Patents to Products: Innovation, Product Creation, and Firm Growth*

David Argente[†]
FRB of Minneapolis

Salomé Baslandze[‡]
EIEF & CEPR

Sara Moreira[§]
Kellogg

Douglas Hanley[¶]
U. of Pittsburg

First version: March 2019
PRELIMINARY

Abstract

How do patents relate to product innovation? To study this question, we construct a new patent-to-product data set combining data from the U.S. Patent and Trademark Office with the RMS Nielsen data covering the universe of firms in the consumer product goods sector. Using the textual analysis of patent documents together with product descriptions from Nielsen data, we link specific patents to finely defined product categories within firms and time periods. Our findings indicate that there is a substantial amount of product innovation that comes from firms that have never patented. Nevertheless, for patenting firms, standard patent-based measures of innovation are strongly related to product innovation measured both by the quantity and the quality of new products. We find that market leaders use patents differently than followers. In particular, patents of large firms have a weaker association with the quality and quantity of product innovations. Nevertheless, larger firms are able to generate higher revenues from those patents. This is largely because their patents are better at deterring competition than those from smaller firms. We decompose the value of a patent considering these effects in our theoretical framework. We show that the value of a patent increases as firms become market leaders. This increase is mostly driven by an increasing value derived from protective patenting as opposed to productive patenting.

JEL Classification Numbers: O3, O4

Keywords: Innovation, patents, productivity, patent value, protective patents.

*We are grateful to Ufuk Akcigit, Fernando Alvarez, Ariel Burnstein, Benjamin F. Jones, and seminar participants at Northwestern, EPFL, and SED (Mexico). Empirical results are calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen.

[†]Email: david.argenteamaya@mpls.frb.org. Address: 90 Hennepin Avenue, Minneapolis, MN 55401.

[‡]Email: salome.baslandze@eief.it. Address: Via Sallustiana, 62, 00187 Roma.

[§]Email: sara.moreira@kellogg.northwestern.edu. Address: 2211 Campus Drive, Evanston, IL 60208.

[¶]Email: doughanley@pitt.edu. Address: 230 South Bouquet Street Pittsburgh, PA 15260.

I Introduction

For decades, economists have seen innovation as a key component for productivity gains and economic growth. The introduction of better products to the market is often referred to as the key mechanism through which innovation translates into economic growth (Romer (1990), Aghion and Howitt (1992)). However, the lack of detailed data on the quantity and quality of product introductions has led researchers to use several other metrics to value innovation. Patents have emerged as the primary measure to value innovation, especially since systematic data containing their characteristics became easily available. But how do patents relate to product innovation and how much innovation do we miss out by just looking at patents? What is the value of a patent, and how much of it comes from the pure technological novelty embedded in the patents (*productive component*) as opposed to the protective and strategic role of a patent (*protective component*)? Do larger firms have an incentive to rely more on strategic role of patenting when slowing down on innovation? The main difficulty in answering these questions has been the lack of availability of a large-scale data linking the universe of patenting activities with universe of product innovations. In this paper, we construct such dataset for the Consumer Package Goods (CPG) sector by linking detailed barcode level data from Nielsen RMS with the universe of patents from the United States Patent Office (USPTO).

To do so, we first link the firms in USPTO with those present in the Nielsen RMS dataset. Next, using the textual analysis of patent documents together with product descriptions from Nielsen expanded with additional information obtained from Wikipedia, we classify specific patents into finely defined product categories. With this algorithm, we are able to build a dataset at the firm-product category-year level with information on product characteristics on the one hands and patent characteristics that protect them, on the other hand.

The advantages of this data match are multiple. First, we are able to measure innovation directly using both the quantity of innovation – new products introduced by a firm – and the quality of innovation which we identify based on detailed attributes of new products introduced to the market. This allows us to evaluate standard patent-based measures of innovation against "true" product innovation measures. Second, we have detailed data on product sales and prices. Hence, we are able to link specific patents to the revenue generated from the products associated with them. This allows us to calculate the value of a patent, explore its components and understand how it differs for firms of different sizes. Third, our match at the firm $\times$ product category $\times$ year level is particularly useful to study large firms that have products in multiple categories, allowing us to study the relationship between products and patents using variation within the firm in a year and controlling for product-level aggregate trends.

The first part of the paper is a measurement exercise showing how well the patent-based metrics proxy for product innovation in the CPG sector. In particular, we document relationship between patenting at the firm and firm $\times$ category level and product innovation – measured by the total number of new products and their quality.

We observe that patent-based metrics do indeed go hand in hand with true product innovations by firms. Firms introduce more and higher-quality new products exactly at the time of the patent application. Importantly, this observation is only true for patents that are eventually granted and patents that receive a lot of forward citations, which suggests that the standard quality metrics of patents are strongly associated with product innovation rates. We also find, however, that more than half of the innovations in the sector comes from firms that have never patented. This indicates that the amount of innovation we miss by only studying the behavior of patents may be substantial.

Hence, our first fact is:

Fact 1: *Patenting is strongly positively associated with product innovation – measured by the total number of new products and their quality.*

In the second part of our paper, we try to understand both theoretically and empirically firms' activities in the product and patent space as a function of their market position. Do firms change their growth strategies – acquiring leadership through product innovation vs maintaining leadership by protecting it – as they become market leaders?

To help to think about this question systematically, we first construct a simple theoretical model of firm dynamics with firm's innovation and patenting decisions. In the model, firms can exert *productive* innovation effort, innovate, patent, and push the technology frontier. In addition, firms can also spend resources on *protective* patents that do not push the technology frontier, but rather help firms to protect their technology by limiting competition and creative destruction.

The model predicts that as firms grow and become market leaders, they shift their innovation more towards protective strategies. Another implication of our model is that average patent of a market leader is associated less with product innovation and more with competitors' deterrence. As a result, the value of a patent increases with firm size. After further decomposing this value into *productive* and *protective* components, we see that patent value of market leaders is driven to a larger extent by the protective component. We test predictions from our model in the data.

Consistent with the model, we document that larger firms have lower product innovation rates – both in terms of number and quality of new products that they introduce to the market.

While firms are slowing down on product innovation, they increase patenting. In fact, we show that larger firms have higher patents per new products, and this does not translate into higher innovation by them. Nevertheless, they generate higher revenues from patented products than smaller firms. These results suggest that the monetary value of a patent is larger for market leaders but this value is not coming from the quality of the new products launched. We see that market leaders are more likely to launch very successful products in terms of revenue but the innovative content of these products is lower than the products of followers.

The intuition from our model suggests that the increase in the value of a patent for larger firms should mainly come from the protective part of patenting – the ability to deter entry. We measure the deterrence component using information on the product introduction of other firms within the same product category, and we explore whether protected innovations by market leaders deter competition more. Thus, empirically, the value of the deterrence component comes from the additional revenue generated by new products that results from restricting competition. We find that this component is substantial both because the product innovation rate of the market significantly declines if market leaders patent new products and also because larger firms are more likely to launch very successful products – those that last longer in the market and generate more revenue. As a result, the value of a patent for market leaders mainly comes from the deterrence component, whereas for followers the quality component is more important.

We summarize these results here:

Fact 2: *Larger firms have lower product innovation rates.*

Fact 3: *Larger firms have higher patents per new product, and this does not translate into higher-quality products.*

Fact 4: *Patents of larger firms deter competitors from new product introduction more than patents of the smaller firms.*

The paper is related to the recent work that attempts to rationalize the increasing research and patenting trends with the limited effects on overall innovation and growth (J.Gordon (2016), Bloom et al. (2017)). We show that there may be yet another explanation for this trend. Our results indicate that large incumbents may rely on protective and defensive strategies instead of innovation to maintain their market leadership. In the environment where economic activities reallocate towards high-market power firms (De Loecker and Eeckhout (2017), Autor et al. (2017)), incentives of large incumbents to direct their innovation effort towards productive rather than protective strategies may be very limited.

The remainder of this paper is organized as follows. In section II, we present the description of the different datasets we use, the procedure to match the datasets along with our text analysis algorithm, and presents the summary statistics. In section III, we present several event studies to show the relationship between patenting and new product introductions. Section V presents results on the relationship between patenting, innovation and deterrence over firm's size/market rank; and IV presents a simple firm dynamics model with firm's innovation and patenting decisions to provide the theoretical mechanism at play in our value estimation. The appendix provides additional empirical findings, and extensions of the model.

Related Literature

Patents represent the most common metric for innovation in the literature. Firms with larger number of patent filings are considered to be more innovative. Patent filings are standardized, comparable across firms and countries, and they contain rich information on characteristic of invention – hence, providing the most systematic and large-scale dataset proxying innovation. However, in the absence of independent measures of innovation (for example, as in Alexopoulos (2011) or Moser (2012) considering alternative measures, but for aggregate innovation), it is hard to evaluate how good the patents are as proxy for innovation.[1]

[1] As Boldrin and Levine (2013) put it, "there is no empirical evidence that they [patents] serve to increase innovation and productivity, unless productivity is identified with the number of patents awarded – which, as evidence shows, has no correlation with measured productivity".

There are at least two main challenges with this proxy. First, not all innovations are patented (Moser, 2012); and second, patents by their nature – in addition to their role to reflect certain technological novelties – give the right to exclude others from using same or similar technology (Hall and Harhoff, 2012). These negative competitive spillovers from patenting has been long recognized in the literature (Lanjouw (1998),Lanjouw and Schankerman (2001), Jaffe AB (2004), Bloom et al. (2013)). Since most of innovation is cumulative in nature – this feature of the patents may prevent others to "stand on the shoulders of the giants" (Scotchmer (1991), Furman and Stern (2011)) and may induce large social costs. For example, various studies have studied the effect of patenting on follow-on innovation (see Williams (2013), Heller and Eisenberg (1998),Sampat and Williams (2019) for the papers on biomedical research and Cockburn and J. MacGarvie (2011) for software industry, and Lampe and Moser (2015) for more general discussion).

In this paper, the unique match of patents to products in the CPG sector allows us to partly address these challenges. First, by having an independent and direct measure of product innovation from the product-level data, we examine how much product innovation is associated with patent introductions. Second, we try to decompose the value of a patent that comes from its technological component – patent representing novelty, and its strategic component – the ability to protect market from competitors. Importantly, we provide the evidence that these different components that lead to the revenue premium to the firm crucially depend on firm's current market position. Patents by larger incumbents contain more strategic component, while patents of followers contain more productive component. Hence, we contribute to the literature that tries to estimate patent value. For example, earlier studies (Hall and Harhoff (2012)) tried to infer monetary value of a patent using patent renewals (Schankerman and Pakes, 1986), the direct survey questions (Gambardella et al. (n.d.), Harhoff et al. (2003)), market value estimations (Hall et al. (2005), Toivanen et al. (2002), Seru et al. (2017)), and patent sales (Abrams et al., 2013).

In our case, we *directly* observe revenues from products linked to patents, and we directly observe the behavior of competitors – how they react to patents through product introduction.

Our results indicate that patents may help market leaders to further strengthen their position without necessarily introducing a lot of high-quality subsequent products. In their classical paper, Gilbert and Newbery (1982) propose the possibility of preemptive patenting that may contribute to the persistence of monopoly. Empirically, Blundell et al. (1999) show that the impact of patenting on market value is larger for market leaders arguing that partly it comes from pre-emptive patenting. Indeed, large firms may rely on various protective or defensive strategies – like firm acquisitions (Cunningham et al., 2018) or connections with politicians (Akcigit et al. (2018)) – as they slow down on innovation (Cavenaile and Roldan (2019),Akcigit and Kerr (2018)). In this paper, we look at changes in firm incentives from productive patenting to protective patenting, hence contributing to our understanding of firm's growth strategy as a function of their market lead.

By better understanding these growth strategies we can potentially understand the link increasing market concentration, markups and dominance of large firms on one hand (Rossi-Hansberg et al. (2018), Gutierrez and Philippon (2017), De Loecker and Eeckhout (2018), De Loecker and Eeckhout (2017), Autor et al. (2017)) and declining growth and business dynamism (Decker et al., 2016) on the other.

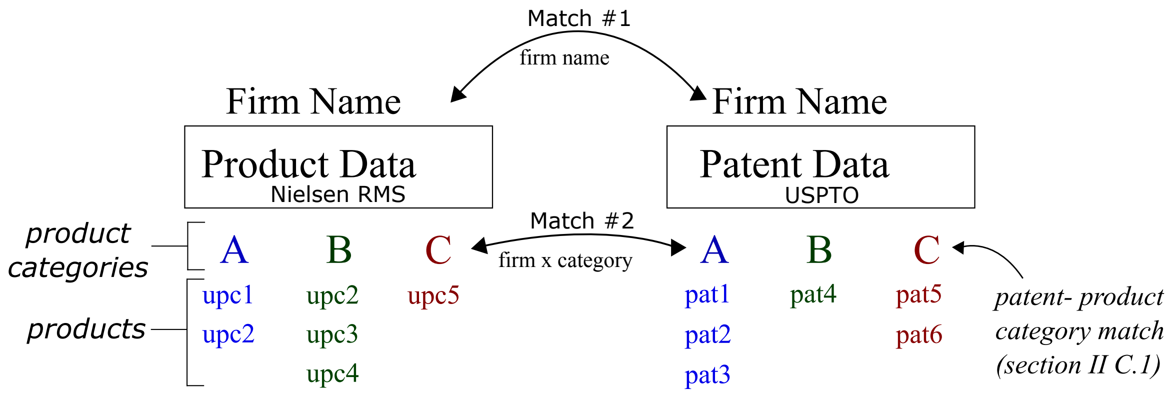
II Product and Patents Dataset and Measurement of Innovation

In this paper, we study the relationship between new products and patents. There are two important challenges associated with this: identification of product portfolio of firms and the match of products and patents.

We tackle the first challenge by making use of scanner data and the richness of the data on the product portfolio available for firms in the Consumer Package Good (CPG) sector. Figure 1 schematically represents our data construction. For each firm in the CPG sector, we have a full portfolio of various *products* – UPC's – that we classify into different *product categories*. These product categories are constructed in Section B by grouping similar products using textual analysis of product descriptions. At the firm and firm $\times$ product category levels, we then define our measures of *product innovation* based on quantity and quality of new products, as defined in Section B.

We address the second challenge by matching all United States Patent and Trademark Office (USPTO) patents of CPG firms to their product portfolios. Our *Match 1 – at the firm level* – simply matches patents to products based on firm identifiers in product and patent datasets. Our preferred *Match 2 – at the firm* $\times$ *product category level* – goes a step beyond by classifying firms' patents into product categories using textual analysis of patent documents.

Figure 1: Product and Patents Dataset



This is a crucial step in our analysis and is described in Sections C and C.1. As a result, at the firm and firm × product category levels, on the one hand, we have product innovation measures from CPG data and, on the other hand, various patenting measures, such as patent and quality-adjusted patent counts.

This section describes data and these procedures in details. We first describe our product-level data from Nielsen Retail Measurement Services (RMS) in Section A and next introduce definition of product innovation in Section B. In Section C, we discuss our patent data construction and the algorithm that matches patents to products. Finally, Section D discusses our two-way match of patents to products: at the firm level and at the firm × product category level.

A Product Data

We rely primarily on the Nielsen Retail Measurement Services (RMS) scanner data set that is provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The data is generated by point-of-sale systems in retail stores. Each individual store reports weekly sales and quantities of every UPC code that had any sales volume during that week.

The main advantage of this dataset is its size and coverage. Overall, the RMS consists of more than 100 billion unique observations at the week × store × UPC level. Our sample period covers the period 2006-2015. Per year, the data set comprises around 12 billion transactions worth, on average, \$220 billion of dollars. Over our sample period the total sales across all retail establishments are worth approximately \$2 trillion and represent 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, 2% in convenience stores, and 1% in liquor stores. A key distinctive feature of this database is that the collection points include more than 40,000 distinct stores from around 90 retail chains, across 371 MSAs and 2,500 counties. As a result, the data provide good coverage of the universe of products and of the full portfolio of firms in this sector. In comparison to other scanner data sets collected at the store level, the RMS covers a much wider range of products and stores. In comparison to scanner data sets collected at the household level, the RMS also has a wider range of products because it reflects the universe of transactions for the categories it covers as opposed to the purchases of a sample of households.

The original data consist of more than one million distinct products identified by UPC. The data is organized into 1,070 detailed product modules that are aggregated into 114 product groups that are then grouped into 10 major departments. The ten major departments are: Health and Beauty aids, Dry Grocery (e.g., baby food, canned vegetables), Frozen Foods, Dairy, Deli, Packaged Meat, Fresh Produce, Non-Food Grocery, Alcohol, and General Merchandise). For example, 31-ounce bag of Tide Pods has UPC 037000930389 is produced by Procter & Gamble and it is mapped to product module "Detergent-Packaged" in product group "Detergent", which belongs to the "Non-Food Grocery" department.[2] Each UPC contains information on the brand, size, packaging, and a rich set of product features including the weight and the volume of the product which we use to compute unit values.

Our data set combines all sales at the national and annual level. For each product j in year t, we define sales y_{jt} as the total sales across all stores and weeks in the year. Likewise, quantity q_{jt} is defined as total quantities sold across all stores and weeks in the year. Price p_{jt} is defined by the ratio of revenue to quantity, which is equivalent to the quantity weighted average price.[3]

[2]In the product group "Detergent" several product modules include: automatic dishwasher compounds, detergents heavy duty liquid, detergents light duty, detergents packaged, dishwasher rinsing aids, and packaged soap.

[3]We use the weight and the volume of the product to compute unit values.

Assigning Products to Firms

We link firms and products with information obtained from GS1 US, the single official source of UPCs. This allow us to perform the analysis at the parent company level, as opposed to at the level of the manufacturing firm.[4] In order to obtain a UPC, firms must first obtain a GS1 company prefix. The prefix is a five- to ten-digit number that identifies firms and their products in over 100 countries where the GS1 is present.[5] Given that the GS1 US data contains all of the company prefixes generated in the US, we combine these prefixes with the UPC codes in the RMS.

Lastly, for part of our analysis we obtain firm level characteristics from Compustat and NETS. To combine the Nielsen data with the Compustat database, we matched the names provided by the GS1 to those in Compustat using the string matching algorithm described in Schoenle (2017). After applying the algorithm we matched around 500 publicly traded companies over our sample period. Our matched sample represents 22% of the total sales in Compustat and 45% of the total revenue in the RMS. Approximately 21% of the total number of products in the data belong to publicly traded firms. We mostly use information contained in the Compustat 10-k. A 10-k is a comprehensive summary report of a firm's performance that must be submitted annually to the Securities and Exchange Commission (on top of the annual report). It includes a section with an overview of the firm's main operations, including its products and services. We use it to classify firms into firms that mostly operate in the CPG industry and firms that operate in CPG and other sectors. Details available in Data Appendix.

To combine the Nielsen data with the National Establishment Time Series (NETS), which is provided by Walls & Associates, and comprises annual observations on specific lines of business at unique locations over the period 1990-2014. In particular, NETS data allow us to observe and track sales, employment and industry classification of establishments. We matched the names provided by the GS1 to those in NETS using the string matching algorithm described in Schoenle (2017). We use information on industry of each establishment to classify firms that mostly operate in the CPG industry and firms that operate in CPG and other sectors. Details available in Data Appendix.

As a result of a firm-product match, we have a panel data at the firm level and firm ×

[4] Argente, Lee and Moreira (2018) provide more details on this data.
[5] More details available in the Appendix.

product category level on a detailed product portfolio.

B Defining Products and Product Innovation

Product definition

In our baseline analysis we define products at its finest level of disaggregation by using barcode data. A barcode is a 12-digit Universal Product Code (UPC), a code consisting of 12 numerical digits that is uniquely assigned to each specific good available in grocery and drug stores. Defining products as barcodes has some important advantages.[6] First, barcodes are by design unique to every product: changes in any attribute of a good (e.g. forms, sizes, package, formula) result in changes in barcode.[7] The most common alternative is to define goods by industry classifications, which will aggregate potentially very heterogeneous barcodes, and not allow us to identify changes in the product offerings of firms. Second, barcodes are so widespread that our data is likely to cover all products in the Consumer Packaged Goods sector. We will only miss the products that are not sold in grocery and drug stores. Finally, because firms and products are included in the sample provided that a sale occurs, we observe a wide range of products, and we can explore several dimensions of heterogeneity.

Product categories

In the Nielsen data, each product is classified into one of 1,070 low-level product modules. These product modules are further aggregated into a set of 114 product groups. For the purposes of our analysis, we would like to use a product classification scheme that groups together products that are relatively substitutable (meaning firms are plausibly directly competing with one another for customers) and does not group together products that are not substitutable. Of course, the precise definition of "relatively substitutable" is somewhat subjective. Nonetheless, through experimentation with different classification schemes, we

[6] This is also consistent to previous works, such as Broda and Weinstein (2010); Argente and Yeh (2017).

[7] Firms have strong incentives not to reuse barcodes. Assigning more than one product to a single barcode can interfere with a store's inventory system and pricing policy. Moreover, it is reasonable to assume that all goods with different UPCs differ in some way that might cause consumers to pay a different price for them and that it is rare for a meaningful quality change to occur that does not result in a change of UPC.

have found that the module level specification is far too fine for our purposes, while the group level is too coarse.

Thus there is a need for an intermediate level of aggregation, between modules and groups, which we generically call product category. As discussed in Section C.1, we use Wikipedia article texts to match each patent to a particular product category (be it module, group, etc.). As part of this analysis, for each product module, we assign a set of Wikipedia articles that characterize it. This turns out to be useful for creating new aggregations of modules (product category). By vectorizing the text of its associated Wikipedia articles, each module can be mapped into a high dimensional vector space, where the dimensionality M is the size of the vocabulary used for vectorization (see Appendix A.3 for details on this procedure).

Finally, we can aggregate these module vectors into clusters using a popular technique known as k-means clustering. This procedure allows one to specify the desired number of clusters k beforehand and yields a clustering assignment that minimizes the within group vector variance (average squared distance from cluster mean). That is, letting x be a given module vector and set S_i be cluster i, we choose our cluster sets S_i so as to minimize

$$\sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2$$

where $\mu_i = \frac{1}{|S_i|} \sum_{x \in S_i} x$.

In our main analysis, we use $k = 400$ for our cluster size. However, we also investigate various other choices for this particular parameter. See Section C.1 for a related discussion on our method of matching patents into product categories.

Product Innovation

We measure *product innovation* of a firm in a product category by using information on the number and nature of barcodes. A critical part of this exercise is the identification of entries. We define entry as the first quarter of sales of a product.[8] Thus, product introduction is simply measured as the count of new UPC's in a year by a firm in that product category.

[8] We cannot determine entry for products that are already active in the first year of the sample (2006).

$$\mathrm{N}_{jt}^{m} = \sum_{i=1}^{T_{jt}^{m}} \mathbb{1}[i \text{ is entry}].$$

To account for differences in quality across new barcodes, we use different approaches. First, we identify *newness* by using detailed product attributes of each barcode. We follow Argente and Yeh (2017) and construct a newness index that proxies for a quality improvement in a new product. The index uses detailed information about the observable characteristics of each UPC provided in the Nielsen RMS data set and counts the number of new and unique attributes a product has at the time of its introduction relative to all of the other products ever sold within the same product module. We define a product i in product module m as a vector of characteristics $V_{mi} = [v_{mi1}, v_{mi2}, .., v_{miA_m}]$ where A_m denotes the number of attributes we observe in product module m in our data.[9] Let Ω_{mt} contain the set of product characteristics for each product ever sold in product module m at time t, then the *newness index* of a product i in product module m, launched at time t is defined as follows:

$$\text{Newness}_{it}^{m} = \frac{1}{A^m} \sum_{k=1}^{A^m} \mathbb{1}[v_{mik} \notin \Omega_{mt}].$$

For example, if a new product within the soft drinks category enters with a flavor and size that has never been sold in any store before, its newness index is $(1+1)/N_{\text{soft drinks}} = 2/8$. On average, we observe 7.2 product characteristics in each product module.[10] We assume that each attribute is equally weighted in order to remain agnostic about the relative importance of each attribute to the degree of newness of a product.[11]

[9]For example, the product module "soft drinks - carbonated" consists of $A_{\text{soft drinks}} = 8$ attributes for each barcode: brand, flavor, firm, size, type (sparkling soda or natural soda), container (e.g. can or bottle), formula, generic (i.e. private label).

[10]Comparing the newness index of different products across distinct modules depends not only on the number of new attributes of each product but also on the total amount of observable characteristics the Nielsen data provides for each module. The minimum characteristics we observe for each module is 5 and the maximum is 12.

[11]See Argente, Lee, and Moreira (2018) for more details on the the average of the index for each of the product groups in the Nielsen data.

C Patent Data

Data Sources

Our main data source for patent analysis is United States Patent Office (USPTO) data on the universe of published patent applications, granted or not. We make use of the original bulk data files provided by USPTO and distributed through the PatentsView platform by the USPTO Office of the Chief Economist in addition to our own parsing of the original files from the Bulk Data Storage System. Our data contains information on 10 million patent applications filed by more than 500 thousands assignees for the period 1975-2017. Advantage of using all patent applications, as opposed to just granted applications, is twofold. First, since patent grant usually on average happens with a 2-years lag, latest sample years suffer from the severe truncation. Having all patent applications significantly reduces this problem. Second, we have a larger sample and can differentiate between patents that are granted, pending, or abandoned – this can also serve as one of the quality measures of patents as discussed below.[12]

For each patent, we utilize information on the following main variables of interest: patent application year, patent status (granted, pending, abandoned), patent technology classifications (IPC), forward patent citations received, number of claims on a patent, and information on patent assignees. For our textual analysis of patent documents described below, we use the full patent texts.

Patent Assignees

To assign patents to firms, we proceed in the following steps. First, since our product-level data assigns products to firms as of 2015 , we need to treat patents in a similar way. Hence, in the first step, we utilize patent assignment dataset together with the Thomson Reuters Mergers & Acquisition data[13] to designate each patent to its most current holder. The details on this step are delegated to Appendix A.1. Second, since firm names in these datasets often have misspellings or various abbreviations, it is challenging to cleanly identify from raw data

[12] In fact, adding non-granted patents information increases number of patents in the data from 6 million to 10 million.

[13] Thomson Reuters M& A provides complete coverage of global mergers and acquisitions activity, including more than 300,000 US-target transactions, since 1970. The data covers mergers of equals, leveraged buyouts, tender offers, reverse takeovers, divestitures, stake purchases, spinoffs , and repurchases. It also provides detailed information of the target, the acquirer, and the deal terms.

which companies are the same. To overcome this challenge, we develop a company name cleaning algorithm to clean and standardize company names. This procedure builds on and extends cleaning algorithms from the NBER Patent Data Project (Hall et al., 2001) and Akcigit et al. (2016). We delegate the details of this procedure to Appendix A.2.

C.1 Matching Patents to Product Categories

Our goal is to classify each patent into a product category, meaning one of the Categories whose construction we described in Section B. To do this, we use a text similarity approach. That is, for each patent and product category, we construct a representative document (a set of words), and base our classification on those pairs yielding the highest similarity.

The first task is to construct these representative documents. Our distinguishing data on the patent side consists primarily of the text of the patent application (or publication), which includes: title, abstract, and list of claims. We also have US and international patent classification codes for each patent, each of which has an associated short text description. The product side is much more limited. Each category has an associated title, and the products within the categories have descriptors, but these are primarily abbreviations which are often hard to interpret.

To get around this limitation of the Nielsen data, we utilize the text of Wikipedia entries as an intermediary. Specifically, for each product category, we manually assign a list of one to three Wikipedia entries that most closely represent the category. This task is much more manageable than the patent-category match. We then use the text of these entries to facilitate this full match.[14]

Having constructed these documents, we then convert them into vectors (one for each document) representing word frequencies and calculate numerical similarities between these vectors to generate a match. The precise details of this match are described in Appendix A.3. There are a number of design decisions and metaparameters that go into this construction, and in each case we use those prevalent in the well-established natural language processing (NLP) and machine learning literature and associated software environments.

The final result of the vectorization procedure is an ℓ^2-normalized word frequency vector for each patent and each product category. Each document vector is of length M, which is

[14] In the case where two modules have precisely the same set of associated Wikipedia entries, we aggregate them together. Regardless, the k-means algorithm would trivially group them together as well.

the number of tokens that we include in our vocabulary. Multiplying any two such vectors together yields a similarity metric between two documents. This is guaranteed to be in the range $[0, 1]$ with zero corresponding to zero word overlap and one corresponding to the case in which the documents are identical (or are multiples of one another).

Generating the patent-product match is then a matter of finding, for each patent, which modules among those where the firm has any products (in any year) are most similar to it in a textual sense. To this end, we entertain multiple notions of a match. One is simply looking at the generally unique maximally similar module. Another is using the previous notion but also having an absolute threshold for a valid match of 0.05. This last option allows for the possibility of a non-match. The results presented here rely on this last notion.

C.2 Patent quantity and quality measures

We construct several measures of patenting by firms or firm × product category over time. A simple count of patent applications constitutes our basic dynamic measure. Notice that we count all patents in the year they were applied for since this should be year when the innovation occurred. Hence, we denote by $Patents_{it}$ number of patents applied by firm i at time t. Likewise, we denote by $Patents_{imt}$ number of patent applications by firm i in category m at time t.

It is well known that patents are very heterogeneous in their quality. While the first basic measure places all importance on the quantity, we also construct measures that rely on value of a patent. As the first option, we designate abandoned patents[15] as low-quality patents and define count of abandoned patents by firm i (and category m) applied in year t as $Patents\ abn_{i(m)t}$. Similarly, we will denote by $Patents\ grant_{i(m)t}$ number of patent applications that eventually got granted.

Another measure of patent quality is patent citations. Citations received have traditionally been used as measures of the economic and technological significance of a patent (see Pakes (1986), Schankerman and Pakes (1986), Trajtenberg (1990), Harhoff et al. (1999), Hall et al. (2001), Bessen (2008), Kogan et al. (2012), Moser et al. (2015)). We define citations-adjusted patent count of a firm as total number of patents weighted by forward citations received by

[15] Abandoned patents account for 16% of all patent applications in our data. Patents may get abandoned before they have been granted or they may get abandoned after the grant if they are not renewed every 4 years.

patent j. Note that this does not depend on time t, as it counts all the forward citations that patent will ever receive, not the citations received until time t. For truncation adjustment, which takes into account the fact that more recent patents have less time to accumulate citations, we focus on the citations received in the first 5 years since the application time. Hence, another variable we construct is $Patents\ cit_{i(m)t}$ which is the total number of 5-year forward citations received by all patents filed by firm i (in category m) year t.

D Matching Products and Patents

In order to study the relationship between new products and patents, we develop algorithms that match patents to product. Our first algorithm uses information on the patents assignee name and firm name from the product data. It basically links all patents of a firm to all products of a firm in a year. With this algorithm we build a dataset at the firm-year level that allow us to explore the relationship between products and patents using variation across firms over time. Our second algorithm, uses text analysis of patent documents and product descriptions expanded with Wikipedia dictionaries to classify patents across product categories, allowing us to match sets of patents to sets of products within a firm. With this algorithm we are able to build a dataset at the firm-category-year level with information on patents and product innovation. This dataset is particularly useful to study large firms. Large firms have products and patents in multi-categories, allowing us to study the relationship between products and patents using variation within the firm.

Match 1: Firm × Year Level

Our basic match at the firm level uses firm × year level datasets constructed on products (Section A) and patents (Section C) side and matches them using unique company identifiers derived from name cleaning in Appendix Section A.2. Our matched Nielsen-USPTO sample consists of 5,170 firms which can be divided into those firms that issue a patent during our sample period for Nielsen data 2006-2015 (a total of 3,284 firms) and those that issue a patent prior to 2006 (a total of 1,879 firms).

Match 2: Firm × Category × Year Level

Our preferred match at the firm × category × year level is possible now because in Section B we classified products into various product categories and in Section C.1, using textual analysis of patent documents, we assigned firms' patents to product categories. Hence, our benchmark data, on the one hand, has product innovation measures and other variables related to product portfolio, and, on the other hand, has patent-related variables for each firm within a category over time. The matched sample has 5,170 firms operating in 400 product categories with more than 250 thousand observations.

E Summary Statistics

In this section, we provide descriptive statistics of our basic firm-year dataset on patents and products. Our baseline product data has approximately 35 thousand firms and an average of 300 thousand active products every quarter. To define firm's patenting status, we divide firms into three categories: (i) firms that have never patented, (ii) firms that patented last before 2006 (the beginning of the Nielsen RMS data set) and (iii) firms that have patents between 2006 and 2015. Table I shows the share of each type of firm in our data. It also shows the share of total products and the share of total revenue accounted by different type firms. The table shows that a large amount of products in the market belong to firms that have never patented. This accounts for approximately 46% of the total revenue generated in the CPG sector. In our matched sample, 9.5% of the firms have patents during our sample period. However, they account for disproportionately large number of products in the data – 34%, and account for almost half of the total revenue in the sector. These firms represent only 2.27% of the total number of U.S. firms filing for a patent in USPTO in 2006-2015. However, their patents represent about 20% of all patents filed – showing the importance of the CPG patenting in the universe of U.S. patents.

Table I: Number of Firms, Products, and Revenue in Nielsen RMS by Patenting Status

	Firm's patenting status		
	No Patents	Patents before '06	Patents in '06-'15
Number of Firms	85.05%	5.44%	9.50%
Total Number of Products	57.35%	8.9%	33.7%
Total Revenue	45.70%	10.81%	43.49%

Notes: The table shows the shares of all firms, of total products, and of total revenue in Nielsen RMS data accounted by firms with different patenting status. The first column is for firms that have no patents, the second column is for firms that have patents but before they first appear in Nielsen RMS (before 2006), and the third column is for firms that have patents in our main period 2006-2015. The total number of firms in RMS Nielsen data is 34,536. The total number of products active every quarter is 300 thousand.

Table II shows more detailed statistics for firms of different patenting status. It shows that patenting firms are larger (in revenue) and they not only introduce more products but these products also generate more revenue. They are also more diversified than firms that never patented or that patented last before 2006. On average, firms that patent have products in more than 4 different product categories. Interestingly, although patenting firms introduce more high-revenue products, a big chunk of those products is not very novel according to our newness index: on average, firms that never patent introduce fewer, but more novel products.

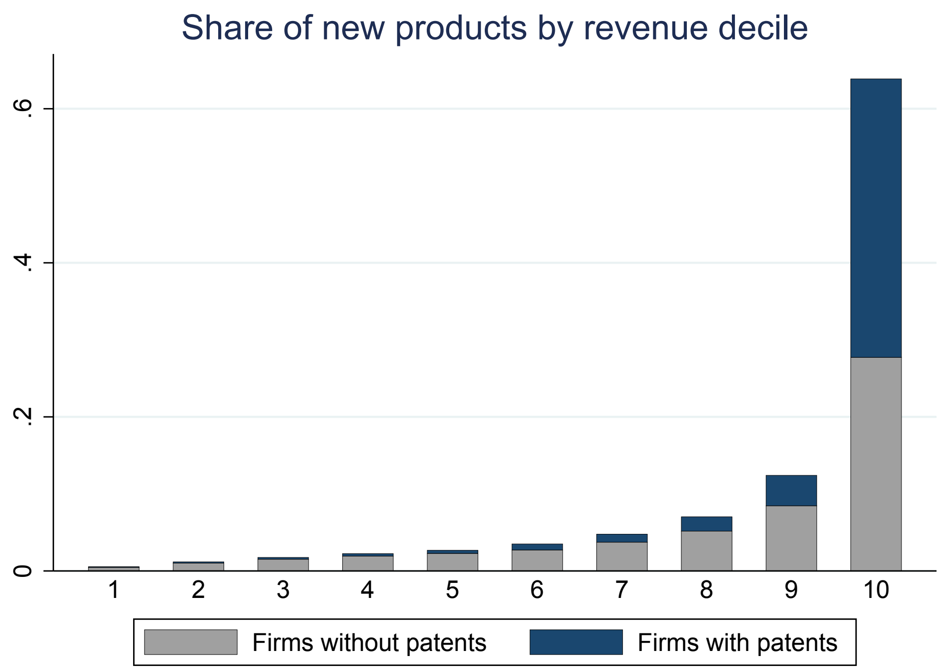
Table II: Summary Statistics by Patenting Status

	Firm's patenting status		
	(1)	(2)	(3)
	No Patents	Patents before '06	Patents in '06-'15
Revenue all products	3708.75	12275.72	27598.79
Revenue new products	243.27	988.62	1670.12
Revenue new products, post entry period	386.62	2064.76	3955.06
Number of products	16.11	35.09	74.09
Number of new products	2.60	6.73	12.91
Product entry rate	0.19	0.17	0.22
Number of product categories	2.36	3.11	4.11
Share of new products lasting more than 4 quarters	0.74	0.70	0.75
Share of new products lasting more than 16 quarters	0.44	0.40	0.42
Average newness of new products	0.13	0.09	0.10
Newness-weighted number of new products	0.57	0.70	1.07
Number of patent applications	0.00	0.00	6.14
Number of granted patent applications	0.00	0.00	4.47
Number of citations-weighted patent applications	0.00	0.00	8.87
Stock of patent applications until year t	0.00	10.88	125.93
Stock of granted patent applications until year t	0.00	10.59	115.39
Number of different technology classes (IPC3) on patents	.	.	5.56
NumFirms	29373	1879	3284
Observations	188118	15285	29030

Notes: The table shows descriptive statistics for a pooled sample of firms for the period 2006-2015 by firms with different patenting status. The first column is for firms that have no patents, the second column is for firms that have patents but before they first appear in Nielsen RMS (before 2006), and the third column is for firms that have patents in our main period 2006-2015. Observations are at the firm × year level.

Figure 2 shows that in the CPG sector most product creation occurs in large high-revenue firms. The figure shows the share of new products introduced by revenue decile. More than 60% of new products launched in the sector belong to firms in the top decile. The figure also shows that patenting is more common in large firms. However, the figure also shows that, even within large firms, almost half of new product innovations do not come from patenting firms. That is, more than half of the product innovations in the sector are not patented. This is more clearly shown in Appendix Figure A.3. The figure shows our main measures of product innovation – number of new products and newness-weighted number of new products – accounted by firms with different patenting statuses. Around 55% of new product innovation comes from never-patenting firms. This number is even larger if we consider products weighted by newness index – our measure of novelty of the products. This indicates that, by considering innovation only based on patenting, we are missing more that half of the product innovation occurring in this sector.

Figure 2: New Products and Patents by Revenue Decile

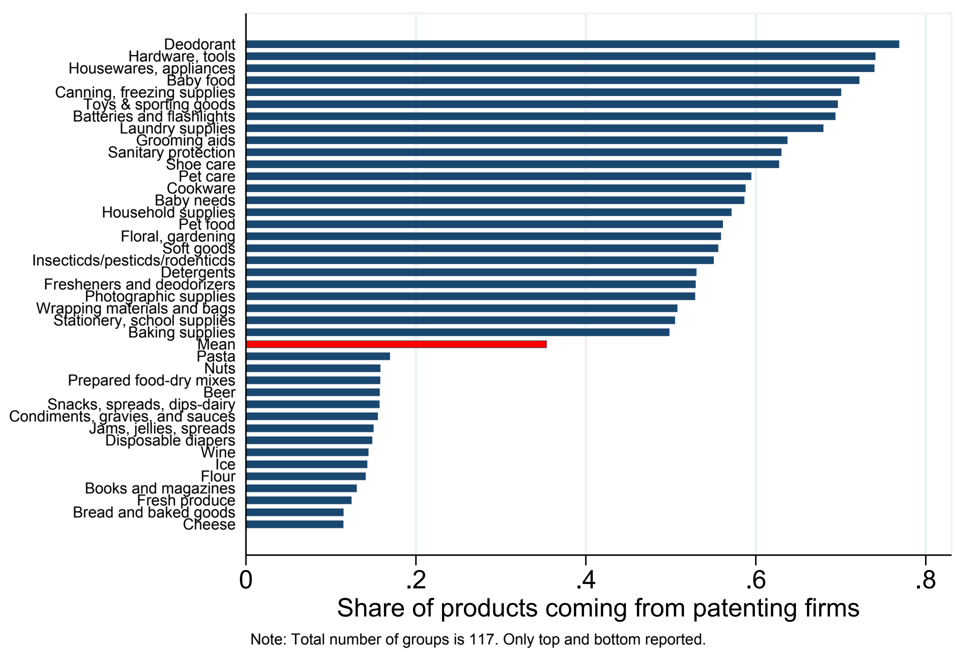


Note: The figure shows the share of new products by revenue decile. The blue part of the bars indicate the share of new products introduced by firms that do not have a patent. The red part of the bars indicate the share of new products that is introduced by firms with patents.

In which product groups are firms more likely to patent? One advantage of the Nielsen data is that it allows us to explore a wide range of product categories, from perishables to semi-durable products. Interestingly, patenting firms are present in every product group. Figure 5 shows the product groups according to their patenting intensity – the share of products coming from patenting firms. In our sample, the average product group has patenting intensity of 50%. The figure shows that product groups with more durable products have high patenting intensity. Some examples are: deodorants, detergents, shaving needs, cookware, and kitchen gadgets. On the other hand, product groups such as wine, flour, ice, and cheese show low patenting intensity.

To confirm this intuition we construct a measure of the durability of a product based on the Nielsen Consumer Panel. For each product group we count the average number of shopping trips made by households in a given year to purchase products in each product group. Our assumption is that if households take longer to purchase products for a certain category those products must last longer. Thus, we call categories with few trips per year durable categories. Examples of durable categories are sun exposure trackers (1.00), bathroom scales (1.03) and printers (1.03), where the average number of shopping trips per year is in parenthesis. Examples of non-durable categories are refrigerated milk (23.61), cigarettes (19.19) and fresh bread (18.76). Our measure of durability is the inverse of the average number of trips per year in a given product group. Panel (a) in Appendix Figure A.4 shows the relation between the share of patenting firms and our durability measure. The figure shows a clear positive relationship between the fraction of firms patenting in a group and the durability of the products sold in that group. Panel (b) shows that the relationship holds if we focus on the share of new products by patenting firms. These figures suggest that firms are more likely to rely on patenting in sectors where products are more likely to last longer after a household purchase.

Figure 3: Patenting Intensity by Product Groups



Notes: the figure shows the share of products belonging to patenting firms (patenting intensity) for a sample of product groups in the Nielsen data. The figure shows the intensity of the top 25 groups, the average intensity, and the intensity of the bottom 15 groups.

III Patents and Product Innovation

How good are the patent-based measures of innovation in capturing true product innovation? To shed light on this question, we start by exploring how strongly patenting is associated with product innovation in our data. For this, we utilize both layers of the data matches – at the firm × year and firm × category × year levels. In Section A, we present regression analysis for the behavior of firm's product innovation before and after the first patent application, by using variation within firm × product category and time × product category. In Section B, we present a similar analysis but for the intensive margin of patenting. We show how firm's product innovation behaves around firm's patent application times. In this section, we find that:

Fact 1: *Patenting is strongly positively associated with product innovation by the firm.*

A Extensive Margin of Patenting

Do firms introduce new or more innovative products after they start patenting? In this section, we establish that firms introduce more products and more novel new products after they switch to patenting.

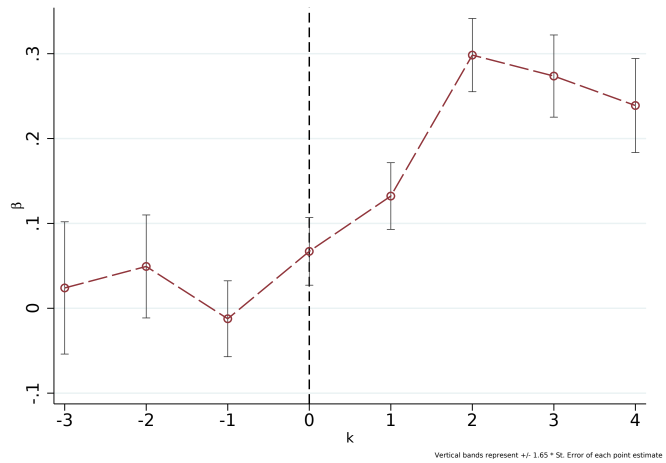
A.1 Exploring firm-level variation

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 if there are changes in the firm's product introduction following the application of a patent (likewise, can be done for after a patent is granted). To evaluate if firm's outcomes change after a firm starts patenting, we estimate the following specification:

$$\ln Y_{j,t+k} = \beta(k) dP_{j,t} + \alpha_j + \gamma_{t+k} + u_{j,t+k} \ , \ k = -3, ..., 0, ..., 4 \tag{1}$$

where $Y_{j,t+k}$ is the outcome of firm j in $t+k$, $dP_{j,t}$ is a dummy that takes value 1 after a firm gets its first patent, α_j are firm fixed-effects, and γ_{t+k} are time fixed-effects. Firm fixed -effects ensure that all firm-specific influences are accounted for, provided they are invariant over time. Year fixed effects are also included to reflect time-varying factors common to all firms. With firm and time fixed effects, the approach is akin to a difference-in-difference model. Identification rests on the dispersion across firms (and time) of first patenting, i.e. uses conditional variation of firms that started patenting between 2006 and 2015.[16]

Figure 4: Product Innovation and Change in Patenting Status



Notes: The figure plots the estimated coefficients after estimating equation (1) on log number of new products. The graph should be read as follows: Firms that become patentees in t=0, change product creation by β percent in t=-3,...,4.

Figure 4 presents the estimated change of number of new products (logs) associated with a switch in status from being a non-patentee to a patentee. Firms that become patentees in t=0, change product creation by approximately 10-30 percent after becoming a patenting firm. The estimated effects before zero are non-significant, indicating that product creation does not changes prior to becoming a patentee. The significant estimated effect on product creation after 3 and 4 years indicate that the effect is persistent (as opposed to temporary).

These dynamic effects lead to an average increase in product introduction by about 20 percent after the switch to patenting as seen in the first column of Table III. This effect is mostly driven by high-quality patents (using their granted versus abandoned status as a proxy) as seen from columns 2 and 3. Following Acemoglu et al. (2019), we tested the strength of the correlation by estimating a dynamic panel model with firm and time fixed effects, to better account for confounding effects. The regression consists of adding to the specification one lags of the outcome variable – log new products – to control for the dynamics of new

[16] In the Appendix we develop alternative approaches using alternative sources of variation.

products. Columns (4)–(6) of Table III shows that while the number of new products are serially autocorrelated, the relationship between becoming a patentee and product introduction remains very similar.

As discussed above, our measure of product introduction does not measure degree of newness of the new products. This is at odds with the nature of patents, as they may represent a novel product/process. In the second panel of Table III we estimate the average impact on the (log) quality adjusted number of new products. The results show a similar relationship between becoming a patentee and product introduction.

A.2 Exploring product category and firm-level variation

The results above show that firms introduce more products and more novel new products after they switch to patenting, thus consistent with patents representing some technological innovation that gets reflected in firms' product offerings. While we cannot establish causality, these results are robust when we control for product category confounding effects. We use our data at the firm $\times$ product category $\times$ year level and find that patenting in a specific product category is positively associated with product introduction in the same product category.

Table IV shows that after the first patent introduction in product category, firms introduce on average 20% more products in the same product category, conditional on sector specific time and firm effects. Moreover, what matters most are high-quality patents – patents that are eventually get granted.

To control for time-varying firm-level unobservables that may affect both patenting decision and product introduction, in columns (4)–(6), we test if the relationship exists after we account for serial correlation of product introduction. Across the different specifications, we see a strong positive relationship between the first patent application and product innovation afterwards. The relationship is however weaker when we account for the quality of product introduction (Panel B of Table IV). One prior that one may have is that patents reflect explicit innovations that are likely to be associated with new products that have a large number of novel features. This hypothesis does not find strong support in the aggregate data. In Section V, we will explore heterogeneity of this effect and try to understand for which firms, this association is stronger.

Table III: Product Innovation after the First Patent: Firm-level evidence

	Baseline			Dynamic Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New Products (Log N)						
After I patent(t)	0.1738**			0.2004***		
	(0.053)			(0.052)		
After I granted patent(t)		0.1605**			0.1013*	
		(0.056)			(0.052)	
After I non-granted patent(t)			0.0696			0.2683**
			(0.127)			(0.128)
Log N(t-1)				0.0004***	0.0004***	0.0004***
				(0.000)	(0.000)	(0.000)
Observations	195,686	195,686	195,686	158,554	158,554	158,554
R-squared	0.897	0.897	0.896	0.899	0.899	0.899
Time Effect	Y	Y	Y	Y	Y	Y
Firm Effect	Y	Y	Y	Y	Y	Y
Panel B: Quality-adjusted New Products (Log q-N)						
After I patent(t)	0.1806*			0.2678**		
	(0.089)			(0.102)		
After I granted patent(t)		0.2668**			0.2583*	
		(0.125)			(0.157)	
After I non-granted patent(t)			-0.1656			-0.0168
			(0.220)			(0.211)
Log q-N(t-1)				0.0013	0.0013	0.0013
				(0.002)	(0.002)	(0.002)
Observations	50,146	50,146	50,146	24,495	24,495	24,495
R-squared	0.836	0.836	0.836	0.800	0.800	0.800
Time Effect	Y	Y	Y	Y	Y	Y
Firm Effect	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 ($Log\ q - N$) in a firm $\times$ year as a function of a dummy equal to one after the first patent application in a firm $\times$ year. Quality of a product is based on our $Newness$ index defined in Section B *After I patent* is a dummy equal to one after any patent application; *After I granted patent* is a dummy equal to one after a patent application that is granted; and *After I non-granted patent* is a dummy equal to one after a patent application that has not been granted (abandoned or pending).

Table IV: Product Innovation after the First Patent: firm × product category-level evidence

	Baseline			Dynamic Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New Products (Log N)						
After I patent(t)	0.1946**			0.1947*		
	(0.092)			(0.105)		
After I granted patent(t)		0.2496**			0.2785**	
		(0.108)			(0.134)	
After I non-granted patent(t)			0.1553			0.1036
			(0.127)			(0.128)
Log N(t-1)				0.0384***	0.0383***	0.0388***
				(0.011)	(0.011)	(0.011)
Observations	494,546	494,546	494,546	390,963	390,963	390,963
R-squared	0.897	0.897	0.896	0.899	0.899	0.899
Time-Category Effect	Y	Y	Y	Y	Y	Y
Firm-Category Effect	Y	Y	Y	Y	Y	Y
Panel B: Quality-adjusted New Products (Log q-N)						
After I patent(t)	-0.0994			0.0649		
	(0.161)			(0.179)		
After I granted patent(t)		0.1831			0.4125***	
		(0.125)			(0.157)	
After I non-granted patent(t)			-0.3734*			-0.2127
			(0.220)			(0.211)
Log q-N(t-1)				-0.0360*	-0.0361*	-0.0369*
				(0.020)	(0.020)	(0.021)
Observations	55,704	55,704	55,704	20,052	20,052	20,052
R-squared	0.836	0.836	0.836	0.800	0.800	0.800
Time-Category Effect	Y	Y	Y	Y	Y	Y
Firm-Category Effect	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 ($Log\ q-N$) in a firm × category × year as a function of a dummy equal to one after the first patent application in a firm × category × year. Quality of a product is based on our $Newness$ index defined in Section B. *After I patent* is a dummy equal to one after any patent application; *After I granted patent* is a dummy equal to one after a patent application that is granted; and *After I non-granted patent* is a dummy equal to one after a patent application that has not been granted (abandoned or pending).

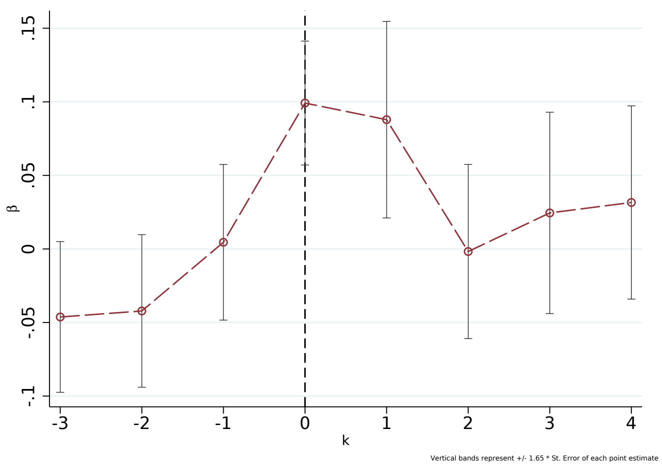
B Intensive Margin of Patenting

We now explore how the performance of the firm changes as the firm accumulates patents (granted and non-granted). As before, we are interested in evaluating the timing of an increase in the number of patents and the outcomes using the following specification:

$$\ln N_{j,t+k} = \beta \ln NP_{j,t} + \alpha_j + \gamma_{t+k} + u_{j,t+k}\ ,\ k = -5, ..., 0, ..., 5 \tag{2}$$

where $N_{j,t+k}$ is the number of new products of firm j in $t+k$, $NP_{j,t}$ are the new granted patents of the firm j in t, α_j are firm fixed-effects, and γ_{t+k} are time fixed-effects. Figure 5 shows the estimated effects. We observe that the new granted patents has contemporaneous effects on product creation. Contrary to the extensive margin, we do not observe that the accumulation of patents has long-lasting effects. Firms introduce new products on impact – the same year or a year after the patent applications.

FIGURE 5: PRODUCT INNOVATION AND PATENTING INTENSITY STATUS



Notes: The figure plots the estimated coefficients after estimating equation (1) on log number of new products. The graph should be read as follows: Firms that become patentees in t=0, change product creation by β percent in t=-3,...,4.

Table V: Product Innovation: Intensive Margin

	Baseline			Dynamic Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New Products (Log N)						
Patents(t-1)	0.0411***			0.0334***		
	(0.011)			(0.012)		
Patents granted(t-1)		0.2496**			0.0443***	
		(0.013)			(0.014)	
Patents non-granted(t-1)			0.0193			0.0195
			(0.015)			(0.016)
Log N(t-1)				0.0357***	0.0357***	0.0357***
				(0.002)	(0.002)	(0.002)
Observations	455,307	455,307	455,307	392,477	392,477	392,477
R-squared	0.673	0.673	0.673	0.693	0.693	0.693
Time-Category Effect	Y	Y	Y	Y	Y	Y
Firm-Category Effect	Y	Y	Y	Y	Y	Y
Panel B: Quality-adjusted New Products (Log q-N)						
Patents(t-1)	0.0628**			0.0522		
	(0.027)			(0.039)		
Patents granted(t)		0.0633*			00.0494	
		(0.033)			(0.045)	
Patents non-granted(t-1)			0.0493			0.0433
			(0.034)			(0.052)
Log q-N(t-1)				-0.1202***	-0.1203***	-0.1203***
				(0.010)	(0.010)	(0.010)
Observations	41,254	41,254	41,254	16,801	16,801	16,801
R-squared	0.566	0.566	0.566	0.624	0.624	0.624
Time-Category Effect	Y	Y	Y	Y	Y	Y
Firm-Category Effect	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 ($Log\ q - N$) in a firm × year in Panel B as a function of log number of patents. Quality of a product is based on our *Newness* index defined in Section B. *Patents* is the number of any patent applications in firm × year; *Patents granted* is the number of granted patent applications; and *Patents non-granted* is the number of patent application that have not been granted (abandoned or pending).

IV Innovation, Patent Value, and Firm Heterogeneity: Conceptual Framework

How do firms decide on patenting and product innovation? Where does the value of a patent come from? On the one hand, patent is meant to represent a technological improvement that can be commercialized to generate returns for a firm. On the other hand, however, it is well-known that patents can be used for other reasons not directly related to innovation. Many patents, for example, are filed purely for defensive reasons – strengthening existing patents' position and with the intention to exclude competitors from using a particular invention or building a new one around it (e.g., Cohen et al. (2014), Abrams et al. (2013)). In this case, the value of a patent for firms in terms of revenue should capture both, the value from innovation improvements, as well as market value from deterring competition. The latter could be large even if the patent is not used by the firm to launch new products or improve its processes.

In this section, we present a simple model of firm dynamics with firm's innovation and patenting decisions to speak to these issues. In the model, firms can exert *productive* innovation effort, innovate, patent, and push the technology frontier. In addition, firms can also spend resources on *protective* patent that does not push the technology frontier but rather helps firms protect their technology by limiting competition and creative destruction.

We use the model to generate several predictions that are then tested in the data. The model predicts that as firms grow and become market leaders, they shift their innovation more towards protective strategies. The average patent of a market leader is associated less with product innovation and more with competitors' deterrence, relative to smaller firms. Value from a patent increases with firm size. After decomposing this value into *productive* and *protective* components, we see that patent value of market leaders is driven more by protective component. Our goal (TBD) is to calibrate this model to our data and infer the monetary value of a patent and its decomposition into productive and protective components depending on market position of a firm.

Production Consider a partial equilibrium model with the following sectoral production function. We should think of this sector as a product category, or module analog in our model.

$$Y = \frac{1}{1-\beta}\left[\sum_{m=1}^{M} q_m^{\frac{\beta}{1-\beta}} y_m\right]^{1-\beta} \tag{3}$$

In this production function, y_m denotes the output of a producer of vintage m. Different vintages differ by their qualities q_m. The newest vintage M has the highest quality. Different vintages are perfect substitutes after adjusting for their quality. Sectoral output producers are perfectly competitive and we normalize the price of sectoral good to one. The demand function faced by vintage$-m$ producer is given by:

$$p_m = q_m^{\frac{\beta}{1-\beta}}\left[\sum_{m=1}^{M} q_m^{\frac{\beta}{1-\beta}} y_m\right]^{-\beta} \tag{4}$$

Producers of different vintages compete in prices in order to capture the full market. In equilibrium, the firm with the highest quality will serve the market. Moreover, the producing firm will also charge the monopoly price. Hence, we will study the monopolist problem throughout this section. Production technology for each monopolist is one-for-one in labor:

$$y = l \tag{5}$$

The wage rate of the worker is given by w. Hence, the marginal cost of production is simply w.

Monopolist's static maximization A monopolist maximizes its profit subject to the demand function it faces as follows:

$$\begin{aligned} \pi &= \max_{l}\{py - wl\} \\ s.t. &\quad p = q^{\beta}y^{-\beta} \text{ and } (5) \end{aligned}$$

This maximization delivers the following equilibrium labor (l), output (y), revenue (Rev), profits (Π) for firms that face regulatory burden

$$\begin{aligned} y = l &= \left(\frac{1-\beta}{w}\right)^{\frac{1}{\beta}} q, \qquad Rev = \left(\frac{1-\beta}{w}\right)^{\frac{1-\beta}{\beta}} q \\ \Pi &= \beta\left(\frac{1-\beta}{w}\right)^{\frac{1-\beta}{\beta}} q \equiv \pi q, \end{aligned}$$

Firms with higher-quality products are larger, employ more labor, generate higher revenue and profits.

Productive and Protective Innovation Incumbent may choose to invest in R&D to come up with a patent. Patent may be *productive* and increase productivity as in usual innovation-based growth models (Aghion and Howitt (1992)). If firm pays the cost

$$C = \frac{x^2}{2}q$$

it gets a Poisson arrival rate of innovation x. If innovation is successful, it increases firm's productivity from q to $q(1+\lambda)$, where $0 < \lambda < 1$. At each point in time, incumbent firm faces probability of creative destruction from entrants τ. For simplicity, we consider τ to be exogenous for now. If firm is creatively destroyed by a competitor, it is replaced and exits the market.

Alternatively, incumbent's patent may be *protective* – directed to protect innovation and deter entry. To file for a protective patent, firm needs to pay a fixed cost c. In return, competitors will find it harder to replace an incumbent (for example, by expecting higher infrindgement possibility and the litigation costs). For simplicity, let the competitors' innovation rate be τ/γ when incumbent has a protective patent, where $\gamma > 1$ and is the strength of protection offered by a protective patent.

Value Functions and Solution Let us conjecture that $\exists\ q^*$, such that when $q > q^*$ firm decides to file for a protective patent, while below q^* firm only conducts a productive innovation. We will verify this conjecture below. In what follows, we focus on BGP where all aggregates and value functions grow at constant rate g (we will also utilize a household Euler equation $r = g + \rho$ from a household with logarithmic preferences).

Consider $q < q^*$:

Denote by $V_1(q)$ a value of being an incumbent with a product quality q. Then

$$\rho V_1(q) = \max_x \left[\pi q - \frac{x^2}{2}q + x(V_1(q(1+\lambda)) - V_1(q)) - \tau V_1(q)\right] \tag{6}$$

Guess that value function is linear in quality: $V_1(q) = v_1 q$

$$\rho v_1 q = \max_x \left[\pi q - \frac{x^2}{2} q + x v_1 q \lambda - \tau v_1 q \right]$$

$$\begin{aligned} FOC &: \quad xq = v_1 q \lambda \\ x^* &= v_1 \lambda \end{aligned}$$

substituting yields that v_1 should be a solution to the following quadratic equation.

$$\frac{\lambda^2}{2} v_1^2 - v_1(\tau + \rho) + \pi = 0 \tag{7}$$

Consider $q > q^*$:

Denote by $V_2(q)$ a value of being an incumbent with a product quality q. Then

$$\rho V_2(q) = \max_x \left[\pi q - \frac{x^2}{2} q - c + x(V_2(q(1+\lambda)) - V_2(q)) - \frac{\tau}{\gamma} V_2(q) \right] \tag{8}$$

Notice that now firms pay cost of protective patent and face lower destruction rate. We can show that now value function can be written as

$$V_2(q) = v_2 q - \frac{c}{\rho + \tau/\gamma} \text{ , where}$$

$$\frac{\lambda^2}{2} v_2^2 - v_2(\tau/\gamma + \rho) + \pi = 0 \tag{9}$$

Now, let us find q^* that makes incumbents indifferent between investing into protective patent or not.

$$\begin{aligned} v_2 q^* - \frac{c}{\rho + \tau/\gamma} &= v_1 q^* \\ q^* &= \frac{\frac{c}{\rho + \tau/\gamma}}{v_2 - v_1} \end{aligned} \tag{10}$$

Hence, when $q > q^*$, firms decide to get a protective patent in addition to their regular innovation effort. This leads to two theoretical predictions:

Prediction 1: Larger firms have higher incentives to protect their innovation.

Prediction 2: Patents of larger firms limit creative destruction more.

Intuition is simple. Larger firms have higher quality products that generate higher returns. Hence, if larger incumbents are replaced by competitors, they have more to lose, so have high incentives to protect.

Notice that small firms below the threshold size q^* only apply for productive patents. However, in our setting, large firms' productive patents are complemented also with protective patents. Protective patent does not improve productivity, hence an average patent of a large firm contributes less to productivity improvement. Hence, our next prediction, in line with our empirical evidence, is

Prediction 3: As firms grow larger, an average patent of a firm contributes less to innovation.

The Value of a Patent What is the private value of a patent in our model? If $q < q^*$, each patent has a productive value and a value of a patent (PV) is

$$PV_1 = V_1(q(1+\lambda)) - V_1(q) = v_1 \lambda q$$

If $q > q^*$, each productive patent comes also with its protective patent. Hence, a value of the average patent is

$$PV_2 = 0.5 \left[\underbrace{V_2(q(1+\lambda)) - V_1(q(1+\lambda))}_{\text{Protective value}} + \underbrace{V_1(q(1+\lambda)) - V_1(q)}_{\text{Productive value}} \right]$$

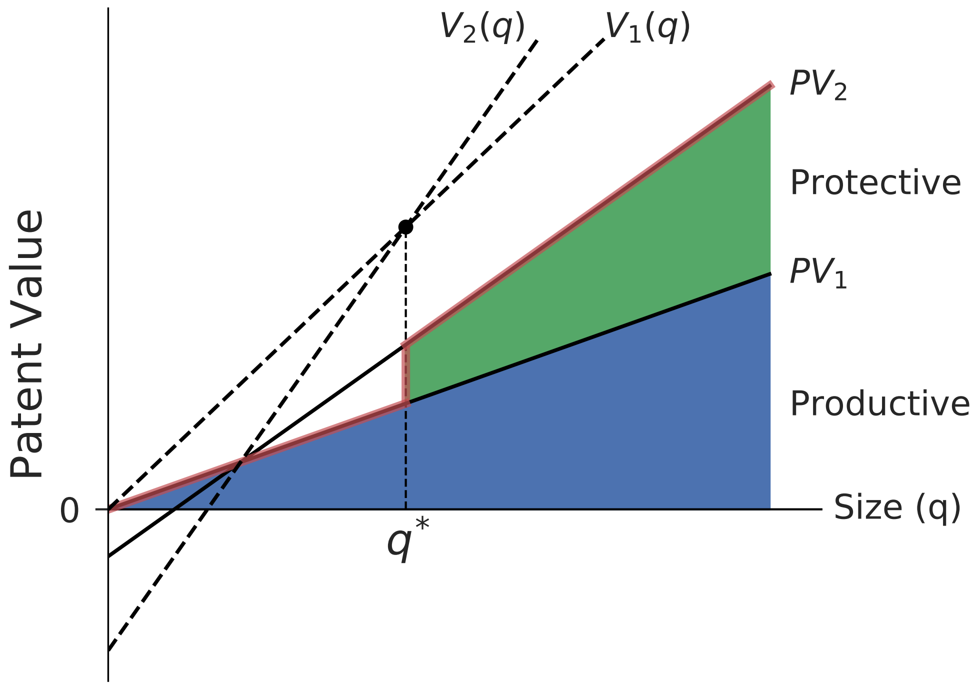
or

$$\begin{aligned} PV_2 &= 0.5 \left[V_2(q(1+\lambda)) - V_1(q) \right] \\ &= \frac{v_2(1+\lambda) - v_1}{2} q - \frac{c}{2(\rho + \tau/\gamma)} \end{aligned}$$

Assume $v_1 < \frac{v_2(1+\lambda)-v_1}{2}$ such that the following inequalities hold $v_1 \lambda < v_1 < \frac{v_2(1+\lambda)-v_1}{2} < v_2$.

To illustrate a change in patent value over firm size and its decomposition into productive and protective components, consider Figure 6.

FIGURE 6: FIRM SIZE AND VALUE OF A PATENT



Dashed lines are value functions $V_1(q)$ and $V_2(q)$ and their intersection defines q^*(10). Lines PV_1 and PV_2 are patent values as defined above. The red curve is the resulting patent value curve as a function of firm size. We see that it increases with size – firms gain more from their patents as they grow because patents protect higher levels of technology. However, larger firms enjoy even higher value from a patent that comes not just from a productive part of a patent (blue dashed area) but from a protective part (red dashed area). Notice that protective part increases with firm size and accounts for an ever increasing share of overall patent value to a firm – value of protection increases higher is the technology level that firm is protecting. Hence, we formulate our final predictions of the model:

Prediction 4: Value of a patent increases with firm size.

Prediction 5: Patent value of larger firms is driven more by protective role of patents rather than their productive role.

V Innovation, Patent Value, and Firm Heterogeneity: Empirics

Our theoretical model predicts that as firms grow, they shift their innovation more towards protective strategies. As a consequence, the average patent of a market leader is associated less with product innovation and more with competitors' deterrence, relative to smaller firms. Conceptually, we define the value of a patent as the revenue premium of products protected by it, and it is the sum of two components:

$$\text{Patent value} = \text{Productive value} + \text{Protective value}$$

where productive component reflects the additional revenue generated by a new product due to the technological improvements that are reflected in the patent that protects them. The protective component captures the additional revenue generated to the firm from the deterrence of competition (e.g. limiting entry of products) by the patent.

Finding the value of a patent and especially its components is a challenging task. Data limitations usually do not permit separating the amount of the revenue or profits of a firm generated by specific patents. More than that, it is hard to disentangle intrinsic quality component of the innovation from an anti-competitive role of a patent. In this section, we make progress in this direction by taking advantage of our data in linking specific patents to specific products and their revenue. We use it to understand if large firms change their innovation incentives and switch more to protective strategies.

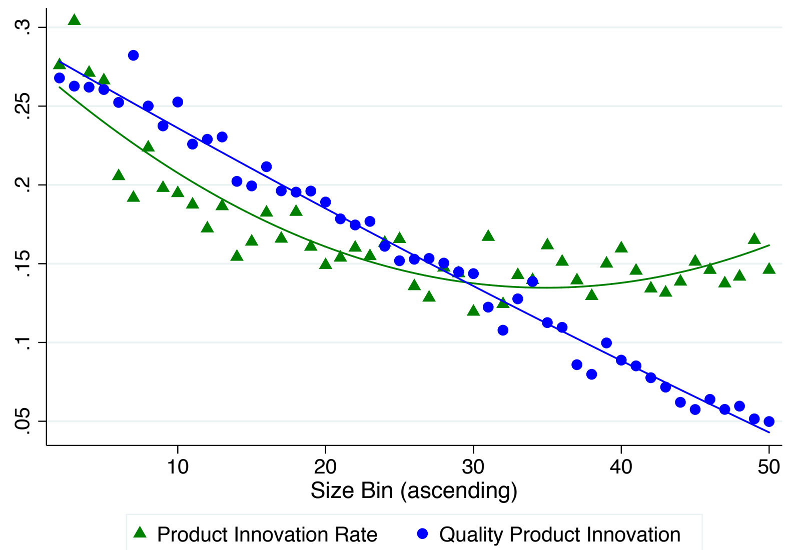
We start by evaluating how product innovation and patenting varies systematically with firm size. We use evidence on how the relationship between patents and product innovation changes systematically with firm size to learn about the role of patents in deterring future innovative activities. We then show how they affect competitors future innovation and the firms' future growth.

A The Nature of Innovation and Firm Size

Fact 2: *Larger firms have lower product innovation rates.*

Figure 7 shows that larger firms have lower product innovation rates, and their new products have lower levels of newness.

Figure 7: Product Innovation Rate by Size

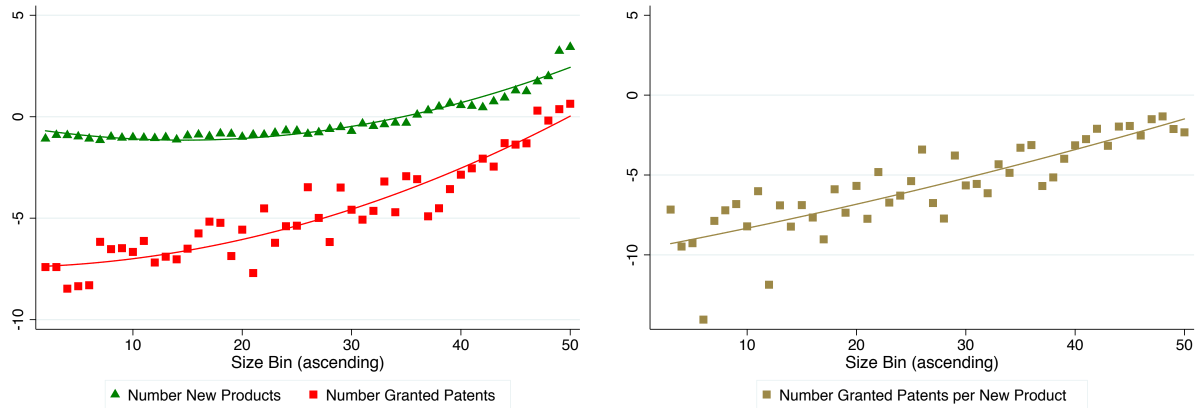


Notes: This figure plots the relationship between product innovation and size of the firm. We use data on product innovation at the firm x product category for the period 2007–2015. For each product category, we compute the average annual entry rate (new products divided by stock of products) and the average quality of the new products (defined by their newness), across 50 bins of size. The figure shows the average after weighting the different product categories by their revenue share.

In Section III we showed that there is a systematic relationship between product introduction and patenting. An increase in granted patents is associated with an increase in the quantity and quality of new products. We now study how this relationship changes with firm size. Figure 8 shows that both the number of new products and new patents increase with firm size, but the increase in patents is substantially larger. The ratio of new patent per new product is larger for larger firms.

Fact 3: *Larger firms have higher patents per new product, and this does not translate into higher-quality products.*

Figure 8: New Products and New Patents by Size



Notes: This figure plots the relationship between product innovation and patenting, and firm size. We use data on product innovation and patents at the firm x product category for the period 2007–2015. For each product category, we compute the number of new products, the number of new granted patents, and the number of new patents per product across 50 bins of size. The figure shows the logs of the average after weighting the different product categories by their revenue share.

This fact is consistent with predictions from our model where as firms grow, the elasticity of product innovation to patenting declines. We formalize this relationship in the next table. We use variation within firm × product category and time × product category according to the following specification:

$$E_{fmt} = \alpha + \beta Patent_{fmt-1} + \gamma Size_{fmt} + \lambda Patent_{fmt-1} \times Size_{fmt} + \eta_{fm} + \theta_{mt} + \epsilon_{jfmt}$$

where E_{fmt} denotes either the entry rate of products launched at time t by firm f in category m, or the average quality of the new products at time t by firm f in category m. $Patent_{fmt-1}$ refers to the number of new patents granted the previous year. $Size_{fmt}$ is the natural logarithm of the revenue of firm f in module m at time t. Table VI reports our results under several specifications. Patenting is strongly related to the rate at which firms introduce new products. Interestingly, $Size_{fmt}$ is positively associated with innovation rate and it is the interaction between $Patent_{fmt-1}$ and $Size_{fmt}$ that decreases the rate at which firms introduce new products. The table shows that patents are less likely to become new products as firms grow either because firms are able to enjoy the rents of their innovations for longer or because they are using their patents for different purposes. Our measurement of total patents and patents associated with product and process innovation, show that this result is even stronger when we only include product patents.

Table VI: Relationship Between Product Innovation Rate and Patenting by Firm Size

	Product Innovation Rate			Quality Product Innovation		
	(1)	(2)	(3)	(4)	(5)	(6)
Patents(t-1)	0.015***	0.013***	0.013**	0.003	0.003	0.012**
	(0.005)	(0.005)	(0.006)	(0.002)	(0.002)	(0.005)
Size(t)		0.055***	0.055***		-0.019***	-0.018***
		(0.001)	(0.001)		(0.001)	(0.001)
Patents(t-1) x Size(t)			0.000			-0.005**
			(0.003)			(0.002)
Observations	452,772	452,772	452,772	83,142	83,142	83,142
R-squared	0.348	0.356	0.356	0.519	0.522	0.522
Time-Category	Y	Y	Y	Y	Y	Y
Firm-Category	Y	Y	Y	Y	Y	Y

Notes: The table shows the relationship between the rate of product innovation, patenting and firm size. The product innovation rate is the rate of product introduction by firm f in category m at time t. The quality product innovation is the product innovation rate defined by their newness. Patent(t-1) is the natural logarithm of granted patents of product category m by firm f applied in year t, using the inverse hyperbolic sine transformation. Size(t) is the natural logarithm of the total sales of firm f in product category m at time t (standardized).

Despite the fact that larger firms are on average less active at launching new products to the market than smaller firms, it could be the case that the products they launch are more novel and the quality of their innovations higher. To measure the amount of innovation contained in each new product, we go back to our innovation quality measures defined in Equation 11. Our results (Table VI) show that larger firms introduce less novelty products. The average quality of new products declines with firm size, and the relationship between patenting and product quality is lower among larger firms. This suggests that while larger firms do relatively more patenting, this does not translate into higher-quality products.

Given that larger firms engage more in patenting, and we do not find evidence that this translated into more or higher-quality product innovation, we evaluate the impact of patenting on competitors' product introduction. Are larger firms using their patents to deter the entry of competitors rather than to innovate? We establish that:

Fact 4: *Patents of larger firms deter competitors' new product introduction more.*

We use the product introduction of other firms within the same product category to study this question. We are interested in exploring whether the product innovation rate of competitors in a given market declines when the market leader introduces a product that is patented. We use the following specification:

$$N_{m_{-f}t} = \alpha + \delta N_{fmt-1} + \beta Pat_{fmt-1} + \gamma Size_{fmt} + \eta N_{fmt-1} \times Size_{mt} + \lambda Pat_{fmt-1} \times Size_{mt} + \epsilon_{jfmt}$$

where $N_{m_{-f}t}$ is the natural logarithm of the number of products introduced by firms other than firm f in module m at time t, N_{fmt} is the natural logarithm of the products introduced by firm f in module m at time t, Pat_{fmt-1} is the natural logarithm of the patents granted by firm f in module m at time t. Our coefficient of interest is λ on the interaction between the size of the firm and the number of granted patents, after accounting for the effect of new products on product introduction by others. Table VII shows our results under several specifications. Column 1 shows that both the introduction of new products and patenting reduce the product innovation rate of the rest of the firms within the same product category. Columns 3-5 show that patents launched by larger firms unambiguously decrease the rate of product introduction of other firms. Our preferred specification to study whether the behavior of market leaders affects the innovation rate of their competitors is shown in column 5. It shows that, controlling for product category $\times$ time and firm $\times$ time effects, new products launched by larger firms have a negative effect on the rate at which competitors introduce new products. Protected innovations by larger firms have the opposite effect, as indicated by the sign of μ. In other words, patents deter innovation of competitors but only if they belong to larger firms. This is particularly important when larger firms introduce products that are patented. Overall, our findings suggest that the value of a patent for a market leader may come from the deterrence of competitors and that an additional revenue generated by leading incumbents' products comes in no small part from this competitive margin, rather than from protecting and launching high-quality innovative products.

Table VII: Deterrence, Patenting, and Market Leadership

	(1)	(2)	(3)	(4)	(5)
New Products by Other Firms					
Log N(t-1)	-0.000***	-0.000***	0.001***	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Patents(t-1)		-0.004***		-0.001	-0.006***
		(0.001)		(0.001)	(0.002)
Size(t)			-0.002***	-0.002***	-0.002***
			(0.000)	(0.000)	(0.000)
Log N(t-1) × Size(t)			-0.002***	-0.002***	-0.008***
			(0.000)	(0.000)	(0.000)
Patents(t-1) × Size(t)				-0.003***	-0.008***
				(0.001)	(0.001)
Observations	389,192	389,192	389,192	389,192	307,417
R-squared	1.000	1.000	1.000	1.000	1.000
Time-Category	Y	Y	Y	Y	Y
Firm-Category	Y	Y	Y	Y	N
Time-Firm	N	N	N	N	Y

Notes: The table shows the relationship between product introduction of other firms $\hat{f}$ within the product categories m and the size of firm f, the product introduction of firm f , and whether these products are related to a patent. The dependent variable is the natural logarithm of the number of products introduced by firms other than firm f. $Log\ N(t-1)$ is the natural logarithm of products introduced by firm f, using the inverse hyperbolic sine transformation. $Patents(t-1)$ is the natural logarithm of granted patents of product category m by firm f applied in year t, using the inverse hyperbolic sine transformation. $Size(t)$ is the natural logarithm of the total sales of firm f in product category m at time t (standardized).

VI Conclusion

tbc

References

AB, Lerner Js Jaffe, "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.

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

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A Robinson, "Democracy does cause growth," *Journal of Political Economy*, 2019, *127* (1), 000–000.

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

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

_ , 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," Working Paper 25136, National Bureau of Economic Research 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, "How do firms grow? The life cycle of products matters," 2018.

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.

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?," Working Paper 23782, National Bureau of Economic Research 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.

Boldrin, Michele and David K. Levine, "The Case against Patents," *Journal of Economic Perspectives*, February 2013, *27* (1), 3–22.

Broda, Christian and David E Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 2010, *100* (3), 691–723.

Cavenaile, Laurent and Pau Roldan, "Advertising, innovation and economic growth," Working Papers 1902, Banco de EspaÃśa;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.

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

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.

Furman, Jeffrey L. and Scott Stern, "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," *American Economic Review*, August 2011, *101* (5), 1933–63.

Gambardella, Alfonso, Dietmar Harhoff, and Bart Verspagen, "The value of European patents," *European Management Review*, *5* (2), 69–84.

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

Gutierrez, GermÃąn 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," Working Paper 8498, National Bureau of Economic Research 2001.

_ , _ , and _ , "Market Value and Patent Citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

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

Harhoff, Dietmar, Francis Narin, Frederic M. Scherer, and Katrin Vopel, "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 1999, *81* (3), 511–515.

_ , Frederic M. Scherer, and Katrin Vopel, "Citations, family size, opposition and the value of patent rights," *Research Policy*, 2003, *32* (8), 1343–1363.

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

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

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," Working Paper 17769, National Bureau of Economic Research 2012.

Lampe, Ryan and Petra Moser, "Patent Pools, Competition, and InnovationâĂŤEvidence 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.

Lanjouw, Jean Olson, "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value," *The Review of Economic Studies*, 10 1998, *65* (4), 671–710.

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.

_ and _ , "Global Market Power," Working Paper 24768, National Bureau of Economic Research June 2018.

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

_ , Joerg Ohmstedt, and Paul Rhode, "Patent Citations and the Size of the Inventive Step - Evidence from Hybrid Corn," Working Paper 21443, NBER July 2015.

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.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration," Working Paper 25066, National Bureau of Economic Research September 2018.

Sampat, Bhaven and Heidi L. Williams, "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome," *American Economic Review*, January 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.

Schoenle, Raphael, "International menu costs and price dynamics," *Review of International Economics*, 2017, *25* (3), 578–606.

Scotchmer, Suzanne, "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *The Journal of Economic Perspectives*, 1991, *5* (1), 29–41.

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

Toivanen, Otto, Paul Stoneman, and Derek Bosworth, "Innovation and the Market Value of UK Firms, 1989âĂŞ1995*," *Oxford Bulletin of Economics and Statistics*, 2002, *64* (1), 39–61.

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.

A Data Appendix

A.1 Patent assignment

To assign patents to their most recent firms, we proceed in the following steps. First, we assign each patent to a *current* assignee(s) (as of 2017 – our patent data vintage). Sometimes this information is missing; in such a case we take the name of an *original* assignee and reassign its patents in case of corporate reorganization. Since most of the time patents get reassigned when a firm is acquired, we track these merger and acquisitions using the Thomson Reuters Mergers & Acquisition data. This is particularly important given that firms that appear both in Nielsen data and USPTO are most likely large firms that undergo many corporate reorganizations.

A.2 Company name cleaning algorithm

We assign each company name to a unique company identifier by using the following procedure that builds on and extends cleaning algorithms from the NBER Patent Data Project and Akcigit et al. (2016).

Step 1. In the first step, we run all company names through name standardization routine and generate unique company identifiers.

1. After capitalizing all letter, we keep the first part of the company name before the first comma.
2. We remove leading and trailing "THE" words; replace different spellings of "AND" words with "&"; and replace accented or acute letters with regular ones.
3. We remove special characters.
4. We standardize frequent abbreviations using dictionaries from the NBER Patent Data Project. For example "PUBLIC LIMITED" or "PUBLIC LIABILITY COMPANY" become "PLC";"ASSOCIATES" or "ASSOCIATE" become "ASSOC"; "CENTER" or "CENTRAL" become "CENT".
5. We delete trailing company identifiers
6. If resulting string is null, we protect it.

7. We redo previous steps on the original company names except for protected strings, for which we now keep the whole string and not just the first portion before comma.
8. If string is protected, we remove company identifiers in any place of the string (not just if trailing as in 5.)
9. We remove spaces to further decrease misspellings.
10. We assign unique company identifiers based on cleaned names from 9.

Step 2. In addition to the extensive cleaning from Step 1, we take advantage of a “dictionary” that resulted from the large effort conducted within the NBER Patent Data Project. In particular, after manual checks and searches of various company directories to identify name misspellings as well as various company reorganization, the NBER files provide mapping between patent assignee names and a company identifier (*pdpass*). Although this data is based on assignees of granted patents before 2006, we use this mapping as a “dictionary” that we combine with our results from Step 1. This helps us leverage both on our algorithm from Step 1 and NBER *pdpass* information to combine strengths of each method to create the new unique company identifier.

For example, Siemens has many different variations of its name in the data. ”SIEMNES AG”, ”SIEMANS ATKIENGESELLSCHAFT”, and ”SIEKENS AG” are just few of such variations that Step 1 does not capture but the NBER files identify as names under the same *pdpass*. In such a case, we will use *pdpass* identifiers to group these three firms together. On the other hand, the NBER file does not identify ”SIEMENS CORP” ”SIEMENS AG” and ”SIEMENS” as the same company and the same as the first three name variations above. In such a case, we use our unique identifiers from Step 1 to group these firms together. As a result, after combining information from NBER files with our cleaning after Step 1, we pool all these six variations into one new company code.

A.3 Document vector similarity metric

For each patent and each product category, we have a text document characterizing them. These are then converted to word vectors, which indicate, for each word, how many times it appears in a documet. Each document vector is of length M, which is the number of tokens that we include in our vocabulary.[17] We use a vocabulary consisting of any words that appear in the Wikipedia entries we consider excluding those in the top 1% of word frequencies (like "the" or "and"). The corpus of documents can be represented by a very sparse matrix c_{ij} of word counts, where $i \in \{1, \ldots, N\} = \mathcal{N}$ represents the document and $j \in \{1, \ldots, K\} = \mathcal{M}$ represents the word.

We also utilize a word-based weighting scheme called total-frequency-inverse-document-frequency (tf-idf) to account for the fact that more common words tend to be of less importance and vice versa. There are a number of possible forms to use here, but we choose the most commonly used

$$w_j = \log\left(\frac{N+1}{d_j+1}\right) + 1 \quad \text{where} \quad d_j = |\{i \in \mathcal{N} | c_{ij} > 0\}|$$

Thus if a word appears in all documents, it gets a weight of one, while those appearing in fewer documents get larger weights, and this relationship is sublinear. For our weighting scheme, we use document frequencies from the patent data, which contains far more documents than the product category side.

Constructing representative documents on the patent side consists of simply concatenating all of the available text into one document. On the product module side, we must aggregate the various Wikipedia entries. In this particular setting, we first vectorize each Wikipedia entry, then average these vectors together (in an an ℓ^2-norm-preserving sense) so as not to overweight longer entries. Additionally, we repeat the first 10% of each Wikipedia entry 10 times (an approximation of the introduction) to emphasize introductory material. Finally, before vectorization, we run each document through a lemmatizer [18], which reduces words to their root form by removing conjugation.

Finally, we are left with a weighted, ℓ^2-normalized word frequency vector for each document, both on the patent and product side. Specifically, these are defined as

$$f_{ij} = \frac{w_j c_{ij}}{\sqrt{\sum_{j'} (w_j c_{ij'})^2}}$$

Multiplying any two such vectors together yields a similarity metric between two documents.

[17] For this exercise, we use 1-grams (a.k.a. words) as tokens. In general one could use n-grams, meaning distinct n-length phrases.

[18] Specifically, the WordNetLemmatizer provided as part of the NLTK (nltk.org) Python module, which utilizes the WordNet lexical database (wordnet.princeton.edu)

This is guaranteed to be in the range $[0, 1]$ with zero corresponding to zero word overlap and one corresponding to the case in which the documents are identical (or are multiples of one another). Notice that this vectorization approach (sometimes referred to as "bag of words") ignores any information about the order of words or even phrases.

B Newness

In order to quantify the novelty of a product, we follow Argente and Yeh (2017) and construct a *newness index* that uses detailed information about the characteristics of each UPC provided in the Nielsen RMS data set. The index counts the number of new and unique attributes a product has at the time of its introduction relative to all of the other products ever sold within the same product module. In contrast to Argente and Yeh (2017) who construct the newness index to capture the novelty of a product from a store's perspective, our measure assigns a higher value to products with more unknown features to the entire market. We define a product j in product module k as a vector of characteristics $V_{kj} = [v_{j1}, v_{j2}, .., v_{jN_k}]$ where N_k denotes the number of attributes we observe in product module k in our data. For example, the product module "soft drinks - carbonated" consists of $N_{\text{soft drinks}} = 8$ attributes for each barcode: brand, flavor, firm, size, type (sparkling soda or natural soda), container (e.g. can or bottle), formula, generic (i.e. private label). Let Ω_{kt} contain the set of product characteristics for each product ever sold in product module k at time t, then the *newness index* of a product j in product module k, launched at time t is defined as follows:

$$\text{NI}_{jt}^{k} = \frac{1}{N_k} \sum_{i=1}^{N_k} \mathbb{1}[v_{kji} \notin \Omega_{kt}]. \tag{11}$$

For example, if a new product within the soft drinks category enters with a flavor and size that has never been sold in any store before, its newness index is $(1+1)/N_{\text{soft drinks}} = 2/8$. On average, we observe 7.2 product characteristics in each product module.[19] We assume that each attribute is equally weighted in order to remain agnostic about the relative importance of each attribute to the degree of newness of a product. Figure A.1 shows the average of the

[19] Comparing the newness index of different products across distinct modules depends not only on the number of new attributes of each product but also on the total amount of observable characteristics the Nielsen data provides for each module. The minimum characteristics we observe for each module is 5 and the maximum is 12.

index for each of the sectors in our data and figure A.2 shows the most common attributes with new characteristics for a sample of sectors. The figure shows, for example, that the most common attribute of new products included in the newness index for the product group carbonated beverages is brands.

Our basic newness index, counts the number of new characteristics (within each attribute) each new product brings to the market. Alternatively, we could determined if a product brings a combination of characteristics for the first time to the market. For example, in the carbonated beverages, a new soft drink will not be the first fusing drink or the first soda with cherry flavor, but the first cherry flavor fusing drink. To account for new products that do not appear with a new characteristic (within a given attribute) but rather a new combination of characteristics, we developed an alternative index. Let c_{jkt} denote the combination of characteristics of UPC j (e.g. container: can, flavor: orange, brand: Coca-Cola, formula: no calories) and let Θ_{kt} contain all the possible combinations of attributes observe up until period t in product module k. Using this information we define an alternative newness index that not only accounts for new characteristics within product attributes but also for new combinations of characteristics:

$$\text{NIC}_{jt}^{k} = \frac{1}{N_k + 1}\left[\sum_{i=1}^{N_k} \mathbb{1}[v_{kji} \notin \Omega_{kt}] + \mathbb{1}[c_{kj} \notin \Theta_{kt}]\right] \tag{12}$$

In this case, we add to the numerator of the baseline newness index one if a new product brings a new combination of characteristics to the market. By definition, if a new product brings a new characteristic within an attribute, it will also bring a new combination. However, new products that may not bring a new characteristic within an attribute to the market might bring a new combination if characteristics (e.g. not the first time we observe a soda with orange flavor or the first time we observe a low calorie soda, but the first time we observe and orange flavor low calorie soda). The advantage of this index, relative to our baseline index, is that it differentiates more clearly products that bring new characteristics to the market (even if marginal) with those that do not.

Figure A.1 reports the fraction of new products that enter the market with a new combination of product characteristics. Overall, 66% of new products enter the market with a new combination of attributes covered by the Nielsen RMS.

Figure A.1: Newness Index by Product Group

Figures A.1 presents the average newness index for a sample of sectors in our data. In particular, it shows the mean newness index by sectors along with the top and bottom sectors as ranked by this measure. We compute the newness index for each product using equation 11. We average across products and product modules to the group level. We focus on cohorts from 2006Q3 to 2014Q4 and on modules with at least 20 barcodes.

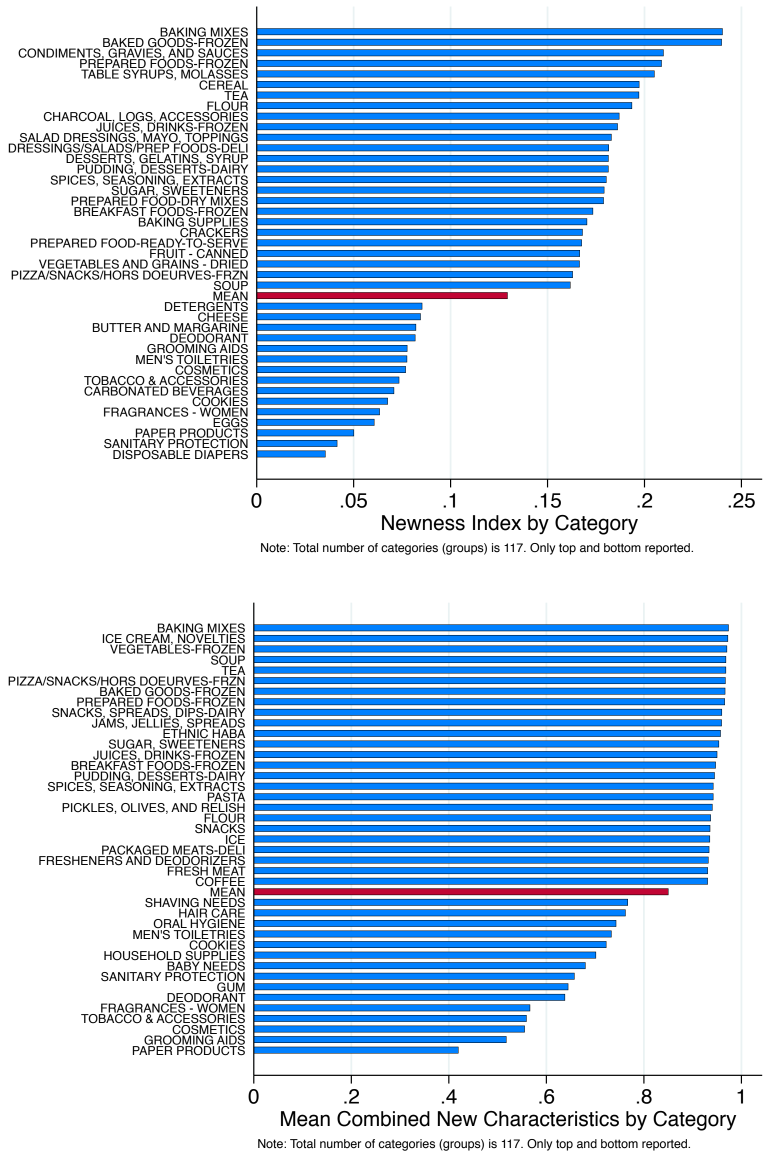
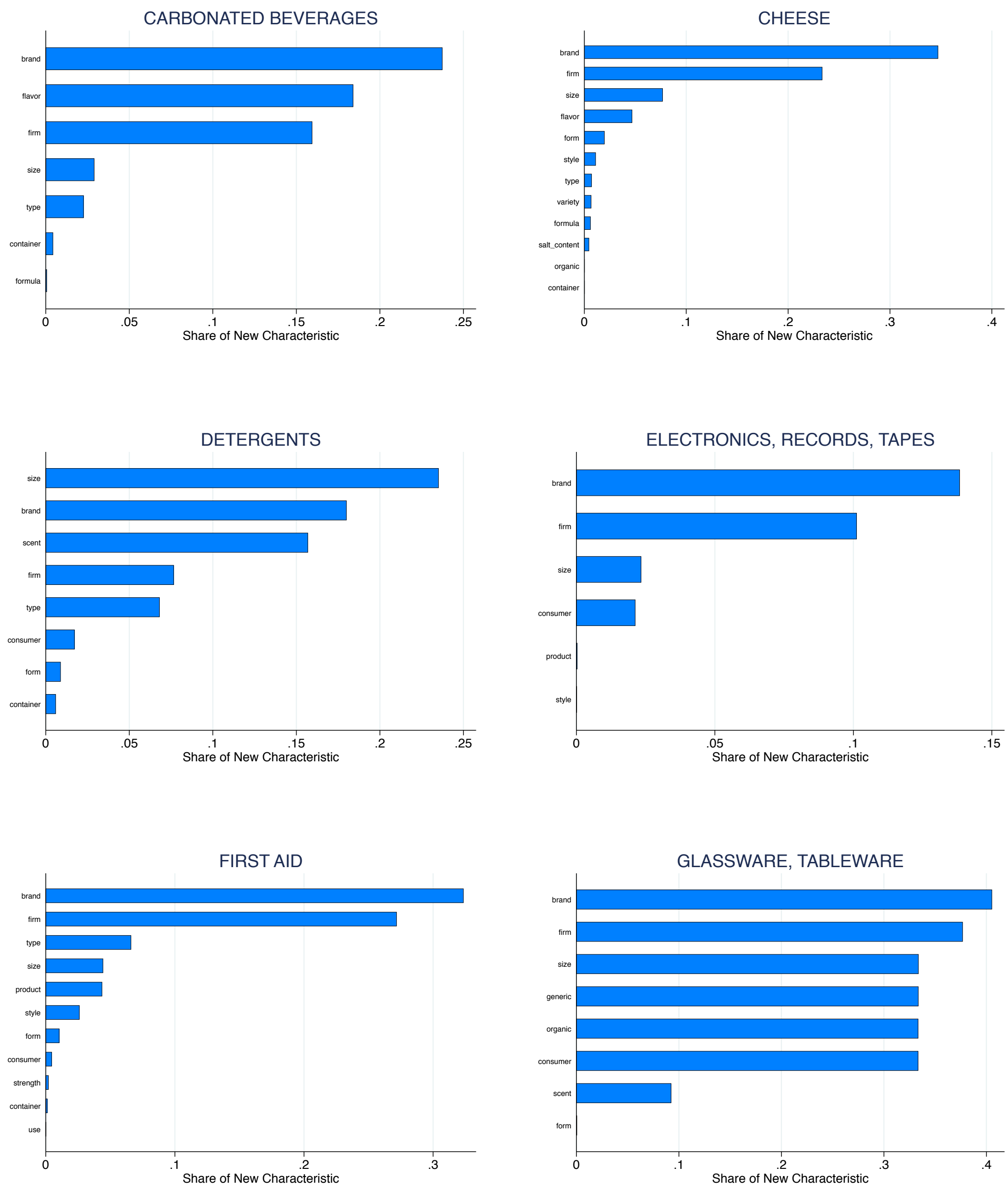


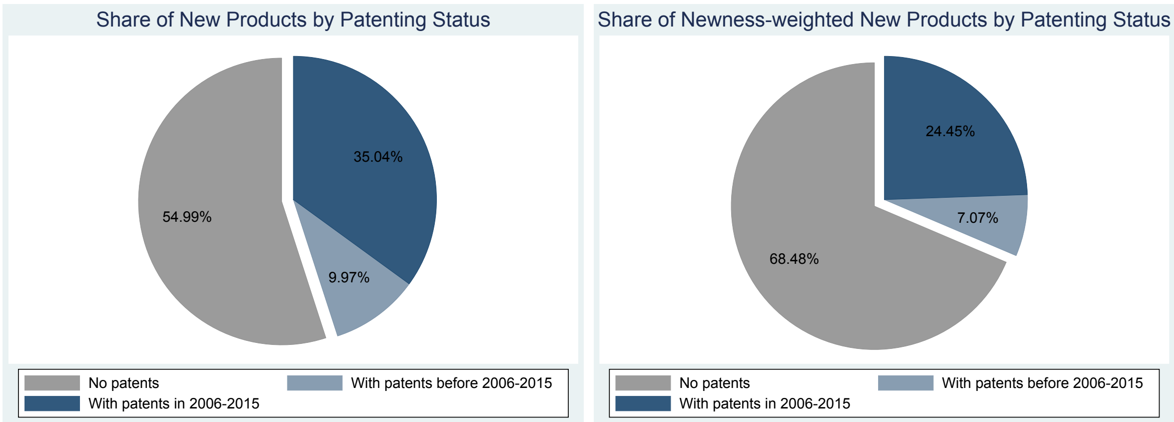
Figure A.2: New Product Characteristics by Product Group

Figures A.2 presents the most common new characteristics appearing in new products for a sample of sectors: carbonated beverages, cheese, detergents, first aid, hardware tools, and electronics, records and tapes. For each product attribute we construct an indicator if it is the first time the attribute appears in a product. Then we average across products within a product module and then across modules to the group level. We focus on cohorts from 2006Q3 to 2014Q4.



C Appendix Tables and Figures

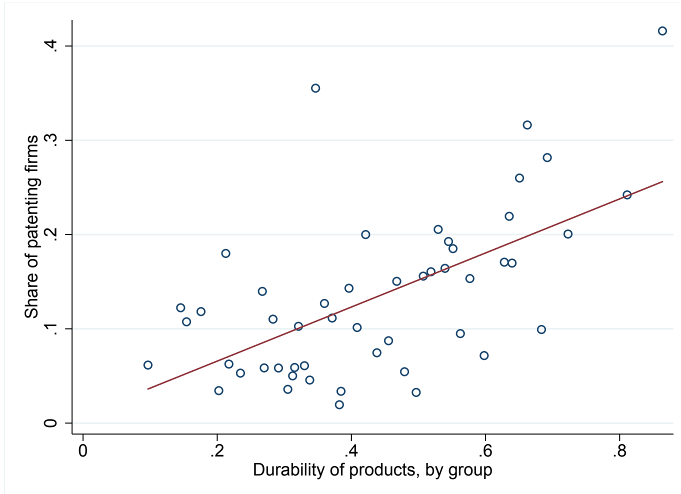
Figure A.3: Share of Product Innovation by Patenting Status



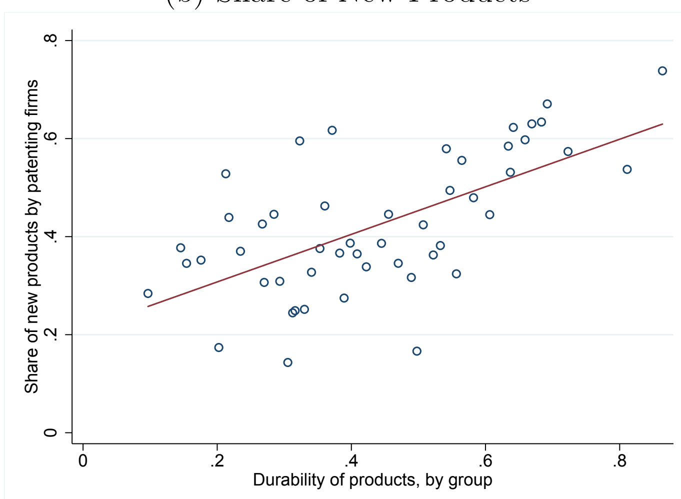
Notes: the figure shows the share of new products and quality(newness)-weighted new products by firms' patenting status. Panel (a) shows the share of entering products launched by firms without patents, with all patents before 2006, and with patents between 2006-2015. Panel (b) shows the newness-adjusted product entry shares accounted by firms with theose patenting statuses.

Figure A.4: Product Durability and Firms' Patenting

(a) Share of Patenting Firms

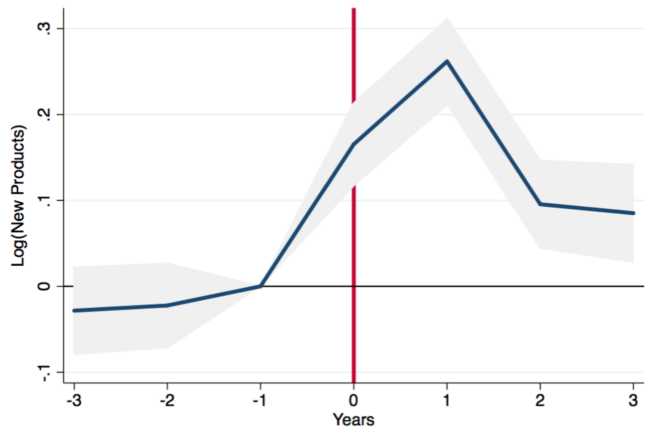


(b) Share of New Products



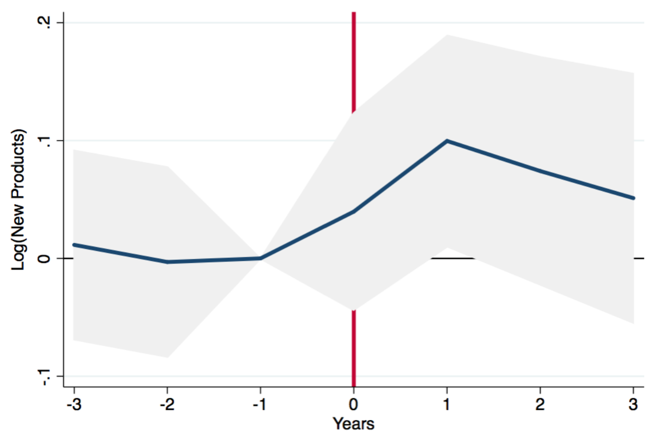
Note: the figure shows the share of patenting firms and the share of new product launched by patenting firms in each product group of the Nielsen data and their relationship with the durability of the products in each group. In order to approximate the durability of each product group we use the Nielsen Consumer Panel Data and count the average number of shopping trips made by households in a given year to purchase products in each product group. We call categories with few trips per year durable categories. Our measure of durability is the inverse of the average number of trips in a given product group.

Figure A.5: Event Study: Patenting Status (Firm level)



Notes: The figure plots the estimated coefficients of the impact of switching to patenting status on new products. The observations are at the firm × year level. The coefficients were estimated using an event study approach.

Figure A.6: Event Study: Patenting Status (Firm × product category level)



Notes: The figure plots the estimated coefficients of the impact of switching to patenting status on new products. The observations are at the firm × product category × year level. The coefficients were estimated using an event study approach.