458	0.057	0.0566
	<i>p</i>	<0.001	<0.001	<0.001	<0.001
Number of Citing Patents	Coefficient	0.0245	0.0251	0.0305	0.026
	SE	0.0147	0.0142	0.013	0.0146
	<i>p</i>	0.113	0.094	0.031	0.092
NTE	Coefficient	0.274	0.2775	0.2596	0.2683
	SE	0.1169	0.1182	0.1087	0.116
	<i>p</i>	0.03	0.03	0.027	0.032
NTF	Coefficient	0.1884	0.1922	0.1905	0.1918
	SE	0.0462	0.045	0.0457	0.046
	<i>p</i>	0.001	<0.001	0.001	0.001
VTF	Coefficient	0.1837	0.1815	0.1829	0.1837
	SE	0.0983	0.0952	0.0992	0.0981
	<i>p</i>	0.077	0.072	0.081	0.077
Number of Keywords	Excluded based on VIF Analysis				

Notes: $N = 533$ applicant–sub-technology–year cells on 20 applicants in every column of both panels. Each column takes the named sub-technology as the exposed group and the remaining sub-technologies as the comparison group. Panel A retains only the group, time and interaction dummies together with applicant fixed effects, which absorb the constant. Panel B adds the patent index and the technology-element change index as specified in Formula (4) and is estimated, as in the originally reported results, without applicant fixed effects, which is why a constant is reported for it. Adding applicant fixed effects to Panel B leaves the pattern unchanged (primary cells 7.3905, bootstrap $p = 0.018$; all other interactions insignificant). Standard errors are clustered at the applicant level, and p (WCB) is obtained from a restricted wild cluster bootstrap with Rademacher weights and 9999 replications. The same sample, rotation and inference are used in Tables 9 and 12 and in Figure 10. The sign pattern of the Panel A interactions is unchanged when sub-technology and year fixed effects are added. Both goodness-of-fit rows in Panel B report unadjusted R^2 , the adjusted values lying between 0.9383 and 0.9390. Panel B is the specification set out ex ante in Formula (4); no variable was retained or dropped on the basis of its p -value (Section 4.3.1). Because the four columns are estimated on the same cells with the same index values, the index coefficients are not comparable across columns.

Table 12. Robustness of the interaction coefficient: in-time placebo, alternative policy onsets, and count specifications.

Sub-Technology (G = 1)	Specification	Panel A. Baseline				Panel B. Formula (4)			
		Coef.	SE	p	p (WCB)	Coef.	SE	p	p (WCB)
(a) In-time placebo on the pre-policy window (2011–2015; counterfeit onset 2013; N = 243)									
Primary Cells (H01M-006)	Placebo 2013	−19.576	8.049	0.026	0.018	−1.590	1.907	0.415	0.444
Fuel Cells (H01M-008)	Placebo 2013	−19.544	8.589	0.035	0.026	−3.933	2.82	0.18	0.203
Secondary Cells (H01M-010)	Placebo 2013	32.188	13.507	0.028	0.003	7.405	4.311	0.103	0.111
Hybrid Cells (H01M-012)	Placebo 2013	−12.175	6.526	0.078	0.072	−4.873	2.358	0.053	0.088
(b) Alternative policy onsets (full sample; N = 533)									
Primary Cells (H01M-006)	Onset 2015	−33.485	8.338	0.001	<0.001	5.662	1.7	0.004	0.01
	Onset 2016 *	−29.460	6.47	<0.001	<0.001	5.89	1.798	0.004	0.008
	Onset 2017	−28.899	7.498	0.001	0.001	4.995	2.107	0.029	0.033
Fuel Cells (H01M-008)	Onset 2015	−20.138	9.59	0.049	0.046	−7.199	4.214	0.104	0.15
	Onset 2016 *	−17.485	9.476	0.081	0.076	−5.553	4.517	0.234	0.309
	Onset 2017	−17.522	10.222	0.103	0.097	−3.312	4.38	0.459	0.501
Secondary Cells (H01M-010)	Onset 2015	31.768	9.595	0.004	0.003	3.792	4.385	0.398	0.435
	Onset 2016 *	29.886	10.648	0.011	0.013	1.839	4.433	0.683	0.689
	Onset 2017	27.062	10.626	0.02	0.022	0.156	4.764	0.974	0.975
Hybrid Cells (H01M-012)	Onset 2015	−17.790	6.701	0.016	0.028	1.907	1.785	0.299	0.347
	Onset 2016 *	−20.500	7.693	0.015	0.031	2.125	2.34	0.375	0.416
	Onset 2017	−19.932	6.801	0.009	0.033	2.536	1.682	0.148	0.137
* The 2016 rows reproduce the interaction coefficients of Table 11.									

* The 2016 rows reproduce the interaction coefficients of Table 11.

Table 12. Cont.

Sub-Technology (G = 1)	Specification	Panel A. Baseline				Panel B. Formula (4)			
		Coef.	SE	<i>p</i>	<i>p</i> (WCB)	Coef.	SE	<i>p</i>	<i>p</i> (WCB)
<i>(c) Count and log-outcome specifications (full sample; N = 533; semi-elasticities)</i>									
Primary Cells (H01M-006)	PPML	−0.703	0.233	0.007	0.01	0.017	0.198	0.935	0.955
	Log OLS	−0.820	0.17	<0.001	<0.001	0.072	0.154	0.645	0.746
Fuel Cells (H01M-008)	PPML	−0.389	0.263	0.155	0.154	−0.052	0.158	0.746	0.738
	Log OLS	−0.529	0.287	0.081	0.082	−0.279	0.149	0.076	0.085
Secondary Cells (H01M-010)	PPML	0.222	0.26	0.404	0.424	−0.228	0.138	0.115	0.111
	Log OLS	0.505	0.274	0.081	0.075	−0.178	0.126	0.175	0.166
Hybrid Cells (H01M-012)	PPML	−0.242	0.328	0.47	0.497	0.268	0.222	0.243	0.341
	Log OLS	−0.245	0.341	0.482	0.492	0.355	0.26	0.189	0.226

Notes: Each row reports the coefficient on the group-by-post-period interaction; panels, rotation of the exposed group, applicant-clustered standard errors and the restricted wild cluster bootstrap (WCB) follow Table 11. (a) The sample is the pre-policy window only (application years 2011–2015; $N = 243$ cells on 19 applicants) with a counterfeit onset at 2013; the number of cells in which the exposed group and the counterfeit post-period coincide ranges from 19 (hybrid) to 56 (secondary), so the test has little power against small differentials. Moving the counterfeit onset to 2014 leaves all four Panel A coefficients significant at the 5% level and all four Panel B coefficients insignificant. (b) Only the onset year changes. * The 2016 rows reproduce the interaction coefficients of Table 11. Adding applicant fixed effects to Panel B preserves the pattern at the 2015 and 2016 onsets; at a 2017 onset the primary-cell interaction falls to the 10% level and the hybrid-cell interaction rises to the 5% level, so no single row should be read as an identified effect. (c) PPML is Poisson pseudo-maximum likelihood with a log link on the untransformed count (1999 bootstrap replications, all converged); Log OLS is OLS on its natural logarithm, defined for every cell because no cell is zero. Coefficients are semi-elasticities and are not comparable in magnitude with those of Table 11: $\exp(b) - 1$ gives −50.5% and 1.7% for primary cells in Panels A and B respectively. Adding applicant fixed effects to Panel B leaves every interaction insignificant under both estimators, the largest remaining being the fuel-cell coefficient in the log specification (−0.2925, bootstrap $p = 0.061$). The Poisson pseudo- R^2 of Panel B lies between 0.794 and 0.830, against 0.940 for the corresponding linear model.

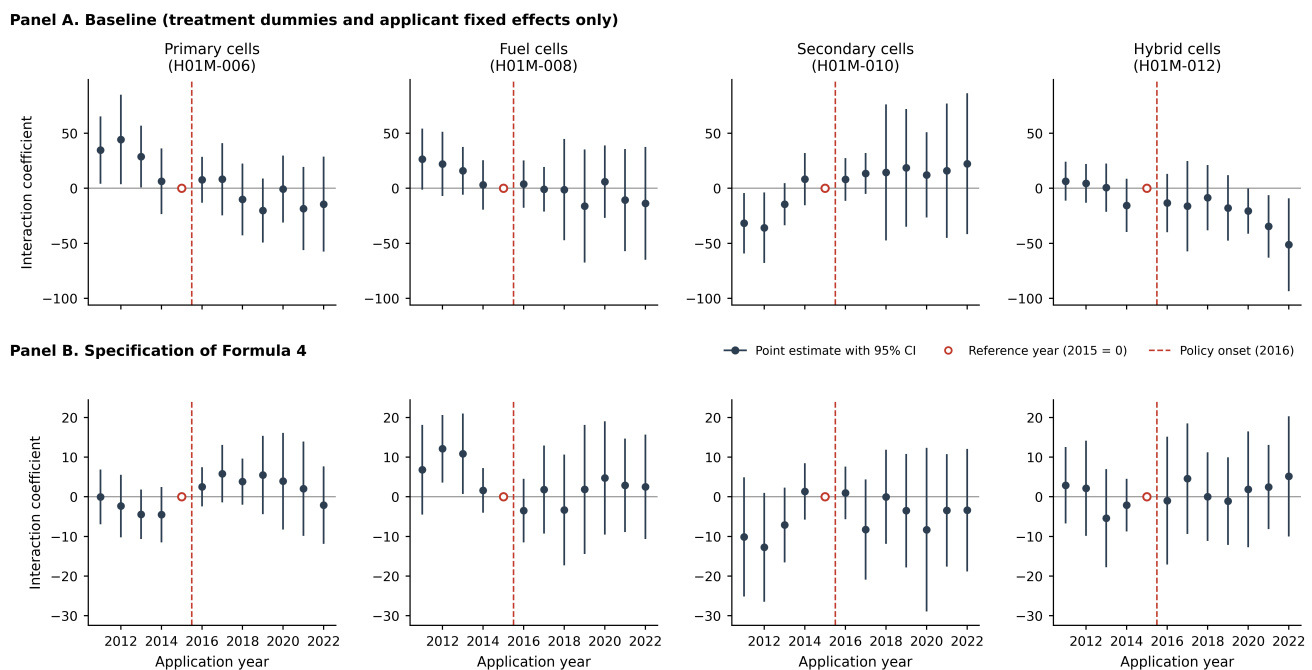


Figure 10. Event-study estimates of the group-by-year interactions, 2011–2022. Dots and vertical bars denote point estimates and 95% confidence intervals; the red hollow circle marks the reference year (2015), normalized to zero, and the red dashed line marks the policy onset (2016). Notes: Each panel plots the coefficient on the interaction between the group dummy and a filing-year indicator, with 2015 as the omitted reference year (hollow marker); vertical bars are 95% confidence intervals and the dashed line marks the policy onset. Sample, panels and inference follow Table 11. The joint test that the four pre-policy leads are zero yields bootstrap p -values of 0.076, 0.336, 0.211 and 0.598 in Panel (A) and 0.470, 0.052, 0.268 and 0.223 in Panel (B), taking the columns in the order shown; the corresponding test on the seven post-policy lags rejects in no column of either panel (all $p \geq 0.18$). Adding applicant fixed effects to Panel B changes nothing except that the joint lead test for fuel cells then rejects at the 5% level ($F = 3.658$, $p = 0.031$).

The two specifications behave in opposite ways. In the covariate-free baseline the placebo interaction is significant at the 5% level for three of the four sub-technologies and at the 10% level for hybrid cells, and the estimates carry the same signs as the corresponding Panel A estimates in Table 11 and roughly two-thirds of their magnitude. The covariate-free specification therefore assigns differentials of almost the same size to an interval in which no policy occurred. This is why Panel A of Table 11 is reported as a decomposition of the model's fit rather than as an identified policy differential.

Under the specification of Formula (4) no placebo interaction is significant at the 5% level. Hybrid cells are the single marginal case ($p = 0.088$) and disappear when the counterfactual onset is moved to 2014. The one interaction that is significant in Panel B of Table 11, that of primary cells, is not reproduced by the placebo, so that estimate is not an artifact of the pre-policy filing path. Read together with Table 10, the two diagnostics point the same way: the pre-policy behavior of the covariate-adjusted specification is consistent with the design and that of the covariate-free specification is not. This asymmetry does not make Formula (4) the preferred model, since the placebo leaves the post-treatment status of its covariates untouched.

A final question concerns the date at which the post-policy window is opened. Because the Agreement was adopted in December 2015 and entered into force in November 2016, 2016 is one defensible onset rather than the only one. We therefore re-estimated both panels with the post-period beginning in 2015 and in 2017. Table 12 reports the interaction coefficients at the three onsets, and they are close to invariant to the choice: under the

specification of Formula (4) the primary-cell interaction remains positive and significant at all three dates while no other column reaches significance at any of them, and in the covariate-free panel the signs and the significance pattern are unchanged throughout. The reported differentials are therefore not an artifact of the date chosen; equally, because no discrete break occurs at any of these dates, the interaction measures a difference in level between two halves of the observation window rather than a response beginning in a given year.

Finally, because the outcome is a heavily overdispersed count and the linear specification returns negative fitted values for about a fifth of the cells, Table 12 re-estimates both specifications by Poisson's pseudo-maximum likelihood [55,56] and on the logarithm of the count. Under Formula (4) no interaction remains distinguishable from zero under either estimator, the primary-cell coefficient of Table 11 included; the linear estimates are retained for comparability with the originally reported specification, not because they fit better.

4.3.2. Developing R&D Strategy

Primary batteries and hybrid cells are excluded from the technology development objectives on the basis of filing volume and trajectory: annual filings in primary batteries fall from 21 in 2011 to 4 in 2022, and in hybrid cells from 15 to 6, leaving both at a scale too small to anchor a development strategy. The proposed strategy therefore focuses on fuel cells and secondary batteries, the two fields that sustain substantial filing volume throughout the window. In primary and hybrid cells the technology-element change index is dominated by NTE (82.6% and 57.9% of the index total, respectively), whereas in fuel cells and secondary cells the index is dominated by NTF (54.7% and 60.2%). Both are the battery types most frequently named as explicit targets of national deployment mandates and support programs in recent energy-transition policy documents [57,58]. Because the interaction coefficients in Table 11 are sensitive to the covariate specification, the strategy is differentiated by the composition of technological change in each field rather than by a ranking of estimated policy effects.

4.3.3. Fuel Cells: Development of Core Technologies Through Collaboration with Research Institutions

Neither panel in Table 11 identifies fuel cells as the sub-technology with the largest policy-associated differential. The interaction is negative in both panels and insignificant under clustered inference (Panel B: -5.5528 , $p = 0.234$); as shown in Section 4.3.1, the negative sign records growth slower than the rest of the battery portfolio, not contraction, and fuel cells are selected for the reasons set out in Section 4.3.2 rather than on the basis of that coefficient. Contrary to the trend for typical patents, an increase in the number of inventors has a negative impact on technology filings, which indicates that a small group of experts is filing most patents. Technologies with a high policy impact are often associated with various regulatory and promotion policies, which necessitate long-term technology development. Therefore, rather than pursuing independent R&D, it is advisable to select research topics that leverage government support, such as joint research with national institutions, and to focus on technology areas that are less sensitive to policy.

The relevant comparison across fields is the internal composition of the index. In fuel cells, 95.6% of the index total is carried by NTE and NTF combined (40.9% and 54.7%), with VTF accounting for only 4.3%—the lowest share among the four sub-technologies. The same contrast survives normalization by filing volume (Table 9). This exploratory, foundation-building pattern makes shared foundational elements—rather than proprietary ones—the natural locus of effort. Table 13 highlights the connections between technology elements in fuel cells. Potential research topics include technologies related to cell maintenance (coolant and wetness), electromagnetic characteristics of cells (voltage, impedance, and measure),

efficient distribution methods in fuel cells (manifold + channel, connect, plate, stack, and surround), and stable management of fuel and oxidizers (gas + leakage, combustor, injector, supply, and burn).

Table 13. Technology elements in fuel cells.

Existing	New	Existing	New	Existing	New
Cell	Coolant	Manifold	Channel	Gas	Leakage
	Wetness		Connect		Combustor
	Voltage		Plate		Injector
	Impedance		Stack		Supply
	Measure		Surround		Burn

In summary, for a technology development company, a long-term plan involving collaboration with research institutions would be more suitable than independent R&D. To this end, companies may seek experts in foundational technologies that broadly apply to fuel cells. Separately, policymakers must review policies that promote or regulate foundational technologies with significant policy impacts. It is also necessary to identify technology elements that have been hindered by regulation and devise strategies for deregulation to facilitate the development of such technology elements.

4.3.4. Secondary Battery: Developing Proprietary Technologies or Benchmarking/Mergers and Acquisitions with Leading Companies in Target Technology Fields

Numerous patents are filed in the field of secondary batteries due to the inherent characteristics of the technology itself. Moreover, as with fuel cells, the connection between new and existing technology elements influences patent filings most prominently. Unlike fuel cells, secondary cells combine active recombination with substantial intensification of established links: VTF accounts for 10.3% of the index total, the highest share among the four sub-technologies, and grows by 417% between the pre- and post-policy windows. This affords companies both the incentive and the flexibility to develop proprietary technologies to gain a competitive advantage. Table 14 presents the connections among emerging technology elements along with the applicants holding patents in this field. Unlike with fuel cells, the connections among technology elements vary significantly across applicants. Detailed classifications, such as lead-acid batteries, battery management systems, and batteries for vehicles, are established as well. The data indicate that companies such as Amperex and LG Chemical focus on the chemical composition aspects of battery technology.

Table 14. Technology elements in secondary battery field by applicant.

Existing	New	Applicant	Existing	New	Applicant
Battery	Acid	GS Yuasa INTERNATIONAL	Formula	BO ₂	CONTEMPORARY AMPEREX TECHNOLOGY
	Information	LG Energy		XO ₄	
	Housing	Ford		Pyrophosphate	LG CHEMICAL
	Bus			Metaphosphate	
	Vehicle			LiMM ₄ O	

Based on this information, technology developers could identify the technology elements and leading players currently relevant in the field to use as benchmarks or consider for mergers and acquisitions. From a policymaker's perspective, policies should be devised to manage potential situations arising from advancements in each technology field or to prepare for technology monopolies.

4.4. Descriptive Corroboration of the Proposed Strategies

Two technology development strategies are proposed herein based on whether the policy impact is statistically significant: (i) developing common technology elements through collaboration with research institutions and (ii) focusing on independent technology elements. To examine whether filing behavior is consistent with these strategies, changes in R&D activity before and after the Paris Climate Agreement are compared for each applicant. Because the pre- and post-policy windows differ in length, all counts are annualized before comparison, and the applicant is the matched unit throughout. A paired-sample *t*-test and, because applicant-level counts are strongly right-skewed, a Wilcoxon signed-rank test are reported for each comparison. Neither test controls for concurrent time trends, industry growth, or changes in applicant composition; the comparisons in this section are therefore reported as descriptive corroboration, and no causal interpretation is drawn from them. The proposed strategy for the fuel cell field is examined by comparing, for each applicant, the number of patents related to the proposed research topics filed before and after 2016. Only filings in which at least one research institution—a university, institute, or research foundation—appears together with a corporate applicant are counted. The test covers all 44 applicants meeting the minimum-volume criterion, of which Table 15 reports the ten with the largest totals. Annualized filings rise from a mean of 2.27 per applicant-year (SD 3.19) before 2016 to 3.44 (SD 3.78) after, an increase of 51.7%, with 28 of the 44 applicants recording an increase. A paired-sample *t*-test gives $t(43) = -2.465$, $p = 0.018$, and the Wilcoxon signed-rank test (Shapiro–Wilk $W = 0.873$, $p < 0.001$ for the paired differences) gives $p = 0.005$.

Table 15. Changes in patent applications per applicant through collaboration with research institutions.

Applicant	Before 2016	After 2016	Before /Year	After /Year	Total
HYUNDAI MOTOR	49	155	8.2	22.1	204
TSINGHUA UNIVERSITY	82	58	13.7	8.3	140
HON HAI PRECISION	79	54	13.2	7.7	133
TOYOTA JIDOSHA	53	52	8.8	7.4	105
SAMSUNG ELECTRONICS	43	60	7.2	8.6	103
KIA MOTORS	5	61	0.8	8.7	66
KYOTO UNIVERSITY	31	30	5.2	4.3	61
LG CHEMICAL	10	49	1.7	7.0	59
KAIST KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY	15	40	2.5	5.7	55
INDUSTRY UNIVERSITY COOPERATION FOUNDATION HANYANG UNIVERSITY	17	33	2.8	4.7	50

Note: Counts are filings in which at least one research institution appears together with a corporate applicant. The ten applicants with the largest totals are shown; the test reported in the text covers all 44 applicants that filed more than ten such patents during 2010–2022. Per-year figures divide the pre-policy count by six years (2010–2015) and the post-policy count by seven years (2016–2022).

No significant relative policy responsiveness was observed in the secondary battery field, where each applicant follows a filing strategy organized around its own research topics. For this examination a broader applicant set extending beyond the top 20—including additional firms and universities—is used. The ten most distinctive research topics for each applicant were identified using the term frequency—inverse document frequency metric (Table 16) and filings containing those topics were counted before and after 2016. The test covers all 47 applicants, of which Table 17 reports the ten with the largest portfolios. Annualized filings on each applicant’s own topics rise from a mean of 14.70 per applicant–

year (SD 27.89) to 23.21 (SD 30.80), an increase of 58.0%, with 32 of the 47 applicants increasing. The paired-sample *t*-test is marginal, $t(46) = -1.949$, $p = 0.057$, and the Wilcoxon signed-rank test (Shapiro–Wilk $W = 0.766$, $p < 0.001$) gives $p = 0.013$; the increase is therefore descriptively supported but not established by the parametric test.

Table 16. R&D themes per applicant.

Applicant	R&D Themes
SAMSUNG SDI	Couple Plate, Plurality Cell, Electrolyte Battery
TOYOTA JIDOSHA	Battery State, Configure Unit, High to Low, Battery Ion
LG CHEMICAL	Electrode Battery, Form Layer, Layer Form, Plurality Cell, Battery Sulfur
LG ENERGY SOLUTION	Module Pack, Plurality Cell, Plurality Module
ROBERT BOSCH	Cell Ion, Hydrogen Particular, Circuit System, Plurality Cell, Battery Ion
HYUNDAI MOTOR	Assembly Electrode, Cathode Stack, Battery Voltage, Plurality Cell
PANASONIC INTELLECTUAL PROPERTY MANAGEMENT	Material Battery, Electrode Battery, Separator Interpose
GM GLOBAL TECHNOLOGY OPERATIONS	Plurality Cell, Layer Dispose, Cell Ion
HONDA MOTOR	Plurality Cell, Electrode Battery, Supply Gas, Direction Rear
CONTEMPORARY AMPEREX TECHNOLOGY	Battery Ion, Collector Layer, Battery Module, Collector Film

Table 17. Changes in patent applications under R&D themes per applicant.

Applicant	Before 2016	After 2016	Before /Year	After /Year	Total
SAMSUNG SDI	1096	598	182.7	85.4	4171
TOYOTA JIDOSHA	172	592	28.7	84.6	3953
LG CHEMICAL	419	834	69.8	119.1	3484
LG ENERGY SOLUTION	11	922	1.8	131.7	2516
ROBERT BOSCH	160	315	26.7	45.0	2091
HYUNDAI MOTOR	86	323	14.3	46.1	1892
PANASONIC INTELLECTUAL PROPERTY MANAGEMENT	72	614	12.0	87.7	1878
GM GLOBAL TECHNOLOGY OPERATIONS	111	241	18.5	34.4	1561
HONDA MOTOR	127	210	21.2	30.0	1516
CONTEMPORARY AMPEREX TECHNOLOGY	6	322	1.0	46.0	1215

Note: “Before 2016” and “After 2016” count filings containing the applicant’s own TF-IDF research topics, whereas “Total” is the applicant’s entire patent portfolio in the sample; the first two columns therefore do not sum to the third. The ten applicants with the largest portfolios are shown; the test reported in the text covers all 47 applicants. Per-year figures divide the pre-policy count by six years and the post-policy count by seven years.

5. Conclusions

This study analyzed factors influencing technology advancement by dividing them into two categories: macro policy and changes in technology-elements. Macro policy—which encompasses various micro policies, such as regulations and financial support—was quantified using the DID approach as a single influencing factor. To utilize the text data contained within patents, which are considered outcomes of technology advancement in terms of novelty, industrial applicability, and inventiveness, we created a

technology element network, interpreting keywords as technology elements. Subsequently, regression analysis was conducted to subdivide technology fields and statistically verify the significance of policy and technology changes. Accordingly, based on filing volume and the composition of technological change in each field, technology strategies were proposed.

This study integrates two research streams that have largely progressed in parallel by combining DID-based policy evaluation with keyword-network analysis of patent texts. While DID-based policy studies are effective in testing whether a policy effect exists, they provide limited insight into how policies shape the internal structure and evolution of technologies. To address this limitation, this study maps the technical meaning embedded in patent texts into a technology-element network and develops a technology change index that captures structural, time-varying change through three complementary components: the emergence of new technology elements (NTE), the formation of new links between elements (NTF), and shifts in link intensity (VTF). For instance, when a sub-technology exhibits a statistically significant post-policy shift relative to control sub-technologies in the DID results, the technology change index clarifies whether the observed change is driven primarily by newly introduced elements (high NTE), by new combinations among elements (high NTF), or by intensified relationships among established elements (high VTF). This design extends policy evaluation beyond effect detection toward mechanism-oriented interpretation, and it operationalizes technology evolution in terms of measurable element-level variation and recombination.

By integrating the technology change index with a DID measure of policy exposure, the proposed methodology links policy-associated shifts in the R&D environment to technology-element dynamics and unifies external-environment analysis and R&D item selection into a single process. The approach constructs coherent technological topics across product types and technology purposes, classifies the forms through which policy influence materializes, and statistically verifies the salience of each type. Building on these outputs, the combined DID–network evidence identifies where post-policy change is heterogeneous across fields and of what it consists, informing policymakers’ prioritization and instrument design. For R&D managers, the resulting map of technological change clarifies an action for R&D activities. In fields whose change consists of new elements being combined for the first time, firms can collaborate around foundational and widely shared elements. In fields where established combinations are being deepened, firms can focus on securing proprietary, performance-critical elements and can additionally rely on targeted benchmarking or mergers and acquisitions to reinforce their competitive positions. In this way, the framework translates macro policy exposure into actionable, field-specific choices.

This study has several limitations. Its design centers on a single policy event in one industry, which constrains generalizability, and it reduces the Paris Agreement to a single onset year: the content and phased implementation of the policy are not modeled, and unobserved shocks coinciding with 2016 cannot be excluded. Because each sub-technology is treated in turn as the exposed group while the remaining three form the comparison group, cross-field spillovers and joint effects are not captured; to the extent such spillovers raise outcomes in the comparison sub-technologies, the reported differentials are lower bounds (Section 3.3.1). Nor can the analysis date the response: under Formula (4) no post-2016 interaction is distinguishable from zero, so the differential reported in Table 11 is an average over seven years rather than a change located at the onset.

Identification is supported only conditionally. Differential pre-policy trends disappear only once the covariates of Formula (4) are included (Table 10), and even then the joint lead test is rejected at the 10% level for primary cells. For fuel cells the joint lead test sits at the conventional boundary in the event study (Figure 10) although not in Table 10; the weaker result should govern, so the fuel-cell interaction in Table 11 is the one estimate in that

Table that should not be read as a policy-associated differential at all (Section 4.3.2). The covariates that secure comparability are themselves post-treatment: the patent index and the technology-element change index are derived from the same documents that generate the dependent variable and may respond to the policy (Section 3.3.1), and removing them increases the interaction coefficients by roughly a factor of five and changes which sub-technology carries the only positive and significant differential. This dual position—necessary for the pre-policy paths to be comparable, yet realized after treatment—is the principal reason the framework should be read as a measurement instrument rather than as an identification strategy. We therefore report both specifications and limit the empirical claim to the presence of heterogeneity across sub-technologies.

The estimates are further limited in magnitude and precision. Under Poisson or log specifications no interaction in Formula (4) remains distinguishable from zero (Table 12); the linear estimates are retained for comparability with the originally reported specification, and none should be converted into a statement about how many filings the policy added. Inference rests on 20 applicant clusters, so the wild cluster bootstrap *p*-values are the appropriate reference and the power to detect modest differentials is limited. The 533 cells are also those in which the technology-element change index is defined, that is, applicants with at least two application years in a sub-technology (Appendix B.3); the sample is thus conditioned on continued filing, describes only applicants already active in a field, and is silent on entry and exit.

Four features of the sample construction warrant note. Applicants were ranked on filings over the full 2010–2022 window rather than on pre-treatment information alone. Recomputing the ranking on 2010–2015 filings returns sixteen of the same twenty applicants (Section 3.1), so the sampling frame is largely insensitive to the window used, but the estimates have not been re-computed on that alternative sample. An extension beyond the top 20 also remains untested and would in any case alter the estimand rather than test it, since the technology-element change index is defined only for cells in which an applicant has at least two application years in a sub-technology (Appendix B.3) and applicants outside the top 20 file too sporadically to populate such cells. The sample is right-truncated at the end of the window: the eighteen-month publication lag leaves a few late-2022 applications unobserved, but the lag is uniform across battery types, is absorbed by the post-2016 indicator, and works against detecting a positive differential. Patents carrying IPC codes in more than one battery-type group enter each of those groups, so the treated and comparison groups are not fully disjoint (Section 3.1). Applicants are identified as the legally separate entities recorded at filing, which is why the LG Energy Solution rows are read at the group level (Section 4.1); a corporate-group-level or assignment-adjusted analysis would be a useful extension. Finally, because the indices are computed from titles and abstracts, they depend on the size of each applicant's corpus and on drafting conventions that may differ across firms and over time.

Future work may accordingly test the framework across other policies and domains, build comparison groups that reflect commonalities across sub-technologies, measure the indices on pre-policy text or model them jointly with filings, and incorporate policy-text mining so that the content of the policy, and not only its onset, enters the analysis.

Author Contributions: Conceptualization, T.R., H.J. and B.Y.; methodology, T.R., H.J. and B.Y.; software, T.R.; validation, T.R. and H.J.; formal analysis, T.R.; investigation, T.R.; resources, B.Y.; data curation, T.R.; writing—original draft preparation, T.R. and H.J.; writing—review and editing, T.R. and H.J.; visualization, T.R. and H.J.; supervision, B.Y.; project administration, B.Y.; funding acquisition, B.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the ANCHOR through the Seoul ANCHOR Center, funded by the Ministry of Education (MOE) and the Seoul Metropolitan Government (2026-ANCHOR-01-007-04) and was also supported by the Korea Institute of Advanced Technology (KIAT) grant funded by the Korea government (MOTIR) (RS-2026-25565742).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The patent data analyzed in this study were retrieved via Wisdomain from the USPTO and WIPO PCT collections. The search criteria and filtering rules used to construct the analytic sample are reported in Section 3.1. The NLP preprocessing steps and the technology-element network construction procedure are reported in Section 3.2.2 and Appendix A.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Text Preprocessing, Network Construction, and Index Computation

The language-filtering rule, the part-of-speech-based stop-word rule, and the network construction procedure are reported in Sections 3.1 and 3.2.2. This appendix records the software, the preprocessing steps not otherwise specified there, and the operational definitions of NTE, NTF, and VTF that underlie Formula (3).

The title and abstract fields were first lower-cased and stripped of non-alphabetic characters with regular expressions; no dedicated cleaning library was used. The cleaned text was tokenized with NLTK (version 3.8.1) and the token sequence was then processed with Stanford CoreNLP (version 4.5.1) for phrase-structure parsing, part-of-speech tagging, and lemmatization, whose output supplies the POS-lemma information illustrated in Figure 6. Each token entered the network as the lemma returned by CoreNLP. No multi-word expressions were merged: a compound such as “light bulb” is represented as two nodes, *light* and *bulb*, joined by an edge, as in the examples of Section 3.2.2, rather than as a single element. No synonym consolidation was applied; lemmatization is the only normalization of the vocabulary.

The three indices are computed as follows. For each applicant, sub-technology, and application year t , a network G_t is built from the titles and abstracts of that cell’s filings: nodes are the retained lemmas, and two nodes are joined by an edge whose weight W_t is the number of times the two lemmas are linked by a dependency relation within the same sentence. W_t is compared with G_{t-1} , the network of the same applicant and sub-technology in its immediately preceding application year (which need not be the preceding calendar year); the 2010 networks serve only as the base for 2011. NTE_t is the number of nodes in G_t that are absent from G_{t-1} . NTF_t and VTF_t are accumulated over the retained nodes, that is, the nodes present in both networks: for each retained node, NTF_t counts the neighbors it has in G_t that it did not have in G_{t-1} , and VTF_t sums $|W_t - W_{t-1}|$ over the edges incident to it that exist in both networks. Because both quantities are accumulated node by node, an edge joining two retained nodes contributes from each endpoint, an edge joining a retained node to a new node contributes once, and an edge joining two new nodes contributes to NTE only; the per-edge quantities illustrated in Table 1 therefore enter the regressions at twice their per-edge value, a constant scaling that does not affect inference.

Appendix B. Search Strategy and Sample Construction

The search criteria and filtering rules used to construct the analytic sample are reported in Section 3.1. The NLP preprocessing steps and the technology-element network construction procedure are reported in Section 3.2.2 and Appendix A.

Appendix B.1. Database and Coverage

Patent records were retrieved from Wisdomain, a commercial patent database indexing the records of the United States Patent and Trademark Office (USPTO) and the Patent Cooperation Treaty (PCT) collection published by the World Intellectual Property Organization (WIPO). No restriction was imposed on the publication stage of the record: both pre-grant publications and granted patents were retrieved. The search was executed in September 2024. Later searches would return additional records for the most recent application years (Section 5).

Appendix B.2. Search Query

The following query was executed against the combined USPTO and PCT collections:

$$\text{ICH} = (\text{H01M}) \text{ AND } (\text{AD} \geq 2010 \text{ AND } \text{AD} < 2023)$$

Field codes are defined as follows. ICH is the IPC search field; a record matches if any IPC code assigned to it, whether in the inventive (main) or the additional position, falls under subclass H01M. AD is the application date, so the restriction selects applications filed in the calendar years 2010 through 2022 inclusive. No restriction was imposed on applicant, country, language, or publication stage.

Appendix B.3. Filtering Sequence

Filters were applied in the following order, each operating on the output of the previous step.

1. Execute the query in B.2. → 231,099 records.
2. Consolidate records sharing an INPADOC patent-family identifier into a single record, retaining the member with the earliest application date and resolving ties by the country precedence US, EP, CN, KR, JP.
3. Where the retained member is not in English, substitute an English-language member of the same family; exclude families with no English-language member. → 103,235 records after steps 2 and 3 (127,864 records removed, 55.3% of step 1). Of these, 69,148 are U.S. pre-grant publications, 27,544 are PCT publications, and 6543 are granted U.S. patents.
4. Rank applicants by number of records over the full 2010–2022 window, using the standardized applicant-name field described in B.4, and retain the top 20. → 29,573 records (28.6% of step 3).
5. Assign each record to every battery-type group whose IPC code it carries, as described in Section 3.1; aggregate to the applicant $\times$ sub-technology $\times$ year level; and retain cells for which the technology-element change index is defined, that is, cells in which the applicant has at least two application years in that sub-technology, the index being computed between successive application years (Section 4.2). → Of 533 observations, 243 fell in the pre-policy window used for the diagnostics in Table 10. Of the 103,235 records retained at step 3, 73,366 carry at least one of the four battery-type codes and 1764 carry codes in more than one of them.

The title and abstract fields of the records retained at step 5 provide the text input for the technology-element network (Section 4.2); the preprocessing applied to them is reported in Appendix A.

Appendix B.4. Applicant Name Consolidation

Applicant identity was taken from the standardized applicant-name field supplied by Wisdomain, which resolves variant renderings of the same recorded name—differences in

transliteration, abbreviation, spacing, punctuation, and spelling—into a single canonical form. The following conventions apply. (i) The database field was used as supplied; no further manual consolidation was performed, and legal-form designators retained in that field (Ltd., Inc., Corp., GmbH, KK, and equivalents) were left in place, so the canonical names are those listed in Table 5. (ii) Records naming more than one applicant were split, and each named applicant was credited with one record, so the counts in Table 5 consequently sum to more than the number of distinct records. (iii) No consolidation was applied across distinct legal entities. In the terminology of the applicant-harmonization literature, the procedure performs name harmonization but not legal-entity harmonization, which would additionally resolve mergers, acquisitions, and subsidiary relations between distinct names [59,60]. LG Chemical and LG Energy Solution, Panasonic and Panasonic Intellectual Property Management, and Panasonic and Sanyo Electric are accordingly retained as separate applicants (Section 4.1). (iv) Applicant identity was taken as recorded at the time of filing. Assignments, mergers, and acquisitions occurring after the application date are not traced. The applicant fixed effects in Formula (4) are defined at the level of the twenty canonical names in Table 5.

References

1. Jang, H.; Roh, T.; Yoon, B. Public values in public R&D through natural language processing. *Sci. Rep.* **2025**, *15*, 37397. [[CrossRef](#)] [[PubMed](#)]
2. Guo, C.; Niu, Y. Research on synergy of science and technology policy types and policy tools based on analysis of policy text. *J. Phys. Conf. Ser.* **2021**, *1774*, 012006. [[CrossRef](#)]
3. Wang, D.; Li, Y. Measuring the policy effectiveness of China's new-energy vehicle industry and its differential impact on supply and demand markets. *Sustainability* **2022**, *14*, 8215.
4. Choung, J.-Y.; Hameed, T.; Ji, I. Catch-up in ICT standards: Policy, implementation and standards-setting in South Korea. *Technol. Forecast. Soc. Change* **2012**, *79*, 771–788. [[CrossRef](#)]
5. Sharma, K.; Al-Zaili, J.; Sayma, A. A system dynamics approach to assess the impact of policy interventions on the market penetration of micro gas turbines. In Proceedings of the ASME Turbo Expo 2023: Turbomachinery Technical Conference and Exposition, Boston, MA, USA, 26–30 June 2023. [[CrossRef](#)]
6. You, T.; Jung, W.-S. A system dynamics analysis of national R&D performance measurement system in Korea. *arXiv* **2019**, arXiv:1901.05447.
7. An, K.; Zhang, S.; Zhou, J.; Wang, C. How can computable general equilibrium models serve low-carbon policy? A systematic review. *Environ. Res. Lett.* **2023**, *18*, 033002. [[CrossRef](#)]
8. Shahriari, S.; Robson, E.N.; Wang, J.; Dixit, V.V.; Waller, S.T.; Rashidi, T.H. Integrating a computable general equilibrium model with the four-step framework. *Transportation* **2023**, *50*, 1213–1260. [[CrossRef](#)]
9. Abadie, A. Semiparametric difference-in-differences estimators. *Rev. Econ. Stud.* **2005**, *72*, 1–19. [[CrossRef](#)]
10. Liu, D.; Xiao, B. Exploring the development of electric vehicles under policy incentives: A scenario-based system dynamics model. *Energy Policy* **2018**, *120*, 8–23. [[CrossRef](#)]
11. Miyata, Y.; Shibusawa, H.; Permana, I.; Wahyuni, A. Computable general equilibrium models for economic and environmental policies. In *Environmental and Natural Disaster Resilience of Indonesia*; Springer: Singapore, 2018; pp. 137–147.
12. Robinson, M.; Imran, A.; Barlow, M. Impact of macro-level regulations on micro-level (individual) use of ICT: A case of mobile services in India. In Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2018), Galway, Ireland, 4–6 April 2018. [[CrossRef](#)]
13. Hassan, T.; Khan, Y.; He, C.; Chen, J.; Alsagr, N.; Song, H. Environmental regulations, political risk and consumption-based carbon emissions: Evidence from OECD economies. *J. Environ. Manag.* **2022**, *320*, 115893. [[CrossRef](#)] [[PubMed](#)]
14. Cabral, A.C.G.; O'Reilly, P.; Appelt, S.; Galindo-Rueda, F.; Hanappi, T. Design features of income-based tax incentives for R&D and innovation. In *OECD Taxation Working Papers*, 60; OECD Publishing: Paris, France, 2023. [[CrossRef](#)]
15. Holt, J.; Skali, A.; Thomson, R. The additionality of R&D tax policy: Quasi-experimental evidence. *Technovation* **2021**, *107*, 102293. [[CrossRef](#)]
16. Hou, X.; Hu, Q.; Liang, X.; Xu, J. How do low-carbon city pilots affect carbon emissions? Staggered difference-in-difference evidence from Chinese firms. *Econ. Anal. Policy* **2023**, *79*, 664–686. [[CrossRef](#)]
17. Wang, K.-L.; Sun, T.-T.; Xu, R.-Y.; Miao, Z.; Cheng, Y.-H. How does internet development promote urban green innovation efficiency? Evidence from China. *Technol. Forecast. Soc. Change* **2022**, *184*, 122017. [[CrossRef](#)]

18. Büttner, B.; Firat, M.; Raiteri, E. Patents and knowledge diffusion: The impact of machine translation. *Res. Policy* **2022**, *51*, 104584. [\[CrossRef\]](#)
19. Pippel, G.; Seefeld, V. R&D cooperation with scientific institutions: A difference-in-difference approach. *Econ. Innov. New Technol.* **2016**, *25*, 455–469. [\[CrossRef\]](#)
20. Gu, J. Effects of patent policy on outputs and commercialization of academic patents in China: A spatial difference-in-differences analysis. *Sustainability* **2021**, *13*, 13459.
21. Gao, K.; Yuan, Y. The effect of innovation-driven development on pollution reduction: Empirical evidence from a quasi-natural experiment in China. *Technol. Forecast. Soc. Change* **2021**, *172*, 121047. [\[CrossRef\]](#)
22. de Rassenfosse, G.; Palangkaraya, A. Do patent pledges accelerate innovation? *Res. Policy* **2023**, *52*, 104745. [\[CrossRef\]](#)
23. Li, G.; Xue, Q.; Qin, J. Environmental information disclosure and green technology innovation: Empirical evidence from China. *Technol. Forecast. Soc. Change* **2022**, *176*, 121453. [\[CrossRef\]](#)
24. Altuntas, S.; Dereli, T.; Kusiak, A. Forecasting technology success based on patent data. *Technol. Forecast. Soc. Change* **2015**, *96*, 202–214. [\[CrossRef\]](#)
25. Petralia, S. Mapping general purpose technologies with patent data. *Res. Policy* **2020**, *49*, 104013. [\[CrossRef\]](#)
26. Yuan, X.; Li, X. A network analytic method for measuring patent thickets: A case of FCEV technology. *Technol. Forecast. Soc. Change* **2020**, *156*, 120038. [\[CrossRef\]](#)
27. Fabrizi, A.; Guarini, G.; Meliciani, V. Green patents, regulatory policies and research network policies. *Res. Policy* **2018**, *47*, 1018–1031. [\[CrossRef\]](#)
28. Long, C.X.; Wang, J. China's patent promotion policies and its quality implications. *Sci. Public Policy* **2019**, *46*, 91–104. [\[CrossRef\]](#)
29. Donis, S.; Gómez, J.; Salazar, I. Economic complexity, property rights and the judicial system as drivers of eco-innovations: An analysis of OECD countries. *Technovation* **2023**, *128*, 102868. [\[CrossRef\]](#)
30. Roh, T.; Jeong, Y.; Yoon, B. Developing a methodology of structuring and layering technological information in patent documents through natural language processing. *Sustainability* **2017**, *9*, 2117.
31. Park, I.; Yoon, B.; Kim, S.; Seol, H. Technology opportunity discovery for safety through topic modeling and opinion mining in the fourth industrial revolution: The case of artificial intelligence. *IEEE Trans. Eng. Manag.* **2019**, *68*, 1504–1519. [\[CrossRef\]](#)
32. Roh, T.; Jeong, Y.; Jang, H.; Yoon, B. Technology opportunity discovery by structuring user needs based on natural language processing and machine learning. *PLoS ONE* **2019**, *14*, e0223404. [\[CrossRef\]](#)
33. Yang, X.; Yu, X. Preventing patent risks in artificial intelligence industry for sustainable development: A multi-level network analysis. *Sustainability* **2020**, *12*, 8667.
34. Kim, S.; Park, I.; Yoon, B. SAO2Vec: Development of an algorithm for embedding the subject–action–object (SAO) structure using Doc2Vec. *PLoS ONE* **2020**, *15*, e0227930. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Wang, X.; Qiu, P.; Zhu, D.; Mitkova, L.; Lei, M.; Porter, A.L. Identification of technology development trends based on subject–action–object analysis: The case of dye-sensitized solar cells. *Technol. Forecast. Soc. Change* **2015**, *98*, 24–46. [\[CrossRef\]](#)
36. Park, H.; Yoon, J.; Kim, K. Identifying patent infringement using SAO-based semantic technological similarities. *Scientometrics* **2012**, *90*, 515–529. [\[CrossRef\]](#)
37. Ree, J.J.; Jeong, C.; Park, H.; Kim, K. Context–problem network and quantitative method of patent analysis: A case study of wireless energy transmission technology. *Sustainability* **2019**, *11*, 1484. [\[CrossRef\]](#)
38. Choi, J.; Hwang, Y.-S. Patent keyword network analysis for improving technology development efficiency. *Technol. Forecast. Soc. Change* **2014**, *83*, 170–182. [\[CrossRef\]](#)
39. Popp, D. Induced innovation and energy prices. *Am. Econ. Rev.* **2002**, *92*, 160–180. [\[CrossRef\]](#)
40. Acemoglu, D.; Aghion, P.; Bursztn, L.; Hemous, D. The environment and directed technical change. *Am. Econ. Rev.* **2012**, *102*, 131–166. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Hassler, J.; Krusell, P.; Olovsson, C. Directed technical change as a response to natural-resource scarcity. *J. Political Econ.* **2021**, *129*, 3039–3072. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Aghion, P.; Dechezleprêtre, A.; Hémous, D.; Martin, R.; Van Reenen, J. Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *J. Political Econ.* **2016**, *124*, 1–51. [\[CrossRef\]](#)
43. Porter, M.E.; van der Linde, C. Toward a new conception of the environment-competitiveness relationship. *J. Econ. Perspect.* **1995**, *9*, 97–118. [\[CrossRef\]](#)
44. Peters, M.; Schneider, M.; Griesshaber, T.; Hoffmann, V.H. The impact of technology-push and demand-pull policies on technical change—Does the locus of policies matter? *Res. Policy* **2012**, *41*, 1296–1308.
45. Liu, K.-t.; Chen, C.-H. Formulation of research and development strategy by analysing patent portfolios of key players in the semiconductor industry according to patent strength and technical function. *World Pat. Inf.* **2022**, *70*, 102125. [\[CrossRef\]](#)
46. Malerba, F. Sectoral systems of innovation and production. *Res. Policy* **2002**, *31*, 247–264.
47. Fleming, L. Recombinant uncertainty in technological search. *Manag. Sci.* **2001**, *47*, 117–132. [\[CrossRef\]](#)
48. March, J.G. Exploration and exploitation in organizational learning. *Organ. Sci.* **1991**, *2*, 71–87. [\[CrossRef\]](#)

49. Weitzman, M.L. Recombinant growth. *Q. J. Econ.* **1998**, *113*, 331–360. [[CrossRef](#)]
50. Autor, D.H. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *J. Labor Econ.* **2003**, *21*, 1–42. [[CrossRef](#)] [[PubMed](#)]
51. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-based improvements for inference with clustered errors. *Rev. Econ. Stat.* **2008**, *90*, 414–427. [[CrossRef](#)]
52. Goodman-Bacon, A. Difference-in-differences with variation in treatment timing. *J. Econom.* **2021**, *225*, 254–277. [[CrossRef](#)]
53. de Chaisemartin, C.; D’Haultfœuille, X. Two-way fixed effects estimators with heterogeneous treatment effects. *Am. Econ. Rev.* **2020**, *110*, 2964–2996. [[CrossRef](#)]
54. Bertrand, M.; Duflo, E.; Mullainathan, S. How much should we trust differences-in-differences estimates? *Q. J. Econ.* **2004**, *119*, 249–275. [[CrossRef](#)]
55. Santos Silva, J.M.C.; Tenreyro, S. The log of gravity. *Rev. Econ. Stat.* **2006**, *88*, 641–658. [[CrossRef](#)]
56. Wooldridge, J.M. Distribution-free estimation of some nonlinear panel data models. *J. Econom.* **1999**, *90*, 77–97. [[CrossRef](#)]
57. International Energy Agency. *The Future of Hydrogen: Seizing Today’s Opportunities*; IEA: Paris, France, 2019; Available online: <https://www.iea.org/reports/the-future-of-hydrogen> (accessed on 18 September 2026).
58. International Energy Agency. *Batteries and Secure Energy Transitions*; IEA: Paris, France, 2024; Available online: <https://www.iea.org/reports/batteries-and-secure-energy-transitions> (accessed on 18 September 2026).
59. Magerman, T.; Van Looy, B.; Song, X. Data production methods for harmonised patent statistics: Patentee name harmonisation. In *Eurostat Working Papers and Studies*; Publications Office of the European Union: Luxembourg, 2006.
60. Thoma, G.; Torrisi, S.; Gambardella, A.; Guellec, D.; Hall, B.H.; Harhoff, D. *Harmonizing and Combining Large Datasets: An Application to Firm-Level Patent and Accounting Data*; National Bureau of Economic Research: Cambridge, MA, USA, 2010. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.