Inventing while you work: knowledge, non-R&D learning and innovation

You-Na Lee and John P. Walsh
Georgia Institute of Technology

ABSTRACT

“Intuition, judgment, creativity are basically expressions of capabilities for recognition and response based upon experience and knowledge (p. 128-9)” (Simon, 1997). Workers gain experience and knowledge in the course of their normal jobs. Therefore, innovative ideas can be generated from knowledge built from learning opportunities across the firm (not just the R&D lab). Employees working for different functions (R&D and outside of R&D) in an organization have different work practices and build their learning through different processes. Moreover, the relative effectiveness of learning by different work practices for innovation is contingent on nature of knowledge, characterized by generality (i.e., high mobility/transferability) and visibility (i.e., tighter links between actions and outcomes). Using multiple datasets combining public and private data and focusing on births of innovations, this study shows how the nature of knowledge affects differences in the innovation productivity of R&D and non-R&D work, with implications for how work reorganization can affect innovation. The paper concludes with a discussion of the implications of these insights for innovation management and policy.

Keywords: knowledge; learning; R&D innovation; non-R&D innovation

Acknowledgment: We thank Ashish Arora, Doug Noonan and the participants in the American Sociological Association Annual Meeting, Chicago and seminars at Georgia Tech School of Public Policy and Indiana University School of Informatics and Computing. The Inventor Survey data collection was supported by the Research Institute of Economy, Trade and Industry. This research project was supported by NSF (grant #1262418).

Inventing while you work: knowledge generality, visibility and non-R&D innovation

1. Introduction

Most research on the innovation process has focused on the results of R&D projects (Cohen, 2010). The positive relation between R&D intensity as an input and innovative performance as an output has become the canonical image for research on innovation (Cockburn and Griliches, 1987; Wakelin, 2001). While R&D is an important input to innovation, there is growing evidence that a significant share of innovation is not born from R&D (Arundel et al., 2008; Heidenreich, 2009; Rosenberg, 1982).

Recent work on European firms finds that about half of innovating firms do no formal in-house R&D (Arundel et al., 2008; Huang et al., 2011). Following on such observations, we have seen a recent growth in studies of non-R&D innovators, primarily in the context of low and medium-tech environments (Barge-Gil et al., 2011; Hervas-Oliver et al., 2012; Hervas-Oliver et al., 2011; Santamaria et al., 2009). These studies have found, for example, that non-R&D innovators, compared to their R&D counterparts, rely relatively more on innovation management tools such as human resource management and team work or marketing and organizational innovations to compensate for the lack of R&D capabilities (Heidenreich, 2009; Hervas-Oliver et al., 2015; Moilanen et al., 2014; Rammer et al., 2009).

In the US, according to NSF's Business R&D and Innovation Survey (BRDIS) 2011, out of all US firms, only 5% conduct R&D. Furthermore, out of all US product innovating firms, about 72% are non-R&D innovators. At the same time, R&D-active firms do have a higher probability of generating a product innovation than non-R&D-active firms (58% vs. 7%) (NSF, 2014). These prior studies and statistics suggest that non-R&D activity is an important input for innovation and that non-R&D innovators have different characteristics from R&D innovators.

However, the prior literature has overlooked within-firm heterogeneity in the sources of innovation (R&D and non-R&D). Generally, prior work (using firm-level measures) treats any innovation from R&D performing firms as R&D innovation and does not show if a given innovation originated from R&D or non-R&D work in the firm. This firm-level perspective also leads to a tendency to equate the R&D vs. non-R&D innovation comparison with that of large vs. small firms or high-tech vs. low-tech. However, innovation outside of R&D is not only a small firm or low-tech industry phenomenon, but also common in large firms or high-tech industries.

Therefore, taking a more balanced view that incorporates non-R&D innovation generated in R&D performing firms as well as innovation in non-R&D performing firms, we examine the factors that affect the relative composition of different types of innovation in an organization. We posit that there are different learning paths to innovate in an organization, which are explained by the relation between the work (i.e., R&D and non-R&D) and the nature of knowledge. We argue that two main characteristics (i.e. generality and visibility) of knowledge environments change the relative effectiveness of learning by R&D and non-R&D for innovation in an organization. To explain this relation, we adopt insights from sociology of work and

organizational learning literatures and combine them with the economics of innovation perspective.

This study has the following features. First, we focus on the originator of the innovation, linking the invention with the origin activity in the firm. While some prior research sees innovators in terms of ownership of innovation, for example, in the classification of technology adopter, innovator that conducts contract R&D or in-house R&D (Arundel et al., 2008; Huang et al., 2011), this study sees innovators in terms of origin of the innovation. In our sample, all innovations are internally generated in the inventing firm, allowing us to analyze learning and innovation within an organization (distinct from markets for technology).

Second, prior work emphasizes the factors that affect the likelihood of generating an innovation, for example, the impact of greater investment in R&D and/or activities outside of R&D (such as training, marketing and design) (Barge-Gil et al., 2011; Hervas-Oliver et al., 2011; Lööf and Johansson, 2014; Spithoven et al., 2011). However, our meaning of “non-R&D” in this study is the routine work activity of those outside of R&D, which is somewhat different from what these prior studies mean by non-R&D activity. Moreover, contrary to the prior work, our study focuses on estimating the relative likelihood of R&D versus non-R&D innovation in a firm, as a function of the nature of knowledge, which affects the relative effectiveness of R&D and non-R&D learning for innovation within a firm. We characterize knowledge environments in terms of generality (with high mobility or transferability) and visibility of knowledge (with tighter links between actions and outcomes). We test this relative innovation productivity using the unique structure of our invention-level data, rather than the more common firm-level data. This will also help guide future studies of how we can change the effectiveness of learning by different work practices in an organization, through the introduction of more effective firm strategy, management tools or organizational structures (e.g., Hervas-Oliver and Alborns-Garrigos, 2009; Lundvall and Johnson, 1994; Rammer et al., 2009).

Finally, we contribute to international statistics on non-R&D innovation. Non-R&D innovation in the US has not been widely studied. Using novel data on the US, this study will add new evidence on non-R&D innovation in comparison to R&D innovation, and also add data from a relatively high-tech environment such as the US (cf. Barge-Gil et al., 2011; Hervas-Oliver et al., 2011).

The structure of this paper is as follows. We first describe learning rooted in non-R&D work, distinct from learning in R&D, and discuss how this can explain an underexplored category of innovation, i.e., non-R&D innovation. Second, we discuss how different knowledge environments affect different innovation productivities, given R&D and non-R&D work, through their distinct modes of learning. Third, we test hypotheses built from the theory using an integrated dataset built from several data sources. Lastly, we conclude with results, implications, and future studies.

2. An extended view: distributed loci of innovation within an organization

Organizations can learn and innovate through a variety of processes (Malerba, 1992). Figure 1 characterizes innovation by locus of internal activity and type of innovation. The framework in Figure 1 does not draw a pure typology by origin and type of innovation, because, for example, sometimes innovation can be related to both product and process and also because there can be ambiguity about where the solution is conceptualized. However, this ontological framework with concepts and contrasts in innovation types is still useful to understand how innovations are created through different mechanisms (Winter, 2003).

First, organizations learn from direct experience. Learning by doing increases efficiency because workers improve their competencies in repeated procedures, which increases the frequency of successful output (Levitt and March, 1988). Libertyship building, aircraft production, and nuclear plant operation are well-known examples in the learning by doing literature (Benkard, 2000; Joskow and Rozanski, 1979; Thompson, 2012). Although simple adaptation or improved competencies over time in the production process enhance productivity, workers sometimes develop better ways to solve errors in the procedures beyond autonomous learning by doing (Argote and Darr, 2000; Darr et al., 1995). These examples of process innovation through learning by doing are positioned around area (A).

While the organizational learning literature discusses learning by cumulative experience, there is another stream of literature that argues that learning in the production process may be through deliberate R&D work, rather than direct production experience (Area (B) in Figure 1). Hatch and Mowery (1998) show that learning by doing in the semiconductor industry is the result of allocating engineering resources to learning in the context of new process introduction and analyzing production data to solve problems, rather than a by-product of production experience. Similarly, Sinclair et al. (2000) show in the case of a specialty chemical manufacturer that learning by production experience, measured by cumulative past output, does not directly affect cost reduction. It rather increases incentives to do process R&D or influences the choice of R&D projects, which generate cost reductions. They argue that R&D personnel or chemical engineers perform the actual experiments and modify production processes. Therefore, they conclude that even so-called learning by doing is largely the result of R&D activity.

Thus, organizational learning and innovation literatures have suggested that learning by cumulative experience is mostly related to process improvement, and even that might be by R&D. On the other hand, product innovation is seen as mainly driven by formally organized innovative effort, or R&D, as represented by area (C) (Aoki, 1991; Cohen, 2010; Kemp and Pearson, 2007).

-----INSERT FIGURE 1 ABOUT HERE-----

However, these perspectives neglect the possibility of informal deliberate learning and innovation (including product innovation) in non-R&D work. For example, for process innovations, Halle (1984) finds in his study of chemical engineers and production workers in a chemical plant that production workers routinely experiment and innovate in the production process. Production workers can make improvements in production and modify their equipment, which can eventually create significant changes from the original design (Kusterer, 1978).

Thomas (1994) also finds in an aluminum company that operators in the factory sometimes keep a “little black book” documenting their work history, saving original blueprints and recording the tricks they had used to run a difficult operation. These examples suggest that process innovation may be the result of deliberate study by non-R&D workers, beyond simple improvements in skill gained from experience. Moreover, Thomas (1994) points out that conventional thinking in organization theory, which sees manufacturing or other functions outside of R&D as relatively powerless in innovation, leads firms to underuse valuable knowledge in manufacturing that could provide competitive advantage to the R&D unit. Therefore, the conventional view does not recognize the potential for manufacturing to also generate product innovation, not just process improvements (Area (D) in Figure 1). Thus, not only R&D, but also non-R&D work such as production contributes substantially to firm innovation (see Lee, 2015b, for a more detailed discussion of several examples).

Thus, non-R&D innovation needs to be examined more comprehensively, including product as well as process innovation, and examining within-firm differences in innovation from R&D and non-R&D work. To develop this expanded understanding of innovation, we examine the relative effectiveness of R&D and non-R&D learning for innovation and how these vary by knowledge environments. The next section discusses how the nature of knowledge shapes innovation processes differently for R&D and non-R&D work in an organization.

3. Nature of knowledge, learning and innovation in R&D vs. non-R&D work

To explore innovations from non-R&D compared to R&D in an organization, we need to expand our understandings of learning and innovation. As we discussed in the previous section, learning and innovation are often thought of as something that happens outside the ordinary workplace. However, Jensen et al. (2007) argue that firms have two different modes of learning: 1) Science, Technology and Innovation (STI) mode, which relies on science and technical knowledge and 2) Doing, Using and Interacting (DUI) mode, which relies on informal processes of learning, experience-based know-how and user needs. Moreover, Brown and Duguid (1991), based on the practiced-based view of learning, contend that learning is not separate from working and spans across working and innovating.

Thus, the dominant view of learning by R&D, which values abstract knowledge over actual practice, does not recognize learning in working and misses many potential innovations generated from learning in working (Brown and Duguid, 1991; Jensen et al., 2007). In addition, subsequent improvements in technology, which account for the bulk of technological changes and the pace of improvement, result from feedback from earlier experience in working (Rosenberg, 1976). This feedback will be more effective if the production process and products are more visible. The ongoing activity of production, sales, marketing and other work besides R&D will also inspire creativity in the non-R&D workers (Brown and Duguid, 1991; Rosenberg, 1982; Smith, 1776; Solow, 1994). For example, Barge-Gil et al. (2011) show that the activities of design, training, use of advanced machinery, and technology forecasting are important drivers of innovation in firms that do not conduct R&D.

Therefore, “learning by doing is an important part of the process by which new technology gets *created*, modified, and broken in (p. 1047) [emphasis added]” (Nelson, 1981). Although generating inventions is not non-R&D workers’ primary role, while doing their normal job, some of them develop new ideas. Such extra-role behavior is common in organizations and seen as a key to organizational effectiveness (Organ, 1988; Walsh and Tseng, 1998). Hervas-Oliver et al. (2011) and Santamaria et al. (2009) find that such non-R&D activity is important for any firm, but especially in low and medium tech industries that mostly consist of small and medium firms. Focusing on innovation in service firms, Trigo (2013) finds that non-R&D service innovators, in addition to adopting new innovation from outside, generate innovation from hidden innovation activities, which grow from daily learning and elicit creative ideas. Furthermore, Hervas-Oliver et al. (2011) show that the effects on innovation of absorptive capacities separately built from R&D and non-R&D vary by the different external information sources used (e.g., supplier, customer, university). Similarly, Santamaria et al. (2009) show that the effects of different innovation activities (e.g., formal R&D, training, design, use of advanced machinery) are different between high and low-tech industries. Extending this prior work, we argue below that it is not high-tech versus low-tech per se, but rather differences in the knowledge environments that drive the relative effectiveness of R&D and non-R&D work for generating innovations.

Thus, firms have two different modes of learning. The nature of knowledge will affect different intensities of the two modes of learning in firms, enhancing the role played by one relatively more than the other (Jensen et al., 2007; McIver et al., 2013). The effect size of R&D and non-R&D learning on innovation is not constant, but varies by nature of knowledge. In the next sections, we discuss particular knowledge environments and their impact on differences in the effects of R&D and non-R&D learning on innovation.

3.1. General knowledge environments and R&D vs. non-R&D innovation

Asheim and Coenen (2006) contend “the actual knowledge base of various industries strongly shapes the innovation processes of firms”. In particular, we argue the nature of knowledge in various industries should affect the innovation processes of different activities in firms. Prior research characterizes knowledge bases slightly differently, although these representations share similar underlying characteristics. Asheim and Coenen (2006) and Asheim and Hansen (2009) distinguish “analytical” knowledge, which is science-based, formal and codified, from “synthetic” knowledge, which is relatively more engineering-based and path-dependent. Jensen et al. (2007) characterize forms of knowledge into explicit, global knowledge, which enhances the role of STI mode learning, and implicit, local knowledge, which enhances the role of DUI mode learning. Pavitt (1984) and Winter (1984) distinguish between entrepreneurial regime, favorable to science-based innovative activity, and routinized regime, favorable to innovative activity by cumulative learning. Zuboff (1988) contrasts “action centered” skill, which is direct, experiential and local, from “intellective” skill, which is more abstract, mediated and general.

These characterizations by prior work are related to the generality of knowledge. Context-specific knowledge is sticky and hence difficult to apply in different contexts. In contrast, general and abstract knowledge, articulated in universal terms and based on codified scientific and technical information, is less context dependent, more readily applicable in diverse contexts, and potentially moves the locus of problem-solving (Arora and Gambardella, 1994; Asheim and Isaksen, 2002; Kenney and Dossani, 2005; Pavitt, 1984; von Hippel, 1994). Therefore, an environment where general knowledge is more important matches work practices and enhances the role of R&D workers (Arora and Gambardella, 1994; Jensen et al., 2007; Stinchcombe, 1990). This is because they are more likely to apply principles they learn in their higher science and engineering education and test hypotheses. Furthermore this application of science and engineering principles is supplemented by the ability to use sophisticated instruments and devices.

On the other hand, in an environment where sticky knowledge is more important, innovative activity by skilled, non-R&D workers becomes more effective. In this environment, computer simulation and laboratory analyses are less likely to anticipate problems (Asheim and Coenen, 2006; Lüthje et al., 2005; Malerba, 1992). In the same way that Barge-Gil and López (2014) argue that development draws more heavily on synthetic knowledge while research draws relatively more on analytic knowledge, we argue that non-R&D activity will draw less on analytic knowledge than does R&D. Therefore, along a continuum of importance from general to sticky knowledge in an environment, higher importance of general knowledge in an environment will enhance the role of R&D more than that of non-R&D.

We analyze how these knowledge environments create differences in the effects of two different learning modes on innovation in a firm with a given level of investment in R&D and non-R&D. The importance of general knowledge in an environment may increase the productivity of R&D, thereby eliciting more investment in R&D given sales revenue (cf. Cohen, 2010; Cohen and Levinthal, 1989). Furthermore, it will affect the relative effectiveness of learning by R&D and non-R&D for innovation (i.e., changing the innovation rates by R&D and non-R&D), controlling for R&D intensity (Adler and Clark, 1991; Dutton and Thomas, 1984; Joskow and Rozanski, 1979). Understanding this underlying mechanism (further described in the empirical model below), we will show that, given R&D intensity, higher importance of general knowledge in an environment increases the relative effectiveness of R&D (versus non-R&D) learning for innovation. This leads to the following hypothesis:

Hypothesis 1: The higher the importance of general knowledge in an environment, the greater the rate of innovation given R&D work, compared to the rate of innovation given non-R&D work.

3.2. Visible knowledge environments and R&D vs. non-R&D innovation

The effectiveness of learning also changes in the face of more opportunities to apply knowledge earned through work, i.e. the more visible is technology of production. By visibility,

we mean the extent to which those engaged in the production of the product/service can see problems and can see how their own actions affect the outcomes (i.e., there are tighter links between actions and outcomes). Organizations that assume learning as an information-transmission process (i.e. non-R&D workers are told what they need to know) tend to produce and adopt opaque technologies (Brown and Duguid, 1989). This makes non-R&D learning more difficult because the technology or problem is not visible to them (Zuboff, 1988). There is a loss of “cause-and-effect” knowledge (Zuboff, 1988). This reinforces the assumption that non-R&D workers cannot learn on their own, making technology more and more opaque and creating a negative feedback loop (Brown and Duguid, 1989; Zuboff, 1988). Furthermore, more visible problems recurring frequently will provide more opportunities for creative thinking and utilizing learning by doing or in working, and get solved faster. (Kenney and Tanaka, 2003; Winter, 2003). Therefore, visibility contributes to the introduction of more innovations given investment in learning activity in an organization. A production technology with low visibility of problems and high uncertainty in the production process makes the links between activities and their consequences harder to discern, making innovation opportunities less likely to be recognized (Brown and Duguid, 1989).

This visibility of problems is affected by organizational structure (Lundvall and Johnson, 1994; Van de Ven et al., 1976). Organizational structures that make problems more visible are associated with more opportunities for enhancing and utilizing learning by doing, using and interacting (Jensen et al., 2007). For example, the Toyota Production System, which specifies activity for each worker and tightly-linked sequential processes, makes problems more visible and provides more opportunities for learning to be utilized in the organization (Spear and Bowen, 1999).

However, not only organizational structure, but technology or knowledge itself can also have more visible characteristics in certain activities or industries. Argote and Darr (2000) give two contrasting examples in their study of learning by doing in pizza franchises. In the example of the ‘cheese spreader’, the problem of unevenly distributed cheese is obvious to all those who work in the store. However, developing a ‘proofing’ method for pizza dough requires more time and experimentation because of the weak links and low visibility between the dough making process, the nature of the dough and the outcome (i.e. good or bad pizza). In the case of many mechanical industries, the product components and their interactions may be relatively more visible, when compared to, for example, chemicals, materials or electronics products (Seymore, 2009). For example, in the process of designing the new V-Rod motorcycle, Harley-Davidson production engineers could readily spot production problems inherent in early designs simply by looking at a steel and clay prototype (Sichterman, 2001). Owan and Kim (2013) show that in semiconductor chip production, opportunities for problem solving are greater in some product segments than others because of the visibility of problems in those products. For example, most semiconductor products are “black box”, so that problems are not visible when they work poorly. However, the elements of a memory chip are individually addressable, so the specific nature of problems is more easily spotted and solved (Owan and Kim, 2013). They show higher rates of

learning by doing in memory chip production than in other segments, consistent with this visibility argument. Furthermore, the choice of technology and its implementation in an industry can also affect visibility (Noble, 1984, Zuboff, 1988).

Therefore, greater ease of seeing problems will be associated with more opportunities for utilizing learning by non-R&D, increasing the effect of non-R&D work on innovation. The high visibility of problems will also affect learning by R&D. However, the effect of visibility on learning and innovation of non-R&D will be relatively much larger than that of R&D because R&D, whose main job is inventing, will be relatively less sensitive to visibility of problems in production (and also be more removed from direct interaction with the production process). Therefore, the relative difference in the effects of R&D and non-R&D learning on innovation will be smaller in high visibility environments. Accordingly, we have the following hypothesis:

Hypothesis 2: The greater visibility of knowledge in an environment, the greater the rate of innovation given non-R&D work compared to the rate of innovation given R&D work.

4. Data and methods

4.1. Data

The focal data in this study are from a survey of US inventors (hereafter, the Inventor Survey). The Inventor Survey is a survey of US inventors on triadic patents (patents filed in Japan and the EPO and granted by the USPTO) in the application period 2000 to 2003, and collects information on the projects that generated the patent (Walsh et al., 2015). This allows us to code whether this was an R&D or a non-R&D invention. The survey sampled triadic patents stratified by NBER technology classes.¹ The number of patents belonging to each unique inventor was recorded to use as a weight for later survey data estimation. The survey received 1919 responses with a response rate of 24.2% (31.9% adjusted for undelivered, deceased, etc.).² After limiting data to patents assigned to firms (i.e., excluding universities and hospitals, government labs and individual inventions), the sample used in this study includes 1738 triadic patents. We combined the Inventor Survey data with Carnegie Mellon Survey (CMS), NSF, and US Census data. The

¹ To limit respondent burden, we randomly selected one patent out of multiple patents belonging to the same inventor

² To test for non-response bias, we used data from the patent documents to compare respondents to non-respondents. We find little evidence of non-response biases that were either statistically or substantively significant. In particular, measures of collaboration (solo inventions: 27% for respondents, 26% for non-respondents; average number of inventors: 2.71 for respondents, 2.80 for non-respondents), links to universities (citations to non-patent literature: 2.4 for respondents, 2.7 for non-respondents) and measures of patent value (forward citations, 2.2 for respondents, 2.4 for non-respondents) are all similar (none are significantly different, $p < .05$, $N = 7933$). The only significant differences are that inventors for which we only had a company address (instead of home address) are less likely to respond (4% of respondents had a company address v. 6% for non-respondents, $p < .001$) and those with more patents are more likely to respond (mean of 1.18 patents for respondents, 1.13 for non-respondents, $p < .001$), although the absolute differences are quite small. Thus, despite the modest response rate, we have some confidence that our sample is representative of the underlying population of US-based inventors on triadic patents.

CMS is a 1994 survey of R&D managers in US R&D units as part of a manufacturing firm, and includes information about knowledge sources, which allows us to create measures of the knowledge environment (Cohen et al., 2000). The NSF R&D in Industry data in 1999 (NSF, 2002) and US Census provide industry R&D and sales data prior to the Inventor Survey.

To combine those data, we create industry concordances among the three datasets building from the US Census SIC-NAICS concordance between the CMS and NSF/US Census, and the USPTO US patent class-product industry NAICS concordance between the Inventor Survey and NSF/US Census.³ Therefore, environment, or industry, variables for the invention represent the characteristics of the industry that is related to the patented technologies (which may be different from the NAICS classification of the firm). These “projects in firms” data allow us to control for firm-level and technology-level industry characteristics in the same model. All knowledge environment variables are at the 3-digit 2007 NAICS covering US manufacturing industries. The following sections explain the empirical model with its links to our hypotheses, and describe the measures of the variables for tests.

4.2. Empirical model

Our hypotheses relate knowledge environments to the relative invention productivity of R&D and non-R&D (see H1 and H2 above). The Inventor Survey data provide information about whether a given invention is from R&D or non-R&D (i.e. R&D or non-R&D invention). Thus, to operationalize these hypotheses, our empirical strategy is to link the observables (whether a given invention comes from R&D or non-R&D) to our theoretical concepts (the relatively likelihoods of R&D and non-R&D to generate an invention), using our theories of how this latter ratio varies by knowledge environments. To examine the underlying mechanisms of observing the rates of R&D and non-R&D invention, the probability of R&D invention over the probability of non-R&D invention (i.e. odds) can be decomposed using Bayes’ theorem and canceling out the probability of invention from both the numerator and denominator, producing equation (1).

$$\frac{P(RD|invention)}{P(NRD|invention)} = \frac{\frac{P(RD)P(invention|RD)}{P(invention)}}{\frac{P(NRD)P(invention|NRD)}{P(invention)}} = \frac{P(RD)P(invention|RD)}{P(NRD)P(invention|NRD)} \quad (1)$$

Taking the log of equation (1), we have equation (2).

$$\ln\left(\frac{P(RD|invention)}{P(NRD|invention)}\right) = \ln\left(\frac{P(RD)}{P(NRD)}\right) + \ln\left(\frac{P(invention|RD)}{P(invention|NRD)}\right) \quad (2)$$

We then link the 2nd term in the right hand side (RHS) of equation (2) to our characteristics of knowledge environments.

³ Some patent classes have multiple relevant product industries. In this case, we randomly sample one industry.

First, let:

$$\ln(P(\text{invention}|RD)) = \ln\theta_R = \alpha_R GEN + \beta_R VIS + \gamma_R \mathbf{X} + \epsilon_R \quad (3)$$

$$\ln(P(\text{invention}|NRD)) = \ln\theta_N = \alpha_N GEN + \beta_N VIS + \gamma_N \mathbf{X} + \epsilon_N \quad (4)$$

That is, generality and visibility differentially affect the numerator and denominator of the second ratio on the RHS of equation (2), net of controls.

Then:

$$\begin{aligned} \ln\left(\frac{P(\text{invention}|RD)}{P(\text{invention}|NRD)}\right) &= \ln\left(\frac{\theta_R}{\theta_N}\right) = \ln\theta_R - \ln\theta_N \\ &= (\alpha_R - \alpha_N)GEN + (\beta_R - \beta_N)VIS + (\gamma_R - \gamma_N)\mathbf{X} + (\epsilon_R - \epsilon_N) \\ &= \alpha GEN + \beta VIS + \gamma \mathbf{X} + \epsilon \quad (5) \end{aligned}$$

For the 1st term in the RHS of equation (2), in general, the non-R&D portion of a firm's activity is larger than the R&D portion. In other words, in most firms, R&D intensity (i.e., the share of employees or sales dedicated to R&D) is relatively small. In fact, mean R&D intensity in US manufacturing (measured as budget share from domestic sales and R&D performance data in BRDIS 2011) is about 3.9% (NSF, 2014), meaning that the average firm invests significantly more effort in "production" than in "R&D". However, given the relative size of R&D and non-R&D in a firm, the base-line productivity (e.g., inventions per person-year or inventions per unit cost) of R&D in inventing will be higher than that of non-R&D because R&D is an activity focused on invention.

Accordingly, the change in the relative effectiveness of learning for invention between R&D and non-R&D (2nd term in the RHS of the equation (2))⁴, controlling for a firm's R&D size relative to non-R&D (1st term in the RHS of equation (2)), will change the ratio of R&D to non-R&D inventions we observe, given invention outputs (term in the LHS of the equation). Prior studies already analyzed the determinants of the share of R&D in firms or the decision to do R&D (Cohen and Levinthal, 1989; Cuervo-Cazurra and Annique Un, 2010), which is not our primary interest. Controlling for R&D size (relative to non-R&D), i.e., our definition of R&D intensity in this study, we focus on the effects of knowledge environments on invention productivity by R&D and non-R&D.

Equation (5) contains the observables, with GEN being generality, VIS being the visibility and $\mathbf{X}$ being a vector of control variables. This equation highlights our theoretical predictions. The parameter α is the difference of α_R (the effect of generality on the R&D learning rate, θ_R in equation (3)) and α_N (the effect of generality on the non-R&D learning rate, θ_N in equation (4)). Similarly, β is the difference of β_R (the effect of visibility on the R&D

⁴ Relative effectiveness of learning for invention, relative invention productivity, and relative size of learning rates, as we are using them here, all mean the same. Different readers (e.g., organizational learning theorists vs. economists) may be more comfortable with one term than the other.

learning rate) and β_N (the effect of visibility on the non-R&D learning rate). Our theory of general knowledge environments above hypothesizes that $\alpha_R > \alpha_N$, so that, H1: $\alpha > 0$. Similarly, our theory of visible knowledge environments hypothesizes that $\beta_R < \beta_N$, so that, H2: $\beta < 0$. For example, the effect of visibility is negative on the relative effectiveness of learning by R&D for invention to that by non-R&D, and therefore we will observe, controlling for R&D intensity, relatively fewer R&D-based inventions (i.e. a decrease in the ratio of R&D-sourced inventions to non-R&D-sourced inventions).

Thus, we take observables that are hypothesized to change the relative effectiveness of learning for invention between R&D and non-R&D, and show how, controlling R&D intensity, changes in the observables (knowledge environment) changes the intensity of R&D invention over non-R&D invention in a firm. As an alternative specification, for the multi-invention firms in the sample, we also control for firm-specific characteristics such as management practices as well as the relative size of R&D over non-R&D, using firm dummy variables. This will help us see the extent to which nature of knowledge explains the variation in relative effectiveness of different learning (thereby observing variation in rates of different types of invention), controlling for the mediating effect of firm characteristics (Nelson, 1981).

Based on this model, we use a logit specification to test our hypotheses. In addition, the Inventor Survey used a stratified random sampling procedure (with equivalent sampling rates across strata) and drew randomly one patent per each inventor when inventors have multiple patents. Therefore, each inventor has a different weight. Because of the sampling strategy, the parameters in all statistics and regression outputs in this paper are estimated taking the survey design into account (Kalton, 1983; Lee and Forthofer, 2006).

4.3. Measures of R&D vs. non-R&D invention

Innovation in this study is measured at the project level, using triadically patented inventions as a proxy for innovation. Using patents as a measure of innovation may be problematic. However, in this case, these are significant inventions, as these patents are on novel technologies that were filed in three jurisdictions, suggesting they have high importance (Grupp and Schmoch, 1999). Thus, while many (non-patented or US-only patented) firm inventions will not be included, we have some confidence that we are capturing significant inventions. Using these data, we can create a measure of internally-generated R&D v. non-R&D inventions.

Defining R&D and non-R&D invention requires understanding how people work in a firm to produce inventions. An R&D worker's job is to create new knowledge by doing R&D. However, non-R&D workers such as sales, marketing, and production employees sometimes join an R&D project with R&D workers through a cross-functional team. Thus, not only inventions by R&D workers but also inventions from cross-functional R&D projects are defined as R&D inventions. However, if non-R&D workers who do not join an R&D project conceptualize the solution of a problem (i.e., "invent" something, see footnote 5 below), it becomes a non-R&D invention. Therefore, the type of employees and their creative process in producing the invention should be considered together when defining R&D and non-R&D invention at the project level.

To define non-R&D invention, we employ two questions from the Inventor Survey: one about the creative process that led to their invention and one about the type of unit to which they belonged at the time of the invention.

The question about the creative process is as follows:

Which of the following scenarios best describes the creative process that led to your invention?

- a. The **targeted achievement** of a research or development project
- b. An **unexpected by-product** of a research or development project
- c. An **expected by-product** of a research or development project
- d. Directly **related to your normal job** (which is not inventing), and was then further developed in a (research or development) project
- e. From **pure inspiration/creativity** or from your normal job (which is not inventing), and was not further developed in a (research or development) project

The inventors' affiliation is categorized by the following units:

1. An independent R&D unit or its sub-unit
2. R&D sub-unit attached to a unit with its primary focus on non-R&D such as manufacturing
3. Manufacturing
4. Software development
5. Other (e.g. Sales/marketing)

The inventions from (a, b, c) OR (1) are defined as R&D invention. Therefore, if people in non-R&D units produce an invention as a result of joining an R&D project, it is classified into R&D invention (OECD, 2002). In contrast, inventions from (d, e) AND (2, 3, 4, 5) are defined as non-R&D invention (see Seymore, 2009, for a discussion of the requirements for "inventor").⁵ The counts of inventions in each cell are displayed in Table 1.

-----INSERT TABLE 1 ABOUT HERE-----

⁵ Some may argue that if R&D helped in the subsequent development of a non-R&D invention, then it should be categorized as an R&D invention. However, we argue that it is consistent with the USPTO definition of "inventor" to treat these as non-R&D inventions. According to the USPTO Manual of Patent Examining Procedure [Chapters 2137 and 2138], the invention must be clearly conceived of in the inventor's mind (such that she could clearly describe it to another), but she does not have to carry out all the steps in the process of reducing the invention to practice (<http://www.uspto.gov/web/offices/pac/mpep/mpep-2100.html>). While "invention" requires "conception" and "reduction to practice", reduction to practice can be accomplished by the written description of the invention (Seymore, 2009), or by giving direction and guidance for others to carry out (USPTO guidelines). In our case, since we are dealing with patented inventions, with the (non-R&D) inventor certified by the patent office, this requirement has been fulfilled. Therefore, as long as the solution is clearly conceptualized in the course of the normal jobs of non-R&D personnel, although it is further developed in an R&D project later, following USPTO guidelines, this should be defined as non-R&D invention.

In the cases of an R&D subunit attached to a unit with its primary focus on non-R&D (such as manufacturing), if those inventions are from their normal job (which is not inventing) or from pure inspiration/creativity, they were included as non-R&D inventions. Subordinate R&D units (item 2 in the unit question) are likely heavily involved in technical services, compared to independent R&D units (item 1 in the unit question). Therefore, we should not assume that all inventions from workers in such subunits should be classified *a priori* into R&D inventions (if the answer to the question about the invention process suggests these were not part of an R&D project). This is consistent with the Frascati manual and NSF's classification, which exclude "production and related technical activities" from R&D (OECD, 2002). According to the CMS data, when comparing R&D subunits located in production facilities and stand-alone R&D units, among business units with 10+ employees, the percentage of technical service (providing manufacturing support, troubleshooting, etc.) out of total "R&D" effort is substantially higher in subordinate R&D units than in stand-alone R&D units (23% v. 14%, $p < .01$). The Inventor Survey data also show that, on average, inventors from subordinate R&D units spend substantially more effort on technical service than those from independent R&D units (14% v. 7%, $p < .01$). Therefore, reflecting the different characteristics between subordinate R&D units and stand-alone R&D units, if, based on the answer to the invention process question, workers from subordinate R&D units are not involved in a planned, official R&D projects leading to the invention, their inventions can be defined as non-R&D inventions, e.g., from their technical service or pure inspiration. Table 1 shows that, of the 313 inventions from R&D subunits attached to non-R&D units, 82 (26%) are non-R&D inventions, likely due either to technical service or to pure inspiration.

4.4. Variables

The dependent variable is a dummy variable that has 1 for R&D invention and 0 for non-R&D invention using measures described in section 4.3. In this section, we describe our explanatory and control variables created by multiple datasets.

Importance of general knowledge in industries

The generality of knowledge in the firm's environment is measured by several items. First, more codified knowledge is relatively more generalizable (Arora and Gambardella, 1994; Jensen et al., 2007; Kenney and Dossani, 2005). Using the CMS, the codifiability of knowledge is measured by the industry mean of a 4-point scale asking the importance to the R&D unit of publications and reports a) from other firms, or b) from universities or government research labs (the maximum score of two items, one asking about firm sources and the other about universities/government labs): *codifiable knowledge*. Also, if the knowledge is more upstream, it would be more likely to be generalized. The production of basic knowledge in an industry is measured by the industry average percent of basic research produced by firms, from CMS: *basic knowledge*. Finally, university or government lab-driven knowledge can be applied broadly because it is more related to abstract principles. The contribution of university-driven knowledge

is measured by the industry average percent of firms' R&D projects in the last three years that made use of university knowledge, using the CMS: *university-driven knowledge*. Each characteristic measures an aspect of the generality of knowledge and is not sufficient alone for a proxy for general knowledge. For example, codifying knowledge alone does not necessarily make knowledge more general and accessible to others because using secret codes can undermine transferability (Asheim, 2002; Jensen et al., 2007). Therefore, we create a generality index using the sum of the standardized values of these three measures. This index represents industry knowledge characteristics built prior to the time when the firms in the Inventory Survey generate inventions and allows testing the effect of industry knowledge base on the innovation process in firms. The higher value in the index means the more relevance or importance of general knowledge in an industry.

Visibility of knowledge in industries

Our second measure of knowledge environment is visibility. Visibility is measured with a dummy variable where industries with high visibility of knowledge have 1 and those with low visibility of knowledge have 0. This industry classification, although a coarse measure of visibility, is created from two criteria.

First, Stinchcombe (1965) shows that past organizational forms determine the present structure of organizations. For example, the present construction industry still has aspects of a pre-industrial craft form of organizations, distinguished from more "modern" industries such as chemicals. He classifies industries according to vintage of the establishment of that industry. We argue that, following his classification, "prefactory" industries (e.g. printing, ship building, construction, etc.) and "early nineteenth century" industries (e.g. woodworking, glass, leather, apparel, textile, etc.) would likely still have less complex organizational structures. Those industries perhaps have more directly visible production processes, than "railroad age" and "modern" industries (e.g., metal, petroleum, chemicals, rubber, electrical equipment, transportation equipment etc.), because the past craft form of organizations is more adapted to problems in that industry (Stinchcombe, 1965).

Second, as suggested by Seymore (2009), mechanical industries should have higher knowledge visibility, compared to chemical or electronic industries. The contrast between high problem visibility in designing transportation equipment such as motorcycles (Sichterman, 2001) and the low-visibility "black-box" semiconductors (except for memory chips) (Owan and Kim, 2013) is an example. To determine what industries are more mechanical, we used the CMS and classified industries into two groups: one that has 3-digit NAICS industry means that are above the overall average for the importance of mechanical engineering as a knowledge source, and the other whose means are below average.

Using these two criteria (prefactory/early industrial vs. railroad age/modern; and mechanical knowledge-based vs. other), prefactory/early industries or mechanical knowledge-based industries are grouped into industries with high visibility, and the rest into industries with low visibility. Accordingly, food, textile, wood product, fabricated metal, machinery, computer

and electronic product (except semiconductor), transportation equipment, furniture and miscellaneous manufacturing (including medical equipment and supplies) are coded as industries with high visibility (i.e., NAICS 311-316, 321-323, 332-334 (except 3344), 336-339); while chemical, pharmaceutical, petroleum and coal, plastics and rubber, non-metallic mineral, metal, semiconductor, and electrical equipment industries are defined as industries with low visibility (i.e., NAICS 324-327, 331, 3344, 335).⁶

Inventor patent propensity

Since all inventions from the Inventor Survey are patented inventions, the effect of different patent propensity by inventors needs to be considered. Higher patent propensity by inventors working in certain units could affect the ratio of R&D v. non-R&D patented inventions observed in the firm. For example, non-R&D inventions might have to meet a higher threshold in order to get patented compared to R&D inventions, which affects the comparison between these two types of inventions. Moreover, R&D and non-R&D personnel may receive different rewards for invention, which affects disclosure and patenting of their inventions. Therefore, we control for inventor patent propensity, measured by the percent of inventions disclosed by the respondent to her firm that resulted in a patent application, from the Inventor Survey.

Industry annual growth

Rapidly growing industries can create higher payoff for innovation. In the literature, this is seen as increasing the incentives for conducting R&D (Cohen and Walsh, 2000), and hence might elicit relatively more R&D inventions than non-R&D inventions. More generally, this is a control for the demand for innovation (from both R&D and non-R&D, although perhaps with

⁶ Some may suggest that this high vs. low visibility classification equals low vs. high-tech industries. High-tech/low-tech is based on industry R&D intensity (OECD, 2011). However, we use prior studies of organization structure and technology characteristics and classify high/low visibility deductively. Therefore, our high/low visibility is theoretically distinct from a low-tech/high-tech classification. They are also empirically distinct. In our classification, the low visibility industries include several low-tech industries. For example, based on BRDIS 2011 data, the R&D intensity of the plastics and rubber industry is 1.4%; that of the metals industry 0.4%; and that of the chemicals (except pharmaceuticals) industry 1.2%, much lower than the manufacturing mean R&D intensity, 3.9%. Similarly, non-metallic minerals and the metals industries are classified into a low-tech industry by OECD (2011) and Hervas-Oliver et al. (2011), but are included in low visibility in our classification. The high visibility industries also include high-tech industries. For example, the R&D intensity of the computer and electronic product (except semiconductor) industry is 9.9%, greater than the manufacturing mean R&D intensity, and the R&D intensity of the machinery industry is 3.8%, close to the mean. Moreover, machinery and medical equipment are classified into high-tech industry by OECD (2011) and Hervas-Oliver et al. (2011), but are included in the high visibility industries by our classification. Overall, the correlation between industry R&D intensity and our high v. low visibility measure is .03 (Table 2). Therefore, our measure is not equivalent to a low/high-tech classification.

We, however, acknowledge there is some measurement error in this measure. For example, as noted above, some segments of semi-conductors (memory chips) have relatively high visibility (Owan and Kim, 2013). Similarly, much contemporary innovation in the food industry is based on chemistry, often with low visibility (Warner, 2014).

differential effects on one or the other) and hence might draw more invention productivity from one or the other of these (although it is not clear which). Therefore, we control for the patent's industry average real annual sales growth for 5 years between 1997 and 2002, using US Census data. Furthermore, this variable, along with firm size and industry R&D intensity, is used to control for firm-level R&D intensity (in equation 2).

Firm size

We control for firm size, which can be associated with cumulative experience or incentive to introduce innovation (Cohen and Klepper, 1996a, b; Pavitt, 1984). Firm size is measured by log of midpoints in categories of employees: less than 100 employees, 100-250 employees, 250-500 employees, and 501 or more employees (using 750 for the midpoint of the top category), from the Inventory Survey. Furthermore, this also helps control for firm R&D intensity (see below).

R&D intensity/Firm dummies

As discussed in the empirical model section, controlling for a firm's R&D intensity (i.e., relative size of R&D to non-R&D), we can examine the underlying relationship between knowledge environment and relative invention productivity by R&D and non-R&D, which leads to the change in the ratio of R&D to non-R&D inventions we observe, given invention outputs. To control R&D intensity in each firm, we use product industry annual growth, firm size and (disaggregate) industry R&D intensity together. We could not control each firm's R&D intensity due to limitations on firm-level data. Therefore, following equation (2) we create a variable from the log of the ratio of R&D to non-R&D spending for assignee firms' NAICS,⁷ if possible, at a more disaggregate-level NAICS, using R&D funds (except federal funds) and domestic net sales data from the NSF R&D in Industry data in 1999 (NSF, 2002), temporally prior to the Inventor Survey (with inventions from the period of 2000 to 2003). Then, we run models jointly using this measure, product industry annual growth and firm size as a control for firm-level R&D intensity, which together should capture much of the impact of firm R&D intensity.⁸

As an alternative specification, we repeat these tests controlling for assignee firm dummies, to rule out firm heterogeneity more strictly. In particular, this model controls for differences in R&D intensity and in firm-level patent propensity, compensating for some of the shortcomings of the first model. However, the model with firm dummies also controls for differences in the relative invention productivity of R&D and non-R&D units due to organizational practices and structures. Furthermore, these firm dummies are likely to capture

⁷ Firms with multiple businesses were assigned to industry based on the modal industry of the patents assigned to that firm (choosing at random in the case of ties).

⁸ We regressed our dependent variable on the (disaggregate) industry-level R&D intensity, and found that it has a positive but not significant effect ($p=.33$), likely due to measurement error. When we regress our dependent variable on all three variables: industry R&D intensity, industry annual growth and firm size, as a joint proxy for firm-level R&D intensity, the joint significant test (Wald test) of the three variables is significant ($p < .10$). Hence, we use these three variables jointly as a proxy for firm R&D intensity, which is one of the components in the RHS of the equation (2).

some of the industry-level knowledge environment effects. Hence, this model is likely to bias downward (toward zero) the effects of knowledge environment because the firm dummy controls will mediate some of the impact of knowledge environment.⁹ Thus, while we suspect that the first model (because of the weak controls for R&D intensity) may overestimate the effects of knowledge environment, the second model will underestimate the effects. Hence, the true effect is likely to be bounded by these two models. For this analysis, we create independent dummy variables for all assignees that have three or more patents in the Inventor Survey data. If we limit the sample to these cases, we have 974 patents representing 121 firms.¹⁰

The descriptive statistics and correlation matrix of the variables are in Table 2. From this table we can see that 88% of our inventions are R&D inventions (and hence 12% are non-R&D inventions). We discuss this in detail in the next section. In addition, prior work suggests that general knowledge may be more important in high-tech than in low-tech industries (cf. Cohen and Levinthal, 1989; Jensen et al., 2007). Table 2 shows the correlation between industry R&D intensity and the general knowledge environment measure. We find they are correlated ($r = .31$). This correlation is consistent with the interpretation that high-tech industries have higher use of general knowledge, but also with our claim that there can be an additional effect, controlling for the R&D intensity-general knowledge environment relation.

----- INSERT TABLE 2 ABOUT HERE-----

5. Results

5.1. Comparing R&D and non-R&D inventions

Based on measures of R&D vs. non-R&D inventions created in section 4.3, Table 3 shows that out of all triadically patented inventions, about 12% of **inventions** (as distinct from innovating firms) are non-R&D inventions. According to the BRDIS 2011, out of all patented inventions, 6% of inventions are non-R&D inventions (NSF, 2014). However, since BRDIS categorizes all inventions from R&D performers as R&D invention, it would underestimate the percentage of non-R&D inventions because it does not distinguish non-R&D inventions from R&D inventions in the R&D performing firms. Moreover, Lee (2015a) shows that using innovation data in US manufacturing industries based on a survey by Arora et al. (2014), out of all internally generated

⁹ Note that the industry of the firm is based on firms' sales, while the industry of the invention is based on the technology class of the invention (and the technology class-industry concordance). Hence, controlling for firms does not completely account for our patent industry knowledge environment variables. To see the degree of collinearity, we regressed our knowledge environment variables on the set of firm dummies used in our models, taking survey design into account. We find that the R-squared in the OLS model for generality is .29, and for visibility is .20. Hence, the firm dummies likely capture some of the effects of knowledge environment (more so for generality), but still leaving some independent variance for the direct effect of knowledge environment on the ratio of R&D to non-R&D invention.

¹⁰ These patents account for 56% of our sample of patents and these firms represent 15% of the total number of assignee firms in this study sample. However, in the regressions, some observations are omitted, mostly because many firms have only one kind of inventions (i.e. no variation, with perfect prediction) and partially due to missing values on other variables.

new-to-market **innovations** by US manufacturing firms, 11% are non-R&D **innovations**, which is comparable to 12% non-R&D inventions in this study. Therefore, our estimate of 12% non-R&D inventions is broadly consistent with these statistics, showing the advantage of using a finer-grained project-based measure of R&D v. non-R&D invention. Of these non-R&D inventions in our data, about a third comes from manufacturing units, about 20% comes from sales, service or other units, and the rest comes from R&D subunits attached to manufacturing, etc. (i.e., technical service), or software development units. These rates of non-R&D invention vary across industries, implying industry-associated characteristics can drive relative differences in the rates of R&D and non-R&D innovation. We can see higher rates of non-R&D inventions in textiles, apparel and leather; wood and paper; fabricated metals; and machinery industries. In contrast, the rates of non-R&D invention are low in food; chemicals; and pharmaceuticals.

-----INSERT TABLE 3 ABOUT HERE-----

-----INSERT TABLE 4 ABOUT HERE-----

The descriptive analyses of non-R&D and R&D invention in Table 4 highlight their similarities and differences. First, R&D workers are usually those who are highly educated, often with a PhD degree, and their role is to search knowledge and develop something from it; while non-R&D workers gain more practical knowledge, anchored in their normal task. In Table 4, we can see that it takes significantly less time for R&D inventors to apply for their first patent compared to non-R&D inventors (mean age 34 vs. 37). The longer tenure for the first patent for non-R&D inventors suggests that non-R&D workers depend on their accumulated work knowledge more than R&D workers (Kenney and Tanaka, 2003). Furthermore, age at highest degree for R&D inventors is slightly higher than for non-R&D inventors (28 vs. 27), presumably because of R&D inventors' advanced degrees. The rate of R&D inventors with PhD as their highest degree is twice as high as that of non-R&D inventors (48% vs. 24%). The rate of inventors with science and engineering as their highest degree is also significantly higher for R&D inventors than for non-R&D inventors, although the majority have S&E degrees in both groups. According to Science and Engineering Indicators 2012 (NSB, 2012), in 2008, 75% of scientists and engineers are employed in non-S&E occupation or S&E related occupation, not pure S&E occupation, consistent with the high rate of S&E degree holders even among non-R&D inventors. Those non-R&D personnel who have an S&E degree (not necessarily their terminal degree) can combine technical knowledge learned in their education with their experiences in production, sales, and other non-R&D work, resulting in innovations growing out of their non-R&D work. This suggests that STEM training, combined with local knowledge, may be an important source of innovative capacity even in the non-R&D units of firms.

Second, there is a significant difference in the number of information sources used (e.g., patent literature, conferences, universities, suppliers, customers, competitors, etc.). Non-R&D inventors use fewer sources than R&D inventors, which is consistent with the claim that non-R&D workers utilize more task-specific knowledge while R&D workers deal with a wider range of solutions or search broadly (Gavetti et al., 2012; Haunschild and Sullivan, 2002; Lüthje et al., 2005; Malerba, 1992; March and Simon, 1958).

We also see that, in our sample, about 80% of triadically patented inventions are product inventions and there is no significant difference between R&D and non-R&D inventions in terms of which one is more associated with product invention ($p = .90$). Therefore, our data mostly represent product invention, allowing a stricter test of our theory, because while much of the literature about non-R&D innovation is about process improvement, we can show that these theories can also explain non-R&D product innovations.

Table 4 also presents comparisons of value between the two types of inventions: claims, forward citations, and commercialization of the invention (Lerner, 1994; Trajtenberg, 1990). The counts of claims and forward citation between R&D and non-R&D inventions are not significantly different in our sample, while the rate of any commercialization is higher in non-R&D invention than R&D invention. Given that triadically patented inventions have passed a high threshold, and hence are likely to be highly valuable inventions, we might not expect major differences between the two groups. If we consider the total population of R&D and non-R&D invention including both non-patented and patented inventions, R&D inventions might be relatively more valuable than non-R&D inventions. Yet, what Table 4 shows is that there are a significant number of non-R&D inventions that are at least as valuable as firms' R&D inventions (cf. Arundel et al., 2008), which can be easily missed by the dominant perspective of innovation research centering on R&D. This table suggests that non-R&D inventions are important beyond the stereotypical image of process improvements or marginal shop-floor inventions.

5.2. *Nature of knowledge and R&D vs. non-R&D learning for invention*

R&D and non-R&D learning generate innovation. However, depending on the characteristics of the knowledge environment in which they operate, the *relative* effectiveness of learning by R&D and non-R&D can vary, which helps us better understand non-R&D innovation. First, we test the effect of different knowledge environments on the relative probability of R&D invention to non-R&D invention (based on equation (3) above). In Table 5, Model 1 shows that importance of general knowledge in an environment increases the probability of R&D invention over non-R&D invention, while greater visibility of knowledge increases that of non-R&D invention over R&D invention. We also test each indicator of our generality index separately in Models 2 to 4, and the results are robust, with each indicator showing a positive significant effect by itself. If we control for firm R&D intensity, the predicted ratio of R&D inventions to non-R&D inventions implies the relative effectiveness of R&D and non-R&D learning (based on equation (2) above). Using joint controls for the patent's industry annual growth, firm size and disaggregate industry R&D intensity to proxy for firm R&D intensity, although allowing for some measurement error, Model 5 shows the positive effect of importance of general knowledge.¹¹ The result implies that a more general (i.e., less sticky) knowledge environment increases the productivity of R&D over non-R&D in inventing, thereby

¹¹ Even after adding the knowledge variables in Model 5 of Table 5, the proxies for firm R&D intensity (industry R&D, firm size and industry growth) are jointly statistically significant ($p < .10$).

resulting in relatively more R&D inventions being observed than non-R&D inventions, supporting Hypothesis 1.

Moreover, the higher visibility of problems affects building and exploiting learning by non-R&D through more problem-solving opportunities in the production process. Therefore, it results in a relative increase in the invention productivity by non-R&D compared to R&D, given firm R&D intensity, leading to a relative increase in the probability of non-R&D invention being observed rather than R&D invention, supporting Hypothesis 2.

Based on Model 5 in Table 5, through the underlying mechanism of different learning rates by different work practices, Table 6 shows the marginal effects of nature of knowledge. For example, a change in generality from minimum to maximum increases the probability of R&D invention by 11.8 percentage points while a change in visibility from low to high visibility decreases the probability of R&D invention (i.e., increases the probability of non-R&D invention) by 6.2 percentage points, holding all others at their mean. The relative change in the odds due to change from low to high visibility is 51.5%.

Finally, we control for firm-level R&D intensity and other firm-level characteristics, as well as some aspects of firm knowledge environment (see footnote 9), using firm dummies. First, we test the same model of Model 5 in Table 5 for those selected cases of patents from assignee firms with at least three patents to see if there is any selection problem between the full set of cases and those selected cases we will use for the models with firm dummies. Consistent results between Model 5 and Model 6 in Table 5 show that there is no serious selection problem. After controlling firm heterogeneity (Model 7), comparing Model 6 to Model 7,¹² the direct effect of general knowledge becomes not significant (though still positive). This implies that the effect of general knowledge on invention productivity is mediated by firm-level characteristics. However, the visibility of knowledge in the invention's industry still has a strong direct effect on relative invention productivity of non-R&D compared to R&D, even after considering mediation by firm-level characteristics. This result implies that higher visibility of knowledge is important to increasing effectiveness of learning by non-R&D in inventing, even controlling for individual organizational practices and structures or firm R&D intensity or patent propensity.

Thus, using, admittedly weak, controls for R&D intensity, we find that knowledge generality and visibility affect the relative effectiveness of learning for invention between R&D and non-R&D. Using, possibly overly strong, controls for firm characteristics, including R&D intensity, we find that firm controls mediate the generality effect, but that the effect of visibility remains. We suspect that the effect of generality is between these two extremes. For controls, the firm size variable by itself is not significant although its direction is positive. Industry annual growth and inventor patent propensity are positive, but not consistently significant across models.¹³

¹² After firms that have only one kind of inventions and their patents are excluded (see footnote 10), matching observations in Model 6 and those in Model 7 leaves 444 patents representing 41 firms.

¹³ There may be some concern that the non-R&D inventions may be trivial inventions while the R&D inventions are major inventions, making the comparison inappropriate. We note that the R&D and non-

-----INSERT TABLE 5 ABOUT HERE-----
-----INSERT TABLE 6 ABOUT HERE-----

5.3. Robustness tests

5.3.1. Alternative classifications of R&D and non-R&D invention

We estimated the models with alternative classifications of R&D and non-R&D invention from our original definition, for example by redefining cases with possible ambiguity in our original classification (e.g. R&D subunits, or inventions from one's normal non-R&D job but further developed into an R&D project) from non-R&D to R&D invention or excluding those ambiguous cases from the analysis. We display the results of models equivalent to Model 5 in Table 5.¹⁴ First, when we redefine all inventions from normal job further developed in an R&D project regardless of work units as R&D invention (Model 1 in Table 7) or excluding all those from the analysis (Model 2), the results are consistent with Model 5 in Table 5.

Next, when we redefine all inventions from R&D subunits attached to a non-R&D unit as R&D invention (Model 3 in Table 7), although some of them are likely to be non-R&D technical service, or exclude all those inventions from the analysis (Model 4), the results are qualitatively consistent, but losing some significance in generality of knowledge. To further explore the sensitivity of this category and to decompose the results of Models 3 and 4 in Table 7, especially for the reduced effect of generality, we estimate a multinomial probit model with all inventions from R&D subunits attached to a non-R&D unit (i.e., the highly heterogeneous case that likely includes some R&D and some non-R&D inventions) as a base group. Generality affects the difference between the base group and R&D invention (Model 5), but visibility does not have a significant effect. On the other hand, visibility affects the difference between the base group and non-R&D invention (Model 6), but generality does not have a significant effect. These results tell us that invention from R&D subunits attached to a non-R&D unit has some similarities and differences with R&D invention and non-R&D invention, and suggest that our original definition (including a part of this middle category in R&D invention and a part of it in non-R&D invention, based on whether the inventor joins an R&D project or not) is a reasonable strategy for dealing with this ambiguity.

Thus, comparing Table 5 and Table 7, our results are largely robust to alternative boundaries between R&D and non-R&D invention, although significance levels are sensitive to

R&D inventions in our sample are all triadically patented inventions, which suggests that they have to some degree been filtered based on value. Moreover, in our sample, R&D and non-R&D inventions are not significantly different in forward citations, number of claims, or the probability of being a product (versus process) invention ($p > .10$), which undermines the conjecture that the different characteristics of the types of inventions may be driven by non-R&D inventions being trivial and R&D inventions being important. This similarity in the selection process (leading to triadic patents) produces a more matched sample for comparison (than if we had data on all inventions from R&D and non-R&D). Hence, if we see significant differences in how each type of invention responds to different knowledge environments, this would be a stronger test of the sensitivity of R&D and non-R&D productivity to differences in the knowledge environment.

¹⁴ For more elaborate robustness checks, see Lee (2015b).

the specific operationalization of R&D v. non-R&D. We argue that our main operationalization (Table 5) is the most consistent with OECD technical service and USPTO inventor definitions. Furthermore, controlling for firm heterogeneity (which captures R&D intensity, firm characteristics as well as some aspects of knowledge environment) mediates the effect of generality, though the effect of visibility is robust even to this control.

-----INSERT TABLE 7 ABOUT HERE-----

5.3.2. Other robustness tests

Additionally, we tested our models limiting the sample to product inventions only and obtained consistent results, which reemphasizes that non-R&D work also generates product inventions and our arguments on knowledge environments explain the different rates of product inventions by R&D and non-R&D work. We also tested the effect of a continuous measure of visibility only using the industry scores of use of mechanical engineering knowledge, although having the shortcoming of capturing only one aspect of visibility, and obtained consistently significant negative effects for this measure of visibility, as in Table 5. Moreover, we aggregated patents to the parent firm-level (not the assignee-firm level) and tested the equivalent model to Model 7 in Table 5 using the parent firm-level dummies. We had consistent results, without any significant difference between using assignee-level firms and parent-level firms for our analysis. Lastly, out of 1738 patents used in this study, 29% (N=503) have more than one industry assigned in the patent class-product industry concordance. In this case, we randomly sampled one industry. However, to rule out potential bias by measurement error from these patents, we also tested the predictions using only patents having one industry in the concordance (although we lose many cases). The results are consistently significant, in particular, with generality showing even stronger effects than those in Table 5.

6. Conclusion and implications

6.1. Conclusion

R&D-centered innovation studies in the economics of innovation have undervalued or neglected innovation outside of R&D. Recognizing this gap, earlier studies of non-R&D innovation show the difference between R&D and non-R&D performing firms in innovating (e.g., Arundel et al., 2008) and the relative importance of activity outside of formal R&D (compared to R&D activity) in firms' being innovative, comparing high and low-tech industries, or between large and small firms (e.g., Barge-Gil et al., 2011; Hervas-Oliver et al., 2012; Hervas-Oliver et al., 2011; Santamaria et al., 2009). Although this prior work addresses important questions in this underexplored area of non-R&D innovation, most studies have been conducted at the firm level. These firm-level studies often neglect that R&D performing firms generate non-R&D innovation as well as R&D innovation and how these different types of innovation are generated in the same firm. Furthermore, much of the prior work on "non-R&D innovation" focuses on activities intending to produce innovation, such as design, training or marketing. However, in addition to such innovation-focused non-R&D activity, we also argue for the

importance of innovations that grow out of the experience and problem solving of non-R&D workers in their routine work activities. In this study, we tried to answer this question with an invention-level analysis, adopting insights from the sociology of work and organizational learning literatures. Bringing the work and organizations literatures and economics of innovation literatures together allows us to observe a gap in the current innovation research, as discussed in Section 2, and further helps us to analyze workers' different learning paths to innovate in an organization.

Using the novel, invention-level data and a finer-grained measure of R&D v. non-R&D invention, we first show variation in invention-level rates of non-R&D innovation across industries, and further examine these rates by location within the firm (in Table 3). Many of these inventions come from manufacturing, sales or similar functions in the firm. These data also emphasize that the rate of non-R&D *innovators* (i.e., firms with no R&D but having any innovation) and that of non-R&D *innovations* are different statistics and need to be carefully used depending on the purpose of the study (Lee, 2015a). The former is commonly used in firm-level studies, while our study focuses on the latter. Moreover, we compare the characteristics of the two different invention types, and the two different groups of inventors (in Table 4), to better characterize the profile of non-R&D inventions compared to R&D inventions (among the set of triadically patented inventions). These descriptive data show that non-R&D invention is not rare (about 12% of all triadically patented inventions). Furthermore, we show that non-R&D inventions are comparable to R&D inventions in terms of product (versus process) invention and in terms of patent value (citations, claims, commercialization rate). However, the inventors differ significantly (in addition to location in the firm) in terms of age and education (with non-R&D inventors generally older, and R&D inventors more likely to have PhDs and somewhat more likely to have STEM degrees).

With these comparable sets of R&D and non-R&D inventions as our data, we analyze differences in the effectiveness of learning for innovation by R&D and non-R&D. We develop an argument on the relation between learning by different work (i.e., R&D and non-R&D) in an organization and nature of knowledge (generality, visibility), based on existing theories of knowledge and learning. We show that the relative invention productivity of non-R&D to R&D is higher in high visibility of knowledge than low visibility environments, while the relative invention productivity of R&D to non-R&D increases as generality of knowledge goes up. Our study overcomes some limitations in the existing organizational learning research, which usually deals with only R&D work or only non-R&D work (e.g., production work) in a given study. Moreover, the results contribute to adding evidence to the argument of knowledge characteristics and learning effectiveness, in particular the visibility argument, which has received limited empirical testing. Finally, our results move beyond estimates of the learning rates (the main focus of much of the learning literature) to examine factors that cause variation in learning rates, by showing how relative rates increase or decrease with variations in knowledge environments.

6.2. Limitations and future research

This study suggests several follow-on research questions that grow from the limitations of the current project. First, future research needs to develop more sophisticated measures of visibility and other characteristics of knowledge environments. Although our measure of visibility is a multi-indicator measure, it is a crude dummy variable. The more elaborate measure will contribute to developing the visibility arguments that have received limited empirical testing. We also need more accurate R&D spending data to reduce measurement error. In addition, direct measures of the relative invention productivity of R&D and non-R&D workers would add to the results in this model. Furthermore, observational studies of non-R&D workers (in the workplace ethnography tradition) focusing on non-R&D worker innovation could help elaborate our understandings of how knowledge environments affect non-R&D innovation.

In addition, this study mainly focuses on the effect of nature of knowledge, net of firm-level capabilities to change the relative effectiveness of R&D and non-R&D learning for innovation. Future research needs to explore firm strategy and management tools to better understand how firm capabilities affect rates of non-R&D innovation, along with how these capabilities interact with the effects of knowledge environments.

Lastly, there is a need to analyze the relative values of non-R&D innovation (compared to R&D innovation) or rates of participation in the market for technology, beyond whether non-R&D activity can generate innovation or not (Arora et al., 2013; Lee, 2015b). Our results suggest there may be underexplored high-value non-R&D innovation, unlike common perception that non-R&D innovation is mostly trivial compared to R&D innovation.

6.3. Organizational, managerial and policy implications

Our results motivate work in how to build a learning organization by providing training or reorganizing structure to cultivate non-R&D workers' creativity, as well as that of R&D workers (Kenney and Tanaka, 2003; Vallas, 2003). For example, one benefit of such programs as lean manufacturing and similar re-engineering programs is that, not only do they improve quality, but, by making the production process and problems more visible, these organizational innovations should also generate higher rates of non-R&D invention (based on Table 5). This perspective also suggests that firms may want to encourage stronger intellectual links between the R&D and non-R&D parts of the firm (Thomas, 1994). Constructing organizational structures that help problems become more visible and provide more learning opportunities; organizational policies that encourage employees to disclose their inventions regardless of their work role; and investing in R&D drawing on an understanding of the knowledge environment to increase net value, can all be means to manage innovation more effectively by building learning both from R&D and non-R&D work in the firm.

In addition to organizational restructuring, choices in technology design can make the production process more visible or more opaque (Brown and Duguid, 1991). For example, Noble (1984) documents the choices in the design of numerical control machine tools that made programming the tools either more or less visible to production workers. Similarly, Zuboff (1988) argues that using automation technology to “informatize” work processes in a paper mill

changes the knowledge environment from one characterized as local, experiential and highly visible to one that was more abstract and general and with low-visibility, thereby limiting the ability of shop-floor production workers to apply their knowledge to production problems. Therefore, choices in the development of technology may also affect the learning rates of non-R&D.

Furthermore, prior work has found that firms' investment in activities outside of formal R&D such as training can increase firm innovation (e.g., Barge-Gil et al., 2011; Santamaria et al., 2009). Our results suggest that this result is likely due to the improvement in the learning rate of the non-R&D workforce, and that, not only will there be more innovators, but that there will be a relatively higher share of non-R&D innovations in such firms. Thus, training programs will not only improve worker productivity, but may also increase rates of non-R&D innovation.

This study also has important implications in managing innovation in an organization and developing future innovation policy. First, innovations from non-R&D work can be an alternative or complementary strategy to R&D innovations and may have a significant economic value that has been underappreciated. This also means that policies that encourage or allow outsourcing of manufacturing are also facilitating outsourcing of innovation, and, furthermore, that this effect is greatest for high visibility knowledge environments (such as machinery industries). Disentangling different mechanisms for developing learning by R&D and non-R&D indicates the importance of non-R&D units as another source for innovation, and guides firm strategy for nurturing the creative potential in the non-R&D as well as the R&D workforce and for developing different types of innovations.

Our results highlight the importance of non-R&D innovation for firm and national innovation strategy and suggest the need for developing innovation policies that focus on the non-R&D segment of the economy in order to better tap the large potential for non-R&D innovation. For example, while there may be a general policy goal of encouraging innovation, the policy apparatus of R&D tax credits specifically subsidizes one part of the firm's innovative activity over other parts, although this was not necessarily the intent of the policy. Similarly, policies that either encourage or discourage outsourcing of manufacturing facilities can affect rates of innovation, as non-R&D production activity is an important locus of innovation. Developing policies that promote innovation more broadly requires recognizing the variation in innovative activity in a firm, and further how that variation is sensitive to differences in the knowledge environments firms face. The importance of non-R&D work as a resource for innovation varies by industry due to variations in knowledge environments. Hence, understanding the environmental and organizational drivers of non-R&D innovation can help provide a more developed innovation policy and firm strategy.

Lastly, patent policy struggles with the definition of "inventor" (see Seymore, 2009). Our work on non-R&D inventions highlights the need for a patent policy that recognizes inventors that develop their inventions outside of formal R&D projects. More generally, this work will ultimately help build a more integrated and nuanced model of the innovation process. Our results show that the loci of innovation in the firm are more diffused than the standard R&D-based

model suggests. This suggests the need for focus on, training in and recognition of innovation in the rest of the organization, both for managers and for policy-makers trying to track and encourage innovation.

References

- Adler, P.S., Clark, K.B., 1991. Behind the learning curve. *Management Science* 37, 267-281.
- Aoki, R., 1991. R&D competition for product innovation. *The American Economic Review* 81, 252-256.
- Argote, L., Darr, E., 2000. Repositories of knowledge in franchise organizations. The nature and dynamics of organizational capabilities, 51-68.
- Arora, A., Cohen, W.M., Walsh, J.P., 2014. The Acquisition and Commercialization of Invention in American Manufacturing. National Bureau of Economic Research.
- Arora, A., Fosfuri, A., Rønde, T., 2013. Managing licensing in a market for technology. *Management Science* 59, 1092-1106.
- Arora, A., Gambardella, A., 1994. The changing technology of technological change. *Research Policy* 23, 523-532.
- Arundel, A., Bordoy, C., Kanerva, M., 2008. Neglected innovators, INNO-Metrics Thematic Paper.
- Asheim, B.T., 2002. Temporary organisations and spatial embeddedness of learning and knowledge creation. *Geografiska Annaler: Series B, Human Geography* 84, 111-124.
- Asheim, B.T., Coenen, L., 2006. Contextualising regional innovation systems in a globalising learning economy. *The Journal of Technology Transfer* 31, 163-173.
- Asheim, B.T., Hansen, H.K., 2009. Knowledge bases, talents, and contexts. *Economic Geography* 85, 425-442.
- Asheim, B.T., Isaksen, A., 2002. Regional innovation systems. *The Journal of Technology Transfer* 27, 77-86.
- Barge-Gil, A., Jesus Nieto, M., Santamaría, L., 2011. Hidden innovators. *Technology analysis & Strategic management* 23, 415-432.
- Barge-Gil, A., López, A., 2014. R&D Determinants. *Research Policy* 43, 1634-1648.
- Benkard, C.L., 2000. Learning and forgetting. *American Economic Review* 90, 1034-1054.
- Brown, J.S., Duguid, P., 1989. Innovation in the workplace, CMU Conference on Organizational Learning, Pittsburgh, PA.
- Brown, J.S., Duguid, P., 1991. Organizational learning and communities-of-practice. *Organization science* 2, 40-57.
- Cockburn, I.M., Griliches, Z., 1987. Industry effects and appropriability measures in the stock markets valuation of R&D and patents. National Bureau of Economic Research Cambridge, Mass., USA.
- Cohen, W.M., 2010. Fifty years of empirical studies of innovative activity and performance. *Handbook of the Economics of Innovation* 1, 129-213.
- Cohen, W.M., Klepper, S., 1996a. Firm size and the nature of innovation within industries. *The Review of Economics and Statistics* 78, 232-243.
- Cohen, W.M., Klepper, S., 1996b. A reprise of size and R & D. *The Economic Journal* 106, 925-951.
- Cohen, W.M., Levinthal, D.A., 1989. Innovation and learning. *The Economic Journal* 99, 569-596.
- Cohen, W.M., Nelson, R.R., Walsh, J.P., 2000. Protecting their intellectual assets. National Bureau of Economic Research.
- Cohen, W.M., Walsh, J.P., 2000. R&D spillovers, appropriability and R&D intensity. *Economic Evaluation of Technological Change*, 22-29.
- Cuervo-Cazurra, A., Annique Un, C., 2010. Why some firms never invest in formal R&D. *Strategic Management Journal* 31, 759-779.
- Darr, E.D., Argote, L., Eppler, D., 1995. The acquisition, transfer, and depreciation of knowledge in service organizations. *Management Science* 41, 1750-1762.
- Dutton, J.M., Thomas, A., 1984. Treating progress functions as a managerial opportunity. *Academy of Management review* 9, 235-247.
- Gavetti, G., Greve, H.R., Levinthal, D.A., Ocasio, W., 2012. The behavioral theory of the firm. *The Academy of Management Annals* 6, 1-40.
- Grupp, H., Schmoch, U., 1999. Patent statistics in the age of globalisation. *Research Policy* 28, 377-396.
- Halle, D., 1984. *America's Working Man*. University of Chicago Press.

Hatch, N.W., Mowery, D.C., 1998. Process innovation and learning by doing in semiconductor manufacturing. *Management Science* 44, 1461-1477.

Haunschild, P.R., Sullivan, B.N., 2002. Learning from complexity. *Administrative science quarterly* 47, 609-643.

Heidenreich, M., 2009. Innovation patterns and location of European low-and medium-technology industries. *Research Policy* 38, 483-494.

Hervas-Oliver, J.-L., Albors-Garrigos, J., 2009. The role of the firm's internal and relational capabilities in clusters. *Journal of Economic Geography* 9, 263-283.

Hervas-Oliver, J.-L., Albors-Garrigos, J., Baixauli, J.-J., 2012. Beyond R&D activities. *Economics of Innovation and New Technology* 21, 55-81.

Hervas-Oliver, J.-L., Albors Garrigos, J., Gil-Pechuan, I., 2011. Making sense of innovation by R&D and non-R&D innovators in low technology contexts. *Technovation* 31, 427-446.

Hervas-Oliver, J.-L., Sempere-Ripoll, F., Boronat-Moll, C., Rojas, R., 2015. Technological innovation without R&D. *Technology analysis & Strategic management* 27, 19-38.

Huang, C., Arundel, A., Hollanders, H., 2011. How firms innovate. working paper.

Jensen, M.B., Johnson, B., Lorenz, E., Lundvall, B.Å., 2007. Forms of knowledge and modes of innovation. *Research Policy* 36, 680-693.

Joskow, P.L., Rozanski, G.A., 1979. The effects of learning by doing on nuclear plant operating reliability. *The Review of Economics and Statistics* 61, 161-168.

Kalton, G., 1983. *Introduction to survey sampling*. SAGE Publications, Incorporated.

Kemp, R., Pearson, P., 2007. Final report of the MEI project measuring eco innovation, UM MERIT, Maastricht.

Kenney, M., Dossani, R., 2005. Offshoring and the future of US engineering. *National Academy of Engineering The Bridge* 35(3).

Kenney, M., Tanaka, S., 2003. "Transferring the Learning Factory to America? The Japanese Television Assembly Transplants." In W. Cooke (Ed.) *Multinational Companies and the Transnational Workplace: Strategies for Human Resource Management and Labor Relations* (Westport, CT: Quorum Books): 123-142.

Kusterer, K., 1978. *Know-how on the job*, Boulder, CO: Westview.

Lee, E.S., Forthofer, R.N., 2006. *Analyzing complex survey data*. Sage.

Lee, Y.-N., 2015a. Evaluating and extending innovation indicators for innovation policy. *Research Evaluation*. Published online July 30, 2015, doi:10.1093/reseval/rvv017

Lee, Y.-N., 2015b. Expanding understanding of the innovation process, PhD dissertation. Georgia Institute of Technology, Atlanta, GA.

Lerner, J., 1994. The Importance of Patent Scope. *RAND Journal of Economics* 25, 319-333.

Levitt, B., March, J.G., 1988. Organizational learning. *Annual review of sociology*, 319-340.

Löf, H., Johansson, B., 2014. R&D Strategy, Metropolitan Externalities and Productivity. *Industry and Innovation* 21, 141-154.

Lundvall, B.-ä., Johnson, B., 1994. The learning economy. *Journal of industry studies* 1, 23-42.

Lüthje, C., Herstatt, C., von Hippel, E., 2005. User-innovators and "local" information. *Research Policy* 34, 951-965.

Malerba, F., 1992. Learning by firms and incremental technical change. *The Economic Journal* 102, 845-859.

March, J.G., Simon, H.A., 1958. *Organizations*. Wiley, New York.

McIver, D., Lengnick-Hall, C., Lengnick-Hall, M., Ramachandran, I., 2013. Understanding work and knowledge management from a knowledge-in-practice perspective. *Academy of Management review* 38, 597-620..

Moilanen, M., Østbye, S., Woll, K., 2014. Non-R&D SMEs. *Small Business Economics* 43, 447-462.

Nelson, R.R., 1981. Research on productivity growth and productivity differences. *Journal of Economic Literature*, 1029-1064.

Noble, D., 1984. *Forces of production*, New York: Knopf.

NSB, 2012. Science and Engineering Indicators 2012, Arlington VA: National Science Foundation (NSB 12-01).

NSF, 2002. Research and Development in Industry: 1999. NSF 02-312. Division of Science Resources Statistics, Arlington, VA.

NSF, 2014. Business R&D and Innovation: 2011. Detailed Statistical Tables NSF 15-307. National Center for Science and Engineering Statistics, Arlington, VA. Available at <http://www.nsf.gov/statistics/2015/nsf15307/>.

OECD, 2002. Frascati Manual 2002. OECD Publishing: Paris.

Organ, D.W., 1988. Organizational Citizenship Behavior. Lexington Books, Lexington, MA.

Owan, H., Kim, J.-H., 2013. Differentiation, interface technology, and the economies of scope as the determinants of product outsourcing. Hitotsubashi University, Tokyo, Japan 18 March 2013, Presentation at Research Workshop on Innovation Process.

Pavitt, K., 1984. Sectoral patterns of technical change. *Research Policy* 13, 343-373.

Rammer, C., Czarnitzki, D., Spielkamp, A., 2009. Innovation success of non-R&D-performers. *Small Business Economics* 33, 35-58.

Rosenberg, N., 1976. On technological expectations. *The Economic Journal*, 523-535.

Rosenberg, N., 1982. Inside the black box. Cambridge University Press.

Santamaria, L., Nieto, M.J., Barge-Gil, A., 2009. Beyond formal R&D. *Research Policy* 38, 507-517.

Seymore, S.B., 2009. Serendipity. *NCL Rev.* 88, 185.

Sichterman, J., 2001. Harley Davidson. Original Productions.

Simon, H.A., 1997. Administrative behavior. The Free Press, New York, NY.

Sinclair, G., Klepper, S., Cohen, W., 2000. What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. *Management Science* 46, 28-45.

Smith, A., 1776. The wealth of nations. London.

Solow, R.M., 1994. Perspectives on growth theory. *The Journal of Economic Perspectives* 8, 45-54.

Spear, S., Bowen, H.K., 1999. Decoding the DNA of the Toyota production system. *Harvard Business Review* 77, 96-108.

Spithoven, A., Clarysse, B., Knockaert, M., 2011. Building absorptive capacity to organise inbound open innovation in traditional industries. *Technovation* 31, 10-21.

Stinchcombe, A.L., 1965. "Social structure and organizations" p142-193 in *Handbook of organizations* edited by James G. March. Oxford University Press, USA.

Stinchcombe, A.L., 1990. Information and organizations. University of California Press.

Thomas, R.J., 1994. What machines can't do. University of California Press.

Thompson, P., 2012. The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. *The Journal of Economic Perspectives* 26, 203-224.

Trajtenberg, M., 1990. A Penny for Your Quotes. *RAND Journal of Economics* 21, 172-187.

Trigo, A., 2013. The Nature of Innovation in R&D-and Non-R&D-Intensive Service Firms. *Industry and Innovation* 20, 48-68.

Vallas, S.P., 2003. Why teamwork fails. *American sociological review*, 223-250.

Van de Ven, A.H., Delbecq, A.L., Koenig Jr, R., 1976. Determinants of coordination modes within organizations. *American sociological review*, 322-338.

von Hippel, E., 1994. "Sticky Information" and the Locus of Problem Solving. *Management Science* 40, 429-439.

Wakelin, K., 2001. Productivity growth and R&D expenditure in UK manufacturing firms. *Research Policy* 30, 1079-1090.

Walsh, J.P., Lee, Y.-N., Nagaoka, S., 2015. Openness and innovation in the US. Working paper.

Walsh, J.P., Tseng, S.-f., 1998. The effects of job characteristics on active effort at work. *Work and Occupations* 25, 74-96.

Warner, M., 2014. Pandora's Lunchbox. Simon and Schuster, New York.

Winter, S.G., 1984. Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization* 5, 287-320.

Winter, S.G., 2003. Understanding dynamic capabilities. *Strategic Management Journal* 24, 991-995.
Zuboff, S., 1988. *In the age of the smart machine: The future of work and power*. Basic Books, New York.

Figure 1. Framework of origin and type of innovation

		Type of innovation	
		Process	Product
Origin of innovation	R&D	(B)	(C)
	Non-R&D	(A)	(D)

tables

Table 1. Measures of R&D and non-R&D invention

Creative process that led to the invention	Location of the Inventor						Total
	Independent R&D unit or its subunit	R&D subunit attached to a non-R&D unit	Manufacturing	Software development	Other (e.g., sales)	(missing)	
Targeted achievement of a R&D project	626	147	63	19	33	7	895
Unexpected by-product of a R&D project	132	35	12	10	10	2	201
Expected by-product of a R&D project	123	49	6	10	3	1	192
Related to your normal job (not inventing)	97	43	31	7	21	0	199
Pure inspiration/creativity	129	39	37	10	31	0	246
(missing)	5	0	0	0	0	0	5
Total	1112	313	149	56	98	10	1738

Colored cells defined as non-R&D invention.

Table 2. Descriptive statistics and correlation matrix

Variable	N	Mean	SD	Min	Max	Correlation								
						1	2	3	4	5	6	7	8	9
1 R&D (vs. Non-R&D) inv	1738	0.88	0.33	0.00	1.00	1.00								
2 Generality of knowledge	1738	0.00	2.32	-3.44	8.34	0.08	1.00							
3 Codifiable	1738	0.00	1.00	-3.05	2.46	0.03	0.71	1.00						
4 Basic	1738	0.00	1.00	-1.01	4.18	0.10	0.82	0.35	1.00					
5 University-driven	1738	0.00	1.00	-2.34	1.70	0.05	0.80	0.29	0.56	1.00				
6 Visibility of knowledge	1738	0.57	0.50	0.00	1.00	-0.10	-0.12	0.02	-0.35	0.05	1.00			
7 Inventor patent propensity	1384	68.24	32.44	0.00	100	0.04	0.03	0.02	0.02	0.04	-0.03	1.00		
8 Industry annual growth	1738	-2.94	2.80	-6.42	6.90	0.04	0.18	-0.09	0.52	-0.01	-0.11	-0.01	1.00	
9 Firm size	1738	6.23	0.89	3.91	6.62	0.03	0.00	-0.03	0.02	0.00	-0.05	-0.04	-0.05	1.00
10 Log R&D intensity	1718	-3.03	0.87	-7.42	-0.75	0.02	0.31	0.34	0.11	0.28	0.03	-0.09	0.06	-0.14

Bold at $p < .05$

Table 3. Statistics of non-R&D invention and innovation in US manufacturing industries

NAICS	N	NRD Invention	Non-R&D inventors' work units			
			R&D subunit attached to non-R&D unit (e.g., technical service)	Manufacturing	Software development	Others (e.g., sales)
		%	(% of non-R&D inventions)			
		(1)	(2)	(3)	(4)	(5)
311-2 Food, beverage and tobacco product manufacturing	5	0.0	0.0	0.0	0.0	0.0
313-6 Textile, apparel and leather	10	21.4	0.0	100.0	0.0	0.0
321-3 Wood product, paper, printing and related support activities	21	20.8	80.0	20.0	0.0	0.0
325 Chemical manufacturing (except pharmaceutical and medicines)	257	4.1	41.7	41.7	0.0	16.7
3254 Pharmaceutical and medicine manufacturing	56	4.5	33.3	33.3	0.0	33.3
326 Plastics and rubber products manufacturing	78	9.1	25.0	37.5	0.0	37.5
327 Nonmetallic mineral product manufacturing	36	11.9	60.0	20.0	0.0	20.0
331 Primary metal manufacturing	16	10.5	50.0	50.0	0.0	0.0
332 Fabricated metal product manufacturing	74	19.5	31.3	50.0	0.0	18.8
333 Machinery manufacturing	266	18.6	42.1	33.3	1.8	22.8
334 Computer and electronic product manufacturing (except semiconductor)	441	12.9	28.4	28.4	16.4	26.9
3344 Semiconductor and other electronic component manufacturing	177	11.7	53.9	23.1	15.4	7.7
335 Electrical equipment, appliance, and component manufacturing	121	11.5	52.9	23.5	11.8	11.8
336 Transportation equipment manufacturing	63	9.5	50.0	37.5	0.0	12.5
337-9 Furniture and related product, and miscellaneous manufacturing	117	14.8	25.0	30.0	5.0	40.0
All	1738	12.1	38.6	32.1	7.6	21.7

Reproduced from Lee (2015a, p12).

Table 4. Descriptive statistics

	Invention type		
	R&D inv (N=1519)	Non-R&D inv (N=219)	t
Inventor characteristics			
Age at first patent application	34	37	-5.19 ***
Age at highest degree	28	27	1.97 *
Highest degree = PhD	0.48	0.24	7.62 ***
Highest degree major = Science/Engineering	0.98	0.92	2.81 ***
Invention process			
No. of information sources (university, customer, supplier etc., max = 11)	5.07	4.47	2.61 ***
Invention output			
Product (vs. Process) invention	0.79	0.80	-0.15
Value of invention			
Any commercialization (Inhouse, start-up, or licensed)	0.53	0.64	-2.75 ***
No. of claims	23.18	22.64	0.47
Forward citation	3.16	3.44	-0.75

*** at .01, ** at .05, * at .10

MBA and JD excluded for comparison of highest degree major

Forward citation as of July 1st 2008

Table 5. Knowledge environment and R&D vs. non-R&D invention

Variables	R&D invention vs. non-R&D invention						
	Logit						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Generality of knowledge	0.151 ** (0.065)				0.141 ** (0.066)	0.215 ** (0.108)	0.030 (0.104)
Codifiable		0.210 * (0.119)					
Basic			0.433 ** (0.183)				
University-driven				0.198 ** (0.098)			
Visibility of knowledge	-0.704 *** (0.203)	-0.713 *** (0.203)	-0.550 *** (0.204)	-0.757 *** (0.208)	-0.724 *** (0.205)	-0.881 *** (0.332)	-0.810 ** (0.397)
Inv. patent propensity	0.003 (0.003)	0.003 (0.003)	0.003 (0.003)	0.003 (0.003)	0.003 (0.003)	0.005 (0.005)	0.009 * (0.005)
Industry annual growth	0.105 ** (0.051)	0.102 ** (0.050)	0.055 (0.041)	0.088 ** (0.041)	0.113 ** (0.053)	0.129 (0.089)	0.143 (0.101)
Firm size	0.111 (0.094)	0.113 (0.095)	0.100 (0.095)	0.108 (0.094)	0.102 (0.098)	0.025 (0.219)	
Log R&D intensity					0.010 (0.105)	-0.215 (0.236)	
Firm dummies	No	No	No	No	No	No	Yes
Observations	1384	1384	1384	1384	1370	444	444
<i>Model F</i>	F(5, 1379)	F(5, 1379)	F(5, 1379)	F(5, 1379)	F(6, 1364)	F(6, 438)	F(44, 400)
	4.11 ***	4.08 ***	3.92 ***	4.31 ***	3.44 ***	1.95 *	1.47 **

*** at .01, ** at .05, * at .10

Table 6. The effects of nature of knowledge

	Generality	Visibility
Marginal probability effect	0.013 (0.006)	-0.062 † (0.017)
Discrete Changes in the predicted probabilities		
min → max	0.118	-0.062
25th → 75th	0.053	-0.062
-+ SD/2 from the mean	0.029	-0.032
Odds (Changes with $\Delta x = 1$)		
Factor change (Odds ratio)	1.151 (0.076)	0.485 (0.099)
Relative change	15.1% (7.6%)	51.5% (9.9%)

† This is for discrete change of dummy variable from 0 to 1.

Table 7. Robustness tests for alternative measures of R&D vs. non-R&D invention.

Variables	Creative process criteria		Affiliation criteria			
	Logit		Logit		Multinomial probit	
	All normal job = RDinv	All normal job = missing	All R&D sub = RDinv	All R&D sub = missing	Base group = All R&D sub (to non-R&D) inv	
	(1)	(2)	(3)	(4)	RDinv (5)	NRDinv (6)
Generality	0.189 ** (0.088)	0.203 ** (0.090)	0.068 (0.082)	0.076 (0.081)	0.055 * (0.029)	0.007 (0.048)
Visibility	-0.633 ** (0.263)	-0.666 ** (0.266)	-0.724 *** (0.256)	-0.731 *** (0.256)	-0.024 (0.126)	0.430 ** (0.175)
Inv. patent prop	0.007 * (0.004)	0.007 * (0.004)	0.005 (0.004)	0.006 (0.004)	0.005 ** (0.002)	0.000 (0.003)
Industry growth	0.129 * (0.076)	0.139 * (0.077)	0.123 * (0.070)	0.124 * (0.069)	0.027 (0.022)	-0.049 (0.039)
Firm size	0.162 (0.128)	0.156 (0.130)	0.300 *** (0.106)	0.247 ** (0.108)	-0.286 *** (0.078)	-0.434 *** (0.093)
Log R&D intensity	0.085 (0.142)	0.069 (0.143)	0.029 (0.145)	0.035 (0.142)	0.048 (0.068)	0.024 (0.101)
Observations	1370	1226	1370	1133	1370	
Model F	F(6, 1364) 2.49 **	F(6, 1220) 2.62 **	F(6, 1364) 3.80 ***	F(6, 1127) 3.33 ***	F(12, 1358) 4.18 ***	

*** at .01, ** at .05, * at .10