Innovation and Firm Productivity: Evidence from the US Patent Data*

Jingbo Cui and Xiaogang Li†

Selected Paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-August 2

Abstract

In this paper, we examine the relationship between productivity and innovation, using the U.S. manufacturers' patent data from 1976-2006. First, we investigate whether productive firms actively participate in innovation in terms of having more patents, and then examine whether their innovation activities are involved in a wide spectrum of technological categories. Moreover, we are interested in, to successfully develop a new patent in a certain technological field, whether productive firms need to cite more or less patents within the field and/or across various related fields. The firm-level productivity is estimated as the total factor productivity (TFP). We find that: (i) productivity is positively correlated with the number of patents granted and the number of technological categories for these patents; and (ii) productivity is positive correlated with the number of citations per granted patent, and is also positively correlated with the number of technological categories for cited patents per granted patent. Whereas the former finding indicates that productive firms actively conduct research and their innovation is involved in different technological fields, the latter suggests that, to develop a new patent, productive firms are capable of learning from cited patents in various technologies fields.

Keywords: Productivity, Backward Citations, Innovation, Knowledge Stock
JEL Classification: D22, O31

†Jingbo Cui is an assistant professor in the Economics and Management School of Wuhan University, China. Xiaogang Li is a Ph.D. student in the Department of Economics at Iowa State University, the United States. The corresponding author: Jingbo Cui, email: jbcui2013@gmail.com, personal website: https://sites.google.com/site/jbcui2013/. We would like to thank the Business School at Iowa State University for providing access to Wharton Research Data Services when Jingbo Cui was a Post-doctoral Research Associate in the Center for Agricultural and Rural Development at Iowa State University.

1 Introduction

The quest to explain the sources of productivity growth has led to a large body of empirical literature that investigates the relationship between firm-level productivity and innovative activities by using patent data. Early work suggested a positive correlation between productivity and innovation, but was less conclusive about how firms' productivity is related with the spectrum of technological categories associated with patents. The process of producing a patentable innovation is similar to the combination of several pieces of familiar and unfamiliar technological components (Fleming, 2001). Entering in an unfamiliar technological field would incur additional costs to overcome technical barriers. A firm's productivity may be associated with the probability of choosing to bear such costs, and may be important in expanding the spectrum of technological fields and spurring innovation by combining new ideas and inventions.

In this paper we examine the empirical relationship between firm-level productivity and innovation from a variety of perspectives. First, we investigate whether productive firms actively participate in innovation in terms of patents. Second, we examine whether firms' patents are involved in a wide spectrum of technological categories (extensive innovation scope), and whether their innovative activities focus on inventing within a narrow set of fields (intensive innovation scope). Third, we investigate whether, to successfully develop a patentable innovation in a certain technological field, productive firms cite patents across various related fields (extensive citation scope) or whether they only cite more patents from a narrow set of technological categories (intensive citation scope). Lastly, this paper explores how the novelty of invented patents is correlated with firm-level productivity.

The data that we use pertain to US public firms in the manufacturing sector over the period 1976-2006. This dataset is compiled from two main sources: the NBER patent database, and Compustat data (Wharton Research Data Services). The former provides comprehensive information for patents granted by the US Patent and Trademark Office, while the latter includes detailed financial information for all firms traded in the US stock market during the study period. Since the main focus of this paper is on the manufacturing sector, we further restrict our sample to industries within this sector, and match the data with the industry-level NBER-CES manufacturing productivity database.

Productivity, a key variable in our analysis, is estimated as total factor productivity (TFP), using the Levinsohn and Petrin (2003)'s method. To link productivity with innovation, our dependent variables include measures of innovation quantity measured by the number of patents. Each patent is classified into technological fields by using the International Patent Classification (IPC) code. We use technological categories associated with patents to define extensive and intensive innovation scope, capturing the technology spectrum in which a firm has produced patentable inventions. The extensive innovation scope is simply proxied by the number of different technological fields associated with these patents, while the intensive innovation scope is the number of different technological fields divided by the number of patents. Similarly, we count the number of different technological fields for cited patents, and divide it by the number of patents. This measure captures the number of cited technical fields per patent that a firm learns from cited patents. Furthermore, two additional measures on the "generality" and "originality" of patents are used, following Trajtenberg et al. (1997).

Our findings are consistent with the empirical literature on the relationship between firm-level productivity and innovation (Hall, 2011; Mohnen and Hall, 2013). First, we document a consistent and positive correlation between productivity and innovation quantity – the more productive a firm, the more patentable inventions are produced. Second, there are some novel insights about how productivity is related to extensive and intensive innovation scope at the firm level. A positive correlation between productivity and extensive innovation scope is documented. Productive firms tend to engage in a wide spectrum of technological fields. Moreover, intensive innovation scope is also positively correlated with productivity, indicating that on average productive firms appear to invent relatively less patents in a wide area of technical fields. Lastly, our findings show that productivity is positively correlated with both average citations per patent and average cited technical fields per patent. The former indicates productive firms tend to cite more patents per patented invention, while the latter suggests that these productive firms are more capable of learning from cited patents across various technologies fields. Furthermore, the "generality" and "originality" measures of novel patents are positively correlated with firm productivity.

This paper is related to the growing literature that study innovation using firm-level patent. One strand is the (causal) relationship between productivity and patent innovation. Recent review papers (Nagaoka et al., 2010; Hall, 2011; Mohnen and Hall, 2013) summarize the pioneering work by Griliches (1996, 1998), which attempt to employ patents stock to explain the residual growth in productivity. Most of the followed-up studies that have examined the effects of innovation on labor productivity are based on the well-known CDM (Crepon, Duguet, and Mairesse, 1998) model. This model is generally presented as a recursive system of three blocks of equations with endogenous choices of innovation, and hence handles some of the endogeneity problem related to R&D expenditures and product or process innovation. Whereas a few studies document a negative correlation between innovation and productivity (Roper et al., 2008; Van Leeuwen and Klomp, 2006), this strand of work have found the positive relationship between innovation and productivity (Crepon et al., 1998; Janz et al., 2004; Loof and Heshmati, 2006; Castellani and Zanfei, 2007; Hall et al., 2009; Raymond et al., 2013; Hall and Sena, 2014). When the growth rather than the level of productivity is chosen as the dependent variable (Geroski, 1989; Miguel Benavente, 2006) or human capital is added as a control variable (Crepon et al., 1998; Castellani and Zanfei, 2007; Therrien and Hanel, 2009), the effects of innovation on productivity is weakened.

This paper is also related to existing literature that use the matched data between NBER patent database and firm-level financial information. Among the first wave of practitioners that link NBER patent with firms' innovation expenditure, Bound et al. (1982); Hall et al. (1984) analyze the R&D and patenting behavior of manufacturing firms. Hall et al. (2001) summarize how NBER patent data are merged and matched with public firms in Compustat. Using this matched firm-level patent data, one of emerging studies by Hall and Helmers (2010) examines the role of patent protection in technical changes, and another work by Hall and Harhoff (2012) investigate whether patent protection encourages technology transfer from dirty to clean. Arora et al. (2015) emphasizes the decline of scientific publications in corporate R&D.

The remainder of the paper is organized as follows. The next section describes data sources, data construction and description. The empirical strategy, results and robustness checks are provided in section 3. The last section concludes the paper.

2 Data

2.1 Data Sources

The data pertain to the US public firms in the manufacturing sector over the period of 1976 - 2006. We compile the data from two main sources: the NBER patent database and the Compustat data. The former provides comprehensive information for patents granted by the US Patent and Trademark Office (USPTO), while the latter maintained by Wharton Research Data Services (WRDS) includes detail financial information for all firms traded in the U.S. stock market during the study period. The industry-level variables that are used to construct variables of interests are obtained from the NBER-CES manufacturing industry database (Becker et al., 2013).

The NBER patent data comprise detail information on US patents granted from 1976 - 2006, and all citations associated with these patents.[1] The patent information file records patent application year, granted year, technological classification in terms of 8-digit IPC code, and the unique assignee identification number. There are about 2.87 million patents associated with 0.20 million firms in this original database. Moreover, for each granted

[1] Please see the data from the new NBER data project via link https://sites.google.com/site/patentdataproject/Home.

patent, the patent citing file contains information about all citations made to each patent (backward citations) and total cites received by patents (forward citations) issued during 1976-2006. There are about 23 million citation observations with 2.8 million citing patents and 2.5 million cited patents.

The US Compustat database provides historical archive of annual financial statement and balance sheet for all firms traded in the US stock market from 1975 to 2007. The firm-level variables include gvkey as an identification code, standard industry classification (SIC), net sales, number of employees, book value of asset value, capital expenditure, wage, R&D expenses, and may others. We use this firm-level data to estimate TFP, and construct other firm-level variables of interest, including capital intensity, firm age, size, and R&D expenditure per value of assets. The construction method mainly follows Keller and Yeaple (2009), which will be discussed in details in the following subsection of variable construction.

Bessen (2009) has matched patent data with public firms in Compustat, using an algorithm involved with firm names.[2] As owners of patent change over time, the patent and Compustat matched file provides dynamic match of patent assignee to corporate entity in the Compustat. The matching results uniquely link assignee identification number from patent data with public firms' permanent identification number (i.e., gvkey) in Compustat database. In the matched sample, there are around 1.8 million patents associated with 6,575 unique corporate identities and 132,585 firm-by-year observations.

Since the main focus of this paper is on the manufacturing sector, we further restrict our sample within this industry, and match with the industry-level NBER-CES manufacturing productivity database. This step of data matching and merging leads to a final sample of 52,055 firm-by-year observations with 3,182 unique firms across the study period from 1976 to 2006. These firms have 528,620 granted patents with over 5 million citations.

[2]For the matching algorithm and matched results, please refer to the new NBER patent project via link https://sites.google.com/site/patentdataproject/Home.

2.2 Variables Construction

Table 1 presents the definition and construction of each dependent and independent variables of interests used in our empiric work.

2.2.1 Innovation

The main dependent variable is innovation. The existing studies in related literature have employed patents to measure firm-level innovation output (Griliches, 1998; Hall et al., 2001; Amore and Morten, 2011; Aghion et al., 2013). Following this strand, we use the log number of patents granted to proxy firm's innovation quantity,

$$\text{Innovation}_{it} = log(1 + \# \text{ of Patent}_{it})$$

Each patent is assigned with a 8-digit IPC code provided by the World Intellectual Property Organization (WIPO). The IPC separates the whole body of technical knowledge using a hierarchical classification system, i.e., section, class, subclass, group and subgroup, in descending order of hierarchy. There are eight sections with each section symbol designated by one of the capital letters A through H. Each section is subdivided into classes which are the second hierarchical level of the IPC. The class symbols consist of the section symbol followed by a 2-digit number. Each class comprises several subclasses with the symbol of the class symbol followed by a capital letter. Lastly, each subclass is then broken down into groups, either main groups or subgroups. Take one IPC code for example, $H01S3/00$ is the code for "Lasers", where "H" denotes the section of Electricity, "H01" stands for the class of "Basic Electric Elements", while "H01S" references the subclass of "Devices using Stimulated Emission." The IPC has been periodically revised to take account of technical improvement with minor adjustments for the definition of class and subclass technology fields. Using the 2006 edition of the IPC, the whole body of technology spectrum is divided into 129 classes and 633 subclasses.[3] We use 4-digit IPC code (i.e., the subclass) to define technological fields. For each year, we sum up the number of different technological categories in which a firm's patents are classified. Hence, the *extensive innovation scope* is measured by the log number of unique technological fields, capturing the wide spectrum of a firms' technological categories in which it has invented patents.

$$\text{Extensive Innovation Scope}_{it} = log(1 + \#\text{ of TechField}_{it})$$

To further understand the intensiveness of a firms' technological spectrum, we measure the *intensive innovation scope* by the log number of unique technological fields divided by the number of patent,

$$\text{Intensive Innovation Scope}_{it} = log(1 + \frac{\#\text{ of TechField}_{it}}{\#\text{ of Patent}_{it}})$$

Definition of technology categories is of importance to our study. To check whether results are robust to the alternative IPC code, we use the first 3-digit IPC code (i.e., class) to separate technology categories from each other. An alternative measure of extensive innovation scope is the count of different technological fields associated with patents in terms of the first 3-digit IPC code. Thus, $\text{Alt. Extensive Innovation Scope}_{it} = log(1 + \#\text{ of Alt. Tech Field}_{it})$. Similarly, an alternative measure of intensive innovation scope is defined as $\text{Alt. Intensive Innovation Scope}_{it} = log(1 + \frac{\#\text{ of Alt. TechField}_{it}}{\#\text{ of Patent}_{it}})$.

Forward citations appear to be correlated with the value of patents (Trajtenberg, 1987), whereas backward citations reflect knowledge spillover of "prior art."[4] For each firm,

[3] Please refer to the summary statistics table provided by the WIPO http://www.wipo.int/classifications/ipc/en/ITsupport/Version20060101/transformations/stats.html.

[4] Patent examiners, rather than applicants, are ultimately responsible for for the citations made(Hall et al., 2001; Nagaoka et al., 2010). Unfortunately, we could not distinguish citations added by patent examiners with citations referenced by inventors.

we calculate average backward citations per patent, that is,

$$\text{Average Citation}_{it} = log(1 + \frac{\text{\# of Citation}_{it}}{\text{\# of Patent}_{it}})$$

Similar to the innovation scope, for each granted patent owned by a firm, we classify all patents that were cited by the focal patent into different technological categories in terms of 4-digit IPC code. We then calculate the number of unique technological fields associated with cited patents for all patents granted by the firm. This measure denotes *extensive citation scope*.

$$\text{Extensive Citation Scope}_{it} = log(1 + \text{\# of CitedTechField}_{it})$$

Furthermore, the *intensive citation scope*, computed as the number of unique cited technological fields divided by the number of patents granted, captures how many different fields on average a firm must reference to innovate a new patent.

$$\text{Intensive Citation Scope}_{it} = log(1 + \frac{\text{\# of CitedTechField}_{it}}{\text{\# of Patent}_{it}})$$

In the section of robustness check, we broaden the definition of technological fields from 4-digit IPC (i.e., subclass) to 3-digit IPC (i.e., class) code on the one hand, and recalculate measures of citation scope by using the number of nonself-citations, which is the number of backward citations from other firms but itself.

A wide variety of citations-based measures is defined to examine different aspects of patent innovation. Following Trajtenberg et al. (1997), we use "generality" and "originality" measures, computed in the original patent file, to proxy patent quality. The former is computed based upon a patent's forward citations, while the latter is calculated based upon its backward citations. The "generality" measure renders a lower score if a patent's forward citations are concentrated in a few technology fields, while a higher score if its forward citations belong to a wide range of fields. Similarly, the "originality" measure takes a high score if a patent's backward citations in a wide range of fields, but a low score if its backward citations belonging to a narrow set of technologies. We further calculate the firm-level generality and originality measures by computing the weighted average patents scores of "generality" and "originality," respectively.

2.2.2 Other Firm-Level Controls

The main variable of interest is the estimated firm-level productivity, denoted by TFP_{it}. The estimation of productivity is full of challenges. Olley and Pakes (1996) (the OP method for short) solves the simultaneity of output and capital stock when estimating the TFP. In addition, this method corrects the selection bias problem that only firms with high productivity can survive and be continually observed in panel data sample. Based on the OP method, Levinsohn and Petrin (2003) (the LP method for short) propose a similar method by introducing intermediate inputs to the estimation to solve the simultaneity problem. Compared with OLS and fixed effects, both the OP and LP methods use investment or intermediate inputs to control for correlation between input and unobserved productivity. Because of their special advantages, the OP method and the LP method are widely used in applied researches relative to firm's productivity (Keller and Yeaple, 2009; Smarzynska Javorcik, 2004; Fernandes, 2007; Kasahara and Rodrigue, 2008; Petrin and Levinsohn, 2012). In this paper, we use the TFP estimated from the LP method in the baseline, and choose the TFP estimated from the OP method as a robustness check.

As noted in the growing literature (Konar and Cohen, 2001; Bloom et al., 2010; Carrion-Flores and Innes, 2010; Amore and Morten, 2011; Kock et al., 2012), other firm-level characteristics play important roles in explaining firms innovation activities. Specifically, we use firm's (log) deflated net sales to measure firm size. Capital intensity is calculated as the ratio of deflated capital to employment. Moreover, we also control for firm's age, which is the difference between the current year and its beginning year reported in Compustat.

R&D expense is generally regarded as a driving force for firm's innovation. We employ R&D expense per asset value as a control variable. Last, but not least, patent stock and technology stock represent firms' knowledge stock. Firms tend to create ideas and invent new patents if they have accumulated knowledge in certain technological fields that they are familiar with. Thus, the stock of patents and technologies have expected positive impacts on firms' innovative behavior. For each firm, we sum up all patents and unique technological categories occurred in previous years, and use these variables to measure the spillover effects of knowledge stock on innovation.

2.3 Descriptive Statistics

The upper panel of Figure 1 depicts average TFP and innovation across year during the study period of 1976 - 2006, while the lower panel of Figure 1 captures annual changes in average TFP and innovation measures. For each year, we take average of firm-level TFP, innovation, innovation scope, citation, and citation scope. This figure shows that innovation activities are positively correlated with TFP across years. We further take a close look at the correlation between TFP and innovation at firm level. Figure 2 provides box plots for firm-level innovation and TFP. We rank firms by their estimated TFP, and divide the TFP distribution by 10 deciles. For each decile of the productivity distribution, a series of box plot for firm-level innovation, innovation scope, citation, and citation scope are plotted. As shown in Figure 2, the horizontal line is the 10 deciles of TFP distribution. From left to right, the TFP increases from the lowest 10 percentile to the top 10 percentile. Clearly, there is a positive correlation between productivity and innovation activities.

Table 2 presents summary statistics for variables of interests at firm level. On average, each firm has around 12 patents granted per year, covering roughly 3 different technological fields in terms of 4-digit IPC code. The total backward citations for all patents granted by each firm reach more than 120 and are involved in 9 different technological fields. To develop a new patent, each firm on average needs to cite around 5 patents across two different technological fields.

3 Empirics and Results

We are interested in investigating how firm-level productivity is correlated with innovation activities in terms of the number of patents granted and the number of technological fields associated with these patents. Moreover, to successfully develop a new patent in a certain technological field, we explore how firm-level productivity is associated with the number of citations per patent and the number of technological fields associated with these cited patents. Furthermore, we include a handful of firm-level controls while estimating the correlation between productivity and innovation activities at firm-level. To this end, the following specification with a set of fixed effects for industry, year, and firm is presented:

$$Y_{it} = \beta_0 + \beta_1 TFP_{it-1} + \alpha \mathbf{X}_{it-1} + \theta_{jt} + \delta_t + \mu_i + \varepsilon_{ijt} \tag{1}$$

where i indexes a firm, j indicates an industry, and t references a year. In (1), θ_{jt} is a vector of industry variables that remove the time-variant shocks common to all manufacturers in the same industry, δ_t is year fixed effect, μ_i is firm-level fixed effect controlling for unobservable firm heterogeneity, and ε_{ijt} is the stochastic error term.

The outcome variable Y_{it} includes a series of measures on firm-level innovation, innovation scope, citation, and citation scope. TFP_{it-1}, which is critical to our study, denotes the estimated one-year lagged firm-level productivity using the LP method. Other one-year lagged firm-level controls that are related to innovation activities are absorbed in a vector of $\mathbf{X}_{it-1}$. The parameter of interest, denoted by β_1, captures how firm-level productivity is correlated with firms innovation activities.

Based upon the baseline specification (1), we first seek to explore how productivity and firms characteristics are correlated with innovation, extensive innovation scope, and intensive innovation scope. We then test how firm-level productivity is related with citations and citation scope in both extensive and intensive fashions.

3.1 Innovation and Innovation Scope

Table 3 presents the OLS estimation results for innovation and innovation scope. In each column, a set of fixed effects for year, industry, and firm as well as industry year trend is added in specification (1) as noted in the bottom of the table. Industry fixed effects are measured at 4-digit SIC level. Standard errors presented in the parenthesis are clustered at firm level.

Columns (1) - (3) in Table 3 provide the estimated results for the correlation between TFP and innovation, controlling for firms' characteristics that are related to innovation choices. The estimated coefficients for TFP are positive and statistically significant at 1% level in all columns, capturing the positive correlation between firm-level productivity and innovation in terms of the number of patents granted. Productive firms tend to be more actively engaged in innovation. One percent increase in productivity is associated with roughly 10 percent increase in the number of patents granted as noted in column (3). Firms characteristics other than productivity also plays an important role in innovative activities. The effect of firm size on innovation is positive and statistically significant for 1% level. Firms with bigger size in terms of larger value of sales are more likely to bear sunk costs of investment in innovation. In addition, there is a positive estimate of capital intensity, which is statistically significant at 1% level. Firms with intensive capital relative to labor appear to have more patents granted. In column (1), the estimated coefficient for firms age is negative and statistically significant at 1% level, suggesting that younger firms are more actively involved in innovation. With firm fixed effects added in columns (2)-(3), the negative estimates of firm age lose statistical significance at any conventional level, suggesting little evidence on the negative effects of firm age on innovation. When it comes to R&D expenses per value of asset, the estimated coefficient is positive and statistically significant at 1% level. As one of driving forces for innovation, the more expenses in R&D per value of asset, the more patents a firm would have. Lastly, the estimated coefficients for patent stock and technology stock are positive. These positive estimates with statistical significance at 1% level lend support to the strong evidence that knowledge stock has a positive and statistically significant spillover effects on innovation.

Columns (4)-(6) in Table 3 show the estimated results for extensive innovation scope, which is measured by the log number of unique technological categories for patents granted at the firm level. First of all, positive estimates of TFP are documented. These estimates are statistically significant at 1% level, reflecting the positive correlation between productivity and extensive innovation scope. Productive firms tend to develop patents in a wide spectrum of technological categories. The point estimate of β_1 coefficient is 0.068 in column (8), indicating that one percent increase in TFP is associated with 6.8 percent increase in the number of unique technological fields. Secondly, firms attributes including size and capital intensity are found to have positive and statistically significant effects on extensive innovation scope. The estimated effect of R&D expenses on innovation is positive and statistically significant at 1% level. Lastly, knowledge stock, either in a form of patent stock or technology stock, appears to have significantly positive spillover effects on broadening firms' technology spectrum.

Last three columns of Table 3 present the corresponding results for intensive innovation scope. The estimated coefficients for TFP are positive and statistically significant at 1% level. These positive estimates suggest that productive firms have less patents per technological field. Other firms' attributes including capital intensity, R&D expenses, and knowledge stock appear to facilitate innovating firms to focus on certain technological fields, as shown by negative and statistically significant coefficients in columns (9)-(12) of Table 3.

Together with strong evidence on the positive correlation between productivity and innovation or innovation scope, we find that productive firms have more patents, and their innovative activities cover a wide area of technological fields.

3.2 Citation and Citation Scope

We are interested in, to develop a new patent, how many citations are needed for patentable inventions, and how many different technological fields associated with these cited patents the firm must reference. Moreover, we seek to exaime how firms attributes are related with citation, average citation, and measures of citation scope. Table 4 presents the OLS estimation results for specification (1) with dependent variables constructed based upon citation and citation scope.

So far as the effect of productivity on citation is concerned, we find consistently positive coefficients of β_1 for the estimated TFP in columns (1)-(3) of Table 4. These estimated coefficients are statistically significant at 1% level, lending a strong support on the positive correlation between productivity and citation. Together with the positive correlation between patents and productivity, the more productive a firm, the more patents it has been granted, and the more citations associated with granted patents the firm references. As moving on to the average citation, shown in columns (4)-(6), we document positive and statistical significant estimates on the productivity. To develop a new patent, firms with higher productivity on average cite more patents than those with lower productivity. Put it differently, productive firms are more capable of learning from a relatively large number of other patents to develop a new one.

When it comes to citation scope, we first look at the extensive citation scope, which is the number of unique technological fields associated with cited patents. As reported in columns (7)-(9) of Table 4, the estimated coefficients for TFP are positive and statistically significant at 1% levels. These positive estimates suggest that the more productive a firm, the more the number of technological fields associated with cited patents the firm references. The remaining columns of Table 4 show the results for intensive citation scope. A consistently positive estimate for TFP is documented with statistical significance at 1% level, suggesting productive firms on average cite patents from a wider spectrum of technological fields than their competing counterparts.

3.3 Originality and Generality

Table 5 presents the OLS results for measures regarding novelty of invented patents. Two additional firm-level measures of originality and generality are examined as dependent variables. The measure of originality captures a firm's ability of inventing a new patent by combining pieces of knowledge from different technological fields, whereas the measure of generality reflects how popular patents of a firm are widely used (hence cited) in other future invention of patents.

In the first three columns of Table 5, estimates of TFP are positive and statistically significant at 1% level, lending a strong support on the positive correlation between productivity and the novelty of invented patents. The higher productivity, the more novel the invented patents owned the firm are. Moreover, consistently positive estimates of TFP are documented in the last three columns of Table 5. These statistically significant estimates imply that productive firms tend to invent patents which are generally cited by future inventors.

In addition, firm-level characteristics other than productivity significantly contribute to the novelty of innovation. Size, capital intensity and R&D expenditure per value of asset have positive and statistically significant impacts on both originality and generality of innovation. Age of firms, however, still is negatively correlated with the novelty of innovation. The negative estimates of firm age are statistically significant at 1% level with year and industry fixed effects added, while the significance loses at any conventional level when firm fixed effects are controlled to absorb unobservable firm-level heterogeneity. Lastly, knowledge stock, in a form of patent stock or technology stock, is positively correlated with the novelty of innovation, as suggested by the positive estimates with statistical significance at 1% level in all columns.

3.4 Robustness Checks

We conduct a series of robustness checks on alternative model specification, technological classification, nonself-citations, and productivity measure.

3.4.1 Poisson Method

Because the number of patents and the number of unique technological fields associated with patents are count data, the Poisson model is an alternative specification suitable to dealt with this type of data. Instead of taking logarithm, we use the number of patents and the number of unique technological fields served as alternative measures for innovation and extensive innovation scope, respectively. Moreover, the number of citations and the number of unique technological fields associated with cited patents are adopted for dependent variables. The alternative specification is given by,

$$Y_{it} = exp(\beta_1 TFP_{it-1} + \alpha X_{it-1})\eta_{it} + \varepsilon_{it} \tag{2}$$

where $Y_{it} \in \{0, 1, 2, 3, \cdots\}$ is the count data of interests. TFP_{it-1} is the estimated one-year lagged productivity using the LP method, and X_{it-1} is a vector of one-year lagged firm attributes including size, age, capital intensity, R&D expenses per value of asset, and knowledge stock in a form of patent and technology stock. η_{it} captures all unobservable, time-varying attributes of firm i and may be correlated with some of variables X_{it-1}. ε_{it} is an error term satisfying $E(\varepsilon_{it}|X_{it-1}, \eta_{it}) = 0$.

Table 6 presents the corresponding estimation results for the Poisson specification that employs the count data. Columns differ in the choice of fixed effects (i.e., year, industry, and firm level) as noted in the bottom of the table. Standard errors presented in the parenthesis are clustered at firm level. The first four columns of Table 6 show how firms attributes are related with innovation and extensive innovation scope, while the remaining columns present how firms attributes are correlated with citations and extensive citation scope. The positive coefficients for productivity are statistically significant at 1% level in all columns, but column (7), providing the corroborating evidence that TFP is positively correlated with innovation, extensive innovation scope, citation, and extensive citation scope when the count data are employed.

The results for firms other attributes are also robust to the alternative specification of Poisson model. There are consistently positive and statistically significant correlation between firms size and innovation or innovation scope. Similarly, the positive correlations are documented for other firm-level characteristics including capital intensity, R&D expenses and knowledge stock (either in a form of patent or technology stock). One exception is firm age. The estimated coefficients for firm age are negative and statistically significant without firm-level fixed effects. When firm fixed effects are controlled to absorb unobservable firm heterogeneity, firm age has no statistically significant effects on extensive innovation scope, citation and extensive citation scope.

3.4.2 Alternative Technology Classification

To test the robustness of our empirical results, we further relax patents' technological classification from the primary 4-digit IPC code to 3-digit IPC code. Hence, technological categories are broadened from subclass to class in the ascending IPC hierarchy. With this alternative definition of technology categories, associated dependent variables of innovation scope and citation scope in forms of both extensive and intensive fashions are recalculated. In addition, an alternative measure for technology stock is computed when the 3-digit IPC code is used to define technology fields.

Columns (1)-(4) in Table 7 present the corresponding OLS estimation for specification (1) with the implementation of the alternative technology classification. The main results regarding the positive correlation between productivity and dependent variables of interests are robust to this alternative measure. Productive firms are involved in a wide spectrum of technology categories, and have less patents per technology field. To develop a new patent, productive firms cite patents from a wide area of technological fields.

3.4.3 Nonself-Citation

In this robustness check, we subtract citations made by firms themselves when counting the number of citations, and the number of unique technological fields associated with citations. Hence measures for average citation, extensive and intensive citation scope are recalculated without accounting for self-citations. Columns (5)-(7) in Table 7 present the corresponding results for the relevant dependent variables. In the last two columns of Table 7, both the alternative technology classification and nonself-citation are taken into consideration for the relevant dependent variables. We still find consistent evidence on the positive and significant correlation between productivity and dependent variables of interests.

3.4.4 Alternative Productivity Measure

Lastly, but not least, we use the OP method to estimate firm's productivity. Under a series of assumptions, the OP method can get consistent estimations, using the firm-level product function. One of these assumptions is the monotonic relationship between investment and output. However, not every firm in every year conducts investment, thereby having zero investment in some years. Under the OP method, many observations with zero investment have to be deleted.

Table 8 shows the estimated results for the baseline specification with an alternative productivity measure estimated from the OP method. A set of fixed effects on year and firm is controlled as noted in the bottom of the table. Compared with the baseline results in Table 3 and 4, the estimates for the alternative productivity measure remain the same sign with smaller magnitude, but lose statistical significance in any conventional level when it comes to intensive innovation scope and intensive citation scope. For the remaining firm-level characteristics, the main results are robust to the alternative productivity measure.

4 Conclusion

This paper employs firm-level patent data from the US manufacturing public firms over the period of 1976-2006. We document several interesting and novel results regarding firm-level productivity and innovation. First, there is a positive correlation between productivity measured by the TFP estimates and innovation proxied by the number of patents. Unsurprisingly, productive firms tend to have more patents. Second, when taking closer look at technological fields associated with these patents, there is strong evidence on the positive correlation between productivity and extensive innovation scope. Productive firms are more actively engaged in innovation not only by having more patents, but also being involved in a wide spectrum of technological fields. Third, a positive correlation between productivity and intensive innovation scope, measured by innovation scope per patent, is documented. Productive firms appear to conduct innovation intensively across a wide spectrum of technological fields. Lastly, when it comes to developing a new patent, we find consistently evidence supporting a positive correlation between productivity and extensive citation scope, and a positive correlation between TFP and intensive citation scope. The more productive a firm, the more patents it must cite to develop a new one, the wide spectrum of technological filed associated with cited patents the firms must reference.

References

Aghion, P., Howitt, P., Prantl, S., 2013. Patent Rights, Product Market Reforms, and Innovation. Tech. rep., National Bureau of Economic Research.

Amore, M. D., Morten, B., 2011. Corporate Governance and the Environment: Evidence from Clean Innovations. Working paper.

Arora, A., Belenzon, S., Patacconi, A., 2015. Killing the Golden Goose? The Decline of Science in Corporate R&D. Working Paper 20902, National Bureau of Economic Research.

Becker, R., Gray, W., Marvakov, J., 2013. NBER-CES Manufacturing Industry Database: Technical Notes. Working paper.

Bessen, J., 2009. User Documentation: Matching Patent Data to Compustat Firms. NBER PDP Project.

Bloom, N., Genakos, C., Martin, R., Sadun, R., 2010. Modern Management: Good for the Environment or Just Hot Air? The Economic Journal 120 (544), 551–572.

Bound, J., Cummins, C., Griliches, Z., Hall, B. H., Jaffe, A. B., 1982. Who Does R&D and Who Patents? Working Paper 908, National Bureau of Economic Research.

Carrion-Flores, C. E., Innes, R., 2010. Environmental Innovation and Environmental Performance. Journal of Environmental Economics and Management 59 (1), 27–42.

Castellani, D., Zanfei, A., 2007. Internationalisation, Innovation and Productivity: How Do Firms Differ in Italy? World Economy 30 (1), 156–176.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level. Working Paper 6696, National Bureau of Economic Research.

Fernandes, A. M., 2007. Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries. Journal of International Economics 71 (1), 52–71.

Fleming, L., 2001. Recombinant Uncertainty in Technological Search. Management Science 47 (1), 117–132.

Geroski, P. A., 1989. Entry, Innovation and Productivity Growth. The Review of Economics and Statistics 71 (4), 572–578.

Griliches, Z., 1996. The Discovery of the Residual: A Historical Note. Journal of Economic Literature (3), 1324–1330.

Griliches, Z., 1998. R&D and Productivity the Econometric Evidence. University of Chicago Press.

Hall, B. H., 2011. Innovation and Productivity. Working paper, National Bureau of Economic Research.

Hall, B. H., Griliches, Z., Hausman, J. A., 1984. Patents and R&D: Is There a Lag? Working Paper 1454, National Bureau of Economic Research.

Hall, B. H., Harhoff, D., 2012. Recent Research on the Economics of Patents. Annual Review of Economics 4 (1), 541–565.

Hall, B. H., Helmers, C., 2010. The Role of Patent Protection in (Clean/Green) Technology Transfer. Tech. rep., National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Working paper, National Bureau of Economic Research.

Hall, B. H., Lotti, F., Mairesse, J., 2009. Innovation and Productivity in SMEs: Empirical Evidence for Italy. Small Business Economics 33 (1), 13–33.

Hall, B. H., Sena, V., 2014. Appropriability Mechanisms, Innovation and Productivity: Evidence from the UK. Working Paper 20514, National Bureau of Economic Research.

Janz, N., Loof, H., Peters, B., 2004. Firm Level Innovation and Productivity - Is there a Common Story Across Countries? Working Paper Series in Economics and Institutions of Innovation, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

Kasahara, H., Rodrigue, J., 2008. Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence. Journal of Development Economics 87 (1), 106–118.

Keller, W., Yeaple, S. R., 2009. Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States. Review of Economics and Statistics 91 (4), 821–831.

Kock, C. J., Santaló, J., Diestre, L., 2012. Corporate Governance and the Environment: What Type of Governance Creates Greener Companies? Journal of Management Studies 49 (3), 492–514.

Konar, S., Cohen, M. A., 2001. Does the Market Value Environmental Performance? Review of Economics and Statistics 83 (2), 281–289.

Levinsohn, J., Petrin, A., 2003. Estimating Production Functions Using Inputs to Control for Unobservables. The Review of Economic Studies 70 (2), 317–341.

Loof, H., Heshmati, A., 2006. On the Relationship between Innovation and Performance: A Sensitivity Analysis. Economics of Innovation and New Technology 15 (4-5), 317–344.

Miguel Benavente, J., 2006. The Role of Research and Innovation in Promoting Productivity in Chile. Economics of Innovation and New Technology 15 (4-5), 301–315.

Mohnen, P., Hall, B. H., 2013. Innovation and Productivity: An Update. MERIT Working

Paper 021, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

Nagaoka, S., Motohashi, K., Goto, A., 2010. Patent Statistics as an Innovation Indicator 2, 1083–1127.

Olley, G. S., Pakes, A., 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. Econometrica 64 (6), 1263–1297.

Petrin, A., Levinsohn, J., 2012. Measuring Aggregate Productivity Growth Using Plant-Level Data. The RAND Journal of Economics 43 (4), 705–725.

Raymond, W., Mairesse, J., Mohnen, P., Palm, F., 2013. Dynamic Models of R&D, Innovation and Productivity: Panel Data Evidence for Dutch and French Manufacturing. Working Paper 19074, National Bureau of Economic Research.

Roper, S., Du, J., Love, J. H., 2008. Modelling the Innovation Value Chain. Research Policy 37 (6-7).

Smarzynska Javorcik, B., 2004. Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages. American Economic Review 94 (3), 605–627.

Therrien, P., Hanel, P., Nov. 2009. Innovation and Productivity. In: Innovation in Firms. Organisation for Economic Co-operation and Development, pp. 139–156.

Trajtenberg, M., 1987. Patents, Citations and Innovations: Tracing the Links. Working Paper 2457, National Bureau of Economic Research.

Trajtenberg, M., Henderson, R., Jaffe, A., 1997. University versus Corporate Patents: A Window on the Basicness of Invention. Economics of Innovation and New Technology 5 (1), 19–50.

Van Leeuwen, G., Klomp, L., 2006. On the Contribution of Innovation to Multi-Factor Productivity Growth. Economics of Innovation and New Technology 15 (4-5), 367–390.

Figures and Tables

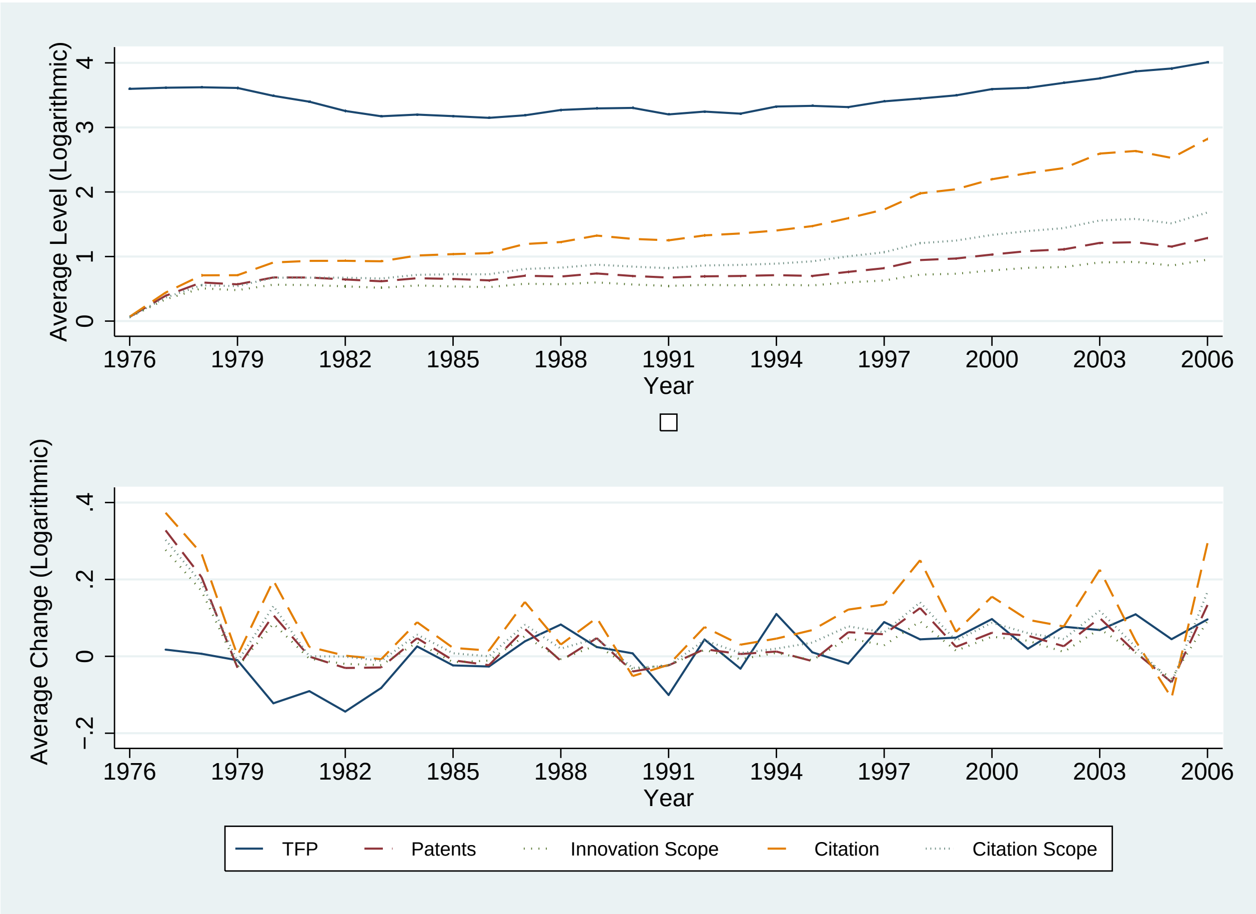

Figure 1: Average TFP and Innovation during 1976-2006

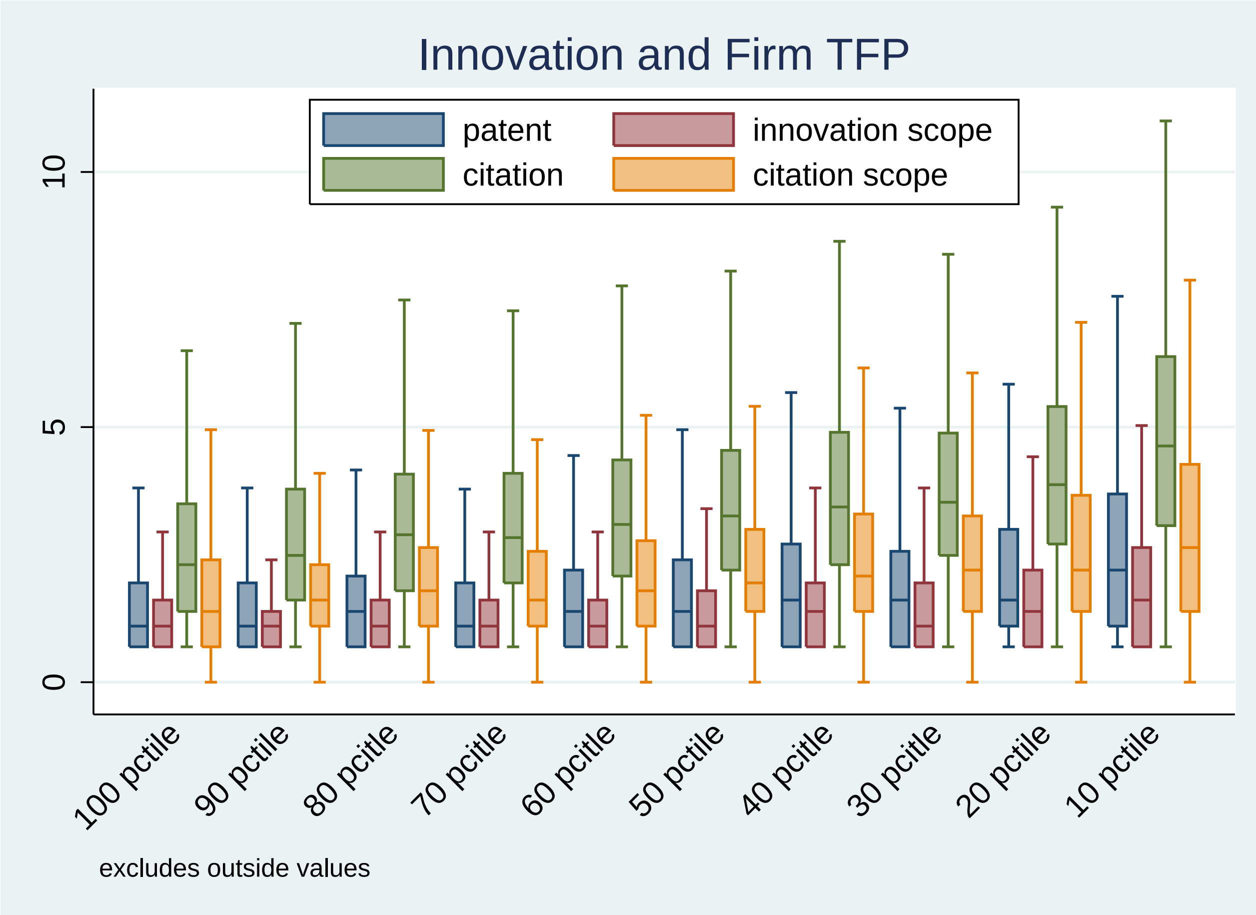

Figure 2: Box Plots for Firm-level Innovation and TFP Percentile Distribution

Table 1: Variable Description

Variables	Description
Firm-Level Patent Data from NBER Patent Database	
Patent	Number of patents
Tech Fields	Number of unique technology categories associated with patents under 4-digit IPC
Citation	Number of forward citations for firms' patents
Cited Tech Fields	Number of unique technology categories associated with cited patents under 4-digit IPC
Innovation	$\log(1 + \text{\# of Patent})$
Extensive Innovation Scope	$\log(1 + \text{\# of Tech Field})$
Intensive Innovation Scope	$\log(1 + \frac{\text{\# of Tech Field}}{\text{\# of Patent}})$
Citation	$\log(1 + \text{\# of Citation})$
Average Citation	$\log(1 + \frac{\text{\# of Citation}}{\text{\# of Patent}})$
Extensive Citation Scope	$\log(1 + \text{\# of Cited Tech Field})$
Intensive Citation Scope	$\log(1 + \frac{\text{\# of Cited Tech Field}}{\text{\# of Patent}})$
Patent Stock	Stock of patents
Technology Stock	Stock of patents' unique technology categories under 4-digit IPC
Firm-Level Financial Statement from Compustat of WRDS	
TFP	Productivity estimated by the Levinsohn and Petrin (2003)'s method
Alt. TFP	Alternative productivity estimated by the Olley and Pakes (1996)'s method
Sales	Value of net sales/turnovers (item 12 at Compustat), deflated by industry-level value of shipment deflator
Capital	Total net value of property, plants and equipment less depreciation (item 8), deflated by industry-level shipment deflator
Labor	Number of employees (item 29)
Wage Expenditure	Total payrolls, calculated labor multiplied by average industry wage
Materials	Costs of goods sold (item 41) plus administrative and selling expenses (item 189) less depreciation (item 14) and wage expenditure, deflated by industry-level material costs deflator
R&D Expenses	Research and development expenses (item 46)
Asset Value	Book value of total assets (item 29)
Investment	Capital expenditures (item 128), deflated by industry-level investment deflator
Revenue	Total revenue
Age	Firm's age (current year - the first year)
Size	Log sales, deflated by industry-level value of shipment deflator
Capital Intensity	Log ratio of capital to labor
Industry-level Data at the 4-Digit SIC from the NBER-CES Manufacturing Data	
Value of Shipment Deflator	Deflator for value of shipment
Investment Deflator	Deflator for investment value
Material Cost Deflator	Deflator for material costs

Table 2: Summary Statistics

Variables	N	Mean	Std. Dev.	Min	Max
Patent	37340	12.22	71.80	0	2344
Extensive Innovation Scope	37340	3.241	9.854	0	152
Intensive Innovation Scope	37340	0.344	0.422	0	5
Citation	37340	127.9	931.1	0	60010
Average Citation	37340	5.415	13.14	0	505.8
Extensive Citation Scope	37340	9.261	25.89	0	359
Intensive Citation Scope	37340	1.094	1.856	0	44
Originality (ln)	37340	1.398	1.956	0	10.38
Generality (ln)	37340	1.148	1.763	0	9.356
Patent Stock	37340	117.8	759.5	1	30937
Technology Stock	37340	11.37	28.92	1	397
TFP	36663	3.521	1.536	-6.844	11.22
Alt. TFP	33485	37.53	26.22	-1.375	113.3
Capital (ln)	37216	3.112	2.754	-6.908	11.64
Employment (ln)	36849	-0.116	2.240	-6.908	6.776
Capital Intensity (ln)	36815	3.235	1.073	-2.977	8.352
RD Expenditure per Asset Value (ln)	30899	-2.804	1.983	-10.23	9.435
Sale (ln)	37239	4.753	2.419	-0.0192	12.72
Age	37340	11.22	7.083	1	31

Note: TFP is the estimated productivity, using the LP method, while Alt. TFP is the estimated productivity with the OP method. All variables of interests presented in the above table are not implemented in log fashions unless noted.

Table 3: Baseline Results for Innovation and Innovation Scope

VARIABLES	Innovation $log(1+\#\text{ of Patent}_{it})$			Extensive Innovation Scope $log(1+\#\text{ of TechField}_{it})$			Intensive Innovation Scope $log(1+\frac{\#\text{ of TechField}_{it}}{\#\text{ of Patent}_{it}})$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TFP	0.124***	0.174***	0.096***	0.081***	0.113***	0.068***	0.017***	0.018***	0.021***
	(0.015)	(0.023)	(0.022)	(0.011)	(0.017)	(0.016)	(0.005)	(0.006)	(0.007)
Size	0.054***	0.122***	0.169***	0.050***	0.098***	0.125***	0.005	0.037***	0.036***
	(0.010)	(0.022)	(0.023)	(0.007)	(0.016)	(0.017)	(0.003)	(0.006)	(0.006)
Capital Intensity	0.071***	0.085***	0.059***	0.039***	0.057***	0.041***	0.007*	0.009*	0.008*
	(0.012)	(0.016)	(0.015)	(0.009)	(0.012)	(0.011)	(0.004)	(0.005)	(0.005)
Age	-0.029***	0.019	-0.000	-0.021***	0.012	-0.004	-0.004***	-0.008**	-0.010***
	(0.002)	(0.022)	(0.020)	(0.001)	(0.014)	(0.012)	(0.001)	(0.003)	(0.003)
R&D Expenditure	0.115***	0.138***	0.126***	0.090***	0.098***	0.090***	0.025***	0.026***	0.026***
	(0.009)	(0.013)	(0.012)	(0.006)	(0.010)	(0.009)	(0.003)	(0.004)	(0.004)
Patent Stock	0.457***	0.257***	0.271***	0.268***	0.121***	0.133***	-0.005	-0.067***	-0.065***
	(0.023)	(0.028)	(0.026)	(0.015)	(0.020)	(0.019)	(0.004)	(0.005)	(0.005)
Tech Stock	0.248***	0.284***	0.255***	0.267***	0.240***	0.217***	0.055***	0.034***	0.032***
	(0.035)	(0.039)	(0.033)	(0.024)	(0.030)	(0.025)	(0.005)	(0.007)	(0.007)
Observations	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066
Adjusted R^2	0.705	0.290	0.314	0.681	0.232	0.256	0.079	0.033	0.038
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y			Y			Y		
Firm FE		Y	Y		Y	Y		Y	Y
Industry Year Trend			Y			Y			Y

Note: All independent variables are in one-year lagged fashion. TFP is the estimated productivity, using the LP method. Size is log deflated value of sales. Capital intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D expenditure is the log R&D expenses per deflated value of asset. Patent Stock is the stock of a firm's patent in past years. Tech Stock is the stock of unique technological categories associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 4: Baseline Results for Citation and Citation Scope

VARIABLES	Citations $log(1 + \#\text{ of Citation}_{it})$			Average Citation $log(1 + \frac{\#\text{ of Citation}_{it}}{\#\text{ of Patent}_{it}})$			Extensive Citation Scope $log(1 + \#\text{ of CitedTechField}_{it})$			Intensive Citation Scope $log(1 + \frac{\#\text{ of CitedTechField}_{it}}{\#\text{ of Patent}_{it}})$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
TFP	0.213***	0.229***	0.153***	0.102***	0.076***	0.077***	0.115***	0.138***	0.091***	0.009***	0.017***	0.014***
	(0.025)	(0.037)	(0.038)	(0.017)	(0.023)	(0.025)	(0.015)	(0.023)	(0.024)	(0.003)	(0.004)	(0.004)
Size	0.038**	0.265***	0.317***	-0.012	0.169***	0.175***	0.047***	0.163***	0.198***	0.006***	0.017***	0.019***
	(0.017)	(0.037)	(0.040)	(0.011)	(0.021)	(0.023)	(0.010)	(0.023)	(0.025)	(0.002)	(0.003)	(0.003)
Capital Intensity	0.142***	0.139***	0.111***	0.074***	0.059***	0.056***	0.074***	0.086***	0.069***	0.002	0.006*	0.004
	(0.020)	(0.027)	(0.028)	(0.014)	(0.018)	(0.019)	(0.012)	(0.017)	(0.017)	(0.002)	(0.003)	(0.003)
Age	-0.054***	-0.012	-0.039	-0.028***	-0.038***	-0.046***	-0.032***	-0.002	-0.021	-0.001***	-0.002	-0.004
	(0.004)	(0.031)	(0.029)	(0.002)	(0.014)	(0.014)	(0.002)	(0.020)	(0.017)	(0.000)	(0.002)	(0.002)
R&D Expenditure	0.198***	0.232***	0.225***	0.102***	0.117***	0.120***	0.121***	0.143***	0.139***	0.013***	0.017***	0.015***
	(0.015)	(0.023)	(0.022)	(0.010)	(0.015)	(0.015)	(0.009)	(0.015)	(0.014)	(0.002)	(0.002)	(0.002)
Patent Stock	0.720***	0.439***	0.454***	0.251***	0.097***	0.100***	0.393***	0.210***	0.224***	-0.010***	-0.057***	-0.055***
	(0.038)	(0.039)	(0.037)	(0.020)	(0.019)	(0.019)	(0.022)	(0.026)	(0.025)	(0.002)	(0.003)	(0.003)
Tech Stock	0.321***	0.414***	0.377***	0.104***	0.148***	0.138***	0.305***	0.308***	0.280***	0.028***	0.015***	0.014***
	(0.057)	(0.058)	(0.052)	(0.028)	(0.027)	(0.027)	(0.034)	(0.039)	(0.035)	(0.003)	(0.004)	(0.004)
Observations	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066
Adjusted R^2	0.619	0.288	0.299	0.344	0.164	0.169	0.622	0.255	0.269	0.060	0.038	0.044
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y			Y			Y			Y		
Firm FE		Y	Y		Y	Y		Y	Y		Y	Y
Industry Year Trend			Y			Y			Y			Y

Note: All independent variables are in one-year lagged fashion. TFP is the estimated productivity, using the LP method. Size is log deflated value of sales. Capital intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D expenditure is the log R&D expenses per deflated value of asset. Patent Stock is the stock of a firm's patent in past years. Tech Stock is the stock of unique technological categories associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 5: Alternative Measures on the Novelty of Innovation

VARIABLES	Originality			Generality		
	(1)	(2)	(3)	(4)	(5)	(6)
TFP	0.192***	0.187***	0.131***	0.184***	0.224***	0.158***
	(0.022)	(0.033)	(0.033)	(0.020)	(0.030)	(0.030)
Size	0.017	0.199***	0.239***	0.027*	0.167***	0.203***
	(0.015)	(0.032)	(0.035)	(0.014)	(0.029)	(0.032)
Capital Intensity	0.138***	0.126***	0.106***	0.093***	0.114***	0.094***
	(0.018)	(0.023)	(0.024)	(0.017)	(0.022)	(0.022)
Age	-0.046***	-0.013	-0.037	-0.034***	-0.016	-0.034
	(0.003)	(0.026)	(0.025)	(0.003)	(0.028)	(0.025)
R&D Expenditure	0.168***	0.191***	0.186***	0.165***	0.203***	0.192***
	(0.013)	(0.020)	(0.019)	(0.012)	(0.019)	(0.018)
Patent Stock	0.660***	0.515***	0.529***	0.536***	0.229***	0.245***
	(0.034)	(0.034)	(0.032)	(0.028)	(0.032)	(0.031)
Tech Stock	0.283***	0.360***	0.327***	0.305***	0.358***	0.329***
	(0.050)	(0.050)	(0.045)	(0.041)	(0.043)	(0.039)
Observations	31,066	31,066	31,066	31,066	31,066	31,066
Adjusted R^2	0.623	0.331	0.341	0.570	0.254	0.264
Year FE	Y	Y	Y	Y	Y	Y
Industry FE	Y			Y		
Firm FE		Y	Y		Y	Y
Industry Year Trend			Y			Y

Note: All independent variables are in one-year lagged fashion. TFP is the estimated productivity, using the LP method. Size is log deflated value of sales. Capital intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D expenditure is the log R&D expenses per deflated value of asset. Patent Stock is the stock of a firm's patent in past years. Tech Stock is the stock of unique technological categories associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 6: Robustness Check for Poisson Method

VARIABLES	Innovation # of Patent$_{it}$		Extensive Innovation Scope # of TechField$_{it}$		Citation # of Citation$_{it}$		Extensive Citation Scope # of CitedTechField$_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFP	0.377***	0.338***	0.216***	0.224***	0.208***	0.170	0.158***	0.194***
	(0.025)	(0.078)	(0.013)	(0.049)	(0.047)	(0.128)	(0.013)	(0.044)
Size	0.079***	0.161***	0.108***	0.178***	0.058*	0.127	0.102***	0.175***
	(0.014)	(0.050)	(0.008)	(0.036)	(0.031)	(0.094)	(0.008)	(0.034)
Capital Intensity	0.217***	0.176***	0.072***	0.132***	0.121***	0.274***	0.051***	0.144***
	(0.020)	(0.055)	(0.011)	(0.041)	(0.036)	(0.065)	(0.011)	(0.035)
Age	-0.030***	0.051**	-0.028***	0.026	-0.033***	0.036	-0.028***	0.019
	(0.003)	(0.022)	(0.002)	(0.017)	(0.003)	(0.023)	(0.002)	(0.017)
R&D Expenditure	0.333***	0.238***	0.243***	0.208***	0.277***	0.162*	0.200***	0.213***
	(0.020)	(0.070)	(0.009)	(0.040)	(0.038)	(0.097)	(0.010)	(0.036)
Patent Stock	0.537***	0.300***	0.264***	0.096***	0.640***	0.391***	0.290***	0.137***
	(0.019)	(0.046)	(0.012)	(0.031)	(0.038)	(0.082)	(0.012)	(0.028)
Tech Stock	0.334***	0.440***	0.480***	0.412***	0.241***	0.401***	0.399***	0.373***
	(0.021)	(0.044)	(0.015)	(0.036)	(0.029)	(0.058)	(0.016)	(0.033)
Observations	31,066	30,133	31,066	30,133	31,066	30,133	31,066	30,133
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y		Y		Y		Y	
Firm FE		Y		Y		Y		Y

Note: All independent variables are in one-year lagged fashion. TFP is the estimated productivity, using the LP method. Size is log deflated value of sales. Capital intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D expenditure is the log R&D expenses per deflated value of asset. Patent Stock is the stock of a firm's patent in past years. Tech Stock is the stock of unique technological categories associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 7: Robustness Check for Alternative Technology Classification and Nonself-Citation

	Atl. Tech Class				Nonself Citation			Both	
VARIABLES	Extensive Innovation Scope (1)	Intensive Innovation Scope (2)	Extensive Citation Scope (3)	Intensive Citation Scope (4)	Extensive Citation Scope (5)	Average Citation (6)	Intensive Citation Scope (7)	Extensive Citation Scope (8)	Intensive Citation Scope (9)
TFP	0.055***	0.019***	0.070***	0.012***	0.088***	0.074***	0.034**	0.069***	0.027**
	(0.014)	(0.007)	(0.020)	(0.003)	(0.024)	(0.025)	(0.014)	(0.020)	(0.011)
Size	0.118***	0.031***	0.176***	0.015***	0.199***	0.173***	0.076***	0.171***	0.059***
	(0.015)	(0.006)	(0.022)	(0.003)	(0.025)	(0.023)	(0.012)	(0.021)	(0.010)
Capital Intensity	0.033***	0.007	0.055***	0.003	0.068***	0.055***	0.025**	0.055***	0.018**
	(0.011)	(0.005)	(0.015)	(0.003)	(0.017)	(0.018)	(0.010)	(0.015)	(0.009)
Age	-0.004	-0.010***	-0.018	-0.003	-0.022	-0.046***	-0.025***	-0.019	-0.020***
	(0.009)	(0.003)	(0.014)	(0.002)	(0.017)	(0.013)	(0.008)	(0.014)	(0.006)
R&D Expenditure	0.079***	0.022***	0.118***	0.013***	0.139***	0.120***	0.053***	0.117***	0.041***
	(0.008)	(0.004)	(0.012)	(0.002)	(0.014)	(0.015)	(0.008)	(0.012)	(0.007)
Patent Stock	0.130***	-0.066***	0.215***	-0.048***	0.279***	0.110***	-0.061***	0.211***	-0.076***
	(0.016)	(0.005)	(0.020)	(0.003)	(0.023)	(0.017)	(0.010)	(0.020)	(0.008)
Alt. Tech Stock	0.143***	0.022***	0.178***	0.005	0.198***	0.099***	0.062***	0.172***	0.045***
	(0.024)	(0.006)	(0.032)	(0.003)	(0.036)	(0.026)	(0.014)	(0.031)	(0.011)
Observations	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066	31,066
Adjusted R^2	0.213	0.035	0.241	0.048	0.261	0.161	0.070	0.239	0.053
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry Year Trend	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: All independent variables are in one-year lagged fashion. TFP is the estimated productivity, using the LP method. Size is log deflated value of sales. Capital intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D expenditure is the log R&D expenses per deflated value of asset. Patent Stock is the stock of a firm's patent in past years. Alt. Tech Stock is the stock of unique technological categories by 3-digit IPC associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 8: Robustness Check for Alternative Productivity Measure

VARIABLES	Innovation	Extensive Innovation Scope	Intensive Innovation Scope	Citation	Average Citation	Extensive Citation Scope	Intensive Citation Scope	Originality	Generality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Alt. TFP	0.115***	0.071***	0.002	0.114**	0.004	0.060*	0.008	0.089**	0.111***
	(0.032)	(0.023)	(0.008)	(0.050)	(0.030)	(0.032)	(0.005)	(0.044)	(0.039)
Size	0.223***	0.164***	0.047***	0.398***	0.213***	0.247***	0.028***	0.307***	0.301***
	(0.020)	(0.015)	(0.005)	(0.032)	(0.018)	(0.021)	(0.003)	(0.028)	(0.026)
Capital Intensity	0.051***	0.033***	0.003	0.082***	0.030	0.051***	0.001	0.080***	0.070***
	(0.017)	(0.012)	(0.005)	(0.028)	(0.018)	(0.018)	(0.003)	(0.024)	(0.023)
Age	-0.405***	-0.246***	-0.015	-0.438**	-0.057	-0.226*	-0.029*	-0.347**	-0.437***
	(0.121)	(0.088)	(0.030)	(0.188)	(0.111)	(0.120)	(0.017)	(0.166)	(0.147)
R&D Expenditure	0.087***	0.066***	0.020***	0.162***	0.089***	0.099***	0.011***	0.134***	0.135***
	(0.015)	(0.011)	(0.004)	(0.026)	(0.016)	(0.017)	(0.003)	(0.022)	(0.021)
Patent Stock	0.259***	0.123***	-0.069***	0.450***	0.103***	0.218***	-0.059***	0.531***	0.256***
	(0.030)	(0.022)	(0.006)	(0.042)	(0.020)	(0.028)	(0.004)	(0.037)	(0.035)
Tech Stock	0.293***	0.246***	0.037***	0.426***	0.154***	0.318***	0.018***	0.369***	0.356***
	(0.040)	(0.031)	(0.007)	(0.061)	(0.029)	(0.041)	(0.004)	(0.052)	(0.046)
Observations	27,848	27,848	27,848	27,848	27,848	27,848	27,848	27,848	27,848
Adjusted R^2	0.295	0.238	0.035	0.298	0.170	0.265	0.040	0.345	0.264
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: All independent variables are in one-year lagged fashion. Alt. TFP is the calculated productivity, using the alternative OP method. Size is log deflated value of sales. Capital Intensity is the log ratio of capital to the number of employment. Age is firm's age. R&D Expenditure is the log R&D expenses per deflated asset value. Patent Stock is the stock of a firm's patents in past years. Tech Stock is the stock of unique technological categories associated with a firm's patents in past years. Standard errors presented in the parenthesis are clustered at plant level. Industry dummies are measured at the 4-digit SIC level. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.