Firm Growth through Product Offerings[*†]

Leo Liu, Elvira Sojli, Wing Wah Tham

Work in Progress

Do not circulate or cite without permission

Abstract

We use textual analysis of high-dimensional data from USPTO trademark documents to develop novel measures of product offerings for public and private firms across all industries from 1926-2020. We classify products, both goods and services, into three categories: 1) business stealing, 2) quality improving, and 3) new variety products and measure the economic value of each product. The USPTO's stringent requirements of 1) evidence through a specimen and detailed description on how the mark is used in commerce; and 2) the date of first use of the mark in commerce for trademark applications, maintenance, and renewal, allow us to accurately track firms' evolving product portfolios. Our measures capture the evolution of product waves over the last half century and the emergence of breakthrough (new variety) products. Stock market reactions around the products' first-use-in-commerce align with future product sales. Firms introduce fewer new variety products, yet these offerings contribute significantly to firm growth. Firms focusing on business stealing grow through acquisitions, whereas firms make more R&D investments to fuel new variety creation.

Keywords: Trademarks, Product Value, New Variety, Quality Improvement, Business Stealing, Endogenous Growth

[*]We thank Jan Bena, Kai Li, Gordon Phillips, Scott Stern (discussant) and participants at the University of Luxembourg, UBC Summer Conference, AIEA-NBER Innovation and Entrepreneurship Conference for very helpful suggestions.

[†]Sojli and Tham at University of New South Wales, Liu at University of Technology Sydney. Contact: leo.liu@uts.edu.au (Liu), e.sojli@unsw.edu.au (Sojli), w.tham@unsw.edu.au (Tham).

1 Introduction

The portfolio of goods and services produced by firms, collectively referred to as products henceforth, and its associated economic value are crucial for the economic modeling of firms' production and investment decisions, resource allocation, product innovation, boundaries and competitiveness of firms, and aggregate output. In economic models, firms produce business stealing, quality improving, and new variety products (see, e.g., Schumpeter, 1939; Romer, 1990; Aghion and Howitt, 1992; Acemoglu, 2003; Klette and Kortum, 2004; Akcigit and Kerr, 2018). The contribution of each of these product types to firm growth matters for resource allocation, investment, R&D, and financing decisions. However, it is often challenging to directly test the rich predictions related to the "extensive" margin of firm adjustment and allocation of resources across these product types because of the lack of directly observable measures of the set and the economic value of firms' products for representative samples.

Economists have attempted to obtain direct measures of product types using scanner data for nondurable consumer goods and U.S. Census data for manufacturing firms.[1] However, they recognize the limitations inherent in these data sets, including issues with sample representativeness, the brevity of time series, and their suitability for evaluating the quality of service provided by newly established and exiting firms in sectors such as business services. Given these measurement challenges, some economists opt for second-best indicators and employ a range of indirect measures, such as analyzing patterns of job creation and destruction, see Garcia-Macia, Hsieh, and Klenow (2019).

To determine product types, an ideal data set would include: 1) market entry dates of new products; 2) detailed descriptions of the products to enable classification into categories; 3) standardized metric of product economic value that allows for comparisons across firms and over time; and 4) extensive coverage of products across time, industries, and firms. While achieving a perfect measure may remain elusive, this paper introduces new metrics

[1] Bernard et al. (2010) highlight the significant role of product creation and destruction in contributing to aggregate output in the manufacturing sector, using U.S. Census data. Studies utilizing scanner/supermarket data, such as Chevalier, Kashyap, and Rossi (2003), Broda and Weinstein (2006), and Hottman, Redding, and Weinstein (2016), have demonstrated the potential to analyze products at more granular levels of disaggregation, although these datasets are typically limited to one sector. Hoberg and Phillips (2016) introduce a time-varying industry classification based on text-based cosine similarity of the 10-K business description sections, which covers US-domiciled listed firms from 1996.

and classifications for firm products that we believe meet these criteria and address many of the shortcomings associated with traditional measures. Leveraging textual analysis of previously untapped examiner-validated information in United States Patent and Trademark Organization (USPTO) trademarks spanning from 1926 to 2020, we develop measures for the key goods and services produced by firms, alongside a market-based valuation for each product. In our data, "a product" is one of approximately 52,000 categories equivalent to nine-digit Standard Industrial Classification (SIC) categories, e.g., basketball shoes. We demonstrate that these measures and economic values positively correlate with major award-winning products, products disclosed in Form 10-K, and future sales performance.

In particular, we introduce a new direct measure of firm-level products (physical goods and services) for all firms operating in the US economy (private and public), categorizing them into: 1) new variety, 2) quality improving, and 3) business stealing products, and obtain an economic value estimate for each product introduced by US publicly listed firms from 1926 to 2020. We use these measures to shed light on two hotly debated issues in the economic and finance literature, first, the impact of each of these product types on firm growth, employment, productivity, and capital investment; second, how corporate investment react to changes in product portfolio composition across these three product types.

Our analysis reveals that while the significance of product types varies, new variety products, such as desktops in the 70s, have a substantial influence on firm growth over the past half-century. Contrary to the current assumptions that incumbents and more mature firms produce little or no new varieties, we find that both new and more mature incumbents generate new variety products, and they are critical for growth. Business stealing products also play an important role for growth, capital investment, and employment. The introduction of business-stealing products tends to have a delayed yet more sustained effect (3-10 year time frame) on profitability, employment, and capital investments. In contrast, quality-improving products, which often cannibalize existing offerings within a firm, typically have a short-term impact on firm growth, observable within three years. Furthermore, expansions through business stealing are significantly associated with increases in acquisitions, while increases in R&D investments are positively associated with new variety products.

Our proposed measures and the categorization of data at individual firm (both private and public), industry, and economy level over an extended time series provide fresh avenues for examining various product-related questions in industrial organization, international trade,

finance, and economic growth. For instance, our findings can inform the selection of firm growth models that most consistently align with the empirical data and provide new insights on product-related firm financing decisions.

Our measures are based on USPTO registered trademarks. A trademark is a type of intellectual property, which identifies and distinguishes goods or services of one party from those of others. Most firms trademark their key products, because trademarks have low granting requirements, are not expensive to acquire, and do not expire as long as the product is still used in commerce. Trademarks are applicable to essentially all products and are the most widely used form of IP protection (Hall, Helmers, Rogers, and Sena, 2014), thus trademarked products cover a large cross-section of firms and industries. For example, Johnson and Johnson states in its 2017 10-K "*The Company's subsidiaries have made a practice of selling their products under trademarks and of obtaining protection for these trademarks by all available means. These trademarks are protected by registration in the U.S. and other countries where such products are marketed. The Company considers these trademarks in the aggregate to be of material importance in the operation of its businesses.*" Over 88% of U.S. listed firms have at least one trademark. Empirically, we find extensive coverage of firms' products through trademarks, where more than 87.6% of the products disclosed by firms in annual 10-K filings are trademarked.[2]

The rich information in the USPTO trademark data (containing trademark classes, detailed product descriptions, application, first use-in-commerce, renewal, and expiry dates) facilitates the construction and tracking of a firm's product portfolio across a long time series.[3,4] Trademarks are established by use in commerce, and trademark data provide two pieces of crucial information: the exact date a product has entered the market (first use-in-commerce date), and the complete description of the product to enable categorization into different product types (Goods and Services, G&S, description). The USPTO requires applicants to provide evidence through a specimen on how the mark is used in commerce and the date of first use of the mark anywhere and in commerce for trademark applications,

[2]The result is consistent with Axarloglou and Theoharakis (2003), which shows that the number of new products advertised on the Wall Street Journal is highly correlated (0.84) with new trademark registrations.

[3]We carefully separate trademarks related to products from those related to marketing and advertising efforts, as described in Section 2.2.

[4]Prior work has not extensively made use of this data, due to its inaccessibility relative to securities filing records and other more conveniently-available data. The trademark data has been cleaned, digitized, and recently made available by the USPTO.

maintenance, and renewal. The first key feature of the data, first use-in-commerce date, coupled with stock market information allows for the construction of a measure of product economic value comparable across firms and time. The information on use-in-commerce and renewal dates also allows for more precise tracking of firms' product portfolios and product life cycles.

The second set of key data features comes from the requirement to clearly and concisely describe the trademark's associated goods and services according to the USPTO ID Manual, which consists of 52,848 categories (term-IDs) of goods and services. New goods and services are regularly added on the ID Manual by the USPTO to record new product varieties.[5] Although each trademark has a unique term ID description, the USPTO organizes and classifies each trademark into broad categories of goods or services (Goods and Services Classification Category), as determined by 1) the Nice International Agreement and 2) the US classification system. The Nice category of goods or services has 45 classes. The Nice and US trademark classifications are analogous to the 2-digit SIC industry classification, while the ID Manual classifies a trademark at the most granular level similar to a nine-digit SIC category. Despite its importance, the mapping of each trademark to a term ID is missing from the USPTO trademark data. Using state-of-the-art NLP techniques, we match each trademark based on the G&S description to the term ID in the ID Manual to provide a finer G&S classification and to construct measure of product variety at the firm, industry, and economy level.

This finer term ID facilitates a categorization of each trademarked product into one of three product economic types: quality improving, business stealing, and new variety products. The three categories are motivated by endogenous growth and industrial organization models in the economics literature. New products that add to a firm's own existing product lines are categorized as quality improving. This classification is motivated by the work of Krusell (1998), Lucas Jr and Moll (2014), and Akcigit and Kerr (2018), where firms can grow through improvements in their own products rather at the expense of their competitors. Products that add a new product line to a firm, but the product line already exists among the firm's competitors are categorized as business stealing or creative destruction. This classification is in line with the theory of creative destruction, which argues that firms can grow

[5]See Chapter 1400 of the USPTO Trademark Manual of Examining Procedure (TMEP) for the classification and identification of goods and services and Section 2.3 for further details.

at the expense of their competitors (see, e.g., Schumpeter, 1939; Aghion and Howitt, 1992; Klette and Kortum, 2004). Any trademarked products that are new to both the firm and any other existing firms are classified as new varieties. This categorization is in line with the work of Romer (1990), Acemoglu (2003), Jones (2016), Broda and Weinstein (2006), where a firm produces a new breakthrough product that is not in its own and competitors' product lines.

We construct these measures for all firms operating in the US (private and publicly-listed) over the sample period from 1926 to 2020. One can readily merge these measures and the USPTO trademark data to financial databases such as CRSP and Compustat for studies that require accounting information and industry classification, using a name-matching algorithm between the trademark applicant names and CRSP firm names.

We estimate the private economic value of each trademarked product, defined as product value, using stock returns around the first use-in-commerce date of each product, in the spirit of Kogan, Papanikolaou, Seru, and Stoffman (2017). This product value is only available for listed firms but allows for the use of economic-value-based weights in constructing aggregated product value at the firm, industry, and economy level. The construction of the economic value-weighted product value relies on the insight that asset prices are forward-looking. The stock market reaction to trademarking a new product provides an estimate of the private economic value of this product to the firm through the market expectation of its future value. We exploit the "first use-in-commerce" requirement in US trademarks to extract a product's first date of commerce and to accurately track firms' product portfolios.

Using the Consumer Electronics Show (CES) Innovation Awards, which acknowledges a firm's significant innovations, our analysis confirms the effectiveness of our product measures in recognizing major award-winning product introductions from 21st century. Moreover, we investigate the relation between the estimated private value of products and their subsequent market sales. Using individual product sales data from Compustat Segments, we establish a positive correlation between the ex ante private market value of a product and its future sales performance.

We explore the effect of product market considerations on firms' decisions and outcomes by examining the relation between firms' product portfolio structure and (i) firm growth and capital allocation, and (ii) the firms' investment decisions. We find that younger, smaller, and less profitable firms in more competitive industries are positively associated with launching

more business stealing products.

Over a ten-year horizon, a one standard deviation increase in a firm's business stealing products' value is associated with an increase of about 12.9% in profits, 11.5% in sales, 12.1% in capital, and 11.3% in employment, controlling for patents. This is consistent with models of creative destruction and previous related empirical tests (Aghion and Howitt, 1992; Lentz and Mortensen, 2008). Quality improving products are associated with short-term growth (1-5 years), profitability (4.9%), sales (4.3%), capital (4.6%), and increases in firm employment (5.2%). On the other hand, new variety products are associated with medium to long-term growth (1-10 years), profitability (6.6%), sales (6.3%), capital (5.4%), and increases in firm employment (4.4%). Our results provide novel insights on the sources of product-driven growth for public firms that improving quality and business-stealing products not only influence firm growth in different ways, but also exhibit distinct short-term and long-term impacts on growth, employment, and capital investments. While the existing literature downplays the role of new variety, our results - leveraging the new measures - highlight the impact of new varieties on growth, challenging the assumptions in the current literature.[6] This finding has important implications for industrial and firm innovation policy.

Furthermore, expansions through business stealing and quality improvement are significantly associated with an increase in acquisitions. This is supported by cases such as Apple's purchase of the fingerprint sensor company AuthenTec Inc., which aimed to enhance the digital security features of iPads and iPhones, and Google's acquisition of YouTube in 2006 to break into the online video sharing market. Our findings suggest that firms can effectively delegate R&D responsibilities to smaller companies for business stealing and quality enhancement, and then acquire those that manage to innovate successfully, rather than conducting all R&D in-house. We observe that increases in R&D investment are positively correlated with the introduction of new product varieties by firms. Factors such as the desire to maintain control over the innovation process, secure a sustained competitive advantage through proprietary technologies, the limited number of target firms at the product frontier, and concerns over intellectual property rights might prompt firms to invest in in-house R&D for developing brand new varieties.

Our work is broadly related to various economic literatures studying the effect and de-

[6]For example, Garcia-Macia, Hsieh, and Klenow (2019) estimated the impact of new variety on aggregate growth to be small relative to quality improvement and business stealing.

terminants of product market dynamics, such as the endogenous growth, industrial organization, economics of innovation, international trade, and finance literatures. One of the challenges in these literatures is the lack of measures of changes in product portfolios; how they are created, compete, and evolve across a broad set of public and private firms over a long time. The measurement issue has hindered the investigations of many important economic questions across these fields on product market dynamics. We contribute to the general economic literature by presenting new metrics and classifications for firm products that overcome many limitations inherent in traditional measures. In particular, we provide a new product data set that has 1) the broadest product coverage of public and private firms, 2) economic value of these products for public firms, 3) the ability to classify products into quality improvement, new variety, and business stealing for both services and physical goods provided by firms, 4) the ability to track the entry and exit of products, 5) the longest time series coverage of firm product dynamics.

More specifically, our paper speaks to the literature on growth models that aims to study the role of product variety, product quality, and business stealing for firm outcomes. Understanding how much each channel contributes to firm growth matters for innovation policy because different channels have different welfare effects. It is empirically challenging to estimate the importance of these different channels. Our comprehensive direct measure over a long time series coupled with the estimate of product private value allows us to directly estimate the contribution of each channel as we observe the changes in the product portfolio across different product types for each firm. We document the following new findings. First contrary to the widely-used assumption that new variety products are developed by new entrants, we find that incumbents play an important role in new variety creation. Second, the estimated significance of new varieties on firm growth higher than the effect of quality improving products. In the Garcia-Macia, Hsieh, and Klenow (2019) analysis of aggregate growth, new varieties come from only new entrants and have little importance for aggregate growth. Third, we find that business stealing products have the most significant impact on firm growth. New variety products contribute to firm growth in the medium to long run, and have a higher impact than quality improvement products. Our analysis also adds additional insights that not only quality-improving and business stealing products affect firm growth differently, but also they have different short and long-run effects on growth, employment, and capital investments.

We add to the important work of Kogan et al. (2017), which estimates the private value of patents and uses it as a proxy for product innovation in various applications of the relation between products and growth. We extend their work through the introduction of a direct measure of the outcomes of innovation, product commercialization, through goods and services and its associated value. Our measures have broad coverage of firms, over 87% of the listed firms across all industries (manufacturing and service-based) and capture product outcomes of both observable (patent-based) and unobservable (trade secret) innovation.[7]

Our analysis sheds light on corporate investment and financing decisions, contributing to the industrial organization literature within finance. Specifically, our research addresses how multiproduct firms' product mix decisions influence corporate finance. We provide new insights on how firms' adjustments in their product portfolio can affect corporate investment and financing decisions. Our study also contributes to the literature on acquisition motives.[8] Historically, much of the industrial organization research has concentrated on specific industries—for instance, studies like those by Alexander (1997), Chevalier, Kashyap, and Rossi (2003), and Broda and Weinstein (2010) have examined product variety and market structure through detailed analyses of individual music records and scanner/supermarket data for consumer goods. Our study complement the work of Hoberg and Phillips (2023), who utilize information from annual reports of public firms spanning from 1996, to explore how increases in product scope are valued by financial markets and influence corporate strategies. Our work extends this narrative by incorporating an additional dimension of firm adjustment based on varying product types. Furthermore, we introduce alternative but more comprehensive measures of product dynamics over an extended time series that

[7]A granted patent must be renewed every four years. More than 17% of patents are never renewed, 20% are renewed only once in their fourth year, 19% are renewed in their eighth year and only 44% are renewed beyond 12 years. Furthermore, few patents are (easily) linked to specific products. More importantly, over 67% of US listed firms in our sample never file for a patent. Therefore assessing product variety and value via patents leaves out a significant proportion of firms, including almost all the service industry. Notable non-patented but trademarked products include the Coca-Cola recipe, McDonald's Big Mac special sauce, WD-40, Listerine, Kentucky Fried Chicken, etc.

[8]Some recent papers study the relation of trademarks to stock market outcomes. Hsu, Li, Li, Teoh, and Tseng (2021) and Faurel, Li, Shanthikumar, and Teoh (2020) find that the number of new trademarks predicts significantly higher stock returns and that CEO compensation structures are related to product trademark registrations, respectively. Heath and Mace (2020) study the effect of the 1996 Federal Trademark Dilution Act on firms' profits and strategy. Scanlon (2019) uses the US trademark stock to proxy for the number of goods in the economy and uses changes in trademark stock to proxy for product variety, in the calibration of a new consumption capital asset pricing model with product variety.

includes both private and public firms, thereby broadening the scope for research into the interplay between corporate policies and product market dynamics.

2 Construction of Product Measure

Our main objective in this section is to measure a firm's product space. In addition, we want to categorize a firm's products into three different economic types (quality improving, new product variety, business stealing) and to estimate the private economic value of these products. To do this, we use detailed information of goods and services from the USPTO trademark data and firm stock price movements. We proceed in the following three steps.

First, addressing the scarcity of detailed product-level information within firms poses a significant challenge. Previous studies often rely on scanner data, which is limited in cross-sectional and temporal scope and may not be representative of the product offerings in the economy. We capitalize on the prevalent practice of firms to trademark their key commercial products, which are applicable to virtually all goods and services. Trademarks offer several advantages for economic analysis: they are the most common form of intellectual property protection, they provide comprehensive coverage across firms and industries, and the data set contains rich information including product classes, detailed descriptions, application timelines, initial commercial usage, and renewal and expiry dates. Evidence of commerce must be provided and is authenticated by USPTO examiners. This wealth of information enables us to construct and monitor extensive product portfolios for a large number of all firms over extended time periods. In other words, our measure captures how and the timing of when firms add and drop products from their portfolios.

The ability to observe these dynamics facilitates the categorization of products into three economic types — quality improvements within a firm, business-stealing from competitors, and the introduction of entirely new product varieties, such as the emergence of personal computers in the late 1970s. By tracking a firm's and its competitors' time-varying product portfolio, we are able to identify the replacement of an existing product by a new product within a firm because of quality improvement. We can also identify the introduction of a new product in the firm's portfolio already produced by another firm, which we term as business stealing. A product is categorized as a new product variety when a new goods and services term ID is introduced in the USPTO ID Manual, such as smart phone.

Finally, we obtain an empirical estimate of the economic value of each product based on the present value of the rent associated with the trademarked product. To estimate the value, we combine information of the first-use-date in commerce (evidence provided by trademark applicant and verified by the USPTO examiner) and firm stock price movements in the spirit of Kogan et al. (2017). This allows us to study the contribution of each channel to firm and aggregate economic growth, which is critical for corporate and industrial policies. Below, we provide a detailed description of the three steps.

2.1 Financial and Accounting Data

For the empirical analyses that require asset prices, industry classification, and firm financial and accounting information, we merge trademark data to financial data using a name-matching algorithm described in Appendix A1 for US listed firms. For financial data, we use stock price information from the Center for Research in Security Prices (CRSP) and firm accounting information from Compustat.[9] Table B1 provides a list and definitions for all the variables used in the paper. The CRSP sample period covers 1926-2020, while the Compustat sample period starts in 1950, but it provides the most comprehensive coverage from 1976. Our firm level analysis focuses on listed firms from 1976-2020 with additional results for the 1950-2020 sample presented in Appendix C.

2.2 Trademark Data and Product Measure

We first describe how we use the USPTO trademark data to measure the dynamics of firms' key product (goods and services, G&S) portfolios. According to Chapter 22 of Title 15 of the US Code, a trademark is a word, phrase, symbol, design, color, smell, sound, or combination thereof that identifies and distinguishes the goods and services of one party from those of others (15 USC. § 1127). Trademark law serves to prevent consumer confusion and preserve consumer expectations regarding the source and quality of the goods and services they purchase. It also protects trademark owners' investments from misappropriation by preventing one seller from free-riding on another seller's reputation by offering similar goods

[9]We exclude firm-years with missing total assets and SIC classification codes. For firm and industry analysis we also exclude the heavily regulated industries of utilities (SIC code 4900-4949) and financial firms (SIC code 6000-6799). We winsorize all unbounded variables at the 1% level using annual breakpoints.

or services under a confusingly similar mark.[10]

The trademark data provides the following information, as exemplified in Figure 1 and detailed in Table B2: 1) the trademark's mark characters (the words/texts in the trademark), e.g., Office; 2) important dates in the trademark's life cycle, such as filing date, first use in commerce date, publication date, registration date, and cancellation date; 3) the trademark owner's name, address, and postcode; 4) Goods & Services (G&S) class(es); and 5) the owner's description of the use of the G&S provided by the trademarked good. Figure 1 provides an example from a Microsoft Office trademark registration.

A trademark right is established by use, therefore, for a trademark to be registered the applicant must submit evidence that, as of the filing date, the mark is being used in commerce or is intended to be used in commerce within six months. The trademark cannot be registered without showing evidence of the *actual* G&S use in commerce. The use in commerce information is encapsulated in the "first use in commerce" date, referred to hereafter as the commercialization date. Therefore, the commercialization date provides the most reliable and accurate timing of the introduction of a product to the market, validated by trademark attorneys and USPTO examiners.

The trademark application follows three stages to registration, presented as a flow chart in Figure 2. First, a randomly assigned examining attorney assesses the validity and uniqueness of the filing. The attorney also examines the goods and services classification of the mark. Specifically, registration may be refused if the descriptive terms of the goods and services used in commerce are not capable of identifying the source of the goods and services. This examination process forms the basis of our finer classification of trademarks into distinct product categories. Second, after the examination is completed and the trademark is deemed registrable, the USPTO publishes the trademark for opposition. Third, if there are no objections within 30 days after the USPTO publication, or the applicant successfully overcomes any opposition, the trademark proceeds to registration.[11]

The universe of USPTO trademark case files for the period 1870-2020 comprises 7.3

[10]For further information, see the USPTO documentation on "Basic facts about trademarks" (`https://www.uspto.gov/trademarks-getting-started/trademark-basics`) and the detailed data description in Graham, Hancock, Marco, and Myers (2013).

[11]A large proportion of trademarks reach the publication stage. In our sample, 79% of trademark applications are published for opposition, and 81% of trademarks published for opposition are registered. Graham et al. (2013) provides a detailed description of the trademark application and registration process.

million registered trademarks for US public and private firms, foreign firms, government entities, and individuals. We match the trademark owner names to US-listed firm names in CRSP using name matching algorithms. We match 0.97 million trademarks to 18,467 US publicly listed firms (US and non-US headquartered).[12] For other firms, we apply the same name matching algorithm to match firms in ORBIS, which includes US private firms and non-US publicly listed and private firms. We match 1.02 million trademarks to 141,610 firms with more than 10 million EUR operating revenue ("Large" and "Very Large") in ORBIS.[13]

In the publicly-listed firm sample, 85% of firm-years of firms listed on AMEX, NYSE and NASDAQ with common equity have at least one live trademark during for the 1976 to 2020 sample. Each firm holds on average 45 trademarks during the sample period. Table 1 presents the sample characteristics. Firms with registered trademarks during the sample period are larger, older, and less levered than the average firm in the sample.

2.3 Defining Product Categories for Trademarked Products

The USPTO organizes and classifies each trademark into broad categories of goods or services (Goods and Services Classification Category), as determined by: 1) the Nice International Agreement, and 2) the US classification system. The Nice Classification (NCL), established by the Nice Agreement in 1957, is an international classification of goods and services for the registration of trademarks. The USPTO adopted the Nice Classification in September 1973. Prior to that, it used a US-specific classification, which it still provides as supplementary information.[14] Currently, there are 45 classes, 34 goods and 11 services, in the NCL Goods and Services classification. The classes are broadly defined, for example, WIPO (World

[12]See Appendix A1 for a detailed description of the name matching algorithm and disambiguation approach. For robustness, we randomly select 500 firms to manually check the accuracy of the trademark-to-firm matching process.

[13]Companies in ORBIS are classified as "Large" when they meet at least one of the following criteria: operating revenue of >10 million EUR (approx. $13 million), total assets of >20 million EUR (approx. $26 million), a workforce of >150 employees, or they are publicly listed. It is important to note that the "large" firms threshold in ORBIS corresponds to fairly small firms in the Compustat sample. For instance, a firm with $13 million asset value is at the bottom 3% of the asset distribution of firms in Compustat in 2020. Excluding very small firms allows us to focus on more economically important products.

[14]Details are available from `https://www.wipo.int/classifications/nice/en/preface.html`. The USPTO has populated the NCL international classification equivalent for trademarks prior to 1973, see `https://www.uspto.gov/trademarks/trademark-updates-and-announcements/nice-agreement-current-edition-version-general-remarks` and Graham et al. (2013). The US classification remains important for referencing pre-1973 trademarks.

Intellectual Property Organization) comments on Class 9 as:

Class 9 is often described as the class of electronic equipment but, in fact, it includes a wide variety of goods. There are five general areas of goods: apparatus and instruments for scientific or research purposes, information technology and audiovisual equipment, apparatus and devices for controlling the distribution and use of electricity, optical apparatus and instruments, and safety equipment.

Obviously, the G&S classes recorded in the trademark are too broad to define each class as an individual product line.[15] To obtain a more granular classification of products, we utilize the text data from the trademark G&S statements written by the applicants. During the trademark registration process, the USPTO asks applicants to consult the Acceptable Identification of Goods and Services Manual (ID Manual) to describe the goods and services (G&S statement).[16] The ID Manual is based on the Nice Classification and is maintained by the USPTO's Office of the Administrator for Trademark Classification Policy and Practice. Since 2013, a new edition has been published annually, while updates occurred every five years prior to that.

Figure 3 provides an example page for a search in the manual related to mobile telephones. Although entries in the ID Manual are acceptable as written, the examining attorney may require amendments to an otherwise definite identification in certain circumstances. The listing is not exhaustive but is intended to serve as a guide for examining attorneys in acting on applications and for the public in preparing applications. An example of a recent product category is "Self-driving cars" from NCL version 11-1997 and USPTO Term ID (012-1525), introduced in the ID Manual in 2019.

The G&S statements provided by the applicant are validated by trademark examiners. The registration may be refused if the descriptive terms of the goods and services used in commerce cannot identify the source of the goods and services. This feature of the trademark examination process forms the basis for our finer classification of trademarks into distinct

[15]Consider an analogy with firm industry classifications. If one wants to define the industry at a granular level, one would choose four or more SIC digits. The broad G&S class recorded in each trademark is analogous to 2-digit SIC industries, of which there are 74.

[16]The ID Manual is available at `https://idm-tmng.uspto.gov/id-master-list-public.html`. It is derived from the Nice Classification (NCL) and contains over 52,000 unique product terms, representing granular product categories. This classification is more granular than the 8-digit SIC industry classification, which has about 18,000 categories. In electronic applications (TEAS Plus) for trademarks, applicants are asked to use the exact entry contained in the ID Manual.

product categories by mapping the product phrases from trademark G&S statements to the product terms in the ID Manual.

Each G&S statement listed in the trademark is matched with an entry in the product term from the Trademark ID Manual based on its semantic meaning. Here, we briefly discuss the mapping process, which is illustrated in Figure 4 and detailed in Appendix A2. First, since each G&S statement is written in paragraphs and typically includes multiple products, we convert each trademark G&S statement into individual product terms (e.g., shirts, shoes, mobile phones, tablets) using the large language model ChatGPT. Each product phrase is then converted into a vector representation using embeddings.[17] Similarly, we convert each product term in the ID Manual into a vector representation.

Next, we compute pairwise cosine similarities between each trademark G&S statement and each product term listed in the Trademark ID Manual. Direct comparisons are avoided due to variations in terminology (e.g., lager vs. beer). Instead, we match each product term in the G&S statement with the corresponding term in the ID Manual that has the highest cosine similarity.[18] This method effectively groups similar product terms into a fixed set of product categories defined by the NICE classification. Consequently, each product term in the G&S statement is assigned a Term ID indexing different product categories from the USPTO Trademark ID Manual.

For example, if the G&S statement is "Mobile Phone" as in Figure 3, then the procedure outlined in Figure 4 will match it with an entry in the ID Manual called "Mobile Telephone," because it is the closest item in the manual. It will not, for example, be matched to "Mobile telephone batteries" because it is an inferior match to "Mobile Telephone" in terms of semantic meaning.

2.4 Validation of Trademarked Products as Key Products

Trademarks generate value for a firm by indicating the origin of goods and services and by facilitating product differentiation. Thus, most firms will likely trademark their key goods and services, because trademarks have low granting requirements and are not expensive to

[17] We experimented with several embedding models and ultimately selected the MXBAI model (https://huggingface.co/mixedbread-ai/mxbai-embed-large-v1) due to its superior performance.

[18] Cosine similarity measures the distance between two vectors, ranging from -1 to 1. A value close to one indicates that the two vectors have very similar meanings according to the NLP model.

acquire. However, trademarked products are only a proxy for the product space of a firm. We assess the trademark coverage of individual firm key goods and services by comparing individual firm's trademarked products to: The major products a firm reports as a product line in the SEC filings in their annual reports through their 10-K filings and the Compustat Segment product database (Cohen, Diether, and Malloy, 2013), as well as award-winning products in the Consumer Electronics Show (CES). We find that a vast majority of listed-firms' products as captured by SEC filings and award-winning products are trademarked.

2.4.1 10-K Filings

When firms introduce important products or refer to trademarked products in their annual reports, they annotate these products with ® or ^{TM} for the readers. We expect a good coverage of these products as reported in the 10K filing. We search the 10-K filings of all CRSP firms in EDGAR for trademarks mentioned in the filings. Specifically, we download the 10-Ks for all sample firms for the available sample period 1996 to 2020. We search the 10-K reports of each firm for ® or ^{TM}, as trademark symbols, to identify firm's trademarked products. We find 53,996 unique trademark names from 7,766 firms' 10Ks. We find that 74% of the trademarks reported in 10-K filings are captured in the USPTO trademark data. A manual check of the remaining trademarks shows that they belong to firms with non-US operations or headquarters, and mostly refer to trademarks filed with intellectual property (IP) offices outside the US.

2.4.2 Compustat Product Segments

We use product information from firms' annual reports in 10K filings, as captured by the segment-level product data from the Compustat Segment data file (Segment product henceforth). The Segment product data provides business and geographic details for over 70% of the companies in the North American database. Our analysis focuses on the Business segment data, which presents industry segments or product lines reported by a company, to capture firm-level product information.

The Segment product database includes all the major products a firm discloses as a product line. The vast majority of firms report product lines rather than individual product names, for example, pharmaceutical companies often report "Drugs" as a category without

mentioning the name of the individual drug; car companies often report general model names such as "Hatchback," rather than detailed model descriptions. In this section we refer to product lines and individual products collectively as products. The Segment product data file records a unique product number for each new product and carries that product number through time (e.g., XBOX is product number 7 and Microsoft Office is product number 4 for Microsoft). The database provides the following information: GVKEY, product identifier number, NAICS code, product name, sales, and sales date. There are 9,705 firms in the sample period 1998-2018, and each firm reports nine products on average in this period.[19]

We use the Segment product data to carry out two types of analysis. First, we assess the relation between the number of new product lines and trademarks for each firm, through a regression analysis. Second, we extract all the individual products identified by name and examine their coverage in the USPTO trademark data.

Table 2 presents a firm-year panel regression of the log number of segment products reported by firm i in year t on the (log) number of trademarked products held by firm i at year t, where a product remains in a firm's portfolio as long as the trademark is in use, i.e., between the trademark's first-use-commerce date and the cancellation date.[20] The regression coefficients for the (log) number of products indicate a positive contemporaneous relation between trademarked products and products reported in firms' annual reports. This relation is statistically and economically significant after accounting for firm, time, and industry×time fixed effects. Economically a 1% increase in trademarked products is related to a 0.173% increase in 10-K reported products (column 1). Given that firms do not generally report all of their new products in 10-K filings and firms disclose most individual products only as part of product lines, we are likely to be underestimating the true magnitude of the relation. The positive and significant relation suggests that trademarked products provide a good indication of firms' new products.

To further gauge the coverage of trademarks for firms' product space, we take the list of individual products report in Segment data and match them to trademarks. We use a name matching algorithm to match product names to trademark characters. We match individual products for 250 firms in the period 1998 to 2018. The smaller number of firms

[19]The segments product data sample starts in 1996, however it is very sparsely populated prior to 1998.

[20]Where the cancellation date is incomplete, we use the last trademark renewal date plus five years. The results remain qualitatively similar to different imputations of cancellation dates and are available from the authors upon request.

in this trademark-CRSP-Compustat-Segment product sample is driven by the practice of a vast majority of firms of reporting just broad product categories or terms (such as cloth, glassware, product, etc.) rather than the exact product names in their 10K statements.[21] We find that 87.6% of the identified individual products reported in the Segment product data are trademarked.

2.4.3 Award Winning Products

A firm's significant and groundbreaking products are often acknowledged by both their peers and the marketplace. One such form of recognition comes from external awards, such as the CES (Consumer Electronics Show) Innovation Awards. CES, held annually in Las Vegas, Nevada, showcases the latest advancements in technology and consumer electronics across various industries, including automotive, health and wellness, home appliances, entertainment, and more. Regarded as one of the world's largest and most influential tech events, CES unites industry leaders, innovators, and consumers to explore new products, services, and ideas. In recent years, CES has highlighted groundbreaking technologies such as self-driving cars, smart home devices, 5G networks, virtual reality, and augmented reality systems.

The CES Innovation Awards is an annual contest that celebrates exceptional design and engineering in consumer technology products. Open to all CES exhibitors, entries are assessed by a panel of independent experts using criteria such as engineering, design, functionality, and innovation. Winning a CES Innovation Award is a prestigious accolade for companies. Award recipients are announced at CES and featured in a showcase during the event. Furthermore, winners receive recognition through various media channels, including online and print publications, social media, and industry events.

Considering the prominence of trademarked award-winning products, they are vital for firms to protect. We examined the extent of CES winning products' coverage by trademarked

[21] Specifically, we first create a list of all unique product names in the Segment product sample and strip off all generic terms, which do not represent firms' individual products and services. To remove generic terms, we recursively look up product names in the English dictionary and remove the product names that are in the dictionary. This allows us to capture unique product names such as XBOX in the 10-K filing of Microsoft Inc. We acknowledge that there may be shortcomings to this method. For example, Microsoft reports sale revenues for the Windows operating system as "Windows." Since "window" (after stemming) is in the English dictionary, the product is excluded from the list. However, these cases are relatively rare because firms often use unique names to distinguish their products. We also remove all chemical compound names as these compounds are generally sold under associated drug names.

products and find 86.5% (333 out of 385) of the CES award-winning products in 2021 are trademarked.[22] The products not covered by trademarks are primarily from non-US firms, or the products have not yet been commercialized. Overall, this section's evidence demonstrates that trademarked products offer extensive coverage of firms' key products, including those reported in 10-K filings, and recognized by awards.

2.5 Classifying Product Types

Using the product list created from the procedure above, we classify products into three categories: new variety (NV), quality improving (QI), and business stealing (BS) products, in the spirit of Garcia-Macia, Hsieh, and Klenow (2019). We define a new variety product as one that is new to both the firm and its industry competitors as a product line. These products are first time appearances in the trademarked product database. An example of a new variety in our data is mobile telephones. The first mobile telephone is trademarked under PULSAR from Motorola Inc.[23] The product category "Mobile Telephone" (Trademark ID Manual Term ID 009-1000) starts when Motorola filed the trademark for "PULSAR," which we categorize as a new variety in our data.

A quality improving product of a firm is a new product that falls into an existing product category. It improves the current products within the firm's product space, but it does not create a new product category for either the firm or the industry. For example, iPhone 6 is a quality improving extension of the earlier iPhone product line of Apple Inc. iPhone 6 belongs to the Smart Phone category and is not a new product line to Apple Inc. and its competitors (i.e., Samsung).

Finally, we define products that add a new product to a firm but this product already exists in the market as business stealing or creative destruction. An example of a business stealing product is the Samsung GT-I7500 Galaxy, a first generation Samsung smart phone, released on 29 June 2009, which adds a new product (smart phone) for Samsung to compete with Apple's iPhone second and third generation iPhone 3G/3GS. All subsequent Samsung smart phones are categorized as quality improving for Samsung.

We classify each trademarked product as follows: for each trademarked product of firm i at time t, we track firm i's product vector up to $t-1$ to find whether firm i has already

[22] At the time of writing the paper, the trademark data provided by the USPTO was available until 2021.
[23] `https://en.wikipedia.org/wiki/History_of_mobile_phones\#Handheld_mobile_phone`.

trademarked products in its product portfolio. If the product already exists in the firm's product portfolio, the trademark is categorized as quality improvement over the existing products. If not, we search whether other firms (firm $j \in J$) have already introduced the same product. If other firms have the same product line as the trademarked product, we assign it to business stealing, as this trademarked product is an improvement over other firms' product and potentially steals business from competitors. If no other firm has such product lines, we assign it to a new variety/product line.

2.6 Economic Value of Products

In this section, our primary goal is to obtain an empirical estimate of the economic value of a firm's products upon their initial commercialization. This measure corresponds to the present value of the anticipated future cash flows derived from products at the time of their first commercialization. Our approach to this estimation aligns closely with the methodology outlined in Kogan et al. (2017).

The central challenge in gauging market reaction—or the economic value of a product—lies in identifying the precise date when new product information is released to the market and maintaining consistency across different firms and periods. If one could determine the initial date of commerce, they could use the market reaction to calculate the private economic value of a new product. However, pinpointing a product's first usage in commerce often proves difficult. In recent work, Mukherjee et al. (2021) use Factiva searches to identify a firm's new product launches. However, not all firms publicly announce their new products, especially if they are services or intermediary goods.

To address this issue, we leverage the USPTO's stipulation that a product must have been used in commerce for a trademark application to be valid. To register a trademark, applicants must provide a legal basis for the trademark's use, either "used in commerce" or "intended to use in commerce," and the date of this usage commencement, hereinafter referred to as the commercialization date. This date offers the most accurate and reliable timing for a product or service's introduction to the market, corroborated by trademark attorneys and the USPTO examiner. The commercialization date typically marks the product's first release to the market when market participants can assess the product's features and potential.

This date may fall either before or after the publication for opposition date, as products

can be trademarked based on evidence of future commercialization, up to 6 months in advance. Only 16% of our sample has a publication date preceding the first use in commerce date. For any given trademark, we use the earlier of the commercialization and publication for opposition dates, but results are remarkably similar when using only the first-use-in-commerce date.

For the construction of the private economic value of each trademarked product, we use the "first use in commerce" date discussed previously. When the "first use in commerce" date is poorly populated, we use the publication date to estimate the economic value for these trademarked products because the publication date is the date when the product first becomes publicly known.[24] The trademark and CRSP merge leaves us with 0.97 million trademarks (Trademark-CRSP sample).

Finally, to obtain the economic value of trademarked products, we also exclude trademarks issued to CRSP-firms before their exchange listing, because we cannot calculate their market value without stock price information. Our product-value sample consists of 469,169 trademarks with associated CRSP information, which is used to calculate the product economic value. Consequently, all value-weighted product value aggregated at the firm, industry, and economy level is based on this product-value sample. The trademarked product sample is smaller than the sample of patents in Kogan et al. (2017), probably due to two main factors. First, some patents never result in a product (only 63% of patents are renewed beyond their eight years). Second, some products use technologies and components covered by multiple patents (e.g, reportedly the iPhone involves more than 200 patents, `https://www.patentlyapple.com/patently-apple/tech-iphone/`). However, the firm coverage with trademarks includes almost the full universe of listed firms.

Moreover, it is necessary to separate the product-related market reaction from other factors. The net firm return around this date, after accounting for aggregate market fluctua-

[24]Over 75% of the trademarks have well-populated dates with information on the day, month, and year of the first use in commerce date. Poorly populated dates often contain information on only the month and the year. This does not affect the equally-weighted product measure, because it does not require a "use in commerce" date. However, the absence of the exact day of first use in commerce affects the calculation of the stock returns and economic value-weighted product measures. For robustness, we conduct analysis only using trademarks with complete "use in commerce" dates. We also check the differences with the economic value-weighted measures, where we replace the poorly populated date with the publication date, which often falls in the same month. We find our product measures across these procedures do not have significant differences and the results of our empirical applications remain qualitatively similar.

tions, should represent the market's perceived value of the product. We conduct the analysis within a three-day window starting on the event day, namely, the commercialization date.

On the product commercialization date, the market learns that the firm has a new product and reacts conditioning on the product information. The market reaction ΔV on the commercialization day of product p is given by:

$$\Delta V_p = (1 - \pi_p)\xi_p, \tag{1}$$

where π is the probability the market attaches to the firm successfully commercializing product p, and ξ_p is the dollar value of the commercialized product p. For example, if the market anticipates a firm will produce and successfully commercialize a product with probability π, this should be accounted for in the estimation of ξ_p. However, the unconditional probability of the commercialization of a new key product for a particular firm on a particular day is likely to be very small. Therefore, we approximate equation (1) by $\Delta V_p \approx \xi_p$.

To isolate returns related to only product commercialization, we first remove observations with commercialization dates that coincide with scheduled announcements like earning announcements. Second, we focus on stock's abnormal return R_i, firm return minus market return (proxied by CRSP valued-weighted returns). We further decompose the abnormal return R_i into return related to the product value v_p plus unrelated factors ϵ_i:

$$R_i = v_p + \epsilon_i.$$

From equation (1), the estimate economic value $\hat{\xi}_p$ of the product p for $\pi = 0$ is,

$$\hat{\xi}_p \approx \frac{1}{N_i} E[v_p|R_i]M_i, \tag{2}$$

where N_i is the total number of products commercialized by firm i on the event date. $\hat{\xi}_p$ is the expected product value v_p times the firm market capitalization M_i. If multiple products are commercialized on the event date, these products are assigned equal value, which is a fraction $\frac{1}{N_i}$ of the total value. In our sample almost all firms commercialize only one product on event days.

To implement equation (2), we make the following assumptions about the distribution of v and ϵ. First, we allow v and t to vary across firm i and across time t. Second, we assume

a strictly positive value of v, which is normally distributed truncated at 0, $v_p \sim N^+(0, \sigma^2_{v,it})$; we also assume a normal distribution for $\epsilon_i \sim N(0, \epsilon^2_{\epsilon,it})$. Under these assumptions, the expected value of product p produced by firm i for abnormal return R (denoted as $v_p|R_i$) is

$$E[v_p|R_i] = \delta_{it}R_i + \sqrt{\delta_{it}}\sigma_{\epsilon,it}\frac{\phi\Big(-\sqrt{\delta_{it}}\frac{R_i}{\sigma_{\epsilon,it}}\Big)}{1-\Phi\Big(-\sqrt{\delta_{it}}\frac{R_i}{\sigma_{\epsilon,it}}\Big)}, \tag{3}$$

where ϕ and Φ are the standard normal pdf and cdf functions, respectively; and δ is the signal-to-noise ratio

$$\delta_{it} = \frac{\sigma^2_{v,it}}{\sigma^2_{v,it} + \sigma^2_{\epsilon,it}}. \tag{4}$$

Next, we estimate parameters $\sigma_{v,it}$ and $\sigma_{\epsilon,it}$. To make the estimation feasible, the signal-to-noise ratio is constant across firms and time. This assumption implies that $\sigma^2_{\epsilon,it}$ and $\sigma^2_{v,it}$ vary across firms and time, but in constant proportions to each other. To estimate δ, we regress the log squared returns on the product commercialization date dummy C:

$$log(R_{id})^2 = \gamma C_{id} + cZ_{id} + u_{id}, \tag{5}$$

where R_{id} refers to the 3-day abnormal return of firm i, starting on commercialization day d. Z is a vector of control variables indicating the day of week and firm×year fixed effects to control for seasonal return fluctuations and volatility. Then, the signal-to-noise estimate can be recovered from the estimated value of γ using $\hat{\delta} = 1 - e^{-\hat{\gamma}}$. We estimate $\hat{\gamma} = 0.0162$ and $\hat{\delta} = 0.016$. We use these values to compute the benchmark results.

Finally, we estimate $\sigma^2_{\epsilon,it}$ using the sum of squared abnormal returns, which allows the estimate to vary at an annual frequency. Specifically, we first estimate the realized mean abnormal squared returns, $\hat{\sigma}^2_{it}$ and the fraction of trading days that are commercialization dates, d_{it}. The variance of the measurement error can be recovered through

$$\sigma^2_{\epsilon,it} = 3\hat{\sigma}^2_{it}(1 + 3d_{it}(e^{\hat{\gamma}} - 1))^{-1}. \tag{6}$$

To construct the firm-year level measure of product value, we sum up the commercial-

ization value of all trademarks of a particular type commercialized by firm i in year T:

$$\Theta^{P}_{i,T} = \sum_{p=1}^{P_{i,T}} \hat{\xi}_p, \tag{7}$$

where $\hat{\xi}_p$ denotes the market value in 1982 \$ calculated for each trademarked product type (BS, QI, NV), and $P_{i,T}$ denotes total trademarked product type commercialized by firm i in year T.

2.6.1 Product Values and Future Sales

We validate the estimated product private economic value against future product sales. We use the Compustat Segments products dataset that compiles sales data at the product line or individual product level when firms report them in their 10-K fillings. We anticipate that the value of trademarked products will mirror the cash flows generated by each product.

We first analyze the trademarked products which are also covered in Segments Compustat as described in Section 2.4.2 and footnote 21. Figure 5 presents the relationship between product value and product sales as reported in the Segments product data. We aggregate sales at the product level for the reporting period of the sample, as product value represents the net present value of all future cash flows from the product. Therefore, a product that records sales for the period from 1998-2017 is represented by a single data point that sums all the sales in 1982 dollars. The Y-axis displays the log of total product sales in 1982 dollars, and the X-axis shows the log of product value in 1982 dollars. The Figure illustrates a strong positive correlation between product value and product sales. The relation is approximately log-linear, and a 1% increase in market expected product value is associated with a 0.82% increase in realized total sales.

2.6.2 Descriptive Statistics

Figure 6 shows the time-series of the evolution of different product types in firm portfolios across industries from 1976 to 2020. A product enters a firm's portfolio in the year of the use in commerce date and it is removed from the portfolio when it is canceled or not renewed. Firms have increased business stealing and quality improving products over the sample period, while decreasing new varieties. This is not surprising, because as the baseline

of available products increases, it becomes harder for firms to generate entirely new products, much like ideas, e.g. Bloom, Jones, Van Reenen, and Webb (2020). Furthermore, firms cancel trademarks related to new products as they replace them with higher quality new versions. At the end of the 2010s, on average firms have more quality improving products than any other product types in their portfolios. Firms in the Chemical and Allied Products industry hold most products: business stealing, quality improvement and new variety in their portfolios, until the turn of the century. Over the last 20 years, they have been overtaken by firms in the Consumer Nondurables and Telephone and TV Transmission industries for BS and Consumer Nondurable for QI products. Firms in the Healthcare, Medical Equipment and Drugs, and Oil, Gas, and Coal Extraction industries have the least number of different type of products in their portfolios, which is in line with the low product differentiation in the oil and gas industry, and the very high cost of developing new or improved products in the healthcare and pharmaceutical industry.

Table 3 presents univariate characteristics of firms that commercialize different product types, in the year of their commercialization. The same firm can appear in all three categories in a given year, if it introduces three different product types in that year. New varieties are introduced by the largest (total assets and PPE), oldest, and most profitable (ROA) firms operating in more concentrated markets (HHI). These firms also hold the most patents. Younger, less profitable, growth firms (Tobin's Q) operating in less concentrated markets introduce the most business stealing products. Firms with lower sales and asset growth tend to introduce more quality improving products.

Economic theory provides insights on the type of products firms introduce at different times in their life span. The expectation is that new entrant firms introduce new or better products to displace incumbents. Figure 7 provides some empirical evidence based on publicly listed firms, through the breakdown of product types by firm age group, where age is determined from Jay Ritter's dataset of firm IPOs (Loughran and Ritter, 2004; Field and Karpoff, 2002).[25] Panel A presents the average number of new products introduced by firms in different age groups each year, and Panel B presents the average number of product types in a firm's portfolio, by age group. Younger and older firms introduce new varieties at almost equal rates. Surprisingly, on average, young firms (0-5 years) commercialize more

[25]The data is available for free download from `https://site.warrington.ufl.edu/ritter/files/founding-dates.pdf`.

business stealing products, at double the rate of quality improving ones. Firms continue to commercialize more business stealing products, and their product portfolios are dominated by business stealing products until reaching 11-20 years. After that, firms do more quality improvements, and they retain more quality improving products in their portfolios. Older firms (above 21 years) introduce the most quality improvement products.

Finally, we turn to the private market value of the different product types. Table 4 presents the summary statistics and the distribution of product value $\hat{\xi}_p$ at the product type level. QI products have the highest average value of $29.31 million, followed by BS products at $23.69 million and NV products with the lowest value of $13.55 million. There is a large variation in market value across products, with a standard deviation of $61 million for BS products. The distribution of $\hat{\xi}$ is heavily right-skewed, where some products generate extremely high value. Notable examples include a version of MS Office 365 in 2011, with a market value of $155 million in 1982 dollars.

3 Product Types, Firm Growth and Corporate Decisions

In this section, we examine the relation among the composition of firms' product portfolios, firm growth, and corporate financing strategies. This inquiry intersects with several key issues for understanding corporate finance, as well as matters pertinent to industrial policy. Specifically, we investigate how different types of products are related to future firm growth and whether a rise in products that capture market share from competitors (business stealing) correlates with increased merger and acquisition activities. Additionally, we explore the dynamics of research and development expenditures in relation to enhancements in firms' product quality and the introduction of new product varieties. We consider whether firms opt for acquisitions and capital expenditures as strategies for entering new product categories, traditionally dominated by business-stealing tactics. Finally, this section studies the financing mechanisms for these investments to understand how firms align their financial strategies with their product development objectives.

3.1 Product Types Commercialization and Firm Growth

Several economic theories (see, e.g., Schumpeter, 1939; Klette and Kortum, 2004; Aghion, Akcigit, and Howitt, 2015) emphasize the crucial role of product innovation in overtaking competitors and reshaping market dynamics. These theories highlight how new or improved products allow firms to capture and maintain market share, underscoring the competitive advantages in dynamic markets. The threat of competition compels incumbents to either introduce new products or improve existing ones to maintain or enhance their market positions, see Lucas Jr and Moll (2014). Models like Akcigit and Kerr (2018) integrate creative destruction and quality improvement. Other theories, such as Romer (1990) and Jones (2016), stress the impact of new product varieties.

However, the link between product types and firm growth remains unexplored. We use our direct measure of product types to test their relation to firm growth. To our knowledge, this is the first empirical evidence that connects a direct product type measure with firm growth. For each firm-year, we aggregate the market value of commercialized products within types and standardize them by firm total assets to account for firm size:

$$\theta_{i,t}^{P} = \frac{\hat{\Theta}_{i,t}^{P}}{B_{i,t}}, \quad for\ P\ \in (BS,\ QI,\ NV). \tag{8}$$

The product market value serves as a weight for the importance of a product in a firm's growth. More valuable products will carry higher weight when using this measure. However, results presented below are not limited to market value weighted products and remain qualitatively similar when using the number of different product types instead of product market value.

We use Local Projection Regression (LPR) or Jordà regressions (Jordà, 2005) to estimate the effect of different product types on firm growth. LPR is a simple, flexible and robust (to model misspecification and structural breaks) method with excellent statistical properties to estimate the exact same impulse responses in population as VAR, see Jordà (2005), Plagborg-Møller and Wolf (2021), and Stock and Watson (2018) for other advantages of using LPR. We estimate the impulse responses across one to ten years horizon following Jordà (2005) and Kogan et al. (2017). We test the relation between products types and firm growth with

regressions of the following form:

$$ln(\frac{Y_{i,t+\tau}}{Y_{i,t}}) = \beta_0 + \beta_1 \theta_{i,t}^{BS} + \beta_2 \theta_{i,t}^{QI} + \beta_3 \theta_{i,t}^{NV} + \gamma_1 ln(Y_{i,t-1}) + \gamma_2 ln(Y_{i,t-2}) + \Gamma' X_{i,t} + \alpha_{sic3} + \delta_t + \epsilon_{i,t}, \quad (9)$$

where the dependent variables Y_i are firm level outcome variables: profit (sales minus cost of goods sold), sales, capital (property, plant and equipment) and employment. Equation (8) describes $\theta_{i,t}^{BS/QI/NV}$ scaled by total assets, $X_{i,t}$ is a vector of control variables including log capital stock, log of employment, log total asset, and idiosyncratic volatility, and $\tau \in 1-10$. α_{sic3} and δ_t are industry and year fixed effects. All right-hand-side variables are scaled to unit standard deviation. The estimated coefficients β across τ represent the impulse responses at each horizon.

Table 5 presents the results of the relation between product commercialization at time t and firm growth up to ten years in the future. Figure 8 shows the time series graph of the three coefficients for profits and sales. We find that new varieties, quality improving, and business stealing products are all positively correlated with firm growth, albeit with different magnitudes and timing. Quality improving products have the largest contribution within the first five years. A one standard deviation increase in a firm's quality improving products corresponds to a 4.9% profit, 4.3% sales, and 4.6% capital investment increase and 5.2% employment growth. The short-term effect is in line with quality improving products self-cannibalizing the firm's product space (Dhingra, 2013).

The introduction of business-stealing products tends to have a more sustained effect on firm growth. A one standard deviation in business stealing is correlated with approximately 9.2% (12%) increase in profits, sales, and capital and 7.6% (11%) growth in employment after five (ten) years. Finally, new varieties, which are introduced by both new and incumbent firms (see Figure 7), are important for longer term growth. A standard deviation increase in new varieties has a sustained and high with future profit, sales, capital and employment growth which is higher than quality improvement but smaller than business stealing.

Our analysis contributes significant new insights into the mechanisms through which product innovation drives growth in public companies. Products that enhance quality and those that capture market share ("business-stealing" products) not only fuel firm growth differently but also have differentiated impact over short and long-term horizons on growth, employment, and capital expenditures. Contrary to prevailing academic consensus, which

often overlooks the role of new product varieties, our new measures and findings underscore their significant, contrasting effects on growth - a vital element for formulating both industrial and firm-level innovation policies. These results have important implications for firms' investment in product innovation and provide support to all three classes of growth theories related to creative destruction, quality improvement, and new varieties.

These product measures offer novel insights into the outcomes of prior inventions' commercialization. By incorporating them into our analysis, we aim to gauge the economic effects of successful commercialization of various product types on growth and resource allocation. Moreover, the value of patents in the stock market, serving as an indicator of new inventions, shows a positive correlation with growth, aligning with the Kogan et al. (2017) results, see table B5. Thus, our measures extend beyond merely tracking a firm's inventive capacity to include their ability to commercialize these inventions within the innovation process. When included in the same empirical model, these product measures remain both economically substantial and statistically significant, underscoring the critical role of R&D—both invention and commercialization—in promoting firm growth. This provides a novel adjustment margin within the innovation process that can be integrated into economic models.

3.2 Corporate Finance Policies and Product Expansion

An additional important dimension of firm behavior related to the product space are corporate finance policies, which we turn to next. We focus on the relation between product composition on ex-post investments, including acquisitions, divestitures (target of acquisition), R&D and CAPX. We follow the relevant literature in introducing controls that are pertinent to the different outcome variables (Bena and Li, 2014; Frésard, Hoberg, and Phillips, 2020). In all regressions, we control for size, firm age, Tobin's Q, HHI, cash flow, and industry $\times$ year fixed effects.

In Table 6, we find that firms which launch more business stealing products tend to expand through acquisitions and not own investment. These firms acquire more, are significantly less likely to be targets, and invest less in R&D and capital expenditure, holding everything else constant. They acquire more than other types of firms (t-statistic of 9.74) and the economic magnitudes are quite large. If we vary predicted BS by one standard deviation, the probability of acquiring another firm increases by 8%, which is 44% of the

mean and 21% of the standard deviation of the probability of making an acquisition. The decreases in becoming a target is also significant with a t-statistic of -2.47. Business stealing firms seem to focus on growth through acquisitions, as they invest less in R&D and in CAPX. Results are robust to controlling for product scope (Hoberg and Phillips, 2024) in Table C2.

These findings underscore the strategic role of acquisitions in penetrating previously unexplored product segments. Our results are consistent with firms acquiring entities to swiftly enter and potentially dominate new markets, bypassing the slower processes of organic growth and internal development. This approach not only accelerates a firm's product space diversification but also enhances its market power, by absorbing competitors or complementary assets that could have posed threats or offered synergies. The observation that acquisitions have been pivotal in shaping firms' product expansion strategies provides a deeper understanding of the dynamics driving M&A activity over the past three decades. Economically, this suggests that M&As have been used not merely for consolidation, but also as strategic tools for competitive positioning and market penetration. This finding has implications for regulatory discussions on the consequences of excessive acquisitions and the balance of innovation and competition against the risk of market monopolization.

We find that firms with high incentives to develop new product varieties increase R&D and capital expenditures. The increase in R&D is significant with a t-statistic of 2.31. Economically, if we vary new varieties by one standard deviation, firms are predicted to increase their R&D/Assets ratio by 2.6 percentage points, which is 52% of the mean and 26% of the standard deviation of the R&D/Assets ratio. The commercialization of entirely new product categories typically stems from in-house R&D, mainly because acquisition targets that conduct frontier research and innovation are scarce. Furthermore, conducting R&D internally offers firms complete control over their innovation processes, aligning developments with strategic objectives and existing product lines. This approach circumvents the integration issues and cultural clashes often associated with acquisitions.

Quality improvements can arise due to improvements of new varieties the firm has introduced in the market, or improvements of products the firm took from competitors. If a firm is at the product frontier through new varieties, it continues to improve on these varieties using its own R&D, as it is alone in the frontier. However, if it is improving products in 'stole' from competitors, then it either needs to invest in more internal R&D or it can continue to acquire other firms. Consistent with these strategies, we find that firms expand their

quality improvement range through a mix of acquisitions and R&D investments, depending on which product range is being improved on.

Finally, as firms are investing in the different types of products, we expect these products to be reflected in future growth opportunities, as captured by Tobin's Q. Results in column (5) show that business stealing and quality improvement products are reflected in higher expected growth opportunities and firm value. These results are in line the period one findings in Table **??**, where quality improvement have the highest correlation and new varieties have the lowest. A one standard deviation increase in business stealing products is reflected in 0.1 higher Tobin's Q, which is 4.8% of its mean and 5.7% of its standard deviation. Quality improvement products have a higher impact on future growth opportunities and firm value. A one standard deviation increase in quality improvement products is reflected a 0.45 higher Tobin's Q, which is 26% of its standard deviation. New varieties also have a positive albeit insignificant impact on Tobin's Q. Their importance might not be immediately reflected in firm value due to the time lag in market perceptions and the uncertainty surrounding market acceptance, production costs, and profitability.

4 Conclusion

In economic theory, the array of products firms create—categorized into business-stealing, quality-improving, and new varieties—plays a crucial role in influencing a firm's strategic decisions around production, resource allocation, and growth. These product types significantly impact investment, R&D, and financing strategies. Empirical work in this area is often hindered by the difficulty in obtaining measurable data on firm products and their economic value, posing challenges for validating theoretical economic models.

In this paper, we introduce new metrics and a comprehensive product dataset that addresses limitations of traditional approaches by providing the most extensive coverage of products from both public and private firms. Our dataset includes economic values of products for public firms and classifies products into categories such as quality improvement, new variety, and business stealing across services and goods. We also tracks product entry and exit and offer a long-term perspective on firm product dynamics. We use our data set to explore: 1) the effect of different product types on firm growth, employment, productivity, and capital investment; 2) how firms adjust their investment and financing strategies in

response to changes in their product portfolio composition among these product categories.

We provide three new insights from our analysis based on our new product measures. First, we find that business stealing products contribute most significantly to firm growth over a longer-run from year 5 to 10, while quality improvement is more critical for firm growth over a shorter horizon (year 1 to 5). Second, new variety products are introduced by large and mature firms (incumbents) at the same frequency as by new entrants. These products have a large and prolonged impact on firm growth. Third, business stealing firms are more likely to expand through acquisition, while other firms invest in more R&D to stimulate radical product innovation. These measures of economic types provide the opportunity to expand our knowledge of the interplay between economic outcomes and different types of innovation at the firm level. Further analysis can be conducted at the economy level in testing extant economic theory.

Tables and Figures

Figure 1
Example trademark page

The figure provides an example page (screenshot) of a trademark publication in the Trademark Electronic Search System (TESS, `https://tmsearch.uspto.gov/`). We underline in red the most important information on the trademark publication, including the date of first use and the first use in commerce.



Word Mark	OFFICE
Goods and Services	IC 009. US 021 023 026 036 038. G & S: Computer software and computer programs for use in database management, electronic spreadsheets, designing, creating, editing and publishing documents, note taking, email, calendaring, scheduling, contact information storage and management, accounting, customer relationship management, project management, presentation graphics, desktop publishing, web publishing, document management, database management, data collection, data analysis, graphics, presentation graphics, word processing, instant messaging, Voice over Internet Protocol (VoIP), video conferencing, audio conferencing, web conferencing and document sharing, application sharing, computer desktop sharing, file transfer, sensing and providing user presence information, telephony, real-time communications featuring instant messaging, audio, video and data exchange, mapping geographical locations, manipulating video, audio and graphics; communications software for managing the exchange of communications and data and information over computer networks, wireless networks, and global communications networks; computer communications software for use in accessing web sites and exchanging information and data and obtaining services from web sites; computer software for use in developing web sites, operating electronic mail and providing computer security; computer software for recording, organizing, editing and transmitting information in typed written and voice format; computer communications software for accessing web messaging and shared documents from web sites on the Internet; computer utility programs; computer software development programs; computer security and authentication software for controlling access to and communications with computers; computer software for computer network security, anti-virus protection, and intrusion detection and prevention; computer game software; house mark for a full line of business software applications. FIRST USE: 20121212. FIRST USE IN COMMERCE: 20121212
Mark Drawing Code	(3) DESIGN PLUS WORDS, LETTERS, AND/OR NUMBERS
Design Search Code	26.05.21 - Triangles that are completely or partially shaded 26.11.02 - Plain single line rectangles ; Rectangles (single line)
Serial Number	85799081
Filing Date	December 10, 2012
Current Basis	1A
Original Filing Basis	1B
Published for Opposition	May 21, 2013
Registration Number	4456462
Registration Date	December 24, 2013
Owner	(REGISTRANT) **Microsoft** Corporation CORPORATION WASHINGTON One **Microsoft** Way Redmond WASHINGTON 980526399
Attorney of Record	Matthew E. Moersfelder
Disclaimer	NO CLAIM IS MADE TO THE EXCLUSIVE RIGHT TO USE "OFFICE" APART FROM THE MARK AS SHOWN
Description of Mark	Color is not claimed as a feature of the mark. The mark consists of a three-dimensional geometric design with two vertical sides and two horizontal sides. The vertical sides and top side are three-dimensional images of rectangular solids and the bottom side is a three-dimensional image of a triangle. To the right of the geometrical design is the word "Office".
Type of Mark	TRADEMARK
Register	PRINCIPAL
Affidavit Text	SECT 15. SECT 8 (6-YR).
Live/Dead Indicator	LIVE

Figure 2
Process of trademark registration

The figure shows the typical timeline in the trademark application process. The process entails three steps: (1) An application is filed on the filing date to be examined by a randomly assigned attorney. (2) If it passes the examination, it is approved for publication and published for opposition in the Official Gazette on the publication date. (3) If there is no opposition the trademark is registered on the registration date.

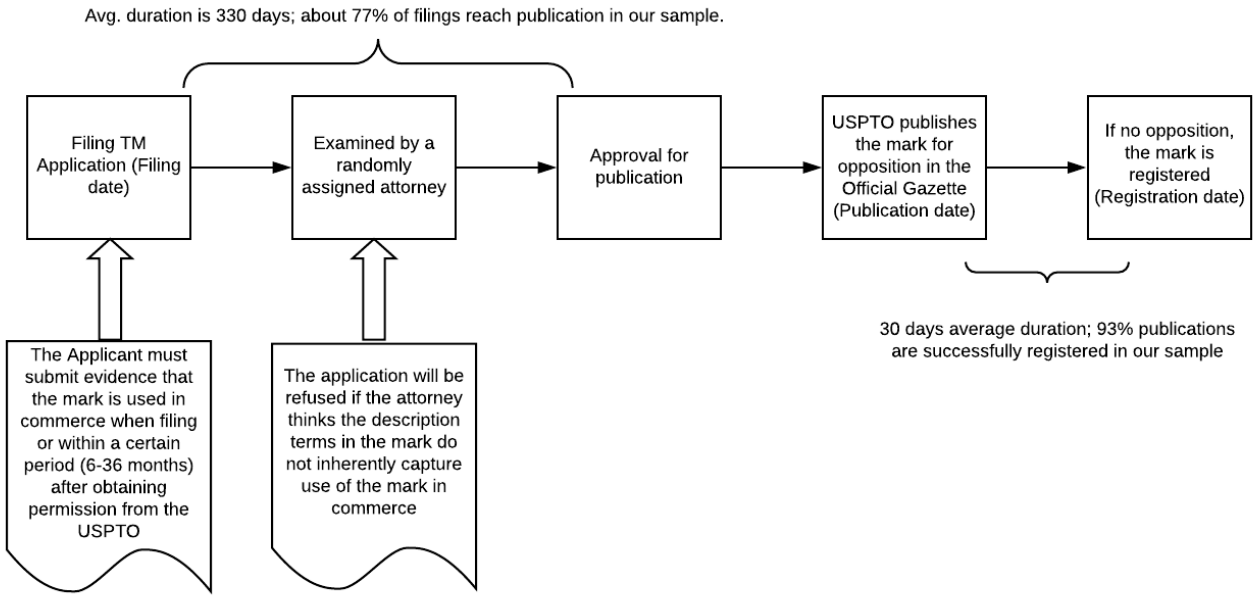


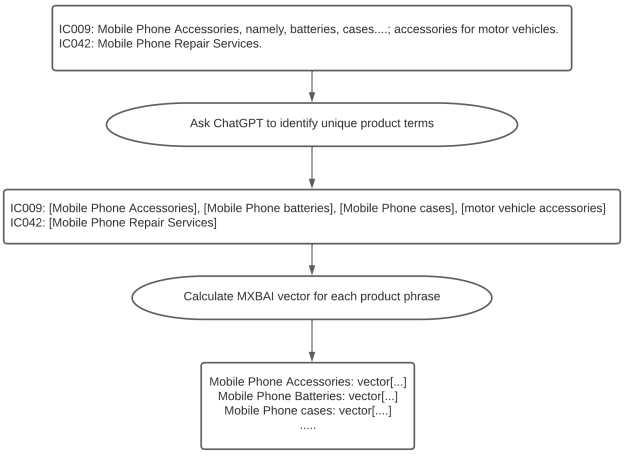
Figure 3
Example page for Trademark ID manual

The figure presents the output page for the search of "mobile telephone" in the Trademark ID Manual. The manual contains over 52,848 unique product description.

Term ID	Class	Description	Type	NCL Version
009-1000	009	Mobile telephones	GOODS	"07-1997"
009-1925	009	Mobile telephone batteries	GOODS	"08-2002"
009-1924	009	Mobile telephone apparatus with built-in facsimile systems	GOODS	"08-2002"
009-4023	009	Fitted plastic films known as skins for covering and providing a scratch proof barrier or protection for electronic devices, namely, MP3 players, mobile telephones, smart telephones, digital cameras, global positioning systems and personal digital assistants [excluding game apparatus]	GOODS	"10-2013"
009-3907	009	Mobile telephone accessories, namely, belt clips	GOODS	"09-2007"
009-617	009	Encoded electronic chip cards for {indicate use, e.g., identifying particular users of computers and mobile telephones}	GOODS	"07-1997"
009-3435	009	Headsets for mobile telephones	GOODS	"09-2007"

Figure 4
Process of matching trademark statement text to ID Manual

(1) This figure shows the procedure of cleaning trademark text and calculate word/phrase vector



(2) This figure shows the procedure of matching each trademark statement to the product description in the ID Manual via cosine similarity

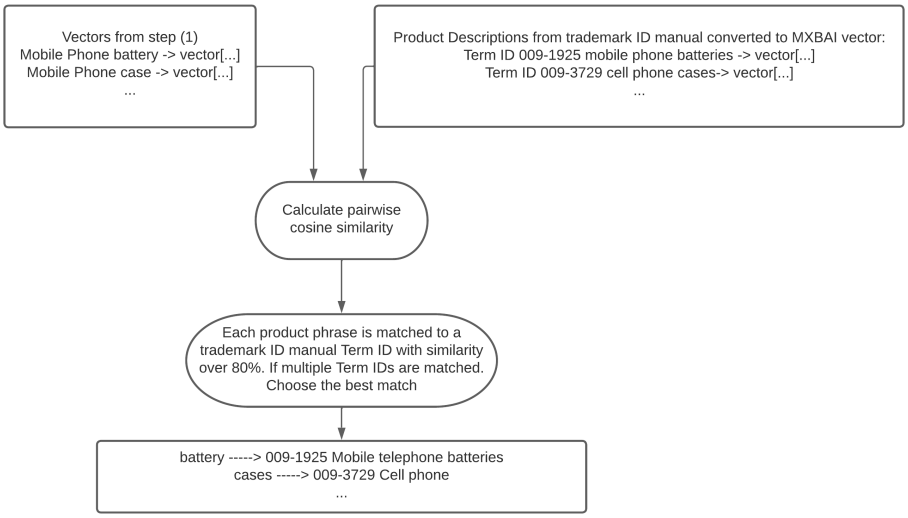


Figure 5
Trademark product value and product sales

The figure presents a scatter plot of the relation between product sales and trademark product value. The aggregate value of the product sales (1982 dollars in millions) is from the Compustat Segments product database. The trademark value is the market value (1982 dollars in millions) of individual trademarks as per the trademark commercialization date. The y-axis presents the log of total product sales from 1998 to 2018, the x-axis presents the log trademarked product value ξ.

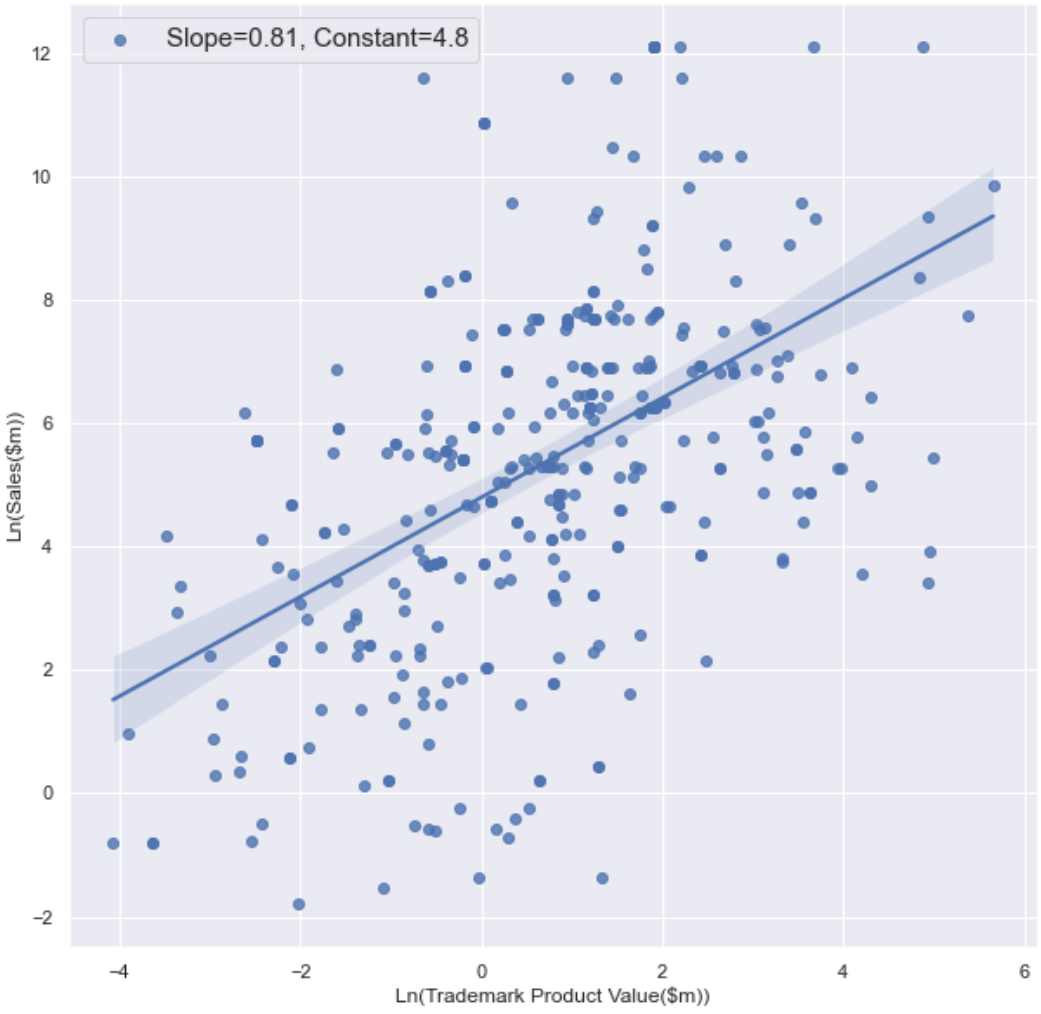
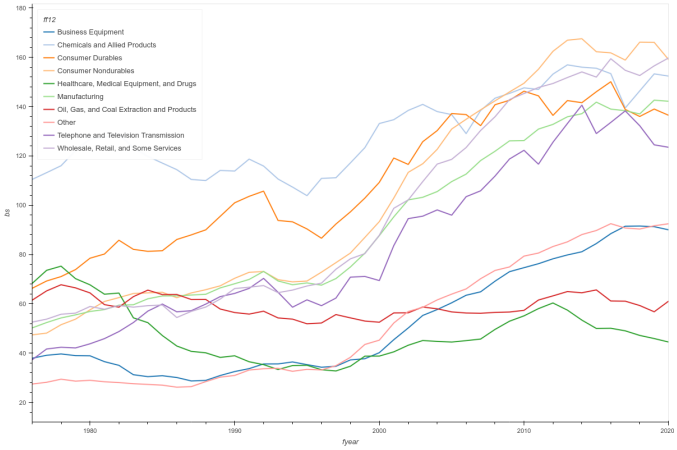
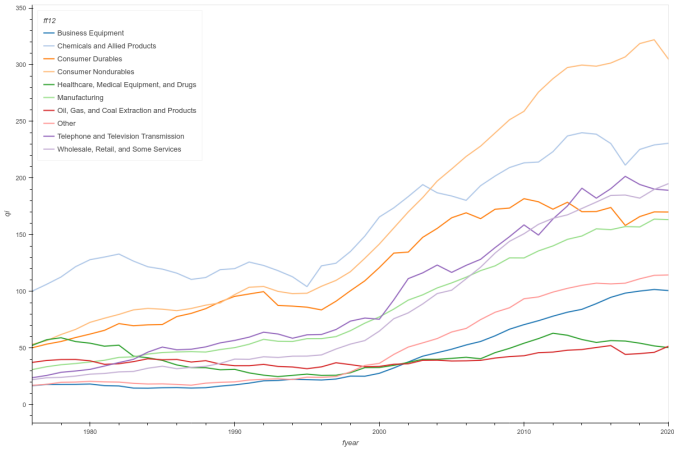


Figure 6
Time series variation of product types

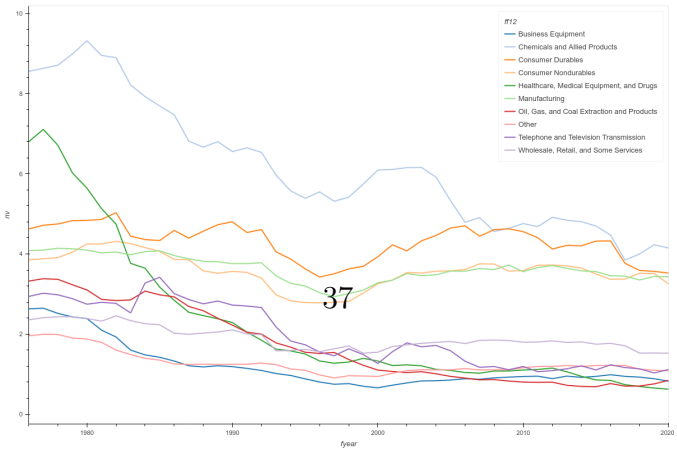
This figure presents the time series variation in product type stock by industry for publicly listed firms. The figure presents the average number of products by type in a firm's portfolio for each Fama-French industry and year. Products remain in a firm's portfolio from the use in commerce date until they are canceled or not renewed. Panel A presents the time series for business stealing (BS), Panel B for quality improvement (QI), and Panel C for product variety (NV). Product types are defined in Section 2.3.



Panel A. Business Stealing



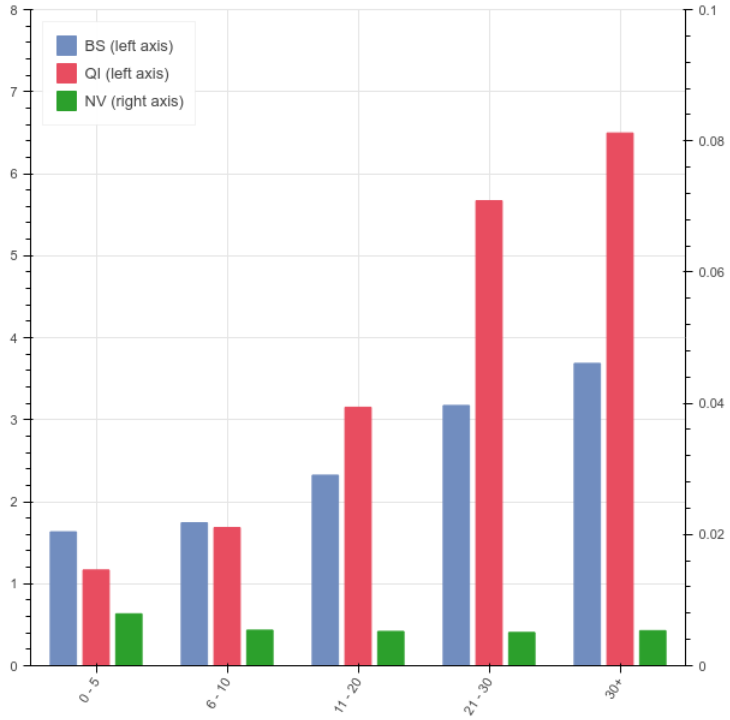
Panel B. Quality Improving



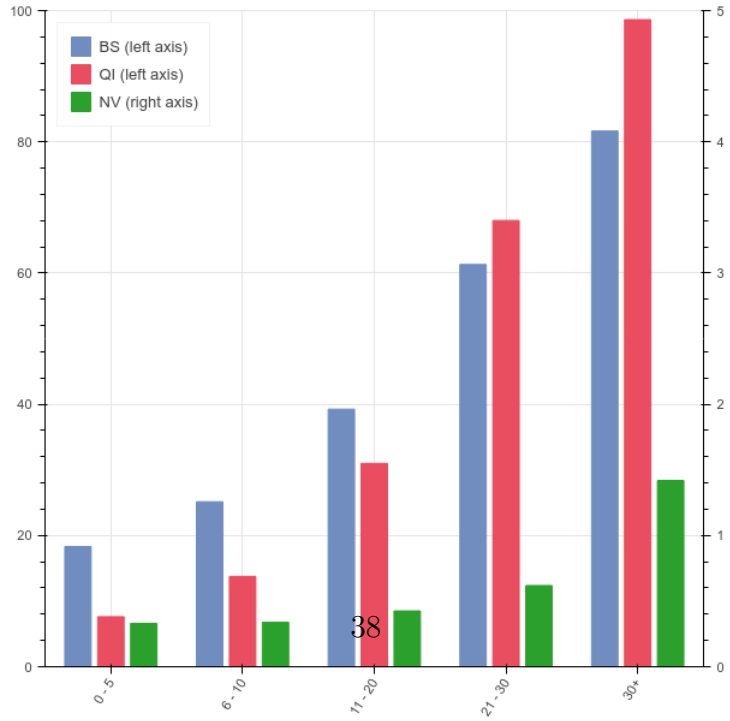
Panel C. New Variety

Figure 7
Product types by firm age

This figure presents the product types held by firms of different age. We present the average number of products held by each firm in each age buckets: business stealing (BS), quality improvement (QI), and product variety (NV). Product types are defined in Section 2.3. Firm age is calculated from the firm founding year from Jay R. Ritter's website. The sample period is 1976 to 2020.



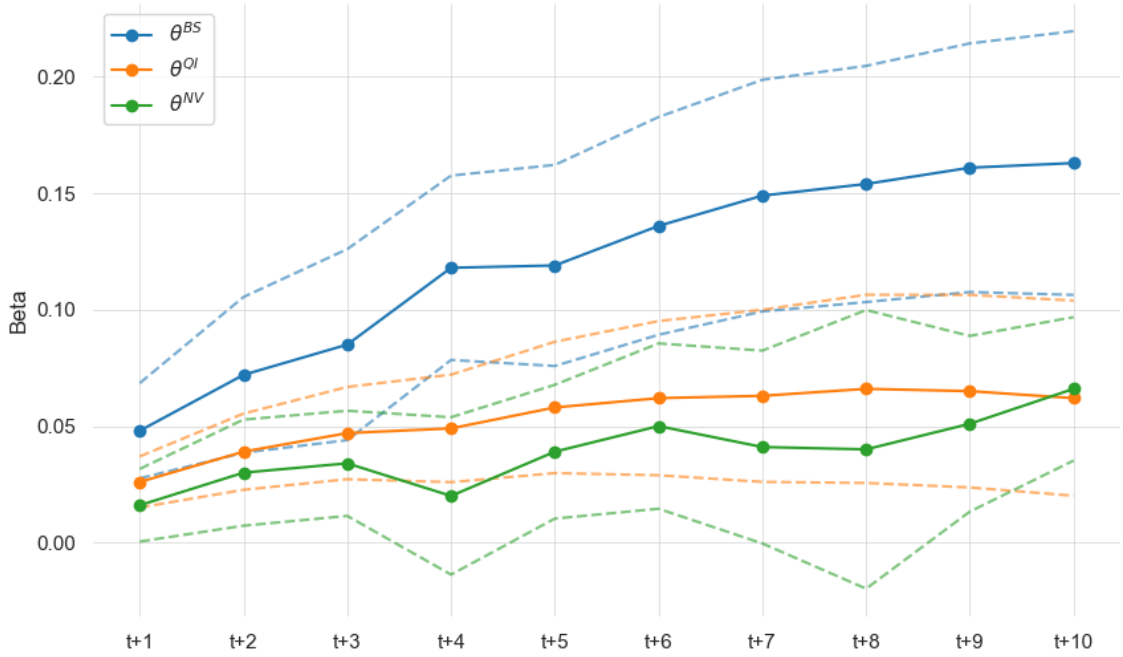
Panel A. New Products Commercialized



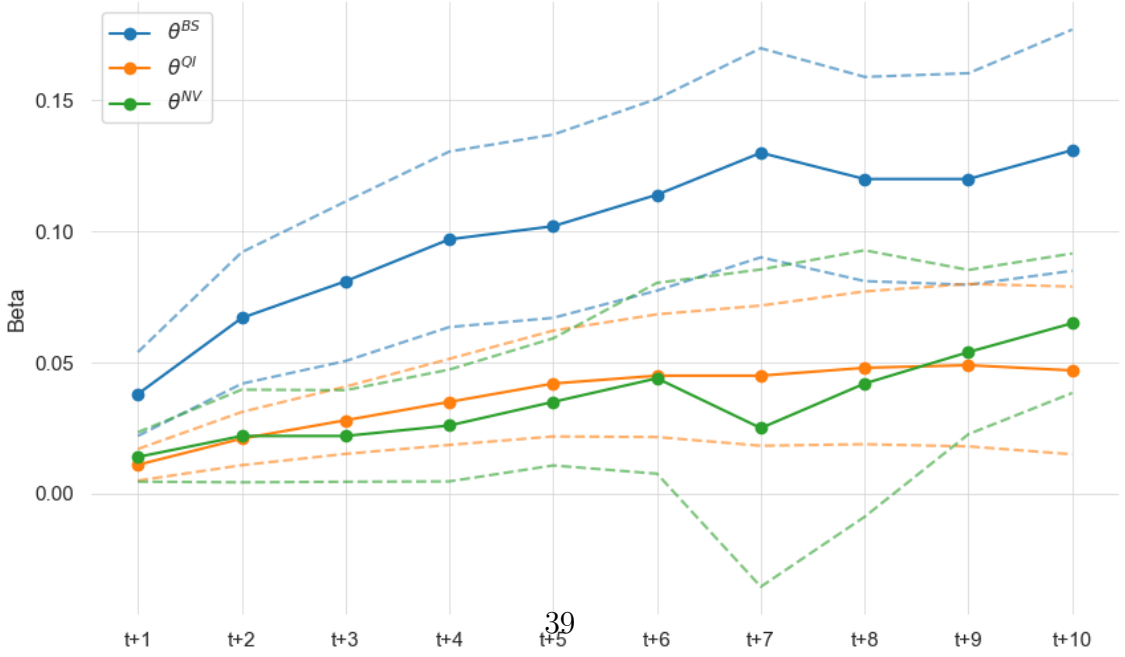
Panel B. Product Portfolio

Figure 8
Firm growth and product types estimated coefficients

This figure present the coefficients from the regression of firm growth outcomes on product types in Table 5. The figure shows the coefficient estimates and 95% confidence intervals for a 10-year window around the product commercialization date. The sample period is 1976 to 2020.



Panel A. Profits



Panel B. Sales

Table 1
Trademark firm characteristics

This table reports univariate characteristics of firms with and without trademarks, and their difference. TM presents the characteristics of firm-years, where each firm has at least one live trademark. Non-TM presents the characteristics of firm-years where the firm does not have any live trademarks. Diff. is the difference between TM and non-TM, and t-stat is the associated t-statistic. All variable definitions are presented in Table B2. ***, **, * indicates significance level at 1%, 5% and 10%. All potentially unbounded variables are winsorized at 1% level. The sample period is 1976 to 2020.

	All				**TM**				**All − TM**	
	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.	Diff.	t-stat
ln(Assets)	168,597	4.91	4.77	2.21	143,972	5.18	5.04	2.15	-0.27***	-34.39
ROA	168,594	-0.05	0.03	0.27	143,969	-0.04	0.04	0.26	-0.01***	-10.00
CAPX	168,597	0.07	0.04	0.07	143,972	0.06	0.04	0.06	0.00***	20.49
PPE/AT	168,397	0.28	0.22	0.22	143,817	0.26	0.21	0.21	0.02***	20.32
R&D	168,597	0.05	0.00	0.10	143,972	0.05	0.00	0.10	-0.00***	-7.09
Leverage	168,588	0.71	0.34	2.15	143,965	0.68	0.33	2.08	0.03***	3.99
Tobin's Q	166,654	2.02	1.42	1.75	143,027	2.02	1.44	1.71	-0.00	-0.16
Acquirer	168,597	0.21	0.00	0.40	143,972	0.22	0.00	0.42	-0.02***	-11.55
Target	168,597	0.01	0.00	0.09	143,972	0.01	0.00	0.09	0.00	0.68
Divest.	168,597	0.06	0.00	0.24	143,972	0.07	0.00	0.25	-0.00**	-2.34
Sales Growth	154,387	0.19	0.09	0.60	133,361	0.18	0.09	0.55	0.01***	6.23
Asset Growth	154,387	0.14	0.06	0.42	133,361	0.14	0.06	0.41	0.00	1.29
Equity Issue	168,597	0.11	0.00	0.32	143,972	0.11	0.00	0.31	0.00	0.89
Debt Issue	168,597	0.09	0.00	0.28	143,972	0.10	0.00	0.29	-0.01***	-8.13
Dividends	168,341	0.01	0.00	0.02	143,758	0.01	0.00	0.02	-0.00***	-7.13
Equ. Repur.	168,597	0.39	0.00	0.49	143,972	0.41	0.00	0.49	-0.02***	-10.32
Firm age	168,597	14.17	9.00	15.37	143,972	15.31	10.00	15.96	-1.13***	-20.14
HHI	168,597	0.27	0.22	0.20	143,972	0.28	0.22	0.20	-0.00***	-3.46
No. Patents	168,597	5.69	0.00	25.01	143,972	6.62	0.00	26.91	-0.92***	-9.86

Table 2
Product segments and product lines

This table present the results regressing number of product segments on the number of product lines measured by trademarks. T-statistics adjusted for clustered standard errors at firm level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All potentially unbounded variables are winsorised at 1% level. The sample period is 1999 to 2018.

	OLS			Poisson		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Product Lines (NICE))	0.159***	0.049***	0.041***	0.224***	0.042***	0.042***
	(16.76)	(4.41)	(3.40)	(10.97)	(2.80)	(3.30)
SIC4 FE	✓			✓		
Year FE	✓	✓		✓	✓	
Firm FE		✓	✓		✓	✓
SIC4 by Year FE			✓			✓
Obs.	37,238	36,248	34,377	37,238	36,248	34,377
Adj. R^2	0.32	0.87	0.88			

Table 3
Univariate Firm Characteristics by Product Type

This table reports univariate characteristics by product type: business stealing (BS), quality improvement (QI), and product variety (NV). Product types are defined in Section 2.3. Each firm is categorized under a product type, each year it introduces that product type. All variable definitions are presented in Table B2. The sample period is 1976 to 2020.

	BS			**QI**			**NV**		
	Mean	Median	Std. Dev	Mean	Median	Std. Dev	Mean	Median	Std. Dev
ln(Assets)	6.11	6.14	2.18	6.28	6.30	2.06	6.38	6.56	2.20
ROA	-0.00	0.05	0.22	0.01	0.05	0.19	0.03	0.06	0.16
CAPX	0.06	0.04	0.06	0.06	0.04	0.05	0.07	0.06	0.05
PPE	0.25	0.21	0.19	0.26	0.22	0.19	0.28	0.25	0.17
R&D	0.04	0.01	0.09	0.04	0.01	0.08	0.03	0.01	0.06
Leverage	0.69	0.38	2.04	0.70	0.39	2.06	0.68	0.45	1.71
Tobin's Q	2.07	1.52	1.66	2.02	1.51	1.57	1.93	1.45	1.44
Sales Growth	0.18	0.09	0.53	0.15	0.08	0.44	0.18	0.11	0.45
Asset Growth	0.15	0.07	0.39	0.14	0.07	0.37	0.18	0.10	0.38
Firm age	18.99	13.00	18.83	19.96	14.00	19.03	22.99	17.00	20.42
HHI	0.28	0.22	0.20	0.28	0.23	0.21	0.29	0.24	0.20
No. Patents	24.74	0.00	169.94	24.87	0.00	168.53	37.07	1.00	114.33

Table 4
Commercialization Value of Product Types

The table presents the distribution of commercialization value of the trademarked products, measured as described in Section 2.6. The sample period is 1976 to 2020.

	BS	QI	NV
Mean	23.69	29.31	13.55
Std	90.98	101.33	46.66
10%	0.24	0.33	0.17
25%	0.82	1.20	0.59
40%	2.01	2.89	1.40
50%	3.42	4.82	2.35
60%	5.73	8.03	4.07
75%	13.54	18.53	8.67
90%	47.90	63.69	25.51
Obs.	140,144	250,918	1,020

Table 5
Product Types and Firm Growth

The table presents the regression results for the relation between product value and firm-level outcomes. We estimate regressions of the form:

$$ln(\frac{Y_{i,t+\tau}}{Y_{i,t}}) = \beta_0 + \beta_1\theta^{BS}_{i,t} + \beta_2\theta^{QI}_{i,t} + \beta_2\theta^{NV}_{i,t} + \gamma_1 ln(Y_{i,t-1}) + \gamma_2 ln(Y_{i,t-2}) + \Gamma' X_{i,t} + \alpha_{sic3} + \delta_t + \epsilon_{i,t},$$

where Y are the firm outcomes, $\theta^c_{i,t}$ are the aggregated market value (ξ) of each product type introduced (use in commerce year) by firm i in year t scaled by total assets, $X_{i,t}$ is a vector of control variables including log capital stock, log of employment, log total asset, and idiosyncratic volatility, and $\tau \in 1-10$. α_{sic3} and δ_t are industry and year fixed effects. All right-hand-side variables are scaled to unit standard deviation. Panel A presents results for firm gross profits (sales minus cost of good sold), Panel B for firm sales, Panel C for firm capital stock (property, plant, and equipment), and Panel D for firm employment. Standard errors are clustered at the firm level. ***, **, * indicates significance level at 1%, 5% and 10%, respectively. All unbounded variables are winsorized at 1% level using annual breakpoints. The sample period is 1976 to 2020.

	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
					Panel A. Profits					
θ^{BS}	0.031***	0.049***	0.061***	0.090***	0.093***	0.107***	0.118***	0.122***	0.127***	0.129***
	(3.85)	(3.58)	(3.69)	(5.50)	(5.20)	(5.51)	(5.65)	(5.77)	(5.80)	(5.45)
θ^{QI}	0.020***	0.030***	0.039***	0.042***	0.049***	0.053***	0.056***	0.058***	0.056***	0.054***
	(4.29)	(4.12)	(4.22)	(3.84)	(3.72)	(3.38)	(3.08)	(2.94)	(2.77)	(2.58)
θ^{NV}	0.015**	0.027***	0.032***	0.019	0.036**	0.046**	0.040*	0.038	0.049***	0.066***
	(2.34)	(3.11)	(3.33)	(1.11)	(2.49)	(2.57)	(1.94)	(1.30)	(2.67)	(4.62)
					Panel B. Sales					
θ^{BS}	0.033***	0.059***	0.072***	0.087***	0.091***	0.101***	0.115***	0.106***	0.106***	0.115***
	(4.59)	(5.18)	(5.21)	(5.71)	(5.83)	(6.21)	(6.47)	(6.06)	(5.83)	(5.53)
θ^{QI}	0.010***	0.021***	0.029***	0.036***	0.043***	0.046***	0.045***	0.048***	0.049***	0.048***
	(3.41)	(3.88)	(4.15)	(4.13)	(4.01)	(3.67)	(3.18)	(3.08)	(2.94)	(2.74)
θ^{NV}	0.013***	0.020***	0.020**	0.024**	0.032**	0.040**	0.024	0.040	0.051***	0.063***
	(3.09)	(2.59)	(2.43)	(2.17)	(2.47)	(2.09)	(0.81)	(1.61)	(3.25)	(4.91)
					Panel C. Capital					
θ^{BS}	0.041***	0.060***	0.078***	0.091***	0.096***	0.105***	0.120***	0.116***	0.116***	0.121***
	(6.29)	(5.89)	(5.94)	(6.04)	(5.32)	(5.67)	(5.71)	(5.34)	(5.14)	(5.75)
θ^{QI}	0.011***	0.024***	0.030***	0.036***	0.046***	0.053***	0.055***	0.059***	0.066***	0.068***
	(3.31)	(3.21)	(2.97)	(3.19)	(3.70)	(3.74)	(3.48)	(3.72)	(3.88)	(3.40)
θ^{NV}	0.012**	0.012	0.011	0.024*	0.027	0.037	0.037	0.042*	0.047**	0.054***
	(2.13)	(1.42)	(1.04)	(1.74)	(1.44)	(1.51)	(1.45)	(1.69)	(2.28)	(3.13)
					Panel D. Employment					
θ^{BS}	0.031***	0.050***	0.062***	0.071***	0.076***	0.090***	0.106***	0.104***	0.103***	0.110***
	(4.80)	(4.67)	(4.72)	(5.07)	(4.90)	(5.17)	(5.31)	(5.10)	(4.95)	(4.99)
θ^{QI}	0.013***	0.026***	0.036***	0.044***	0.052***	0.057***	0.058***	0.061***	0.063***	0.061***
	(4.54)	(5.68)	(5.81)	(5.39)	(5.39)	(5.08)	(4.52)	(4.37)	(4.08)	(3.39)
θ^{NV}	0.012	0.012	0.013	0.021**	0.026***	0.029	0.026	0.031	0.044***	0.044***
	(1.43)	(1.41)	(1.42)	(2.16)	(2.65)	(1.30)	(1.15)	(1.36)	(3.10)	(3.17)
Obs.	126,228	113,907	103,041	93,644	85,315	77,859	71,154	65,110	59,630	54,664

Table 6
Investment

This table presents the relationship between product types and investment choices. We run regressions of the following form:

$$Y_{i,t} = \beta_0 + \beta_1 BS_{i,t-1} + \beta_2 QI_{i,t-1} + \beta_3 NV_{i,t-1} + \beta_c' X_{i,t-1} + \alpha_{sic3} \times \delta_t + \epsilon_{i,t},$$

where Y is a different investment outcome variable, BS is the number of business stealing products, QI is the number of quality improvement products, and NV is the number of new variety products introduced by firm i in year t. The outcomes variables are: M&A outcomes (Acquirer or Target indicator variable) in columns (1) and (2), R&D and capital expenditures (CAPX) in columns (3) and (4). Acquirer and target are indicator variables equal to one when the firm is an acquirer or a target in an MA deals, otherwise zero. R&D and CAPX are R&D are scaled by total assets. T-statistics adjusted for clustered standard errors at firm level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All variable definitions are presented in Table B2. All potentially unbounded variables are winsorized at 1% level. The sample period is 1976 to 2020.

	Acquirer	Target	R&D	CAPX
	(1)	(2)	(3)	(4)
No. BS	0.247***	-0.009**	-0.007	-0.001
	(9.74)	(-2.47)	(-1.35)	(-0.36)
No. QI	0.145***	-0.003	0.017***	-0.002
	(7.45)	(-1.21)	(3.86)	(-0.81)
No. NV	-0.854	0.137	0.256**	0.129
	(-1.06)	(1.26)	(2.31)	(1.36)
ln(Assets)	4.776***	0.039**	-0.705***	0.140***
	(44.45)	(2.36)	(-16.72)	(5.80)
ln(Age)	-1.566***	-0.065***	-0.532***	-0.572***
	(-10.43)	(-2.62)	(-10.05)	(-18.53)
HHI	0.375	-0.080	-2.455***	0.292
	(0.29)	(-0.42)	(-4.92)	(1.07)
Cash Flow	3.269***	-0.009	-2.598***	1.908***
	(4.13)	(-0.05)	(-9.70)	(14.66)
Tobin's Q	0.879***	-0.054***		
	(9.58)	(-3.56)		
Sale growth	7.205***	-0.029		
	(33.72)	(-0.62)		
ROA	8.582***	-1.066***		
	(14.21)	(-6.72)		
Leverage	-0.124**	0.001		
	(-2.18)	(0.08)		
R&D	-0.206***	-0.014***		
	(-11.64)	(-4.08)		
Industry × Year FE	✓	✓	✓	✓
Obs.	148,442	148,442	150,339	150,339
Adj. R^2	0.19	-0.00	0.39	0.32

REFERENCES

Acemoglu, D. 2003. Labor-and capital-augmenting technical change. *Journal of the European Economic Association* 1:1–37.

Aghion, P., U. Akcigit, and P. Howitt. 2015. Lessons from Schumpeterian growth theory. *American Economic Review* 105:94–9.

Aghion, P., and P. Howitt. 1992. A model of growth through creative destruction. *Econometrica* 60:323–51.

Akcigit, U., and W. R. Kerr. 2018. Growth through heterogeneous innovations. *Journal of Political Economy* 126:1374–443.

Alexander, P. J. 1997. Product variety and market structure: A new measure and a simple test. *Journal of Economic Behavior & Organization* 32:207–14.

Axarloglou, K., and V. Theoharakis. 2003. Diversity in economics: An analysis of journal quality perceptions. *Journal of the European Economic Association* 1:1402–23.

Bena, J., and K. Li. 2014. Corporate innovations and mergers and acquisitions. *The Journal of Finance* 69:1923–60.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb. 2020. Are ideas getting harder to find? *American Economic Review* 110:1104–44.

Broda, C., and D. E. Weinstein. 2006. Globalization and the gains from variety. *Quarterly Journal of Economics* 121:541–85.

———. 2010. Product creation and destruction: Evidence and price implications. *American Economic Review* 100:691–723.

Chevalier, J. A., A. K. Kashyap, and P. E. Rossi. 2003. Why don't prices rise during periods of peak demand? Evidence from scanner data. *American Economic Review* 93:15–37.

Cohen, L., K. Diether, and C. Malloy. 2013. Misvaluing innovation. *The Review of Financial Studies* 26:635–66.

Dhingra, S. 2013. Trading away wide brands for cheap brands. *American Economic Review* 103:2554–84.

Faurel, L., Q. Li, D. Shanthikumar, and S. H. Teoh. 2020. Bringing innovation to fruition: Insights from new trademarks. Working Paper.

Field, L. C., and J. M. Karpoff. 2002. Takeover defenses of IPO firms. *The Journal of Finance* 57:1857–89.

Frésard, L., G. Hoberg, and G. M. Phillips. 2020. Innovation activities and integration through vertical acquisitions. *The Review of Financial Studies* 33:2937–76.

Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow. 2019. How destructive is innovation? *Econometrica* 87:1507–41.

Graham, S. J., G. Hancock, A. C. Marco, and A. F. Myers. 2013. The USPTO trademark case files dataset: Descriptions, lessons, and insights. *Journal of Economics & Management Strategy* 22:669–705.

Hall, B., C. Helmers, M. Rogers, and V. Sena. 2014. The choice between formal and informal intellectual property: A review. *Journal of Economic Literature* 52:375–423.

Heath, D., and C. Mace. 2020. The strategic effects of trademark protection. *Review of Financial Studies* 33:1848–77.

Hoberg, G., and G. Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124:1423–65.

———. 2024. Scope, scale and concentration: The 21st century firm. *The Journal of Finance* Forthcoming.

Hsu, P.-H., D. Li, Q. Li, S. H. Teoh, and K. Tseng. 2021. Valuation of new trademarks. *Management Science* .

İmrohoroğlu, A., and Ş. Tüzel. 2014. Firm-level productivity, risk, and return. *Management Science* 60:2073–90.

Jones, C. I. 2016. Life and growth. *Journal of Political Economy* 124:539–78.

Jordà, O. 2005. Estimation and inference of impulse responses by local projections. *American Economic Review* 95:161–182–.

Klette, T. J., and S. Kortum. 2004. Innovating firms and aggregate innovation. *Journal of Political Economy* 112:986–1018.

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

Krusell, P. 1998. How is R&D distributed across firms? a theoretical analysis. *Working Paper* .

Lentz, R., and D. T. Mortensen. 2008. An empirical model of growth through product innovation. *Econometrica* 76:1317–73.

Loughran, T., and J. Ritter. 2004. Why has IPO underpricing changed over time? *Financial Management* 33:5–37.

Lucas Jr, R. E., and B. Moll. 2014. Knowledge growth and the allocation of time. *Journal of Political Economy* 122:1–51.

Olley, G. S., and A. Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64:1263–97.

Plagborg-Møller, M., and C. K. Wolf. 2021. Local projections and VARS estimate the same impulse responses. *Econometrica* 89:955–80.

Romer, P. M. 1990. Endogenous technological change. *Journal of Political Economy* 98:S71–S102.

Scanlon, P. 2019. New goods and asset prices. *Journal of Financial Economics* 132:140–57.

Schumpeter, J. A. 1939. *Business cycles*, vol. 1. McGraw-Hill New York.

Stock, J. H., and M. W. Watson. 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal* 128:917–48.

Appendix A. Additional Information

A1 Name Matching Algorithm

To match trademark owners to CRSP firms, we perform a name matching exercise using a name-matching algorithm. It has two steps, name standardization and calculate text similarity. In name standardization, we uses extensive list of name variations to standardize individual words in names, e.g., changing "Corp" to "Corporation," "Ltd" to "Limited," etc. The list comes from several sources. The first is from CRSP supplied firm name abbreviations used in their database; the second is from recursive searches in Wikidata to find out common firm abbreviations (e.g., IBM for International Business Machines); and the third is manual collections of name abbreviations from several databases including CRSP, Compustat, Google Patents, and USPTO Trademark case file. Then, we compute the pairwise text similarity among trademark owner's name and the firm name using a Levenshtein-based string distance. To account for possible spelling errors, the text similarity is performed both at the phrase and word level and similarity threshold are set to allow for spelling errors. In addition, based on the uniqueness of the words/phrases presented in the name, we decide on keeping/dropping the match. Non-unique words/phrase that commonly observed in the English dictionary are matched when overall similarity is high and unique words will attract a low similarity score to be a match (similar to TF-IDF approach used in the other matching algorithm). We benchmark existing name matching databases conducted by other researchers like KPSS. The authors also use a name-matching algorithm to match patent assignees to CRSP firms and matched 1.8 million patents for the period 1926 to 2010 out of total 7.8 million patents. We use the same dataset from Google Patents and perform name-matching on this dataset. We matched over two million patents. We manually inspect those additional matches and confirm they are mapped due to our extensive name standardization efforts. There is a non-matched portion of firms matched using our algorithm compared to KPSS's matched sample (less than 20%). We also manually inspect that sample and find that the unmatched patents often belong to subsidiaries with significantly different names from parent companies, and they are outside the scope of a text matching algorithm. People often conduct hand-match of those observations. We conclude that our matching algorithm is of good quality to conduct fuzzy name matching in firm names.

A2 Mapping Product Phrase in Trademark to ID manuals

We sketch the process of cleaning the statement text and mapping the statement to the Trademark ID Manual in Figure 4. Here, we provide more technical details.

We start with the trademark G&S statement text recorded for each trademark. Each trademark may include one or multiple broader G&S classes. For example, in Figure 4, we take the mobile phone trademark, which has broader G&S classifications such as IC009 and IC042. We split the paragraph by semi-colons.[26] In this example, we will be left with "Mobile Phone Accessories, namely, batteries, cases ..." and "Accessories for motor vehicles" and "Mobile Phone Repair Services". Each part is then fed into ChatGPT 3.5 Turbo through the OpenAI API using the following prompt:

```
For the below product(s) or service(s),
clearly identify the specific "kind" of product(s) or service(s) being provided.
Clearly identify any descriptive information. Do not add your own information.
For each of them, return them in the format "kind (description)" nothing else, on separate lines.
--- {part of the statement text here}
```

The model will return the product phrases. In our example, the output is:

- batteries (Mobile Phone accessories)
- cases (Mobile Phone accessories)
- accessories (for motor vehicles)
- repair services (mobile phone)

The prompt is designed to extract unique product terms from the statement text, as even within each part split by the semi-colon, there could be multiple products. For each extracted term, we embed them into vectors using MXBAI, specifically employing the sentence transformer (SBERT) to achieve this. We then embed each product term indexed by Term ID in the USPTO Trademark ID Manual using SBERT and calculate the pairwise cosine similarity between each product term extracted from the statement text and the ID Manual.

[26] It is officially required to use semi-colons to separate distinct goods and services, see `https://www.bitlaw.com/source/tmep/1902_02_f.html`.

For each product term in the statement text, we choose the Term ID with the highest cosine similarity. Matches with a cosine similarity below 0.8 are discarded. After this process, each trademark carries a set of Term IDs, effectively categorizing the trademarked products into the predefined categories from the USPTO Trademark ID Manual.

Appendix B. Additional Tables and Figures

Table B1
Variable description

Variable Name	Abbrev.	Description
Firm growth		
Profits		Sales minus cost of goods sold [COMPUSTAT: SALE - COGS deflated by CPI]
Output		Sales (COMPUSTAT: sale) plus change in inventory [COMPUSTAT: INVT deflated by CPI]
Capital		[COMPUSTAT: PPEGT deflated by the NIPA price of equipment]
Employment		[COMPUSTAT: EMP]
Total factor productivity	TFP	Revenue-based TFP constructed using methodology of Olley and Pakes (1996) applied using procedure in İmrohoroğlu and Tüzel (2014)
Firm characteristics		
Log Assets		Natural log of (1 + Firm's total asset [COMPUSTAT AT])
Tobin's Q		Sum of total assets plus market value of equity minus book value of equity divided by total assets [COMPUSTAT (AT+CSHO× PRCC_F - CEQ) / AT)].
Log Market Cap		Natural log of total market capitalisation [COMPUSTAT CSHO×PRCC_F]
Leverage		Firm's total debt divided by total assets [COMPUSTAT (DLTT + DLC)/AT]
HHI		Herfindahl-Hirschman Index of firm's SIC4 industry sales concentration in a given year [COMPUSTAT: SALES, SIC]
Exporter		Indicator variable equal to one if the firm exports in a given year, zero otherwise [COMPUSTAT SEGMENTS]
OI/AT		Firm profitability [COMPUSTAT OIADP / AT]
Return on Assets	ROA	Earnings before interest and taxes divided by total assets [COMPUSTAT EBIT / AT]
Sales Growth		Sales growth from last fiscal year end [COMPUSTAT ($SALE_t/SALE_{(t-1)}$ -1)].
Asset Growth		Sales growth from last fiscal year end [COMPUSTAT ($AT_t/AT_{(t-1)}$ -1)].
Dividends		Firm's total dividends issued [COMPUSTAT DVT/AT]
Dividend Payer		Firm's total dividends issued ¿ 0 [COMPUSTAT DVT]
Capital expenditure	CAPX	Capital expenditures divided by total assets [COMPUSTAT CAPX/AT]
Property, plant, and equipment	PPE	Property, plant, and equipment scaled by asset [COMPUSTAT PPENT/AT]
Cash holding		Cash assets scaled by asset [COMPUSTAT CHE/AT]
Cash Flow		Cash Flow scaled by asset [COMPUSTAT CHECH/AT]
R&D	R&D	R&D Research and development expenses divided by total assets [COMPUSTAT XRD/AT]
Missing R&D		Indicator variable equal to 1 if R&D expense is missing, zero otherwise
Advertising	R&D	Advertising expenses divided by total assets [COMPUSTAT XAD/AT]
Firm Age		Number of years since founding year (Jay R. Ritter website)
Listing Age		Number of years since the firm first appears in CRSP database
Acquirer		Indicator variable equal to one if the firm acquired another firm in a given year, zero otherwise [SDC]
Target		Indicator variable equal to one if the firm is an acquisition target in a given year, zero otherwise [SDC]
Divest		Indicator variable equal to one if the firm divested in a given year, zero otherwise [SDC]
Equity Issue		Indicator variable equal to one if the firm issued equity in a given year, zero otherwise [SDC]
Debt Issue		Indicator variable equal to one if the firm issued debt in a given year, zero otherwise [SDC]
Repurchase		Indicator variable equal to one if the firm repurchased equity in a given year, zero otherwise [SDC]
Vertinteg		Firm's vertical intergration level [Hoberg-Phillips Data Library]
Idiosyncratic volatility	$\sigma_{i,t}$	Sum of squared abnormal returns [CRSP DSF: $\sum_{d\in t}(r_i - r_m)^2$]
Patent Value		Estimated patent value [from Kogan et al. (2017)]

Table B2
Trademark data variable description

The table presents definitions for the key variable in the trademark database.

Variable	Notes	Example
Word Mark	Actual characters of mark text	OFFICE
Good and Service Statement	Owner's description of the use of the mark	Computer software and computer programs . . .
Primary G&S Class	G&S Classification primary code for the use of the mark under 45 G&S classes	Class IC 009 Electronics
Mark Drawing Code	Identifies mark physical characteristics	Design plus words, letters and/or numbers
Design Search Code	Identifies Mark Graphic Design characteristics	26.05.21 Triangles that are completely or partially shaded
Serial Number	Unique Identifier for trademark application	85799081
Registration Number	Registration number	December 24, 2013
First Use in Anywhere Date	First time for the G&S classification used in anywhere	2012-12-12
First Use in Commerce Date	First time for the G&S classification used in commercial activities (first commercialization date)	2012-12-12
Filing Date	Date application is filed	December 10, 2012
Filing Basis (current)	Whether trademark is used in commerce	1A: Used in Commerce
Filing Basis (original)	Whether trademark is used in commerce when filed	1B: Intent to use
Publication Date	Date mark published for opposition	May 21, 2013
Registration Cancel Date	Date cancellation of entire registration recorded	
Renewal Date	Date Trademark registration renewed	
Owner Name	Name of owning party	Microsoft Corporation
Owner Address	Owner address	Microsoft Way, Redmond Washington 98052-6399
Attorney		Matthew E.Moersfeider

Table B3
List of stopwords

a	d	here	myself	shouldn't	wasn't
about	did	hers	needn	should've	we
above	didn	herself	needn't	so	were
after	didn't	him	no	some	weren
again	do	himself	nor	such	weren't
against	does	his	not	t	what
ain	doesn	how	now	than	when
all	doesn't	i	o	that	where
am	doing	if	of	that'll	which
an	don	in	off	the	while
and	don't	into	on	their	who
any	down	is	once	theirs	whom
are	during	isn	only	them	why
aren	each	isn't	or	themselves	will
aren't	few	it	other	then	with
as	for	its	our	there	won
at	from	it's	ours	these	won't
be	further	itself	ourselves	they	wouldn
because	had	just	out	this	wouldn't
been	hadn	ll	over	those	y
before	hadn't	m	own	through	you
being	has	ma	re	to	you'd
below	hasn	me	s	too	you'll
between	hasn't	mightn	same	under	your
both	have	mightn't	shan	until	you're
but	haven	more	shan't	up	yours
by	haven't	most	she	ve	yourself
can	having	mustn	she's	very	yourselves
couldn	he	mustn't	should	was	you've
couldn't	her	my	shouldn	wasn	

Table B4
Commercialization Value of Product Types — Full sample

The table presents the distribution of commercialization value of the trademarked products, measured as described in Section 2.6. The sample period is 1929 to 2020.

	BS	QI	NV
Mean	21.56	27.11	9.63
Std	85.27	96.35	32.82
10%	0.24	0.33	0.21
25%	0.81	1.15	0.64
40%	1.94	2.73	1.38
50%	3.25	4.52	2.25
60%	5.38	7.46	3.46
75%	12.29	16.81	7.41
90%	42.61	57.16	19.18
Obs.	160,882	279,561	2,277

Table B5
Product Types, Patents and Firm Growth

The table presents the regression results for the relation between product value and firm-level outcomes. We estimate regressions of the form:

$$ln(\frac{Y_{i,t+\tau}}{Y_{i,t}}) = \beta_0 + \beta_1\theta^{BS}_{i,t} + \beta_2\theta^{QI}_{i,t} + \beta_3\theta^{NV}_{i,t} + \beta_4\theta^{Pats}_{i,t} + \gamma_1 ln(Y_{i,t-1}) + \gamma_2 ln(Y_{i,t-2}) + \Gamma' X_{i,t} + \alpha_{sic3} + \delta_t + \epsilon_{i,t},$$

where Y are the firm outcomes, $\theta^c_{i,t}$ are the aggregated market value (ξ) of each product type introduced (use in commerce year) by and all patents granted to firm i in year t scaled by total assets, $X_{i,t}$ is a vector of control variables including log capital stock, log of employment, log total asset, and idiosyncratic volatility, and $\tau \in 1-10$. α_{sic3} and δ_t are industry and year fixed effects. All right-hand-side variables are scaled to unit standard deviation. Panel A presents results for firm gross profits (sales minus cost of good sold), Panel B for firm sales, Panel C for firm capital stock (property, plant, and equipment), and Panel D for firm employment. Standard errors are clustered at the firm level. ***, **, * indicates significance level at 1%, 5% and 10%, respectively. All unbounded variables are winsorized at 1% level using annual breakpoints. The sample period is 1976 to 2020.

	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
					Panel A. Profits					
θ^{BS}	0.028***	0.041***	0.047***	0.074***	0.075***	0.089***	0.100***	0.104***	0.108***	0.110***
	(3.49)	(3.15)	(3.14)	(4.99)	(4.69)	(5.12)	(5.31)	(5.50)	(5.59)	(5.28)
θ^{QI}	0.018***	0.025***	0.030***	0.030***	0.037***	0.041***	0.044**	0.046**	0.044**	0.041**
	(3.74)	(3.46)	(3.32)	(2.86)	(2.85)	(2.63)	(2.43)	(2.33)	(2.17)	(1.98)
θ^{NV}	0.015**	0.027***	0.033***	0.021	0.038***	0.048***	0.042**	0.041	0.051***	0.068***
	(2.42)	(3.27)	(3.62)	(1.21)	(2.70)	(2.78)	(2.11)	(1.40)	(2.88)	(5.04)
θ^{Patent}	0.014***	0.028***	0.047***	0.059***	0.061***	0.061***	0.061***	0.063***	0.063***	0.064***
	(5.54)	(6.26)	(7.75)	(7.53)	(6.53)	(5.80)	(5.40)	(5.27)	(4.89)	(4.74)
					Panel B. Sales					
θ^{BS}	0.030***	0.051***	0.060***	0.072***	0.076***	0.085***	0.098***	0.089***	0.089***	0.098***
	(4.21)	(4.76)	(4.75)	(5.25)	(5.40)	(5.93)	(6.18)	(5.65)	(5.42)	(5.28)
θ^{QI}	0.008***	0.015***	0.020***	0.025***	0.032***	0.034***	0.034**	0.036**	0.037**	0.036**
	(2.58)	(2.96)	(3.03)	(3.08)	(3.09)	(2.81)	(2.41)	(2.35)	(2.27)	(2.10)
θ^{NV}	0.013***	0.020***	0.021***	0.025**	0.034***	0.042**	0.026	0.042*	0.054***	0.065***
	(3.31)	(2.83)	(2.72)	(2.44)	(2.73)	(2.26)	(0.90)	(1.71)	(3.49)	(5.39)
θ^{Patent}	0.014***	0.029***	0.046***	0.053***	0.056***	0.057***	0.057***	0.059***	0.059***	0.058***
	(6.11)	(7.13)	(9.97)	(9.37)	(7.99)	(6.96)	(6.48)	(6.55)	(6.18)	(5.86)
					Panel C. Capital					
θ^{BS}	0.037***	0.053***	0.065***	0.076***	0.079***	0.085***	0.099***	0.094***	0.095***	0.100***
	(5.87)	(5.43)	(5.40)	(5.42)	(4.72)	(5.03)	(4.98)	(4.62)	(4.51)	(5.20)
θ^{QI}	0.009***	0.019***	0.021**	0.026**	0.034***	0.039***	0.040***	0.044***	0.051***	0.053***
	(2.65)	(2.58)	(2.17)	(2.30)	(2.81)	(2.83)	(2.61)	(2.84)	(3.09)	(2.70)
θ^{NV}	0.012**	0.012	0.011	0.025**	0.028	0.040*	0.039	0.044*	0.050**	0.057***
	(2.25)	(1.55)	(1.18)	(1.97)	(1.60)	(1.66)	(1.60)	(1.84)	(2.51)	(3.52)
θ^{Patent}	0.012***	0.026***	0.044***	0.055***	0.063***	0.070***	0.072***	0.073***	0.073***	0.074***
	(6.17)	(7.03)	(9.17)	(9.85)	(9.30)	(8.67)	(8.14)	(7.85)	(7.51)	(7.30)
					Panel D. Employment					
θ^{BS}	0.028***	0.044***	0.052***	0.059***	0.063***	0.076***	0.091***	0.089***	0.088***	0.095***
	(4.45)	(4.24)	(4.24)	(4.53)	(4.43)	(4.82)	(4.94)	(4.73)	(4.63)	(4.75)
θ^{QI}	0.011***	0.022***	0.029***	0.035***	0.043***	0.047***	0.048***	0.051***	0.053***	0.050***
	(3.90)	(4.92)	(4.96)	(4.54)	(4.60)	(4.28)	(3.77)	(3.66)	(3.42)	(2.77)
θ^{NV}	0.012	0.012	0.014	0.022**	0.028***	0.031	0.028	0.032	0.046***	0.046***
	(1.46)	(1.51)	(1.58)	(2.46)	(3.01)	(1.41)	(1.26)	(1.46)	(3.40)	(3.53)
θ^{Patent}	0.010***	0.022***	0.036***	0.044***	0.047***	0.050***	0.049***	0.050***	0.051***	0.053***
	(5.58)	(6.26)	(8.63)	(10.00)	(9.68)	(8.73)	(7.97)	(7.63)	(7.41)	(7.18)
Obs.	126,228	113,907	103,041	93,644	85,315	77,859	71,154	65,110	59,630	54,664

Appendix C. Additional Robustness

Table C1
Investment long sample

This table presents the relationship between product types and investment choices. We run regressions of the following form:

$$Y_{i,t} = \beta_0 + \beta_1 BS_{i,t-1} + \beta_2 QI_{i,t-1} + \beta_3 NV_{i,t-1} + \beta_c' X_{i,t-1} + \alpha_{sic3} \times \delta_t + \epsilon_{i,t},$$

where Y is a different investment outcome variable, BS is the number of business stealing products, QI is the number of quality improvement products, and NV is the number of new variety products introduced by firm i in year t. The outcomes variables are: M&A outcomes (Acquirer or Target indicator variable) in columns (1) and (2), R&D and capital expenditures (CAPX) in columns (3) and (4). Acquirer and target are indicator variables equal to one when the firm is an acquirer or a target in an MA deals, otherwise zero. R&D and CAPX are R&D are scaled by total assets. T-statistics adjusted for clustered standard errors at firm level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All variable definitions are presented in Table B2. All potentially unbounded variables are winsorized at 1% level. The sample period is 1950 to 2020.

	Acquirer	Target	R&D	CAPX
	(1)	(2)	(3)	(4)
No. BS	0.226***	-0.009**	-0.010**	-0.003
	(9.44)	(-2.44)	(-2.28)	(-0.89)
No. QI	0.173***	-0.003	0.003	-0.008***
	(8.66)	(-1.27)	(0.64)	(-2.81)
No. NV	-1.822***	0.093	0.219***	0.228***
	(-3.69)	(1.50)	(2.82)	(3.32)
ln(Assets)	4.379***	0.038**	-0.522***	0.202***
	(43.13)	(2.45)	(-14.39)	(8.58)
ln(Age)	-1.606***	-0.059***	-0.257***	-0.495***
	(-12.08)	(-2.70)	(-5.64)	(-17.12)
Tobin's Q	0.753***	-0.049***	0.945***	0.404***
	(8.47)	(-3.31)	(22.00)	(20.44)
HHI	0.056	-0.076	-1.948***	0.328
	(0.05)	(-0.44)	(-4.60)	(1.25)
Cash Flow	3.495***	-0.025	-3.301***	1.520***
	(4.29)	(-0.13)	(-12.63)	(11.44)
Sale growth	7.932***	-0.034		
	(35.68)	(-0.69)		
ROA	9.452***	-1.062***		
	(15.66)	(-6.73)		
Leverage	-0.121**	-0.001		
	(-2.07)	(-0.07)		
R&D	-0.220***	-0.014***		
	(-11.94)	(-4.03)		
Industry × Year FE	✓	✓	✓	✓
Obs.	169,657	169,657	169,811	169,811
Adj. R^2	0.21	-0.01	0.42	0.33

Table C2
Investment control scope

This table presents the relationship between product types and investment choices, controlling for product scope Hoberg and Phillips (2024). We run regressions of the following form:

$$Y_{i,t} = \beta_0 + \beta_1 BS_{i,t-1} + \beta_2 QI_{i,t-1} + \beta_3 NV_{i,t-1} + \beta_c' X_{i,t-1} + \alpha_{sic3} \times \delta_t + \epsilon_{i,t},$$

where Y is a different investment outcome variable, BS is the number of business stealing products, QI is the number of quality improvement products, and NV is the number of new variety products introduced by firm i in year t. Product scope is measured by d2vscope. The outcomes variables are: M&A outcomes (Acquirer or Target indicator variable) in columns (1) and (2), R&D and capital expenditures (CAPX) in columns (3) and (4). Acquirer and target are indicator variables equal to one when the firm is an acquirer or a target in an MA deals, otherwise zero. R&D and CAPX are R&D are scaled by total assets. T-statistics adjusted for clustered standard errors at firm level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All variable definitions are presented in Table B2. All potentially unbounded variables are winsorized at 1% level. The sample period is 1976 to 2020.

	Acquirer	Target	R&D	CAPX
	(1)	(2)	(3)	(4)
No. BS	0.246***	-0.009**	-0.015***	-0.003
	(9.68)	(-2.47)	(-3.03)	(-1.04)
No. QI	0.146***	-0.003	0.007*	-0.006**
	(7.50)	(-1.21)	(1.72)	(-2.38)
No. NV	-0.810	0.137	0.271**	0.108
	(-1.00)	(1.26)	(2.46)	(1.14)
Scope	0.155***	0.000	0.179***	-0.023***
	(4.04)	(0.05)	(14.22)	(-3.71)
ln(Assets)	4.679***	0.039**	-0.695***	0.198***
	(42.65)	(2.30)	(-17.38)	(8.09)
ln(Age)	-1.482***	-0.065**	-0.238***	-0.510***
	(-9.80)	(-2.57)	(-4.62)	(-16.15)
Tobin's Q	0.869***	-0.054***	0.957***	0.374***
	(9.46)	(-3.56)	(20.90)	(19.14)
HHI	0.531	-0.080	-2.062***	0.342
	(0.42)	(-0.42)	(-4.40)	(1.26)
Cash Flow	3.309***	-0.009	-3.409***	1.516***
	(4.18)	(-0.05)	(-12.58)	(11.62)
Sale growth	7.199***	-0.029		
	(33.69)	(-0.62)		
ROA	8.676***	-1.065***		
	(14.35)	(-6.71)		
Leverage	-0.128**	0.001		
	(-2.24)	(0.08)		
R&D	-0.213***	-0.014***		
	(-12.07)	(-4.08)		
Industry × Year FE	✓	✓	✓	✓
Obs.	148,442	148,442	148,592	148,592
Adj. R^2	0.19	-0.00	0.42	0.33

Table C3
Investment Robustness — SIC4

This table has the same specification with table 6 replacing SIC 3 FE with SIC 4 FE. T-statistics adjusted for clustered standard errors at Industry level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All potentially unbounded variables are winsorized at 1% level. The sample period is 1976 to 2020.

	Acquirer	Target	R&D	CAPX
	(1)	(2)	(3)	(4)
No. BS	0.239***	-0.008**	-0.002	-0.004
	(9.27)	(-2.20)	(-0.53)	(-1.20)
No. QI	0.148***	-0.003	0.007*	-0.005*
	(7.41)	(-1.09)	(1.73)	(-1.76)
No. NV	-0.650	0.130	0.204**	0.141
	(-0.79)	(1.20)	(1.96)	(1.43)
ln(Assets)	4.814***	0.035**	-0.623***	0.175***
	(43.14)	(2.05)	(-15.97)	(6.89)
ln(Age)	-1.576***	-0.060**	-0.256***	-0.483***
	(-10.33)	(-2.32)	(-5.01)	(-15.12)
Tobin's Q	0.865***	-0.053***	0.896***	0.375***
	(9.27)	(-3.42)	(19.52)	(19.20)
HHI	-227.518**	81.552	13.786**	22.835
	(-2.27)	(1.01)	(2.17)	(1.09)
Cash Flow	3.294***	-0.009	-3.561***	1.529***
	(4.09)	(-0.05)	(-13.03)	(11.70)
Sales growth	7.132***	-0.031		
	(32.86)	(-0.65)		
ROA	8.421***	-1.047***		
	(13.57)	(-6.44)		
Leverage	-0.123**	0.003		
	(-2.10)	(0.15)		
R&D	-0.187***	-0.015***		
	(-10.39)	(-4.17)		
Industry × Year FE	✓	✓	✓	✓
Obs.	147,537	147,537	147,685	147,685
Adj. R^2	0.20	-0.00	0.45	0.34

Table C4
Investment HP300 FE

This table has the same specification with table 6 with HP300 fixed effects. T-statistics adjusted for clustered standard errors at Industry level are reported in the parentheses. ***, **, * indicates significance level at 1%, 5% and 10%. All potentially unbounded variables are winsorized at 1% level. The sample period is 1976 to 2020. This table has the same specification with table 6 replacing SIC3 FE with HP300 FE.

	Acquirer	Target	R&D	CAPX
	(1)	(2)	(3)	(4)
No. BS	0.246***	-0.006	-0.007*	-0.006*
	(8.51)	(-1.51)	(-1.72)	(-1.92)
No. QI	0.116***	-0.002	0.022***	-0.006**
	(5.58)	(-1.07)	(6.89)	(-2.25)
No. NV	-1.395	-0.145	0.014	-0.071
	(-1.24)	(-0.80)	(0.11)	(-0.69)
ln(Assets)	5.159***	0.015	-0.808***	0.207***
	(38.27)	(0.70)	(-19.36)	(8.50)
ln(Age)	-1.450***	-0.058*	-0.229***	-0.380***
	(-7.75)	(-1.90)	(-4.26)	(-11.20)
Tobin's Q	0.902***	-0.052***	0.738***	0.347***
	(8.20)	(-2.88)	(15.58)	(17.20)
HHI	1.136	-0.223	-1.548***	-0.437**
	(0.99)	(-1.41)	(-5.32)	(-2.24)
Cash Flow	3.515***	0.012	-3.651***	1.295***
	(3.82)	(0.06)	(-12.61)	(9.68)
Sales growth	8.760***	-0.003		
	(30.71)	(-0.06)		
ROA	9.533***	-1.067***		
	(13.05)	(-5.81)		
Leverage	-0.322***	0.017		
	(-4.71)	(0.87)		
R&D	-0.151***	-0.013***		
	(-7.53)	(-3.25)		
Industry × Year FE	✓	✓	✓	✓
Obs.	105,699	105,699	105,809	105,809
Adj. R^2	0.14	0.01	0.51	0.33