Is firm-level clean or dirty innovation valued more?

Antoine Dechezleprêtre

Grantham Research Institute on Climate Change and the Environment, London School of Economics*

Cal B. Muckley

UCD School of Business and Geary Institute, University College Dublin†

Parvati Neelakantan

UCD School of Business University College Dublin‡

Abstract

We study the relation between Tobin's Q and 'clean' (carbon saving) and 'dirty' (fossil-based) innovation and innovation efficiency at the firm level using a global patent dataset, covering over 15,000 firms across 12 countries. Clean innovation relates to patented technologies in areas such as renewable energy generation and electric cars, whereas dirty innovation relates to fossil-based energy generation and combustion engines. We find that innovation in clean technologies is priced higher than dirty innovation. The stock market recognizes the value of clean innovation and innovation efficiency and accords higher valuations to those firms that engage in successful clean research and development activities. The results are substantively invariant across innovation and innovation efficiency, United States and European patents, model specifications and estimators adopted.

JEL Classification: G35, G32, C58

Keywords: Innovation, research and development, patents, citations, clean technologies, dirty technologies, market value

*E-mail: A.Dechezlepretre@lse.ac.uk.

†E-mail: cal.muckley@ucd.ie **Corresponding author:** Cal Muckley, University College Dublin, Michael Smurfit Graduate Business School, Dublin, Ireland. The authors would like to thank for suggestions Don Bredin, Matt Spiegel, Luca Taschini and Heather Tookes. We also thank the participants in seminars at Babson College finance department, University College Dublin's Geary Institute, Dublin City University Business School, Loughborough University, the London School of Economics and the Organisation for Economic Co-operation and Development. Cal Muckley would like to acknowledge the financial support of Science Foundation Ireland under Grant Number 16/SPP/3347 and 17/SP/5447. This manuscript was substantively advanced while Cal Muckley was a CRH Fulbright scholar at Yale University and while he was a visiting researcher at the London School of Economics, Grantham Research Institute. This research has benefitted from the financial support of the Grantham Foundation for the Protection of the Environment and the UK Economic and Social Research Council through the Centre for Climate Change Economics and Policy. The usual disclaimer applies.

‡E-mail: parvati.neelakantan@ucdconnect.ie

1 Introduction

According to the latest Assessment Report by the Intergovernmental Panel on Climate Change, stabilising global carbon emissions in 2050 requires a 60% reduction in the carbon intensity of global GDP compared with a business-as-usual scenario[1] (IPCC, 2014). In order to achieve this long-term decarbonisation of the economy, while meeting growing energy demand, the world needs to implement a radical change in the mix of technologies used to produce and consume energy. This, in turn, requires massive investments in research and development activities. For this reason, one of the most pressing challenges for climate change policies today is how to provide an adequate economic-incentive for firms to redirect innovation away from fossil fuel ('dirty') and towards low-carbon ('clean') technologies, despite multiple market failures associated with environmental externalities and R&D provision (Jaffe et al., 2005).

In this paper, we avail of the capital markets to determine if there is an adequate economic incentive for firms to redirect innovation away from fossil fuel ('dirty') and towards low-carbon ('clean') technologies.[2] We associate 'dirty' innovation with fossil-based energy generation and ground transportation, and 'clean' innovation with renewable energy generation and electric vehicles, but discuss carefully issues around this definition and consider alternatives. We regress firm-level Tobin's Q on firm-level clean and dirty innovation, together with other measures of innovation and firm traits. Our main analyses avails of a global firm-level data set of United States patents, covering more than 15,000 firms across 12 countries. To ascertain the expected economic performance of 'clean' and 'dirty' investment activities, we, specifically, adopt a firm's intangible stock of knowledge value function as used in Hall et al. (2005). Our principal models comprise accounting-based asset valuation firm-level data in line with Ohlson (1989) and used,

[1] Assuming a 2.5 per cent annual GDP growth.

[2] Investments in R&D activities can affect firms' future financial performance and are thus reflected in firms' market value (see Grandi et al. (2009)).

most recently, in respect to the evaluation of a firm's knowledge stock in Hirshleifer et al. (2013).

A limitation of existing studies of directed technological change, towards clean innovation, is that a multitude of drivers can determine companies' decisions to conduct R&D activity. These drivers include the relative prices of production factors (Hicks, 1932; Popp, 2002; Acemoglu et al., 2012), the quality of environmental policy instruments (Johnstone et al., 2010), the extent of market demand and a path-dependency in knowledge creation (Acemoglu et al., 2012; Aghion et al., 2016), which can all influence the prospective economic returns of clean and dirty innovation. Most importantly, a variety of policies and drivers can coexist in a given jurisdiction - for example, carbon markets, fuel taxes, energy efficiency standards and renewable energy mandates - making it difficult to measure the overall impact of these policies and drivers taken together or considered in isolation. An additional complexity arises from the fact that it is the expected realization of these policies and drivers in the future which determines innovation decisions, rather than current observed realizations, but these expectations are inevitably not directly observed and may vary markedly across firms.[3] A major advantage of our approach, relative to extant studies, is that the stock market evaluation of patented innovation in clean and dirty technologies can reveal market expectations with respect to the prospective economic performance of these investments.

Our main data are drawn from the European Patent Office's (EPO) World Patent Statistical database (PATSTAT). Our database reports the name of patent applicants which allows us to match clean and dirty patents with distinct patent holders. The global nature of the database means that we can test our hypothesis on several measures of patenting activity, including patents taken out in the world's major patents offices such as the United States Patents and

[3]Despite these empirical challenges, several studies present evidence that firms may redirect innovation away from fossil fuel towards low carbon technologies when faced with change in policies and energy prices. For instance, Calel and Dechezlepretre (2016) investigate the impact of the European Union Emissions Trading System on regulated companies and note that among the regulated companies there is a 1% increase in the amount of low-carbon technology innovation. Similarly, Newell et al. (1999)and Popp (2002) report a substantial increase in the production of energy-efficient technologies following increase in energy prices.

Trademark Office (USPTO), irrespective of the jurisdiction of the innovating firm. Our data also includes information on patent citations, allowing us to partly address the well-known issue of heterogeneity in patent value.

We first verify the capital market value accorded by the capital market to generic innovation productivity (Chan et al., 2001; Deng et al., 1999) and innovation efficiency (Hirshleifer et al., 2013) internationally using firm-level market value models. This serves to extend the non-linear least squares regression model findings in Hall et al. (2005) and the Fama-MacBeth (1973) regression findings in Hirshleifer et al. (2013) methodologically.[4]

We then disaggregate innovation productivity measures that are similar to those used in Deng et al. (1999) and Hirshleifer et al. (2013) innovation efficiency measures to account for 'clean' and 'dirty' innovation production and efficiency, respectively. Our main finding is that 'clean' innovation is associated with an economically important and positive Tobin's Q influence, relative to the influence of dirty innovation.

We provide a brief outline of our main findings. Consistent with the view that the capital market evaluates clean innovation positively, we find that the economic impact of generating a clean patent, per million dollars of book value, is 3.77%. With respect to generating a citation in respect to a clean patent, per million dollars of book value, is 1.27%. Similarly, we infer an economic impact of a citation on a clean patent, per million dollars of R&D expense, is 1% while noting that the comparable efficiency of R&D investments, in generating dirty patents, decreases the market value of the firm to the tune of 0.97% economic value.

We implement a series of additional robustness tests and show that our main findings remain qualitatively unchanged. Our robustness checks are based on a variety of dimensions: (i) we

[4]The initial findings concur with a large body of research which provides compelling evidence that the patent productivity of R&D and the citations received by these patents have a statistically and economically significant positive impact on firms' market value (e.g. Griliches (1981), Cockburn and Griliches (1987), Hall (1993, 2000), Hall et al. (2005), Jaffe (1986), Pakes (1985); Griliches et al. (1986), Chan et al. (2001), Eberhart et al. (2004).

use an array of complementary innovation measures (innovation productivity and efficiency) to proxy for clean and dirty innovation; (ii) we test if the results can be accounted for by including emerging technology innovation and a range of firm traits from the accounting based asset pricing literature; (iii) we test if the results are robust to an alternative estimator, the Fama-Macbeth two step regression estimator and (iv) we conduct a Heckman two-stage analysis to account for sample selection concerns. Our main findings are substantively unchanged across all these tests.

Our paper relates to the extensive literature that links firm-level environmental performance with its financial performance. Earlier papers include Gupta and Goldar (2005) which shows that capital markets can create financial and reputational incentives for pollution control in both developed and emerging market economies (see also Hamilton (1995) and Dasgupta et al. (2001)). More recent papers such as that of Guenster et al. (2011) show that eco-efficiency relates positively to operating performance and market value (see also, Ziegler et al. (2007) and Von Arx and Ziegler (2014)). Prior studies, however, suffer from several problems including small samples and the lack of objective environmental performance criteria. We do not rely on subjective analysis to characterize environmental performance. Instead, we study the documented environmental patenting activity and the efficiency of this patenting activity of publicly traded firms around the world. In addition this prior literature, unlike our paper, does not look at the critically important performance criterion of environmentally friendly patented innovation (IPCC 2014), with a view to improving the mix of technologies used to produce and consume energy. It does not, hence, examine whether this type of environmental performance can be related to financial performance and capital market values.

The remainder of the paper is organized as follows. The next section presents a discussion of the related literature in respect to possible mechanisms which can inter-relate market valuations and environmental innovation. Section 3 presents our data sources and characterizes our sample.

Section 4 presents our econometric methodology. Section 5 presents our results. Section 6 concludes.

2 Theoretical background: Market valuation implications of 'green' business decisions

It is well established that investments in R&D activities affect firms' future financial performance and are thus reflected in firms' market value (Grandi et al., 2009).[5] As the returns to R&D investments will typically accrue over a number of years, stock prices or market value should provide, given market information efficiency arguments, useful information on their expected future benefits. Empirical studies analysing the relationship between R&D investments and market value typically model the market value relative to tangible assets (Tobin's Q) as a function of intangible assets (R&D capital), and show that the R&D-market value relationship is consistently positive (Ballardini et al., 2005). There exists, however, significant differences in the market value of R&D investments across time, sectors and countries (Grandi et al., 2009). With regards to clean and dirty innovation specifically, the literature identifies two potentially countervailing relationships between investments in environmental innovation and financial performance.

Low-carbon and more generally environmental innovation by firms can be evaluated positively in the capital market as it can increase expected firm-level cash-flows (revenues less costs) and/or reduce the risk of these cash flows. There is a variety of potential mechanisms which can link firm-level environmental innovation and financial performance. Due to the plethora of emissions trading systems, climate and energy policies around the world (Ellerman, Marcantonini, and Zaklan, Ellerman et al.) , such innovation not only has generic research and development expenditure implications for future firm operating cash flows and risks (Hall, 2000) and (Czar-

[5]Following the seminal contribution of Griliches (1981), a large number of papers have used stock market value as an indicator of firms' expected performance from investing in R&D (e.g. Blundell et al. (1999); Toivanen et al. (2002); Czarnitzki et al. (2006); Hall and Oriani (2006); Bloch (2008).

nitzki et al., 2006). It also reflects recipient firms' expected environmental taxes and subsidies and financial penalties for environmental policy violations.

First, to the extent that environmental innovation is a measure of environmental performance, investors can link pro-active environmental innovation to lower firm risk. For instance, environmental performance can proxy for (i) high-skilled management (Bowman and Haire, 1975) and labour conditions at the firm and thus the firm's capacity to attract high-quality employees (Turban and Greening, 1997) and increasing employee morale and productivity (Dowell et al., 2000); (ii) operational efficiency (Porter and Van der Linde, 1995); and (iii) sales benefits in existing markets (Klassen and McLaughlin, 1996) and in new markets (Porter and Van der Linde, 1995) due to improved corporate and brand reputation with regulators, employees and the public (Corbett and Muthulingam, 2008; Russo and Fouts, 1997). More generally, (iv) environmental innovation can be regarded as a less risky investment (Narver, 1971; Shane and Spicer, 1983; Spicer, 1978). There is also evidence that firms with high commitments towards corporate social responsibility offer lower wage and enjoy higher employee productivity due to better recruitment, higher intrinsic motivation (many employees prefer a socially responsible employer and will accept a lower wage to achieve this), and a more effort-promoting corporate culture (Nyborg and Zhang, 2013; Brekke and Nyborg, 2008).

Second, climate change innovation can serve to mitigate risks of losses from crises or new regulation[6] (Reinhardt, 1999) and prevent expenses due to lawsuits and legal settlements (Karpoff et al., 2005). Investors can, hence, assign a lower discount rate to firms which are high environmental performers which would accord the firm a higher market value (and lower expected stock returns).

Finally, climate change innovation can attract funds from ethical investors who can prefer

[6] Calel and Dechezlepretre (2016) show that the European Union Emission's Trading System has had a quick causal impact on technological change in the form of new patenting activity.

firms with good track records of environmental performance (Heinkel et al., 2001). This interest on the part of ethical investment funds can reduce the cost of capital for the firm when it seeks to raise finance in the capital markets.

To the contrary, it is also possible that corporate investment in environmental innovation can deteriorate a firm's financial performance (Walley and Whitehead, 1994; Palmer et al., 1995). Climate change innovation can also, thus, be associated with a negative stock market valuation impact. Fisher-Vanden and Thorburn (2011) and Jacobs et al. (2010) show that emissions reductions can be associated with significant negative market reactions. In particular, the stock market may respond negatively to such innovation due to the possibility that the capital budget of the firm is deteriorated by such investment. For instance, it may be interpreted by participants in the capital market that pertinent environmental legislation is not binding at present or in the future. Environmental subsidies which are sought or the avoidance of financial penalties in respect to the emission of pollutants, which has motivated the environmental patenting activity, can be ascribed a lower probability by capital market participants, than by firm management.

Two results from the broad empirical literature on the market valuation of R&D are, in addition, worth mentioning here. First, firms' market share positively impacts on the valuation of R&D (Blundell et al., 1999), and firms conducting 'dirty' innovation are typically large incumbents, while firms engaged in clean innovation are more likely to be new entrants. New firms are often the vehicle through which radical, game-changing innovations enter the market. Second, a decreasing relationship between market uncertainty and the valuation of R&D investments has been observed - although only up to a certain point (Oriani and Sobrero, 2008). Since the demand for clean innovation fundamentally depends upon environmental policies, which are naturally uncertain, this could lower the premium associated with pursuing environmental R&D investments.

Given these opposing theoretical predictions, the net effect of clean innovation on market value is an empirical question.

3 Data and Variables

This section presents our sample of firm level traits and patenting data. It also presents the construction of our key innovation productivity and efficiency variables of interest and our set of control variables.

3.1 Data on firm-level traits

Our dataset comprises firm trait data sourced in the Worldscope Database, which presents information on the largest firms internationally. The original sampled database comprises 47,420 firms in 40 countries. From the original sample of firms, we eliminate firms for which the ISIN No. is missing, and we retain firms in the home market where the ISIN No. is the same for two firms in two different markets. Next, we drop firms with negative total assets, market capitalization or common cash dividend paid. We also drop firms for which we have less than 5 consecutive firm-year observations between 1995 and 2012 across all the firm-level variables used in the basic regression - year-end market capitalisation, capital expenditure, and earnings before interest, tax and amortisation. Until this point all the data is sourced from Worldscope on an annual frequency. The final firm-count is 25,255 firms from Worldscope. Next, we get data from Datastream for these 25,255 firms.

3.2 Data on firm-level patenting

We use patent data to identify innovation in clean and dirty technologies. To construct our innovation variables, we have drawn data from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office. PATSTAT is the largest international patent database and includes patent data from over 80 patent offices, including all of the major offices such as the United States Patent and Trademark office (USPTO), the Japan patent office (JPO) and the China National Intellectual Property Administration office (CNIPA). Therefore PATSTAT data cover close to the population of all worldwide patents since the 1980s. In PATSTAT, patent documents are categorized according to the new Cooperative Patent Classification system (CPC), the International Patent Classification (IPC) and national classification systems.

Our selection of patent classification codes for clean technologies relies on previous work by the OECD Environment Directorate.[7] We examine areas of clean patenting activity related to energy generation from renewable and non-fossil sources (wind, solar, hydro, marine, biomass, geothermal and energy from waste), combustion technologies with mitigation potential (for example combined heat and power), other technologies with potential contribution to emissions mitigation (in particular energy storage), electric and hybrid vehicles and energy conservation in buildings. We refer to these areas as climate change mitigation innovation or in short 'clean' innovation.

Our selection of patent classification codes for dirty technologies relies on Noailly and Smeets (2015) for electricity generation technologies and on Aghion et al. (2016) for the automobile industry. Our dirty environmental innovation pertains to IPC codes in different technological classes, including steam engine plants, gas turbine plants, combustion engines, steam generation, combustion apparatus and furnaces.

For each patent we know at which date it was filed (the application date), when it was first published (the publication date) and, if it was ever granted by the patent office, when the granted patent was published. In our study we focus on patent publication date as it is reasonable to expect that capital market participants will become aware of the new patents at

[7] See www.oecd/environment/innovation

this date. However, we also use grant publication date in our sensitivity analyses as this is the time when the future ability of the invention to generate revenues becomes known to market participants.

Our database in providing the identity of the patent applicants also facilitates matching clean and dirty patents with distinct patent applicants. Our analysis focuses on a sample of published patents and citations filed by around 15,000 firms belonging to the top 12 country leaders in clean innovation [8] over the period 1995-2012. We primarily study the patents and citations that are published by the USPTO, however for robustness we also conduct our analysis to the patents and citations published by the European patent office (EPO).

3.3 Variables

This section reports our principal variables of interest, including Tobin's Q and innovation productivity and innovation efficiency variables.

3.3.1 Key variables of interest

Dependent variable

The dependent variable in all our Model specifications is the natural logarithm Tobin's Q ratio which is the market value of firm i in year t to its replacement cost

$$Tobin'sQ = Q = \frac{Total\_assets - Book + Market\_Value}{Total\_assets} \tag{1}$$

where $Book$ is the book value of equity and $Market\_Value$ is the Market Capitalization.

Innovation productivity variables

Our innovation productivity variables are inspired by prior literature (Chan et al., 2001; Deng et al., 1999). We use R&D expense over book value of equity, $RDBE$ (worldscope # 05491 is

[8] The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain.

book value per share) (Chan et al., 2001), patents over book value of equity, $Pat/Book$ (Deng et al., 1999) and adjusted patent citation (Gu, 2005) over book value of equity, $Cit/Book$, as our innovation productivity variables.

RDBE is defined as the ratio of the R&D expense of firm i in year t scaled by the book value of equity in year t

Similarly, we define Pat/Book as the ratio of firm i's patents published in year t scaled by the book value of equity

$$\frac{Pat_{i,t}}{Book_{i,t}} = \frac{Patents_{i,t}}{Book_{i,t}} \tag{2}$$

To construct our citation productivity variable we ensure that the citations count is observable to investors in the market when they make investment decisions, Gu (2005) uses citations received in the year t with respect to patents granted in the previous five years. C_{ik}^{t-j} is the number of citations received in year t by patent k for firm i which is granted in year $t-j$ (j=1...5). This number is scaled by the average number of citations received in year t by all patents of the same subcategory granted in year $t-j$ (j=1...5).[9] N_{t-j} is the total number of patents granted in year $t-j$ to firm i. This method for adjusting citations is in line with Gu (2005) and Hirshleifer et al. (2013). It is an attempt to adjust for citation propensity attributed to differences in technology fields, grant year and the year in which the citation occurs. We define Cit/Book as follows

$$\frac{\mathrm{Cit}_{i,t}}{\mathrm{Book}_{i,t}} = \frac{\sum_{j=1}^{T}\sum_{k=1}^{N_{t-j}} C_{ik}^{t-j}}{Book_{i,t}}. \tag{3}$$

We further dissaggregate our patent and citation productivity variables as 'clean', 'dirty' and 'other'. For example 'clean' patent productivity is defined as follows

[9]Patent subcategories are defined based on the International Patent Classification.

$$\frac{Pat\_clean_{i,t}}{Book_{i,t}} = \frac{\text{Clean Patents}_{i,t}}{Book_{i,t}} \tag{4}$$

where Clean Patents denote the number of clean patents of firm i published in year t.

Innovation efficiency variables

We do not wish to focus exclusively on clean or dirty innovation productivity variables, but rather we wish to also focus on the efficiency with which research and development ($R\&D$) expenditure is used to generate that output. We use two proxies for the measurement of clean/dirty innovation efficiency which are tailored variants on those proxies used in Hirshleifer et al. (2013). First, we study clean/dirty patents scaled by R&D capital, Pat_clean/RDC and Pat_dirty/RDC.[10] Second, we study adjusted clean/dirty patent citations scaled by R&D expenses, Cit_clean/RD and Cit_dirty/RD. Hence, whereas Hirshleifer et al. (2013) study innovation efficiency, we focus on clean and dirty innovation efficiency.

Pat_clean/RDC is defined as the ratio of firm i's clean patents published in year t, scaled by its R&D capital in year $t-2$. It can be defined as

$$\frac{\text{Pat\_clean}_{i,t}}{\text{RDC}_{i,t-2}} = \frac{\text{Clean Patents}_{i,t}}{R\&D_{i,t-2} + 0.8 * R\&D_{i,t-3} + 0.6 * R\&D_{i,t-4} + 0.4 * R\&D_{i,t-5} + 0.2 * R\&D_{i,t-6}}. \tag{5}$$

The R&D capital is the five year cumulative R&D expenses assuming an annual depreciation rate of 20 % (Chan et al., 2001; Lev et al., 2005). In line with Lev and Sougiannis (1996), we assume a 5 year technology cycle with respect to the benefits of R&D.[11] The time lag between

[10] Research and development expense represents all direct and indirect costs related to the creation and development of new processes, techniques, applications and products with commercial possibilities; worldscope # 01201

[11] We set missing R&D to zero throughout but when we repeat our tests with variables with no missing R&D observations we have similar findings.

the innovation input (R&D capital) and output (patents) is to account for the average two year application to grant lag documented with respect to US patents (Hall et al., 2001). The use of cumulative R&D expenses in this innovation efficiency measurement is informed by R&D expenses over the preceding five years contributing to successful patent applications in t-2.

As the number of citations made to a firm's clean/dirty patents can reflect the patents' technological or economic importance, we also follow Hirshleifer et al. (2013) to define a new variable which is adjusted clean/dirty patent citations scaled by R&D expenses, Cit_clean/RD and Cit_dirty/RD. Specifically, Cit_clean/RD is defined as

$$\frac{\text{Cit\_clean}_{i,t}}{\text{RD}_{i,t}} = \frac{\sum_{j=1}^{T}\sum_{k=1}^{N_{t-j}} C_{ik}^{t-j}}{(R\&D_{i,t-3} + R\&D_{i,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7})}. \tag{6}$$

C_{ik}^{t-j} is defined above. The denominator, RD, is the summation of R&D expenses in years $t-3$ to $t-7$. This denominator is informed by the assumption that there is a 2-year application-grant time lag and that only R&D expenditure in year $t-2$ contributes to successful patent applications which are granted in year t.

Pat_dirty/RDC and Cit_dirty/RD are defined similarly, focusing on dirty patents only.

Control variables

The adopted set of control variables comprises firm traits that can play a role in the market's accordance of stock price value. The set of firm trait variables includes the inverse of book equity, $1/BE$, capital expenditure (Worldscope # 04601) to market value, $CEME$ and advertisement expenditure to market value, $Advert$ (Worldscope # 01101). We control for capital expenditure and advertising expenditure because they are found to explain firm operating performance (e.g., Lev and Sougiannis (1996); Pandit et al. (2011)). The set of firm trait variables also includes *abnormal* earnings, $Earning_{abnormal}$ (the earnings, E is defined as earnings before interest tax depreciation and amortisation, Worldscope # 18198). To obtain abnormal earnings, $Earning_{abnormal}$, earnings, E, is adjusted by the corporate income tax rate, $\tau_{i,t}$ (Worldscope # 08346) on firm earnings and the annualised risk free rate, r_t (Datastream annualised 90/91 day annualised Treasury bill rate), multiplied by the book value of equity is deducted (Ohlson, 1995).

We also include the tax shelter associated with R&D expenditure, $taxRDBE$, as a control variable (Hirshleifer et al., 2013) and substantial R&D growth, RDG, (Eberhart et al., 2004). An episode of R&D growth (RDG) is captured in a dummy variable which is equal to one if there is an episode of growth (R&D expenditure is greater than 5% of total assets (worldscope # 02999) and of total sales (worldscope # 01001) and there is a growth of at least 5% in R&D expenditure and a growth of at least 5% in R&D expenditure scaled by total assets relative to the prior year) and is zero otherwise. Eberhart et al. (2004) report significantly positive abnormal stock returns following substantial R&D expenditure growth.. Finally, we include time and industry dummies in our regression specifications.

4 Econometric methodology and hypotheses tests

4.1 Methodology

4.1.1 Estimation of Market value as a function of Innovation productivity and efficiency variables using Firm-level Market-value Model

We follow Hall et al. (2005) and adopt the firm-level market-value model to evaluate the relationship between R&D investment and the market value of the firm. The chief novelty in our approach consists in the way we apply the model to assess if the stock market recognizes the value of innovation productivity and efficiency in the production of 'clean' and 'dirty' technologies. The market-value model used in Hall et al. (2005), Hall and Oriani (2006)and many other studies on valuation of R&D investments assume that a firm is valued as a combination of both tangible and intangible assets by the stock market. However, the intangible assets that are created by the R&D investments are often not factored in the computation of the dependent variable, Tobin's Q. The model represents the market value, V, of the firm i at a time t as a function of book value of tangible assets, $A_{i,t}$, replacement value of firm's knowledge assets, $K_{i,t}$, and the replacement value of the other intangible assets, $I^j_{i,t}$ and can be represented as below.

$$V_{i,t} = V(A_{i,t}, K_{i,t}, I^1_{i,t}, \ldots, I^n_{i,t}) \tag{7}$$

Assuming assets can be written in an additive and linearly separable fashion and neglecting the other intangible assets, the market-value model is expressed as

$$V_{i,t} = b(A_{i,t} + \gamma K_{i,t})^{\sigma} \tag{8}$$

where σ accounts for the non-constant scale effects in the market-value function, γ represents the shadow value of knowledge assets relative to a firm's tangible assets and b denotes the average market valuation coefficient of total assets of a firm and can be interpreted to account for a firm's monopoly position and its differential risk (Grandi et al., 2009). Simplifying the representation of the model by taking the natural logarithm on both sides of the equation and assuming that σ=1 we get the following model

$$logV_{i,t} = logb + log(A_{i,t}) + log(1 + \gamma\frac{K_{i,t}}{A_{i,t}}). \tag{9}$$

which further simplifies to

$$logQ_{i,t} = log(\frac{V_{i,t}}{A_{i,t}}) = logb + log(1 + \gamma\frac{K_{i,t}}{A_{i,t}}) \tag{10}$$

where $Q_{i,t}$ stands for Tobin's Q. From the above model, one can estimate the average effect of a unit currency invested in knowledge assets on the firm's market value.

In creating our innovation productivity and efficiency variables, we consider that the full value of R&D investments can be captured from investment in R&D to creation of patents to efficiency of R&D investment in generating patents, to the generation of citation and finally the efficiency of R&D investment in creating citations. So, in our specifications we use R&D over book value of equity ($RDBE$) as a proxy for R&D productivity; patents over book value of equity ($Pat/Book$) and patents over R&D Capital (Pat/RDC) as proxies for patent productivity and efficiency; and citations over book value of equity ($Cit/Book$) and citations over RD (Cit/RD) as proxies for citation productivity and efficiency. We further disaggregate these variables into 'clean', 'dirty' and 'other' components to determine their relative importance in assessing the market value of the firm.

We first assess the impact of each individual innovation productivity and efficiency variable on the Tobin's Q of the firm by estimating various specifications derived from the Models

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \tag{11}$$

and

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \tag{12}$$

We then disaggregate these variables as 'clean' and 'dirty' components and examine whether the stock market attaches any importance to these technology classes separately and analyze the relative importance of each productivity and efficiency variable. For this, we estimate the various specifications of following Models:

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \tag{13}$$

and

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \tag{14}$$

where Pat^* and Cit^* denote the 'clean', 'dirty' or 'other' knowledge asset.

4.1.2 Estimation of Market value as a function of Innovation productivity and efficiency stocks using Ohlson's accounting based asset valuation Model

We adapt the Ohlson (1989) accounting-based asset valuation model to examine whether, and, if so, to what extent, the stock market assimilates the information content in clean and dirty innovation production and efficiency.[12] This model allows a test of whether clean and dirty innovation expenses explain market value and of any difference between their market value contributions. Ohlson (1989) derives the following valuation equation:

$$M_{i,t} = BE_{i,t} + \beta_0[E_{i,t}(1 - \tau_{i,t}) - r * BE_{i,t}] + \beta_1[\tau_{i,t} RD_{i,t}] + \alpha * Z_{i,t}, \tag{15}$$

where $M_{i,t}$ is the market value of the i^{th} firm at time t. $[E_{i,t}(1-\tau_{i,t})-r*BE_{i,t}]$ is a measure of abnormal earnings discussed above and initially defined in Ohlson (1989); $[\tau_{i,t} RD_{i,t}]$ accounts for the tax shelter associated with $R\&D$ expenditure; $Z_{i,t}$ is a vector of other information variables. Other variables are as defined above.

[12]This general asset pricing framework is also used in Barth et al. (1998); Sougiannis (1994); Ohlson (1995) and Hirshleifer et al. (2013) among others. It is recommended in Brennan's 1991 review paper (Brennan, 1991)

In our adaptation of this accounting-based asset valuation model, we use natural logarithm of Tobin's Q as the dependent variable and we include 'clean' and 'dirty' innovation productivity and efficiency variables, and the control variables used in Hirshleifer et al. (2013) as our vector of controls (RDG, Earning$_{abnormal}, invBE, CEME, Adverts, taxRDBE$[13])

We run non-linear least squares regressions in line with Hall et al. (2005) as well as Fama-MacBeth (1973) annual cross-sectional regressions at the firm level. Our robustness tests regression specifications are derived from the following models:

$$\begin{aligned} logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} \\ + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \end{aligned} \tag{16}$$

$$\begin{aligned} logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} \\ + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it} \end{aligned} \tag{17}$$

$$\begin{aligned} logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} \\ + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it} \end{aligned} \tag{18}$$

[13]See Table 1 of the definition of these variables

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it} \quad (19)$$

Pat^*, and Cit^*, are 'clean', 'dirty' or 'other' patents and citations.

4.2 Testable Hypotheses

Our point of departure is the well-established notion that stock markets can provide useful information on the value and expected performance of R&D investments (Griliches, 1981; Hall, 1993; Hirshleifer et al., 2013). Assuming efficient capital markets, traded security prices can provide an unbiased estimate of the present value of discounted future cash flows (Malkiel and Fama, 1970). What we examine in this paper, which has not been studied previously, is whether clean firm-level innovation productivity and efficiency are valued in capital markets around the world, in particular compared to dirty innovation productivity and efficiency.

As a result of the two conflicting potential outcomes of climate change related innovation (positive/ negative stock price impact) we test the following major hypothesis: Is clean climate change innovation priced positively or negatively in the capital market?

Furthermore, we test if 'dirty' innovation is positively or negatively priced in the capital market. The price impact of dirty innovation can provide some insight into whether the market response to clean innovation is due to climate change related innovation or whether it is due to capital budget efficiency on the part of the firm. We can conduct a difference-in-differences test with respect to the relative stock price impacts of clean and dirty innovation. This method can help to isolate the price impact with respect to clean innovation and abstract from generic price impacts of $R\&D$ investment.

Although our main focus in this paper is the productivity of patents in these areas of innovation, we also start by analysing the efficiency of the productivity of patents more generally, in order to verify that our results are in line with the previous literature which has looked at patents in all technological classes.

5 Descriptive Statistics and Empirical findings

5.1 Descriptive Statistics

The global rate of growth of production of environmentally friendly 'clean' technologies, vis-a-vis 'dirty' technologies, can be observed in Figure 1, which compares the aggregate clean and dirty patents and citations published by the US Patent office. This Figure reports a slight increase in the number of dirty patents published during the period 1995-2002, though there is no substantial change in the number of patents published yearly from 2002 to 2012. In contrast, there is a considerable increase in the number of clean patents published with an average growth of 13.58% per year. Turning to Figure 2, it identifies the top 12 country leaders in clean and dirty innovation.[14] These countries are ranked based on the number of clean and dirty patents published by the US Patent office. All the dirty technology producing countries, except Italy, are also among the clean technology producing countries. So, if there is a high level of innovation both dirty and clean innovation tend to prevail. A comparison of the aggregate clean and dirty patents published to these countries underscores the rising importance of environmentally friendly technologies in these nations.

[Please insert Figure 1 and Figure 2 about here.]

[14] The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain. The top 12 dirty innovation producing countries in descending order are: Japan, USA, Germany, Korea, France, Sweden, Finland, Italy, Taiwan, Great Britain, Canada, Netherlands.

To assess whether firms have a net incentive or disincentive to produce clean technologies, we construct our innovation productivity ($RDBE$, $Pat/Book$ and $Cit/Book$) and innovation efficiency variables (Pat/RDC and Cit/RD) and further disaggregate these variables into 'clean', 'dirty' and 'other' components for investigating their distinct influences on the Tobin's Q of the firm. The descriptive statistics for these variables are shown in Table 2.

[Please insert Table 2 about here.]

For our dataset, firms on an average allocate 4% of their book value of equity to R&D investments. Also, the clean and dirty innovation relative to book value of equity and R&D is a small fraction of total innovation. For instance, while clean and dirty patents over book value of equity account for 3.74% and 0.49%, these same patents over R&D Capital account for 2.62% and 0.68% respectively.

5.2 Empirical findings

5.2.1 Association between Tobin's Q and Innovation productivity and efficiency variables

Tables 3 and 4 report the results for the non-linear regression specifications which are derived from the firm-level market value model and are similar to those reported in Hall et al. (2005). We first determine the impact of the innovation productivity and efficiency variables on a firm's Tobin's Q (Table 3), and, then, disaggregate these variables into clean, dirty and other components to assess their distinctive impact on a firm's Tobin's Q (Table 4). All our Model specifications include time and industry dummies. Since R&D productivity is highly correlated with the firm's individual effect (Hall et al., 2005), we exclude firm-fixed effect to sidestep overcorrection.

Table 3 reports the results for specifications derived from equations (12) and (13). The results suggest that on an average, R&D, patent and citation productivity ($RDBE$, $Pat/Book$

and $Cit/Book$) positively correlate to Tobin's Q. In the light of the new international data, this corroborates the main findings reported in Hall et al. (2005). We also assess the impact of the efficiency of R&D investments in generating patents and citations on the Tobin's Q (Hirshleifer et al., 2013) to find that innovation efficiency variables (Pat/RDC, Cit/RD) are also positively associated with Tobin's Q. To determine the economic impact of these variables, we estimate the corresponding semi-elasticities, the results of which can be found in Appendix A. For example, the semi-elasticities with respect to citation over book ($Cit/Book$) for specification 3 suggest that an additional citation per million dollars of book value of equity is associated with an increment of 1.1% in Tobin's Q, respectively. Similarly, for specification 4 and 5, we find that the patents over R&D capital (Pat/RDC) and citations over RD (Cit/RD) are positively associated with the Tobin's Q at an economic impact of 1%.

[Please insert Table 3 about here.]

To determine whether the capital markets incentivize clean innovation vis-a-vis dirty innovation, we disaggregate patents over book ($Pat/Book$), citations over book ($Cit/Book$), patents over R&D capital (Pat/RDC), and citations over RD (Cit/RD) into clean, dirty and other components. For the first specification reported in Table 4, the clean patents over book ($Pat\_clean/Book$) is positively associated with the Tobin's Q at an economic impact of 3.77%. We also find that the clean citation over book ($Cit\_clean/Book$) is positively associated with Tobin's Q at an economic impact of 1.27% (specification 2 of Table 4). Additionally, we disaggregate our innovation efficiency variables and find that the clean citations over RD (Cit_clean/RD) is positively related to the dependent variable with an economic impact of 1% (specification 4 of Table 4). We find that the clean patents over R&D capital (Pat_clean/RDC) is positively related to Tobin's Q, though this result is not significant (specification 3 of Table 4). However, efficiency of R&D investments in generating dirty patents decreases the market value of the firm to the tune of 0.97% economic value (specification 3 of Table 4). Significantly, the t-test for the difference between coefficients of clean and dirty patents over book ($Pat\_clean/Book - Pat\_dirty/Book = 0$), patents over R&D capital ($Pat\_clean/RDC - Pat\_dirty/RDC = 0$), citations over book ($Cit\_clean/Book - Cit\_dirty/Book = 0$), and citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are statistically different from zero at a 5% level. The results for semi-elasticities for Tables 4 and t-test can be found in Tables A2 and A5 (Panel A) in the Internet Appendix A, respectively.

[Please insert Table 4 about here.]

5.2.2 Robustness tests

We are skeptical of the results reported in Table 4 as it possible that the estimates of clean innovation productivity and efficiency may be relaying the effect of emerging technologies on the firm's Tobin's Q. Emerging technologies are the new and disruptive innovations such as Information technologies, that are positively associated with both the firm's Tobin's Q as well as with clean technologies. Hence, the omission of emerging technologies may upwardly bias the estimates of clean innovation productivity and innovation efficiency. Therefore, we extend the Models reported in Table 4 to include the patent and citation productivity ($Pat\_emtech/Book$, $Cit\_emtech/Book$) in emerging technologies and the corresponding efficiency variables (Pat_emtech/RDC, Cit_emtech/RD) as controls and find no substantial change in the estimates of clean innovation productivity and innovation efficiency. This substantiates the results reported in Table 4.[15] We also find that the t-test for the difference between coefficients of clean and dirty patents over book ($Pat\_clean/Book - Pat\_dirty/Book = 0$), patents over R&D capital ($Pat\_clean/RDC - Pat\_dirty/RDC = 0$), citations over book ($Cit\_clean/Book - Cit\_dirty/Book =$

[15]The baseline regression results hold even when we consider the sample of firms with non-zero patents and the sample of firms producing both clean and dirty technologies (See Internet Appendix C).

0), and citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are statistically different from zero at a 5% level. The results for semi-elasticities for Tables 5 and t-test can be found in Tables A3 and A5 (Panel B) in the Internet Appendix A, respectively.

[Please insert Table 5 about here.]

As a robustness test, we extend the non-linear regression models reported in Table 4. We extend these models to include firm traits in line with the Ohlson's accounting based asset valuation model cited in Hirshleifer et al. (2013). In this heavily parameterized setting, our main results hold well in respect to clean and dirty citations over RD (Cit_clean/RD, Cit_dirty/RD), as indicated in specification 4 of Table 6. We also include patent and citation productivity and efficiency with respect to emerging technologies (Specifications 5-8 of Table 6), and again find that our results are robust with respect to clean and dirty citations over RD (Cit_clean/RD, Cit_dirty/RDC), as indicated in specification 8 of Table 6. The estimates of clean citation efficiency, Cit_clean/RD, reported in specifications 4 and 8 of Table 6 are similar to the one reported in Table 4 having the same economic impact of 1.04% on the firm's Tobin's Q. We also find that the efficiency of R&D investments in generating dirty citations (Cit_dirty/RD) decreases the Tobin's Q of the firm to the tune of 0.99%. For specifications 4 and 8 we find that difference between coefficients of clean and dirty citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are statistically different from zero at a 5% level. The results for semi-elasticities for Tables 6 and t-test can be found in Tables A4 and A5 (Panel C and D) in the Internet Appendix A, respectively.

[Please insert Table 6 about here.]

We also adopt the Fama-MacBeth estimator to assess the Models in Tables 4 and 5[16] and

[16] Please refer Table D2 in the Internet Appendix D

this confirms clean innovation premium reported in those Tables. The economic impact of clean innovation productivity and efficiency is similar to those derived from Tables 4 and 5, with the exception of clean patent productivity (Pat_clean/Book), which is three fold of the corresponding clean patent productivity (Pat_clean/Book) derived from Table 4.

[Please insert Table 8 about here.]

Further, we conduct robustness of the specifications included in Table 4 by including patent and citation productivity and efficiency with respect to emerging technologies ($Pat\_emtech/Book$, $Cit\_emtech/Book$, Pat_emtech/RDC, Cit_emtech/RD), and controlling for firm traits in line with the Ohlson's accounting based asset pricing model. These Models are estimated using the Fama-MacBeth estimator and our main results of the stock market that yield significantly more value to clean as opposed to dirty innovation productivity and innovation efficiency remain unchanged[17][18].

As demand in the market and generic government policies inform a firm's decision to innovate in a particular area, we posit that the 5-year change in the Environmental policy stringency score (Botta and Koźluk, 2014) would proxy for the appetite of the investors and consumers. Therefore, we add the difference between one-year and six-year lag of Environmental policy stringency score of the US (EPSlag1-EPSlag6) and emerging technology variants of innovation productivity and efficiency variables to the baseline regression models (Models in Table 4) and find that there is clean innovation premium with respect to efficiency of R&D investment in generating citations. We argue that this finding is economically relevant as citations show the importance of a particular innovation and further propel innovation in that area[19].

[17] Please refer Table D3 in the Internet Appendix D

[18] The main results hold even when we construct the innovation productivity and efficiency variables with respect to the grant date instead of publication date.

[19] Please refer Table C1 in the Internet Appendix C

5.2.3 Test to address Endogeneity concern

In our study sample selection bias may arise if managers choose to innovate in clean technologies more relative to dirty technologies. Therefore, to address sample selection we adopt the Heckman two stage 1979 regression approach. We model the likelihood of a firm to conduct clean innovation in the first stage of the Heckman Model. In the first stage we regress the variable Clean_firm, which is a dummy variable that takes the value 1 if a firm has a clean patent published by the USPTO and 0 otherwise, on RDG, Adverts, CEME, taxRDBE, $Earning_{\text{abnormal}}$, invBE, Total_assets, EPSlag1-EPSlag6, year and industry dummies and Emtech_firm which takes the value 1 if a firm has an emerging technology patent published by the USPTO and 0 otherwise.

$$Clean\_firm = \alpha + \gamma_1 Emtech\_firm_i + \gamma_2 RDG_{it} + \gamma_3 invBE_{it} + \gamma_4 taxRDBE_{it} + \gamma_5 CEME_{it} \quad (20)$$
$$+\gamma_6 Earning_{\text{abnormal it}} + \gamma_7 Adverts_{it} + \gamma_8 Total\_assets + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

The inverse Mills ratio is then computed from the first stage and is included as an explanatory variable in Models 1 and 2 of Tables 4 and 6. We find that our inference of a clean innovation premium remains, despite this correction. These results could be found in Table 7.

[Please insert Table 7 about here.]

5.2.4 Does the clean innovation premium hold in the European markets?

To check if our results hold in a different jurisdiction, we run the robustness tests for the patents and citations published by the European Patent Office (EPO). We find a positive association between clean patent productivity ($Pat\_clean/Book$) and Tobin's Q and this result is statistically significant at 5%. We also find a negative and significant association between dirty citation productivity ($Cit\_dirty/Book$) and efficiency variables (Cit_dirty/RD) with the Tobins'Q[20].

[20] Please refer Table B2 in the Internet Appendix B

To check if clean innovation and innovation efficiency are biased upwards, we include the patent and citation productivity and efficiency with respect to emerging technologies and thus find that the results do not change substantially[21]. These Models were estimated using Non-linear least squares method.

Additionally, we estimate these Models using Fama-MacBeth estimator and find that our main results hold with regard to clean and dirty patent productivity and efficiency. We also extend these models to include a patent and citation productivity and efficiency with respect to emerging technologies ($Pat\_emtech/Book$, $Cit\_emtech/Book$, Pat_emtech/RDC, Cit_emtech/RD) and thus find that the results do not change substantially[22].

Further, we estimate the influence of 'clean' and 'dirty' innovation productivity and efficiency variables on the Tobin's Q of the firm while controlling for emerging technology variants of innovation productivity and efficiency variables and firm traits in line with the Ohlson's accounting based asset pricing model cited in Hirshleifer et al. (2013). In this heavily parameterized setting, our main results hold well in respect to clean and dirty patent productivity and efficiency ($Pat\_clean/Book$, $Pat\_dirty/Book$, Pat_clean/RDC and Pat_dirty/RDC), as indicated in specifications 1, 3, 5 and 7 of Table B7 in the Internet Appendix B.

6 Conclusion

Innovation productivity is critically important for firm- and national-level competitiveness in international markets (Porter, 1992). Innovation productivity to curtail, and ultimately reverse, environmental degradation (*i.e.* 'clean' innovation) can prove vital to establish a sustainable market economy around the world (Allen and Yago, 2011). Such a sustainable market economy will mitigate market failures and serve to protect air, water, fisheries, wildlife, and biodiversity.

[21] Please refer Table B3 in the Internet Appendix B

[22] Please refer Tables B5 and B6 in the Internet Appendix B

In this paper, we raise the question of whether there is a sufficiency of incentive for firms to pursue strategies of clean environmentally-supportive innovation, as opposed to carbon-emitting dirty innovation activities.

We use a unique dataset covering 15,217 listed firms across 12 countries to measure the relationship between market value and innovation activity. We disaggregate annual patent counts by technology, distinguishing between clean, dirty and other technologies. Our dataset also includes patent citation data which is used to proxy for patent quality.

We start by verifying the value accorded by the capital market to generic innovation and innovation efficiency internationally, in a non-linear regression model setting (Hall et al., 2005). This serves to establish the validity of our data and empirical set-up.

Our main contribution is that we elicit capital market associations with the disaggregated Deng et al. (1999) innovation productivity measure and Hirshleifer et al. (2013) innovation efficiency measures to account for 'clean' and 'dirty' innovation production and efficiency, respectively. We report that 'clean' innovation efficiency is typically associated with an economically important and positive Tobin's Q influence, while the capital market ascribes no (or a negative) market value influence to 'dirty' innovation efficiency. The relative Tobin's Q association of 'clean' vis-a-vis 'dirty' innovation is significant and economically important across measurements. These main results are robust with respect to variation in model specification, estimator and control variables.

References

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). The environment and directed technical change. *American economic review 102*(1), 131–66.

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy 124*(1), 1–51.

Allen, F. and G. Yago (2011). Environmental finance: Innovating to save the planet. *Journal of Applied Corporate Finance 23*(3), 99–111.

Ballardini, F., A. Malipiero, R. Oriani, M. Sobrero, and A. Zammit (2005). Do stock markets value innovation? A meta-analysis. *A Meta-Analysis (January 2005).*

Barth, M. E., W. H. Beaver, and W. R. Landsman (1998). Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics 25*(1), 1–34.

Bloch, C. (2008). The market valuation of knowledge assets. *Econ. Innov. New Techn. 17*(3), 269–284.

Blundell, R., R. Griffith, and J. Van Reenen (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Botta, E. and T. Koźluk (2014). Measuring environmental policy stringency in OECD countries.

Bowman, E. H. and M. Haire (1975). A strategic posture toward corporate social responsibility. *California management review 18*(2), 49–58.

Brekke, K. A. and K. Nyborg (2008). Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics 30*(4), 509–526.

Brennan, M. J. (1991). A perspective on accounting and stock prices. *The Accounting Review 66*(1), 67–79.

Calel, R. and A. Dechezlepretre (2016). Environmental policy and directed technological change: Evidence from the european carbon market. *Review of economics and statistics 98*(1), 173–191.

Chan, L. K., J. Lakonishok, and T. Sougiannis (2001). The stock market valuation of research and development expenditures. *The Journal of Finance 56*(6), 2431–2456.

Cockburn, I. M. and Z. Griliches (1987). Industry effects and appropriability measures in the stock markets valuation of R&D and patents.

Corbett, C. and S. Muthulingam (2008). An empirical investigation of the depth of adoption of the LEED green building standards. Technical report, Working Paper, UCLA.

Czarnitzki, D., B. H. Hall, and R. Oriani (2006). The market valuation of knowledge assets in US and European firms. *The Management of Intellectual Property, Cheltenham Glos*, 111–131.

Dasgupta, S., B. Laplante, and N. Mamingi (2001). Pollution and capital markets in developing countries. *Journal of Environmental Economics and management 42*(3), 310–335.

Deng, Z., B. Lev, and F. Narin (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal 55*(3), 20–32.

Dowell, G., S. Hart, and B. Yeung (2000). Do corporate global environmental standards create or destroy market value? *Management science 46*(8), 1059–1074.

Eberhart, A. C., W. F. Maxwell, and A. R. Siddique (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance 59*(2), 623–650.

Ellerman, A. D., C. Marcantonini, and A. Zaklan. The EU ETS: Eight Years and Counting. *Robert Schuman Centre for Advanced Studies (RSCAS) Research Paper Series* (Paper No. 2014/04).

Fisher-Vanden, K. and K. S. Thorburn (2011). Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and management 62*(3), 430–445.

Grandi, A., B. H. Hall, and R. Oriani (2009). R&D and financial investors. *Evaluation and Performance Measurement of Research and Development, Cheltenham, UK: Edward Elgar*, 143–165.

Griliches, Z. (1981). Market value, R&D, and patents. *Economics letters 7*(2), 183–187.

Griliches, Z., A. Pakes, and B. H. Hall (1986). The value of patents as indicators of inventive activity.

Gu, F. (2005). Innovation, future earnings, and market efficiency. *Journal of Accounting, Auditing & Finance 20*(4), 385–418.

Guenster, N., R. Bauer, J. Derwall, and K. Koedijk (2011). The economic value of corporate eco-efficiency. *European Financial Management 17*(4), 679–704.

Gupta, S. and B. Goldar (2005). Do stock markets penalize environment-unfriendly behaviour? Evidence from India. *Ecological economics 52*(1), 81–95.

Hall, B. (2000). Innovation and market value,[w:] Productivity, innovation and economic performance, edsd. R. Barro, G. Mason, M. O'Mahoney.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *RAND Journal of economics*, 16–38.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The NBER patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H. and R. Oriani (2006). Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany, and Italy. *International Journal of Industrial Organization 24*(5), 971–993.

Hall, P. A. (1993). Policy paradigms, social learning, and the state: The case of economic policymaking in britain. *Comparative politics*, 275–296.

Hamilton, J. T. (1995). Pollution as news: Media and stock market reactions to the toxics release inventory data. *Journal of environmental economics and management 28*(1), 98–113.

Heinkel, R., A. Kraus, and J. Zechner (2001). The effect of green investment on corporate behavior. *Journal of financial and quantitative analysis 36*(4), 431–449.

Hicks, J. R. (1932). The theory of wages. *London: Macmillan*.

Hirshleifer, D., P.-H. Hsu, and D. Li (2013). Innovative efficiency and stock returns. *Journal of Financial Economics 107*(3), 632–654.

Jacobs, B. W., V. R. Singhal, and R. Subramanian (2010). An empirical investigation of environmental performance and the market value of the firm. *Journal of Operations Management 28*(5), 430–441.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value.

Jaffe, A. B., R. G. Newell, and R. N. Stavins (2005). A tale of two market failures: Technology and environmental policy. *Ecological economics 54*(2-3), 164–174.

Johnstone, N., I. Haščič, and D. Popp (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental and resource economics 45*(1), 133–155.

Karpoff, J. M., J. R. Lott, Jr, and E. W. Wehrly (2005). The reputational penalties for environmental violations: Empirical evidence. *The Journal of Law and Economics 48*(2), 653–675.

Klassen, R. D. and C. P. McLaughlin (1996). The impact of environmental management on firm performance. *Management science 42*(8), 1199–1214.

Lev, B., B. Sarath, and T. Sougiannis (2005). R&D reporting biases and their consequences. *Contemporary Accounting Research 22*(4), 977–1026.

Lev, B. and T. Sougiannis (1996). The capitalization, amortization, and value-relevance of R&D. *Journal of accounting and economics 21*(1), 107–138.

Malkiel, B. G. and E. F. Fama (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance 25*(2), 383–417.

Narver, J. C. (1971). Rational management responses to external effects. *Academy of Management Journal 14*(1), 99–115.

Newell, R. G., A. B. Jaffe, and R. N. Stavins (1999). The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics 114*(3), 941–975.

Noailly, J. and R. Smeets (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management 72*, 15–37.

Nyborg, K. and T. Zhang (2013). Is corporate social responsibility associated with lower wages? *Environmental and Resource Economics 55*(1), 107–117.

Ohlson, J. A. (1989). Accounting earnings, book value, and dividends: The theory of the clean surplus equation. *Unpublished paper*.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary accounting research 11*(2), 661–687.

Oriani, R. and M. Sobrero (2008). Uncertainty and the market valuation of R&D within a real options logic. *Strategic Management Journal 29*(4), 343–361.

Pakes, A. (1985). On patents, R&D, and the stock market rate of return. *Journal of political economy 93*(2), 390–409.

Palmer, K., W. E. Oates, and P. R. Portney (1995). Tightening environmental standards: the benefit-cost or the no-cost paradigm? *Journal of economic perspectives 9*(4), 119–132.

Pandit, S., C. E. Wasley, and T. Zach (2011). The effect of research and development (R&D) inputs and outputs on the relation between the uncertainty of future operating performance and R&D expenditures. *Journal of Accounting, Auditing & Finance 26*(1), 121–144.

Popp, D. (2002). Induced innovation and energy prices. *American economic review 92*(1), 160–180.

Porter, M. and C. Van der Linde (1995). Green and competitive: Ending the stalemate. *The Dynamics of the eco-efficient economy: environmental regulation and competitive advantage 33*.

Porter, M. E. (1992). Capital disadvantage: America's failing capital investment system. *Harvard business review 70*(5), 65–82.

Reinhardt, F. L. (1999). Bringing the environment down to earth. *Harvard business review 77*(4), 149–57.

Russo, M. V. and P. A. Fouts (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of management Journal 40*(3), 534–559.

Shane, P. B. and B. H. Spicer (1983). Market response to environmental information produced outside the firm. *Accounting Review*, 521–538.

Sougiannis, T. (1994). The accounting based valuation of corporate r&d. *Accounting review*, 44–68.

Spicer, B. H. (1978). Investors, corporate social performance and information disclosure: An empirical study. *Accounting Review*, 94–111.

Toivanen, O., P. Stoneman, and D. Bosworth (2002). Innovation and the market value of UK firms, 1989–1995. *Oxford Bulletin of Economics and Statistics 64*(1), 39–61.

Turban, D. B. and D. W. Greening (1997). Corporate social performance and organizational attractiveness to prospective employees. *Academy of management journal 40*(3), 658–672.

Von Arx, U. and A. Ziegler (2014). The effect of corporate social responsibility on stock performance: New evidence for the USA and Europe. *Quantitative Finance 14*(6), 977–991.

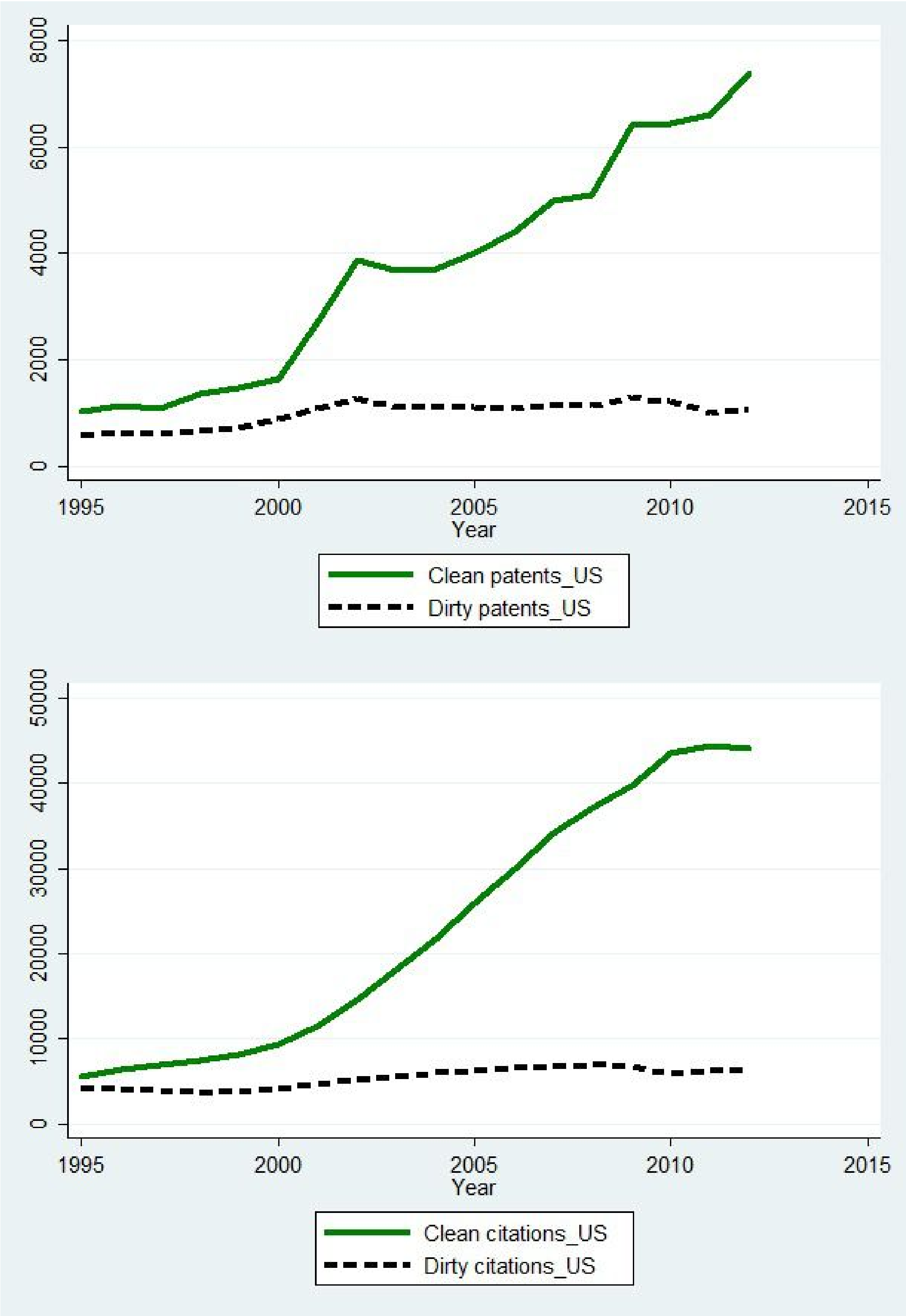
Walley, N. and B. Whitehead (1994). It's not easy being green. *Reader in Business and the Environment 36*, 81.

Ziegler, A., M. Schröder, and K. Rennings (2007). The effect of environmental and social performance on the stock performance of European corporations. *Environmental and Resource Economics 37*(4), 661–680.

Table 1: Variable Definitions

Variable	Definition
<u>Measures of firm value</u>	
Tobin's Q	Market value of the firm to the book value of tangible assets $(Total\_assets - Book + Market\_Value)\ /(Total\_assets)$.
Total_assets (millions of $)	Total Assets represents the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets.
Market_Value	Total market value of the company based on year end price and number of shares outstanding converted to U.S. dollars using the year end exchange rate.
Book (millions of $)	Book value of equity.
<u>Measure of R&D Intensity</u>	
RDBE	Research and Development expense divided by Book.
<u>Measures of Innovation Intensity</u>	
Pat/Book	Number of US patents of the firm, in any patent category, divided by Book.
Pat*/Book	As per Pat/Book but US patent category is *: clean, dirty, other or emerging technologies.
Cit/Book	The numerator is the number of citations received in year t by US patent k, granted in year t-j (j=1-5) scaled by the average number of citations received in year t by all patents of the same subcategory granted in year t-j (j=1-5). This number is summed over the total number of patents granted in year t-j to firm i. The numerator is divided by the book value of equity.
Cit*/Book	As per Cit/Book but US patent category is *: clean, dirty, other or emerging technologies.
<u>Measures of Innovation Efficiency</u>	
Pat/RDC	Number of US patents of the firm divided by the 5-year cumulative R&D expenses, observed in year t-2, assuming a depreciation rate of 20% per annum.
Pat*/RDC	As per Pat/RDC but US patent category is *: clean, dirty, other or emerging technologies.
Cit/RD	The numerator is the number of citations received in year t by US patent k, granted in year t-j (j=1-5) scaled by the average number of citations received in year t by all patents of the same subcategory granted in year t-j (j=1-5). This number is summed over the total number of patents granted in year t-j to firm i. The numerator is divided by the summation of R&D expenses in years t-3 to t-7.
Cit*/RD	As per Cit/RD but US patent category is *: clean, dirty, other or emerging technologies.
<u>Firm traits</u>	
invBE	Inverse of Book.
CEME	Capital expenditure (funds used to acquire fixed assets other than those associated with acquisitions) to Market Value of Equity.
Adverts	Advertising expenditure to Market Value of Equity.
RDG	R&D growth; An episode of R&D growth (RDG) is captured in a dummy variable which is equal to one if there is an episode of growth (R&D expenditure is greater than 5% of total assets and of total sales and there is a growth of at least 5% in R&D expenditure and a growth of at least 5% in R&D expenditure scaled by total assets relative to the prior year) and is zero otherwise (Total sales measured in millions of $, is the gross sales and other operating revenue less discounts, returns and allowances).
$\mathrm{Earning}_{\mathrm{abnormal}}$	Abnormal earnings; earnings before interest tax depreciation and amortization, E, is adjusted by the corporate income tax rate, $\tau_{i,t}$ on firm earnings and the annualized risk free rate, r_t, multiplied by the book value of equity is deducted.
taxRDBE	Tax shelter associated with R&D expenditure
<u>Regulation</u>	
EPS	Environmental Policy Stringency Index (Botta and Kozluk, 2014); This index takes the value from 0 (least stringent) to 6 (most stringent) and is a country-specific stringency measure.

Figure 1: Clean and dirty patents and citations



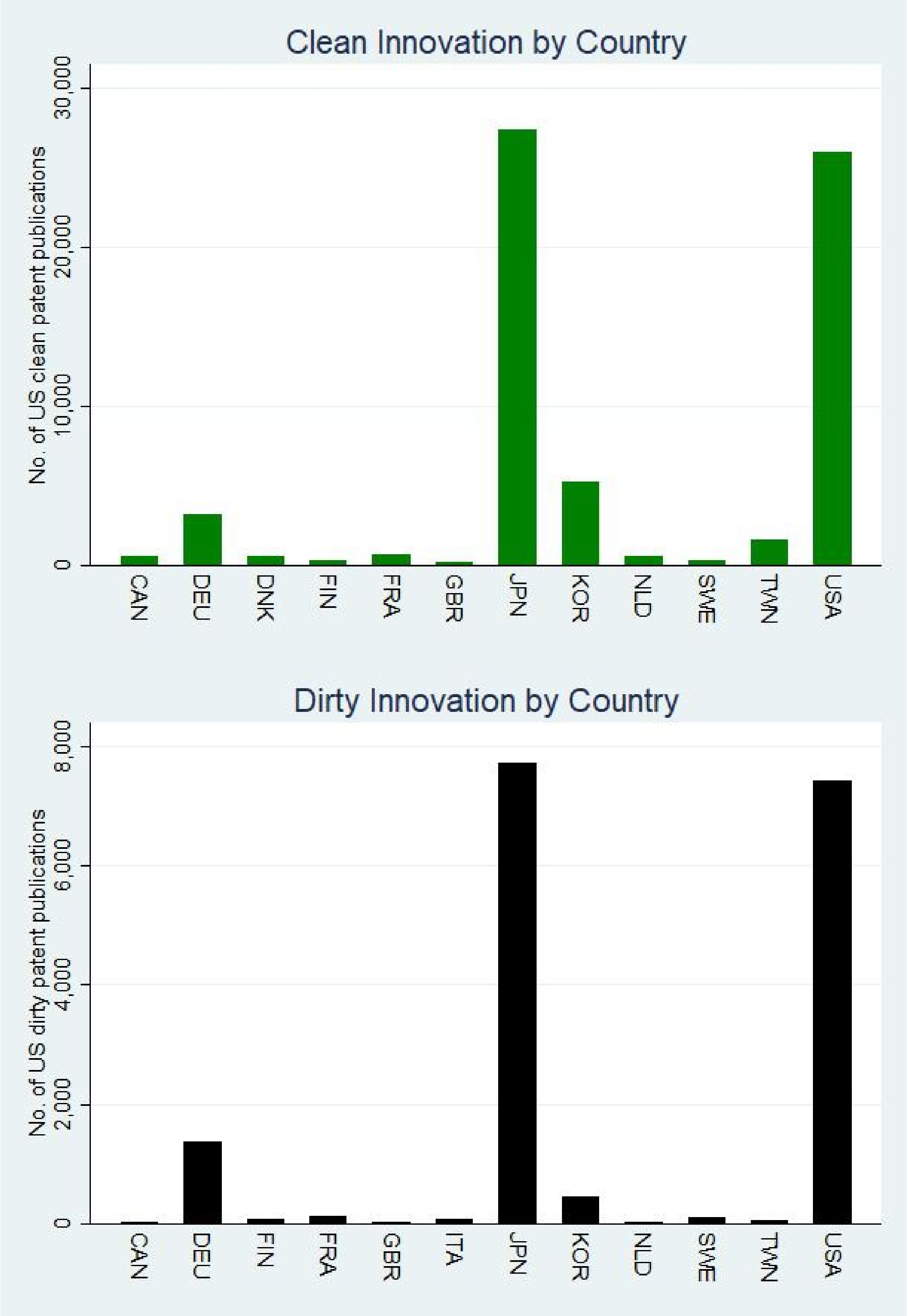
Notes. The Figure shows, over time, the number of published patents in clean and dirty technologies in the US (upper Panel) and shows related citations, accumulated in a 5-year window, in regard to clean and dirty innovations (lower Panel). We refer to Clean (Dirty) patents_US as the total number of clean (dirty) patents published by the USPTO during the period 1995-2012. We refer to Clean (Dirty) citations_US as the number of clean (dirty) patent citations of the firm, related to patents granted in the past 5 years by the USPTO.

Table 2: Summary Statistics

VARIABLES	N	Mean	Standard deviation
Innovation intensity			
RDBE	283,254	0.0426	1.451
Pat/Book	186,710	0.0267	2.210
Pat_clean/Book	186,710	0.000998	0.187
Pat_dirty/Book	186,710	0.000130	0.00595
Pat_emtech/Book	186,710	0.00621	0.718
Pat_other/Book	186,710	0.0256	2.039
Cit/Book	186,710	0.132	7.829
Cit_clean/Book	186,710	0.00475	0.494
Cit_dirty/Book	186,710	0.000561	0.0298
Cit_emtech/Book	186,710	0.0330	3.547
Cit_other/Book	186,710	0.127	7.504
Innovation efficiency			
Pat/RDC	283,253	0.0855	7.889
Pat_clean/RDC	283,254	0.00224	0.118
Pat_dirty/RDC	283,254	0.000585	0.0595
Pat_emtech/RDC	283,254	0.00734	0.250
Pat_other/RDC	283,253	0.0827	7.886
Cit/RD	283,254	0.210	8.506
Cit_clean/RD	283,254	0.00793	0.468
Cit_dirty/RD	283,254	0.00227	0.346
Cit_emtech/RD	283,254	0.0263	0.976
Cit_other/RD	283,254	0.200	8.451
Firm traits			
RDG	283,254	0.0377	0.190
invBE	283,254	-0.00775	0.945
taxRDBE	283,254	0.136	1.225
CEME	283,254	-0.0167	0.944
$\text{Earning}_{\text{abnormal}}$	283,254	-0.00287	0.939
Adverts	283,254	0.257	2.665

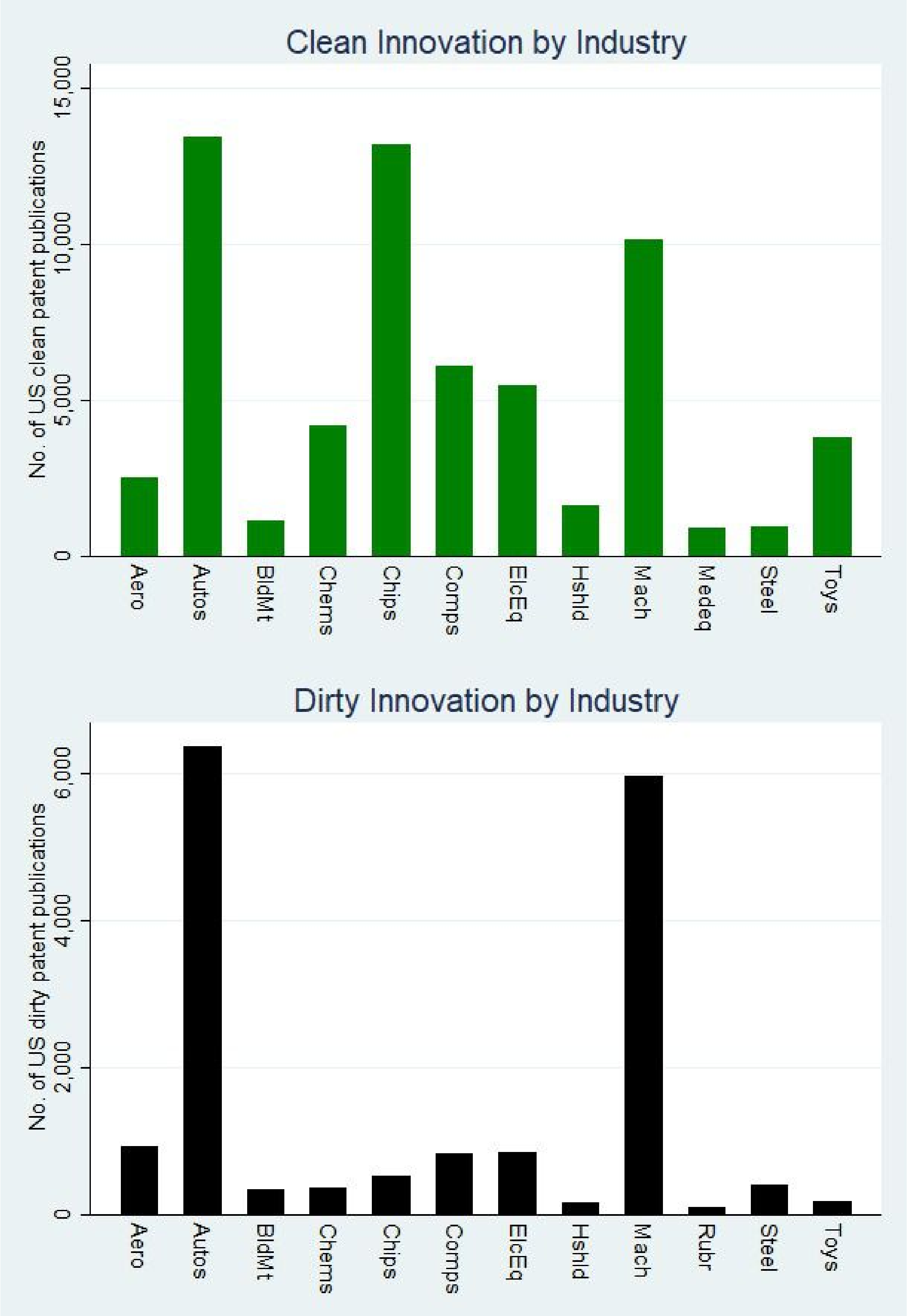
Notes. The Table presents summary statistics for Innovation productivity variables (RDBE, Pat/Book, Pat*/Book, Cit/Book and Cit*/Book), Innovation efficiency variables (Pat/RDC, Pat*/RDC, Cit/RD and Cit*/RD) and variables controlling for firm traits (Hirshleifer et al., 2013) during the period 1995-2012. The Variables are defined in Table 1.

Figure 2: Clean and dirty patent productivity by country



Notes. The Figure shows the number of published patents in clean and dirty technologies held by 12 leading clean technology producing countries (upper Panel) and 12 leading dirty technology producing countries (lower Panel).
The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain.
The top 12 dirty innovation producing countries in descending order are: Japan, USA, Germany, Korea, France, Sweden, Finland, Italy, Taiwan, Great Britain, Canada, Netherlands.

Figure 3: Clean and Dirty Patent productivity by Industry



Notes. The Figure shows the top 12 leading clean technologies producing industries (upper Panel) and the top 12 leading dirty technologies producing industries (lower Panel) in the 12 leading clean technology producing countries.

The top 12 clean innovation producing industries in descending order are: Autos (Automobile), Chips (Electronic equipment), Mach (Machinery), Comps (Computers), ElcEq (Electrical equipment), Chems (Chemicals), Toys (Recreation), Aero (Aircraft), Hshld (Consumer goods), BldMt (Construction materials), Steel (Steel) and Medeq (Medical equipment). The top 12 dirty innovation producing industries in descending order are: Autos (Automobile), Mach (Machinery), Aero (Aircraft), ElcEq (Electrical equipment), Comps (Computers), Chips (Electronic equipment), Steel (Steel), Chems (Chemicals), BldMt (Construction materials), Toys (Recreation), Hshld (Consumer goods) and Rubr (Rubber and Plastic).

Table 3: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.243***	0.243***	0.243***	0.244***	0.245***	0.245***
	(0.0363)	(0.0363)	(0.0362)	(0.0369)	(0.0369)	(0.0369)
RDBE	0.540***	0.533***	0.559***	0.259***	0.260***	0.260***
	(0.0377)	(0.0371)	(0.0389)	(0.0199)	(0.0199)	(0.0199)
Pat/Book	0.914***		0.456***			
	(0.0926)		(0.102)			
Cit/Book		0.170***	0.0999***			
		(0.0172)	(0.0184)			
Pat/RDC				0.00482**		0.00123
				(0.00186)		(0.000992)
Cit/RD					0.00874***	0.00847***
					(0.00204)	(0.00203)
Time FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO	NO	NO
Observations	87800	87800	87800	87799	87800	87799
Adjusted R^2	0.202	0.203	0.204	0.190	0.191	0.191

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.195***	0.195***	0.195***	0.195***
	(0.0393)	(0.0393)	(0.0391)	(0.0391)
RDBE	1.072***	1.072***	1.258***	1.256***
	(0.0778)	(0.0779)	(0.0814)	(0.0813)
Pat_clean/Book	1.803**			
	(0.615)			
Pat_dirty/Book	-0.972			
	(0.552)			
Pat_other/Book	0.170			
	(0.109)			
Cit/Book	0.144***			
	(0.0277)			
Cit_clean/Book		0.322**		
		(0.117)		
Cit_dirty/Book		-0.0876		
		(0.105)		
Cit_other/Book		0.139***		
		(0.0291)		
Pat/Book		0.216*		
		(0.108)		
Pat_clean/RDC			0.0588	
			(0.0375)	
Pat_dirty/RDC			-0.0355**	
			(0.0137)	
Pat_other/RDC			0.000536	
			(0.000684)	
Cit/RD			0.0144***	
			(0.00280)	
Cit_clean/RD				0.0505*
				(0.0236)
Cit_dirty/RD				-0.00551
				(0.00478)
Cit_other/RD				0.0136***
				(0.00266)
Pat/RDC				0.000608
				(0.000700)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	79285	79285	79284	79284
Adjusted R^2	0.215	0.215	0.212	0.212

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.195***	0.195***	0.195***	0.196***
	(0.0392)	(0.0393)	(0.0390)	(0.0391)
RDBE	1.071***	1.069***	1.252***	1.242***
	(0.0778)	(0.0777)	(0.0810)	(0.0807)
Pat_clean/Book	1.788**			
	(0.605)			
Pat_dirty/Book	-0.942			
	(0.567)			
Pat_emtech/Book	0.638			
	(0.355)			
Pat_other/Book	0.0829			
	(0.107)			
Cit/Book	0.141***			
	(0.0275)			
Cit_clean/Book		0.316**		
		(0.113)		
Cit_dirty/Book		-0.0820		
		(0.107)		
Cit_emtech/Book		0.249**		
		(0.0819)		
Cit_other/Book		0.115***		
		(0.0332)		
Pat/Book		0.207		
		(0.108)		
Pat_clean/RDC			0.0459	
			(0.0399)	
Pat_dirty/RDC			-0.0336*	
			(0.0131)	
Pat_emtech/RDC			0.195***	
			(0.0431)	
Pat_other/RDC			0.0000337	
			(0.000455)	
Cit/RD			0.0123***	
			(0.00266)	
Cit_clean/RD				0.0470*
				(0.0227)
Cit_dirty/RD				-0.00511
				(0.00481)
Cit_emtech/RD				0.0756***
				(0.0158)
Cit_other/RD				0.00822***
				(0.00242)
Pat/RDC				0.000532
				(0.000712)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	79285	79285	79284	79284
Adjusted R^2	0.216	0.215	0.213	0.213

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.244***	0.244***	0.244***	0.245***	0.244***	0.244***	0.245***	0.245***
	(0.0331)	(0.0331)	(0.0329)	(0.0330)	(0.0331)	(0.0331)	(0.0329)	(0.0330)
RDBE	0.325***	0.325***	0.229***	0.229***	0.324***	0.325***	0.228***	0.228***
	(0.0361)	(0.0361)	(0.0274)	(0.0274)	(0.0359)	(0.0361)	(0.0274)	(0.0274)
Pat_clean/Book	0.649				0.648			
	(0.487)				(0.485)			
Pat_dirty/Book	0.491				0.494			
	(0.929)				(0.930)			
Pat_emtech/Book					0.305			
					(0.272)			
Pat_other/Book	0.221*				0.207*			
	(0.0941)				(0.0945)			
Cit/Book	0.0620***				0.0619***			
	(0.0163)				(0.0162)			
Cit_clean/Book		0.144				0.144		
		(0.0877)				(0.0877)		
Cit_dirty/Book		-0.0151				-0.0151		
		(0.0423)				(0.0421)		
Cit_emtech/Book						0.0585		
						(0.0421)		
Cit_other/Book		0.0595***				0.0597**		
		(0.0167)				(0.0195)		
Pat/Book		0.236*				0.236*		
		(0.0936)				(0.0939)		
Pat_clean/RDC			0.0422				0.0316	
			(0.0281)				(0.0311)	
Pat_dirty/RDC			-0.0218				-0.0203	
			(0.0146)				(0.0140)	
Pat_emtech/RDC							0.144***	
							(0.0313)	
Pat_other/RDC			0.000974				0.000191	
			(0.000852)				(0.000536)	
Cit/RD			0.00662***				0.00532***	
			(0.00174)				(0.00159)	
Cit_clean/RD				0.0446*				0.0429*
				(0.0209)				(0.0208)
Cit_dirty/RD				-0.00159*				-0.00152**
				(0.000629)				(0.000585)
Cit_emtech/RD								0.0394***
								(0.00971)
Cit_other/RD				0.00614***				0.00329*
				(0.00165)				(0.00139)
Pat/RDC				0.00110				0.00109
				(0.000872)				(0.000885)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Firm-level controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	87800	87800	87799	87799	87800	87800	87799	87799
Adjusted R^2	0.248	0.248	0.244	0.245	0.248	0.248	0.245	0.245

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3, 4, 7 and 8)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005 and Hirshleifer et al., 2013 with the inclusion of firm-level control variables, year and industry fixed-effects. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7: Heckman sample selection 2^{nd} stage Model: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.383***	0.383***	0.435***	0.435***
	(0.0784)	(0.0784)	(0.0723)	(0.0723)
RDBE	0.0593***	0.0563***	0.0622***	0.0606***
	(0.00988)	(0.00949)	(0.00929)	(0.00892)
Pat_clean/Book	0.351***		0.218***	
	(0.0625)		(0.0579)	
Pat_dirty/Book	0.0883		0.0569	
	(0.263)		(0.242)	
Pat_other/Book	-0.0689***		-0.0646***	
	(0.00929)		(0.00859)	
Cit/Book	0.00712***		0.00983***	
	(0.00169)		(0.00157)	
Cit_clean/Book		0.0276***		0.0248***
		(0.00557)		(0.00516)
Cit_dirty/Book		0.0508		0.0952
		(0.107)		(0.0994)
Cit_other/Book		0.00255		0.00669***
		(0.00161)		(0.00150)
Pat/Book		-0.0228***		-0.0337***
		(0.00445)		(0.00414)
Inverse Mills Ratio	-0.10731***	-0.10720***	-0.06192***	-0.0620553***
	(0.00662)	(0.00663)	(0.00623)	(0.00624)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	YES	YES
Observations	130,798	130,798	130,798	130,798
Censored observations	115,432	115,432	115,432	115,432
Uncensored observations	15,366	15,366	15,366	15,366
Wald Chi^2	3847.09	3827.85	7254.20	7245.59
Prob > Chi^2	0.0000	0.0000	0.0000	0.0000
Rho	-0.23543	-0.23509	-0.14898	-0.14927
Sigma	0.45579	0.45600	0.41564	0.41573

Notes. The Table presents the regression results of various specifications of the 2^{nd} stage Heckman Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

The likelihood of a firm to conduct clean innovation is modeled in the 1^{st} stage of Heckman sample selection Model

$$Clean\_firm = \alpha + \gamma_1 Emtech\_firm_i + \gamma_2 RDG_{it} + \gamma_3 invBE_{it} + \gamma_4 taxRDBE_{it} + \gamma_5 CEME_{it} + \gamma_6 Earning_{\text{abnormal it}} + \gamma_7 Adverts_{it} + \gamma_8 Total\_assets + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

where Clean_firm and Emtech_firm are indicator variables that take the value 1 if a firm has a USPTO published patent and 0 otherwise. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable in the 2^{nd} stage Model is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)
Intercept	0.0409*	0.0410*	0.0362	0.0365
	(0.0219)	(0.0219)	(0.0228)	(0.0227)
RDBE	0.289***	0.288***	0.309***	0.309***
	(0.0267)	(0.0264)	(0.0365)	(0.0365)
Pat_clean/Book	2.319**			
	(0.930)			
Pat_dirty/Book	-0.329			
	(1.259)			
Pat_other/Book	0.0690			
	(0.0783)			
Cit/Book	0.0324*			
	(0.0164)			
Cit_clean/Book		0.289**		
		(0.104)		
Cit_dirty/Book		-0.115		
		(0.252)		
Cit_other/Book		0.0321*		
		(0.0165)		
Pat/Book		0.0770		
		(0.0739)		
Pat_clean/RDC			0.121**	
			(0.0455)	
Pat_dirty/RDC			0.0181	
			(0.0327)	
Pat_other/RDC			0.00190*	
			(0.00109)	
Cit/RD			0.00390***	
			(0.00107)	
Cit_clean/RD				0.0224**
				(0.00921)
Cit_dirty/RD				-0.00985
				(0.00857)
Cit_other/RD				0.00422***
				(0.00117)
Pat/RDC				0.00254*
				(0.00137)
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	79,285	79,285	79,284	79,284
avg. R-squared	0.193	0.192	0.188	0.188

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 3 and 4)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Internet Appendices A, B and C (supplementary material) for "Is firm-level clean or dirty innovation valued more?"

Internet Appendix A

Table A1: Semi-elasticities for determining the impact of aggregated Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 3

	(1)	(2)	(3)	(4)	(5)	(6)
RDBE	.5217474	.5157097	.5372324	.2705109	.2694972	.2374112
	(.0295228)	(.0292468)	(.030325)	(.0182262)	(.0181668)	(.0151506)
Pat/Book	.8830685		.4383902			
	(.0801154)		(.0953855)			
Cit/Book		.1644729	.0960732			
		(.0149254)	(.0172613)			
Pat/RDC				.0050269		.0012913
				(.001932)		(.0009818)
Cit/RD					.0090763	.0086154
					(.0020831)	(.0019414)
Observations	87,800	87,800	87,800	87,799	87,800	87,799

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

Table A2: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 4

	(1)	(2)	(3)	(4)
RDBE	.7887771	.7889395	.926713	.9253292
	(.0493505)	(.0493836)	(.0493264)	(.0493022)
Pat_clean/Book	1.326956			
	(.4503582)			
Pat_dirty/Book	-.7155582			
	(.4055185)			
Pat_other/Book	.125084			
	(.0799847)			
Cit/Book	.1060888			
	(.0199707)			
Cit_clean/Book		.2370153		
		(.0854426)		
Cit_dirty/Book		-.0644859		
		(.0769865)		
Cit_other/Book		.1022499		
		(.0210762)		
Pat/Book		.1588235		
		(.0795991)		
Pat_clean/RDC			.0433553	
			(.0275895)	
Pat_dirty/RDC			-.0261331	
			(.0100539)	
Pat_other/RDC			.000395	
			(.0005039)	
Cit/RD			.0106228	
			(.0020175)	
Cit_clean/RD				.037232
				(.0173066)
Cit_dirty/RD				-.004062
				(.0035169)
Cit_other/RD				.0100111
				(.0019199)
Pat/RDC				.0004483
				(.0005153)
Observations	79285	79285	79284	79284

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

Table A3: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 5

	(1)	(2)	(3)	(4)
RDBE	.7891391 (.0493976)	.7885194 (.0493652)	.9253519 (.0493095)	.9191813 (.0492154)
Pat_clean/Book	1.317652 (.4437327)			
Pat_dirty/Book	-.6944672 (.4174329)			
Pat_emtech/Book	.4699748 (.261107)			
Pat_other/Book	.0610632 (.0791405)			
Cit/Book	.1042786 (.0198877)			
Cit_clean/Book		.2332346 (.0830839)		
Cit_dirty/Book		-.0604374 (.079217)		
Cit_emtech/Book		.1837248 (.06009)		
Cit_other/Book		.0851515 (.0242553)		
Pat/Book		.1527549 (.0792739)		
Pat_clean/RDC			.033916 (.0294584)	
Pat_dirty/RDC			-.0248023 (.0096386)	
Pat_emtech/RDC			.1439807 (.0315708)	
Pat_other/RDC			.0000249 (.0003367)	
Cit/RD			.0091208 (.0019304)	
Cit_clean/RD				.0347744 (.0167299)
Cit_dirty/RD				-.0037782 (.0035559)
Cit_emtech/RD				.0559705 (.0115567)
Cit_other/RD				.006084 (.0017702)
Pat/RDC				.000394 (.0005264)
Observations	79,285	79,285	79,284	79,284

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

Table A4: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 6

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RDBE	.29763	.29777	.21502	.21484	.29662	.29782	.21371	.21339
	(.03110)	(.03110)	(.02450)	(.02450)	(.03102)	(.03109)	(.02442)	(.02448)
Pat_clean/Book	.59483				.59395			
	(.44610)				(.44392)			
Pat_dirty/Book	.45009				.45245			
	(.85132)				(.85242)			
Pat_emtech/Book					.27972			
					(.24796)			
Pat_other/Book	.20273				.18952			
	(.08567)				(.08622)			
Cit/Book	.05680				.05674			
	(.01476)				(.01473)			
Cit_clean/Book		.13162				.13165		
		(.08027)				(.08029)		
Cit_dirty/Book		-.01386				-.01389		
		(.03873)				(.03862)		
Cit_emtech/Book						.05363		
						(.03850)		
Cit_other/Book		.05453				.05474		
		(.01521)				(.01775)		
Pat/Book		.21614				.21606		
		(.08514)				(.08538)		
Pat_clean/RDC			.03956				.02958	
			(.02630)				(.02907)	
Pat_dirty/RDC			-.02040				-.01901	
			(.01367)				(.01309)	
Pat_emtech/RDC							.13505	
							(.02891)	
Pat_other/RDC			.00091				.00018	
			(.00080)				(.00050)	
Cit/RD			.00620				.00498	
			(.00161)				(.00147)	
Cit_clean/RD				.04181				.04014
				(.01950)				(.01939)
Cit_dirty/RD				-.00149				-.00142
				(.00059)				(.00055)
Cit_emtech/RD								.03690
								(.00901)
Cit_other/RD				.00575				.00308
				(.00153)				(.00130)
Pat/RDC				.00103				.00102
				(.00082)				(.00083)
Observations	87,800	87,800	87,799	87,799	87,800	87,800	87,799	87,799

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1, 2, 5 and 6) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3, 4, 7 and 8)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

Table A5: The Table reports the nonlinear hypothesis for the coefficients from the Models reported in Tables 4, 5, and 6

t Test for Non-linear least squares estimation	Chi^2	$Prob > Chi^2$
Pat_clean/Book - Pat_dirty/Book = 0	8.58	0.0034
Cit_clean/Book - Cit_dirty/Book = 0	6.28	0.0122
Pat_clean/RDC - Pat_dirty/RDC = 0	6.36	0.0116
Cit_clean/RD - Cit_dirty/RD = 0	5.81	0.0160

(a) Panel A: Test for Non-linear hypotheses after estimation for the Models reported in Table 4

t Test for Non-linear least squares estimation	Chi^2	$Prob > Chi^2$
Pat_clean/Book - Pat_dirty/Book = 0	8.38	0.0038
Cit_clean/Book - Cit_dirty/Book = 0	5.99	0.0144
Pat_clean/RDC - Pat_dirty/RDC = 0	3.44	0.0638
Cit_clean/RD - Cit_dirty/RD = 0	5.41	0.0201

(b) Panel B: Test for Non-linear hypotheses after estimation for the Models reported in Table 5

t Test for Non-linear least squares estimation	Chi^2	$Prob > Chi^2$
Pat_clean/Book - Pat_dirty/Book = 0	0.02	0.8822
Cit_clean/Book - Cit_dirty/Book = 0	2.62	0.1057
Pat_clean/RDC - Pat_dirty/RDC = 0	5.74	0.0166
Cit_clean/RD - Cit_dirty/RD = 0	4.90	0.0268

(c) Panel C: Test for Non-linear hypotheses after estimation for the Models (1-4) reported in Table 6

t Test for Non-linear least squares estimation	Chi^2	$Prob > Chi^2$
Pat_clean/Book - Pat_dirty/Book = 0	0.02	0.8848
Cit_clean/Book - Cit_dirty/Book = 0	2.62	0.1054
Pat_clean/RDC - Pat_dirty/RDC = 0	2.78	0.0957
Cit_clean/RD - Cit_dirty/RD = 0	4.57	0.0326

(d) Panel D: Test for Non-linear hypotheses after estimation for the Models (5-8) reported in Table 6

Internet Appendix B

Robustness tests for the patents and citations published by the European Patent Office (EPO). We study the influence of the innovation productivity and efficiency variables on the Tobin's Q of firms. The firms in this study have their headquarters in any of the top 12 clean innovation producing countries

Table B1: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.242***	0.239***	0.241***	0.244***	0.243***	0.244***
	(0.0361)	(0.0359)	(0.0360)	(0.0369)	(0.0370)	(0.0369)
RDBE	0.648***	0.809***	0.716***	0.261***	0.260***	0.261***
	(0.0432)	(0.0545)	(0.0502)	(0.0199)	(0.0199)	(0.0199)
Pat/Book	1.931***		1.555***			
	(0.186)		(0.241)			
Cit/Book		0.106***	0.0282*			
		(0.0121)	(0.0139)			
Pat/RDC				0.0262***		0.0255***
				(0.00586)		(0.00575)
Cit/RD					0.00158	0.000682
					(0.000917)	(0.000814)
Time FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO	NO	NO
Observations	87800	87800	87800	87800	87800	87800
Adjusted R^2	0.204	0.200	0.204	0.190	0.190	0.190

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.241***	0.241***	0.244***	0.244***
	(0.0360)	(0.0360)	(0.0369)	(0.0369)
RDBE	0.715***	0.730***	0.261***	0.261***
	(0.0503)	(0.0514)	(0.0199)	(0.0199)
Pat_clean/Book	3.315*			
	(1.362)			
Pat_dirty/Book	0.869			
	(1.104)			
Pat_other/Book	1.502***			
	(0.243)			
Cit/Book	0.0301*			
	(0.0139)			
Cit_clean/Book		0.00111		
		(0.00240)		
Cit_dirty/Book		-0.0624***		
		(0.0102)		
Cit_other/Book		0.0330*		
		(0.0148)		
Pat/Book		1.503***		
		(0.243)		
Pat_clean/RDC			0.0304	
			(0.0239)	
Pat_dirty/RDC			-0.0336	
			(0.0448)	
Pat_other/RDC			0.0259***	
			(0.00624)	
Cit/RD			0.000759	
			(0.000762)	
Cit_clean/RD				0.000403
				(0.00246)
Cit_dirty/RD				-0.00108***
				(0.000116)
Cit_other/RD				0.00114
				(0.000994)
Pat/RDC				0.0253***
				(0.00572)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	87800	87800	87800	87800
Adjusted R^2	0.204	0.204	0.190	0.190

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.241***	0.241***	0.244***	0.244***
	(0.0360)	(0.0360)	(0.0369)	(0.0369)
RDBE	0.714***	0.736***	0.260***	0.261***
	(0.0503)	(0.0524)	(0.0199)	(0.0199)
Pat_clean/Book	3.235*			
	(1.359)			
Pat_dirty/Book	0.881			
	(1.120)			
Pat_emtech/Book	2.205***			
	(0.640)			
Pat_other/Book	1.357***			
	(0.254)			
Cit/Book	0.0354*			
	(0.0144)			
Cit_clean/Book		0.00116		
		(0.00244)		
Cit_dirty/Book		-0.0623***		
		(0.0101)		
Cit_emtech/Book		0.00293		
		(0.0304)		
Cit_other/Book		0.0345*		
		(0.0152)		
Pat/Book		1.497***		
		(0.243)		
Pat_clean/RDC			0.0282	
			(0.0218)	
Pat_dirty/RDC			-0.0265	
			(0.0435)	
Pat_emtech/RDC			0.166**	
			(0.0630)	
Pat_other/RDC			0.0196***	
			(0.00576)	
Cit/RD			0.000279	
			(0.000771)	
Cit_clean/RD				0.000318
				(0.00244)
Cit_dirty/RD				-0.00108***
				(0.000115)
Cit_emtech/RD				0.00161*
				(0.000787)
Cit_other/RD				0.000817
				(0.00162)
Pat/RDC				0.0253***
				(0.00571)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	87800	87800	87800	87800
Adjusted R^2	0.204	0.204	0.191	0.190

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B4: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.0381	0.0382	0.0381	0.0381	0.0381	0.0381
	(0.0249)	(0.0249)	(0.0249)	(0.0249)	(0.0249)	(0.0249)
RDBE	0.0867***	0.0868***	0.0860***	0.0879***	0.0878***	0.0879***
	(0.0177)	(0.0188)	(0.0177)	(0.0189)	(0.0189)	(0.0189)
Pat/Book	0.343***		0.406***			
	(0.115)		(0.135)			
Cit/Book		0.0251***	-0.00203			
		(0.00597)	(0.0103)			
Pat/RDC				0.00892***		0.00948***
				(0.00218)		(0.00234)
Cit/RD					0.000429	-9.45e-05
					(0.000508)	(0.000593)
Industry FE	YES	YES	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO	NO	NO
Observations	87,800	87,800	87,800	87,800	87,800	87,800
avg. R-squared	0.181	0.177	0.182	0.177	0.176	0.177

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Table B5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)
Constant	0.0381	0.0381	0.0381	0.0381
	(0.0249)	(0.0249)	(0.0249)	(0.0249)
RDBE	0.0860***	0.0862***	0.0879***	0.0879***
	(0.0177)	(0.0177)	(0.0188)	(0.0189)
Pat_clean/Book	4.109***			
	(1.397)			
Pat_dirty/Book	0.777			
	(0.908)			
Pat_other/Book	0.410***			
	(0.134)			
Cit/Book	-0.00293			
	(0.0101)			
Cit_clean/Book		0.197		
		(0.143)		
Cit_dirty/Book		-0.228		
		(0.230)		
Cit_other/Book		-0.00133		
		(0.0107)		
Pat/Book		0.421***		
		(0.138)		
Pat_clean/RDC			0.152*	
			(0.0855)	
Pat_dirty/RDC			0.0870	
			(0.0954)	
Pat_other/RDC			0.0102***	
			(0.00288)	
Cit/RD			-1.04e-06	
			(0.000613)	
Cit_clean/RD				-0.00932
				(0.00688)
Cit_dirty/RD				-0.0145
				(0.0161)
Cit_other/RD				0.000181
				(0.000577)
Pat/RDC				0.0101***
				(0.00271)
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	87,800	87,800	87,800	87,800
avg. R-squared	0.183	0.183	0.178	0.177

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 3 and 4)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)
Intercept	0.0380	0.0380	0.0380	0.0381
	(0.0249)	(0.0249)	(0.0249)	(0.0249)
RDBE	0.0883***	0.0870***	0.0879***	0.0879***
	(0.0171)	(0.0176)	(0.0188)	(0.0189)
Pat_clean/Book	4.097***			
	(1.381)			
Pat_dirty/Book	0.693			
	(0.907)			
Pat_emtech/Book	0.794**			
	(0.301)			
Pat_other/Book	0.389**			
	(0.138)			
Cit/Book	0.00331			
	(0.0106)			
Cit_clean/Book		0.186		
		(0.142)		
Cit_dirty/Book		-0.228		
		(0.230)		
Cit_emtech/Book		0.0959**		
		(0.0403)		
Cit_other/Book		-0.0111		
		(0.0107)		
Pat/Book		0.431***		
		(0.139)		
Pat_clean/RDC			0.145*	
			(0.0823)	
Pat_dirty/RDC			0.0922	
			(0.0961)	
Pat_emtech/RDC			0.0750**	
			(0.0276)	
Pat_other/RDC			0.0097***	
			(0.00306)	
Cit/RD			-0.000269	
			(0.000612)	
Cit_clean/RD				-0.0106
				(0.00734)
Cit_dirty/RD				-0.0140
				(0.0151)
Cit_emtech/RD				0.00280**
				(0.00122)
Cit_other/RD				-0.000327
				(0.000746)
Pat/RDC				0.0103***
				(0.00275)
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	87,800	87,800	87,800	87,800
avg. R-squared	0.185	0.183	0.178	0.177

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 3 and 4)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Table B7: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.161***	0.161***	0.161***	0.161***	0.161***	0.161***	0.161***	0.161***
	(0.0319)	(0.0319)	(0.0319)	(0.0319)	(0.0319)	(0.0319)	(0.0319)	(0.0319)
RDBE	0.0550***	0.0553***	0.0610***	0.0610***	0.0569***	0.0561***	0.0610***	0.0610***
	(0.0145)	(0.0146)	(0.0155)	(0.0155)	(0.0144)	(0.0146)	(0.0155)	(0.0155)
Pat_clean/Book	3.413***				3.433***			
	(1.053)				(1.043)			
Pat_dirty/Book	0.703				0.643			
	(0.774)				(0.768)			
Pat_emtech/Book					0.811***			
					(0.195)			
Pat_other/Book	0.388***				0.385***			
	(0.116)				(0.122)			
Cit/Book	-0.00404				-0.000504			
	(0.0103)				(0.0105)			
Cit_clean/Book		0.191				0.181		
		(0.136)				(0.134)		
Cit_dirty/Book		0.0905				0.0917		
		(0.245)				(0.245)		
Cit_emtech/Book						0.0719*		
						(0.0397)		
Cit_other/Book		-0.00326				-0.0111		
		(0.0110)				(0.0103)		
Pat/Book		0.397***				0.407***		
		(0.120)				(0.120)		
Pat_clean/RDC			0.119*				0.114*	
			(0.0657)				(0.0634)	
Pat_dirty/RDC			0.0924				0.0962	
			(0.0810)				(0.0816)	
Pat_emtech/RDC							0.0581**	
							(0.0220)	
Pat_other/RDC			0.00922***				0.00922***	
			(0.00249)				(0.00287)	
Cit/RD			-3.31e-05				-0.000253	
			(0.000644)				(0.000633)	
Cit_clean/RD				-0.0110				-0.0117
				(0.00778)				(0.00810)
Cit_dirty/RD				0.00623				0.00736
				(0.0104)				(0.0107)
Cit_emtech/RD								-0.000450
								(0.00140)
Cit_other/RD				2.31e-05				-0.000307
				(0.000571)				(0.000670)
Pat/RDC				0.00908***				0.00921***
				(0.00233)				(0.00237)
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Firm-level controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	87,800	87,800	87,800	87,800	87,800	87,800	87,800	87,800
avg. R-squared	0.303	0.303	0.299	0.298	0.304	0.303	0.299	0.298

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 2, 4, 7 and 8)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Internet Appendix C

This appendix presents additional robustness tests. All the patents and citations are published by USPTO. We study the influence of disaggregated innovation productivity and efficiency variables on the Tobin's Q of firms. The firms in this study have their headquarters in any of the top 12 clean innovation producing countries

Table C1: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables

	(1)	(2)	(3)	(4)
Intercept	0.0375	-0.385***	-0.304***	-0.317***
	(6412.6)	(0.0466)	(0.0469)	(0.0469)
RDBE	0.291	1.099***	0.426***	0.430***
	(1864.3)	(0.0853)	(0.0376)	(0.0381)
Pat_clean/Book	-0.301			
	(1930.8)			
Pat_dirty/Book	2.794			
	(17910.6)			
Pat_emtech/Book	-0.0812			
	(520.2)			
Pat_other/Book	-0.0397			
	(254.8)			
Cit/Book	0.00966			
	(61.95)			
Cit_clean/Book		0.235		
		(0.206)		
Cit_dirty/Book		-0.0579		
		(0.459)		
Cit_emtech/Book		0.169		
		(0.0916)		
Cit_other/Book		0.251***		
		(0.0454)		
Pat/Book		0.752***		
		(0.199)		
Pat_clean/RDC			0.0617	
			(0.0591)	
Pat_dirty/RDC			-0.0369	
			(0.0265)	
Pat_emtech/RDC			0.382***	
			(0.0747)	
Pat_other/RDC			0.000317	
			(0.000974)	
Cit/RD			0.0110***	
			(0.00313)	
Cit_clean/RD				0.0771*
				(0.0370)
Cit_dirty/RD				-0.00287*
				(0.00112)
Cit_emtech/RD				0.111***
				(0.0242)
Cit_other/RD				0.00678*
				(0.00274)
Pat/RDC				0.00210
				(0.00176)
EPSlag1-EPSlag6	0.831	1.048***	1.144***	1.153***
	(5328.1)	(0.197)	(0.191)	(0.193)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	66697	66697	66696	66696
Adjusted R^2	0.188	0.212	0.198	0.198

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4(EPSlag1 - EPSlag6) + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4(EPSlag1 - EPSlag6) + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. The Innovation productivity and efficiency variables are defined in Table 1 and we refer to EPSlag1 and EPSlag6 as the one year and six year lag of EPS. And we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C2: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables. We run the baseline regression for firms with positive patents only.

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.304^{***}	0.307^{***}	0.307^{***}	0.301^{***}	0.303^{***}	0.303^{***}
	(0.0648)	(0.0645)	(0.0645)	(0.0649)	(0.0659)	(0.0659)
RDBE	0.838^{***}	0.772^{***}	0.762^{***}	0.980^{***}	0.979^{***}	0.980^{***}
	(0.0848)	(0.0830)	(0.0824)	(0.0926)	(0.0929)	(0.0930)
Pat/Book	0.559^{***}		0.145			
	(0.110)		(0.0922)			
Cit/Book		0.142^{***}	0.123^{***}			
		(0.0240)	(0.0249)			
Pat/RDC				0.00211^{*}		0.000218
				(0.00105)		(0.000489)
Cit/RD					0.0107^{***}	0.0107^{***}
					(0.00245)	(0.00245)
Time FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO	NO	NO
Observations	49343	49343	49343	49342	49343	49342
Adjusted R^2	0.227	0.230	0.230	0.222	0.225	0.225

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables. We run the baseline regression for firms with positive patents only.

	(1)	(2)	(3)	(4)
Intercept	0.306***	0.307***	0.303***	0.303***
	(0.0645)	(0.0645)	(0.0658)	(0.0658)
RDBE	0.760***	0.761***	0.980***	0.978***
	(0.0822)	(0.0823)	(0.0929)	(0.0928)
Pat_clean/Book	1.781**			
	(0.567)			
Pat_dirty/Book	-0.865			
	(0.512)			
Pat_other/Book	0.105			
	(0.0924)			
Cit/Book	0.121***			
	(0.0250)			
Cit_clean/Book		0.319**		
		(0.114)		
Cit_dirty/Book		-0.0718		
		(0.0936)		
Cit_other/Book		0.115***		
		(0.0260)		
Pat/Book		0.151		
		(0.0930)		
Pat_clean/RDC			0.0526	
			(0.0351)	
Pat_dirty/RDC			-0.0291*	
			(0.0127)	
Pat_other/RDC			0.000165	
			(0.000474)	
Cit/RD			0.0105***	
			(0.00242)	
Cit_clean/RD				0.0449*
				(0.0213)
Cit_dirty/RD				-0.00509
				(0.00431)
Cit_other/RD				0.00971***
				(0.00228)
Pat/RDC				0.000219
				(0.000487)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	49343	49343	49342	49342
Adjusted R^2	0.231	0.231	0.226	0.226

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C4: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables. We run the baseline regression for firms with positive patents only.

	(1)	(2)	(3)	(4)
Intercept	0.320***	0.320***	0.320***	0.322***
	(0.0685)	(0.0685)	(0.0687)	(0.0686)
RDBE	0.435***	0.437***	0.276***	0.275***
	(0.0464)	(0.0466)	(0.0305)	(0.0305)
Pat_clean/Book	0.744			
	(0.489)			
Pat_dirty/Book	0.593			
	(0.957)			
Pat_emtech/Book	0.299			
	(0.255)			
Pat_other/Book	0.219*			
	(0.0870)			
Cit/Book	0.0688***			
	(0.0158)			
Cit_clean/Book		0.157		
		(0.0854)		
Cit_dirty/Book		-0.0181		
		(0.0347)		
Cit_emtech/Book		0.0744		
		(0.0427)		
Cit_other/Book		0.0644***		
		(0.0185)		
Pat/Book		0.248**		
		(0.0863)		
Pat_clean/RDC			0.0335	
			(0.0312)	
Pat_dirty/RDC			-0.0183	
			(0.0147)	
Pat_emtech/RDC			0.125***	
			(0.0314)	
Pat_other/RDC			-0.0000739	
			(0.000403)	
Cit/RD			0.00403**	
			(0.00140)	
Cit_clean/RD				0.0427*
				(0.0203)
Cit_dirty/RD				-0.00147*
				(0.000574)
Cit_emtech/RD				0.0352***
				(0.00957)
Cit_other/RD				0.00227
				(0.00117)
Pat/RDC				0.000505
				(0.000637)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	53557	53557	53556	53556
Adjusted R^2	0.227	0.227	0.220	0.220

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables. Firms with both clean and dirty patents

	(1)	(2)	(3)	(4)
Intercept	1.496***	1.505***	1.481***	1.483***
	(0.0173)	(0.0130)	(0.0156)	(0.0150)
RDBE	0.0123	0.00329	0.0256	0.0240
	(0.0171)	(0.0114)	(0.0150)	(0.0142)
Pat_clean/Book	0.616*			
	(0.277)			
Pat_dirty/Book	-0.103			
	(0.214)			
Pat_other/Book	-0.0394			
	(0.0529)			
Cit/Book	0.0238*			
	(0.00995)			
Cit_clean/Book		0.124***		
		(0.0221)		
Cit_dirty/Book		-0.000688		
		(0.00882)		
Cit_other/Book		0.0152		
		(0.00823)		
Pat/Book		-0.00995		
		(0.0291)		
Pat_clean/RDC			0.00259	
			(0.00595)	
Pat_dirty/RDC			-0.00678*	
			(0.00327)	
Pat_other/RDC			-0.00167	
			(0.00249)	
Cit/RD			0.00125	
			(0.00185)	
Cit_clean/RD				0.0179
				(0.0136)
Cit_dirty/RD				-0.00313**
				(0.000998)
Cit_other/RD				-0.000313
				(0.000870)
Pat/RDC				-0.000527
				(0.000832)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	6593	6593	6593	6593
Adjusted R^2	0.215	0.218	0.197	0.204

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables. Firms with both clean and dirty patents

	(1)	(2)	(3)	(4)
Intercept	1.502***	1.506***	1.482***	1.485***
	(0.0153)	(0.0128)	(0.0153)	(0.0147)
RDBE	0.00549	0.00200	0.0248	0.0230
	(0.0146)	(0.0112)	(0.0146)	(0.0137)
Pat_clean/Book	0.549*			
	(0.277)			
Pat_dirty/Book	-0.155			
	(0.202)			
Pat_emtech/Book	-0.169*			
	(0.0778)			
Pat_other/Book	-0.0219			
	(0.0506)			
Cit/Book	0.0303**			
	(0.0112)			
Cit_clean/Book		0.123***		
		(0.0220)		
Cit_dirty/Book		-0.000478		
		(0.00868)		
Cit_emtech/Book		0.0106		
		(0.00661)		
Cit_other/Book		0.0205		
		(0.0144)		
Pat/Book		-0.0148		
		(0.0287)		
Pat_clean/RDC			-0.00340	
			(0.00559)	
Pat_dirty/RDC			-0.00463	
			(0.00285)	
Pat_emtech/RDC			0.0339*	
			(0.0161)	
Pat_other/RDC			-0.00409	
			(0.00212)	
Cit/RD			0.00118	
			(0.00184)	
Cit_clean/RD				0.0182
				(0.0135)
Cit_dirty/RD				-0.00278**
				(0.000914)
Cit_emtech/RD				0.00635
				(0.00347)
Cit_other/RD				-0.00126
				(0.000642)
Pat/RDC				-0.000409
				(0.000757)
Time FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	6593	6593	6593	6593
Adjusted R^2	0.216	0.218	0.199	0.206

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

and the Model (columns 3-4)

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \epsilon_{it}$$

that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Internet Appendix D

Table D1: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.0380	0.0380	0.0380	0.0378	0.0369	0.0361
	(0.0249)	(0.0249)	(0.0249)	(0.0249)	(0.0253)	(0.0255)
RDBE	0.0860***	0.0890***	0.0884***	0.0878***	0.0878***	0.0878***
	(0.0157)	(0.0166)	(0.0175)	(0.0189)	(0.0189)	(0.0189)
Pat/Book	0.150***		0.141**			
	(0.0473)		(0.0664)			
Cit/Book		0.0234***	0.0124			
		(0.00714)	(0.0104)			
Pat/RDC				0.00446***		0.00361**
				(0.00152)		(0.00142)
Cit/RD					0.00185***	0.00175***
					(0.000486)	(0.000464)
Industry FE	YES	YES	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO	NO	NO
Observations	87,800	87,800	87,800	87,799	87,800	87,799
avg. R-squared	0.180	0.181	0.183	0.176	0.177	0.177

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 4-6)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Table D2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)
Intercept	0.0410*	0.0410*	0.0363	0.0373
	(0.0219)	(0.0220)	(0.0226)	(0.0225)
RDBE	0.288***	0.288***	0.308***	0.308***
	(0.0266)	(0.0258)	(0.0364)	(0.0365)
Pat_clean/Book	2.301**			
	(0.920)			
Pat_dirty/Book	-0.348			
	(1.257)			
Pat_emtech/Book	0.309*			
	(0.160)			
Pat_other/Book	0.0422			
	(0.0724)			
Cit/Book	0.0357**			
	(0.0158)			
Cit_clean/Book		0.286**		
		(0.100)		
Cit_dirty/Book		-0.117		
		(0.252)		
Cit_emtech/Book		0.164***		
		(0.0486)		
Cit_other/Book		0.0268		
		(0.0161)		
Pat/Book		0.0461		
		(0.0707)		
Pat_clean/RDC			0.107**	
			(0.0478)	
Pat_dirty/RDC			0.0246	
			(0.0314)	
Pat_emtech/RDC			0.0787***	
			(0.0181)	
Pat_other/RDC			-0.000548	
			(0.00112)	
Cit/RD			0.00355***	
			(0.00104)	
Cit_clean/RD				0.0240**
				(0.00837)
Cit_dirty/RD				-0.00954
				(0.00848)
Cit_emtech/RD				0.0278**
				(0.00961)
Cit_other/RD				0.00259**
				(0.000924)
Pat/RDC				0.00249*
				(0.00140)
Industry FE	YES	YES	YES	YES
Firm-level controls	NO	NO	NO	NO
Observations	79,285	79,285	79,284	79,284
avg. R-squared	0.194	0.194	0.189	0.190

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

and the Model (columns 3 and 4)

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

Table D3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.161***	0.161***	0.159***	0.159***	0.161***	0.161***	0.160***	0.159***
	(0.0320)	(0.0320)	(0.0323)	(0.0323)	(0.0320)	(0.0320)	(0.0322)	(0.0323)
RDBE	0.0569***	0.0516***	0.0609***	0.0609***	0.0593***	0.0537***	0.0608***	0.0609***
	(0.0121)	(0.0120)	(0.0155)	(0.0155)	(0.0124)	(0.0128)	(0.0155)	(0.0155)
Pat_clean/Book	1.564*				1.549**			
	(0.739)				(0.726)			
Pat_dirty/Book	0.234				0.0946			
	(1.217)				(1.195)			
Pat_emtech/Book					0.291**			
					(0.116)			
Pat_other/Book	0.140***				0.169**			
	(0.0670)				(0.0595)			
Cit/Book	0.0160				0.0146*			
	(0.00939)				(0.00781)			
Cit_clean/Book		0.209**				0.205**		
		(0.0979)				(0.0975)		
Cit_dirty/Book		0.0775				0.0688		
		(0.179)				(0.175)		
Cit_emtech/Book						0.0301**		
						(0.0138)		
Cit_other/Book		0.0153				0.0158		
		(0.00926)				(0.00946)		
Pat/Book		0.142**				0.143**		
		(0.0638)				(0.0599)		
Pat_clean/RDC			0.0976**				0.0888**	
			(0.0393)				(0.0397)	
Pat_dirty/RDC			0.0169				0.0208	
			(0.0421)				(0.0420)	
Pat_emtech/RDC							0.0499***	
							(0.0132)	
Pat_other/RDC			0.00276***				0.000791	
			(0.000937)				(0.000928)	
Cit/RD			0.00152***				0.00133***	
			(0.000381)				(0.000331)	
Cit_clean/RD				0.0172**				0.0177**
				(0.00652)				(0.00626)
Cit_dirty/RD				0.00364				0.00387
				(0.0109)				(0.0110)
Cit_emtech/RD								0.0110***
								(0.00369)
Cit_other/RD				0.00232**				0.00186***
				(0.000829)				(0.000614)
Pat/RDC				0.00309**				0.00291**
				(0.00114)				(0.00115)
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Firm-level controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	87,800	87,800	87,799	87,799	87,800	87,800	87,799	87,799
avg. R-squared	0.305	0.305	0.299	0.299	0.306	0.306	0.300	0.300

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model

$$
\begin{aligned}
logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} \\
+\gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \\
\sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}
\end{aligned}
$$

and the Model (columns 2, 4, 7 and 8)

$$
\begin{aligned}
logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \\
\gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{\text{abnormal it}} + \gamma_9 Adverts_{it} + \\
\sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \epsilon_{it}
\end{aligned}
$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.