Patents to Products: Product Innovation and Firm Dynamics*

David Argente — Penn State

Salomé Baslandze — FRB Atlanta & CEPR

Douglas Hanley — U. of Pittsburgh

Sara Moreira — Northwestern

April 15, 2020

Abstract

We study the relationship between patents and actual product innovation in the market, and how this relationship varies with firms' market share. We use textual analysis to create a new data set that links patents to products of firms in the consumer goods sector. We find that patent filings are positively associated with subsequent product innovation by firms, but at least half of product innovation and growth comes from firms that never patent. We also find that market leaders use patents differently from followers. Market leaders have lower product innovation rates, though they rely on patents more. Patents of market leaders relate to higher future sales above and beyond their effect on product innovation, and these patents are associated with declining product introduction on the part of competitors, which is consistent with the notion that market leaders use their patents to limit competition. We then use a model to analyze the firms' patenting and product innovation decisions. We show that the private value of a patent is particularly high for large firms as patents protect large market shares of existing products.

JEL Classification Numbers: O3, O4
Keywords: Product innovation, patents, creative destruction, growth, productivity, patent value.

---

*Emails: dargente@psu.edu; salome.baslandze@atl.frb.org; doughanley@pitt.edu; sara.moreira@kellogg.northwestern.edu.

We are grateful to Antonin Bergeaud, Pete Klenow, and Jesse Perla for very helpful discussions. We also thank Ufuk Akcigit, Fernando Alvarez, Toni Braun, Benjamin F. Jones, Claudio Michelacci, Juan Rubio-Ramirez and numerous seminar and conference participants at Harvard Business School, University of Austin, Chicago Fed, Atlanta Fed, Bank of Portugal, EIEF, EPFL, Northwestern, Rochester, University of Lugano, and University of Milan; CEPR Symposium, NBER SI Innovation, NBER SI Economic Growth, NBER Productivity Group, SED (Mexico), Midwest Macro. Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Data sets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. The views expressed in this paper do not necessarily represent the views of the Federal Reserve System.

1 Introduction

Product innovation – the introduction of new and improved products to the market – is a key contributor to economic growth and a central element of endogenous growth models (Romer, 1990; Aghion and Howitt, 1992). However, the paucity of detailed data about the introduction and quality of new products has led researchers to use other metrics to measure innovation. As a result, patents have emerged as the primary metric of innovation, especially after comprehensive data sets with information about their timing and characteristics were made readily available (Griliches, 1981).

Indeed, many great inventions like solar panels and liquid-crystal displays (LCD) have been patented. Yet no patents have been filed for other inventions that have transformed our lives in important ways like the World Wide Web and the magnetic strip behind modern-day credit cards. In other cases, firms file patents that never turn into new products in the market.[1] These examples suggest that the relationship between patents and innovation is complex and that patents are a crude measure of innovation. Patents are also a protective tool that firms can use to preempt competitors from entering their product market space. This protection may be especially advantageous for large market leaders because of greater incentives to defend their existing lead (Gilbert and Newbery, 1982; Jaffe and Lerner, 2004). While the protective role of patents is important in shaping firm dynamics and industry competition, we lack large-scale systematic evidence on the nature of the relationship between patents and product innovation.

In this paper we use textual analysis to create a unique data set that links patents to products of firms in the consumer goods sector. We use these data to study the relationship between patents and actual product innovation, and how this relationship varies with firms' market share. After documenting a set of new facts, we complement our empirical analysis with a theoretical model of the patenting and product innovation decisions of firms.

Our key empirical findings can be summarized as follows:

- Fact 1: More than half of product innovation comes from firms that do not patent.
- Fact 2: On average, patents are positively associated with subsequent product innovation by firms.

[1] In fact, in recent years, patenting activity has skyrocketed whereas innovation and productivity growth have not (Bloom, Jones, Van Reenen and Webb, 2017).

- Fact 3: Larger firms have lower product innovation rates (quantity and quality), but file more patents for each new product.
- Fact 4: Patenting by larger firms is strongly associated with an increase in revenue above and beyond the patents' effect on product innovation.
- Fact 5: Patenting by larger firms is associated with a decline in product introduction by competing firms.

There are two main challenges in studying the relationship between patents and product innovation. First, while patent data are broadly available, measures of product innovation in the market are rarely available at large scale. To address this challenge, we use comprehensive data for firms and products in the consumer goods sector collected from Nielsen-Kilts point-of-sale systems in retail locations. This data set includes detailed information about the characteristics of each consumer-goods product sold from 2006 to 2015, along with sales and price information. We exploit this rich data set to construct measures of product innovation. Our simplest measure is the number of new products (barcodes) introduced at the firm and product category level in a given year. Since many new products represent only minor innovations relative to existing products, we also construct measures of the quality-adjusted number of new products. We infer quality improvements by tracking the new attributes (e.g. formula, style, content) that a product brings to the market and by exploiting variation in product prices and sales.

The second challenge involves linking product innovations to their respective patents. We address this challenge by developing two distinct matching procedures. The first procedure maps each firm's patents to its full product portfolio using the names of firms in the patent and product data sets. This results in a yearly firm-level data set (Match 1). The matching procedure in this step is simple and parsimonious, but it is too coarse. Most firms in our data are active in several product categories and could be patenting products in some product categories and not in others, so we need a more granular procedure to match patents with products. We leverage the richness of the information about product and patent characteristics in our data and use modern methods from the field of natural language processing and information retrieval to link firms' patents with sets of its products (Manning et al., 2008). For this match, we first define product categories – sets of similar products – by applying clustering analysis to the short product descriptions included in the Nielsen data extended with text from Wikipedia articles about the products. We then analyze the text of patent applications and assign each patent to the product category with which it has the highest text similarity.[2] This classification of firms' products and patents into various product categories of the firm results in our benchmark patent-to-products data set at the yearly firm × product category level (Match 2).

The resulting granular data set tracks patents and products for firms in the consumer goods sector. The patenting intensities and product introduction rates of these firms are, on average, comparable to those of other manufacturing sectors. Out of 35 thousand firms covered in our data set, 15% applied for a patent at least once (9% applied during the period covered by Nielsen). This value is in line with that of the manufacturing sector and is substantially higher than that of other sectors in the economy (Graham et al., 2018). The consumer-goods sector also covers a wide range of product categories with distinct patenting intensities. The share of patenting firms varies from zero in some food categories to more than half in printers or water purification products.

We begin our analysis by documenting that never-patenting firms account for a large share of product innovation. Over our sample period, they introduced more than 54% of new products and more than 65% quality-adjusted new products. These shares are larger if we rely on the patent-to-products link at the firm × category level. These statistics are corroborated by similar statistics about sources of growth in the sector. We decompose the 10-year sales growth of the sector into growth coming from patenting and non-patenting firm × categories, and find that although non-patenting firms are smaller, they account for 58% of growth in the sector.

Nonetheless, we find that patenting is positively associated with product innovation both at the extensive margin – when firms switch to patenting – and at the intensive margin. Firms introduce more and better-quality products around the time of a patent application, with the largest correlation in the year following the application. Exploiting our matched firm × category-level data over time, which allows us to control for product category-specific trends and firm-category specific effects, we find that the elasticity of product innovation to the number of patents filed a year before ranges from 0.02 to 0.04. We find similar patterns when we focus our attention on granted patents or on patents that receive many forward citations, but not when we consider non-granted patents or uncited patents. This evidence suggests that commonly used measures of the quality of patents are informative about product innovation rates. These elasticities are instructive in the context of various

[2] Younge and Kuhn (2016) and Kelly et al. (2018) use similar techniques when evaluating textual similarities between patents.

policies meant to encourage innovation.[3] The evaluation of these policies often relies on the estimated elasticity of patents to R&D inputs. However, by and large, patents are not the main policy target of such policies – innovation is. Hence, to study how policies encouraging R&D affect product innovation, for instance, one needs to take into account not only the R&D-to-patents elasticity, but also the patents-to-product innovation elasticity.

The importance of patents for protecting firms' products and deterring competition draws attention to the question of how different firms use this strategic role of patenting. Gilbert and Newbery (1982) and Blundell et al. (1999) suggest, for instance, that market leaders have greater incentives to use preemptive patenting to protect their market lead. Survey results from Cohen et al. (2000) report that the motives behind large firms' patenting often go beyond the direct commercialization of patented innovations and extend to strategic deterrance of rivals. For this reason, a major focus of our paper is to understand how the relationship between patents and product innovation changes with a firm's market leadership.

Using variation across firms within product categories, we estimate that firms at the bottom quintile of the size distribution in a given year, as measured by total sales in a product category, introduce one new product for every five existing products in their portfolio, on average. Firms at the top quintile of the size distribution, on the other hand, introduce one new product for every seven existing products in their portfolio. Though larger firms' innovation rates are lower, they are patenting more intensively. But the patent filings of larger firms have significantly weaker association with their product introduction. Moreover, the average quality improvements of new products decline more steeply with firm size than the rate of product introduction does. Overall, these empirical patterns indicate that the disconnect between patent-based measures of innovation and firms' actual product innovation in the market is bigger for firms with large market shares.

Our results suggest that the main role of patents for these market leaders is to constrain product innovation of competitors and thereby protect sales of their existing products. First, we find that patents filed by market leaders carry a larger revenue premium, even after controlling for the quantity and quality of new products these firms introduce. To the contrary, for smaller firms the revenue premium is fully accounted for by product innovation associated with these patents. Second, we show that patent filings by market leaders are associated with a decline in competitors' product introduction in shared product categories.

[3] For example, R&D tax incentives and subsidies as in Dechezlepretre et al. (2016) or Akcigit et al. (2016a).

The same is not true if we consider patent filings of smaller firms.[4]

In addition to the empirical analysis, we consider the patenting and product innovation decisions of firms in a model. The model builds on quality-ladder models that feature creative destruction (Aghion and Howitt, 1992), but it allows for separation between the decision to innovate and the decision to patent – a distinction we can discipline with the data set we have constructed. In the model, both innovation and patenting are costly activities. Introducing higher quality products increases a firm's profit, while patenting decreases the firm's chances of being displaced by entrants. The model can replicate key empirical facts from our data. Larger firms (market leaders) shift their product innovation towards protective strategies, meaning an increase in the number of patents by large firms restricts competition and innovation and does not translate into higher consumer welfare. We use the model to provide a back of the envelope calculation for the private value of a patent and to decompose it into its protective and productive components.[5] The productive component represents the option value of implementing a patented idea into higher-quality products in order to gain additional profits. The protective component represents the gains for the firm from impeding creative destruction by competitors. After calibrating the model to our data, we estimate that 43% of the average patent value is accounted by the protective component of the patent, and this share increases substantially with firm size.

Our new data set that combines information on product innovation and patents contributes to our understanding of the usefulness of patent statistics for measuring innovation. In the absence of direct measures of innovation, the literature has relied on indirect inference approaches using data about employment growth (Garcia-Macia et al., 2019) or valuing innovation from patent statistics themselves (e.g., Akcigit and Kerr, 2018). Other researchers have looked at innovations that occur outside of the patent system by examining the number of new books on technical topics (Alexopoulos, 2011) or innovations featured at World Fairs between 1851 and 1915 (Moser, 2012). While we document an overall positive relationship between patents and product innovation, we highlight that the usefulness of patent metrics in inferring innovation significantly hinges on the size of the firms that own the patents.

[4] Various studies have analyzed the effect of patenting on follow-on innovation by other firms: for example, see Williams (2013), Heller and Eisenberg (1998), Sampat and Williams (2019) for biomedical research; Cockburn and J. MacGarvie (2011) for the software industry; and Lampe and Moser (2015) for a more general discussion.

[5] This decomposition is possible because we directly observe both the sales from products linked to patents as well as the behavior by competitors. The previous approach in the literature to infer the (total) monetary value of a patent using surveys, samples of patent sales, or patent renewals is discussed in Section 6.

The findings in this paper contribute to our understanding of firms' growth strategies. Recent studies have shown that large firms rely on other protective strategies such as acquiring potential competitors (Cunningham et al., 2018) or forging political connections (Akcigit et al., 2018) as they slow down on innovation (Akcigit and Kerr, 2018; Cavenaile and Roldan, 2019). We suggest that patenting is yet another protective tool that firms substitute for actual product innovation as they grow.

Additionally, our findings regarding the patenting and innovation decisions of firms can potentially speak to several puzzling macroeconomic trends in recent data: patenting is soaring, but productivity growth is stagnating (Gordon, 2016; Bloom et al., 2017); large firms funnel more resources into intangible capital – including intellectual property, but this is manifested in the increasing dominance of those firms instead of perceptible improvements in aggregate innovation in the economy (Crouzet and Eberly, 2019). Our results show that the incentives of large incumbents to direct their efforts towards productive rather than protective strategies may be limited, which is particularly relevant as more economic activities are reallocated towards firms with a large degree of market power (De Loecker and Eeckhout, 2017; Autor et al., 2017; Gutiérrez and Philippon, 2017; Akcigit and Ates, 2019).

The rest of the paper is organized as follows: in Section 2, we present a description of the data sets along with our data matching procedures. We also discuss the validation exercises and present summary statistics. In Section 4, we explore the relationship between patents and product innovation. Section 5 explores the role of firm size. Section 6 presents the theoretical framework and the patent value calculation. Section 7 concludes.

2 Patent and Product Data

2.1 Overview

We face two main challenges in our study of the relationship between patents and product innovation. First, while data about patents are broadly available, information about the introduction of new products is rarely available at large scale. Second, the link between patents and the related new products is challenging to create. This section overviews the empirical strategies we use to address these challenges.

We construct a data set about product introduction by beginning with product-scanner data that cover the product portfolio of firms in the consumer goods sector between 2006 and 2015. This data set allows us to identify new products by their barcodes and to observe their detailed characteristics from which we can compute various measures of innovation for a large representative sector. We draw patent information from the United States Patent and Trademark Office (USPTO). The combination of these two data sets gives us information about patents and product innovations covering a large sector of the economy.

To address the second challenge of linking patents to products, we develop several matching procedures. We begin by using the names of the firms in the patent and product data sets to produce a mapping between firms' patent portfolios and their respective products. We refer to this firm-level data set as Match 1. This matching procedure is simple and parsimonious, but is too coarse to allow us to connect patents with specific products.

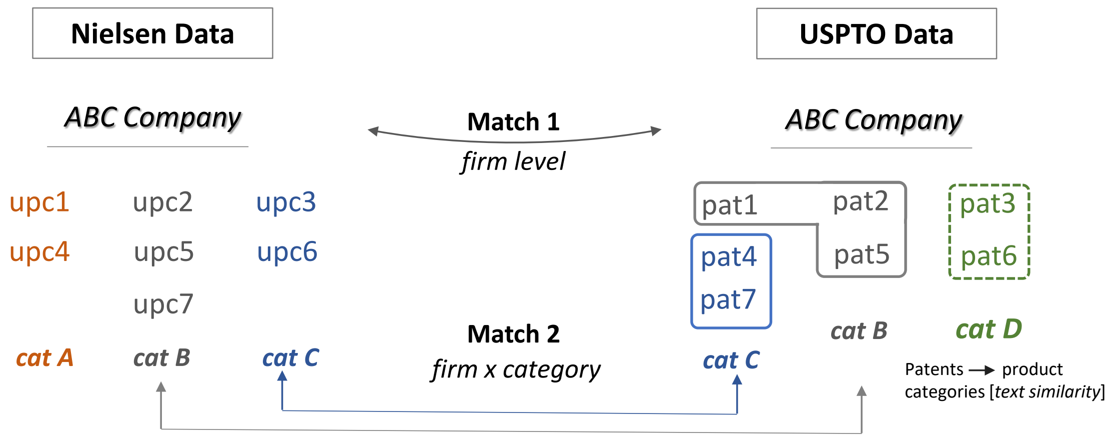
In turn, our second matching procedure leverages the richness of product and patent characteristics using methods from the natural language processing literature to create systematic links between sets of patents and sets of products within a firm. A patent may generate no products or multiple products, and a product may have benefited from multiple patents or from none at all. Therefore, forcing a one-to-one matching between a specific narrowly defined product and a specific patent is neither possible nor desirable.

Hence, our approach is to first define product categories as sets of similar products, which are identified using clustering analysis of product descriptions extended with Wikipedia-based dictionaries. We then assign each specific patent to the product category with which it has the highest text similarity. This classification of a firm's products and patents into the various product categories offered by that firm yields our benchmark patent-to-products data set, which we will refer to as Match 2. Figure 1 illustrates our data schematically, and our matching algorithms are described in detail below.

2.2 Data

Product Data.—Our primary source of product information is the scanner data set from Nielsen Retail Measurement Services (RMS), provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. This data is collected from point-of-sale systems in grocery, drug, and general-merchandise stores. The original data set consists of more than one million distinct products identified by Universal Product Codes (UPCs), which are scanned at the point of sale. Each UPC consists of 12 numerical digits that are uniquely assigned to each product, and we use these to identify products. UPCs carry information about the brand and a rich set of product attributes like its size, packaging, formula, and flavor.

Figure 1: Product and Patent Data Sets



Notes: This diagram exemplifies the construction of the two data sets linking products and patents. In this example, under Match 1, all products of a firm with name "ABC Company" match to all the patents with assignee name "ABC Company". Under Match 2, upc2, upc5, and upc7 match to pat1, pat2, and pat5 under product category B; upc3 and upc6 match to pat4 and pat7 under product category C; upc1 and upc4 of category A do not match to any patents of the firm; pat3 and pat6 of category D do not match to any products of the firm.

The data focuses on the consumer product goods (CPG) sector, which accounts for 14% of the total consumption of goods in the U.S.[6] The Nielsen RMS dataset covers about 40% of the CPG sector, and nearly covers the universe of firms and product introductions in the sector (Argente et al., 2020). Our sample period covers the years from 2006-2015, and combines all sales, quantities, and prices at the national and annual levels. We use the panel structure of each UPC to measure its entry year. Online Appendix A.1 provides more detail about types of products and store coverage represented in our sample.

Patent Data.—Our main source of data for patent analysis is the USPTO data on the universe of published patent applications, granted or not. We use the original bulk data files provided by USPTO's Bulk Data Storage System for our analysis. Our sample initially contains information on more than 7 million patent applications filed by more than 500 thousand patent assignees in the years 1975-2017. For each patent, we use information about the patent application year, patent status (granted, pending, or abandoned), patent technology classifications, forward patent citations received, the number of claims on a patent, and whether it is a utility or design patent. For our textual analysis of patent documents, we extract patent titles, the text of patent abstracts, the text of corresponding patent classification titles, claims text, and the titles of citing patents. Online Appendix A.2 gives more

[6]This sector includes non-durables (also known as consumer packaged goods) and semi-durable goods. It excludes consumer durables, producer intermediates, and producer capital.

detail about our sample and the variables we use.

2.3 Matching Firms

In our firm-level data set (Match 1), we match patents to products at the firm level using the firm names in both patent and product data sets. To match firms to patents, we obtain the firm names for each product using UPCs and data set from the GS1 US, which is the single official source of UPCs. This data set links barcodes with the names of firms that sell the product. For the patent data, we begin with the assignee name(s) of each patent. This name is typically the original assignee of the patent and may not represent the current user of the patent because of sales or company reorganizations. We combine the USPTO patent re-assignment data with Thomson Reuters Mergers & Acquisition data to re-assign each patent to its most current holder. This step relies on the assumption that when a firm acquires (or merges with) another firm, the new firm will own all patents that the firms owned before the acquisition (merger). The details of these steps are described in Online Appendix A.1 and A.2.

A firm's name could be formatted or abbreviated differently in the product and patent data sets, or it may even be misspelled, which presents a challenge in joining information from the two data sets. We developed a name-cleaning algorithm to clean and standardize the firm names to overcome this challenge. This procedure builds on and extends cleaning algorithms from the NBER Patent Data Project (Hall et al., 2001) and Akcigit et al. (2016b) and is described in detail in Online Appendix A.3.

2.4 Matching Patents to Product Categories

In this section, we describe the details of the algorithm we used to build Match 2 (Firm $\times$ Category level). The algorithm has three crucial steps. The first step creates product categories at a level of aggregation such that they collect distinct and sufficiently large sets of similar products that are meaningfully related to a distinct set of patents. This step yields a set of product categories, a vector of terms used to describe each product category, and a mapping of products into categories. In the second step, similarity scores between patents and product categories are computed. We use various text descriptions to build a vector of terms that describe each patent. We then compute similarity scores between each patent and every product category. These scores represent the overlap between the texts in patents and the text associated with each product category. The final step of our patent-product matching algorithm consists in using the similarity scores and information about the production of the respective patenting firms to generate a procedure that systematically classifies each patent into a product category.

Defining Product Categories.—We define product categories by exploring the product classification scheme used by Nielsen. In the original data, each product is classified into one of 1,070 detailed product modules. These product modules are further aggregated into a set of 114 product groups, and those are further aggregated into ten departments. For example, "disposable cups" and "disposable dishes" are two distinct product modules that are part of the group "paper products" which is part of the department "non-food grocery". Nielsen's modules aggregate products that are close in their technological characteristics. However, there are some sets of distinct modules that have very similar products. At the same time, many Nielsen's groups include products that are quite distinct. For example, "disposable cups", "disposable dishes", "pre-moistened towelettes" and "paper napkins" are all part of the group "paper products", but only "disposable cups" and "disposable dishes" are technically similar. Hence, we seek an intermediate categorization of products – more aggregated than modules and less aggregated than groups – to be able to meaningfully associate patents to a well-defined set of products.

To this end, we apply a clustering procedure to aggregate the Nielsen modules into distinct product categories. Each module is paired with a vector of descriptive terms[7] that are weighted by their importance. We expand short module descriptions from Nielsen data with the text of hand-collected Wikipedia articles to get to the comprehensive description of product content of the modules. The resulting vectors of descriptive terms collect all the words from the Wikipedia and Nielsen texts, after applying standard parsing and lemmatizing algorithms. When building term vectors, one must appropriately weight terms by their importance. We use the leading approach in textural analysis – the "term-frequency-inverse-document-frequency" sublinear transformation – that accounts for both the frequency with which a term appears describing a module and how commonly it is used to describe other modules (Aizawa, 2003). This approach ensures that we under-weight common terms that appear in many documents as these are less diagnostic of the content of any individual document.

We then aggregate these module vectors into clusters using a popular technique known as k-means clustering (Lloyd, 1982). This procedure allows one to specify the desired number

[7] We use single words and two-word phrases as terms.

of clusters and yields a clustering assignment that minimizes the within-cluster term vector variance. As a baseline, we use an aggregation of modules into 400 clusters which we refer to as product categories. We find that this partition strikes a balance between aggregating very similar products while maximizing the difference between products across categories. Appendices A.4.1 and A.4.2 provide extensive descriptions of methods we have taken from the literature on natural language processing, including the details of clustering, quality assessment, and alternative methods to encourage robustness.

After defining the level of aggregation, we build term vectors describing each product category. We use the same methodology that we used to build the term vectors for modules, but now we use the titles of the clustered module(s) and all the text from their corresponding Wikipedia articles. We ensure that when a product category aggregates multiple modules, we first vectorize each module description and then average these vectors together so that we do not overweight longer entries. The final product category vectors are normalized to have unit length.

Patent Vectors and Similarity Scores.—This subsection describes how we measure the amount of overlap between the texts of patent applications and product categories. For the patent description, we use the following fields from the patent applications: the title, abstract, international patent classification description, and the titles of cited patents. We create vectors of terms by concatenating all these fields into one document, followed by the same parsing and lemmatizing algorithms. As before, we adjust the weights of each term according to the "term-frequency-inverse-document-frequency" sublinear transformation and normalize patent vectors to have unit length.

Finally, we construct a similarity score for each patent p and each product category j by computing the cosine similarity between two normalized vectors, $s_{jp} = f_j \times f_p$. This similarity score is guaranteed to be in the range $[0, 1]$ with zero indicating no word overlap and one indicating that the documents are identical. Online Appendix A.4.3 provides technical description of this step.

Classifying Patents into Product Categories.—The final step of our patent-products matching algorithm consists in using the similarity scores and determining which product categories are valid matches for each patent. We must, however, make some adjustments because we use all patents of each firm with products in the consumer goods sector, and some patents may relate to goods outside the consumer goods sector or correspond to more general process/method patents. Hence, we should allow for the possibility that a patent will not be assigned to any product category. After an extensive review of patent texts and a great deal of testing, we identified systematic adjustments to the algorithm that ensure that irrelevant patents remain unmatched with products.

We first adjust the algorithm to include a similarity score threshold. We tested different threshold levels and, in our baseline algorithm, we restrict the set of potential categories for each patent p to the product categories whose similarity score exceeds 0.025. The idea is that patents with low text similarity are unrelated to the product categories that we consider. The implication of this adjustment is that patents whose highest similarity are below that threshold are more likely to be classified as "non-matched".

Second, we use information about the set of product categories sold by the firm. For each patent, we define the set of potential matches, whose elements consist of all product categories in which the patenting firm ever sold a product, according to our product data. Together, these criteria imply that patent p will be classified as unmatched if no product categories satisfy the threshold similarity and belong to the set of categories the firm produces in. For the patents that have more than one product category satisfying those conditions, we assign the final patent-product category match so that the patent matches to the product category with the highest similarity score.

Our methodology assumes one product category match for each patent. However, some patents may be more general in nature so that they relate to multiple categories. Our baseline algorithm abstracts from this possibility. However, our procedure to define product categories is designed to ensure that the product categories would encompass a broad range of products that are technically similar such that one patent plausibly relates to this and only this range of products.[8] In Online Appendix A.4.4, we present the details of this procedure and all the robustness exercises with which we tested our baseline algorithm.

2.5 Match Statistics and Validation

Table 1 provides statistics of the baseline data used in our analysis. The data set Match 1 (Firm level) includes annual data for all 34,665 firms that sold at least one product in our consumer goods sector data (CPG firms). The raw USPTO patent data covers information from 1975 to 2017, but because our product data only covers years from 2006 to 2015, our

[8]In this sense, the methodology delivers a many-to-many patent-to-product match, where each patent can be matched to multiple products of the firm.

analysis can only consider annual variation for the period 2006-2015. In this shorter period, the USPTO data includes about 3.4 million patent applications in total, and about 500 thousand patent applications filed by CPG firms. The data set Match 2 (Firm $\times$ Category level) includes 40% of those patent applications. The remaining 60% of patents, while filed by CPG firms, could not be associated with products in the consumer goods sector.

Table 1: Match Statistics

	Period	
	1975-2017	2006-2015
Number of patent applications		
All assignees in USPTO	7,304,072	3,386,208
CPG firms (Match 1)	1,046,030	505,544
CPG firms in product categories (Match 2)	399,684	190,575
Number of firms		
All CPG firms		34,665
CPG with at least a patent applied in 1975-2017		5,209
CPG with a patent applied in 2006-2015		3,266

Notes: Match statistics for the baseline data sets Match 1 (Firm level) and Match 2 (Firm $\times$ Category level). Match 1 is described in Section 2.3 and Match 2 is described in Section 2.4.

We perform an extensive set of validation exercises to evaluate the robustness and quality of our match. Online Appendix A.5 presents details on these validation exercises, while here we focus on summarizing the most important. We use four main types of validation exercises: manual checks, external validations using online-collected data on patent markings, analysis of the robustness of the algorithm-implied similarity scores and placebo tests, and validation of non-matches.

Manual checks.—We manually checked many of the patent-to-products matches and some examples are listed in Table 1 of Online Appendix. The table lists 100 patent applications by the top-selling firms in the largest product categories according to Nielsen. One can easily see that the patent titles reflect well the product categories to which the patents were assigned. For most patents we analyzed, we found that our manual choices of product categories also coincide with the product categories chosen by our matching algorithm using similarity scores.

Virtual patent markings.—We next use virtual patent markings to validate our matches. Using virtual patent markings, firms may give a notice to the public that their product is patented by publishing their products and the patents protecting them online. Website searches showed that very few firms in our data used virtual patent markings, and even when they did, only a selection of products and patents appeared in the markings. Nevertheless, these data give a unique opportunity for an external validation of our matching algorithm.

For Procter & Gamble (P&G) and Kimberly Clark (KC), we manually collected virtual patent markings from the company websites and mapped them to our product categorization. We then validate our patent-product category matches for these firms against this information. Online Appendix A.5.2 shows that the patent-product category mapping from virtual markings is also selected by our matching algorithm in about 70% of cases.

Robustness of the match and placebo tests.—We evaluate robustness of the product category choice by our matching algorithm to potential small perturbations in the algorithm. For the algorithm to be robust against small changes, we should observe that highest-ranked product categories have substantially higher similarity scores with the patents than lower-rank product categories do. Section A.5.3 of Online Appendix shows this is the case. Next we verify that we are indeed carving out well-defined neighborhoods in the technological space by matching patents into distinct categories. For that, we compare the actual distribution of similarity scores between patents classified in the same product category versus a placebo group of patents drawn at random. Section A.5.4 of Online Appendix shows that the distribution of similarity scores between pairs of patents within product categories is indeed very different and first order stochastically dominates that of the placebo group.

Validating non-matches.—In our last step of the algorithm for Match 2, multiple criteria are used to allow for the possibility that some patents filed by CPG firms are not associated with any of the consumer-good product categories. A valid "non-match" can arise for two main reasons. First, a patent may relate to goods that the firm may be producing outside the CPG sector; second, a patent may be about a general process or method that does not affect the introduction of new products. In the spirit of Hoberg and Phillips (2016), we use information from publicly traded companies' 10K reports to identify firms whose output is mostly in the consumer-goods sector, and we find that only a minority of their patents are classified as "non-match", contrasting with patents held by firms who mostly sell products outside the consumer goods sector. Next, we follow Bena and Simintzi (2017), and use patent claims to create proxies for process-related and product-related patents. We find that the share of "non-matches" is significantly higher among process-related patents. These exercises, which are presented in Section A.5.5 of Online Appendix, offer reassurance that our algorithm successfully filters out patents that are not directly related to the products in our data.

3 Measures of Product Innovation and Patenting

3.1 Product Innovation

Our measures of product innovation are based on the number of products that firms introduce to the market and the quality improvements in those products. We use the product data described above to identify the entry dates of products in the market and their respective characteristics and performance. We create separate measures of innovation for the firm-level (Match 1) and firm×category level (Match 2) data. Our first measure is the number of **new products** of firm i (in product category j) in year t:

$$\mathrm{N}_{ijt} \equiv \sum_{u=1}^{T_{ijt}} \mathbb{1}[u \text{ is entrant}],$$

where product u is sold by firm i in product category j, T_{ijt} is the number of products that firm i sells in j as of period t, and $\mathbb{1}[u \text{ is entrant}]$ is an indicator that takes the value of one if u is a new barcode in t. This measure is simple and parsimonious but does not distinguish major product innovations from innovations that make relatively minor changes to a product's characteristics. Our second set of measures of **quality-adjusted new products** deals with this potential drawback by explicitly accounting for differences in characteristics across new products:

$$\mathrm{qN}_{ijt} \equiv \sum_{u=1}^{T_{ijt}} \mathrm{q}_u \mathbb{1}[u \text{ is entrant}],$$

where $\mathrm{q}_u \in [0, 1]$ is a measure of quality that we describe below. Together these two metrics allow us to account for differences in both the quantity and quality of product innovation across firms and over time.

Our baseline measure of product quality aims at capturing differences in novelty and economic impact across new products. We build on Argente and Yeh (2017) and use detailed information on product attributes that is available from the product data. Products can then be compared on the basis of characteristics associated with their attributes $\{v_{u,1}, ..., v_{u,A}\}$.[9] We test if each new product has characteristics distinct from those of all existing products

[9] For example, "children" and "regular" are two mutually exclusive characteristics associated with the attribute "formula" for "pain remedies-headache" products. Naturally, the number and type of attributes varies across product categories. For example, the product category "pain remedies-headache" includes 10 attributes: brand, flavor, container, style (i.e. children, regular), form, generic, formula (i.e. regular, extra strength, rapid release), type (i.e. aspirin), consumer (i.e. trauma, migraine), and size. On average, we observe that the different product categories include between 5 to 12 attributes. Online Appendix B gives details.

available in the market, and we compute the quality of a new product as a weighted sum of its novel characteristics across all product attributes:

$$q_u \equiv \sum_{a=1}^{A} \omega_a \mathbb{1}[v_{ua} \text{ is new}].$$

where ω_a are weights that reflect the economic value associated with a particular attribute. We develop a novel approach to estimate weights that capture the importance of each attribute by using "shadow prices" from hedonic pricing regressions (Bresnahan and Gordon, 1996). The underlying assumptions here are that the degree of novelty of a product should be reflected in the price of a product and that the price of a product reflects its embodied characteristics as valued by implicit or shadow prices. A new product has a high novelty score if it has many new characteristics and/or if its characteristics are associated with high implicit prices. We provide details on the properties of this procedure in Online Appendix B, along with some evidence that the novelty score is strongly associated with the performances of the firm and its products.[10]

We use three alternative measures of new product quality to evaluate the robustness of our empirical results. First, we use a simpler version of the quality measure that weighs each attribute equally (quality *q1*). This measure only captures variation in the share of new product characteristics contained in a product. Second, we develop a measure that is computed much like our baseline measure with the exception that it uses weights that reflect "shadow sales" (quality *q2*). This measure assigns lower quality to new products that are associated with high shadow prices but do not reach many customers. Finally, we use a measure of residual demand taken from Hottman et al. (2016) and Argente et al. (2020) (quality *q3*). This measure does not use information about the degree of novelty of a product and instead captures the relative appeal of new products relative to other products sold by the firm, under some functional-form assumptions. Overall, our baseline measure and these alternative metrics allow us to consider many critical dimensions of the quality of new products, and allow us to assess the robustness of our results.

3.2 Patent Measures

Using an approach similar to how we measured product innovation, we compute measures that allow us to account for differences in the quantity and quality of patent applications

[10]We show that our measure is correlated with the growth rate of the firm, the share of sales generated by new products, and the average duration of new products in the market even after conditioning on the number of products being introduced by the firm (Table 2 in the Online Appendix).

across firms and over time. Our baseline measure is the number of **patent applications** (P_{it}). Using our patent-product category match, we are also able to measure the number of patent applications filed by firm i in product category j in year t as follows:

$$\mathrm{P}_{ijt} \equiv \sum_{p=1}^{P_{it}} \mathbb{1}[p \text{ is match to } j].$$

Throughout the paper, we use information about whether a patent was granted and information about patent citation counts to compute our measures of patent quality. Patent applications that become **granted patents** (gP_{ijt}) are perceived as high-quality patents because the patent office deemed them novel enough to not be rejected. We compute the number of patent applications that are granted as:[11]

$$\mathrm{gP}_{ijt} \equiv \sum_{p=1}^{P_{it}} \mathbb{1}[p \text{ is granted}] \times \mathbb{1}[p \text{ is match to } j].$$

We also define **patent citations** (cP_{ijt}) as the total number of patents weighted by forward citations received in the first five years since the application was filed:[12]

$$\mathrm{cP}_{ijt} \equiv \sum_{p=1}^{P_{it}} c_p \times \mathbb{1}[p \text{ is match to } j].$$

Measures based on forward citations have traditionally been used to assess the economic and technological significance of a patent (for earlier contributions, see Pakes (1986), Schankerman and Pakes (1986), Trajtenberg (1990)).

3.3 Summary Statistics

Table 2 provides summary statistics about the product- and patent-related variables for the firms in our sample, grouped by their patenting activity. We split firms into three groups: (i) firms that have never filed a patent application, (ii) firms whose last patent application was filed before 2006 (the beginning of the Nielsen RMS data set) and (iii) firms that filed a patent application between 2006 and 2015.

[11] The condition $\mathbb{1}[p \text{ is match to } j]$ is only used for Match 2.

[12] A 5-year citations measure attempts to reduce the truncation issue inherent to citations – the fact that patents filed more recently have had less time to accumulate citations (Hall et al., 2001).

Table 2: **Summary Statistics by Firm's Patenting Status**

	No Patents	Patents before 2006	Patents 2006-2015
Product data			
Number of products	15.49	31.08	78.35
Number of new products (N)	2.58	5.26	13.45
Average quality of new products (q)	0.27	0.20	0.20
Quality-adjusted number of new products (qN)	0.46	0.62	1.48
Product introduction rate (n)	0.19	0.17	0.22
Quality-adjusted product introduction rate (qn)	0.07	0.04	0.06
Sales from all products	2371.59	9392.09	37094.71
Sales from new products	454.74	1811.01	8130.00
Number of product categories	2.36	3.07	5.46
Average quality of new products ($q1$)	0.13	0.10	0.10
Average quality of new products ($q2$)	0.18	0.11	0.12
Average quality of new products ($q3$)	0.06	0.32	0.10
Patent data			
Number of patent applications (P)	0.00	0.00	6.34
Number of granted patent applications (gP)	0.00	0.00	4.57
Number of citations-weighted patent applications (cP)	0.00	0.00	5.88
Stock of patent applications	0.00	11.33	125.36
Stock of granted patent applications	0.00	11.02	107.63
Stock of citations-weighted patent applications	0.00	17.97	215.24
Number of firms	29215	1943	3266
Observations	186934	15803	29052

Notes: The table shows the average of product-based and patent-based variables of the Match 1 data set. The first column groups firms that have no patents; the second column considers firms that have patents, but filed them before they first appear in Nielsen RMS (before 2006); and the third column is for firms that have patents in our focus period of 2006-2015. The statistics regarding product introduction can only be computed for the period 2007-2015 because we cannot determine entries for products first introduced in 2006 (left censored). The statistics for sales are given in thousands of dollars, deflated by the Consumer Price Index for all urban consumers.

The share of patenting firms and product introduction rates in the consumer goods sector are comparable to those of other manufacturing sectors. Table 2 shows that more than 5 thousand firms (15%) applied for at least one patent and more than 3 thousand firms (9.5%) filed a patent application during the period 2006-2015. For comparison, Graham et al. (2018) links Census data to the USPTO and finds that 6.3% of manufacturing firms have at least one granted patent application between 2000 and 2011.[13] The corresponding number in our data is 7.6%, which is only slightly higher.[14] Table 2 indicates that product introduction rates are on average 20%. While there is no equivalent comprehensive product data for other sectors, Goolsbee and Klenow (2018) use the Adobe Analytics data on online transactions covering

[13]The incidence of patenting in the rest of the economy is lower, at 1%.

[14]Notice that Graham et al. (2018)'s patent data includes only granted patents, while our data also includes unsuccessful patent applications. If we count only granted applications, we would have 2629 patenting firms.

multiple products, and report product introduction rates that are comparable to those of other non-durable consumer manufacturing sectors.[15] Our data covers product categories that exhibit substantial heterogeneity in patenting intensity and entry rates. The share of patenting firms varies from zero in some food categories to a maximum of about 60% in water purification products, and the product introduction rates vary from about 5% in some food categories to a maximum of about 30% in printer supplies.

As expected, patenting firms are larger: they sell more products, operate in more product categories, and have higher sales. Firms that filed patents between 2006 and 2015 account for 61% of sales in our sample. Patenting firms also introduce more products, but this relationship is weaker once we account for scale and instead focus on the rates with which new products are introduced. Interestingly, our four different quality measures indicate that the average novelty of new products sold by patenting firms is not higher than that of non-patenting firms, conditional on product introduction.

Firms with patent applications between 2006 and 2015 file more than six patents per year, on average.[16] Because many patents receive no citations, especially in the first five years, the average number of citation-weighted patent applications, cP_{ijt}, is very similar to the average raw number of patent applications, P_{it}. These firms may hold some design patents, but the majority of patents in our sample are utility patents. Unsurprisingly, the summary statistics show that firms who filed a patent between 2006 and 2015 hold a larger stock of patents than firms who last filed a patent application before 2006. Actively patenting firms hold, on average, 125 patents in stock every year whereas firms that last patented before 2006 hold approximately 11 patents.

4 Relationship Between Product Innovation and Patents

How do patents relate to actual product introduction to the market? How much product innovation in the consumer goods sector is captured by patent-based metrics of innovation? We document the relationship between patents and product innovation using following exercises. First, we show the cross-sectional allocation of product innovation between patenting and non-patenting firms. Second, we consider how a firm's product introduction changes after it files a first patent application. Third, we quantity the strength of the relationship

[15]Goolsbee and Klenow (2018) show that some durable consumer goods (e.g. furniture), not covered in our data set, have entry rates that are larger than those of non-durables (e.g. food).

[16]Patent statistics are very skewed, and we present averages after winsorizing patent-based variables at top 0.1%.

between the changes in the number of patents filed and the amount of product innovation. Finally, we explore the dynamics of these effects. Our findings can be summarized in the following two empirical facts:

Fact 1: More than half of product innovation comes from firms that do not patent.

Fact 2: On average, patents are positively associated with subsequent product innovation by firms.

Product Introduction and Firm's Patenting Status.—We begin our analysis of the relationship between patents and product innovation by exploring cross-sectional variation across firms according to their patenting status. Table 3 shows that in our data, 54% of new products were introduced by firms that never applied for a patent. If we account for the degree of novelty of new products, we estimate that about 65% of quality-adjusted product introduction comes from never-patenting firms. This indicates that, on average, patenting firms introduce more products that make only an incremental improvement over existing products on the market.[17]

Since they rely on the firm-level match, the above statistics implicitly attribute all new products introduced by a patenting firm to some of its patents. However, highly diversified firms might be patenting in one product category, while introducing many products that have no relation to the patents they are filing in other categories. Thus we may be attributing too much product introduction to patents if we rely only on the firm's overall patenting status. This observation exemplifies the importance of establishing a closer link between patents and products using the Match 2 data set. To make these more granular links, we replicate the above exercise but define patenting status at the firm $\times$ category level. As seen from Table 3, firms that never patented in a category are responsible for a greater share of new products introduced in that category.

It is not surprising that a large amount of innovation may not be associated directly with any patents. Even if firms wanted to patent all their new products, some new products represent only small upgrades to existing products, and may not be patentable. Patents are only granted if they exhibit "novelty and non-obviousness", and thus many new products that result from very small changes will not be captured by patent metrics. While it is natural

[17] This observation holds true regardless of the quality adjustment we use. For example, the share of $q1N$ accounted by never-patenting firms is 65%, and the share of $q2N$ by never-patenting firms is 77%. Our residual quality measure of innovation, $q3$, does not allow us to construct a good counterpart to $q3N$, however as seen from Table 2, $q3$ is not necessarily higher for patenting firms.

Table 3: Share of Product Innovation Accounted for by Patenting Firms

	New Products, N	Quality-adjusted New Products, qN
Match 1		
Firms with patents in 2006-2015	.38	.28
Firms with patents before 2006	.08	.07
Firms with no patents	.54	.65
Match 2		
Firm × category with patents in 2006-2015	.23	.16
Firm × category with patents before 2006	.07	.05
Firm × category with no patents	.71	.79

Notes: the table shows the share of product innovation on the market measured by our two benchmark measures – product introduction (column 1) and quality-adjusted product introduction (column 2) – accounted for by firms and firm×categories with or without patents.

that some innovations are not captured by patents, our data offers a unique opportunity to quantify the magnitude of it.

We also evaluate if our measures of product innovation reflect well the sources of growth. Indeed, if we look through the lens of classic innovation-driven growth models, we should expect innovation and growth measures to go hand-in-hand. We conduct simple growth decompositions for our sector to get at this question. We decompose sales growth from 2006 to 2015 into growth that comes from patenting and non-patenting firm × categories as:

$$\underbrace{\text{Growth}_{\text{'}06-\text{'}15}}_{7\%} = \underbrace{\text{Growth}^{\text{Patent}}_{\text{'}06-\text{'}15}}_{4\%} \times \underbrace{s^{\text{Patent}}_{2006}}_{0.72} + \underbrace{\text{Growth}^{\text{No Patent}}_{\text{'}06-\text{'}15}}_{14.4\%} \times \underbrace{s^{\text{No Patent}}_{2006}}_{0.28} \quad (1)$$

where s^{Patent}_{2006} and $s^{\text{No Patent}}_{2006}$ denote sales shares of firm× categories with or without patents, respectively.[18] As with our measures of product innovation, these growth decompositions show that although non-patenting firms are smaller and account for a smaller share of sales in the sector, they contribute more to growth relative to the set of patenting firms – totaling to 58% of the sectoral growth.[19] Hence, the fact that more than a half of the product innovation in the sector is not captured by the patenting status of the firms is corroborated by similar statistics about growth.

[18]We first write $Rev^{CPG}_t = \sum_j \sum_{i\in\Omega^j_{\text{Patent}}} Rev_{ijt} + \sum_h \sum_{i\in\Omega^j_{\text{No Patent}}} Rev_{ijt}$, where the second sum is across product categories and Ω denotes the set of firms with and without patents in category j; and take the percentage changes in sales to arrive at (1).

[19]These observations are not surprising and should not be limited to our sector: as documented above, non-patenting firms are smaller, and smaller firms tend to contribute less to growth on average, as we know from other studies of the overall economy (Haltiwanger et al., 2013).

First-time Patent Filers.—One important feature of our data is that we observe some firms that change their patenting status in the period of analysis 2006–2015. This allows us to evaluate whether a firm's product introduction tends to change after the firm's first patent application.[20] We do so by estimating the following specification:

$$\log Y_{it} = \beta\, dP_{it} + \alpha_i + \gamma_t + u_{it} \tag{2}$$

where Y_{it} is the outcome of firm i in year t, α_i represents firm fixed effects, and γ_t represents year effects. dP_{it} is an indicator variable that equals 1 after the firm's first patent application. Our goal is to understand if the switch to patenting is associated with increased product innovation, which would be the case if patent-based measures were to approximate well product innovation in the market. To uncover this relationship, we estimate the effects of β relative to firms that are already patenting. These have more similar characteristics and thus are likely a more suitable counterfactual for firms that first apply for a patent than those that never apply.[21]

Table 4 presents the estimated change in our two measures of product innovation associated with a firm's transition from non-patenting to patenting. Conditional on firm and year effects, we find an average increase in product introduction of up to 11% after the switch to patenting. Columns (2) and (3) show that the positive correlation is largely driven by high-quality patent applications, if we take the patent's success with the patent office as a proxy of quality. This result is more pronounced if we study the effect on quality-adjusted product introduction, as shown in columns (4) to (6). These exercises reveal a positive correlation between the timing of patent applications and product innovation. One interpretation of this correlation is that firms come up with ideas for new products and then apply for a patent to protect the idea from being copied by competitors; simultaneously, they develop those ideas into new consumer products.[22] Our findings about the changes in the observed product

[20]Although defining the event of the first patent application at the firm level (as opposed to firm-category) avoids potential cross-category spillovers in patenting and sharpens out definition of the event, we also explore the dynamics of this relationship in detail using Match 2 data later on.

[21]The assumption that this group of firms forms a better control − after accounting for time-invariant differences between firms and common year factors − is supported by the summary statistics presented in Table 2. Nevertheless, we find similar estimates when we test if our results are explained by the contrast with the entire sample of non-switching firms, which includes firms already patenting before the beginning of our sample and those that have not yet patented at the end of our sample (Table 3 in the Online Appendix).

[22]It may also be the case that patenting gives firms a preferential status in the economy that allows them to create more products in the future. For example, this status may result from consumer perception that firms are more "innovative" when firms advertise "patent pending" on their products.

introduction around the time that a firm first files a patent add to other papers' findings that patenting is associated with other real changes at the firm level, such as increases in stock market prices or the firm's sales and scope (Hall et al., 2005; Balasubramanian and Sivadasan, 2011; Kogan et al., 2017).

Table 4: Product Innovation after First Patent Application

	Log N			Log qN		
	(1)	(2)	(3)	(4)	(5)	(6)
After patent(t)	0.1168**			0.0352		
	(0.045)			(0.020)		
After granted patent(t)		0.1361**			0.0497**	
		(0.048)			(0.018)	
After non-granted patent(t)			-0.0045			-0.0085
			(0.044)			(0.036)
Observations	29,470	29,470	29,470	29,470	29,470	29,470
Time	Y	Y	Y	Y	Y	Y
Firm	Y	Y	Y	Y	Y	Y

Notes: The table shows regressions of log number of new products (Log N) in Panel A and of log quality-adjusted new products (Log qN) of a firm as a function of a dummy equal to one after the first patent application by the firm. Our benchmark quality measure is defined in Section 3.1. The alternative innovation quality measures ($q1, q2, q3$) produce similar results. Both Log N and Log qN use the inverse hyperbolic sine transformation. *After patent* is a dummy equal to one after any patent application; *After granted patent* is a dummy equal to one after a patent application that is granted; and *After non-granted patent* is a dummy equal to one after a patent application that has not been granted (abandoned or pending). The sample includes 596 firms that switch to patenting and those that already patented before the beginning of our sample. Standard errors robust against heteroskedasticity and serial correlation are reported in parentheses.

Elasticity of Product Innovation to Patents.—We next explore how product innovation varies with the changes in intensive margin of patenting exploiting variation in measures of product innovation at the firm × category level over time. We estimate

$$\log Y_{ijt} = \beta \, \log P_{ijt-1} + \alpha_{ij} + \gamma_{jt} + u_{ijt} \tag{3}$$

where Y_{ijt} is the outcome for firm i in category j in year t and P_{ijt-1} is the log number of patent applications filed by the firm i in category j a year before to allow for a short lag between patent filing and product commercialization. Thanks to the firm × category level data, we can now control for product category-specific trends (e.g., market-wide demand for specific products), and we can control for firm-category specific effects, thus filtering out, for instance, the effects of firm-specific brand power on the sales of specific products.

We seek to provide reduced-form estimates for the elasticity of product innovation to patents (β). One can think of this elasticity along the lines of the knowledge production function approach (Griliches, 1979). In that approach, the estimated elasticity is for patents (output) with respect to R&D (inputs). However, by and large, patent filings are not the main policy target of various innovation policies (e.g., R&D tax incentives) – innovation in the market is. In our case, patents can be considered as inputs to the production function of the output, which is product innovation. Hence, if one is interested, for instance, in how much a policy that encourages additional R&D spending affects product innovation, one needs to account for an additional elasticity that shows the strength of the relationship between the patents and actual innovation, in addition to the patents-to-R&D elasticity.[23]

Table 5 shows the estimates. The rows present results from using different explanatory variables – the log number of patents, granted patents, and non-granted patents. Conditional on firm-category and category-time fixed effects, we find that the observed elasticities of product introduction and quality-adjusted product introduction to patents are 0.04 and 0.02, respectively. As before, the relationship between patenting and product innovation is mainly driven by higher-quality granted patents. Likewise, Table 5 of Online Appendix provides similar results for other quality measures of patents – citations and claims. Overall, these exercises show that we can statistically identify a positive correlation between patenting and product introduction, which corroborates that firms' patenting is positively associated with their product innovation.

The estimated elasticity captures the relationship between product introduction and patents associated with products. Not all patents, however, necessarily relate to product improvements: some patents may relate to cost savings from improvements to the firm's general production processes. Nevertheless, our firm $\times$ category data set filters out patents that are not specifically related to product introductions. Hence, to a large extent, our estimates should be driven by product patents rather than process patents. To support this point, we consider the robustness of our results and employ proxies for product-related and process-related patents drawn from Bena and Simintzi (2017); we find that the coefficient on product-related patents is essentially same as our benchmark coefficient, while process-related patents are wholly unrelated to measures of product innovation (Section A.2 and Table 6 of Online Appendix).

[23] To underline this point with a simple illustration, consider a standard set-up where $Patents \propto R\&D^{\alpha}$, and α is the knowledge production elasticity estimated in the literature (e.g. Dechezlepretre et al. (2016)). At the same time, $Innovation \propto Patents^{\beta}$. Hence, the effect of additional spending on $R\&D$ for $Innovation$ should, in practice, be inferred from the combined elasticity $\alpha\beta$.

Table 5: Product Innovation and Patenting

	Log N			Log qN		
	(1)	(2)	(3)	(4)	(5)	(6)
Patents(t-1)	0.0380***			0.0189***		
	(0.009)			(0.005)		
Patents granted(t-1)		0.0405***			0.0192***	
		(0.010)			(0.005)	
Patents non-granted(t-1)			0.0234*			0.0082
			(0.013)			(0.007)
Observations	409,641	409,641	409,641	409,641	409,641	409,641
R-squared	0.692	0.692	0.692	0.623	0.623	0.623
Time-Category	Y	Y	Y	Y	Y	Y
Firm-Category	Y	Y	Y	Y	Y	Y

Notes: The table shows regressions of the log number of new products (log qN) and of log quality-adjusted new products (log qN) in a firm × category over time as a function of the log number of patents. Our benchmark quality measure is defined in Section 3.1. The alternative innovation-quality measures ($q1, q2, q3$) produce consistent results. *Patents* is the log number of any patent applications in firm × category × year; *Patents granted* is the log number of granted patent applications; and *Patents non-granted* is the log number of patent application that have not been granted (abandoned or pending). The inverse hyperbolic sine transformation is used for logs. Standard errors robust against heteroskedasticity and serial correlation are reported in parentheses.

Dynamics of the Effects.—We are now interested in evaluating the timings of the effects captured in (2) and (3). Thus, we study the relationship between patents and product innovation by running the following separate linear regressions using the firm-category level data set:

$$Y_{ijt+k} = \beta_k \mathrm{E}_{ijt} + \alpha_{ij} + \gamma_{jt} + u_{ijt+k} \ , \qquad k = -4, ..., 0, ..., 4 \tag{4}$$

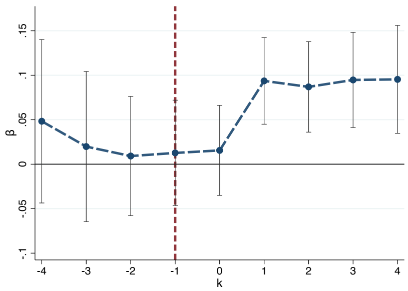
where Y_{ijt+k} is an outcome of firm i in product category j in $t+k$ associated with product introduction and E_{ijt} is either dP_{ijt} (as before, a dummy equal to one after firm starts patenting in category j) or $\log P_{ijt}$, which again denotes the log number of patents filed by firm i in product category j in t. We also include firm-product category and time-product category fixed effects.

Figure 2 plots the estimated coefficients β_k over k. The top panel shows the evolution of N and qN around the time at which the firm starts patenting in a certain product category. The bottom panel is about the intensive margin of patenting, and both are based on patent application years. Consistent with the results above, we find a positive association between patents and product introduction. Our estimates indicate that firms introduce about 10% more products after filing their first patent, with no pre-trends in outcomes before the firm switches to patenting (and 3-4% if we adjust for the novelty of new products– see (a) and (b)). The positive association reaches its maximum magnitude shortly after the first patent is filed in a product category and is fairly persistent thereafter.

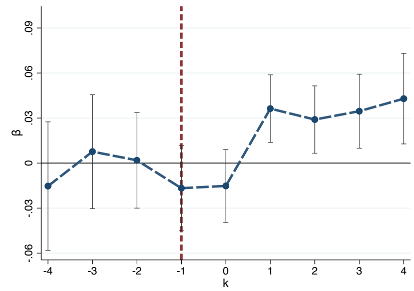
Figure 2: Product Innovation and Patenting: Dynamics

— *Switching to patenting* —

(a) $\beta_{\mathrm{N,dP}}$

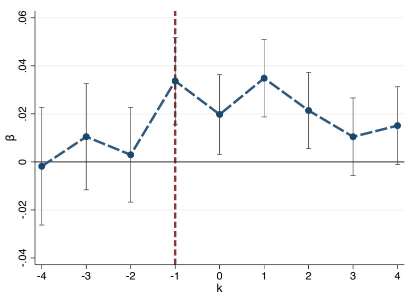


(b) $\beta_{\mathrm{qN,dP}}$

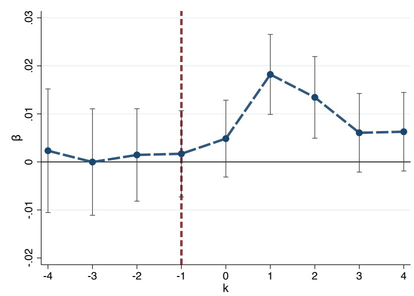


— *Number of patents* —

(c) $\beta_{\mathrm{N,P}}$



(d) $\beta_{\mathrm{qN,P}}$



Note: The figure plots the estimated coefficients after estimating equation (4) for log product introduction, N, in (a) and (c), and quality-adjusted product introduction, qN, in (b) and (d). Our benchmark quality measure is defined in Section 3.1. The main explanatory variable in (a) and (b) is a dummy equal to one after the firm's first patent in a product category and log number of patent applications in (c) and (d). The inverse hyperbolic sine transformation is used for logs. The vertical bands represent +/- 1.65 $\times$ st. error of each point estimate. Standard errors are clustered at the firm $\times$ category level.

Likewise, our results (see (c) and (d)) exploring the co-movement between patent applications and product introduction indicate that product innovation spikes one year after new patent applications. With an exception for product introduction at $k = -1$ in (c), we do not find a significant relation for k below zero. These dynamic specifications are useful for inferring the long-run elasticity of product introduction to patents, in contrast to the instantaneous elasticities discussed previously. Unlike the results with first-time patent filers, the results for the intensive margin of patenting are not persistent over time, which indicates that filing an extra patent application does not lead to an incremental product introduction in the long run. Under exogeneity assumptions in the context of lineal local projections (Jorda, 2005), the implicit long-run elasticity between patents and product introduction is the sum of the β_k coefficients from $k = 0$ onward. Our results point to an elasticity of about 0.1 for product introduction, and about 0.04 for quality-adjusted product introduction in the four years after the patent is filed. In Figure 7 of Online Appendix, we also show that other variables such as the stock of products or sales significantly increase after patents. We also confirm that our results are robust to considering the firm-level data from Match 1.

5 Product Innovation, Patents, and Competition: The Role of Firm Size

The previous sections show that patents are positively associated with product innovation in the market. These associations suggests that patents reflect important technological improvements that firms, on average, commercialize by introducing new products. However, in addition to their role in reflecting certain technological novelties, patents give firms the right to exclude others from using the same or similar technologies (Hall and Harhoff, 2012). These negative competitive spillovers from patenting have long been recognized in the literature (e.g. Lanjouw and Schankerman, 2001; Jaffe and Lerner, 2004; Bloom et al., 2013). Firms can use patents strategically to defend their technologies, reduce competitive pressure, and deter entry (Cohen et al., 2000). How do firms use these two roles of patenting? In their classic paper, Gilbert and Newbery (1982) suggest that monopolists have large incentives toward preemptive patenting. Likewise, Blundell et al. (1999) argue that patents held by market leaders may serve a largely preemptive purpose. This section sheds light on these issues by empirically evaluating how product innovation, patenting, and returns on patents vary systematically with a firm's market lead measured by relative sales in the market. We document the following empirical regularities:

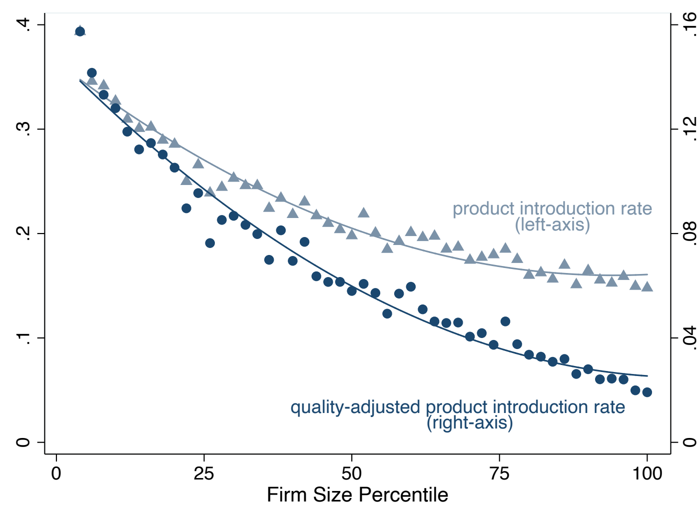
Fact 3: Larger firms have lower product innovation rates (quantity and quality), but file more patents for each new product.

Fact 4: Patenting by larger firms is strongly associated with an increase in revenue above and beyond the patents' effect on product innovation.

Fact 5: Patenting by larger firms is associated with a decline in product introduction by competing firms.

Product Innovation and Patenting by Firm Size. — We begin by exploring how product innovation rates vary with firm size. Figure 3 plots the average product introduction rate – the ratio of product introduction to a firm's stock of existing products – for firms across product categories. Larger firms (within product categories) have lower product innovation rates. On average, firms in the top sales quintile have annual innovation rates of about 16%, while firms in the bottom quintile have rates twice as large. Larger firms do not compensate for this decline in the rate of new product introduction with innovations of higher quality. On average, firms in the top sales quintile have quality-adjusted product introduction rates of 3%, while firms in the bottom sales quintile have rates four times larger. The fact that the quality-adjusted introduction rate declines more steeply than the simple product introduction rate indicates that, on average, new products introduced by larger firms represent only incremental improvements over existing products and are thus less novel. Figure 8 in the Online Appendix confirms similar patterns using alternative measures of innovation quality based on other novelty metrics and residual demand.

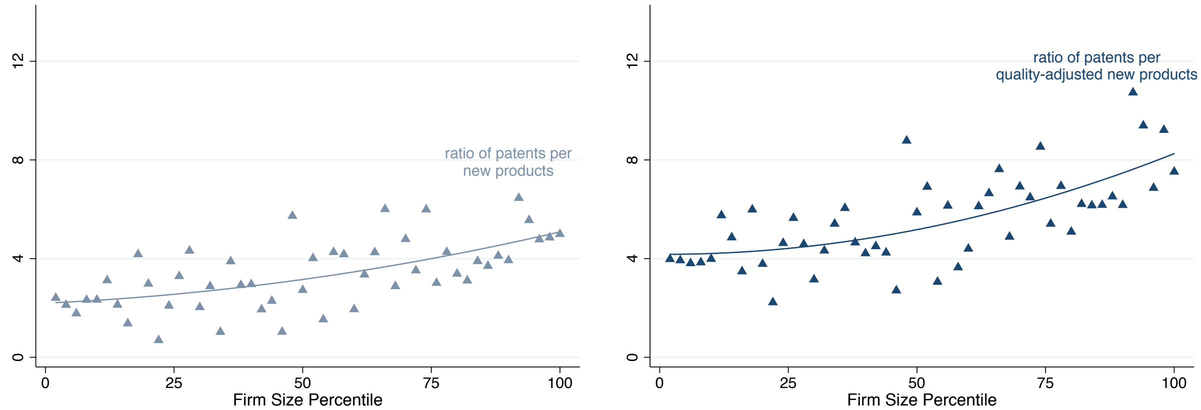
Figure 3: Product Innovation Rate by Firm Size



Notes: This figure plots the relationship between product innovation rates and the relative size of the firm, defined by the firm's sales. We use the firm × product category level data for the period 2007–2015, restricting the analysis to observations with sales above $1,000. For each firm × product category, we compute average sales, the average product innovation rate (new products divided by the total number of products sold), and the quality-adjusted product innovation rate (quality-adjusted new products divided by the total number of products sold). Within each product category, we assign firms to 50 bins for average sales and plot the average product innovation rate and the quality-adjusted product innovation rate for each bin. Each dot/triangle plots the averages after weighting each product category by its importance in the whole sector, as measured by the share of sales accounted for by the category.

This fact speaks to the well-known empirical regularity that larger firms grow slower (e.g. Haltiwanger et al., 2013). We show that this slow-down in growth is a reflection of larger firms' slow-down in the introduction of new and higher-quality products. Keeping this relationship between firm size and product innovation in mind, we now explore how patenting activity varies with firm size. Thanks to our data set that matches patents to products, we can simultaneously measure both product innovation and the associated patent applications by firms. As a results, Figure 4 shows that larger firms, on average, file more patents for each new product introduced.[24] Note that this higher intensity of patenting activity relative to the number of new products introduced is not explained by the possibility that larger firms introduce fewer but more novel products: as one can see, after we adjust for the quality of new products, small and large firms' innovation rates diverge even more.

Figure 4: Patents per New Products, by Size



Notes: This figure plots the relationship between the ratio of patent applications per new products and firm size as defined by sales. We use the firm × product category level data set for the period 2007–2015, restricting the analysis to observations with sales above \$1,000. For each firm × product category, we compute average sales, the average number of patent applications per new products, and the average number of patent applications per quality-adjusted new products. Within each product category, we assign firms to 50 bins of size based on average sales, and compute the average ratio of patents per new products and average ratio of patents per quality-adjusted new products for each bin. Each triangle plots the averages after weighting different product categories by their importance in the whole sector, as measured by their share of sales. The left figure plots the log ratio of patents per new products (× 1000), and the right figure plots the log ratio of patents per quality-adjusted new products (× 1000).

Motivated by this cross-sectional evidence, we now systematically explore how the elasticity of product innovation to the number of patents varies with firm size. We calculate this elasticity as in equation (3) for both product introduction N and quality-adjusted product introduction qN, after controlling for firm-category and category-time fixed effects. Table 6 reports the elasticities for firms in different size groups. In line with the results discussed above, we estimate an average elasticity of 0.038 (column "All"). The table shows that this elasticity varies substantially across the firm size distribution. Smaller firms in the bottom sales quintile have an elasticity twice as large as that of firms in the top sales quintile (0.059 versus 0.030). In the case of very large market leaders with sales in the top sales decile, we find only a non-significant positive association between patenting and product innovation. These market leaders have the highest rates of patenting, but the patents they file do not seem to translate into new products.

[24] Clearly, if we do not scale our measures of patenting down, the results are even starker: the unconditional probability of patenting and the total number of patents filed by large firms are much higher than they are for small firms (see Figure 9 of Online Appendix).

Table 6: **Product Innovation and Patenting: by Size**

	Log N (t)				Log qN (t)			
	All	Small	Large	Leaders	All	Small	Large	Leaders
Log P(t-1)	0.038***	0.059***	0.030**	0.023	0.019***	0.033***	0.017***	0.015*
	(0.007)	(0.018)	(0.013)	(0.017)	(0.003)	(0.007)	(0.006)	(0.009)
Observations	409,641	61,350	86,953	41,325	409,641	61,350	86,953	41,325
R-squared	0.692	0.463	0.742	0.777	0.623	0.407	0.686	0.732
Time-Category	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Category	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table shows regressions of the log number of new products (log qN) and of log quality-adjusted new products (log qN) in a firm $\times$ category over time as a function of the log number of patents. P is the number of patent applications for a firm $\times$ category $\times$ year. For each firm $\times$ product category, we define size based on the average sales over our sample period. The "All" column shows data for all sizes. "Small" column is restricted to the bottom size quintile. "Large" is restricted to the top size quintile. "Leaders" is restricted to the top size decile. The inverse hyperbolic sine transformation is used for logarithms. Standard errors robust against heteroskedasticity and serial correlation are reported in parentheses.

The results are similar for the quality-adjusted product introduction, shown in the rightward columns of the table. We further rule out two additional channels that could explain the weak association between patents and product innovation for large firms. First, we do not find evidence that patents held by larger firms are associated with product innovation with a longer delay. We test the dynamic specifications of equation (3) and do not find that the elasticity of patents to future product introduction is stronger for larger firms.[25] Second, we test whether patents held by larger firms are related more to process innovation rather than to product innovation (Cohen and Klepper, 1996). Building on Bena and Simintzi (2017), we construct proxies for product-related and process-related patents and we find no systematic pattern between a share of process patents filed by a firm and the firm's size (Figure 10 in Online Appendix).[26]

The Role of Patents for Competition by Firm Size.—Our finding that patents are more weakly associated with product innovation for larger firms is by itself important from the perspective of measuring innovation in the market: for market leaders, patents may be misleading proxies for the actual innovation that drives productivity growth. At the same time, this finding may be indicative of market leaders using patents as protective strategy tools. Firms may grow sales not just by introducing new and innovative products, but by accumulating patents that reduce competition and help the firm to capture market share

[25]Indeed, recall that in Section 4, we saw that the increase in product innovation after a new patent filing was short-term.

[26]In addition, if cost reductions due to process innovations are reflected in lower subsequent prices, we can test whether the price changes of larger firms react to patents more. However, we do not find such relationship in the data.

from its competitors.[27] To address this possibility, we take advantage of the information on product sales to understand the revenue premiums from patents for firms of various sizes. We use the following specification:

$$\Delta \log \text{Sales}_{ijt} = \psi \log P_{ijt-1} + \rho \log N_{ijt} + \theta_{ij} + \gamma_{jt} + \varepsilon_{ijt} \tag{5}$$

where the dependent variable is the logarithm of the change in sales at time t, $\log P_{ijt-1}$ is the total number of patent applications until time $t-1$, and $\log N_{ijt}$ is the number of new products introduced at t (we also use the quality-adjusted product introduction $\log qN_{ijt}$). Our coefficient of interest is ψ, which measures the elasticity of sales growth to patents after controlling for the effect that patents may have on sales through increased product innovation.

Table 7 shows the results for all firms and for firms grouped according to size. Overall, we find a positive significant relationship between patents and future growth in sales even after controlling for product innovation (columns "All"). This finding indicates that firms enjoy some additional value from holding a patent beyond the patent's value through new product offerings. Interestingly, this effect is highly heterogeneous across firm sizes. For firms in the bottom sales quintile (columns "Small"), there is no statistical association between patents and sales growth after we control for product introduction. However, for firms in the top quintile (columns "Large"), we find that an increase in total patent applications has a significant positive association with sales growth above and beyond its effect through product introduction. This effect is even larger (elasticity of 0.1) when we further restrict the analysis to firms in the top size decile (columns "Leader"). Note also that the direct impact of product innovation on sales growth (coefficients on $N(t)$ and $qN(t)$) decreases as firms increase in size. Hence, by splitting the sample into small and large firms, we learn that while both patents and new products are associated with increased future sales, the conditional impact of new products is more important for smaller firms, while the impact of patents is important for larger firms.

This additional revenue premium that larger firms draw from a patent may likely operate through patents' effect on competition: if patents discourage competitors from introducing new products, patent holders will benefit by serving a larger market. Thus, we now investi-

[27]The accumulation of patents often creates a web of overlapping intellectual rights which make it difficult for competitors to approach the market leader's technology domain and to leapfrog them. See, for example, Shapiro (2000) for the discussion of patent thickets.

Table 7: Patenting and Sales Growth

	Δ Log Sales (t)				Δ Log Sales (t)			
	All	Small	Large	Leaders	All	Small	Large	Leaders
Log P(t-1)	0.061***	-0.081	0.072***	0.099***	0.073***	-0.101	0.089***	0.111***
	(0.016)	(0.077)	(0.017)	(0.019)	(0.016)	(0.077)	(0.018)	(0.019)
Log N(t)	0.265***	0.316***	0.214***	0.160***				
	(0.003)	(0.011)	(0.003)	(0.004)				
Log qN(t)					0.406***	0.581***	0.310***	0.215***
					(0.006)	(0.029)	(0.007)	(0.007)
Observations	296,320	40,666	131,804	65,680	296,320	40,666	131,804	65,680
R-squared	0.291	0.377	0.294	0.296	0.275	0.368	0.277	0.281
Time-Category	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Category	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table presents estimated outcomes of changes in log sales at the firm × category level as a function of the log number of patents (P) and log number of new products (log qN) and log quality-adjusted new products (log qN), by size groups. We use the firm × product category data set for the period 2007–2015, restricting the analysis to observations with sales above $1,000. For each firm × product category, we define size based on average sales over the sample period. "All" column uses data for all sizes. "Small" column is restricted to the lowest size quintile. "Large" is restricted to the top size quintile. "Leaders" is restricted to the top size decile. The inverse hyperbolic sine transformation is used for logarithms. Standard errors robust against heteroskedasticity and serial correlation are reported in parentheses.

gate whether patents by market leaders are associated with declining product introduction on the part of their competitors, who we will refer to, for simplicity, as market followers. We identify the market leader in each category as the firm with the highest sales in that category and the followers as the remaining firms operating in that market.[28] Then for each year t and market j, we compute the total number of new products introduced by the leader N^{L}_{jt} and by its followers N^{F}_{jt} in t, and we compute the total numbers of patent applications introduced by the leader P^{L}_{jt} and by its followers P^{F}_{jt} until t. We evaluate how product innovation by followers responds to patenting (and product innovation) of the leaders as follows:

$$\log N^{\mathrm{F}}_{jt} = \eta^{\mathrm{F}} \log P^{\mathrm{L}}_{jt-1} + \alpha^{\mathrm{F}} \log N^{\mathrm{L}}_{jt-1} + \theta^{\mathrm{F}}_j + \gamma^{\mathrm{F}}_t + \varepsilon^{\mathrm{F}}_{jt}, \tag{6}$$

where η^{F} is our coefficient of interest, measuring the association of patents of leaders with the product introduction by followers. We control for $\ln N^{\mathrm{L}}_{jt-1}$ to ensure that the relationship between leaders' patents and followers' product introduction is not driven by possible direct interactions between the leader's and followers' product offerings (such as learning from new products on the market).[29] We also include both time- and product-category-fixed effects to

[28] To have a static firm-level measure, we define leaders as of 2006, which is the first year of our data. However, the results are not sensitive to a different choice, like using average sales over all years. Moreover, we consider alternative definitions of market leaders (e.g. top decile) and the results are robust.

[29] We also use quality-adjusted new products in all of these regressions, and the results are similar.

control for time trends and differences in the intensities of patenting and product innovation across product categories. Likewise, we estimate a symmetric regression that measures how leaders' innovation is affected by followers' patenting:

$$\log N^{\mathrm{L}}_{jt} = \eta^{\mathrm{L}} \log P^{\mathrm{F}}_{jt-1} + \alpha^{\mathrm{L}} \log N^{\mathrm{F}}_{jt-1} + \theta^{\mathrm{L}}_{j} + \gamma^{\mathrm{L}}_{t} + \varepsilon^{\mathrm{L}}_{jt} \tag{7}$$

These regressions help us test if the relation between patents of competitors and product introduction is affected by whether we focus on leaders or followers.

Table 8 presents the estimated coefficients. Column 1 shows that product introduction by followers is negatively correlated with the size of the leader's patent portfolio. This means that the product categories in which the leader is intensifying patenting over time are also the product categories in which followers are reducing the introduction of new products. In column 2, we also control for total sales of the market to account for potential shifts over time in the importance of different types of products. In turn, columns 3 and 4 show that product innovation by leaders is not related to the followers' patenting activity. Hence, while patents can be thought of as a protective tool used to hinder competition, our results indicate that this hypothesis is likely to apply when patents are in the hands of large market leaders.

Table 8: Patenting of Market Leaders and Followers

	Followers Log N^{F}			**Leaders** Log N^{L}	
	(1)	(2)		(3)	(4)
Leaders			**Followers**		
Log P^{L} (t-1)	-0.071***	-0.059***	Log P^{F} (t-1)	-0.015	-0.012
	(0.007)	(0.007)		(0.047)	(0.044)
Log N^{L} (t-1)	0.010***	0.005*	Log N^{F} (t-1)	0.215*	0.185*
	(0.002)	(0.002)		(0.112)	(0.094)
Observations	3,192	3,192	Observations	3,188	3,188
Category	Y	Y	Category	Y	Y
Time	Y	Y	Time	Y	Y
Controls	N	Y	Controls	N	Y

Notes: The table shows the relationship between the patents of leaders (followers) and the product introduction of followers (leaders). The leader is defined as the firm with the highest sales in a given category in 2006; the followers are defined as the rest of the firms in the categories. In columns (1) and (2), the dependent variable is the log number of products introduced by followers at time t, and the independent variables are the log number of patent applications by leaders until time $t-1$ and the log number of new products introduced by the leader at time $t-1$. In columns (3) and (4), the dependent variable is the log number of products introduced by leaders at time t, and the independent variables are the log number of patent applications filed by followers until time $t-1$ and the log number of new products introduced by the followers at time $t-1$. Columns (2) and (4) also control for total sales in the category-time. The inverse hyperbolic sine transformation is used for logarithms.

We view the results of this section as providing a consistent story for firms' use of their growth strategies. Firms may use both productive and protective strategies to grow sales. We document that as firms grow, they rely less on productive strategies that encourage the introduction of new and improved products in the market. At the same time, as firms grow, they rely increasingly on protective strategies such as patenting, which is associated with declines in competitors' product introduction and the resulting higher sales for larger firms.

6 Conceptual Framework

In this section, we offer a simple theoretical framework that illustrates the relationship between innovation, patenting, and creative destruction. Our goal is twofold. First, the framework is meant to build intuition about the incentives for patenting, consistent with empirical patterns documented in the previous sections. Second, we use the model to perform a simple back-of-the-envelope calculation of the private value of a patent, and we decompose this private value into its protective versus productive components. The productive component is the option value of implementing the patented idea into higher-quality products in the market, thereby increasing profits. The protective component is the value that the firm gains by impeding creative destruction.

Our framework builds on the quality-ladder model of innovation with creative destruction (e.g. Aghion and Howitt (1992)). In this model, product innovation takes the form of upgrades to the quality of products on the market. These innovations come from either the incumbent leader trying to prolong its lead or from market entrants aiming to become the new leaders. We consider an exercise in which the incumbent firm obtains an idea/blueprint for an innovation and makes a once-in-a-lifetime decision about commercializing that idea into a product and/or patenting it. If the firm decides to commercialize the idea, it will gain additional profits when it introduces higher-quality products to the market; and profits exhibit decreasing returns in quality. In turn, patenting the idea grants the firm extra protection against creative destruction from entrants. Both patenting and product innovation are costly activities. This simple framework can easily rationalize the main empirical facts we have documented. For the same idea/blueprint, smaller firms decide to commercialize ideas into better-quality products, mid-size firms will do both product innovation and patenting, and very large firms will file patents but not upgrade their products on the market. Hence, while larger firms hinder creative destruction more, they are less active in product innovation.

Production. — Consider a partial equilibrium framework that describes innovation in a single sector. There are M potential producers, and aggregate output is a combination of quality-weighted varieties:

$$Y = \frac{1}{1-\beta}\left[\sum_{m=1}^{M} q_m^{\frac{\alpha}{1-\beta}} y_m\right]^{1-\beta}, \quad 0 < \alpha < \beta < 1 \tag{8}$$

where y_m denotes the quantity and q_m is the quality level of variety m. This specification implies that products from different producers are perfect substitutes after adjusting for their qualities. The parameter α captures the consumer's satiation with respect to additional quality. Labor is the only factor of production. Producers use labor to produce output by hiring labor at the common wage rate of w. Output of variety m is then given by $y_m = l_m$, where l_m is labor used to produce variety m. We assume that the overhead cost of production ϵ must be paid before choosing prices and output. Since producers' marginal costs are the same and qualities are different, under Bertrand competition, this overhead cost allows the highest-quality firm to win the market and act as a monopolist.[30]

The monopolist maximizes profits by choosing the price of its product subject to demand from (8),[31] which delivers the following equilibrium objects for output (y), sales (R), and profits (Π), respectively (hereafter, we drop the subscript m):

$$y = \frac{1-\beta}{\beta}\frac{\pi}{w}q^{\gamma}, \qquad R = \frac{\pi}{\beta}q^{\gamma}, \qquad \Pi = \pi q^{\gamma}, \tag{9}$$

where $\pi \equiv \beta\left(\frac{1-\beta}{w}\right)^{\frac{1-\beta}{\beta}}$ and $\gamma \equiv \frac{\alpha}{\beta}$. Hence, firms with higher-quality products are larger, and generate higher sales and profits. Moreover, since $\gamma < 1$ because $0 < \alpha < \beta < 1$, the marginal quantity, sales, and profits decrease with quality.

Dynamic choices: product innovation and patenting.—Now consider the once-in-a-lifetime decision of product quality upgrade and patenting for an incumbent with quality q who exogenously obtains an idea of size λ.[32] If the firm decides to upgrade the quality of its product from q to $q + \lambda$ to generate higher profits, it also has to pay the costs of product development and commercialization c_m. In this case, $q + \lambda$ becomes the largest available quality in the economy. Simultaneously, the firm can also decide to patent the blueprint at a

[30]This assumption simplifies the setup. Alternatively, we would need to work with limit pricing, where the firm with highest quality would still capture the entire market, but the price would be determined by the price of the second highest quality producer.

[31]Price of Y is normalized to one.

[32]For simplicity, we assume that this is a one-time choice. Hence, the idea is either used or disappears afterwards. A more-complete approach with a dynamic decision of patenting and product innovation would bring similar tradeoffs at the expense of tracking the evolution of a firm's position both in the product and patent spaces.

cost c_p.[33] A patent grants the firm additional protection against being replaced by an entrant (more details below). If the firm decides to patent, even if the idea is not commercialized, the highest-available quality in the economy becomes $q + \lambda$ since, by patenting, the firm makes the idea "public". Note that the highest-available quality in the economy could be different than that commercialized by the firm and available to the consumers. In this sense, firms' activities in the product and patent spaces are separated. Product innovation does not necessarily imply patenting activity, and neither does introducing a patent necessarily imply product innovation.

Entry: creative destruction.—Incumbents can be replaced by entrants through creative destruction. The model includes an exogenous arrival rate of entrants at each instant p. Entrants build on "the shoulders of giants" and can replace incumbents by improving upon the highest-quality product available in the economy. The underlying assumption is that entrants can learn from products available in the market and from patents. Hence, "the shoulders of giants" correspond to $q+\lambda$ unless the incumbent neither upgrades nor patents, in which case the highest available quality is q. Entrants draw innovation of step size λ^e from a uniform distribution on $(0, 1)$. Patenting protects the quality level of incumbents $q + \lambda$ by creating a "wall" of height $\varepsilon > 0$ that entrants need to overcome to enter the market. The parameter ε captures the condition that entrants need to come up with an innovation sufficiently different from what has been patented before, which can depend on the strength of intellectual property protections as well as the scope of the patent. Given these assumptions, the probability of creative destruction is p if the incumbent does not patent and is $p(1 - \varepsilon)$ if the incumbent patents (Online Appendix D provides the proof). Notice that, in contrast to standard models of creative destruction, not all product quality improvements by entrants will find their way to the market. The separation between the patent space and the product space introduces the possibility that a better-quality product is not introduced to the market because it is blocked by existing patents.

Value functions and equilibrium.—Let us denote the value of a firm with existing product quality q that both upgrades quality and files patents as V^{11}, the value of only upgrading as V^{10}, the value of only patenting as V^{01}, and the value neither upgrading nor patenting as V^{00}. Then, the value function of the incumbent firm is

$$V(q) = \max\Big\{ V^{11}(q) - c_m - c_p,\ V^{10}(q) - c_m,\ V^{01}(q) - c_p\ ,\ V^{00}(q) \Big\}, \tag{10}$$

[33] We think of c_p as the combination of research, legal filing, and potential patent enforcement costs.

where

$$V^{11}(q) = \frac{\pi(q+\lambda)^\gamma}{r+p(1-\varepsilon)}, \qquad V^{10}(q) = \frac{\pi(q+\lambda)^\gamma}{r+p},$$
$$V^{01}(q) = \frac{\pi q^\gamma}{r+p(1-\varepsilon)}, \qquad V^{00}(q) = \frac{\pi q^\gamma}{r+p}.$$

Notice that incentives for product innovation decline as firm size increases, while the returns to patenting increase. Because marginal profits decrease as q increases, the incremental returns from product innovation decline with firm size, which is the same intuition that underlies the well-known *Arrow-replacement effect*.[34] This effect describes how larger firms and monopolists find it less profitable to replace themselves: innovations might cannibalize their own rents.

On the other hand, the returns to patenting increase with size as larger firms have a higher value to protect. In fact, we show in Online Appendix D that under mild conditions on costs, in this economy there exist cutoffs q^* and q^{**} such that a firm only upgrades when $q < q^*$, engages in both product innovation and patenting in the intermediate region $q^* < q < q^{**}$, and only patents when $q > q^{**}$.[35] As a result, the model delivers an equilibrium that rationalizes the main empirical patterns uncovered in the previous sections:

Implication 1: Many firms (below cutoff q^*) develop product innovations without patenting.

Implication 2: On average, patenting and product innovation are positively correlated.

Implication 3: Larger firms develop relatively fewer product upgrades, but file more patents.

The model also speaks to our empirical facts on the relationship between patenting, firm size, and competition. By construction, patents in the model reduce creative destruction. At the same time, larger firms rely on patenting more. Hence, larger firms also face lower risks of creative destruction.

Implication 4: Patents deter future entry by competitors; larger firms deter entry more.

[34] We provide an empirical estimate of γ and confirm that it is lower than one. However, instead of these decreasing returns that generate a declining relationship between size and innovation, one can generate it through other ways, such as by implementing weaker scalability of R&D technology with increasing size as in Akcigit and Kerr (2018) or an innovation-advertising tradeoff as in Cavenaile and Roldan (2019).

[35] The required conditions on the costs c_m and c_p ensure that at least one firm finds it profitable to introduce a product and at least one firm finds it too costly to patent. This sharp cutoff rule clearly hinges on the simplifying assumption of fixed innovation and patenting costs that buys us tractability. However, the main features of the model – the reduction in product innovation incentives as firms grow and the increase in protection incentives through patenting – can be easily generalized with a more realistic cost structure.

The value of a patent.—If a firm in our model were to sell its patent, what would be its price? Patents embed both productive and protective values. Both values come from the underlying technological innovation contained in the patent. Productive value comes from the option value of commercializing an innovation, while protective value comes from the ability of a patent to protect firms' market lead from competitors. We define the (private) value of a patent as the revenue premium that a patented innovation provides:

$$\begin{aligned} Patent\ Value &= V^{11} - V^{00} \\ &= \underbrace{\frac{\pi(q+\lambda)^{\gamma}}{r+p(1-\varepsilon)} - \frac{\pi(q+\lambda)^{\gamma}}{r+p}}_{Protective} + \underbrace{\frac{\pi(q+\lambda)^{\gamma}}{r+p} - \frac{\pi q^{\gamma}}{r+p}}_{Productive} \end{aligned} \tag{11}$$

The total patent value can be decomposed into productive and protective components by adding and subtracting V^{10}.[36] Productive value is the revenue premium from commercializing a product of upgraded quality if we hold creative destruction fixed. This value from product innovation declines as firms grow, since the same amount of innovation brings marginally lower returns. In contrast, protective value, which is the revenue premium from lower creative destruction holding the technology of a firm fixed, increases as firms grow: the use of patent protection is more relevant as the value of the firm increases. Hence, we formulate our final implication of the model:

Implication 5: The revenue premium from patents comes both from product upgrades and protection. The latter becomes more important as firms grow.

We now set the parameters of the model to estimate the average value of a patent for firms in our data. To estimate (11), we need to assign values to π, λ, γ, p, and ε. First, we normalize the average quality within each product category in our data to one. Notice that we do not observe profits, but given (9), we know that sales are proportional to profits such that $\Pi = \frac{\mu-1}{\mu} \times R$, where μ is the markup. The profit of an average firm is then $\frac{\mu-1}{\mu}$ multiplied by the sales of the average firm, which we take to be equal to the average yearly sales of all firms across all product categories (1.36 million 2015 USD). We take $\mu = 1.21$ drawing upon Barkai (2017)'s average estimate of markups in the U.S. economy in 2014. Assigning values to p and $p(1-\varepsilon)$ involves the following considerations. In the model, if firms do not innovate

[36] An alternative decomposition would be $\underbrace{\frac{\pi(q+\lambda)^{\gamma}}{r+p(1-\varepsilon)} - \frac{\pi q^{\gamma}}{r+p(1-\varepsilon)}}_{Productive} + \underbrace{\frac{\pi q^{\gamma}}{r+p(1-\varepsilon)} - \frac{\pi q^{\gamma}}{r+p}}_{Protective}$.

they face a creative destruction rate that leads to the decline of their expected sales. Hence, we infer the values of p from firms' growth in sales when they do not introduce new products in a given year. In our data, the median firm that does not hold any patents suffers a loss in sales when it does not introduce new products: log sales change is equal to -10.3%. This decline in sales is attenuated if a firm holds a patent, thus giving an estimate for ε. The implied values for creative destruction are $p = 0.098$ and $p(1-\varepsilon) = 0.095$.

Lastly, we jointly estimate λ and γ. Intuitively, λ determines the average growth when the firm innovates, and γ affects how this growth varies with firm size. Specifically, the model implies the following relationship between firm growth and relative size, conditional on product innovation:

$$\Delta \ln R_t = \gamma ln\Big(1 + \lambda\Big[\frac{R_{t-1}}{\bar{R}_{t-1}}\Big]^{-\frac{1}{\gamma}}\Big)$$

We estimate this relationship with a non-linear least squares regression applied to the sample of firms who introduce new products in that year. We define the relative size of firms as sales divided by the average sales of firms in that year and product category. The resulting estimates are $\gamma = 0.899$ (s.e. 0.364) and $\lambda = 0.024$ (s.e. 0.008).

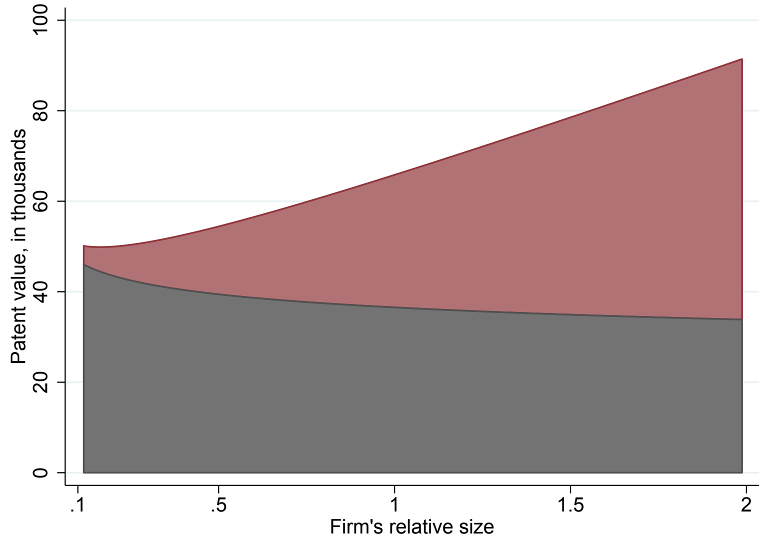
Figure 5 plots the value of a patent against the relative size of firms. The red shaded area depicts the contribution of the protective component of the patent's value, and the gray area depicts the contribution of the value's productive component. For the average firm the value of a patents is around \$65,000.[37] The estimated value increases drastically as firm size increases, mainly due to the contribution of the protective value. For example, for firms ten times smaller than the average firm, only 9% of the value comes from the protective component, while for firms that are twice as large as the average firm, the protective component accounts for 60% of a patent's value.

Our methodology for estimating the value of the patent differs greatly from those used previously in the literature as it relies on the structure of our model, the matched data we constructed between patents and products, and the realized sales of products observed in the data.[38] Nonetheless, we find that our estimates are well in the range of other estimates reported in the literature. Using patent renewal information to infer the private value of

[37] Notice that our calculations do not include sales from the stores not covered by Nielsen. To get at the nationwide sales, we can roughly scale our sales twice (see Online Appendix A.1 for details).

[38] The literature has used various methods to estimate the value of a patent: direct survey questions, inference from observed patent renewals by firms, stock market responses to patent news, as well as direct estimates from patent sales samples. For a comprehensive review, see Hall and Harhoff (2012).

Figure 5: Estimated patent value



U.S. patents issued in 1991, Bessen (2008) estimates a patent's mean value to be \$121,000 (median \$11,000). Interestingly, consistent with our results, Bessen (2008) also finds that the value of patents held by smaller firms is lower, while litigated patents are more valuable. Serrano (2010) estimates the average private value of holding a patent to be \$90,799 (median \$19,184). Using data from a large non-practicing entity, which presumably holds mostly valuable patents, Abrams et al. (2013) find that the mean value of a patent is \$235,723 (median \$47,955).

The advantage of our methodology is that it allows us to decompose the patent value into its two inherent components – productive and protective. The decomposition uncovers the dual role of patenting and the role each component plays for firms of different sizes.

7 Conclusion

Using textual analysis of patent documents and product descriptions, we construct a new patent-to-products data set to study the relationship between patents and product innovation. We find that more than half of the product innovation is not associated with patents. Nonetheless, patent filing is positively associated with subsequent product innovation by firms, on average. We document substantial heterogeneity in this relationship. Patents filed by larger firms reflect less the actual product innovation than other patents do. Instead, we find strong evidence suggesting that the main role of patents for market leaders is to deter product innovation of competitors and protect sales of their existing products. Hence, our results indicate that although on average patents capture product innovation in the market, because the relationship between patents and innovation changes with firm size, patent-based measures distort the differences in actual innovation between firms of different sizes.

Using a simple theoretical framework, we show that for the same patented idea, a larger firm can reap a greater monetary return than a smaller firm can. However, for these large firms, more of this return is derived from the patent's ability to hinder competition than is derived from commercializing a new product using the patented idea. We argue that understanding the contribution of the productive and protective components of patenting and how they vary by firm size has important implications for our understanding of growth, innovation, and intellectual property policy. Specifically, policymakers should pay more attention to the state-dependence of patent and R&D-based policies; these policies should acknowledge that firms' incentives to use intellectual property vary greatly with firm size and market leadership.

In their comprehensive analysis of patent reform in the U.S., Jaffe and Lerner (2004) argue that the seemingly innocent changes in patent policies in the early 1980s significantly affected firms' incentives toward strategic patent filing. In line with this, the survey results published by Cohen et al. (2014) suggest that relative to the early 1980s, large firms now rely somewhat more heavily on patents to protect their sales from competitors. A potential avenue for future research is to understand how large firms' increasing reliance on protective patenting has contributed to the recent trends of increasing dominance of large firms and declining business dynamism on the one hand (Decker et al., 2016; Akcigit and Ates, 2019) and rising top income inequality on the other (Jones and Kim, 2018).

References

Abrams, David S, Ufuk Akcigit, and Jillian Grennan, "Patent Value and Citations: Creative Destruction or Strategic Disruption?," NBER Working Paper 19647 November 2013.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–51.

Aizawa, Akiko, "An information theoretic perspective of TF-IDF measures," *Information Processing Management*, 01 2003, *39*, 45–65.

Akcigit, Ufuk and Sina T Ates, "What Happened to U.S. Business Dynamism?," Working Paper 25756, National Bureau of Economic Research April 2019.

_ **and William R. Kerr**, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374 – 1443.

— , **Douglas Hanley, and Stefanie Stantcheva**, "Optimal Taxation and RD Policies," Working Paper 22908, National Bureau of Economic Research December 2016.

— , **Murat Celik, and Jeremy Greenwood**, "Buy, Keep, or Sell: Economic Growth and the Market for Ideas," *Econometrica*, 2016, *84*, 943–984.

— , **Salome Baslandze, and Francesca Lotti**, "Connecting to Power: Political Connections, Innovation, and Firm Dynamics," NBER Working Paper 25136 October 2018.

Alexopoulos, Michelle, "Read All about It!! What Happens Following a Technology Shock?," *American Economic Review*, June 2011, *101* (4), 1144–79.

Argente, David and Chen Yeh, "Product's Life-Cycle, Learning, and Monetary Shocks," 2017.

— , **Munseob Lee, and Sara Moreira**, "The Life Cycle of Products: Evidence and Implications," 2020.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research May 2017.

Balasubramanian, Natarajan and Jagadeesh Sivadasan, "What Happens When Firms Patent? New Evidence from U.S. Economic Census Data," *The Review of Economics and Statistics*, 2011, *93* (1), 126–146.

Barkai, Simcha, "Declining Labor and Capital Shares," *LBS working paper*, 2017.

Bena, Jan and Elena Simintzi, "Globalization of work and innovation: Evidence from doing business in china," Technical Report, Discussion paper 2017.

Bessen, James, "The value of U.S. patents by owner and patent characteristics," *Research Policy*, 2008, *37* (5), 932–945.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," NBER Working Paper 23782 September 2017.

— , **Mark Schankerman, and John Van Reenen**, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4).

Blundell, Richard, Rachel Griffith, and John Van Reenen, "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *The Review of Economic Studies*, 07 1999, *66* (3), 529–554.

Bresnahan, Timothy F and Robert J Gordon, "Introduction to" The Economics of New Goods"," in "The Economics of New Goods," University of Chicago Press, 1996, pp. 1–26.

Cavenaile, Laurent and Pau Roldan, "Advertising, innovation and economic growth," Working Papers 1902, Banco de Espana;Working Papers Homepage February 2019.

Cockburn, Iain and Megan J. MacGarvie, "Entry and Patenting in the Software Industry," *Management Science*, 05 2011, *57*, 915–933.

Cohen, Lauren, Umit Gurun, and Scott Duke Kominers, "Patent Trolls: Evidence from Targeted Firms," Working Paper 20322, National Bureau of Economic Research August 2014.

Cohen, Wesley M. and Steven Klepper, "Firm Size and the Nature of Innovation within Industries: The Case of Process and Product RD," *The Review of Economics and Statistics*, 1996, *78* (2), 232–243.

Cohen, Wesley M, Richard R Nelson, and John P Walsh, "Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)," Technical Report, National Bureau of Economic Research 2000.

Crouzet, Nicolas and Janice C Eberly, "Understanding Weak Capital Investment: the Role of Market Concentration and Intangibles," NBER Working Paper 25869 May 2019.

Cunningham, Colleen, Song Ma, and Florian Ederer, "Killer Acquisitions," *Academy of Management Proceedings*, 2018, *2018* (1), 11001.

Dechezlepretre, Antoine, Elias Einio, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do tax incentives for research increase firm innovation? An RD Design for RD," Working Papers 73, VATT Institute for Economic Research 2016.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda, "Declining Business Dynamism: What We Know and the Way Forward," *American Economic Review*, May 2016, *106* (5), 203–07.

Garcia-Macia, Daniel, ChangâĂŘTai Hsieh, and Peter J. Klenow, "How Destructive is Innovation?," *Econometrica*, 2019, *87* (5), 1507–1541.

Gilbert, Richard J. and David M. G. Newbery, "Preemptive Patenting and the Persistence of Monopoly," *The American Economic Review*, 1982, *72* (3), 514–526.

Goolsbee, Austan D and Peter J Klenow, "Internet rising, prices falling: Measuring inflation in a world of e-commerce," in "Aea papers and proceedings," Vol. 108 2018, pp. 488–92.

Gordon, Robert J., "The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War.," *Journal of Regional Science*, 2016.

Graham, Stuart JH, Cheryl Grim, Tariqul Islam, Alan C Marco, and Javier Miranda, "Business dynamics of innovating firms: Linking US patents with administrative data on workers and firms," *Journal of Economics & Management Strategy*, 2018, *27* (3), 372–402.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, Spring 1979, *10* (1), 92–116.

— , "Market value, RD, and patents," *Economics Letters*, 1981, *7* (2), 183 – 187.

Gutiérrez, Germán and Thomas Philippon, "Declining Competition and Investment in the U.S.," Working Paper 23583, National Bureau of Economic Research July 2017.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," NBER Working Paper 8498 2001.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *RAND Journal of Economics*, Spring 2005, *36* (1), 16–38.

— **and Dietmar Harhoff**, "Recent Research on the Economics of Patents," *Annual Review of Economics*, 2012, *4* (1), 541–565.

Haltiwanger, John, Ron Jarmin, and Javier Miranda, "Who Creates Jobs? Small versus Large versus Young," *The Review of Economics and Statistics*, 2013, *95* (2), 347–361.

Heller, Michael A. and Rebecca S. Eisenberg, "Can Patents Deter Innovation? The Anticommons in Biomedical Research," *Science*, 1998, *280* (5364), 698–701.

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

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Jaffe, Adam B. and Josh Lerner, "Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It.," *Princeton University Press, Princeton, NJ*, 2004.

Jones, Charles I. and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," *Journal of Political Economy*, 2018, *126* (5), 1785–1826.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," NBER Working Papers 25266 November 2018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth*," *The Quarterly Journal of Economics*, 03 2017, *132* (2), 665–712.

Lampe, Ryan and Petra Moser, "Patent Pools, Competition, and Innovation–Evidence from 20 US Industries under the New Deal," *The Journal of Law, Economics, and Organization*, 08 2015, *32* (1), 1–36.

Lanjouw, Jean and Mark Schankerman, "Characteristics of Patent Litigation: A Window on Competition," *RAND Journal of Economics*, 2001, *32* (1), 129–51.

Lloyd, Stuart, "Least squares quantization in PCM," *IEEE transactions on information theory*, 1982, *28* (2), 129–137.

Loecker, Jan De and Jan Eeckhout, "The Rise of Market Power and the Macroeconomic Implications," Working Paper 23687, National Bureau of Economic Research August 2017.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze, *Introduction to Information Retrieval*, New York, NY, USA: Cambridge University Press, 2008.

Moser, Petra, "Innovation without Patents: Evidence from World's Fairs," *Journal of Law and Economics*, 2012, *55* (1), 43 – 74.

Öscar Jorda, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, March 2005, *95* (1), 161–182.

Pakes, Ariel, "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 1986, *54* (4), 755–784.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, October 1990, *98* (5), 71–102.

Sampat, Bhaven and Heidi L. Williams, "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome," *American Economic Review*, 2019, *109* (1), 203–36.

Schankerman, Mark and Ariel Pakes, "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period," *The Economic Journal*, 1986, *96* (384), 1052–1076.

Serrano, Carlos J., "The dynamics of the transfer and renewal of patents," *The RAND Journal of Economics*, 2010, *41* (4), 686–708.

Shapiro, Carl, "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting," *Innovation Policy and the Economy*, 2000, *1*, 119–150.

Trajtenberg, Manuel, "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *The RAND Journal of Economics*, 1990, *21* (1), 172–187.

Williams, Heidi L., "Intellectual Property Rights and Innovation: Evidence from the Human Genome," *Journal of Political Economy*, 2013, *121* (1), 1–27.

Younge, Kenneth and Jeffrey M. Kuhn, "Patent-to-Patent Similarity: A Vector Space Model," in "in" 2016.