The Role of Generalists in Team-Level Innovation

Kyu Hun Lee

November 2, 2021

Abstract

While technological innovation is becoming an increasingly team-oriented activity, the effect on innovative performance of the composition of team members, in terms of their knowledge profile characteristics, is not well known. This paper analyzes how having more generalist inventors in a research team affects the impact of team innovation. By analyzing U.S. patent data in the pharmaceutical industry, I find evidence that the proportion of generalist inventors has an overall negative effect on innovation impact, and that the effect is positively moderated by domain unfamiliarity. Generalist inventors may suffer from lack of deep expertise and provide limited contribution to team innovation, but can leverage their competence in recognizing and recombining distant knowledge components in an unfamiliar setting.

1 Introduction

It is widely acknowledged in the strategic management literature that technological innovation is a crucial source of competitive advantage (Schumpeter, 1942). As technology advances and the technological knowledge accumulates, the knowledge and skills required to innovate also increases (Jones, 2009). As a result, innovation is increasingly becoming a team-oriented activity (Agrawal et al., 2016; Jones, 2009; Singh & Fleming, 2010; Wuchty et al., 2007).

It naturally follows that team-level knowledge profile can affect team innovation performance. Since innovation is largely a process of knowledge recombination (Ahuja & Morris Lampert, 2001; Hargadon, 2002; Kaplan & Vakili, 2015), the team's innovative output will be shaped by the knowledge components the team possesses. Prior studies have examined the effect of team knowledge base characteristics on team innovative performance (Huo et al., 2019; Lee et al., 2015). For example, Huo et al. (2019) finds that technological knowledge variety of an inventor team has a positive effect on invention impact.

Often overlooked in this stream literature is the importance of member composition (Bercovitz & Feldman, 2011; Harrison & Klein, 2007; Huo et al., 2019). Most studies consider team knowledge profile as an aggregate of individual knowledge, disregarding the distribution of knowledge among team members (Huo et al., 2019). A person's knowledge not only contributes to the team's knowledge pool, but also comprise the knowledge profile of the individual, which can affect the role the person plays within the team and how they interact with other teammates(Melero & Palomeras, 2015; Teodoridis et al., 2019). Therefore, to the extent that team-level innovation depends on successful interaction between team members, it is important to examine the composition of individuals with their own knowledge profile characteristics (Somech & Drach-Zahavy, 2013).

Research on the knowledge portfolio of individuals have focused on whether an individual is a 'specialist' or a 'generalist': whether they possess deep expertise in a narrowly defined domain, or have a wide array of expertise, albeit lacking depth (Teodoridis et al., 2019). Many studies have examined the implications of being a generalist or a specialist on innovative outcomes (Boh et al., 2014; Melero & Palomeras, 2015; Nagle & Teodoridis, 2020; Rulke & Galaskiewicz, 2000; Teodoridis et al., 2019).

However, despite the importance of team-level collaboration in modern day research, most studies on generalists and specialists have been conducted on the individual level. Studies in this stream focus on the effect of an individual being a generalist or a specialist on individual innovative outcomes or career success (Boh et al., 2014; Nagle & Teodoridis, 2020; Teodoridis et al., 2019). Although some studies examine the implications of generalists and specialists on team performance, they still focus on one or two few members of the team. (Melero & Palomeras, 2015; Vakili & Kaplan, 2021). For example, Melero and Palomeras (2015) studies the role of generalists in inventor teams and their effects on innovation performance, but only examines if a generalist inventor is present in a team or not. Vakili and Kaplan (2021) studies a similar subject, and looks at the knowledge profiles of the two most experienced inventors in each team. Studies have yet to examine the effect of the member composition of the team in its entirety.

In this paper, I attempt to close this gap by examining the effect of generalist proportion of a team on its innovation performance. Acknowledging a lack of consensus on whether the effect is positive (Melero & Palomeras, 2015; Nagle & Teodoridis, 2020) or negative (Conti et al., 2014; Jones, 2009; Leahey, 2007), I examine both arguments for positive and negative implications a high proportion of generalist inventors in a team could have on the impact of the resulting innovation, and empirically test the two competing hypotheses.

Furthermore, I propose that the effect will be contingent on the familiarity of the technological domain the team innovates in. Recent findings suggest that since generalists and specialists are good at different types of tasks, the effect of generalist or specialist members on innovative performance depends on the context of the innovation (Mannucci & Yong, 2018; Teodoridis et al., 2019; Vakili & Kaplan, 2021). Following this contingency view, I test the hypothesis that the proportion of generalist inventors in a team will have a more positive effect when the team is innovating in an unfamiliar domain than in a familiar one.

I test my hypotheses using U.S. patent data in the pharmaceuticals industry. I use each patent as an innovation, and inventors listed in the patent document as team members responsible for the innovation. Generalist inventors are identified based on their prior patenting experience in different technological domains. Regression analyses show that the proportion of generalist inventors in a team has an overall negative effect on innovation impact, and that this effect is positively moderated by domain unfamiliarity. The study contributes to the literature on team-level innovation by enhancing our understanding about team member composition in terms of individual knowledge profiles. The managerial implications of this research highlights the importance of considering the context of innovation when allocating human capital in research teams.

2 Theory and hypotheses

2.1 The mixed effects of generalist inventors on team innovation impact

How does the proportion of generalist inventors in a team affect the impact of the team's innovation impact? Prior literature diverges on whether being a generalist has a positive or negative effect on innovation (Boh et al., 2014; Conti et al., 2014; Dane, 2010; Leahey, 2007; Leahey et al., 2017; Nagle & Teodoridis, 2020; Teodoridis et al., 2019). In a team, an inventor acts not only as an individual creator of knowledge, but also as a participant of within-team collaboration. Below I examine how being a generalist inventor affects the two roles.

2.1.1 Generalist inventor as an individual innovator

On one hand, generalists can be better than specialists at creating innovation by combining distant knowledge components. Knowledge recombination is a crucial step in innovation (Ahuja & Morris Lampert, 2001; Hargadon, 2002). In order to come up with novel technological knowledge, inventors should first assess the knowledge they possess, and combine them in a new way. This involves assessing the aggregate knowledge pool of the team and recognizing potentially valuable combinations of those knowledge components. Prior studies point out that generalists are better at knowledge recombination than are specialists (Melero & Palomeras, 2015; Nagle & Teodoridis, 2020). Generalists can utilize their expertise in a wide array of domains to recognize distant knowledge components and recombine them. Research also suggest that generalists can be more cognitively flexible (Audia & Goncalo, 2007; Bilalić et al., 2008; Chai, 2017; Luchins, 1942). Specialists, due to their focus on a narrow expertise, are often subject to paradigmatic rigidity (Audia & Goncalo, 2007; Chai, 2017). The 'einstellung' experiment by Luchins (1942) provides a classic example of such rigidity: After being exposed to a series of problems, all solvable in a similar manner, people tend to stick to the same solution for following problems, even when a much simpler solution exists. Toh (2014) also shows that specialists tend to be limited to their domain of expertise when identifying problems and coming up with solutions. In turn, generalists would be better positioned to evaluate various potential combinations of knowledge components, as they can employ a broader set of knowledge and perspectives.

On the other hand, the lack of deep knowledge in a technological domain may hinder the innovativeness of generalist inventors. By definition, generalists have less expertise in each domain compared to specialists in that particular domain. As a result, a generalist inventor may not have deep enough knowledge to reach the knowledge frontier in a specific technological domain (Jones, 2009). To the extent that "Expertness ... is the prerequisite to creativity" (Simon, 1985), the depth of knowledge a generalist inventor has will not be sufficient to contribute towards significant innovation (Boh et al., 2014; Conti et al., 2014; Dane, 2010). The expertise that generalists lack of includes not only the depth of knowledge on the domain, but also domain-specific problem-solving and memory skills (Larkin et al., 1980; Sweller et al., 1983). The lack of domain-specific problem-solving and memory skills makes the generalist less productive in producing and combining knowledge within the domain. In effect, generalist inventors will have lower innovative capabilities than do specialist inventors.

2.1.2 Generalist inventor as a participant in collaborative innovation

Generalists in a team of inventors can enhance teamwork by bridging inventors from various technological backgrounds (Melero & Palomeras, 2015). When team members have distant domains of expertise, those inventors will have different sets of knowledge, skills, and ways of thinking. While such diversity can enhance the creativity potential of the team, it can also act as an obstacle to collaboration. Dissimilarity in knowledge and perspectives may lead to internal conflicts (Huo et al., 2019), and inventors in different domains may use different jargons, making communication costly (Giuri et al., 2010; Laursen et al., 2005). Since generalist inventors have a broad expertise that spans many different domains, they are more likely to better understand members coming from distant technological domains. Thus, generalists can bring together inventors from dissimilar backgrounds to mitigate conflicts and reduce frictions in communication and knowledge sharing (Rulke & Galaskiewicz, 2000).

However, it is also possible that generalist inventors may not be able to collaborate well with specialist inventors. Literature on transactive memory systems (TMS, see Lewis and Herndon (2011) and Ren and Argote (2011) for a review) highlights team specialization as a dimension of TMS, and suggests that teams with high level of specialization can enhance team performance by allowing to build a deeper knowledge base, reduce redundancy, and enable members to easily locate expertise of others (Austin, 2003; Hollingshead, 2000; Wegner, 1995). In this perspective, the proportion of generalist inventors in a team will be negatively associated with the level of specialization of the team, and thus hinder access to the team's pool of knowledge. Also, prior literature suggests that specialized expertise is associated with various social capital, such as visibility (Leahey, 2007) and credibility (Faulkner et al., 1998). In other words, generalist inventors with their lack of deep expertise may be perceived as less credible by fellow inventors. In effect, generalist inventors' ideas, good or bad, may not be taken seriously by other members. This will limit the contribution generalist inventors can make to team innovation, and can also be a source of conflict.

2.1.3 Proportion of generalist inventors

To sum up, generalist inventors have both strengths and weaknesses in team-level innovation. Generalist inventors are better at knowledge recombination, and can help the team mitigate various problems that emerge during collaborative research. Accordingly, if the proportion of generalist inventors in a team is high, we can expect that the team possess strong knowledge recombination capabilities, and will have better teamwork. As a result, the innovative outcome of the team will have stronger impact.

However, generalist inventors lack deep expertise that specialist inventors have, have lower visibility in collaboration, and may negatively affect the transactive memory system of the team. If the proportion of generalist inventors in a team is high, the team will have less deep-knowledge components to start with, and team-level collaboration may be less productive. Thus, it is also possible that the innovative outcome of the team will have weaker impact. Since the overall effect of the proportion of generalist inventors in a team on team innovation impact is ambiguous, I present the following pair of competing hypotheses:

Hypothesis 1a: The proportion of generalist inventors in a team will have a positive effect on the impact of the team's innovation.

Hypothesis 1b: The proportion of generalist inventors in a team will have a negative effect on the impact of the team's innovation.

2.2 Innovation in unfamiliar domains

I propose that the effects of generalist proportion on team innovation impact, whether positive or negative, will be positively moderated by domain unfamiliarity. Innovating in an unfamiliar domain is, by definition, an exploratory process (March, 1991). It involves departing from the innovator's existing domain of expertise to acquire new knowledge in distant domains, and combining the knowledge components in a novel way. The characteristics of generalist inventors suggest that generalists will be better at exploratory innovation than they are at exploitative innovation. Generalist inventors can utilize their experience in various domains to understand and apply new knowledge from distant fields (Nagle & Teodoridis, 2020), and when they do so, they can combine the new knowledge with a broad set of knowledge components, which can lead to more breakthrough innovations (Ahuja & Morris Lampert, 2001; Fleming, 2001; Huo et al., 2019; Kaplan & Vakili, 2015; Uzzi et al., 2013). Nagle and Teodoridis (2020) finds that on the individual level, generalists have a higher propensity to engage in innovation outside their domain of expertise, and that they produce more high-impact output. I argue that on the team level, this effect scales with the number of generalist inventors in a team. Compared to specialized inventors, generalist inventors are better suited for searching and recombining distant knowledge (Nagle & Teodoridis, 2020; Toh, 2014). Therefore, all other things equal, teams with higher proportion of generalist inventors will possess higher levels of exploration and recombination capabilities. These capabilities will provide higher value to the innovative output when the team is exploring an unfamiliar domain. Thus,

Hypothesis 2: When innovating in an unfamiliar domain, the proportion of generalist inventors in a team will have a more positive effect on the innovation impact than when innovating in a familiar domain.

3 Methodology

3.1 Data and Sample

I test my hypotheses using U.S. patent data granted to global pharmaceutical firms during years 2009 to 2015. I chose the pharmaceutical industry because of a couple of reasons. First, the pharmaceutical firms have high propensity to patent their innovations, because of the importance to protect new discoveries and products (Arundel & Kabla, 1998; Fontana et al., 2013). This allows patent-related measures to be more reliable proxies for innovation. Second, innovation in the pharmaceutical industry spans various technological fields such as chemistry, biology, and computer science. This multidisciplinary nature allows for more variety in team composition and individual profile, in terms of technological domains of expertise. Having a diverse sample is beneficial as a more generalizable result can be attained, and potential bias reduced.

I use each patent as my unit of analysis, regarding each patent as an innovation, and inventors listed in the patent document as the research team responsible. Patent data was retrieved from the PatentsView database. I also used the UVA Darden Global Corporate Patent Dataset to match patent assignee data to publicly listed firms (Bena et al., 2017). I identified pharmaceutical firms as firms with Standard Industrial Classification (SIC) code 2834 - Pharmaceutical preparations. SIC code for each firm was retrieved from the Compustat database.

I also exclude patents with less than 5 team members. Since part of my theory involves not only the individual capabilities of inventors, but also how they interact with other team members, it is important that patents in the sample have enough inventors to capture that effect. If the team size is too small, the regression coefficients may show significance only because of differences in individual competence, and its implications would not be different from those of prior individual-level studies.

I ensured that all inventors in the sample had at least 5 prior patenting experience. As I explain in subsection 3.3, I identified generalist inventors based on their previous patenting experience, and inventors with too little experience could not be properly classified.

Lastly, I dropped data with missing values in any of the variables used for the regression analyses. The final sample comprised of 1,428 patents from 41 pharmaceutical firms.

3.2 Dependent variable

I measure *impact of innovation* by the number of forward citation received by a patent in 5 years from grant date. The use of forward citations as proxy for the impact and value of the innovation has been validated in previous literature (Ahuja & Morris Lampert, 2001; Hall et al., 2005; Hall et al., 2001).

3.3 Explanatory variables.

I measure *the proportion of generalist inventors* for each team based on the prior patent portfolios of each inventor in the team. For each inventor, I constructed a vector representation of the prior patent portfolio for each inventor, so that each element of the vector represents the number of prior patents assigned to the corresponding CPC subclass:

$$V_j^i = (v_{j;1}^i, \ldots, v_{j;M}^i),$$

where M is the number of unique CPC subclasses, and $v_{j;m}^i$ denotes the number of prior patent in the mth subclass inventor j had at the time of filing patent i. Following Melero and Palomeras (2015), I calculated the Herfindahl-Hirschman Index (HHI) of the portfolios to get a measure of expertise concentration for each inventor.

$$\mathrm{HHI}_j^i = \frac{\sum_{m=1}^{M} (v_{j;m}^i)^2}{(\sum_{m=1}^{M} v_{j;m}^i)^2}$$

An inventor with a high-HHI portfolio would have most patents in a few subclasses, and thus be a specialist. In contrast, a low-HHI portfolio would indicate the inventor has more evenly distributed experience in various subclasses, making them a generalist. I labeled an inventor if they had a HHI equivalent to the bottom 10% (or around 0.34) of the population or lower. For each patent, the proportion of generalist inventors was calculated as the number of generalist inventors divided by the total number of inventors listed in the patent document.

Additionally, I measured *domain unfamiliarity* as the proportion of inventors in the team who have no prior patenting experience in the CPC subclass the focal patent is in. The idea is that the more team members who are new to the subclass, the less familiar the domain.

3.4 Control variables

I include control variables related to the profile of knowledge base the team has. By controlling for these variables, the effect of member composition on innovation impact can be isolated from the change in team knowledge base it entails. I control for the *mean technological distance* between inventors in the team. Technological distance, or dissimilarity, is known to affect team innovation performance through knowledge availability and team conflict (Huo et al., 2019). Technological distance between two inventors is determined as 1 minus the cosine similarity between the prior patent portfolios of the inventors. I then calculate the average of the technological distances for all possible pairs of inventors.

$$Dist^i_{jk} = 1 - \frac{V^i_j \cdot V^i_k}{\|V^i_j\| \times \|V^i_k\|}$$

$$\text{Mean technological distance}_i = \frac{\sum_{j \neq k} Dist^i_{jk}}{N_i(N_i - 1)}$$

where N_i is the number of inventors in patent i. I also control for *prior collaboration experience*, which can affect team innovativeness, either by impeding creative process (Skilton & Dooley, 2010), or by facilitating communication and coordination (Seo et al., 2020). The variable was measured following the approach of (Reagans et al., 2005). I count the number of prior collaboration for all possible pairs of inventors within the team, and divide it by the number of pairs:

$$\text{Prior collaboration experience}_i = \frac{\sum_{j \neq k} Collab_{jk}}{N_i(N_i - 1)},$$

where $Collab_{jk}$ denotes the number of patents both inventors j and k are listed in. Additionally, I include *team size*, which is the number of inventors listed in the patent document, *team knowledge scope*, measured as the unique number of CPC subclasses the team (i.e., any of the team members) has prior patenting experience in, and *team experience*, by counting the sum of the number of prior patents each team member has.

I also control for patent-level variables. Patent information reflects the technological characteristics of the innovation, and the knowledge base it draws on. I include *the number of backward citations*, and *the average age of backward citations*. Backward citation refers to the patents the focal patent cites, and can indicate the knowledge components the citing patent combines (Fleming, 2001; Hall et al., 2001; Kaplan & Vakili, 2015). The age of backward citations serve as a proxy for the level of temporal exploration in the knowledge creation process (Nerkar, 2003). Also, I control for the *self-citation ratio* of the patent. Self-citation ratios of patents represent the level of path dependency of the innovation (Song et al., 2003; Sørensen & Stuart, 2000). *The number of claims* the focal patent has is also controlled for, as it is known to be correlated to the quality and value of the patent (Lanjouw & Schankerman, 1999).

Lastly, I control for *firm size*, measured by annual sales of the firm-year. I also add firm, year, CPC subclass dummy variables to account for any related effects.

3.5 Estimation model

I use the negative binomial regression model to obtain estimates of the effects. The negative binomial model is widely used for estimation in data with a count dependent variable with over-dispersion (Cameron & Trivedi, 2001). The negative binomial model is a nonlinear model which assumes that the dependent variable follows a generalized Poisson distribution including a gamma noise variable which accounts for over-dispersion in the data. The regression coefficients determining the mean parameter of the distribution is estimated by maximum likelihood estimation. The estimated coefficients are to be interpreted as the factor by which the expected log number of the dependent variable changes with changes in the corresponding independent variables. The use of patent forward citation counts as the dependent variable, and the over-dispersion of the dependent variable as shown in table 1, warrants using a negative binomial model for the analyses. The full model (Model 3 in table 3) is specified as follows:

$$Pr(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i},$$

where y_i is our observed dependent variable($Innovation\ impact_i$), α is the over-dispersion parameter, and μ_i is the expected mean, determined as:

$$\begin{aligned}\mu_i = exp(&\beta_0 + \beta_1 \text{Proportion of generalist inventors}_i + \beta_2 \text{Domain unfamiliarity}_i \\ &+ \beta_3 \text{Proportion of generalist inventors}_i \times \text{Domain unfamiliarity}_i \\ &+ x_i\gamma + \delta_{firm} + \delta_{subclass} + \tau_{year}),\end{aligned}$$

where x_i is the control variables, and δ_{firm}, $\delta_{subclass}$ and τ_{year} are firm, subclass and year dummies, respectively.

4 Results

Tables 1 and 2 show the descriptive statistics and correlation of the variables. There are no large correlations between the explanatory variables, the largest being between *Proportion of generalist inventors* and *Team scope* (0.584). I conducted the variance inflation factor (VIF) test to further check for multicollinearity. The highest VIF score was 3.23 (mean VIF = 1.69), showing no strong signs of multicollinearity issues within the data.

Table 3 shows the result of the negative binomial regression analyses. Model 1 include the control variables only, thus serving as a benchmark for comparison with the other models derived from my

Table 1: Descriptive statistics

Variable	Mean	Standard Deviation
Impact of innovation	20.200	77.211
Proportion of generalist inventors	0.102	0.187
Domain unfamiliarity	0.024	0.119
Mean technological distance	0.112	0.136
Prior collaboration experience	11.910	14.424
Team size	6.524	2.305
Team knowledge scope	8.597	5.376
Team experience	215.235	142.925
Number of backward citations	86.386	283.942
Average age of backward citations	4040.332	2117.958
Self-citation ratio	0.298	0.327
Number of claims	16.400	12.189
Firm size	30242.58	26670.41

theory. Model 2 tests Hypotheses 1a and 1b, which are competing predictions on the effect of the proportion of generalist inventors on innovation impact. In model 2, the coefficient of *Proportion of generalist inventors* is negative and significant ($\beta = -0.988, p$ value $= 0.039$). This suggests that the proportion of generalist inventors has a negative effect on innovation impact, providing support for Hypothesis 1b. The result indicates that as the proportion of generalist inventor increases, the log number of forward citations is estimated to decrease by a factor of 0.988. This translates to 1 standard deviation (0.187) increase in the proportion of generalist inventors decreasing the estimated number of forward citations by about 16.9%.

Model 3 tests Hypothesis 2, which examines the moderating effect of domain unfamiliarity. The coefficient of *Proportion of generalist inventors* $\times$ *Domain unfamiliarity* is positive and significant ($\beta = 8.433, p$ value < 0.001), lending support for Hypothesis 2: When domain unfamiliarity is high, the proportion of generalist inventors exerts a more positive effect on innovation impact than when domain unfamiliarity is low. The result suggests that if domain unfamiliarity is above around 0.165, the effect of the proportion of generalist inventors on innovation impact turns positive.

4.1 Robustness checks

I conducted additional analyses to check the robustness of the results. First, I performed a sensitivity test with differing sampling thresholds (see table 4,5 in appendix). I used samples with differing minimum team size (4 and 6), and with differing minimum individual patenting experience (4 and 6). Results show that Hypothesis 1b is not supported, or only weakly supported, when using alternative thresholds. Specifically, the coefficient for *proportion of generalist inventors* stays negative, but loses its significance when the minimum team size is 4, and when the minimum individual patenting experience is 4. There could be several reasons for this. One possibility is that the statistical power of the model

Table 2: Correlation matrix of variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Impact of innovation	1.000												
2. Proportion of generalist inventors	-0.064	1.000											
3. Domain unfamiliarity	0.006	0.118	1.000										
4. Mean technological distance	-0.073	0.542	0.167	1.000									
5. Prior collaboration experience	0.015	-0.076	-0.051	-0.247	1.000								
6. Team size	0.021	-0.102	-0.060	-0.068	-0.067	1.000							
7. Team knowledge scope	0.140	0.584	0.081	0.489	-0.103	0.097	1.000						
8. Team experience	0.403	-0.055	-0.052	-0.136	0.541	0.412	0.241	1.000					
9. Number of backward citations	0.741	0.004	-0.003	0.018	0.051	-0.009	0.156	0.389	1.000				
10. Average age of backward citations	0.019	0.105	0.022	0.116	-0.247	-0.096	0.086	-0.165	0.102	1.000			
11. Self-citation ratio	-0.061	-0.133	-0.066	-0.153	0.163	0.047	-0.062	0.139	-0.118	-0.519	1.000		
12. Number of claims	0.041	0.159	0.044	0.022	0.040	0.059	0.082	0.057	0.032	0.033	-0.083	1.000	
13. Firm size	0.261	0.028	0.026	-0.046	0.005	-0.068	0.014	0.004	0.258	0.132	-0.155	0.012	1.000

Table 3: Negative binomial regression results

	DV: Impact of innovation		
	Model 1	Model 2	Model 3
Proportion of generalist inventors		-0.988*	-1.395**
		(-2.06)	(-2.88)
Proportion of generalist inventors $\times$Domain unfamiliarity			8.433***
			(4.22)
Domain unfamiliarity	0.440	0.445	-0.694
	(0.88)	(0.89)	(-1.39)
Mean technological distance	-0.735	-0.479	-1.046*
	(-1.64)	(-1.03)	(-2.21)
Prior collaboration experience	-0.0165*	-0.0163*	-0.0168**
	(-2.51)	(-2.50)	(-2.63)
Team size	0.0352	0.0320	0.0331
	(1.51)	(1.38)	(1.44)
Team knowledge scope	0.0116	0.0253^{+}	0.0207
	(0.97)	(1.84)	(1.54)
Team experience	0.000662	0.000564	0.000504
	(1.19)	(1.02)	(0.92)
Number of backward citations	0.00113***	0.00112***	0.00117***
	(5.94)	(5.96)	(6.32)
Average age of backward citations	-0.000255***	-0.000255***	-0.000263***
	(-8.36)	(-8.39)	(-8.67)
Self-citation ratio	-0.949***	-0.954***	-1.027***
	(-5.24)	(-5.27)	(-5.69)
Number of claims	0.00457	0.00511	0.00575
	(1.24)	(1.38)	(1.55)
Firm size	-0.0000208	-0.0000205	-0.0000191
	(-1.58)	(-1.56)	(-1.46)
Constant	1.254	0.988	1.979
	(0.72)	(0.57)	(1.22)
Firm dummies	Yes	Yes	Yes
Subclass dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
$ln(\alpha)$			
Constant	0.588***	0.583***	0.563***
	(12.12)	(12.00)	(11.55)
Observations	1428	1428	1428

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

is limited due to too many variables. The model includes 13 explanatory and control variables, 6 year dummies, 40 firm dummies, and 41 subclass dummies. The number of firm and dummy variables increases when more relaxed sampling thresholds are applied. The more variables there are, the bigger sample size is required to consistently detect variable effects. This can be mitigated by excluding excess variables. For example, when firm dummy or subclass dummy variables are excluded from the model, consistent support for Hypothesis 1b is observed over all values of thresholds. However, it is not an ideal solution, since it cannot be determined if the effect is due to increased power, or the lack of control for significant firm or subclass effects. The result calls for further investigation using a larger sample. The coefficient for *proportion of generalist inventors* $\times$ *domain unfamiliarity* becomes insignificant when the minimum team size is 6, or when the minimum individual patenting experience is 6. This may be due to the decrease in sample size as a stricter sampling threshold is imposed. Again, a larger sample size may be required to detect significant effects.

Second, I use different cutoff values to determine if an inventor is a generalist or not (see table 6 in appendix). I identify generalist inventors as inventors whose patent portfolios have bottom 5% HHI (around 0.27) and bottom 15% HHI (around 0.39). The coefficients for *proportion of generalist inventors* lose significance in models without the interaction term with domain unfamiliarity when bottom 5% is used as the cutoff, weakening the support for Hypothesis 1b. When bottom 5% is used, the problem with the data is that out of 1,428 observations, 1,147 have generalist proportions of 0. Due to this imbalance, the model would not be able to significantly detect the effect of the proportion of generalist inventors on innovation impact. This should be further tested with a larger sample size.

Third, I exclude data with a generalist proportion of 1 from the sample (see table 7 in appendix). A proportion of 1 means that the team is comprised only of generalist inventors. These all-generalist teams may affect validity of the results, as some of the arguments outlined in the theory section may not apply to such teams. For example, when all inventors in a team are generalists, the inventors may not act as 'bridges' between inventors with distinct expertise, because all inventors may share common technological knowledge bases due to their wide breadth of expertise. Similarly, the argument that generalist lack visibility and credibility compared to specialist inventors may not be applicable to an all-generalist team. To mitigate this potential bias, I test the hypotheses after excluding all-generalist teams from the sample. The results are similar.

Lastly, I tested my hypotheses using an alternative measure of domain unfamiliarity (see table 8 in appendix). Instead of measuring domain unfamiliarity by the proportion of inventors who are new to the focal patent's CPC subclass, I used the mean difference in inventor patent portfolio HHI, before and after the focal patent. If the HHI of an inventor's patent portfolio increases after the focal patent, it means the CPC subclass the patent is in was familiar to the inventor, and if the HHI decreases, it means that the subclass was a relatively unfamiliar domain to the inventor. To illustrate, suppose there are only two domains, A and B. Inventor 1 has 5 prior patents in domain A, and 1 prior patent in domain B. The HHI of inventor 1's patent portfolio is around 0.722. Inventor 1 has more experience in domain A than with domain B. Thus, we can say that it is likely that Inventor 1 is more familiar with technological knowledge associated with domain A than knowledge associated with domain B. If inventor 1 files another patent in domain A, the HHI increases to 0.755. By patenting in a more familiar domain, inventor 1 now has a more concentrated patent portfolio, thus being more specialized in domain A. Alternatively, if inventor 1 files another patent in domain B, the HHI decreases to 0.592. Although inventor 1 has prior patenting experience in domain B, by patenting in a relatively unfamiliar domain, the patent portfolio becomes more 'balanced', and inventor 1 is one step closer to being a generalist. Using the mean difference in HHI instead of new domains allows for a broader definition of exploration, as innovating in a domain in which inventors have relatively little experience, compared to other domains of their expertise. Results show that the coefficient for *Proportion of generalist inventors* $\times$ *Mean difference in HHI* is negative and significant, meaning that the effect of the proportion of generalist inventors on innovation impact is more positive when the focal patent is 'HHI-decreasing', or in a less familiar domain to the inventors. Thus, Hypothesis 2 remains supported.

5 Discussion and conclusion

In this study, I find significant effects of the proportion of generalist inventors in a team on the impact of the resulting innovation. Results from regression analyses suggest that inventor teams with higher generalist proportion is expected to have a lower innovation impact. However, it should be noted that the significance of the effect did not hold in some cases of the sensitivity tests performed. Specifically, the estimated coefficient of *proportion of generalist inventors* stayed negative for all robustness checks, but the coefficient lost significance when different sampling criteria was used, and when generalist inventors were identified by different standards. Due to the inconsistency in the results, Hypothesis 1b ony receives weak support. Lack of statistical power due to a large number of dummy variables and imbalance in the sample is suspected to bring out this inconsistency. The validity of this study could benefit by performing power analysis, and by testing with a larger sample.

Also, the negative effect of generalist proportion on innovation impact should be interpreted with caution, especially when extending the findings outside the specific context of this research. Prior literature demonstrates that the effect of generalists on innovative performance is not straightforward (Melero & Palomeras, 2015; Teodoridis et al., 2019; Vakili & Kaplan, 2021). The effect of generalist proportion on innovative impact should be considered jointly with the characteristic of the domain and the innovation process. Future research could further our understanding on the conditions under which generalists can thrive as innovators.

I also find supporting evidence that the effect of the proportion of generalist inventors on innovation impact turns more positive when the team is innovating in an unfamiliar domain. The ability of generalist inventors to apply a broad set of knowledge and perspectives makes them better explorers. When domain unfamiliarity is high, generalists can contribute more to team innovativeness than in exploitative research. Whether this means that generalist inventors can be better compared to specialists, however, requires further investigation. Since the domain uncertainty variable in the sample is highly skewed to the right, with almost 94% of the sample having zero unfamiliarity, there are not enough data points with high levels of domain unfamiliarity to statistically test if generalist inventors outperform specialists in a highly exploratory setting. One could further examine this possibility by exploiting a research context with more instances of exploratory innovation, or by using quasi-experimental methodology such as propensity score matching.

5.1 Contributions

This study contributes to the literature on team-level innovation by examining the role of member composition, specifically in terms of member knowledge profile, on team innovation performance. Prior literature often focused on team knowledge characteristics, regarding team knowledge as the aggregate knowledge that resides within the team, but less attention has been given to the role of individual-level knowledge profiles on team-level performances (Bercovitz & Feldman, 2011; Harrison & Klein, 2007; Huo et al., 2019). The knowledge base a team can utilize depends not only on the sum of knowledge individual members possess, but also on how those members share their knowledge and interact with each other (Lewis & Herndon, 2011; Ren & Argote, 2011; Rulke & Galaskiewicz, 2000). Therefore, when examining team-level knowledge, the pattern of knowledge distribution among individual members should also be considered. This study fills this gap by examining the proportion of individual members with differing knowledge profile characteristics.

Also, this study extends the literature on generalists and specialists. Whether a person is specialized in a narrow domain, or have expertise spread over broad domains, has been known to affect the knowledge creation processes and outcomes (Boh et al., 2014; Melero & Palomeras, 2015; Nagle & Teodoridis, 2020; Rulke & Galaskiewicz, 2000; Teodoridis et al., 2019). A shortcoming in this stream of research is that most studies focus on the effect of whether a single individual is a generalist/specialist on individual-level outcomes, such as successful exploration (Nagle & Teodoridis, 2020), or career success (Boh et al., 2014). This study advances theoretical arguments that extends the effect to the team level, by acknowledging the two roles of inventors in a team: as an individual innovator, and as a participant of collaborative innovation. I examine how being a generalist inventor affects these two roles, and in turn, the contribution to team-level innovation. The findings of this study adds to our understanding of the innovative capabilities of generalists, and the conditions under which generalists can have a more positive effect on innovative outcomes.

5.2 Limitations and suggestions for future research

This study has a number of limitations. First, this study is based on the assumption that the knowledge base of an inventor is well reflected in the patenting experience of the inventor. However, it may not always be the case. Knowledge and expertise in a technological domain may not always lead to patenting experience in that specific domain, but may manifest itself in other forms, such as multidisciplinary patent classified in a different field, or scientific papers, or it may not result in any traceable record at all. Conversely, it is also possible that an inventor has patenting experience in a technological domain without any actual expertise in that domain. The inventor may have made contributions to the patent that is unrelated to the domain the patent is in. Thus, identifying inventors as generalists or specialists based on their patent records may entail some level of bias. While I attempted to mitigate this bias by choosing a research context with a heavy reliance on patenting activities, further exploration for a valid measure of individual knowledge base may be beneficial. Alternative measures of knowledge, such as scientific papers, field of formal education, and work experience could be examined.

Second, the study assumes that inventors are randomly assigned to innovation projects. However, it is more likely that the allocation of human capital to different teams would be endogenous. Managers would take into account the various traits of inventors and also the types of projects, and assign each inventor to projects they think would maximize the innovative outcome. Therefore, it is plausible that some individuals with particular characteristics would be assigned to innovation projects with higher impact potential more frequently, compared to other individuals with different characteristics. The negative effect of generalist proportion on innovation impact could be interpreted as a result of generalist inventors not being given opportunities to participate in high-value innovation projects. Future research may explore ways to address this endogeneity issue. One approach would be to construct an instrument variable which is not related to innovation impact, but can affect the generalist proportion of the team.

Third, the generalizability of the findings outside the context of the pharmaceutical industry should be examined. Vakili and Kaplan (2021) demonstrates that the ideal team configuration differs by domain, contingent on the technological characteristics of the field. This suggests that the generalist proportion of teams may have different effects on innovation impact when tested in different contexts. However, I argue that while the magnitude and direction of the effects may vary according to the technological domain, underlying mechanisms outlined in the study can be applied across different contexts. As long as we have an understanding of how and why the proportion of generalist inventors in a team exert certain influence on the innovation impact, we can leverage that understanding to infer how the effect would change in different settings. Accordingly, future research should focus on further uncovering the mechanisms through which team composition affects innovative performance.

Lastly, this study leaves the personal traits of team members, other than their technological knowledge profiles, unaccounted for. There are many individual-level factors that can affect team performance. The composition of social, biological and psychological factors, such as race, ethnicity, gender, language, social ties, and personality traits, can affect team innovation performance (**fleming**; Dahlin et al., 2005; Gruenfeld et al., 1996; Perry-Smith & Shalley, 2014). Further examination of such factors, and how they interact with the composition of different knowledge profiles, would deepen our understanding of the antecedents of successful team-level innovation.

References

Agrawal, A., Goldfarb, A., & Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics*, *8*(1), 100–128.

Ahuja, G., & Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, *22*(6-7), 521–543. https://doi.org/10.1002/smj.176

Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? empirical estimates for european firms. *Research Policy*, *27*(2), 127–141. https://doi.org/https://doi.org/10.1016/S0048-7333(98)00033-X

Audia, P. G., & Goncalo, J. A. (2007). Past success and creativity over time: A study of inventors in the hard disk drive industry. *Management Science*, *53*(1), 1–15. https://doi.org/10.1287/mnsc.1060.0593

Austin, J. R. (2003). Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology*, *88*(5), 866–878. https://doi.org/10.1037/0021-9010.88.5.866

Bena, J., Ferreira, M. A., Matos, P., & Pires, P. (2017). Are foreign investors locusts? the long-term effects of foreign institutional ownership. *Journal of Financial Economics*, *126*(1), 122–146. https://doi.org/10.1016/j.jfineco.2017.07.005

Bercovitz, J., & Feldman, M. (2011). The mechanisms of collaboration in inventive teams: Composition, social networks, and geography [Special Section on Heterogeneity and University-Industry Relations]. *Research Policy*, *40*(1), 81–93. https://doi.org/10.1016/j.respol.2010.09.008

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts—reality or myth? quantifying the einstellung effect in chess masters. *Cognitive Psychology*, *56*(2), 73–102. https://doi.org/10.1016/j.cogpsych.2007.02.001

Boh, W. F., Evaristo, R., & Ouderkirk, A. (2014). Balancing breadth and depth of expertise for innovation: A 3m story. *Research Policy*, *43*(2), 349–366. https://doi.org/10.1016/j.respol.2013.10.009

Cameron, A. C., & Trivedi, P. K. (2001). *A companion to theoretical econometrics* (B. Baltagi, Ed.). Wiley. https://books.google.co.kr/books?id=xs55E7FsMHMC

Chai, S. (2017). Near misses in the breakthrough discovery process. *Organization Science*, *28*(3), 411–428. https://doi.org/10.1287/orsc.2017.1134

Conti, R., Gambardella, A., & Mariani, M. (2014). Learning to be edison: Inventors, organizations, and breakthrough inventions. *Organization Science*, *25*(3), 833–849. https://doi.org/10.1287/orsc.2013.0875

Dahlin, K. B., Weingart, L. R., & Hinds, P. J. (2005). Team diversity and information use. *Academy of Management Journal*, *48*(6), 1107–1123. https://doi.org/10.5465/amj.2005.19573112

Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review*, *35*(4), 579–603. https://doi.org/10.5465/amr.35.4.zok579

Faulkner, W., Fleck, J., & Williams, R. (1998). *Exploring expertise: Issues and perspectives* (R. Williams, W. Faulkner, & J. Fleck, Eds.). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-13693-3

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, *47*(1), 117–132. https://doi.org/10.1287/mnsc.47.1.117.10671

Fontana, R., Nuvolari, A., Shimizu, H., & Vezzulli, A. (2013). Reassessing patent propensity: Evidence from a dataset of r&d awards, 1977–2004 [Economics, innovation and history: Perspectives in honour of Nick von Tunzelmann]. *Research Policy*, *42*(10), 1780–1792. https://doi.org/https://doi.org/10.1016/j.respol.2012.05.014

Giuri, P., Ploner, M., Rullani, F., & Torrisi, S. (2010). Skills, division of labor and performance in collective inventions: Evidence from open source software. *International Journal of Industrial Organization*, *28*(1), 54–68. https://doi.org/10.1016/j.ijindorg.2009.07.004

Gruenfeld, D. H., Mannix, E. A., Williams, K. Y., & Neale, M. A. (1996). Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational Behavior and Human Decision Processes*, *67*(1), 1–15. https://doi.org/10.1006/obhd.1996.0061

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *The RAND Journal of Economics*, *36*(1), 16–38. http://www.jstor.org/stable/1593752

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The nber patent citation data file: Lessons, insights and methodological tools* (Working Paper No. 8498). National Bureau of Economic Research. https://doi.org/10.3386/w8498

Hargadon, A. B. (2002). Brokering knowledge: Linking learning and innovation. *Research in Organizational Behavior*, *24*, 41–85. https://doi.org/10.1016/S0191-3085(02)24003-4

Harrison, D. A., & Klein, K. J. (2007). What's the difference? diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228. https://doi.org/10.5465/amr.2007.26586096

Hollingshead, A. B. (2000). Perceptions of expertise and transactive memory in work relationships. *Group Processes & Intergroup Relations*, *3*(3), 257–267. https://doi.org/10.1177/1368430200033002

Huo, D., Motohashi, K., & Gong, H. (2019). Team diversity as dissimilarity and variety in organizational innovation. *Research Policy*, *48*(6), 1564–1572. https://doi.org/10.1016/j.respol.2019.03.020

Jones, B. F. (2009). The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *The Review of Economic Studies*, *76*(1), 283–317. https://doi.org/10.1111/j.1467-937X.2008.00531.x

Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435–1457. https://doi.org/10.1002/smj.2294

Lanjouw, J. O., & Schankerman, M. (1999). *The quality of ideas: Measuring innovation with multiple indicators* (Working Paper No. 7345). National Bureau of Economic Research. https://doi.org/10.3386/w7345

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

Laursen, K., Mahnke, V., Vejrup-Hansen, P., et al. (2005). Do differences make a difference? the impact of human capital diversity, experience and compensation on firm performance in engineering consulting. *Danish Research Unit for Industrial Dynamics*.

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American Sociological Review*, *72*(4), 533–561. https://doi.org/10.1177/000312240707200403

Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research*. *Administrative Science Quarterly*, *62*(1), 105–139. https://doi.org/10.1177/0001839216665364

Lee, Y.-N., Walsh, J. P., & Wang, J. (2015). Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, *44*(3), 684–697. https://doi.org/10.1016/j.respol.2014.10.007

Lewis, K., & Herndon, B. (2011). Transactive memory systems: Current issues and future research directions. *Organization Science*, *22*(5), 1254–1265. https://doi.org/10.1287/orsc.1110.0647

Luchins, A. S. (1942). Mechanization in problem solving: The effect of einstellung. *Psychological Monographs*, *54*(6), i–95. https://doi.org/10.1037/h0093502

Mannucci, P. V., & Yong, K. (2018). The differential impact of knowledge depth and knowledge breadth on creativity over individual careers. *Academy of Management Journal*, *61*(5), 1741–1763. https://doi.org/10.5465/amj.2016.0529

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87. https://doi.org/10.1287/orsc.2.1.71

Melero, E., & Palomeras, N. (2015). The renaissance man is not dead! the role of generalists in teams of inventors. *Research Policy*, *44*(1), 154–167. https://doi.org/10.1016/j.respol.2014.07.005

Nagle, F., & Teodoridis, F. (2020). Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration. *Strategic Management Journal*, *41*(1), 55–85. https://doi.org/10.1002/smj.3091

Nerkar, A. (2003). Old is gold? the value of temporal exploration in the creation of new knowledge. *Management Science*, *49*(2), 211–229. https://doi.org/10.1287/mnsc.49.2.211.12747

Perry-Smith, J. E., & Shalley, C. E. (2014). A social composition view of team creativity: The role of member nationality-heterogeneous ties outside of the team. *Organization Science*, *25*(5), 1434–1452. https://doi.org/10.1287/orsc.2014.0912

Reagans, R., Argote, L., & Brooks, D. (2005). Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science*, *51*(6), 869–881. https://doi.org/10.1287/mnsc.1050.0366

Ren, Y., & Argote, L. (2011). Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals*, *5*(1), 189–229. https://doi.org/10.5465/19416520.2011.590300

Rulke, D. L., & Galaskiewicz, J. (2000). Distribution of knowledge, group network structure, and group performance. *Management Science*, *46*(5), 612–625. https://doi.org/10.1287/mnsc.46.5.612.12052

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. Harper & Row.

Seo, E., Kang, H., & Song, J. (2020). Blending talents for innovation: Team composition for cross-border r&d collaboration within multinational corporations. *Journal of International Business Studies*, *51*(5), 851–885. https://doi.org/10.1057/s41267-020-00331-z

Simon, H. A. (1985). What we know about the creative process. *Frontiers in creative and innovative management*, *4*, 3–22.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management Science*, *56*(1), 41–56. https://doi.org/10.1287/mnsc.1090.1072

Skilton, P. F., & Dooley, K. J. (2010). The effects of repeat collaboration on creative abrasion. *Academy of Management Review*, *35*(1), 118–134. https://doi.org/10.5465/amr.35.1.zok118

Somech, A., & Drach-Zahavy, A. (2013). Translating team creativity to innovation implementation: The role of team composition and climate for innovation. *Journal of Management*, *39*(3), 684–708. https://doi.org/10.1177/0149206310394187

Song, J., Almeida, P., & Wu, G. (2003). Learning–by–hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science*, *49*(4), 351–365. https://doi.org/10.1287/mnsc.49.4.351.14429

Sørensen, J. B., & Stuart, T. E. (2000). Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly*, *45*(1), 81–112. https://doi.org/10.2307/2666980

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, *112*(4), 639–661. https://doi.org/10.1037/0096-3445.112.4.639

Teodoridis, F., Bikard, M., & Vakili, K. (2019). Creativity at the knowledge frontier: The impact of specialization in fast- and slow-paced domains. *Administrative Science Quarterly*, *64*(4), 894–927. https://doi.org/10.1177/0001839218793384

Toh, P. K. (2014). Chicken, or the egg, or both? the interrelationship between a firm's inventor specialization and scope of technologies. *Strategic Management Journal*, *35*(5), 723–738. https://doi.org/10.1002/smj.2128

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472. https://doi.org/10.1126/science.1240474

Vakili, K., & Kaplan, S. (2021). Organizing for innovation: A contingency view on innovative team configuration. *Strategic Management Journal*, *42*(6), 1159–1183. https://doi.org/10.1002/smj.3264

Wegner, D. M. (1995). A computer network model of human transactive memory. *Social Cognition*, *13*(3), 319–339. https://doi.org/10.1521/soco.1995.13.3.319

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039. https://doi.org/10.1126/science.1136099

Table 4: Sensitivity test for minimum team size

	Min. team size 4			Min. team size 6		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Proportion of generalist inventors		-0.316	-0.466		-1.218^{+}	-1.366^{+}
		(-1.05)	(-1.52)		(-1.68)	(-1.86)
Proportion of generalist inventors ×Domain unfamiliarity			2.260*			3.809
			(2.01)			(1.02)
Domain unfamiliarity	0.799*	0.787*	0.331	0.0168	0.0116	-0.421
	(2.46)	(2.43)	(0.89)	(0.03)	(0.02)	(-0.63)
Mean technological distance	-0.877**	-0.785*	-0.932**	-0.921	-0.649	-0.625
	(-2.73)	(-2.35)	(-2.75)	(-1.50)	(-1.01)	(-0.98)
Prior collaboration experience	-0.00595	-0.00549	-0.00600	-0.00933	-0.0103	-0.0101
	(-1.04)	(-0.96)	(-1.05)	(-1.23)	(-1.35)	(-1.33)
Team size	0.00207	0.000763	-0.00153	0.0207	0.0161	0.0171
	(0.10)	(0.04)	(-0.07)	(0.77)	(0.60)	(0.64)
Team knowledge scope	0.0170^{+}	0.0221*	0.0219*	-0.00633	0.00834	0.00833
	(1.77)	(2.05)	(2.05)	(-0.36)	(0.42)	(0.42)
Team experience	0.00122*	0.00117*	0.00119*	0.000324	0.000300	0.000272
	(2.37)	(2.27)	(2.32)	(0.48)	(0.45)	(0.41)
Number of backward citations	0.00137***	0.00137***	0.00137***	0.000783***	0.000763**	0.000774***
	(7.03)	(7.04)	(7.05)	(3.34)	(3.28)	(3.32)
Average age of backward citations	-0.000156***	-0.000156***	-0.000159***	-0.000227***	-0.000229***	-0.000227***
	(-6.46)	(-6.48)	(-6.57)	(-5.48)	(-5.52)	(-5.48)
Self-citation ratio	-0.741***	-0.750***	-0.766***	-1.101***	-1.101***	-1.116***
	(-4.86)	(-4.91)	(-5.01)	(-4.85)	(-4.85)	(-4.92)
Number of claims	0.00798*	0.00812*	0.00824*	0.00396	0.00454	0.00459
	(2.35)	(2.39)	(2.42)	(0.86)	(0.98)	(0.99)
Firm size	-0.0000281*	-0.0000280*	-0.0000288**	-0.0000422**	-0.0000413**	-0.0000408**
	(-2.54)	(-2.54)	(-2.60)	(-2.67)	(-2.62)	(-2.59)
Constant	-2.146	-1.987	-1.756	3.488*	3.166^{+}	3.243^{+}
	(-0.98)	(-0.90)	(-0.81)	(2.02)	(1.83)	(1.87)
Firm dummies	Yes	Yes	Yes			
Subclass dummies	Yes	Yes	Yes			
Year dummies	Yes	Yes	Yes			
lnalpha						
Constant	0.778***	0.777***	0.774***	0.520***	0.515***	0.512***
	(20.31)	(20.26)	(20.17)	(8.30)	(8.19)	(8.15)
Observations	2229	2229	2229	867	867	867

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Sesitivity test for minimum number of prior patents

	Min. prior patent 4			Min. prior patent 6		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Proportion of generalist inventors		-0.537	-0.740^{+}		-1.537**	-1.627**
		(-1.28)	(-1.76)		(-2.98)	(-3.12)
Proportion of generalist inventors ×Domain unfamiliarity			5.322**			4.455
			(2.97)			(1.42)
Domain unfamiliarity	0.130	0.108	-0.556	-0.272	-0.291	-0.727
	(0.32)	(0.27)	(-1.28)	(-0.54)	(-0.59)	(-1.31)
Mean technological distance	-0.423	-0.264	-0.566	-1.173*	-0.801	-0.919^{+}
	(-1.11)	(-0.65)	(-1.39)	(-2.22)	(-1.48)	(-1.67)
Prior collaboration experience	-0.0133*	-0.0128*	-0.0130*	-0.0179**	-0.0179**	-0.0178**
	(-2.09)	(-2.02)	(-2.07)	(-2.73)	(-2.75)	(-2.74)
Team size	0.0190	0.0165	0.0200	0.0401	0.0366	0.0367
	(0.91)	(0.79)	(0.97)	(1.50)	(1.38)	(1.39)
Team knowledge scope	0.0124	0.0192	0.0155	0.0105	0.0315*	0.0304*
	(1.20)	(1.64)	(1.33)	(0.82)	(2.15)	(2.08)
Team experience	0.000898^{+}	0.000847	0.000824	0.000421	0.000285	0.000263
	(1.68)	(1.59)	(1.56)	(0.74)	(0.50)	(0.47)
Number of backward citations	0.00117***	0.00117***	0.00120***	0.00113***	0.00112***	0.00113***
	(5.86)	(5.86)	(6.06)	(6.16)	(6.21)	(6.23)
Average age of backward citations	-0.000243***	-0.000244***	-0.000249***	-0.000268***	-0.000272***	-0.000271***
	(-8.74)	(-8.78)	(-8.96)	(-8.06)	(-8.19)	(-8.18)
Self-citation ratio	-0.967***	-0.976***	-1.015***	-0.984***	-0.989***	-1.014***
	(-5.74)	(-5.78)	(-6.02)	(-5.07)	(-5.10)	(-5.22)
Number of claims	0.00564	0.00600^{+}	0.00630^{+}	0.00592	0.00667^{+}	0.00675^{+}
	(1.63)	(1.73)	(1.82)	(1.47)	(1.66)	(1.67)
Firm size	-0.0000372**	-0.0000372**	-0.0000377**	-0.0000177	-0.0000186	-0.0000181
	(-3.07)	(-3.08)	(-3.13)	(-1.24)	(-1.31)	(-1.28)
Constant	3.317**	3.022**	3.484**	2.745	2.373	2.505
	(2.90)	(2.60)	(2.99)	(1.60)	(1.41)	(1.52)
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes
Subclass dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
lnalpha						
Constant	0.676***	0.674***	0.665***	0.514***	0.502***	0.501***
	(15.48)	(15.44)	(15.20)	(9.68)	(9.44)	(9.41)
Observations	1766	1766	1766	1187	1187	1187

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Test with different HHI thresholds

	Bottom 5% HHI		Bottom 15% HHI	
	Model 2	Model 3	Model 2	Model 3
Proportion of generalist inventors	-0.717	-1.445*	-0.590^{+}	-0.655*
	(-1.04)	(-2.02)	(-1.80)	(-2.02)
Proportion of generalist inventors ×Domain unfamiliarity		10.49***		6.533***
		(3.78)		(3.52)
Domain unfamiliarity	0.468	-0.442	0.363	-0.740
	(0.94)	(-0.89)	(0.73)	(-1.41)
Mean technological distance	-0.600	-1.054*	-0.583	-1.129*
	(-1.28)	(-2.24)	(-1.27)	(-2.38)
Prior collaboration experience	-0.0163*	-0.0167**	-0.0172**	-0.0175**
	(-2.49)	(-2.58)	(-2.64)	(-2.72)
Team size	0.0333	0.0340	0.0329	0.0345
	(1.43)	(1.47)	(1.41)	(1.49)
Team knowledge scope	0.0185	0.0168	0.0222^{+}	0.0162
	(1.35)	(1.24)	(1.66)	(1.22)
Team experience	0.000607	0.000519	0.000567	0.000520
	(1.09)	(0.95)	(1.02)	(0.94)
Number of backward citations	0.00114***	0.00120***	0.00113***	0.00117***
	(6.00)	(6.41)	(5.99)	(6.25)
Average age of backward citations	-0.000254***	-0.000260***	-0.000256***	-0.000259***
	(-8.33)	(-8.55)	(-8.40)	(-8.54)
Self-citation ratio	-0.948***	-1.005***	-0.967***	-1.011***
	(-5.23)	(-5.57)	(-5.33)	(-5.59)
Number of claims	0.00489	0.00570	0.00496	0.00526
	(1.32)	(1.53)	(1.34)	(1.42)
Firm size	-0.0000206	-0.0000194	-0.0000212	-0.0000202
	(-1.57)	(-1.49)	(-1.61)	(-1.54)
Constant	1.080	1.920	1.159	2.006
	(0.62)	(1.16)	(0.67)	(1.22)
Firm dummies	Yes	Yes	Yes	Yes
Subclass dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
lnalpha				
Constant	0.586***	0.570***	0.584***	0.572***
	(12.08)	(11.70)	(12.02)	(11.75)
Observations	1428	1428	1428	1428

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Test with generalist proportion lower than 1

	DV: Impact of innovation		
	Model 1	Model 2	Model 3
Proportion of generalist inventors		-1.301**	-1.770***
		(-2.63)	(-3.58)
Proportion of generalist inventors ×Domain unfamiliarity			8.910***
			(4.39)
Domain unfamiliarity	0.452	0.469	-0.738
	(0.91)	(0.94)	(-1.48)
Mean technological distance	-0.708	-0.336	-0.911+
	(-1.57)	(-0.71)	(-1.91)
Prior collaboration experience	-0.0170**	-0.0171**	-0.0177**
	(-2.59)	(-2.63)	(-2.78)
Team size	0.0349	0.0305	0.0313
	(1.50)	(1.31)	(1.36)
Team knowledge scope	0.0112	0.0287*	0.0245+
	(0.93)	(2.08)	(1.82)
Team experience	0.000698	0.000591	0.000532
	(1.25)	(1.06)	(0.98)
Number of backward citations	0.00113***	0.00112***	0.00117***
	(5.95)	(5.99)	(6.37)
Average age of backward citations	-0.000250***	-0.000250***	-0.000258***
	(-8.20)	(-8.21)	(-8.54)
Self-citation ratio	-0.931***	-0.933***	-1.010***
	(-5.13)	(-5.15)	(-5.60)
Number of claims	0.00421	0.00488	0.00566
	(1.14)	(1.32)	(1.53)
Firm size	-0.0000207	-0.0000202	-0.0000186
	(-1.57)	(-1.53)	(-1.42)
Constant	1.273	0.926	1.945
	(0.73)	(0.54)	(1.20)
Firm dummies	Yes	Yes	Yes
Subclass dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
lnalpha			
Constant	0.592***	0.584***	0.561***
	(12.17)	(11.97)	(11.45)
Observations	1418	1418	1418

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Alternative measure for domain unfamiliarity

	DV: Impact of innovation		
	Model 1	Model 2	Model 3
Proportion of generalist inventors		-1.002*	-1.000*
		(-2.09)	(-2.10)
Proportion of generalist inventors $\times$Mean difference in HHI			-41.55^{+}
			(-1.83)
Mean difference in HHI	6.394*	6.508*	8.442**
	(2.04)	(2.08)	(2.60)
Mean technological distance	-0.590	-0.327	-0.462
	(-1.35)	(-0.72)	(-1.00)
Prior collaboration experience	-0.0187**	-0.0185**	-0.0181**
	(-2.83)	(-2.83)	(-2.76)
Team size	0.0305	0.0272	0.0262
	(1.31)	(1.17)	(1.13)
Team knowledge scope	0.0116	0.0255^{+}	0.0251^{+}
	(0.98)	(1.87)	(1.85)
Team experience	0.000725	0.000627	0.000606
	(1.30)	(1.12)	(1.09)
Number of backward citations	0.00112***	0.00112***	0.00111***
	(5.91)	(5.93)	(5.90)
Average age of backward citations	-0.000249***	-0.000250***	-0.000248***
	(-8.26)	(-8.28)	(-8.24)
Self-citation ratio	-0.907***	-0.912***	-0.930***
	(-5.01)	(-5.04)	(-5.13)
Number of claims	0.00536	0.00592	0.00582
	(1.45)	(1.59)	(1.57)
Firm size	-0.0000211	-0.0000210	-0.0000206
	(-1.61)	(-1.60)	(-1.57)
Constant	1.350	1.095	1.126
	(0.83)	(0.68)	(0.71)
Firm dummies	Yes	Yes	Yes
Subclass dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
lnalpha			
Constant	0.583***	0.578***	0.574***
	(12.00)	(11.88)	(11.79)
Observations	1428	1428	1428

t statistics in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$