Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

Potluck or Chef de Cuisine? Knowledge Diversity and Award-Winning Inventor Teams

Dennis Verhoeven
KU Leuven VAT: BE0419052173
Managerial Economics, Strategy and Innovation
Dennis.Verhoeven@kuleuven.be

Abstract

Is the trend towards specialization by knowledge workers detrimental to the creation of frontier-pushing inventions? The answer to this question is 'yes' under two conditions. First, diverse knowledge is a source of frontier-pushing inventions. Second, teams having a 'chef de cuisine' – an individual with high knowledge diversity – are more effective in generating frontier-pushing inventions as compared to teams obtaining diversity through a combination of specialized team members – a 'potluck' structure. Using the 'R&D100 Awards' as a measure for frontier-pushing performance, this study shows the importance of knowledge diversity for the creation of frontier-pushing inventions. Moreover, 90% of the variance in team knowledge diversity is accounted for by the most diverse team member, indicating that a potluck-structure is rare. Diversity of the most diverse member drives the likelihood of pushing the technological frontier, while the diversity added by the team does not significantly contribute to the effect. These results indicate individual-level diversity cannot be substituted by diversity generated through a combination of specialists. In light of the documented trend towards specialization, these findings raise concerns for policy makers and managers targeting frontier-pushing inventions.

Jelcodes:O31,-

Potluck or Chef de Cuisine?

Knowledge Diversity and Award-Winning Inventor Teams

Dennis Verhoeven

KU Leuven

Faculty of Business and Economics

Department of Management, Strategy and Innovation (MSI)

E-mail: Dennis.Verhoeven@kuleuven.be

Working Paper

March 2016

ABSTRACT

Is the trend towards specialization by knowledge workers detrimental to the creation of frontier-pushing inventions? The answer to this question is 'yes' under two conditions. First, diverse knowledge is a source of frontier-pushing inventions. Second, teams having a 'chef de cuisine' – an individual with high knowledge diversity – are more effective in generating frontier-pushing inventions as compared to teams obtaining diversity through a combination of specialized team members – a 'potluck' structure. Using the 'R&D100 Awards' as a measure for frontier-pushing performance, this study shows the importance of knowledge diversity for the creation of frontier-pushing inventions. Moreover, 90% of the variance in team knowledge diversity is accounted for by the most diverse team member, indicating that a potluck-structure is rare. Diversity of the most diverse member drives the likelihood of pushing the technological frontier, while the diversity added by the team does not significantly contribute to the effect. These results indicate individual-level diversity cannot be substituted by diversity generated through a combination of specialists. In light of the documented trend towards specialization, these findings raise concerns for policy makers and managers targeting frontier-pushing inventions.

Introduction

As inventors are the crux of innovative activity, understanding what drives their performance is pivotal to policy makers and managers. Yet, only recently economists have started to devote attention to individual inventors and inventive teams to explain the rate and direction of technological change (i.a. Audia & Goncalo, 2007; Fleming et al., 2007; Sauermann & Cohen, 2010; Gruber et al., 2013; Conti et al., 2013). Individual-level research on the organization of scientific and inventive activity has established a remarkable and robust trend towards increased reliance on teamwork in the production of knowledge (Wuchty et al., 2007). This trend can be in part explained by the existence of a 'burden of knowledge' (Jones, 2009; Agrawal et al., 2016). As the amount of accumulated knowledge increases, it becomes ever harder for individuals - born without any knowledge - to reach the scientific or technological frontier. As a result, inventors limit their scope of expertise and increase reliance on teamwork. This mechanism is corroborated in the finding that the age of great inventive achievement has increased over last century (Jones, 2010).

The increasing burden of knowledge might particularly affect frontier-pushing inventive outcomes as these are often claimed to be the result of diverse recombination of knowledge. If this is true, an increasing tendency of inventor specialization warrants policy concern for two reasons. First, these few game-changing inventions are arguably responsible for the lion's share of technological progress (Baumol, 2004). With their potential to disrupt existing capabilities of firms and transform industries, they are at the source of creative destruction driving economic growth (Schumpeter, 1939). Second, market failures for innovation are most pronounced for projects targeting to push the frontier because of the uncertainty surrounding their outcomes and appropriation (Arrow, 1962; Fleming, 2001; Hall & Lerner, 2010) . As such, they are a centerpiece of social welfare considerations.

Should the increasing trend towards specialization be a reason for concern about the occurrence of frontier-pushing inventions? A first condition for this to be true, is that diverse knowledge is indeed an important precursor of frontier-pushing technological outcomes. The dominant view among innovation scholars is that technological search over a variety of technological domains is the wellspring of (breakthrough) innovative activity (Trajtenberg et al., 1997; Rosenkopf & Nerkar, 2001; Ahuja & Lampert, 2001). As such inventive teams holding diverse knowledge can be expected to excel at frontier-pushing performance. Behind this view are at least two different mechanisms. First, access to knowledge from different technological domains increases the number of components and principles at hand for recombination, increasing the likelihood to make the new connections that characterize breakthrough inventions. A second mechanism relies on the premise that exposure to diverse knowledge decreases the likelihood of a mental 'lock-in' into an existing paradigm, promoting distant search and breakthrough insights (Jeppesen & Lakhani, 2010). However, diverse knowledge – keeping constant the amount of knowledge – comes at the cost of depth. Hence, a contrasting perspective argues that deep, specialized knowledge allows inventors to identify anomalies of an existing paradigm, increasing the likelihood of breakthrough inventions (Weisberg, 1999; Kaplan & Vakili, 2015).

This study addresses this question by investigating how diversity of a team's knowledge – as evidenced in previous patents – affects the likelihood of generating frontier-pushing inventions. To this end, I analyze a unique dataset of 264 frontier-pushing inventions that received the 'R&D100 Award' between 2002 and 2009. The award-winning inventions are linked to their patents, and compared to patents from the same firms in the same time span. I calculate a knowledge diversity measure which incorporates the three dimensions of diversity – variety, balance and disparity (Stirling, 2007). Controlling for a large number of team-level variables and firm-fixed effects, I find a strong, positive and non-decreasing effect of team knowledge diversity. When distinguishing between diversity that was used in the invention, versus unrelated diversity, I find this effect to be driven by diverse knowledge not incorporated in the invention – supporting a view in which diversity helps paradigm-breaking insights beyond the pure potential for recombination.

Is diverse knowledge more effective when held by an individual inventor? A second condition for increasing specialization to be a concern is that individual knowledge diversity cannot be effectively substituted by diversity attained through a combination of team members. Reliance on teamwork proposes a viable solution to increasing specialization as it enables to form a diverse knowledge base (Singh & Fleming, 2010), while fully reaping the well-understood benefits of specialization (Smith, 1776; Ricardo, 1817). The key assumption behind this view is that the generation and selection of potential new solutions is at least as effective in teams as within an individual. Consistent with theories positing that knowledge recombination is more effective within the individual mind (Simon, 1985; Taylor & Greve, 2006; Melero & Palomeras, 2015), this study argues that diversity held within a single person is essential to envision solutions that break with existing paradigms.

To investigate this second condition, I distinguish between diversity built up by the most diverse individual and the portion of diversity added by team members. I find that 90 percent of the variation in total team diversity is explained by the most diverse individual. The largest part of the diversity effect is driven by individual-level diversity of the most diverse individual, and diversity added by the team does not significantly contribute to the effect. Ergo, added team-level diversity seems not to make up for the absence of individual-level diversity. These findings support the view that, to create frontier-pushing inventions, obtaining diverse knowledge through a mere collection of specialized inventors cannot substitute for reaching the same level of diversity through a single member of a team.

In sum, the results support the presence of the two conditions necessary for concern about the well-documented trend towards specialization and teamwork. They call for further investigation into the drivers of individual-level diversity building and how these decisions are influenced by an inventor's personal traits and her environment. Furthermore, they might inspire policymakers on devising instruments targeting the creation of frontier-pushing inventions, and managers on hiring decisions and human resource practices.

Theoretical Background and Hypotheses

Frontier-Pushing Inventions

Relevance

Not all inventions are equal. Much like the development of scientific knowledge, technology evolves along trajectories defined by paradigms (Kuhn, 1962; Dosi, 1982). Inventions that shift these paradigms – think about inventions such as Google's Pagerank algorithm, the turbojet engine, or the polymerase chain reaction – are typically few and far between. Yet, they are of pivotal importance to free-market economic growth because of their potential to open up new markets, disrupt existing ones, destroy firms' existing capabilities, and shape future technological trajectories (Schumpeter, 1939; Cooper & Schendel, 1976; Henderson & Clark, 1990; Baumol, 2004; Arthur, 2007, 2009). As such, understanding their origins and effects – at any level of economic analysis – is essential to policymakers, managers and entrepreneurs alike.

Besides their potential for stimulating social welfare beyond the private value they generate, frontier-pushing inventions warrant policymakers' attention because of their sensitivity to market failures (Arrow, 1962). R&D projects targeting breakthrough inventions are subject to exceptional uncertainty with respect to their commercial and technological outcomes (Fleming, 2001; Hall & Lerner, 2010). Many projects introducing a novel approach fail, and even the ones that succeed typically go through a lengthy process of refinement before coming to fruition (Rosenberg, 1976; Arthur, 2009). As a result, appropriation of their value is not guaranteed. Because of the uncertainty in outcomes and difficulties of appropriation, it is plausible to argue that the number projects targeting to push the technological frontier, is lower than socially optimal.

Definition

As many labels – breakthrough, radical, discontinuous, novel, high-potential – are given to concepts closely related to what I call 'frontier-pushing' inventions, it seems useful to state what type of inventive outcome is dealt with in this study. Frontier-pushing inventions are those that provide a significant potential to drastically increase performance of the products relying on the invention. As such they present a leap forward on at least one dimension of performance compared to what was possible with previous technologies. They typically present a high-potential novel approach to a problem. While they embody a distinct potential for commercial value, that does not define them, as many other factors influence the realized (private) value.

Search Processes

As opposed to run-off-the-mill, incremental improvements, frontier-pushing inventions typically embody new connections between previously unconnected components, principles and fields of knowledge (Nelson & Winter, 1982; Hargadon, 1998; Weitzman, 1998; Fleming, 2007; Verhoeven et al., 2016). For instance, Google's Pagerank algorithm was first in applying bibliometric principles – such as bibliographic coupling and co-citations – to a web-based search algorithm. Grounded in this perspective about how frontier-pushing inventions come about, a range of studies have identified diverse knowledge sourcing to be at the basis of frontier-pushing (often labelled 'breakthrough') inventive outcomes (Trajtenberg et al., 1997; Schoenmakers & Duysters, 2010; Kelley et al., 2013; Kaplan & Vakili, 2015)

Similar arguments are resonated in research investigating firm-level determinants of breakthrough inventions. This stream of literature investigates technological search strategies that might lead to inventions with high technological impact. A common theme across these studies is the notion that firms have to break away from their existing knowledge base to overcome so-called ‘familiarity traps’ (Ahuja & Lampert, 2001). To do so, they should search for knowledge residing outside the firm’s boundaries (Rosenkopf & Nerkar, 2001; Phene et al., 2006; Jiang et al., 2010).

As individual inventors are at the core of inventive activity, driving both firm and economic performance, it is surprising that this level of analysis has not been at the center stage of economists’ attention when investigating search processes leading to breakthrough performance. Indeed, a large portion, if not all, of a firm’s knowledge is stored insight the heads of its personnel. Likewise, acquiring new knowledge as a source for invention is done by an organization’s personnel (Simon, 1991). As inventors typically have considerable autonomy in directing the search process leading to a new solution (Sauermann & Cohen, 2010), the nature of their knowledge base is arguably an important determinant of frontier-pushing outcomes.

To build up theory on how we should expect the nature of previous knowledge to affect inventive outcomes, I start with characterizing the inventive search process. The work of Arthur (2007, 2009) proves to be a particularly relevant framework in this respect. He characterizes (radical) invention as a recursive process which starts with a new connection between a base principle and a purpose. As an example, Arthur (2007) describes how the turbojet engine introduced the concept of generating thrust by expelling particles to create an opposite force to accelerate an airplane. This was a novel technological approach compared to the paradigmatic approach of using propeller engines that generate a drag in order to drive the plane. Once a viable new principle is found, a lengthy process of follow-on improvement starts in order to reach a working technology and increase its performance. This process consists of finding and configuring suitable elements (which are technologies themselves) to exploit the new principle. More often than not, components and principles have to be used from different technological domains (‘redomaining’) or invented altogether.

Using this framework, I assert that inventive activity largely consists of a search process over technological elements and principles of working in existence (Perkins, 1988; Maggitti et al, 2013). Moreover it consists of shaping and reassembling these elements and principles as to serve the purpose at hand. Clearly, the choice of which elements and principles to use in this combinatory process does not consist of a random ‘mix and match’. More plausibly, the inventor is guided by his/her knowledge and assumptions about technological components and scientific/engineering principles. Provided that this knowledge is to a large extent determined by the nature of previous inventive activities, the search process being applied is shaped by the inventor's existing technological knowledge (Rosenberg, 1982; Cohen & Levinthal, 1990; Schilling et al., 2003).

Are Frontier-Pushing Inventions a Result of Inventor Teams with Diverse Knowledge?

The literature offers two contrasting perspectives that can serve to answer this question (Kaplan & Vakili, 2015). First – and most widespread among innovation scholars – is the so-called 'tension view'. Generally, this perspective sees diverse knowledge as the ultimate way to push the technological frontier – either through enlarging the combination space available, or through decreased sensitivity to 'paradigm-thinking'. A second perspective – the 'foundational view' – sees a paradigm shift as triggered by accumulated anomalies, and a deep immersion into a field as the only way to recognize and solve them. In what follows I structure the arguments of both sides and formulate competing hypotheses.

The Tension View

A large stream of literature in innovation has emerged based on the inherent tension between exploration of new, outside knowledge and exploitation of existing knowledge (March, 1991). Albeit with a focus on the organizational structure, it highlights how organizational routines necessary for exploring new domains are fundamentally different from those suited for the exploitation of existing knowledge and capabilities. Knowledge, components, methods, assumptions and perceived 'truths' accumulate to form clusters and result in the notion of barriers between different technological approaches (von Tunzelmann, 1998; Hargadon, 2006; Arthur, 2009). Then, relying on a recombination perspective (Hargadon & Sutton, 1997; Fleming, 2001), it is believed that frontier-pushing inventions are the result of exploratory, distant and/or boundary-spanning search (Trajtenberg et al., 1997; Rosenkopf & Nerkar, 2001; Phene et al., 2006). This view is (implicitly) based on the idea that creativity is accomplished through an evolutionary process of (blind) variation and retention of combinations of knowledge (Campbell, 1960; Simonton, 1999). By this mechanism, the mere availability of more components and principles considered for recombination by an inventive team, enhances the likelihood of a new one being considered as useful.

Scholarship on (individual) creativity has, however, suggested a different mechanism through which diversity of knowledge can help the creation of high-potential novel solutions. Here, beyond the mere increase of potential candidate components, diverse knowledge helps creativity in a less mechanical way. Having had experience with a larger number of fields, an individual is less prone to mental lock-in (French & Sternberg, 1989) and she is more likely to identify and challenge the assumptions that the common paradigm presents (Gieryn & Hirsch, 1983). A number of studies has empirically established how 'marginality' – being at the margin of a knowledge domain – can be the source of novel solutions to problems considered difficult by insiders (Ben-David 1960; Mullins, 1972; Law, 1973; Jeppesen & Lakhani, 2010). Per this argument, having knowledge on a diverse set of technological domains might help frontier-pushing invention – beyond a purely mechanical increase of number of candidate elements for recombination. As such, a diverse knowledge base might guide the direction of search for new components – not necessarily within the experience base – that are candidates to enhance technological performance.

According to this view, a diverse knowledge base of inventors – because of an enlarged space for recombinant opportunity, and through a decreased sensitivity to search within the existing paradigm – should increase the likelihood of frontier-pushing performance.

Hypothesis 1a: Inventor teams with diverse knowledge are <u>more</u> likely to generate frontier-pushing inventions.

Foundational View

A less known perspective – at least among innovation scholars – can lead to the opposite answer regarding the relationship between diversity and frontier-pushing outcomes. By this perspective, a technological paradigm presents an increasing number of anomalies as technology evolves. It follows a Kuhnian logic in which a paradigm shift addresses (a number of) anomalies in presenting a new solution (Kuhn, 1962). The backbone premise of this theory is that to be able to identify these anomalies an inventor has to immerse herself into a domain completely (Csikszentmihalyi, 1996). Diversity might blur an individual's view and obstruct deep expertise in a technological domain, necessary to identify which rules to break (Taylor & Greve, 2006).

The foundational view is much less tested than the tension view. Kaplan & Vakili (2015) find that inventions relying on a broad recombination of existing knowledge are less likely to result in cognitively novel ideas – measured by the introduction of new topics in the patent literature. However, as this study does not focus on the effect of inventors' experience, this view is hitherto not tested in relation to individual knowledge diversity. Therefore, I formulate the following competing hypothesis:

Hypothesis 1b: Inventor teams with diverse knowledge are less likely to generate frontier-pushing inventions.

Can Diverse Teams Substitute Diverse Individuals?

While being an important topic in the organizational behavior literature (Jackson, 1996), the role teams play in inventive activity has received less attention (Greve & Taylor, 2006; Singh & Fleming, 2010; Melero & Palomeras, 2015). By a logic of division of labor, teams are generally viewed as the ultimate way to collect diverse knowledge without losing the benefits of specialization. The question then is whether the benefits of diversity of knowledge with respect to frontier-pushing inventions are independent of whether it is attained by single individuals or collected through teamwork.

Albeit – in theory – a relatively easy way to gather diversity through collecting specialized individuals, working in team has a number of potential drawbacks (Melero & Palomeras, 2015). I follow Arthur (2007) and Singh & Fleming (2010) in their view that any insight leading to a particular invention comes to a person, not to a team. Yet, insights can be inspired by and shared with team members – affecting the inventive outcome. Note that, to lay out this argument, I make abstraction of the total amount of knowledge and the diversity of the total knowledge base that can serve as the input for the inventive process. Then, the main difference between diversity built up through combining team members and diversity within the individual, is that knowledge needs to be transferred before any recombination can be obtained. This transfer of knowledge might be hampered by its tacit nature. Understanding of a domain is often intuitive and structuring expert knowledge in a way that it can be optimally absorbed by team members is not obvious (think about the amount of time required for preparing a good lecture).

Let us first considering this issue from the perspective that diversity helps frontier-pushing performance through a mere increase in recombinant opportunity. Then, the question becomes whether the same number of combinations envisioned will result when different team members possess different pieces individually, compared to when the pieces are present within an individual. Arguably, the likelihood of a combination being envisioned is higher when the different components reside from one head (Simon, 1985; Melero & Palomeras, 2015). The essential difference between these two scenarios is the existence of communication barriers. The amount of knowledge held by knowledge workers typically surpasses the amount that can be expressed or codified. Moreover, excessive use of jargon specific to technological domains can exacerbate this problem (Maznevski, 1994). In sum, when diversity is obtained through a combination of individuals, it is less likely that the full potential for recombination will be exploited because not all candidate elements for recombination will be considered due to communication barriers.

A second way of looking at this question is through the lens of the 'lock-in' perspective. By this mechanism, diversity helps through a decreased sensitivity to paradigm thinking, leading to less local search. Here, frontier-pushing outcomes are the result of a search process less constrained by the knowledge and assumptions constituting the existing paradigm. The question then becomes whether the benefits of diversity in directing the search strategy are present as strongly when diversity is brought together by several specialized individuals. Arguably, a combination of specialists will not benefit from this mechanism as much as individuals due to a low willingness of individual specialists to depart from their assumptions and knowledge. Indeed, a range of diverse perspectives – an inherent characteristic of diversity – can lead to conflicts inhibiting creative outcomes (Levine & Thompson, 1996; Paulus, 2000) as individuals are reluctant to diverge from their rigid perspectives. As such, from a lock-in perspective, diversity should mainly be beneficial if present within the mind of an individual.

Both mechanisms lead me to formulate following hypothesis:

Hypothesis 2: The diversity in knowledge attained by individual inventors is more effective in generating frontier-pushing inventions compared to diversity in knowledge added by the team.

Methodology

Data

The R&D100 Awards

To proxy the frontier-pushing nature of inventive outcomes I make use of a list of inventions granted with an R&D100 award (other studies using these awards include Carpenter et al., 1981; Scherer, 1989; Block & Keller, 2009; Fontana et al., 2012). Since 1965, R&D Magazine yearly grants 100 awards to new inventions with a significant potential for breakthrough impact. Firms or research institutions apply[1] for the awards, and a jury consisting of the editors of the magazine and a number of outside experts decide on which inventions receive an award. No monetary prize is awarded, but winning the award allows firms and research organizations to obtain visibility for their newly invented product.

The appeal of using award-winning inventions as a proxy of frontier-pushing inventions is twofold. First, it reduces the concern of common method bias (Siemsen et al., 2010). While patent documents arguably provide the best source of information to assess inventive activity, a number of concerns have been known since a long time (Pavitt, 1985; Griliches, 1990). One of these concerns is that unobserved variability in application, granting, classification and citation practices could introduce bias[2]. The concern for such bias is reduced in this setting as the information used to identify remarkable inventive output is drawn from a different source than the information on the independent variables.

Second, it provides us with an accurate measure for the underlying construct we are interested in. The criteria used for the awards closely reflect the frontier-pushing nature of inventions as conceptualized in this study. Each entry is assessed by the editors of R&D magazine and a jury of experts in the field of the invention. The applicant is required to show that the technology is used in a working product. Assessment is based on technological significance and the ultimate goal of the editors is to 'pick the 100 most technologically significant new products from among the entries'. They look for inventions that increase performance drastically – showing an 'orders of magnitude improvement over existing technology'. Moreover, it allows for a relatively large-scale assessment by field experts and it is an almost contemporaneous assessment of frontier-pushing potential (reducing hindsight bias). As such, the outcome measure closely reflects the kind of outcome policymakers would be interested in.

Information on the invention, its inventors and organizations of 1293 R&D100 awards granted between 2002 and 2012 is used to link the award-winning inventions to their patents and select an appropriate control group. Out of this set, only the 910 awards for which information on the inventors is available are used.

---

[1] The information reported about the application procedure, R&D magazine and the selection procedure was retrieved from http://www.rdmag.com, accessed on 3/12/2014

[2] Example: if, because of some unobserved practice by examiners, patents that receive a large number of classes also receive more citations, using forward citations as an output measure would introduce bias when investigating the effect of diverse knowledge as assessed using classification information.

Patent Data

As the data effort heavily depends on a reliable disambiguation of inventor names, I use the inventor database as constructed by Li et al. (2014) which limits the patent information used to USPTO documents. For this study, this seems not to be a severe problem because R&D magazine is a US based company. Hence, we can assume that a large part of patenting winners, will do so in the US jurisdiction. Moreover, I use patent information residing from the focal patents' family members according to the 'DOCDB' definition (Martinez, 2011).

To link R&D100 inventions to patents, I use patent applicants, inventors and abstracts. The criteria used to match the inventions to their patents are 1) having at least one inventor in common 2) having at least one applicant in common and 3) text similarity between the description of the inventions and the patent abstract. After having obtained all matches using this process, extensive manual verification was performed to ensure the award was linked to the correct patent (in the exceptional case where there was doubt, the patent was removed from the analyses). After this process, a set of 329 patents – applied for between 1997 and 2009 – remained, corresponding to over one third of all awards which listed at least one inventor. As the focus of this study is on teams of inventors, we drop all cases in which only one inventors was listed on the patent, resulting in a set of 281 award-winning inventions.

The award-winning inventive teams are compared to a control group of inventions representing 'average' inventive activity. To ensure the control group consists of inventions in a similar technological area (defined by having at least 1 IPC subclass in common with the award-winning group) selected out of all patents within the time span of the award winning group. To be able to control for the organizational environment, I will use organization-fixed effects. For that reason, all patents in the control group are applied for by organizations that won the R&D100 award at least once. Finally, to rule out the possibility that patents from the control group are also related to the award-winning inventions, I exclude all patents from the control group which list an inventor present in the award-winning group. As such, the control group consists of different inventor teams, working in similar technological areas during the same time span and in organizations that at least won an award once. For the construction of the team-level variables, the stock of USPTO patents of all inventors at the time of invention is created, requiring at least one patent by at least one of the team member previous to the invention. To gather all patents of all inventors, I use the USPTO inventor disambiguation as provided by Li et al. (2014). Applying these criteria, leaves me with a final sample of respectively 264 and 36 268 patent families in the award-winning and control group.

To construct the knowledge diversity measures, as well as the control variables, I use information of the retrieved patent families from PATSTAT. This includes information on technological classifications (IPC classification[3] scheme), citations received from other patent documents, patent and scientific non-patent references (Callaert et al., 2011) made in the search reports of the focal patents, applicant names and their sector allocation (Du Plessis et al., 2010) (including companies, not-for-profit organizations, government institutions and hospitals), number of claims, inventors, patent authority and application year.

Measurement

Knowledge Diversity

Assessing the knowledge diversity that an inventor built up through her previous inventive effort, requires a reliable account of the technological components and principles she worked with over her inventive career. To characterize the nature of knowledge acquired, I make use of IPC groups. This level in the classification hierarchy, containing about 8000 classes, is suitable because it is not too aggregated to distinguish relevant pieces of knowledge, yet aggregated enough to make distinctions that are still relevant. However, the 'distance' between different classes in the IPC scheme (nor the USPOC) is not constant for each pair of classes. Moreover, as knowledge domains develop, these different classes might become closer because they become an integral part of a technological solution. Hence, using a standard measure of concentration such as a Herfindahl-index (Melero & Palomeras, 2015) poses a problem as this measure of concentration treats different classes as equidistant.

To remedy this problem, I use a measure of diversity that incorporates variety (number of categories), balance (degree of equal occurrence of these categories), as well as disparity (distance between these categories). This measure was proposed as a general approach to

[3] The International Patent Classification (IPC) was established in 1971 by the Strasbourg Agreement and provides a hierarchical system to classify patents according to the technological areas they belong to. It uses 5 layers of detail to classify a patent documents labeled respectively 'Section' (8), Class (±130), 'Subclass' (±630), 'Group' (±8 000) and 'Subgroup' (±70 000). See http://www.wipo.int/classifications/en/ for more information.

incorporate all three dimensions of the diversity construct and is suitable in a wide variety of applications (Stirling, 1998: 2007). The measure has recently been applied using journal classifications of scientific references in articles to investigate the effects of interdisciplinary research (Wang et al., 2015).

The measure for **team diversity** starts from all unique IPC-group combinations present in the combined total patent stock of the team. Then the knowledge diversity is given by:

$$Knowledge\ Diversity = \sum_{ij(i<j)} d_{ij} * p_i * p_j$$

Where p_i, p_j represents the number of patent-class combinations of class i and class j divided by the total number of class combinations. For example, if two inventors each have one patent, containing respectively class A and B, and class B and C, then $p_A = \frac{1}{3}, p_B = \frac{2}{3}\ and\ p_C = \frac{1}{3}$. d_{ij} represents the distance between class i and class j and is calculated as one minus the cosine similarity index. The cosine similarity index was calculated based on a matrix (treating diagonal elements as zeroes) of co-occurrences of IPC groups in patents applied for during five years preceding the focal invention of the analysis. Table 1 illustrates the calculation of team knowledge diversity with a numerical example.

	Patents	Classes	i - j	pi	pj	dij	summation
Inventor I	1	A	A - B	0,375	0,125	0,01	0,0005
	2	A, B	A - C	0,375	0,125	0,04	0,0019
Inventor II	1	C	A - D	0,375	0,25	0,12	0,0113
	2	A, D	A - E	0,375	0,125	0,02	0,0009
	3	D, E	B - C	0,125	0,125	0,43	0,0067
			B - D	0,125	0,25	0,20	0,0063
			B - E	0,125	0,125	0,40	0,0063
			C - D	0,125	0,25	0,12	0,0038
			C - E	0,125	0,125	0,60	0,0094
			D - E	0,25	0,125	1,00	0,0313
Total number patent-class combinations = 8				**Knowledge Diversity**			**0,0781**
				Herfindahl Index/2			**0,375**

Table 1: Illustration of the calculation of knowledge diversity and comparison to the 1-Herfindahl index. As the theoretical maximum if 0.5, 1-Herfindahl is divided by two for comparison

The analyses also aim at disentangling the extent to which knowledge diversity benefits frontier-pushing inventions through the mechanism of 'recombination' or through the mechanism of decreased sensitivity to mental 'lock-in' (cfr. supra). To do this, I split up the total diversity measure into two parts, based on whether the knowledge used for the invention resided in the previous knowledge stock of the inventors. If pure potential for recombination drives the effect, the part of the diversity that was used in the invention should pick up the effect. On the other hand, if knowledge diversity non-specific to the invention drives the effect, it can be argued the mechanism follows the 'lock-in' logic. Indeed, according to this logic, diversity helps through directing the search process which should lead to use of knowledge not necessarily present in the previous stock of knowledge. To this end, I recalculate knowledge diversity using only these classes that are present in the set of classes assigned to backward references – reflecting knowledge domains drawn from during the search process – of the focal invention ('**specific team diversity**'). Finally, **non-specific team diversity** is calculated as the difference between total team diversity and specific team diversity.

To answer the second question – can diverse teams substitute diverse individuals? – I calculated the knowledge diversity of individual inventors using only the patents present in the inventor's stock. Then, I single out the most diverse member, and compare its diversity score to the team diversity. Note that, as diversity does not depend on the amount of knowledge, team diversity can be lower than or equal to the individual diversity of the most diverse member. To investigate hypothesis one, I compare the effect of the measure '**diversity most diverse**' to the difference between total team diversity and diversity of the most diverse ('**diversity added by team**', which can take on negative values) on the other.

Control variables

Since this sample does not allow a panel set-up to control for team- or inventor-fixed effects (nor do we use a (quasi-)experiment), the empirical strategy heavily relies on controlling for other factors related to our variables of interest, driving the probability of being an award-winner. Of course, many individual characteristics of inventors might determine the diversity of knowledge built up over one's career. If these are related to the probability of frontier-pushing outcomes, they might bias coefficients of interest. The strategy to tackle this concern relies on the assumption that many of these underlying individual characteristics can be observed through previous inventive outcomes of the inventor. As such, I calculate a variety of inventor- and team-level measures that are likely (and often shown in the literature) to affect inventor's propensity for exceptional inventive outcomes. As this study is dealing with outcomes of inventive teams, while most of these variables pick up heterogeneity among individual inventors, I take averages over the team for the measures specific to individuals. To account for the possibility that effects are driven by asymmetry among team members, the standard deviation from the mean is also included in the multivariate analyses.

As general productivity of the inventor is likely to increase the rate of frontier-pushing inventions (Conti et al., 2013), a control is included for the amount of **experience** built up by each inventor by counting the number of patent families (including at least one USPTO member) in an inventor's portfolio. Moreover, as inventors' past success has been shown to positively affect inventive outcomes (Audia and Goncalo, 2007) three control variables are included which reflect previous success of the inventors constituting the team. The first variable is meant to pick up previous **commercial success** and is calculated as the average number of claims in previous patents (Tong & Frame, 1994; Lanjouw & Schankerman, 2004). A second measure aims at picking up average **technological success** of previous inventive effort and uses the number of patent families (Bakker et al., 2016) citing the patent families in the inventor's stock of patents (Albert et al., 1991; Van Zeebroeck, 2011). As citation practices vary over technological fields, and to account for possible truncation, citation counts are normalized by IPC subclass and application year. A third measure proxies **breakthrough success** of previous inventive activity and is calculated as the share of an inventor's previous patent stock that was a 2 standard deviation outlier.

Underlying inventor characteristics – for instance, curiosity, puzzle joy or creativity – could affect how an inventor searches for new solutions which in turn would affect both diversity building and the frontier-pushing nature of inventive outcomes. To control for such underlying characteristics, three control variables related to the nature and outcome of search are included. First, I control for previous **recombinant success** – arguably correlated to creativity (Kaplan & Vakili, 2015). To this end, I use the Novelty in Recombination indicator developed in Verhoeven et al., 2016 (which is similar to the measure for generative creativity in Fleming et al., 2007). To aggregate the measure to the inventor level, the proportion of an inventor's patent stock scoring on this measure is used. Second, I control for the **spread of knowledge sourcing** by calculating the average spread of backward citations across technology classes of previous patents of an inventor (Trajtenberg et al., 1997; Verhoeven et al., 2016). Third, I include a control for the extent to which an inventor engaged in **external knowledge sourcing** (Shane, 2001; Rosenkopf & Nerkar, 2001). This measure consists of the average number of classes referred to in an inventor's patents to which the patent itself was not assigned.

A particular concern might be that inventors with a focus on general purpose technologies are more likely to win the award (as they can show broad relevance to the jury). To address this issue, '**generality**' of previous inventive effort is controlled for (Trajtenberg et al., 1997). This measure calculates the spread over technological classes of patents citing a focal patent and is averaged over an inventor's career.

Mobile inventors are likely to build up a more diverse knowledge stock, as they are active in multiple technological environments. Moreover, mobile inventors might be inherently more productive (Hoisl, 2007) and less prone to paradigm-thinking – hence, more likely to generate frontier-pushing inventions. To proxy the **mobility** of the inventors in our sample, I count the number of distinct applicants they have patented with (Melero & Palomeras, 2015). A similar concern might be raised about inventors with an extensive professional network (Fleming et al., 2007). To control for such **network**, I include a count of the unique previous collaborators of an inventor.

As the length of the inventive career might affect both diversity building and the likelihood to push the frontier (Jones, 2007; Conti et al., 2013), I control for inventors' **tenure** by calculating the time elapsed since their first patent at the time of the focal patent. Moreover, I include the **number of inexperienced** team members to control for the effect of 'fresh' inventors entering the field (Jeppesen & Lakhani, 2010; Audia & Goncalo, 2010; Conti et al., 2013).

Although the measures for diversity are, in theory, not dependent on the number of patents they are calculated over, larger teams might display a larger diversity as the goal of teams might be to increase diversity. Moreover, team size might affect (breakthrough) productivity (Singh & Fleming, 2010). For these reasons, I include a control for **team size**. Moreover, I control for effects due to the **joint experience** of the team by including the number of patents they jointly appeared on previous to the focal invention.

As classification and citation practices might differ across technological domains and time, I also include a set of **technology domain** and **year** dummies. Technological domains are obtained using the (broad) classification scheme developed by Schmoch (2008). Year dummies are based on the application year of the focal patent.

Both diversity building and the nature of inventive outcomes might be determined by the organizational environment of the inventors. Depending on the implications that one wants to draw from the results, these characteristics of the environment should be controlled for. If we assume that the other controls capture the relevant individual characteristics that drive selection into different types of organizations, controlling for the environment might take away relevant (exogenous to the inventor) variability in diversity building. On the other hand, if there are unobserved variables which drive both diversity building and frontier-pushing outcomes captured by the selection into organizations, controlling for the environment is advisable. For these reasons, I report models with and without organization-level controls. When only controlling for the organization type, I use the **sector**-allocation algorithm developed in Du Plessis et al. (2010) for the applicants of the focal patent. To distinguish between small and large firms I use the size of their previous patent stock employing a threshold of 25 patents to distinguish between small and large firms. When including firm-fixed effects, dummies are included for all **applicants** that appeared on at least one R&D100 award. I include a dummy to control for whether a patent was co-applied ('**co-patent**') in all specifications.

Finally, as the measure of experience diversity specific to the invention relies on the classification of backward patent references, I control for the extent to which the invention relied on previous knowledge. Hence I include controls for the number of backward patent references ('**Technology Sourcing**'). Additionally, to control for the extent to which an invention relied on science ('**Scientific Sourcing**'), which might affect the probability to receive the award as the jury might favor frontier-pushing inventions relying on science, I include the number of references to scientific literature (Callaert et al., 2011).

Analyses

Following the two questions outlined in the theory, the analyses are performed in two stages. A first stage investigates the relationship between frontier-pushing inventive outcomes and the knowledge diversity attained by the entire team, using the combined patent stock of all team members to obtain diversity scores. Multivariate analyses present the results of probit regressions where a multitude of variables are controlled for[4]. To disentangle the effect of

[4] Due to the low occurrence of award-winning inventions relative to the control sample, there might be a concern that the results are driven by observations in the control group without common support along a number of covariates. To mitigate this concern, the robustness of the results is confirmed using

diversity specific and non-specific to the focal invention on frontier-pushing outcomes respectively, I split up diversity into the portion specific and non-specific to the invention. In a second stage, I investigate whether diverse teams are able to substitute diverse individuals. To this end, individual diversity measures are calculated and compared to total diversity. As individual diversity of the most diverse member largely drives total team diversity, I continue by disentangling the separate effects of individual diversity of the most diverse member, and the contribution of all other team members to the total team diversity.

propensity score matching based on all control variables (the model with organization type fixed effects).

Variable	Description	Measure	Mean	S.D.
Team Diversity	Knowledge diversity of accumulated patent stock of all inventions	Composite diversity measure based on entire patent stock of the team (see text)	0,26	0,10
Specific Team Diversity	Knowledge diversity calculated on classes used in the focal invention	Composite diversity measure using classes cited by the focal patent	0,05	0,06
Non-Specific Team Diversity	Knowledge diversity calculated based on classes not used in the focal invention	Composite diversity measure using classes not cited by the focal patent	0,20	0,12
Diversity Most Diverse	Knowledge diversity of accumulated patent stock of the most diverse inventor	Composite diversity measure based on the patent stock of the most diverse member	0,26	0,10
Diversity Added by Team	Portion of total team diversity added by members other than the most diverse member	Team Diversity minus Diversity Most Diverse	-0,006	0,029
Team Size	Number of inventors collaborating on the invention	Number of inventors on the focal patent	3,46	1,63
Number Inexperienced	Number of inventors in the team without previous experience	Number of inventors without previous patents on the patent	0,65	0,99
Joint Experience	The extent to which the entire team has collaborated previously	Log(Number of previous patents with same team)	0,37	0,72
Experience	Amount of Inventive Activity	Log(count previous patents)	0,37	0,72
Network Size	The extent to which an inventor has co-operated with a large number of other inventors previously	Log(Number of unique co-inventors)	2,25	0,88
Mobility	The extent to which the inventor has shown mobility in terms of employers	Number of unique previous applicants	1,43	1,32
Tenure	Time elapsed since first invention	Application year invention - Application year first invention	2,25	3,66
Commercial Success	Average commercial success of previous inventive activity	Avg. Number Claims previous patents	16,33	8,44
Technological Success	Average technological success of inventive activity	Standardized Forward Citation Count	-0,01	0,44
Breakthrough Success	Rate of high impact previous inventive activity	Average highly cited patents	0,32	0,97
Recombinant Success	Rate of recombinant novelty of previous inventive activity	Average number of patents with Novelty in Recombination	0,01	0,02
Spread Sourcing	The extent to which the inventor sources from a broad range of technological domains	1-Herfindahl index based on backward references	0,59	0,20
Generality	The general purpose character of previous inventive activity	1-Herfindahl index based on citations received	0,58	0,20
External Sourcing	The extent to which the inventor customarily uses knowledge outside the technological domain of the invention	Number of IPC groups cited not present on focal patent	8,57	7,28
Technology Sourcing	The extent to which the focal invention builds on previous technological effort	Number of backward patent references	14,85	16,3
Scientific Sourcing	The extent to which the focal invention builds on previous scientific effort	Number of scientific articles referred to by the focal patent	2,18	7,69
Co-patent	Whether multiple organizations were involved in the focal invention	Dummy multiple applicants	0,031	0,18
Sector Fixed Effects	Whether the applicant was a small company, large company, university or other type of organization	Dummies type of organization	n.a.	n.a.
Domain Fixed Effects	Technology domain (Schmoch, 2008) of the focal invention	35 'fhg'-dummies	n.a.	n.a.
Year Fixed Effects	Year of the focal invention	Dummies application year	n.a.	n.a.
Firm Fixed Effects	Firm(s) involved in the focal invention	Dummies applicant	n.a.	n.a.

Table 2: Overview of variables, their description and summary statistics

Results

Are Frontier-Pushing Inventions a Result of Inventors with Diverse Knowledge?

Descriptive Statistics

Table 2 provides an overview the key variables with their description and summary statistics. Being highly right-skewed, the variables Joint Experience, Experience, and Network Size are log-transformed in all reported results (the transformation does not affect the conclusions). Note that for all variables that are aggregated from the individual level to the team level by taking the average, also the standard deviation is included for multivariate analyses. The majority of inventions in the sample (34 884 patents from 116 award-winning applicants) originate from companies with at least 25 patents (large firms). Companies with less than 25 patents account for 607 inventions (93 awards), universities and governmental organizations respectively account for 600 (28 awards) and 471 inventions (27 awards).

As the main variable of interest is based on a measure previously unused for assessing knowledge diversity of individual inventors, it is of interest to study its distribution. As can be seen in Figure 1, the distribution is close to a normal distribution with a mean of 0.26, standard deviation of 0.10 and respective minimum and maximum values of 0 and 0.48. The distribution also shows a mass on zero. These are inventor teams where only 1 IPC group occurred in all their previous patents. The maximum value of 0.48 is close to the theoretical maximum value of 0.5 of this measure. All in all, this measure seems to behave better than classical measures such as one minus the Herfindahl-index. The latter showed (not reported) a high skew to the left, meaning a large number of inventor teams displaying very high values. The difference between the measure employed in this study and classical measures mainly lies in the fact that the new measure weighs different classes by their technological distance.

Table 3 shows the results of a contingency table between the award-dummy and presence in different quartiles of the diversity distribution. This unconditional analysis shows that team diversity and the probability to be among the award-winners are positively related ($\chi^2(3)=176.7$, p-value<0.001). This correlation seems to be most present for teams in the upper quartile of the distribution. Out of 264 award-winners, about 57% are among the 25% most diverse teams (2.3 times more than expected). Compared to teams in the lowest quartile, these teams are almost 12 times more likely to be among the award-winners, indicating this positive relationship is strong.

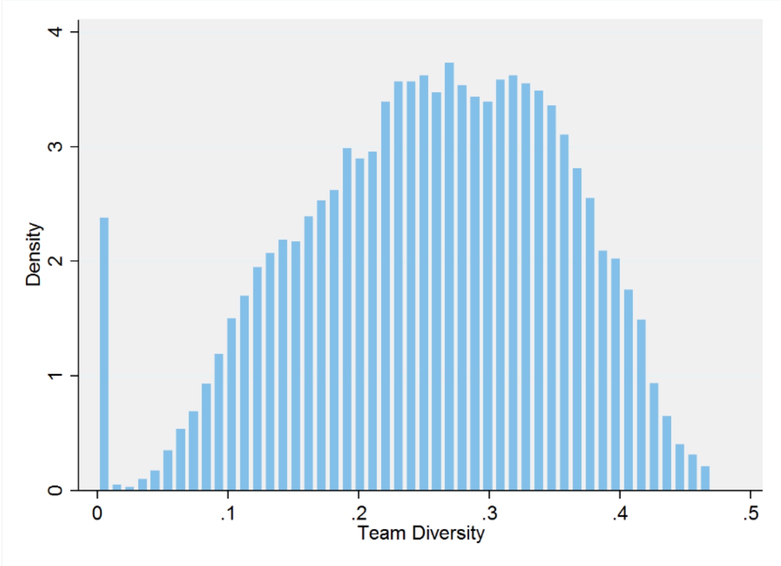

Figure 1: Histogram of the distribution of Team Diversity

	Team Diversity			
	Q1	Q2	Q3	Q4
R&D100 Award				
Observed Co-occurrence	13	30	71	150
Observed/Expected Co-occurrence	**0.20**	**0.44**	**1.10**	**2.30**
Co-occurrence as share R&D100 Awards	4.92%	11.36%	26.89%	56.82%
Co-occurrence as share of Quartile Team Diversity	0.14%	0.32%	0.80%	1.66%

Table 3: Results of contingency analysis Team Diversity and R&D Award. Chi-squared test was performed to test the null-hypothesis of being unrelated.

Table 4 shows pairwise correlations of the key variables used in the multivariate analyses (bold format indicates a correlation significant at the 1% level of confidence). Looking at the correlation coefficients of the award-dummy, confirms a relationship between team diversity and award-winning inventions. A mere count of inventors on the patent, as well as a count of the number inventors without any previous experience on the patent are positively related to winning the award, while a large experience stock decreases the probability (these findings descriptively confirm findings in Audia & Goncalo, 2007; Singh & Fleming, 2010; Conti et al., 2013).

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)
R&D100 Award	(1)	1																						
Team Diversity	(2)	**0,064**	1																					
Specific Team Diversity	(3)	0,001	**-0,038**	1																				
Non-Specific Team Diversity	(4)	**0,055**	**0,878**	**-0,511**	1																			
Diversity Most Diverse	(5)	**0,057**	**0,958**	-0,011	**0,829**	1																		
Diversity Added by Team	(6)	**0,027**	**0,175**	**-0,095**	**0,196**	**-0,115**	1																	
Team Size	(7)	**0,019**	**0,143**	**0,020**	**0,113**	**0,173**	**-0,100**	1																
Number Inexperienced	(8)	**0,019**	**-0,075**	**-0,081**	**-0,026**	**-0,112**	**0,125**	**0,405**	1															
Joint Experience	(9)	**-0,016**	**-0,015**	**0,074**	**-0,049**	-0,003	**-0,043**	**-0,228**	**-0,337**	1														
Experience	(10)	**-0,022**	**0,224**	**0,057**	**0,165**	**0,297**	**-0,246**	-0,006	**-0,471**	**0,434**	1													
Network Size	(11)	**-0,020**	**0,325**	**0,063**	**0,249**	**0,399**	**-0,246**	**0,114**	**-0,456**	**0,281**	**0,855**	1												
Mobility	(12)	**0,029**	**0,348**	**-0,022**	**0,310**	**0,367**	**-0,056**	**-0,028**	**-0,278**	**0,156**	**0,442**	**0,495**	1											
Tenure	(13)	**0,027**	**0,305**	**-0,073**	**0,297**	**0,315**	**-0,026**	**0,112**	**-0,036**	0,011	**0,286**	**0,342**	**0,829**	1										
Commercial Success	(14)	**0,023**	-0,012	**0,133**	**-0,075**	**0,019**	**-0,109**	**-0,037**	**-0,478**	**0,254**	**0,309**	**0,327**	**0,147**	-0,002	1									
Technological Success	(15)	**0,020**	**0,107**	**0,198**	-0,003	**0,111**	-0,009	**0,051**	0,001	**0,014**	0,013	**0,086**	**0,084**	**0,063**	**0,219**	1								
Breakthrough Success	(16)	0,006	**0,167**	**0,125**	**0,084**	**0,179**	**-0,036**	0,004	**-0,055**	**0,147**	**0,304**	**0,296**	**0,302**	**0,256**	**0,123**	**0,478**	1							
Recombinant Success	(17)	**0,042**	**0,227**	**0,040**	**0,176**	**0,223**	**0,019**	-0,006	**-0,051**	**0,015**	0,003	**0,023**	**0,064**	**0,038**	**0,065**	**0,108**	**0,068**	1						
Spread Sourcing	(18)	0,010	**0,310**	**0,164**	**0,188**	**0,339**	**-0,091**	**-0,056**	**-0,726**	**0,322**	**0,448**	**0,499**	**0,299**	**0,078**	**0,537**	**0,100**	**0,105**	**0,129**	1					
Generality	(19)	0,001	**0,273**	**0,151**	**0,163**	**0,305**	**-0,102**	**-0,064**	**-0,713**	**0,312**	**0,453**	**0,495**	**0,297**	**0,081**	**0,546**	**0,174**	**0,128**	**0,103**	**0,872**	1				
External Sourcing	(20)	**0,048**	**0,304**	**0,235**	**0,149**	**0,307**	0,000	**0,018**	**-0,284**	**0,149**	**0,176**	**0,239**	**0,188**	**0,081**	**0,419**	**0,315**	**0,225**	**0,238**	**0,574**	**0,463**	1			
Technology Sourcing	(21)	**0,028**	0,006	**0,400**	**-0,186**	0,012	**-0,023**	**0,088**	-0,001	**0,045**	**0,059**	**0,059**	**0,047**	**0,030**	**0,174**	**0,176**	**0,139**	**0,050**	**0,097**	**0,081**	**0,333**	1		
Scientific Sourcing	(22)	**0,031**	-0,009	**0,151**	**-0,080**	-0,009	0,001	**0,030**	**0,025**	**0,017**	-0,011	-0,005	**0,059**	**0,054**	**0,132**	**0,182**	**0,142**	**0,069**	**0,022**	**0,035**	**0,237**	**0,329**	1	
Co-patent	(23)	0,005	**0,100**	**0,015**	**0,078**	**0,096**	**0,015**	**0,095**	**0,042**	**-0,038**	-0,007	**0,018**	**0,132**	**0,125**	**-0,055**	**0,026**	**0,025**	**0,027**	-0,002	0,003	**0,018**	-0,007	**0,076**	1

Table 4: Correlation table key variables used in the analyses. Correlation coefficients printed in bold are significant at the 99 percent confidence level

Moreover, previous commercial and technological success (but not breakthrough success) are positively related to winning the award, as is previous recombinant success. With respect to the nature of search, a positive correlation with external sourcing is observed. Moreover, inventions winning the award more heavily draw on previous scientific and technological knowledge (Schoenmakers & Duysters, 2010). While the measures of most independent variables are sensitive to the intensity of previous inventive activity (explaining significant correlations), multicollinearity seems not be an issue. An exception is the high observed correlation (0.83) between Tenure and Mobility (not surprising as changing employer naturally leads to a larger number of co-inventors). Yet, I argue that because of the large sample size, there should be enough unshared variance not to distort the main results. Moreover, the results are not sensitive to including only one of these variables[5].

Multivariate Results

Table 5 shows the estimated coefficients resulting from a probit model that explains the probability of being an award-winning invention as a function of the independent variables. To save space, the coefficients of control variables are omitted. Models (1)-(3) serve to test competing hypotheses 1a and 1b. Model 1 does not include organization-specific explanatory variables, Model 2 includes sector dummies (where the sectors are: large firms, small firms, university and governmental institutions), while Model 3 includes a full set of applicant dummies. Across all these models, team diversity positively affects the probability of winning the award – confirming hypothesis 1a. An increase in team diversity of 0.1 leads to an estimated increase in the probability to win the award of about 0.2 percent. To account for a curvilinear relationship, Models (4) to (6) repeat these analyses including a squared term. This squared term is significant across all specifications indicating positive marginal returns to team diversity. Control variables significantly affecting the probability of winning the award across all specifications are team size, network size, mobility and technology sourcing. On average, one extra team member increases award probability with 0.10 percent, as evidenced by the average marginal effect evaluated at actual values of the other control variables. An increase of 10 percent in network size, decreases the probability with about 0.03 percent, while having had one extra employer increases the probability with about 0.13 percent. Finally, an increase of 10 backward patent references, increases award probability with about 0.06 percent.

[5] The conclusion that multicollinearity is not a major concern, can be extended to analyzing correlations between all variables used in the analyses that follow – including standard deviation measures (not reported).

To interpret this effect and the magnitude of the coefficients, Figure 2 shows the predictive margins across 5 quantiles of the distribution of team diversity. It shows these increasing returns are not very outspoken, and rather indicates a linearly increasing effect. It seems there are no values of experience diversity at which the probability is not positively affected by increasing diversity. Moreover, it shows the effect is sizeable, as inventor teams among the top 20% of the distribution are about 3 times more likely to win the award.

To better understand the mechanisms behind how diversity affects frontier-pushing performance, I investigate the separate effects of the portion of diversity that is specific and non-specific to the focal invention. The summary statistics (table 2) show that only a small fraction of the diversity is typically used as an input into the invention (about 20%). As can be read from table 6, the diversity effect is mainly driven by diversity non-specific to the invention. Yet specific diversity seems not to harm the probability of winning the award. The marginal effect (based on model 3, not reported) of non-specific diversity is similar in size as the effect of total diversity, while that of specific diversity is small. In sum, the analyses support hypothesis 1a and indicate that diversity drives frontier-pushing performance beyond the pure increase of the components and principles available for recombination.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Team Diversity	**2.026*****	**1.674*****	**1.483****	**-1.483**	**-1.335**	**-2.054**
	(0.396)	(0.426)	(0.458)	(1.205)	(1.329)	(1.354)
Team Diversity ^ 2				**6.707****	**5.755***	**6.960****
				(2.231)	(2.451)	(2.562)
Team Level Controls	All	All	All	All	All	All
Domain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	**No**	**Yes**	**No**	**No**	**Yes**	**No**
Firm fixed effects	**No**	**No**	**Yes**	**No**	**No**	**Yes**
N	36532	36532	36532	36532	36532	36532
Pseudo R-squared	0.262	0.357	0.387	0.265	0.359	0.389

Table 5: Results from Probit regressions estimating the probability to be among the award-winners. Coefficients in bold, Standard Errors between parentheses. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

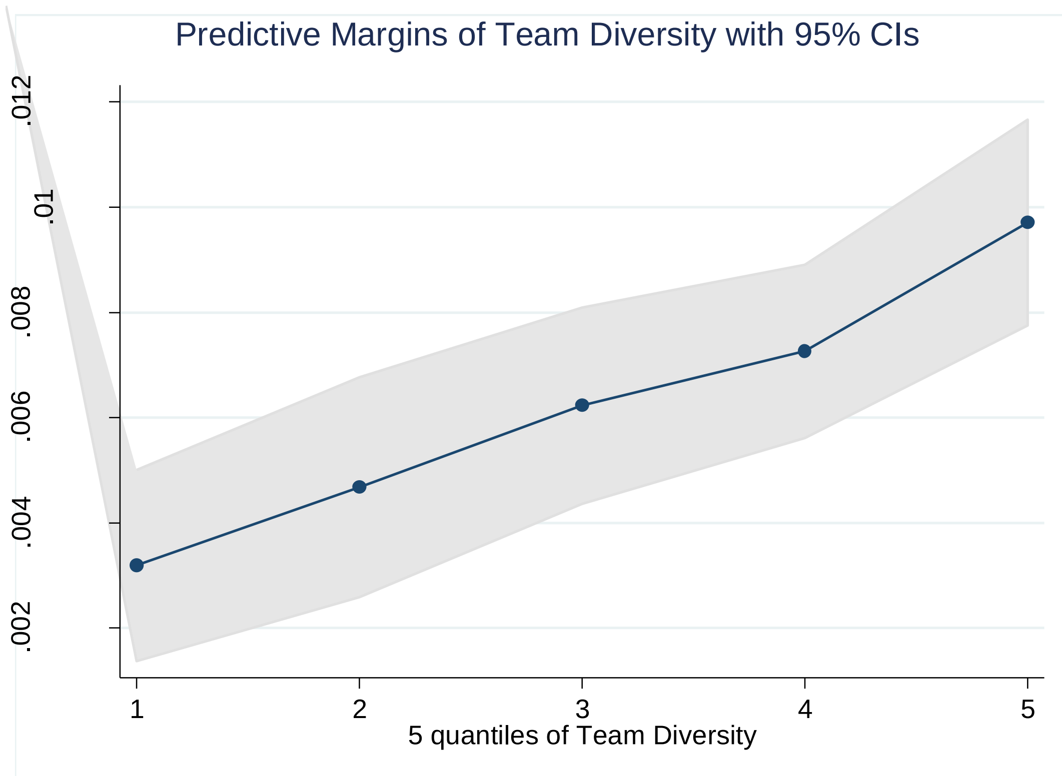

Figure 2: Predictive Margins of the effect of 5 quantiles of Team Diversity on the probability to win the award based on the model including all control variables (regression table not reported). Shaded areas indicate 95% Confidence Intervals using the delta-method.

	(1)	(2)	(3)	(4)	(5)	(6)
Specific Team Diversity	**1.406***	**0.874**	**0.998**	**-0.333**	**-1.224**	**-1.012**
	(0.682)	(0.733)	(0.760)	(1.535)	(1.681)	(1.748)
Specific Team Diversity ^ 2				**0.611**	**3.899**	**1.935**
				(5.104)	(5.479)	(5.807)
Non-Specific Team Diversity	**1.945*****	**1.576*****	**1.438****	**-1.606**	**-1.144**	**-2.211**
	(0.400)	(0.429)	(0.460)	(1.256)	(1.378)	(1.415)
Non-Specific Team Diversity ^ 2				**7.080****	**5.328***	**7.403****
				(2.390)	(2.608)	(2.739)
Team Level Controls	All	All	All	All	All	All
Domain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	**No**	**Yes**	**No**	**No**	**Yes**	**No**
Firm fixed effects	**No**	**No**	**Yes**	**No**	**No**	**Yes**
N	36532	36532	36532	36532	36532	36532
Pseudo R-squared	0.263	0.357	0.387	0.266	0.359	0.389

Table 6: Results from Probit regressions estimating the probability to be among the award-winners. Coefficients in bold, Standard Errors between parentheses. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Can Diverse Teams Substitute Diverse Individuals?

Descriptive Statistics

To answer this question, I calculated individual diversity measures for the most diverse individual, as well as the contribution of all other team members to the total team diversity. Figure 3 shows a scatterplot of total team diversity and diversity of the most diverse team member. It shows total team diversity is in large part driven by the diversity of the most diverse. The explained variance – as measured by the R-squared from regressing total team diversity on diversity of the most diverse – is almost 92%. Even if diversity obtained by a team of specialists could substitute for the same level of diversity as obtained by one team member, this situation does not often occur. This graph also graphically confirms the result of the previous section as it is clear that the award-winning teams are those with high diversity (of their most diverse individual). The remainder of the analyses zeroes in on the effect of separate contributions of the most diverse team member on the on hand, and the other team member on the other.

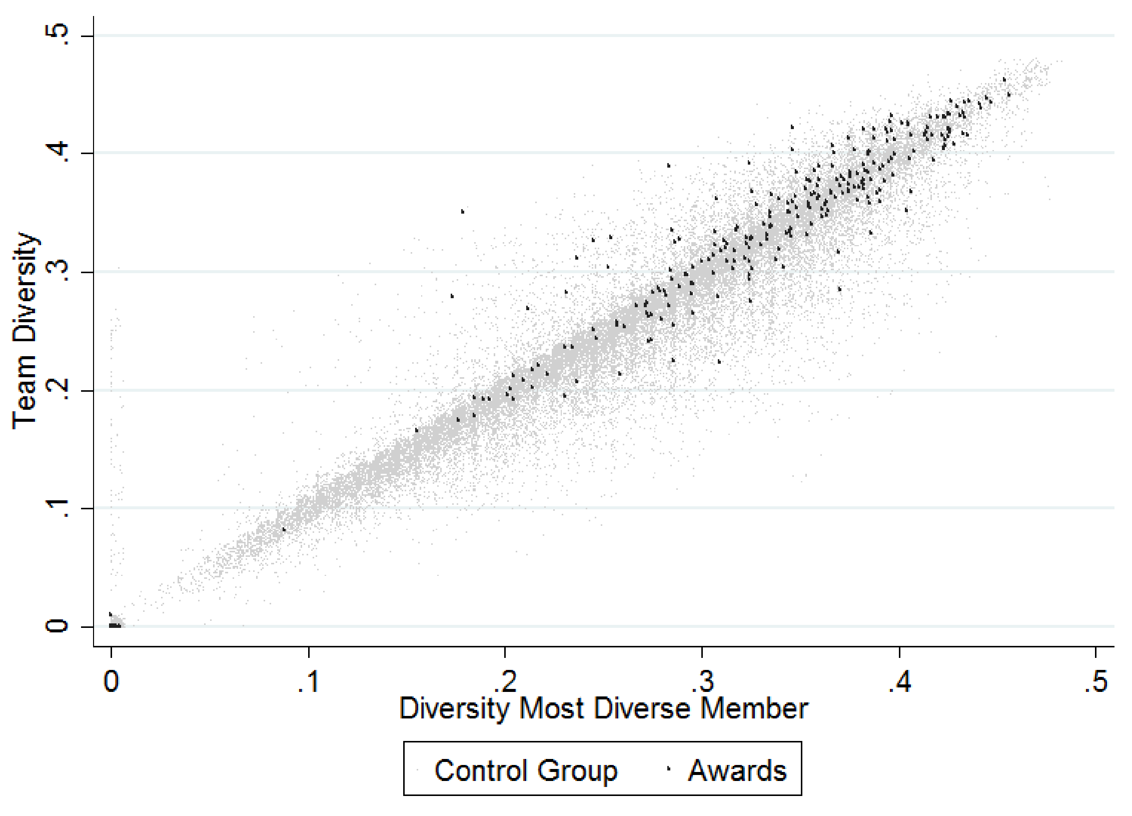

Figure 3: Scatterplot of Diversity Most Diverse Member and Team Diversity. Each dot is randomly perturbed in order to observe the mass of observations.

Table 10 provides summary results of a contingency table analyzing the relationship between winning the R&D100 award and belonging to different groups of diversity of the most diverse team member and diversity added by the team. Groups are defined by belonging to different quartiles of the diversity of the most diverse distribution on the one hand, and whether the other team members decrease, increase or not affect the degree of individual diversity of the most diverse member. These descriptive results indicate diversity of the most diverse strongly relates to the incidence of winning the award. 53% percent of the award-winning teams contain an inventor with an individual diversity score within the highest quartile, as compared to 7%, 13% and 27% for respectively the first, second and third quartile. Moreover, having other team members that add diversity to the most diverse member, seems not to make up for a lack of a diverse inventor. The probability to win the award only increases with a positive diversity contribution of team members when already a diverse member (upper two quartiles of the distribution) is present in the team. In the upper two quartiles of diversity of the most diverse, the null-hypothesis of being unrelated can be rejected (Q3: χ^2=11.30, p-value<0.01; Q4: χ^2=13.99, p-value<0.01).

Diversity Group/R&D100 Award		Observed Co-occurrence	Observed/Expected Co-occurrence	Co-occurrence as share R&D100 Awards	Co-occurrence as share of Diversity Group
Diversity Most Diverse	**Diversity Added by Team**				
Q1	Negative	3	**0.16**	1.14%	0.11%
Q1	Zero	13	**0.33**	4.92%	0.24%
Q1	Positive	2	**0.23**	0.76%	0.17%
Q2	Negative	6	**0.20**	2.27%	0.15%
Q2	Zero	21	**0.76**	7.95%	0.55%
Q2	Positive	7	**0.65**	2.65%	0.47%
Q3	Negative	18	**0.62**	6.82%	0.44%
Q3	Zero	28	**1.16**	10.61%	0.83%
Q3	Positive	26	**1.67**	9.85%	1.20%
Q4	Negative	35	**1.50**	13.26%	1.08%
Q4	Zero	52	**2.5**	19.70%	1.80%
Q4	Positive	53	**3.3**	20.08%	2.39%
Total/Average		264		100%	0.72%

Table 10: Summary contingency table between winning the award and belonging to 4x3 different groups of individual-level/team-level diversity. Groups are defined by the quartile values of diversity of the most diverse member on the one hand, negative/zero/positive values of diversity added by the other team members.

Multivariate Results

Table 11 shows the coefficients of interest to examine hypothesis 2. When looking at the estimated effects of diversity of the most diverse team member, similar results appear compared to the analyses on total team diversity. This should not surprise, given the high correlation between the two variables. Diversity added by the team only shows a significant effect for model 1, indicating that – given the diversity level of the most diverse – added diversity by team members only moderately contributes to the probability to win the award. Yet, because of the low variability of latter variable, the coefficients are not estimated very precisely. For a more intuitive view on these coefficients, I categorize all teams in exclusive categories based on two criteria. The first criterion is the quartile of the most diverse member distribution, the second criterion is whether the team contribution to total diversity is negative, zero, or positive. Table 12 reports the coefficients corresponding to the exclusive categories (where 'Quartile 1 – Negative contribution' is the baseline category). Interestingly, estimated coefficients for categories in which the team adds to the diversity of the most diverse, are only significantly higher than the baseline for the upper 2 quartiles of individual diversity of the most diverse. Hence, if added diversity by team members has a positive effect, it definitely does not substitute diversity of the most diverse – confirming hypothesis 2[6].

[6] Running the analyses on the three different diversity properties separately, reveals this result is not driven by one single diversity property. The multivariate results show no significant effect of the single properties on the probability to win the award of individual diversity of the most diverse team member, nor of diversity added by the team.

	(1)	(2)	(3)	(4)	(5)	(6)
Diversity Most Diverse	**2.022*****	**1.686*****	**1.512****	**-1.625**	**-1.442**	**-2.211**
	(0.412)	(0.442)	(0.475)	(1.219)	(1.348)	(1.372)
Diversity Most Diverse ^ 2				**7.121****	**6.090***	**7.472****
				(2.302)	(2.531)	(2.648)
Diversity Added by Team	**2.053***	**1.569**	**1.243**	**-0.125**	**-0.776**	**-1.175**
	(1.003)	(1.084)	(1.152)	(3.385)	(3.658)	(4.030)
Diversity Added by Team ^ 2				**3.733**	**4.027**	**4.316**
				(5.614)	(6.088)	(6.637)
Team Level Controls	All	All	All	All	All	All
Domain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	**No**	**Yes**	**No**	**No**	**Yes**	**No**
Firm fixed effects	**No**	**No**	**Yes**	**No**	**No**	**Yes**
N	36532	36532	36532	36532	36532	36532
Pseudo R-squared	0.262	0.357	0.387	0.265	0.359	0.389

Table 11: Results from Probit regressions estimating the probability to be among the award-winners. Coefficients in bold, Standard Errors between parentheses. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Diversity Most Diverse	Diversity Added by Team	(1)	(2)	(3)
Q1	Negative	Base Cat.	Base Cat.	Base Cat.
Q1	Zero	**0.0454**	**0.179**	**0.104**
		(0.253)	(0.303)	(0.307)
Q1	Positive	**-0.0155**	**-0.0737**	**-0.0537**
		(0.335)	(0.412)	(0.421)
Q2	Negative	**0.0234**	**0.171**	**-0.139**
		(0.271)	(0.316)	(0.344)
Q2	Zero	**0.352**	**0.498**+	**0.401**
		(0.244)	(0.292)	(0.298)
Q2	Positive	**0.231**	**0.373**	**0.228**
		(0.274)	(0.319)	(0.330)
Q3	Negative	**0.312**	**0.435**	**0.355**
		(0.247)	(0.293)	(0.299)
Q3	Zero	**0.335**	**0.424**	**0.326**
		(0.242)	(0.291)	(0.294)
Q3	Positive	**0.454**+	**0.526**+	**0.412**
		(0.245)	(0.294)	(0.298)
Q4	Negative	**0.522***	**0.630***	**0.527**+
		(0.242)	(0.290)	(0.296)
Q4	Zero	**0.600***	**0.637***	**0.516**+
		(0.240)	(0.288)	(0.294)
Q4	Positive	**0.631****	**0.680***	**0.565**+
		(0.240)	(0.289)	(0.295)
Team Level Controls		All	All	All
Domain Fixed Effects		Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes
Sector fixed effects		**No**	**Yes**	**No**
Firm fixed effects		**No**	**No**	**Yes**
N		36532	36532	36532
Pseudo R-squared		0.265	0.360	0.390

Table 12: Results from Probit regressions estimating the probability to be among the award-winners. Coefficients in bold, Standard Errors between parentheses. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Discussion

The results reveal a number of interesting patterns. First, the tension view is strongly supported: teams with a diverse knowledge base are more likely to generate frontier-pushing inventions. This result is in line with the large literature considering diverse knowledge sourcing as the wellspring of innovative activity (e.g. Trajtenberg et al., 1997; Gruber et al., 2013). However, most studies implicitly assume diversity to benefit breakthrough performance through an increased potential for recombination. This mechanism – at the level of team diversity – does not find support in the data. It seems that team diversity helps frontier-pushing outcomes beyond the pure potential for recombination, as evidenced by the finding that non-specific diversity drives the effect. Second, the perspective considering specialized knowledge to be the only way to identify anomalies (Csikszentmihalyi, 1996; Taylor & Greve, 2006) in order to find new solutions – the foundational view – is not supported. A potential reason for this, is that although specialized knowledge helps to identify anomalies, it obstructs finding high-potential, novel solutions – a second important phase in the process. In light of the finding of Kaplan & Vakili (2015) showing that local knowledge sourcing leads to cognitively novel solutions, this result is surprising. However, it might be the case that the novel ideas produced by specialist teams do not have what it takes to break paradigms. Third, diversity of the most diverse member explains most of the variance of total team diversity. This result indicates that teams that obtain diversity through a combination of non-overlapping specialized knowledge are rarely formed. Anticipation of communication barriers could explain this result. Fourth, when such teams – low diversity of the most diverse, but the team adds diversity – are formed, they seem not effective for frontier-pushing performance. This finding supports the reasoning leading to hypothesis two. In sum, these results provide support for the existence of both conditions that raise concerns about an increasing level of specialization with respect to frontier-pushing outcomes.

Yet, a number of concerns regarding the interpretation of these findings deserve attention. First, this study cannot benefit from a (quasi-) experimental design – a limitation most studies on individual inventors have in common. While the use of archival data to observe actual outcomes of actual inventors benefits external validity, this unavoidably raises a concern about internal validity. More specifically, unobserved heterogeneity in individual traits might be related to the ability to generate frontier-pushing inventions, as well as to the likelihood to build up diverse knowledge. As such, the assumption necessary to have identified a ‘treatment’ effect of knowledge diversity is that all these traits are observed through previous (patented) inventive outcomes. Arguably, this study has gone further than most previous studies in identifying and measuring individual heterogeneity that can be expected (or has been found) to affect the variables of interest. Moreover, imagining an experimental research design to answer questions about individual knowledge diversity seems daunting. Indeed, the ‘treatment’ (building up knowledge diversity) should be in effect for a long period of time as we are dealing with knowledge that is painstakingly built up over a career.

Second, as most information used in this study comes from patent documents, the results can only be extrapolated to inventive activity sensitive to patenting. While a number studies have stressed that not all inventive activity is patented, the richness and comprehensiveness of this data source seems unlikely to be obtained using other information sources (Pavitt, 1985; Griliches, 1990). With respect to the study at hand, an additional concern should be raised. It is conceivable that patent propensity differs between the award and control group. Indeed, individuals seeking recognition might have a higher patent propensity, as both the award and patents can be seen as a recognition of one’s work. This could result in an attrition bias as inventors that do not apply for patents will not end up in the sample. To mitigate this concern, robustness of the results using only inventive teams with at least 5 previous patents were tested and confirmed. The rationale is that this reduced sample only contains inventive teams of sectors where patenting is customary, decreasing the probability of leaving out non-patenting inventors.

Third, while firm-fixed effects are included to test the hypotheses, there might still be differences in strategies within firms which both affect the diversity of the personnel they employ and applying for the award. Indeed, large corporations might have business units with quite different characteristics pertaining to the nature and strategy of their R&D. To make sure this does not drive the results, the analyses were repeated without inventions from very large firms – having more than 100 patents. The results proved not qualitatively different.

Fourth, despite the advantages of using award-winning inventions to proxy for frontier-pushing inventions, a number of concerns might be raised related to this approach. The jury of editors and outside experts might be overly sensitive to technological trends, or favor multi-disciplinary approaches. To mitigate concerns related to the selection procedure, the analyses were repeated using a dependent variable relying on patent information only. As the theoretical mechanisms leading to the hypotheses build on the assumption that frontier-pushing performance is closely related to novelty, robustness of the results is confirmed using the 'Novelty in Recombination' indicator introduced in Verhoeven et al. (2016).

While recognizing these limitations, a number of considerable implications to scholarship on innovation should be highlighted. First and foremost, this study contributes to the discussion about the role of 'generalists' in inventive activity (Melero & Palomeras, 2015; Teodoridis, 2014). While an increasing burden of knowledge adds to a trend towards specialization and teamwork (Jones, 2009), the evidence presented in this paper stresses the role of a diverse knowledge base for the generation of inventions that considerably push the technological frontier. Moreover, it appears to be important that diverse knowledge is built up by individuals, rather than obtained through teamwork. These results raise concern about possible negative welfare implications of the burden of knowledge mechanism.

Second, it contributes to the arguably understudied topic of the relationship between teamwork and inventive activity. While the benefits of organizing inventive activity in teams, rather than leaving it to 'lone' inventors have been shown (Singh & Fleming, 2010), this study suggests the importance of the composition of knowledge among team members and its relationship to remarkable inventive output. It qualifies evidence suggesting teams are the ultimate vehicle to gather a diverse base of knowledge necessary for high-potential output. Indeed, to benefit from diverse knowledge in a team setting, knowledge of different individuals necessarily needs to be transferred. As such, communication barriers constrain the benefits of teams in the process of inventions. Moreover, the results support a view in which diversity adds to frontier-pushing performance beyond the pure, mechanical increase of candidate components available for recombination. It suggests an alternative mechanism in which diversity affects how inventors search as it decreases lock-in into existing paradigms. Then, teams attaining their diversity through a combination of specialists – all of which convinced of their own paradigm – might have difficulties to envision technological solutions that break with existing paradigms.

Finally, the results add to a more general stream of literature about how technological search affects breakthrough inventive performance. Arguing that inventors are key for how knowledge is sourced, this study highlights how individual knowledge is shaped by previous inventive activity. It stresses an important role of individually accumulated knowledge in directing technological search. Yet, a view in which this process acts through the simple recombination of components/principles in the knowledge base of inventors seems overly simplistic. A perspective in which inventors' search strategies – the way in which they tackle problems – are affected by the nature of their previous knowledge seems more plausible.

These findings also raise a number of questions yet to be addressed in future research. What drives individuals to build up diversity in spite of the apparent increasing effort it takes to reach the technological frontier? Which individual traits affect this decision? What role does an inventor's environment play in this decision? What kind of policy instruments could be employed to increase diversity? What determines whether diverse knowledge really leads to frontier-pushing inventions? While high diversity might positively affect the likelihood of coming up with paradigm-breaking inventions, it is far from certain such an outcome will be the result. And even if coming up with a completely new idea, there remains uncertainty about the actual appropriation of rents entailing it. Moreover, diverse knowledge could decrease the likelihood of coming up with run-off-the-mill inventions because of a lack of specialized knowledge – increasing the opportunity cost of building up diversity. This tradeoff and its determinants should be investigated in order to address the questions above. As such, it presents an avenue for future research on career choices of individual inventors.

Although many questions still remain open, these results already hold some practical implications. Given the observed trend of specializing individuals, the finding of an important role of individual knowledge diversity for frontier-pushing inventions raises policy concern about a potential undersupply of individual diversity. As such, innovation stimuli could be directed towards increasing individual-level diversity building. A policy advice, less stringently based on the assumption of a causal relationship, would be to direct existing policy instruments more towards inventors displaying diverse knowledge. Moreover, this study can inspire managers with respect to hiring strategies and human resource practices. Hiring individuals with diverse knowledge could be beneficial to exceptional inventive outcomes – even if they might lack specialized expertise in the technological areas a firm is active in. The results also suggest the beneficial effects of organizational environments that leave individual inventors free to explore technological domains they are unfamiliar with. Moreover, as a lack of deep knowledge seems not to negatively affect frontier-pushing outcomes, a situation where some individuals have a good grasp of different areas the firm is active in seems advisable. Finally, obtaining diverse knowledge using a 'potluck'-structure to obtain team diversity seems to underperform as opposed to teams where one individual – a chef de cuisine – has a diverse knowledge base.

Conclusion

This study investigated how diversity in knowledge built up throughout inventors' careers affects frontier-pushing inventive performance. In the light of a trend towards specialization, it aimed at answering two questions. First, does a diverse stock of knowledge lead to frontier-pushing inventions? Second, can diverse teams substitute diverse individuals? Results suggest the answer to the first question is 'yes'. Total team diversity is found to have a strong effect on the probability to make a frontier-pushing contribution, and the effect is strongest for high levels of diversity. The answer to the second question appears to be 'no'. While total diversity is strongly determined by the most diverse team member, the contribution to total diversity by other members has no (or a small) effect, especially for low values of diversity of the most diverse member. These findings raise concerns about a trend towards increasing specialization, as diversity is most effective to generate frontier-pushing inventions when it is built up by individual inventors. Perhaps even more importantly, it calls for further research scrutinizing these findings and aiming at a better understanding of the drivers of diversity building by inventors and knowledge workers alike.