Shared Culture and Technological Innovation: Evidence from Corporate R&D Teams

Tristan Fitzgerald
Mays Business School
Texas A&M University
tfitzgerald@mays.tamu.edu

Xiaoding Liu
Mays Business School
Texas A&M University
xliu@mays.tamu.edu

May 2024

ABSTRACT

We open the black box of corporate innovation production by examining its most important input: the employees tasked with creating new inventions. Using a novel within-firm research design involving the universe of U.S. corporate inventors across four decades, we find that inventors' shared cultural values are a critical driver of inventor team formation. Moreover, using premature co-inventor deaths as exogenous shocks to team composition, we document both positive and negative impacts of inventor team cultural diversity on team patent production. Less culturally diverse teams produce a higher overall quantity of patents that tend to exploit existing technologies, while more culturally diverse teams produce more risky, exploratory patents with a greater potential for high-impact innovations. Exploring the underlying mechanisms, we present evidence that culturally diverse teams tend to seek new knowledge from more heterogeneous and non-traditional input information sources, but they also face greater knowledge integration challenges. Overall, our results present a more nuanced perspective on diversity, revealing that it does not lead to uniformly better or worse outcomes, but instead impacts the type of R&D output.

Keywords: Inventor Teams; Innovation; Diversity; Culture; R&D; Labor Productivity

JEL Classification: G18; J24; M14; M54; O31; O34

We would like to thank Cesare Fracassi, Danielle Li, Francesco Lissoni, Maria-Teresa Marchica, Yihui Pan, Rui Silva, and Ting Xu (discussants), Lauren Cohen, Lee Fleming, Jarrad Harford, Gerard Hoberg, Josh Lerner, Kai Li, Gordon Phillips, and John Van Reenen, as well as all participants at the American Finance Association (AFA) Annual Meeting, Western Finance Association (WFA) Annual Meeting, American Economic Association (AEA) Annual Meeting, Financial Intermediation Research Society (FIRS) Conference, Society of Labor Economists (SOLE) Annual Meeting, Midwest Finance Association (MFA) Annual Meeting, Edinburgh Corporate Finance Conference, Lone Star Finance Symposium, EPFL Innovation Seminar, University of Massachusetts, and Texas A&M University for their invaluable feedback.

I. INTRODUCTION

Technological innovation plays a pivotal role in the real economy, serving as the primary driver for sustained economic growth over the long run (Schumpeter, 1942; Romer, 1990). Amid the increasing complexities of technological advancement, the rising importance of teamwork in spurring new inventions has become one of the defining features of the modern knowledge economy (Jones, 2009; Lucas and Moll, 2014), where over 80% of patents granted to U.S. public firms in recent years have been produced by teams of two or more inventors. Concurrent with this trend, the topic of diversity in the workplace has attracted significant public attention, with private companies, investors, and regulators making a conscious effort to enact policies aimed at increasing workplace diversity across various dimensions.[1] Enacting such policies has potentially far-reaching social and economic implications given the large and growing contribution of human capital to firm value (Zingales, 2000).

In this paper, we investigate what the intersection of these trends, namely increased teamwork and workplace diversity, means for inventor team formation and team-level labor productivity. To do so, we direct our attention to the innovation process itself and the employees tasked with creating new inventions. The first key part of the innovation process that has been the focus of prior studies is the initial set of technical knowledge, skills, and experiences that each inventor brings to the R&D team (Fleming, 2001). However, there is no guarantee that these initial knowledge inputs can be effectively integrated into useful new technologies. Therefore, the second critical, yet often overlooked, part of the innovation process is how these disparate knowledge inputs are combined to form a new joint invention (Schumpeter, 1934, 1939), which we term as the 'knowledge integration' process. Together, the set of raw knowledge inputs and the integration process governing their relationship create new

[1] For example, several government regulations require firms to take affirmative actions to ensure that equal employment opportunity is provided to all current and prospective employees (e.g., *U.S. Presidential Executive Order 11246*) or enforce explicit diversity quotas (e.g., Norway's *Public Limited Liability Companies Act (1997)*). With respect to corporate actions, 74% of S&P 500 companies have established a dedicated Chief Diversity Officer position or equivalent, with many of these appointments occurring over the last few years (Paikeday, Qosja, Lim, and Flock, 2023).

insights that shape the resulting inventions (Xiao, Makhija, and Karim, 2022). Crucially, even though two R&D teams may possess identical sets of initial knowledge, skills, and experiences, these teams can produce divergent innovation outputs if they adopt distinct thinking styles and problem-solving approaches to the process of integrating their respective knowledge. Unlike knowledge and skills that can be acquired or modified relatively easily, how these raw knowledge inputs are combined during the 'knowledge integration' process depends on inventors' core values, beliefs, and perspectives that are much more ingrained and harder to change (Van den Steen, 2010).[2]

Therefore, we hypothesize that an important aspect of diversity in this context is diversity in the core values, beliefs, and perspectives of R&D team members, which we seek to capture using an exogenous measure based on inherited culture. Culture is defined as the shared values, beliefs, and preferences that characterize a particular group of people, which remain fairly unchanged from generation to generation (Hofstede, 1980; Guiso, Sapienza, and Zingales, 2006). We focus on the inherited, slow moving component of culture that is endowed at birth and is exogenous to personal and career experiences acquired later in life.[3] Unlike other inherited traits, culture is a deeper level construct that is directly related to an individual's core values and beliefs, and thus more likely to have a stronger impact on the creative process (Stahl, Maznevski, Voigt, and Jonsen, 2010).[4] Therefore, cultural diversity within inventor teams captures inherent differences in team members' values, beliefs, and preferences, where such differences can foster a variety of perspectives, thinking styles, and problem-solving approaches within the two-part innovation process. In other words,

[2] Core values, beliefs, and perspectives serve as filters through which people interpret information and make decisions. Therefore, differences in core values, beliefs, and perspectives will result in varying intuitions, belief systems, or mental models, causing individuals presented with identical data to draw divergent conclusions. The pivotal role of belief systems in organizations has been widely studied in several managerial contexts (e.g., Donaldson and Lorsch, 1983; Schein 1985).
[3] Although culture is conceptually antecedent to knowledge, skills, and experiences that inventors acquire later in life, we still control for a rich set of individual and team-level measures of knowledge, skills, and experiences in all our empirical analyses to further distinguish the impact of culture from these raw knowledge inputs.
[4] This contrasts with other inherited traits such as gender and age, which are not as directly related to an individual's core values and beliefs (Stahl et al., 2010). Nevertheless, in later robustness tests, we separately control for the potential impact of gender diversity and age diversity on team innovation output.

distinct from the initial set of knowledge, skills and experiences of each team member that represents *what* the team knows, cultural diversity relates to differences in *how* these team members process, communicate, and apply their technical knowledge to jointly tackle novel technological problems.

The overall net impact of diversity in the cultural values of team members on the knowledge integration process and the resulting team R&D output is both theoretically and empirically unclear. Individuals with different cultural values will tend to approach the problem of synthesizing their respective initial knowledge inputs into novel technological outputs from very distinct perspectives (Gorton and Zentefis, 2023). On the one hand, greater differences in the starting perspectives and cognitive thought processes of more culturally diverse team members will increase the likelihood of intra-team conflicts and miscommunications, as shown in the theoretical models of Van den Steen (2010) and Garlappi, Giammarino, and Lazrak (2017) that study the effect of 'culture clashes' and heterogenous beliefs in corporate investment settings. These differences may impede the ability and/or the efficiency with which disparate knowledge inputs are combined within the team's knowledge integration process to produce new joint inventions. In contrast, these models also suggest that teams sharing similar beliefs and perspectives are unlikely to encounter significant coordination and communication issues, even if they possess divergent initial information sets.

On the other hand, since more culturally diverse teams will more frequently face the challenge of trying to reconcile more diverse perspectives and problem-solving approaches, this will lead them to exhibit an increased willingness to experiment with alternative methods for resolving such disagreements (Donaldson, Malenko, and Piacentino, 2020). For example, they may seek out a broader range of information sources outside of their existing collective knowledge base that, when integrated with their initial knowledge set, can help to produce synthesized technological solutions (Van den Steen, 2010). This process may in turn lead to culturally diverse teams developing more creative and higher quality solutions compared to more culturally homogenous teams.

To explore the broader question of how team cultural diversity affects team productivity, we focus on the unique and economically important context of innovation production within for-profit firms where over $1 trillion is spent globally on R&D each year and where both creativity *and* efficiency are deemed essential. Specifically, we conduct a large-scale study involving approximately 700,000 U.S.-based employee inventors working at more than 5,400 U.S. public firms over 40 years to analyze what factors drive inventors' collaboration decisions and how team cultural diversity affects the team innovation process, especially the team's knowledge integration activities, and their joint outputs. Given that the costs and benefits of cultural diversity may vary across different contexts, understanding the impact of cultural diversity within inventor teams on the creation of commercially viable corporate patents presents an important empirical question that has yet to be explored.[5]

Given the relevance of these questions to both firms and policymakers, there is a considerable body of research on related topics, although it often yields conflicting evidence. With respect to team formation, some researchers find evidence of homophily, namely the tendency of individuals to prefer to work with others who share similar personal characteristics, in other business-related contexts.[6] With respect to diversity and performance, there is much conflicting empirical evidence in the prior literature, for example on the relationship between board diversity and firm economic output[7] and, more relevantly, on the relationship between firm-wide employee diversity and innovation (e.g., Ostergaard, Timmermans, and Kristinsson, 2011; Doran, Gelber, and Isen, 2022).[8]

[5] Our paper differs from studies of culture and creativity in a broader sense in the management literature (e.g., Wang, Cheng, Chen, and Leung, 2019). These studies use small sample sizes, rely mostly on surveys and experiments involving students, and examine laboratory outcomes such as business plan write-ups. In contrast, we conduct a large-scale study involving all inventors across the entire universe of patenting U.S. public firms spanning 40 years. We also use rigorous empirical strategies to address endogeneity concerns and examine commercially viable firm inventions, an important real economic outcome that involves the allocation of significant resources and talent amounting to over $1 trillion each year.
[6] For example, see Ishii and Xuan (2014), Gompers, Mukharlyamov, and Xuan (2016) as well as Calder-Wang, Gompers, and Huang (2023) in the context of directors and executives, venture capitalists, and firm founding teams, respectively.
[7] For example, see Bernile, Bhagwat, and Yonker (2018) and Griffin, Li, and Xu (2021) for studies finding a positive impact and see Adams and Ferreira (2009) and Ahern and Dittmar (2012) for studies finding a negative impact.
[8] Ostergaard et al. (2011) suggests that aggregate firm-wide diversity in employees' inherited traits, calculated across the firm's entire labor force, can promote overall firm innovation, while other studies (e.g., Doran et al. 2022) do not find any

Our study addresses an important gap in the existing literature, which has very limited research focusing on corporate inventor teams. Unlike the conflicting prior literature primarily analyzing firm-level relationships, we focus our analysis on understanding the formation and productivity of distinct inventor teams working *within the same firm* at the *same point in time*.[9] This level of examination offers a unique perspective on the inner workings of corporate innovation production by allowing us to create more granular sets of both input and output variables as well as uncover novel aspects of the team knowledge integration process. Specifically, we develop a detailed set of acquired characteristics and inherited traits for all individual inventors and inventor teams in U.S. public firms. In addition, we create team-level performance metrics, including the quantity, quality, distribution, and type of patent output, which offer richer insights compared to aggregate firm-level output. Moreover, we also use detailed patent data to quantify the initial knowledge that each R&D team member brings to the collaboration *and* the new information acquired by R&D teams during the knowledge integration process. After conditioning on the team's starting amount of knowledge, skills, and experiences as well as controlling for diversity in gender, age, and immigrant status of team members and all unobserved, time-varying characteristics at the management, board of directors and firm levels, our study can better isolate the relative importance of diversity in cultural values in driving the formation, knowledge integration activities, and overall productivity of corporate R&D teams.

To study the impact of cultural diversity within inventor teams on team innovation, we combine and extend over two dozen different data sources to construct a unique database encompassing all U.S.-based patenting inventors and their associated inventor teams working at publicly listed U.S.

---

positive effect of such firm-wide employee diversity on corporate innovation. Importantly, these studies measure diversity and innovation performance at the firm level, making it impossible to link diversity among the individuals that are the most critical and direct inputs into the innovation process (i.e., corporate inventors) to their specific innovative outputs. For example, Ostergaard et al. (2011) relies upon one 2006 survey of 1,600 Danish firms that the authors acknowledge "does not identify the persons who interact with each other or who are involved in the specific innovation process".

[9] Thus, we ensure that pairs of treated and control teams have similar access to physical and financial resources (by being at the same firm) and are exposed to a similar technological and competitive landscape (by working at the same time).

firms from 1981 to 2016. To construct our key measures of inherited cultural values, we introduce a novel multi-layered approach that uses restricted full-count decennial U.S. census data from 1850 to 1940 as well as commercial and public datasets. To the best of our knowledge, we are the first paper to undertake the task of identifying individuals' countries of ancestry by tracing their family trees. Subsequently, we link inventors' countries of ancestry to the well-established six-dimensional cultural framework of Hofstede (1980) and Hofstede, Hofstede, and Minkov (2010) to directly capture the multi-dimensional aspects of inventors' inherited cultural values and beliefs.

Using a sample of over 1.8 million first-time collaborations formed at U.S. public firms between pairs of U.S.-based inventors, we provide new evidence that shared cultural values play a critical role in inventor team formation. In particular, two similarly experienced inventors currently working at the same division within the same firm in the same office are 30% more likely to collaborate with each other if they share similar cultural values. This affinity-based preference is robust across various within-firm specifications and are at least as quantitatively important as a colleague's prior technical experience in explaining the observed collaboration choices of employee inventors.

Given this large-scale evidence documenting the strong preferences of firm R&D employees to work with others that possess similar cultural values, we next examine the important question of how this familiarity bias in co-inventor network formation can impact team-level innovation productivity. To disentangle the selection effect from the treatment effect of inventor team diversity on team performance, we use a quasi-natural experiment involving premature co-inventor deaths that provides an exogenous source of variation in the cultural composition of inventor teams. Crucially, our most restrictive tests allow us to directly compare the output of treated teams working at the same firm that suffer the *same* co-inventor death (and thus the same loss of individual co-inventor human capital) but one team experiences an *increase* in cultural similarity after the focal co-inventor's death while the other treated team experiences a *decrease* in cultural similarity after the same co-inventor's death.

We find that the impact of a co-inventor's death on the surviving team depends in large part on the revised composition of the remaining team members' cultural values. First, we find that treated teams that become less culturally diverse produce a significantly higher quantity of patents than treated teams that become more culturally diverse. Second, we show that more culturally diverse treated teams exhibit a greater ability to develop higher quality, breakthrough innovations that garner citations towards the top of the future patent citation distribution. Third, we show that more culturally diverse teams tend to create more risky, explorative innovations, whereas less culturally diverse teams tend to produce more exploitative patents.

Alternatively stated, our combined results imply that the outputs of less culturally diverse teams tend to fall in the middle of the patenting outcome distribution, while the outputs of more culturally diverse teams tend to fall in the tails of the patenting outcome distribution. Interestingly, we find that these simultaneous changes in the quantity, quality, distribution, and type of output produced by more vs. less culturally diverse teams has offsetting effects on the overall team innovation output. Specifically, we find that total citation-weighted patents in the overall sample (encompassing all industries, time periods, and targeted innovation profiles) show no significant difference between more vs. less culturally diverse teams, suggesting that the innovation output of more culturally diverse teams is not uniformly better or worse compared to less culturally diverse teams; rather, it is simply different in nature.

To better understand the potential mechanisms driving the divergent innovation output profiles of teams with different cultural diversity, we provide novel micro-level evidence on how differences in the gathering, processing, and communication of information across team members from various cultural backgrounds can impact the innovation process. Specifically, we are the first researchers to provide empirical evidence at the inventor team level for two distinct but related channels: differences in the range of input information sources used in the idea generation/knowledge integration process and disparities in the efficiency of synthesizing the perspectives of heterogenous team members. We use a unique and comprehensive dataset of all patent and non-patent citations to identify key sources of information used by teams in their innovation production process.

First, we explore whether R&D teams primarily rely on their existing knowledge or actively seek new knowledge during the knowledge integration process. We find that, compared to culturally homogenous teams, culturally diverse teams exhibit a greater propensity to seek new information outside of their existing collective body of knowledge. Furthermore, when examining the type of 'new' information employed by R&D teams, we find that more culturally diverse teams tend to use more unconventional and risky knowledge resources such as foreign patents and non-patent references including academic and scientific articles. In contrast, less culturally diverse teams tend to rely on more proven technical resources like older U.S. patents. These results are consistent with our earlier theoretical predictions on the potential upsides of increased team cultural diversity such that culturally diverse teams tend to hold disparate initial viewpoints that cannot be easily integrated, prompting them to go beyond their current knowledge base to seek new and unconventional knowledge sources to reconcile their differing perspectives. This greater reliance on more diverse, less proven resources during the knowledge integration process can in turn help to explain why more culturally diverse teams tend to produce more exploratory patents and experience a higher variance in their R&D output, making them more likely to end up in the tails of the patenting outcome distribution (i.e., zero patents vs. very highly cited, high-impact innovations).

Second, we find that more culturally diverse teams are less able to develop subsequent patentable inventions within compressed time frames compared to less culturally diverse teams, after controlling for the type of innovation (i.e., exploitative vs. exploratory) that the team has previously produced. This result is consistent with our earlier theoretical predictions on the potential downsides of increased team cultural diversity, as culturally diverse teams require more time and effort to reconcile more frequent disparities in the initial perspectives of individual inventors. This in turn impedes the efficiency of the team's knowledge integration process and can result in a lower quantity of R&D output.

Overall, our collective findings present a more nuanced view of the relationship between team diversity and team performance. Even for two teams with similar sets of starting inputs, namely the knowledge, skills, and experiences contributed by each co-inventor, more culturally diverse teams appear to have very different knowledge integration processes compared to more culturally homogenous teams that in turn leads to very different innovation outputs. On the downside, diversity in cultural values can impede information sharing and viewpoint integration, thus adversely affecting team output. On the upside, a successful combination of differing perspectives within teams can positively impact the pursuit of technological innovation, particularly more high-risk, high-reward type of inventions. Our results offer the novel perspective that cultural diversity in inventor teams has both positive and negative consequences for team innovation production and that the net impact depends on the type of innovation pursued by the firm. For example, our results suggest that each firm will have a different (and possibly time-varying) optimal mix of culturally diverse vs. culturally homogenous teams to create the desired combination of exploratory vs. exploitative innovations. Notably, these findings contrast with most prior studies that document a uniformly positive or negative impact of diversity on innovation.

Regarding our identification strategy for estimating the causal impact of team cultural diversity on team innovation output, it is important to note that we employ a multi-faceted empirical strategy to mitigate potential selection effects and omitted variables concern. First, we can control for any unobserved, time-varying firm characteristics because all our analysis is based on comparing inventor teams operating within the same firm at the same time. Second, we use exogenous shocks to inventor team composition induced by premature co-inventor deaths to create exogenous variation in the cultural diversity of treated teams. We conduct a variety of tests to support our parallel trends assumption of the comparability of treated and control teams. Third, we show that the effect of these exogenous changes in team cultural diversity on innovation cannot be explained by other changes in team characteristics post-death including team gender, age, and immigrant status diversity as well as the amount and diversity of a team's knowledge, skills, and experiences. Fourth, beyond including team and year fixed effects in our diff-in-diff regressions, we include even more granular dead co-inventor fixed effects in our most restrictive tests. This setup thus compares two sets of treated teams actively working at the same firm where both suffer the loss of the *same* co-inventor but experience *differential* impacts on teams' cultural diversity. Importantly, our specification controls for differences in the relative contribution or value-add of the same dead inventor to each of their treated teams.[10] Finally, to further address any potential remaining concerns with omitted variable bias, we formulate predictions about the two-part innovation production process (specifically the knowledge integration process) that are grounded in theory specific to culture. Consistent with our theoretical predictions, we find empirical support for two underlying mechanisms that can explain *why* team cultural value diversity can have both positive and negative impacts on team innovation. Therefore, while it is impossible to completely prove causality in a large-scale corporate R&D setting, developing a plausible alternative hypothesis that is grounded in theory and that can *simultaneously* explain all our combined empirical findings across an array of selection, treatment, mechanisms, and robustness tests seems challenging.

Our paper's findings contribute to several strands of literature. First, there is a growing body of literature documenting the importance of co-inventor networks for *individual* productivity (e.g., Azoulay, Graff Zivin, and Wang, 2010; Jaravel, Petkova, and Bell, 2018). More broadly, our paper

[10] See Section 4.2.4 and Internet Appendix IA.4 for further details.

relates to studies that examine the role of ethnic scientific communities in facilitating technology transfer to their home countries (e.g., Kerr, 2008), as well as the contributions of first-generation immigrant inventors in creating knowledge spillovers with U.S.-born ("native") inventors (e.g., Bernstein, Diamond, Jiranaphawiboon, McQuade, and Pousada, 2022; Moser, Pasar, and San, 2023). These studies focus on *individual* innovative performance of immigrant vs. native inventors working across different institutions (e.g., self-employed, universities, government agencies, corporations), which does not account for heterogeneity across and within different types of institutions.

In contrast, we focus on corporate inventors and their associated teams *working within the same for-profit company*, which reflects how most individual inventors (and the incentive structures around them) are naturally organized in the modern innovation ecosystem. Specifically, we use a unique within firm setting to study initial *team* formation and compare the performance of inventor *teams* working within the same firm at the same point in time. In addition, we focus on cultural diversity because it directly affects a critical part of the innovation process, namely the 'knowledge integration' process governing how initial knowledge inputs are synthesized together to generate new joint technological outputs. We then develop testable predictions regarding the role of cultural diversity on the innovation process that are grounded in theory specific to culture. Our more granular ancestry country-level culture measures also allow us to explore heterogeneity *within* immigrant and native communities, instead of just comparing immigrants vs. natives as two distinct, homogenous groups. To the best of our knowledge, we are the first paper to conduct a comprehensive large-scale study that explores the role of cultural values in driving intra-firm inventor team formation and its consequent impact on the innovation production process and resulting outcomes of corporate R&D teams. Importantly, in contrast to studies such as Bernstein et al. (2022) that document uniformly positive patenting outcomes of interactions between first-generation immigrant and native inventors, our research provides a more nuanced perspective on diversity by uncovering both positive and negative effects through our within-firm team-based analysis.[11] Moreover, we go one step further by delving into the internal mechanisms of the knowledge integration process, presenting novel granular evidence on *how* cultural diversity impacts the nature of team innovation outputs.

Second, our paper contributes to an extensive literature that examines the relationship between various CEO-, firm-, investor-, and industry-level characteristics and firm-level innovation (e.g., Aghion, Van Reenen, and Zingales, 2013; Bena and Li, 2014; Seru, 2014; Islam and Zein, 2020). We add to studies examining the influence of diversity in management teams, board of directors, and the firm-wide labor force on firm performance. Unlike prior studies that examine firm-level innovation outcomes, we focus directly on the most critical element of the R&D process, namely the individuals ultimately responsible for creating new technologies, and explore how these inventors collaborate within firms. Our research design allows us to identify some unique team- and inventor-level drivers of corporate innovation while accounting for the effects of diversity in other groups and controlling for determinants of innovation at the CEO, firm, and industry levels.

Third, our paper has important policy implications for both firms and government regulators with respect to workplace diversity and labor productivity. Existing efforts to promote workplace diversity tend to primarily focus on the recruitment stage (i.e., the initial hiring of a more diverse set of people into the firm) rather than the later integration of these employees into open-ended teamwork-based activities.[12] Our evidence of strong homophily biases within firms suggests that merely increasing the hiring of workers from more diverse culture backgrounds is not sufficient to ensure that these individuals actually collaborate and share their unique perspectives with a more

[11] We also directly test this effect by controlling for diversity in team members' immigrant status in Section 6.2.1 and find that our results are unaffected.

[12] For example, according the 2011 Workplace Employment Relations Survey of UK managers and employees, approximately one quarter of respondent organizations monitored recruitment and selection policies to ensure that they promoted employee diversity but only a small fraction of these organizations then also implemented other policies to foster employee diversity post the initial recruitment screening stage (Van Wanrooy, Bewley, Bryson, Forth, Freeth, Stokes, and Wood, 2011).

diverse set of colleagues. Instead, firms may need to also enact policies that incentivize existing employees to form collaborations with a more diverse set of co-inventors. Our evidence on team performance also suggests that there are important economic trade-offs from an innovation productivity perspective in the pursuit of greater workplace diversity. While we show that more culturally diverse teams appear to have a greater ability to produce more high impact innovations, it is important to acknowledge that more culturally diverse teams are also relatively more likely to fail to produce patented research output. As such, our results suggest that each firm will have a different (and possibly time-varying) optimal mix of culturally diverse vs. culturally homogenous teams to create the desired combination of exploratory and exploitative innovations.

Finally, our paper contributes to the finance literature on how the cultural values of employees influence corporate decisions (e.g., Liu, 2016; Pan, Siegel, and Wang, 2017; Li, Mai, Shen, and Yan, 2021). More broadly, our paper adds to a growing strand of the economics literature examining the effect of culture on economic outcomes such as the use of financial contracts (Guiso, Sapienza, and Zingales, 2004) and labor choices (Fernández and Fogli, 2009). Since both culture and invention share a reliance on tacit knowledge, the open-ended nature of innovation search activities provides an ideal setting for testing the impact of culture on economic value creation, which is an important topic in the economics and finance literature (e.g., Guiso, Sapienza, and Zingales, 2006, 2015).

We organize the remainder of the paper as follows. Section II presents the data and construction of variables used in the empirical analysis. Section III analyzes the determinants of inventor team formation within an individual firm. Section IV explores the impact of shared culture and other attributes on an inventor team's innovation output. Section V explores the potential mechanisms underlying our main findings while Section VI outlines our various robustness tests.

II. DATA AND VARIABLES

2.1 Sample overview

Our analysis uses a combination of extensive patent-based data and large-scale information on individual inventors' careers and inherited characteristics. Given utility patents are one of the most common measures of innovation used in the prior literature and that all participants in the innovation eco-system (including firms and the USPTO) have strong legal and economic incentives to identify the "true and only" human inventors of a patentable technology (e.g., Gattari, 2005), we first collect information on U.S. patenting from three sources: the United States Patent & Trademark Office (USPTO), PatentsView, and the Berkeley-Fung Patent Database. PatentsView contains detailed disambiguated USPTO patent data from 1976 to 2018 and includes a patent's application and grant date, technology class, inventor names and locations, patent assignee names and locations (where the patent assignee is usually the firm at which the research is conducted) and the number of citations by and to a patent. Then, by using the Berkeley-Fung Patent Database (which extends the existing 1976–2006 NBER Patent–Compustat assignee database through to 2016) in conjunction with PatentsView company assignee ID numbers and our own database extensions,[13] we can identify all granted patents that are applied for between 1976 to 2016 and are assigned to U.S. publicly listed firms.

Next, we incorporate the PatentsView disambiguated inventor database that assigns each inventor one time-invariant ID to track each inventor's patent output and geographic location from 1976 onwards. The PatentsView inventor database encompasses over 3.8 million inventors working on over 6.2 million patents granted between 1976 and 2018. Following the prior literature, we define an inventor's employer or place of employment as the firm that is the assignee on the patent. As such,

[13] Through a combination of algorithms designed to identify similar corporate names as well as manual data checks on firms' time-varying lists of subsidiaries in SEC filings, we augment these existing patent assignee databases by linking patents granted in more recent decades to Compustat firms (enabling greater coverage of U.S. public firms post-2006).

we designate an inventor that files a patent with Firm X in 2005 and another patent with Firm Y in 2006 as an employee of Firm X in 2005 and an employee of Firm Y in 2006. If more than a year elapses between patent filings by the same inventor, we assume an inventor changes employers at the midpoint between the two patent application years following Li and Wang (2023). However, we also impose the requirement that an inventor must have patented at the focal employer at least once in the surrounding three-year period to be classified as an 'active' inventor working at that firm.[14]

Our baseline patent-inventor-firm linked sample includes 698,221 unique U.S.-based inventors employed at 5,403 unique U.S. publicly listed firms and covers the period from 1976 to 2016 to ensure we have at least five years of inventor activity before 1981 and after 2011. We focus on U.S.-based inventors working at U.S. public firms because these firms are likely to have a sufficiently large pool of inventors and distinct inventor teams working in geographically proximate locations to facilitate the large-scale within-firm analysis that is the basis of our identification strategy.

2.2 Identification of inventor teams at individual firms

Using augmented PatentsView inventor IDs (which trace all patents developed by an individual across time),[15] we identify and track all teams of two or more inventors that ever collaborated on a U.S. patent and assign each team a unique identifier. This process yields over 2.7 million distinct teams that co-invent at least one patent during our 40-year sample period. This unique team ID allows us to follow the patent output, citation patterns and technological specialization of each team formed since 1976. We then use this combination of patent, inventor, and team IDs, along with assignee names, to identify all the inventors and inventor teams employed at each U.S. public firm at a given

[14] For example, if a person files a patent with Firm A in 2000 and another patent with Firm B in 2010, we conservatively assume this inventor is an employee of Firm A up to and including 2003 and is an employee of Firm B from 2007 onwards.
[15] We initially use name- and location-based algorithms to identify potential duplicate inventor ID codes (a single inventor is erroneously assigned multiple inventor IDs) and potential 'over-aggregating' inventor ID codes (patent output of two or more distinct inventors is erroneously assigned to only one inventor ID). We then use manual data verification using LexisNexis Public Records and professional networking websites (such as LinkedIn and Relationship Science), as well as USPTO and Google searches, to update the PatentsView inventor database where clear misclassifications are identified.

time and place. This enables us to understand the factors that drive employee collaboration decisions and to compare the performance of teams with differing levels of cultural diversity and experience.

2.3 Variable construction

In this section, we describe the independent and dependent variables used in our analysis. Appendix A provides further details on the construction of each of these variables.

2.3.1 Innovation output

We construct four sets of patent-based measures at both individual inventor and inventor team levels to assess the quantity, quality, distribution, and type of innovation produced.

First, we measure patent quantity using *Total patents*, which represents the number of patents filed (and subsequently granted) in a given year.

Second, we measure patent quality using *Average forward cites per patent*, which is the number of citations that a patent receives divided by the average citations made to patents applied for in the same year and Cooperative Patent Classification (CPC) technology sub-class.[16] We scale raw citation counts to account for potential variation in citation rates across technologies and over time as well as to address truncation bias that results in patents granted towards the end of the sample having less time to garner citations (Hall, Jaffe, and Trajtenberg, 2001). We form the team-year level measure by calculating the average scaled forward citations across all the team's patents applied for in that year.[17]

Third, we consider where each of a team's patents fall in the patent citation distribution. Given that more high-risk, high-reward explorative innovation is more likely to fall in the tail ends of the patent citation distribution (Balsmeier, Fleming, and Manso, 2017), we examine whether the patents developed by a particular team are radical "high impact innovations." Following Islam and Zein

[16] Jointly developed by the USPTO and European Patent Office to create a unified global patent classification system, the CPC has about 650 4-digit technology sub-classes that group patents based on the similarity of their subject matter.
[17] In unreported robustness tests, we also use an alternative measure of average patent quality from Kogan, Papanikolaou, Seru, and Stoffman (2017) that assesses the market value of patents using announcement stock returns. We find very similar results to those reported in the paper when using average forward cites per patent to proxy for patent quality.

(2020), we define "high impact innovations" as those patents that receive citations within the highest decile of patents in the same application year and CPC technology sub-class (*Top 10% cited patent*).

Fourth, we seek to identify the type of innovation undertaken by inventor teams. As discussed in Balsmeier et al. (2017), innovation search strategy can be characterized as the trade-off between exploitative innovation (i.e., the exploitation of known technologies and/or existing capabilities) and exploratory innovation (i.e., the search for technologies and approaches that are distant from pre-existing knowledge sources). As such, we use two different measures to capture an inventor team's relative focus on exploitative innovation, following Balsmeier et al. (2017). The first measure is *Average backward cites per patent*, which is calculated based on the number of citations that a patent makes to relevant prior art (using the same year and technology class scaling adjustment as for forward citations). Patents with more backward citations should correlate with patenting in relatively more crowded, more mature technology areas, which is consistent with a greater (lesser) focus on exploitative (explorative) innovation. The second measure is *Average claims per patent*, computed based on the number of scaled claims made by a patent. A higher number of claims should correlate with a higher amount of effort exerted in delineating the extent of subject matter protection sought by a patent, where such efforts should increase as pressures for immediate and quantifiable results rises, consistent with a greater (lesser) focus on exploitative (explorative) innovation.

Finally, following Kogan, Papanikolaou, Seru, and Stoffman (2017), we measure a team's "net total" or "overall" innovation output using *Total citation-weighted patents*.

For all our patent-based variables, we follow the extensive prior finance literature (e.g., Farre-Mensa, Hedge, and Ljungqvist, 2020) and apply the natural logarithm-transformation of one plus these non-negative patent values. Section 4.2.4 provides more discussion on how our results are robust to using alternative transformations such as the inverse hyperbolic sine transformations.

2.3.2 Cultural values of individual inventors and inventor teams

We create measures of inventors' inherited cultural values as follows, where the full procedure is described in Internet Appendices IA.1 and IA.2. First, we identify inventors' countries of origin using a novel multi-layered method based on the following restricted, commercial and public datasets:

- Restricted full-count decennial U.S. census data from 1850 to 1940 from the Minnesota Population Center.[18] These datasets contain more than 500 million individual records with detailed individual and household information. These records represent the complete set of Census records available to the public in which the respondents' names are disclosed.
- Infutor's Consumer History Plus database, which is a commercial database covering 270 million individuals in the U.S. This database has detailed demographic information including names, year of birth, gender, the first five digits of an individual's social security number as well as current and past residential addresses and the associated dates.
- Berkeley Unified Numident Mortality Database (BUNMD), which is a cleaned and harmonized version of National Archives and Records Administration (NARA) Numident data based on Social Security applications, claims and death records. This data cover more than 49 million death records from 1988 to 2005 and contains information on individuals' social security number, names, gender, race, place of birth, date of birth, date of death, and parents' names (which frequently include mothers' maiden names).
- LexisNexis Public Records, which is a commercial database covering 280 million U.S. individuals. This database has detailed demographic information including names, year of birth, gender, the first five digits of an individual's social security number, current and past residential addresses and the associated dates, emails, education, and employer names.

[18] The 1890 Census database is not available to researchers because most of the 1890 census records were destroyed by a fire in the Commerce Department Building in 1921. Also, the full-count 1950 Census is not yet available to researchers.

- Additional sources for mapping names to countries of origin including Nationalize.io (https://nationalize.io), Forebears.io (https://forebears.io), NamePrism (https://www.name-prism.com), and the Dictionary of American Family Names from Oxford Reference.

Using U.S. Census records, the Infutor database, BUNMD, and LexisNexis Public Records, we identify relatives from each inventor's family tree to determine the exact countries of ancestry for 322,687 unique inventors. For the remaining inventors, we hand collect information on countries of origin using online sources (e.g., LinkedIn, Marquis Who's Who) as well as identify countries of origin for inventors based on their immigrant generations. Specifically, we identify first-generation immigrant inventors based on the first five digits of their social security numbers and determine their countries of origin based on their surnames using several public and commercial sources that are more suitable for recent immigrants.[19] Finally, for the other inventors who are most likely U.S.-born second or higher generation immigrants, we identify their countries of ancestry using the 1850 to 1940 U.S. Census records by finding potential relatives for these inventors as individuals who share the same surnames. Notably, this multi-layered procedure represents a significant methodological advancement compared to relying solely on name-based matching techniques used in prior studies.

Second, to measure the inherited cultural characteristics of inventors from a given country of origin and to capture the multi-dimensional nature of beliefs and values that underlying inventors' holistic perspectives, we use the cultural framework of Hofstede (1980) and Hofstede, Hofstede, and Minkov (2010) that classifies national culture into six dimensions.[20] The first cultural dimension is

[19] While the use of surnames to identify people's cultural background has been relied upon extensively in a wide variety of business disciplines including finance and accounting (e.g., Gompers et al., 2016; Liu, 2016; Pan et al., 2017; Brochet et al., 2019), we acknowledge that this method may not always yield exact matches. For instance, individuals may assume more Anglo-Saxon surnames upon entry into the U.S. In addition, married women may adopt their husbands' last names, although there is a strong tendency for females to marry men from the same cultural background as themselves (Kalmijn, 1998) and cross-cultural marriages were not common among 20th century immigrants (Pagnini and Morgan, 1990). For female inventors, we also use their maiden names provided by Infutor for matching to maximize accuracy where possible.

[20] By measuring specific cultural dimensions, we can capture an individual's core values and beliefs, which are deeper-level constructs that are conceptually more influential than surface-level constructs and go beyond categorical measures of nationalities and ethnicities (as individuals from different ethnicities and nations can share similar values and beliefs).

the individualism index (IDV), which is higher when people in a society are expected to take care of only themselves and their immediate families. The second cultural dimension is the power distance index (PDI), which measures the degree to which members of the society accept the hierarchical distribution of power and obey authority without questioning it. The third cultural dimension is the uncertainty avoidance index (UAI), which measures the extent to which people feel threatened by uncertainty and ambiguity and try to avoid these situations. The fourth cultural dimension is the masculinity index (MAS), which is higher when a society supports more traditional gender roles and emphasizes masculine values such as assertiveness and competitiveness. The fifth cultural dimension is the long-term orientation index (LTO), which is related to the fostering of virtues oriented towards future rewards. The sixth cultural dimension is the indulgence index (IVR), which measures the extent to which people try to control their desires and impulses. As described in Internet Appendices IA.1 and IA.2, we link the six-dimension Hofstede measures to inventors' countries of origin and obtain inherited cultural values for about 98% of the 1.5 million U.S.-based inventors that patented while working at any employer between 1981 and 2016.

We use Hofstede's cultural framework because it is the most influential cultural framework used in various disciplines including finance.[21] The Hofstede values have been replicated by many studies using different populations (e.g., Shane, 1995; Merritt, 2000) and its stability over time has been corroborated by Beugelsdijk, Maseland, and Hoorn (2015). While Hofstede's framework has become widely accepted since 1980, it has faced some criticisms such as the over-reliance on theoretical (over purely statistical) constructs (see Karolyi, 2016). To ensure robustness, we also use other cultural frameworks based on Schwartz (1992) and the World Value Survey and obtain similar results.[22]

[21] Per Karolyi (2016), the Hofstede dimensions are cited/used in 2,083 publications from 1980 to 2015 across 45 journals used in the Financial Times Business School Research Rankings. In finance, this framework has been used in many studies including Chui, Titman, and Wei (2010), Ahern, Daminelli, and Fracassi (2015), and Ahmad, de Bodt and Harford (2021).
[22] In unreported robustness tests, we use two other set of measures to identify inherited cultural values. The first set of measures is from the framework of Schwartz (1992) that classifies national culture into three dimensions: embeddedness

Similar to our usage of Hofstede's national cultural dimensions to construct inherited cultural values for inventors, several recent studies also use Hofstede's national indices to construct inherited cultural values for individuals such as CEOs (e.g., Nguyen, Hagendorff, and Eshraghi, 2018; Pan, Siegel, and Wang, 2020) and security analysts (e.g., Brochet, Miller, Naranjo, and Yu, 2019). This methodology follows the epidemiological approach from the economics literature (Fernández, 2011), which is based on the key idea that when individuals emigrate from their native country to a new country, their cultural beliefs and values travel with them, but their external environment is left behind. Moreover, these immigrants not only bring their cultural beliefs and values to the new country, but they also pass down these beliefs to their descendants. Therefore, inherited cultural values are relevant for both immigrants and their descendants in the U.S., which in our case refer to the U.S.-based inventors that we study. In contrast, non-cultural factors such as institutions and economic environments are indigenous to their native countries, thus are geographically immobile and should not be relevant for immigrants and their descendants in the U.S.

It is important to note that this approach does not imply that inventors have values identical to the average person in their country of ancestry. Instead, this approach explores whether inventors' cultural heritage, as measured by the prevalent preferences in their country of ancestry, significantly impacts their values and preferences. Given our particular interest in interactions among inventors, the underlying premise is that inventors sharing similar cultural backgrounds tend to have more similar values compared to inventors from different cultural backgrounds. In other words, we are assuming that inventors are not systematically drawn from specific parts of the population such that inventors from similar cultural backgrounds in fact have more divergent values than inventors from

vs. autonomy, hierarchy vs. egalitarianism, and mastery vs. harmony. The second set of measures includes trust, hierarchy, and individualism constructed using the World Value Survey following Ahern et al. (2015). Note that two Hofstede dimensions, power distance (i.e., hierarchy) and individualism (i.e., autonomy), overlap with these measures.

different cultural backgrounds. Additionally, if these inventors are fully assimilated to the extent that cultural heritage becomes irrelevant to their values, we would not expect to observe any evidence of homophily or other discernible differences between inventors from diverse cultural backgrounds.

Nevertheless, in Internet Appendix IA.3, we empirically test the relation between the Hofstede cultural values in an individual's country of ancestry and the individual's personal beliefs using individual responses to the U.S. General Social Survey. These results further validate the underlying assumption of our culture measure as well as the persistence of cultural beliefs across generations.

For our subsequent empirical tests, we measure the similarity in cultural values between a pair of inventors (i, j) by first calculating the Euclidean pairwise distance between two inventors' cultural values based on Hofstede's framework according to the following formula:

$$\begin{aligned} &Pairwise\ distance_{ij} \\ &\qquad = [(IDV_{Inventor\ i} - IDV_{Inventor\ j})^2 + (PDI_{Inventor\ i} - PDI_{Inventor\ j})^2 \\ &\qquad + (UAI_{Inventor\ i} - UAI_{Inventor\ j})^2 + (MAS_{Inventor\ i} - MAS_{Inventor\ j})^2 \\ &\qquad + (LTO_{Inventor\ i} - LTO_{Inventor\ j})^2 + (IVR_{Inventor\ i} - IVR_{Inventor\ j})^2]^{\frac{1}{2}} \end{aligned}$$

$Coinventor\ cultural\ similarity$ is then obtained by multiplying this pairwise distance measure by –1. We transform this measure so that its values are bounded between zero and one, where higher values indicate higher similarity of cultural values between inventor i and inventor j.

To measure cultural similarity within a team of N inventors, we calculate the average of pairwise distances in all pairs of team members using the following formula:

$$Team\ cultural\ similarity = -1 \times \frac{\sum_{i,j} Pairwise\ distance_{ij}}{\frac{N(N-1)}{2}} \quad \forall i < j$$

where $Pairwise\ distance_{ij}$ is calculated as above. This measure of $Team\ cultural\ similarity$ is transformed so that its values are bounded between zero and one, where higher values indicate higher cultural similarity or less cultural diversity among the R&D team members.

2.3.3 Acquired knowledge, skill, and experience of individual inventors and inventor teams

Using the patenting history of each individual inventor, we develop several measures of the professional knowledge, skills, and experiences acquired by inventors throughout their career. We then aggregate these measures to the team-level to include as controls in all ensuing empirical tests.

Controls for the size and diversity of an inventor team's initial knowledge base

To measure the breadth of an individual inventor's accumulated knowledge base that they can contribute to the new R&D collaboration, we first compute *Individual technical knowledge base* as the number of distinct CPC sub-classes in which that inventor has patented to date (see generally Verhoeven, 2023). We then define our first set of controls for the size and diversity of a team's initial knowledge base as the team-level average of *Individual technical knowledge base* and the team-level standard deviation of *Individual technical knowledge base*, respectively.

At the team-level, we define two additional variables to measure the range and overlap in the collective technical knowledge of an inventor team. The first measure, *Scope of team technical knowledge,* is defined as the natural logarithm of one plus the number of unique technology classes patented in by at least one team member prior to year t. This variable captures the idea that the higher the number of distinct technology classes worked on by inventors in the team, the greater (and broader) is that team's starting level of technical knowledge. The second measure, *Team technical knowledge overlap*, is defined as the average cosine similarity in team members' pairwise experiences across different technology classes. This measure captures the similarity of the focal team's technical knowledge base, where lower values indicate that each team member's prior technical knowledge is quite distinct compared to the prior technical knowledge of their focal teammates (Jaffe, 1989).

Controls for the amount and diversity of an inventor team's skill

To account for the impact of inventor skill/talent on patenting outcomes, we gauge how accomplished or successful an individual inventor has been relative to their peers in their inventor career to date. We use an inventor's observed productivity to date to proxy for that inventor's latent skill or talent in generating new inventions in the future. We first compute each inventor's *Total patents to date* and their *Average forward cites to date per patent*, where we assume that inventors who have produced a greater quantity and/or quality of patents to date possess relatively higher skill.

For our OLS, treatment, and mechanisms tests, we control for the initial amount of inventor skill that each team member brings to the new collaboration by calculating the team-level average of each inventor's *Total patents to date* and *Average forward cites to date per patent*, respectively. We also control for diversity in skill across team members by computing the team-level standard deviation of each inventor's *Total patents to date* and *Average forward cites to date per patent*, respectively.

Controls for the amount and diversity of an inventor team's experience

To account for the potential effect of prior experience on future patenting output, we measure the cumulative experience acquired by each inventor over their entire career to date. We first calculate *Years of inventor experience* as the number of years between the application date of the first (granted) patent an inventor applies for and the current year. To capture an inventor's relative experience in undertaking more exploitation-focused innovation compared to exploratory innovation, we compute the *Average backward citations per patent* produced by the focal inventor to date.

For our OLS, treatment, and mechanisms tests, we control for the amount of experience that each inventor contributes to the team by calculating the team-level average of each inventor's *Years of inventor experience* and *Average backward citations per patent*, respectively. We also account for diversity in experience among team members by computing the team-level standard deviation of each inventor's *Years of inventor experience* and *Average backward citations per patent*, respectively.[23]

[23] We argue that we can better capture an inventor's relevant (and more recent/practical) technical experience with their patenting record as an inventor (including information on the technology classes they have patented in) compared to their educational attainment in early adulthood. This is especially so given that most U.S. inventors since 1950 possess an undergraduate college degree in a STEM-related field of study (Baumol, Schilling, and Wolff, 2009).

2.3.4 Additional control variables

To account for the role of physical proximity in facilitating team formation, we calculate the geodetic distance (in miles) between each pair of inventors in the focal team. *Team geographic diversity* is then calculated as the team-level average of the pairwise co-inventor geographic distances. For our OLS tests, we also control for the number of inventors in the focal team (i.e., *Team size*).[24]

2.4 Summary statistics

Table 1 provides the mean, median and standard deviation of the various characteristics of our baseline sample of 698,221 U.S.-based inventors working at U.S. publicly listed firms.

In Panel A of Table 1, we present summary statistics detailing the characteristics of all individual inventors working at U.S. public firms across our sample period. In Panel B of Table 1, we provide information on team-level characteristics for the inventor teams in our sample. Consistent with Jaravel et al. (2018), we find that while teamwork is common among inventors employed at large firms, individual inventors usually only collaborate with a small number of other inventors over the course of their career. This implies that co-inventor networks are relatively sticky and exert an important influence on an individual inventor's long-term productivity. We explore the factors influencing the formation of these co-inventor relationships in more detail in Section 3.

In Panel C of Table 1, we show the sample pairwise characteristics of newly formed co-inventor pairs at the time of their first collaboration. One particularly noteworthy feature is that first-time collaborators tend to have greater similarity in their cultural values (or less distance in their cultural values). We explore this univariate relationship further in our co-inventor selection analysis in Section 3 as well as the associated team performance implications in Section 4.

[24] Note that *Team size* cannot be included as a control variable in our difference-in-difference treatment effects tests with team fixed effects because changes in the team size variable at treated teams are collinear with the *After* post-death indicator variable (while there are no changes in inventor team size at matched control firms over the sample window).

III. SELECTION FACTORS IN INVENTOR TEAM FORMATION

In this part, we examine the influence of cultural values in explaining the formation of new R&D collaborations between inventors employed at the same firm.

3.1 Empirical methodology – ex ante selection

Following papers such as Dyck, Morse, and Zingales (2010) and Bena and Li (2014), we estimate the following conditional logit regressions using cross-sectional data:

$$\begin{aligned} Actual\ pair_{ij,t} &= \alpha + \beta_1 Coinventor\ cultural\ similarity_{ij,t-1} \\ &+ \beta_2 Coinventor\ geographic\ distance_{ij,t-1} \\ &+ \beta_3 Coinventor\ difference\ in\ average\ forward\ cites\ to\ date\ per\ patent\ _{ij,t-1} \\ &+ \beta_4 Both\ top\ 10\%\ inventors_{ij,t-1} + \ \beta_5 Coinventor\ technological\ proximity_{ij,t-1} \\ &+ \beta_6 Coinventor\ difference\ in\ average\ backward\ cites\ to\ date\ per\ patent_{ij,t-1} \\ &+ \beta_7 Coinventor\ difference\ in\ years\ of\ inventor\ experience\ to\ date_{ij,t-1} \\ &+ \beta_8 Both\ have\ 5+years\ of\ tenure\ at\ focal\ firm_{ij,t-1} + Group\ FE + \varepsilon_{ij,t} \end{aligned} \tag{1}$$

$Actual\ pair_{ij,t}$, equals one if the inventor pair i, j represents a real, first-time collaboration formed at a U.S. public firm in year t ('treated pair'), and zero otherwise.[25] For each realized inventor pairing, there are two counterfactual co-inventor pairs (i.e., 'control pairs') whose construction is outlined in Section 3.2. As a result, $Group\ FE$ represents the fixed effect for each new realized pairing of co-inventors and its counterfactual control pairs of potential (but ultimately unchosen) co-inventors. As such, our coefficient estimates are based on 'within group' variation in pairwise characteristics.

Our main variable of interest in this analysis is $Coinventor\ cultural\ similarity_{ij,t-1}$ which captures the degree to which shared cultural values is a meaningful driver of realized intra-firm co-inventor pairing, after controlling for the relative difference in the knowledge, skill, work experience, and proximity of the two inventors. Appendix A defines all other variables. Following Bena and Li (2014), we use robust standard errors clustered at the 'group' level.

[25] We focus on the very first collaboration between two individual U.S.-based inventors working at the same U.S. public firm since the decision of two U.S. inventors to collaborate for the first time is unaffected by confounding factors such as experience with past collaborations and accumulated team-specific relationship capital. We use the patent application date as an objective estimate of when two inventors begun their collaboration following Jaravel et al., (2018).

3.2 Counterfactual inventor pairs

To understand which factors influence the formation of new collaborations between inventors, we use our within-firm R&D setting to construct a plausible set of potential co-inventors that were available for collaboration at the time when the focal inventor decided to collaborate with a different co-inventor. This methodology of comparing realized pairings with counterfactual pairings is similar to the one employed by Gompers et al. (2016) in the context of venture capital syndicate formation.

We start the process of generating credible counterfactual inventor pairs by first identifying each pair of inventors that actually initiate a first-time collaboration at a U.S. public firm f in year t (denote as $Inventor_{T1}$ and $Inventor_{T2}$, respectively). Next, we generate pseudo inventor pairs by partnering each of $Inventor_{T1}$ and $Inventor_{T2}$ with one potential (but ultimately unchosen) collaborator that is most comparable to the actually chosen co-inventor on the following dimensions (denote these control co-inventors as $Inventor_{C2}$ and $Inventor_{C1}$, respectively). First, the counterfactual co-inventor and the actual co-inventor must be currently employed at the same firm f in year t. Second, the pseudo co-inventor must not have ever previously collaborated with either inventor in the treated pair.[26] Third, we select the inventor with the same number of (eventually granted) patent applications to date as the actually chosen co-inventor to serve as counterfactual control co-inventors $Inventor_{C2}$ and $Inventor_{C1}$ respectively).[27] This final requirement helps to ensure that the counterfactual inventor has similar innovation experience and patenting productivity to the actually chosen collaborator.

After implementing this procedure, our initial pairwise dataset contains 1.848 million first time collaborations formed between 1981 to 2016 that involve pairs of U.S.-based inventors employed at the same public firm (realized pair $Inventor_{T1}$–$Inventor_{T2}$) and 3.521 million pseudo-control pairs

[26] Our results are very similar if we instead exclude any inventors from the counterfactual control sample who ever collaborate with either inventor in the treated pair, irrespective of whether the collaboration occurs in the future or not.
[27] If more than one inventor at the firm satisfies these three criteria for selection as a counterfactual control inventor, we follow Jaravel et al. (2018) by choosing among these potential counterfactual inventors at random.

($Inventor_{T1}$–$Inventor_{C2}$ and $Inventor_{C1}$–$Inventor_{T2}$).[28] Once we merge in our shared culture measures, we obtain a final dataset of 1.821 million treated pairs and 3.444 million control pairs.

3.3 Determinants of inventor collaboration decisions

Table 2 examines the ability of various pairwise inventor characteristics to explain the observed collaboration decisions of inventors working within the same firm. We find strong evidence that homophily drives partnering decisions in internal skilled labor markets. This result applies to both shared cultural values as well as overlaps in technical skills and relative career experience.

We first document that shared cultural values plays a critical role in the decision of inventors to form new teams with each other. For example, we find that two inventors with similar cultural values are approximately 32% more likely to work together than two otherwise comparable inventors who have a one standard deviation greater distance in their cultural values (see column (1)).[29] These findings are consistent with experimental studies such as Calder-Wang et al. (2023) who find strong homophily biases in team formation among groups of Harvard Business School MBA students.[30]

To further verify the robustness of our selection test results, we incorporate stricter filters in the formation of counterfactual control pairs as follows. First, since geographic proximity is an important driver of inventor partnering decisions (Gera, 2013), as confirmed in column (1) of Table 2, we specify that the potential (but ultimately not chosen) co-inventor must also be located within 50 miles of the actual co-inventor (see column (2) of Table 2).[31] Second, given the possibility that different

---

[28] In less than 2% of cases, we are unable to find any valid counterfactual control pairs that meet our criteria for inclusion in the conditional logit analysis. Given that these rare cases tend to arise in small public firms with a limited pool of R&D personnel and relatively low patenting output, we drop these observations from our final selection sample.

[29] Beyond shared culture, we show that inventors with more similarity in their acquired career experiences are more likely to collaborate with each other. For example, we find that inventors prefer to work with others that have a similar amount of prior experience as an inventor. A one standard deviation reduction (about 5 years) in the difference of general inventor experience between two co-workers raises the probability of an actual collaboration between them by approximately 9%.

[30] It is important to note that our results are not simply driven by two inventors having the same country of origin. In Table IA.2, we re-run equation (1) only on the subsample of inventor pairs where the two inventors are *not* from the same country of origin. We still find that the coefficient on our shared culture measure remains strongly positive and significant.

[31] We use a 50-mile radius cut-off as an estimate of the high likelihood that two co-inventors work in the same office location (Tian, 2011). However, our results are qualitatively unchanged if we instead use a 25-mile or 100-mile threshold.

divisions within the larger corporate entity may have differential access to firm resources and inventor talent, we define an alternative counterfactual sample using only potential (but ultimately unchosen) co-inventors that work in the same division or subsidiary of the firm at the same time (see column (3) of Table 2).[32] Finally, our most restrictive counterfactual sample requires that both treated and control co-inventor pairs must involve two inventors who are actively working: (a) in the same division of the same company *and* (b) at the same geographic location/office (see column (4) of Table 2).

Table 2 presents the results from re-estimating equation (1) with each alternative control sample. Despite using more restrictive definitions of counterfactual co-inventors for comparison purposes, we show that even two inventors working in the same company division at the same office location are approximately 30% more likely to work with each other if there is a one standard deviation increase in the similarity of their cultural values. These are economically and statistically significant effects, especially when assessed in the context of comparing pairs of similarly experienced inventors working in the same corporate division at the same geographic location at the same point in time.

Overall, our results demonstrate the important role of affinity-based personal characteristics such as shared cultural values in shaping within-firm co-inventor networks. Our findings highlight a key challenge for both organizations and policymakers when attempting to design workplace diversity policies. Our evidence on the strong influence of homophily in inventor team formation suggests that merely increasing the hiring of workers from more diverse backgrounds into the firm is unlikely to be sufficient in realizing any benefits of workplace diversity. Instead, firms may need to also enact proactive policies that incentivize existing employees to form a more diverse set of R&D teams. Otherwise, the homophily biases we document may result in the oversupply of relatively homogenous teams within a company, even for those actively targeting a more diverse initial pool of skilled labor.

[32] We identify two inventors as being part of the same division if there is an exact match on the name and location of the patent assignee in their most recent patents. For example, diversified corporations such as Tesla, Inc. will usually file patents under the specific subsidiary that created the invention (e.g., Tesla Motors Inc., Tesla Electronics Inc., etc.).

IV. EFFECT OF INVENTOR DIVERSITY ON INNOVATION OUTPUT

Given our strong evidence of homophily in corporate R&D settings, a natural question to explore is whether these familiarity biases enhance or impede different types of innovation output (the 'ex post treatment effect'). To this end, we first run OLS regressions that show the baseline association between the degree of diversity in team members' cultural values and team R&D output.

However, the identification challenge here is that the selection of co-inventors is not a random process. For instance, inventors may intentionally target new collaborations with individuals from relatively different personal and professional backgrounds to pursue a riskier, exploration-focused innovation search strategy. Therefore, any differences in the average innovation outcomes of diverse and non-diverse teams may be due to *selection* effects that arise from endogenous co-inventor matching or *treatment* effects of inventor team diversity on team performance. To disentangle these two effects, we use exogenous changes in team cultural diversity generated by premature co-inventor deaths to identify the causal impact of inventor team cultural diversity on team innovation outcomes.

4.1 Baseline association between inventor team diversity and team innovation output

In this section, we explore how diversity in the cultural values of individuals working in a corporate R&D team impacts the quantity, quality, distribution, and type of innovation output produced. We start our analysis by identifying all (eventually granted) patents applied for by two or more inventors working at the same U.S. publicly listed firm between 1981 and 2016. We further require all team members to be located within the United States at the time of patenting and that both personal and patent profile characteristics are available for all inventors listed on the patent.

4.1.1 Empirical methodology

With this initial sample of U.S.-based corporate inventor teams, we then construct yearly team-based measures of innovation output, team cultural similarity, and controls as described in Section 2.3. To create a more representative panel of patent output produced by an inventor team across time, we classify a team as 'actively' collaborating together in a given year if all members of the team are 'active' employee inventors at the same focal firm in that year (see Section 2.1 for further details).[33] The advantage of this approach is that it can better capture the entire range of research outcomes of a R&D team (particularly failed research pursuits that do not result in a new patent) rather than only focusing on (relatively successful) R&D projects that produce observable patent output.[34] Our final sample consists of approximately 3.385 million team-year observations spanning 1981 to 2016.

To investigate the baseline relationship between an inventor team's cultural diversity and team innovation output, we run the following ordinary least squares (OLS) regression specification:

$$\begin{aligned} &Team\ innovation_{i,t} \\ &\quad = \alpha + \beta_1 Team\ cultural\ similarity_i + \gamma X_{i,t} + Firm \times Year\ FEs + \varepsilon_{i,t,j} \end{aligned} \quad (2)$$

The dependent variable $Team\ innovation_{i,t}$ is one of the patent-based outcome measures for team i working at U.S. listed employer firm j in year t as discussed in Section 2.3.1. It is set equal to zero if the team has no patent output in that year. The construction of $Team\ cultural\ similarity_i$ is based on the method outlined in Section 2.3.2. The vector $X_{i,t}$ contains various team-level controls including team geographic diversity; team size; the size (and diversity) of a team's knowledge base, the amount (and diversity) of individual inventors' skills in a team, and the amount (and diversity) of experience accumulated by team members in their career to date (see Sections 2.3.3 and 2.3.4 for further details). We use Firm × Year fixed effects in all specifications such that we compare the output of more vs. less culturally diverse teams working within the same firm at the same point in time.[35] The standard errors are clustered at the inventor team level.

---

[33] For example, if Inventor A, B and C (together Team 1) are identified as 'active' inventors at focal firm F between 2000–2006, 2001–2006 and 2002–2007, respectively and Team 1 (successfully) applies for two patents in 2003 only, we include Team 1 in our annual team-year panel from 2002–2006 with zero patent output in 2002 and 2004–2006.

[34] For robustness, we run an alternative specification where we only include team-year observations in which the focal team did in fact formulate at least one patent. We reach the same conclusions with similar levels of statistical significance.

[35] We cannot include Team fixed effects in an OLS setting because a team's cultural composition does not vary over time.

4.1.2 Empirical results

Table 3 presents the results of our baseline OLS regressions. Unlike much of the prior literature, we provide novel evidence that diversity in an inventor team's inherited cultural characteristics has significant positive *and* negative effects on team innovation production.

First, as shown in column (1) of Table 3, we find that less culturally diverse teams produce a significantly higher overall quantity of patents than more diverse teams. This result is consistent with the organizational behavioral theory that it is relatively easier to coordinate production of immediate, quantifiable output when team members share common perspectives (e.g., Van den Steen, 2010).

Second, we document important differences in the observed distribution of innovation quality outcomes between less culturally diverse vs. more culturally diverse teams. We first find that the patents produced by more culturally diverse inventor teams tend to be of significantly higher average quality (measured using average forward patent citations in column (2) of Table 3) than those patents developed by less culturally diverse teams. In addition, more culturally diverse teams have a greater ability to produce radical or breakthrough innovations, as measured by the number of team patents that fall in the top 10% of the future patent citation distribution (see column (3) of Table 3). This implies that while more culturally diverse teams are more likely to engage in failed research endeavors (in terms of a lower raw number of patents generated), the patents that are successfully developed by more culturally diverse R&D teams tend to have a greater impact on future commercial technological development (and thus garner citations towards the top of the future patent citation distribution).

Third, we show in columns (4) and (5) of Table 3 that while less culturally diverse teams tend to produce more exploitative patents, more culturally diverse teams tend to produce more explorative patents. This result is consistent with the notion that teams with greater homogeneity in their cognitive thought processes and beliefs are more likely to search in their common-known areas of technical expertise for more incremental technology improvements. In contrast, more culturally diverse teams are more likely to engage in risky research pursuits outside the boundaries of existing knowledge.

A natural final question to ask is what the net impact on "total" team-level innovation of these differences in the type of innovation produced by teams with differing levels of cultural diversity. Interestingly, we show in column (6) that total citation-weighted patents are insignificantly different between more vs. less diverse teams. This result supports a more nuanced view of the relationship between cultural team diversity and innovation. It suggests that the innovation of more diverse teams is not uniformly better or worse compared to less diverse teams; rather, it is simply different in nature.

Our OLS results imply that an inventor team's cultural diversity is associated with different types of innovation output. We show that more homogeneous teams tend to produce a higher quantity of patents that are more likely to exploit existing technologies and achieve moderate success. In contrast, more culturally diverse teams tend to produce a higher share of risky, more exploratory patents that have a greater chance of becoming high impact innovations. This is consistent with the idea that while more culturally diverse teams may encounter greater difficulties in synthesizing different viewpoints and communication preferences (leading to less total patenting), successfully combining these more disparate perspectives can foster more exploratory and high-impact innovation.

4.2 Quasi-natural experiment involving co-inventor deaths

While the previous results indicate that more (less) culturally diverse teams tend to produce more exploratory (exploitative) innovations, it is still possible that unobserved differences in team characteristics drive both the initial decision to collaborate and subsequent team innovation production. An ideal experiment to establish the impact of inventor team cultural diversity on team innovation output would be to randomly assign certain inventors to form more vs. less culturally diverse teams. Such an experiment would eliminate (endogenous) co-inventor selection effects, thus enabling us to estimate the treatment effect of cultural diversity on team performance. Unfortunately, it is almost impossible to find a corporate R&D setting that convincingly approximates this ideal experiment. However, another useful experiment would be to randomly vary team cultural diversity *after* initial collaboration decisions are made. This would enable us to identify the treatment effect of cultural diversity on team innovation, holding initial team selection effects fixed. If any differences in innovation outcomes between more vs. less culturally diverse teams are driven purely by selection, exogenous variation in cultural diversity post-team formation should have no significant effect on team innovation outcomes. In this paper, we attempt to approximate this second experiment.

Therefore, to address any remaining selection concerns, we exploit a quasi-natural experiment involving premature co-inventor deaths that allows us to observe the change in a team's innovation output after the team experiences an exogenous shock to the team's cultural diversity.

4.2.1 Triple difference-in-differences approach

Relying upon the premature death methodology used in prior studies (e.g., Azoulay et al., 2010; Jaravel et al., 2018), we provide causal estimates of how an inventor team's innovation production would change if there was an exogenous shift in the diversity of the cultural values of its team members. Using data from the USPTO, Infutor, LexisNexis Public Records, and the Fold3 Social Security Death Index, supplemented by extensive manual verification, we exploit the premature deaths of inventors working at U.S. public firms at the time of their passing as a source of exogenous variation in the diversity of a team's cultural values to examine the evolution of treated team innovative output around co-inventor deaths.

We identify the causal effect of inventor team cultural diversity on team innovation performance by utilizing a triple difference-in-differences research design. We start our analysis by identifying a control group of inventor teams working at the same firm at the same time whose co-inventors do not pass away but who have a similar level of cultural diversity as the treated team (pre-death) and who are otherwise similar to the teams that experience the premature death of a co-inventor. However, we do not simply compare the change in the innovative output of treated and control teams around co-inventor deaths to identify the effect of cultural diversity on inventor team output. This is because the difference in the subsequent innovation of treated teams and control teams in the post-treatment period may be due to factors other than changes in the cultural value composition of treated teams induced by co-inventor deaths. These factors may include, for example, the productivity shock to team skill and experience arising from a colleague's unexpected departure.

Importantly, a unique aspect of our setting is that a co-inventor's death can exogenously increase *or* decrease the cultural value similarity of a treated team's surviving inventors. This allows us to compare the difference in innovation output between treated teams whose cultural similarity *increases* post their co-inventor's death and their associated control teams versus the difference in innovation output between treated teams whose cultural similarity *decreases* post their co-inventor's death and their associated control teams. Under the identifying assumption that, conditional on observable team and inventor characteristics, there is no other contemporaneous shock that systematically affects the relative outcomes of the treatment group around co-inventor death date (Gruber, 1994; see also Section 4.2.2), we can use a triple difference-in-differences regression setup to isolate the causal effect of inventor team cultural diversity on team innovation.

We use the following triple difference-in-differences empirical setup to investigate how changes in team cultural diversity impacts subsequent team performance. We estimate this regression using a panel dataset that compares the difference in output between treated teams whose cultural similarity increased (i.e., team cultural diversity decreased) after a co-inventor's death relative to their associated control teams versus the difference in output between treated teams whose cultural similarity decreased after a co-inventor's death relative to their matched control teams:

$$\begin{aligned} Team\ Innovation_{i,t} &= \alpha + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Treated\ team_i \\ &+ \beta_3 After_{i,t} \times Treated\ team_i \times Team\ cultural\ similarity\ change_i \\ &+ \gamma X_{i,t} + Team\ FEs + Year\ FEs + \varepsilon_{i,t} \end{aligned} \tag{3}$$

The dependent variable, $Team\ innovation_{i,t}$, is one of the patent-based outcome measures in year t as described in Section 2.3.1.[36] It is set equal to zero if the team has no patent output in that year. The indicator variable, $After_{i,t}$, equals one for all years after the focal co-inventor's death and zero otherwise. The indicator variable, $Treated\ team_i$, equals one for all teams that experience the shock of losing a co-inventor and zero otherwise. The continuous variable $Team\ cultural\ similarity\ change_i$, equals the difference between the treated team's cultural similarity immediately post the focal co-inventor's death (based on the team's surviving inventors) minus the treated team's cultural similarity immediately prior to the death (which includes the eventually deceased co-inventor).[37] As a result, $Team\ cultural\ similarity\ change_i$ will be positive when the treated team's cultural value similarity increases post their co-inventor's death and will be negative when the treated team's cultural similarity decreases post their co-inventor's death.[38] The key coefficient of interest in this regression is β_3, which compares the relative pre- and post-death impact on team innovation for cases where treated inventor team cultural similarity increases to cases where treated team cultural similarity decreases. We include Team fixed effects to difference away any time-invariant team-level characteristics and Year fixed effects to absorb any common time trends across teams that experience either an increase or a decrease in team cultural similarity.[39] In additional robustness tests, we show in Table IA.4 that our results presented in Table 5 are robust to also including Firm $\times$ Year fixed effects that filter out the effects of any time-varying firm-level characteristics.

[36] Following Jaravel et al. (2018), we examine the change in team innovative output from ten years prior to the focal co-inventor's death to ten years post-death. Our results are qualitatively unchanged with similar levels of statistical significance if we narrow our focus to the 5-year period pre- and post- the focal co-inventor's death.

[37] Treated teams only suffer a change in cultural similarity when the decreased inventor departs the team while control teams don't experience any change in team cultural similarity since control team membership remains constant over time.

[38] In unreported robustness tests, we alternatively define the indicator variable *Team cultural similarity increases* that is equal to one for treated teams that experience an increase in their team's cultural value similarity after their teammate's premature death, and zero otherwise. We reach qualitatively unchanged results with very similar levels of statistical significance to those presented using the continuous variable *Team cultural similarity change*.

[39] Note that Team fixed effects absorbs the *Treated team* dummy variable as well as the coefficients on *Team cultural similarity change* and *After* × *Team cultural similarity change* (noting that no control team observations experience a change in team cultural similarity throughout the pre- and post-death period).

Evidently, a co-inventor's death can induce other key changes in team-related characteristics besides team cultural similarity. For example, a colleague's death may change the amount and/or diversity of team experience. As such, we explicitly control for other changes in non-culture related team variables induced by the focal co-inventor's death. $X_{i,t}$ comprises various control variables including Team geographic diversity, the size and diversity of a team's knowledge base, the amount and diversity of team members' skills, and the amount and diversity of team members' experience (see Section 2.3). As we explain in Section 4.2.4 and Internet Appendix IA.4, this specification effectively controls for the deceased inventor's *relative* contribution to each of their treated teams in terms of knowledge, skill, and experience. This is achieved through our tracking of aggregate team characteristics before and after the inventor's death, combined with the inclusion of team fixed effects.

To implement our difference-in-differences identification approach, we identify active inventor teams working at U.S. publicly listed firms that suffer the 'premature' death of one of their team members (i.e., the "treated teams") and then construct their associated "control teams." We provide a summary of this method below, while the full procedure is described in Internet Appendix IA.5.

To identify treated teams, we use Infutor's Consumer History Plus File, LexisNexis Public Records, USPTO data, and the Fold3 Social Security Death Index to find deceased inventors who:

a) Died between 1981 and 2011 (to ensure we have at least 5 years of pre- and post-death data);
b) Are employed at a U.S. publicly listed firm at the time of their death; and
c) Are no older than 60 years of age at the time of their death, following Jaravel et al. (2018).

Next, to isolate 'active' inventor teams that are likely to be genuinely impacted by the death of their colleague, we identify all teams of 3+ inventors[40] that collaborated with the deceased inventor (at the same firm) on a patent that was applied for within 3 years of the focal co-inventor's death.[41] After

[40] We exclude original two-person inventor teams (i.e., deceased inventor plus one more surviving co-inventor) because it is not feasible to calculate meaningful "team-based" diversity measures in the post-treatment period for such teams.
[41] Our empirical results are very similar if we use a 5-year 'active' cut-off threshold instead of 3 years.

merging with our database of U.S.-based inventors' various characteristics, we identify 2,637 treated teams that were actively collaborating at a U.S. public firm around their co-worker's premature death.

Next, we use the following procedure to match each treated team with a corresponding control team that does not experience a co-inventor death. We specify that each control team must:

(1) Work at the same firm and have the same team size in the pre-death period as the treated team;

(2) Have no team members that are part of an existing collaboration with any treated team member;

(3) Be actively collaborating around the time of the focal co-inventor death (i.e., have successfully patented together at least once in the 3 years leading up to their colleague's death);

(4) Have developed the same number of (eventually granted) patent applications to date as the treated team at the time of the focal co-inventor's death; and

(5) Have the closest proximity to the treated team's pre-death cultural diversity value.

We use this control team's characteristics and patenting activity as the counterfactual for how the relevant treated team would have performed if they did not suffer the loss of their collaborator. After implementing this process, we have a final sample of 2,441 treated teams and 2,441 control teams.[42]

4.2.2 Evidence supporting identification assumptions

As discussed previously, our triple difference-in-differences regressions focus on comparing the changes in innovative output for treated inventor teams whose cultural similarity increases post their co-inventor's death (relative to their associated control teams) with the changes in innovative output for treated inventor teams whose cultural similarity decreases post their co-inventor's death (relative to their associated control teams). Our key identifying assumption is that, conditional on controlling for changes in observable team and inventor characteristics, randomly distributed co-inventor deaths

[42] We are not able to find a suitable counterfactual control team for approximately 7% of our sample. This principally occurs in smaller U.S. public firms with a more limited pool of inventors that are unaffiliated with inventors comprising the treated teams. However, in unreported tests, the unmatched treated teams do not appear to be significantly different (at least on observable characteristics) from those that do find a matching control team. Note also that *all* treated teams are included in our most restrictive heterogeneous treatment effects specification outlined in Section 4.2.4.

do not cause a contemporaneous shock to an unobserved variable that is systematically correlated with the subsequent patenting output of these different sets of treated teams based on whether there is a positive or negative change in team cultural similarity post co-inventor death.

To assess the reasonableness of these underlying assumptions, we first explore whether there are significant differences in observable characteristics between treated and control teams (Roberts and Whited, 2013). As shown in Panels A and B of Table 4, we find no significant differences between the two sets of inventor teams across a range of team innovation outputs (i.e., the quantity, quality, distribution, and type of patents produced) and observable team characteristics (including average experience and technical expertise of team members) leading up to the focal co-inventor's death.

To more directly evaluate the reasonableness of the key identifying assumption underlying our difference-in-differences estimator, we next compare the pre-treatment output and characteristics of treated inventor teams, which are divided by whether the treated team's cultural value similarity does or does not increase post their co-inventor's death. As shown in Panels C and D of Table 4, we find no significant differences between each set of treated teams leading up to their co-inventor's death.

Crucially, we further show that the team-specific loss of technical knowledge and experience arising from the focal co-inventor's death is similar for treated teams whose cultural similarity increases post-death vs. treated teams whose cultural similarity decreases post-death. For example, the total number of unique technology class knowledge areas that is contributed solely by the deceased inventor, as well as the distinctiveness of the deceased inventor's knowledge and experience relative to other co-inventors within the treated team, remain insignificantly different regardless of whether the cultural value similarity of the focal team increased or decreased post their colleague's death.[43] This suggests that the relative importance of the distinct technical contributions of the focal deceased inventor to a given corporate R&D team is relatively similar across both sets of treated teams.

[43] This is captured by the variables, *Percentage of unique tech class knowledge lost due to co-inventor death* and *Change in Team technical knowledge overlap due to co-inventor death*, respectively (see Internet Appendix IA.4 for more details).

Furthermore, in Panel E of Table 4, we present evidence that there are no significant differences in the personal traits, professional characteristics and productivity of the deceased inventors who exogenously depart treated teams where cultural similarity increases after their death compared to the characteristics of deceased inventors at treated teams where cultural similarity does not increase after their death. These combined results are consistent with the notion that the random distribution of inventor deaths across time and individuals represents an exogenous shock to team cultural diversity that is not systematically correlated with changes in team quality and other unobserved attributes.

Finally, we assess the validity of the parallel trend assumption in our setting by investigating the pre-treatment innovation trends between the treated and control groups (see Gao and Zhang, 2017). We define nine dummies from *Year –3* to *Year +4*, and *Year 5+* to indicate the year relative to the co-inventor death. We then re-estimate equation (3) by replacing the $After$ indicator with these nine dummies. In Figure 1, we show that none of the coefficients on the triple interaction terms (i.e., β_3) involving *Year –3, Year –2*, *Year –1*, and *Year 0* indicators are statistically significant across our main outcome variables, suggesting that the parallel trend assumption is not violated in our case.[44]

Overall, the evidence suggests that the treated and control teams in our test samples are comparable in terms of professional accomplishments, personal traits, and innovation potential, allowing us to estimate the causal effect of inventor team cultural diversity on team performance.

4.2.3 Empirical results of triple differences-in-differences tests

Table 5 presents the results of our triple diff-in-diff research design where our focus is on the triple interaction coefficient β_3 in equation (3).[45] First, as shown in column (1) of Table 5, we find

[44] Since innovation is a relatively long-term process (Hirshleifer, Hsu, and Li, 2013) and that random co-inventor deaths are clearly not part of a planned change in a team's patenting strategy, it is unsurprising that we do not observe an immediate change in the innovation trajectory/approach of treated inventor teams post their colleague's death.

[45] As an aside, the somewhat mechanical reason for the strongly negative coefficient on the *After* term in this specification is because we require all treated and control teams to have patented in the pre-treatment period (to identify active teams) but there will be some teams (both treated and control) that will not generate any patents in the post-treatment period.

that teams that experience an exogenous increase in the level of team cultural value similarity (i.e., become less culturally diverse) produce a 5% higher overall quantity of patents relative to teams that experience an exogenous decrease in team cultural similarity (i.e., become more culturally diverse). Second, the result on average forward patent citations in column (2) shows that the average quality of patents tends to decline as inventor teams become less culturally diverse, although this change is not statistically significant at conventional levels. Third, in column (3), we find that a decrease in team cultural similarity (i.e., an increase in team cultural diversity) increases the likelihood of the team producing high impact innovations by approximately 15%, measured as the number of team patents that fall in the top 10% of the patent citation distribution. Fourth, in columns (4) and (5) of Table 5, we provide evidence that teams that experience an increase (decrease) in cultural diversity following a co-inventor's death are relatively more likely to produce more exploratory (exploitative) patents. Furthermore, in additional unreported robustness tests, we construct the 'self-cite ratio' measure outlined in Balsmeier et al. (2017)[46] and find that inventor teams that become more culturally homogenous after a co-inventor death are significantly more likely to cite the firm's previously developed patents when formulating subsequent innovations. This finding aligns with the idea of less culturally diverse teams engaging in greater exploitation of the firm's pre-existing technological expertise. Finally, in column (6), we show that the net overall (citation-weighted) innovation output of more culturally diverse teams is not significantly different from that of less culturally diverse teams.[47] As discussed earlier, our findings are unchanged when we include Firm $\times$ Year fixed effects in our regressions (see Table IA.4).

[46] Following Balsmeier et al. (2017), the 'self-cite ratio' for each patent is computed as the number of backward citations to U.S. patents owned by the team's corporate employer, divided by the total number of backward citations to U.S. patents. We then calculate the 'average self-cite ratio' across all patents developed by the inventor team each year.

[47] Note that if a less culturally diverse team produces a single patent in one year that is subsequently uncited, while a more culturally diverse team does not produce any patents during that year, the comparison in the patent quantity outcome regression would be 1 vs. 0 total patents, respectively. However, the comparison in the total citation-weighted patent outcome regression would be 0 citations for both teams. Almost a quarter of all patents never garner a subsequent citation.

Overall, the results in Table 5 support a causal interpretation of the trends uncovered in the OLS tests of Table 3. Even when comparing groups of inventor teams with a similar amount and diversity of knowledge, skills, and experiences, our evidence implies that the diversity in cultural values within an inventor team exerts a significant causal influence on team innovation outcomes. In contrast to the prior literature that highlights one directional effects of diversity, our results offer a novel perspective that team cultural diversity does not have a uniformly positive or negative effect on technological development. Instead, it significantly affects the type of team innovation output: less culturally diverse teams tend to generate a higher overall quantity of patents that exploit existing technologies and become moderately successful inventions, while more culturally diverse teams produce a higher share of risky, exploratory patents with greater potential to become radical innovations.

4.2.4 Alternative heterogeneous treatment effects specification (treated teams only)

A potential concern with our triple differences-in-differences with matched controls approach is that it may be unreasonable to compare the innovation trajectory of a treated inventor team, which undergoes upheaval and numerical disadvantage due to a co-inventor's death, with the innovation trajectory of a counterfactual control team that does not experience such turmoil. To address this potential concern, we employ an alternative approach with a heterogeneous treatment effects specification. We only focus on the treated team subsample (where all teams in this subsample experience an exogenous shock to their personal and professional composition due to a co-inventor's premature death) and compare the innovation outcomes of treated teams that experience an increase in cultural value similarity after a co-inventor's death with treated teams that experience a decrease in cultural value similarity post a co-inventor's death.

One of the key advantages of this alternative heterogeneous treatment effects specification is that we can include "dead co-inventor fixed effects" to control for a deceased individual's unique accumulated human capital and professional/personal traits. Essentially, this empirical test involves comparing two sets of treated teams actively working at the same firm at the same time where both experience the loss of the *same* co-inventor. However, one treated team experiences an increase in cultural similarity after the co-inventor's death, while the other treated team experiences a decrease in cultural similarity.

Crucially, our novel heterogenous treatment effects specification not only allows us to control for the *absolute* loss of human capital arising from a team member's death (through the inclusion of dead co-inventor fixed effects), but we can also control for the *relative* contribution or value-add of the deceased inventor to each of their respective treated inventor teams.[48] A potential concern regarding our revised identification strategy is that there may be a systematic relationship between changes in team cultural diversity induced by the focal co-inventor's death and the relative contribution/value-add of the deceased inventor to each of their treated R&D teams. This alternative possibility would predict that teams suffering a relatively greater (lesser) loss of value-add will experience uniformly worse (better) subsequent innovation outcomes.

With respect to this possible concern, we first note that, at a more conceptual level, our unique distributional results concerning both the positive *and* negative effects of team cultural value diversity on team innovation output do not appear to be logically consistent with the predictions of this alternative possibility. Moreover, at an empirical level, we note that our univariate comparisons in Panels D and E of Table 4 indicate that the relative importance of the deceased inventor's contribution of knowledge, skills, and experiences to each specific team is similar across both sets of treated teams in this analysis. For example, we show in Panel D of Table 4 that the *Percentage of unique tech class knowledge lost due to the focal co-inventor's death* and the *Change in team technical knowledge overlap due to co-inventor death* is insignificantly different between treated teams whose cultural

[48] For example, it is possible that the loss of the only engineer on a treated team has a relatively greater (negative) impact on team patenting output than the loss of that same engineer on another team comprised solely of fellow engineers.

value similarity increases post-death vs. treated teams whose cultural similarity does not increase post-death.[49] Furthermore, by incorporating the change in each team's amount (and diversity) of knowledge, skills, and experiences pre- and post- the focal co-inventor's death, combined with the inclusion of team fixed effects and dead co-inventor fixed effects, we explicitly control for the unique value-add provided by the deceased inventor in terms of technical expertise to each of their respective treated teams (refer to Internet Appendix IA.4 for further discussion). Thus, by keeping the loss of inventor-specific human capital/skill constant (since both teams experience the same co-inventor death) as well as explicitly controlling for the relative knowledge, skill, and experience contributed by the focal dead inventor to each of their respective treated team, we can better isolate the causal effect of changes in team cultural diversity on team performance.

To implement this alternative treatment test, we use the following regression specification:

$$\begin{aligned} Team\ Innovation_{i,t} &= \alpha + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Team\ cultural\ similarity\ change_i \\ &+ \gamma X_{i,t} + Team\ FEs + Dead\ coinventor\ FEs + Year\ FEs + \varepsilon_{i,t} \end{aligned} \tag{4}$$

The dependent variable, $Team\ innovation_{i,t}$, is one of the patent outcome measures described in Section 2.3.1. The indicator variable, $After_{i,t}$, equals one for all years after the focal co-inventor's death, and zero otherwise. The continuous variable, $Team\ cultural\ similarity\ change_i$, equals the difference between the treated team's cultural similarity post the focal co-inventor's death minus the treated team's cultural similarity pre-death. We include three set of fixed effects: *Team fixed effects* to difference away any time-invariant team characteristics; *Dead co-inventor fixed effects* to control for each deceased inventor's unique traits and experiences; and *Year fixed effects* to absorb any common time trends across teams that experience either an increase or a decrease in cultural diversity. The vector, $X_{i,t}$, comprises the same control variables outlined in Section 4.2.1. We also conduct

[49] The comparability of our treated teams, in terms of team patenting output, is further confirmed by the univariate results presented in Panel C of Table 4 as well as unreported parallel trends tests.

(unreported) tests, analogous to those undertaken in Section 4.2.2 and presented in Figure 1, for our triple difference-in-differences analysis, which support the parallel trends assumption in this setting.

Table 6 presents coefficient estimates from equation (4) using only treated teams that experienced the exogenous shock of losing an active co-inventor. Consistent with the results of the triple difference-in-differences with matched controls approach in Table 5, we see that inventor teams that experience an exogenous decline in team cultural diversity (i.e., the remaining team is more similar in terms of cultural values than the original, pre-death team) tend to produce a higher quantity of more exploitative, moderately cited patents.[50] In contrast, inventor teams that experience an exogenous increase in cultural diversity tend to produce more risky, more explorative patents that have a greater probability of falling into the upper tails of the future patent citation distribution. This suggests that, even among teams with a similar amount and diversity of knowledge, skills and experiences, more culturally diverse inventor teams demonstrate a greater ability to generate riskier yet more novel technological breakthroughs.

We re-emphasize that all our empirical estimates include explicit controls for the amount and diversity of technical knowledge, skills, and experiences across the co-inventors that comprise the focal team. Notably, team-level diversity in knowledge, skills, and experiences does not explain the unique combination of positive and negative effects of team cultural diversity on team innovation outcomes.[51] This is likely explained by the fact that team cultural diversity captures differences in values and perspectives whose effects are concentrated in the second part of the innovation process centered around knowledge integration, while diversity in knowledge, skills, and experiences relates primarily to the first part of the innovation process.

[50] Consistent with the unreported results in Table 5, we find that the average self-cite ratio of inventor teams significantly increases (indicative of more exploitative innovation output) if they become more culturally homogenous post-death.
[51] This is further reinforced by the fact that the pairwise correlation between our team-based cultural value measures and all our control variables is less than 0.05 in absolute terms.

As outlined in Section 2.3.1, our main tests use the natural logarithm-transformation for our patent-based dependent and independent variables. For robustness, we follow the recommendations of the recent econometrics literature (for a summary, see Aihounton and Henningsen, 2021) and re-run all our analysis in Tables 5 and 6 using the inverse hyperbolic sine transformation.[52] We find very similar results to those reported (see, for example, Table IA.5 that replicates Table 6).

V. POTENTIAL MECHANISMS AFFECTING TEAM PRODUCTION

Given our empirical evidence on the heterogenous effects of inventor team cultural diversity on team innovation outcomes, we now seek to develop testable hypotheses regarding the potential underlying mechanisms that may explain why diversity in the cultural values of individual team members significantly affects team innovation production. Specifically, we examine two distinct but related channels. The first channel is related to differences in the range and type of input information sources utilized during the idea generation process. The second channel is related to disparities in the efficiency of integrating the perspectives of different team members. While we cannot definitively establish causality based on the results presented in this section, our evidence does support existing models that underline the crucial role that diversity in the cultural values and perspectives of individual inventors can play in shaping team output in complex technological environments.

5.1 Propensity to incorporate new input information sources

A common rationale for advocating diversity in various workplace settings is that individuals with different backgrounds and experiences tend to approach problems from distinct yet equally valuable perspectives, potentially leading to more rigorously tested and higher quality solutions (Kang, Kim, and Oh, 2022). This concept is particularly applicable in team-based R&D settings when

[52] Unlike log transformations, the inverse hyperbolic sine transformation is still defined at zero while continuing to retain the useful properties of log transformations such as being more robust to outliers in right-skewed distributions.

dealing with complex problems, where the optimal end output or even the optimal search process for discovering a potential solution is often highly uncertain (D'Acunto, Tate, and Yang, 2021).

One potential driver of these differing perspectives among workers with heterogenous cultural values could be large differences in the awareness and/or willingness of individual team members to consider a more diverse range of information sources as inputs into the economic production process (see generally, Kim and Starks (2016) and Bernile et al. (2018) in the context of Boards of Directors). In our specific context, it is possible that inventor teams comprised of more culturally diverse individuals will have greater awareness of and/or exhibit an increased willingness to utilize a broader (and possibly riskier) range of input knowledge sources to generate new innovative ideas. These effects are likely to be amplified by the observation that more culturally diverse teams, who more frequently face the challenge of trying to reconcile more disparate starting viewpoints and integrate more diverse problem-solving approaches, may need to consider materials outside of their existing collective knowledge base to produce synthesized technological solutions. In contrast, more culturally homogenous teams, who share greater similarities in their starting perspectives and cognitive thought processes, may have less incentive to consider more distant information sources and are thus more likely to leverage off their existing bank of codified knowledge (e.g., Van den Steen, 2010).

Therefore, even conditional on the scope of the focal team's existing knowledge/skill set, we hypothesize that:

<u>Prediction 1</u>: *More culturally diverse teams will exhibit a higher propensity to draw on information outside of that team's pre-existing knowledge base when developing new patents together.*

<u>Prediction 2</u>: *More culturally diverse teams will rely on more non-traditional (and thus riskier) knowledge sources as inputs into the team innovation production process.*

If these predictions are true, then it could help to explain our previous finding that the output of more culturally diverse teams is more likely to end up in the tails of the patenting outcome distribution (i.e., relatively more failed searches that produce zero patents vs. the production of a greater number of patents that combine more novel and risky information sources to create more exploratory, path-breaking innovations) compared to less culturally diverse teams.

To test these conjectures, we examine whether more culturally diverse teams are more likely to cite 'prior art' in their patent applications that: (a) has *not* been previously developed or cited by any member of the inventor team in any of their prior patents (i.e., cites to "new knowledge") and (b) is relatively less proven and/or more distant from established theories in that technology field. One unique advantage of our setting is that all patent applicants have a "duty of candor and good faith" to disclose all prior arts that are material to the patentability of their applications (37 CFR 1.56). Given the potential adverse consequences of failing to cite key prior arts and that citation lists are reviewed by external USPTO patent examiners, applicants have a strong incentive to properly cite all relevant materials in their USPTO filings (Hirshleifer, Hsu and Li, 2018). As such, we use the prior art cited on the focal patent document as an objective and unbiased proxy for the information set that was in fact used by the inventor team to create the focal patented technology.[53]

To help determine *how much* an inventor team is drawing on novel information sources relative to their pre-existing knowledge base in order to develop their subsequent innovations, we first estimate the total size and composition of the focal inventor team's existing knowledge base each year. For each team patent applied for in year t, we examine every patent co-invented by any team member up to and including year $t-1$ (whether individually or jointly) and record every backward citation made to U.S. patents, foreign patents, or "other references".[54] Analogous to Ma (2020), we estimate an inventor team's existing knowledge base in year t as any U.S. patent, foreign patent, or

[53] See Hirshleifer et al. (2018) for further discussion of the reasons why the tendency to 'under-cite' or 'over-cite' is relatively negligible in this setting.
[54] 'Other references' includes all non-patent material that the inventor team cite as prior art in their patent application such as scientific/academic journal articles, books, trade publications, public/private sector reports and conference proceedings.

other references that has been previously cited or created by at least one team member prior to year t.[55] A backward citation on the focal year t patent to a U.S. patent, foreign patent, or other references that is outside the team's existing knowledge base is thus classified as a "cite to new knowledge."

We then compute each patent's "*new cite ratio*" as the total number of cites to new knowledge divided by the total number of backward citations to any information source. Thus, a new cite ratio value closer to one indicates a relatively high reliance on "new-to-the-team" information sources when developing the focal innovation idea while a new cite ratio value closer to zero suggests that the inventor team primarily drew on their pre-existing stock of knowledge. It should be noted that our "new cite ratio" measure implicitly controls for the team's general propensity to cite prior art since the denominator of the new cite ratio is total backward citations to any information source.

Next, we assess the *type and risk profile* of input information sources utilized by inventor teams by categorizing each backward citation that an inventor team makes on the focal patent to "new knowledge" into three key information source categories: (1) cites to 'new' foreign patents, (2) cites to 'new' other references, and (3) cites to 'new' mature U.S. patents (where 'mature' U.S. patents are defined as those granted U.S. patents that were developed at least 10+ years ago).

We argue that a team's greater reliance on foreign patents and other non-patent references as their underlying source of "new-to-the-team" knowledge is significantly riskier (albeit potentially more novel) problem-solving strategy than the utilization of older 'mature' U.S. patents. In particular, other (non-patent) references are comprised of some resources that are not as commercially focused and technically relevant as U.S. patents. For example, academic publications (that represent the largest proportion of 'other references') are more likely to contain riskier basic research ideas, where there is inherent uncertainty in translating academic hypotheses and tests to real-world, commercial

[55] In contrast to the prior literature, however, our estimate of the 'existing knowledge base' is measured at the team-level (rather than the firm-level) and we include foreign patents and other references (not just U.S. patent backward citations).

applications (Arora, Belenzon, and Sheer, 2021). Similarly, foreign patents granted in overseas jurisdictions tend to be less well known and cited, and thus not as widely used, tested, and advanced upon by subsequent inventors, compared to U.S. issued patents (e.g., Harhoff, Narin, Scherer, and Vopel, 1999). Conversely, a relatively greater reliance on older U.S. patents as the underlying source of new knowledge is consistent with that inventor team's preference to build upon well-established technologies with more proven commercial utility (Mukherjee, Romero, Jones, and Uzzi, 2017).

We hypothesize that more culturally diverse teams, which exhibit greater variation in the values and perspectives of individual team members that need to be reconciled, will have a relatively:

a) *Higher percentage of back cites to 'new' foreign patents* on the basis that more culturally diverse teams are more likely to have at least one team member that is aware of and willing to rely on technology developed outside of the United States when formulating new ideas
b) *Higher percentage of back cites to 'new' other references* because more culturally diverse teams are more likely to incorporate non-traditional, non-patent materials into the team's innovation search process to help resolve more frequent conflicts in individual perspectives
c) *Lower percentage of back cites to 'new' mature U.S. patents* since less culturally diverse teams, with greater similarity in their problem-solving approaches, may have less incentive to search outside the readily available bank of directly relevant (both commercially and technically) patent-based knowledge, especially more tried and tested 'mature' U.S. patents.

As a final step, we take the average of each of these four $Team\ information\ source\ ratio$ measures across all patents ever developed together by the focal inventor team. By construction, as each team will only have one sample observation, our method implicitly reduces the influence that any single team, particularly more productive ones, may have on our estimates.[56]

[56] Nevertheless, for robustness purposes, we also re-run our analysis at the individual patent level (which implicitly puts more weight on relatively more productive and frequently patenting inventor teams) and obtain very similar results.

Using the same initial sample of inventor teams that patent at U.S. publicly listed firms between 1981 and 2016 (see Section 4.1.1), we run the following OLS regression specification:

$$\begin{aligned}&\mathit{Team\ information\ source\ ratio}_i = \alpha + \beta_1 \mathit{Team\ cultural\ similarity}_i \\ &\quad + \beta_2 \mathit{Scope\ of\ team's\ existing\ knowledge\ base\ across\ all\ information\ sources}_i \\ &\quad + \gamma X_{i,t} + \mathit{Firm} \times \mathit{Year\ FEs} + \varepsilon_{i,t,j}\end{aligned} \tag{5}$$

where subscript i denotes the focal inventor team, t denotes the first year that the focal team patents together and j denotes the U.S. public firm at which the team is employed. We use the same set of controls as in Section 4.1.1, but we also include the *Scope of team's existing knowledge base across all information sources* to control for the possibility that it is more difficult for an inventor team to make a cite to "new knowledge" if the team's pre-existing knowledge base is larger. We use Firm × Year fixed effects in all our regressions and cluster standard errors at the inventor team level.

We present the mechanism results in Table 7. In column (1), we find that, even among inventor teams with a similarly sized existing knowledge base and with a similar amount of knowledge, skill, and experiential diversity, more culturally diverse teams are more likely to draw on information resources outside of their existing knowledge base compared to less culturally diverse teams. In columns (2) to (4), we further show that, when teams are searching for new knowledge to help reconcile differences and synthesize their respective knowledge to create new inventions, more diverse teams are 6% and 15% more likely to rely on riskier, more unconventional knowledge resources such as foreign patents and other non-patent references, respectively. In contrast, less culturally diverse teams tend to rely much more on well-known and conventional resources such as older U.S. patents (whose subsequent history of success/usefulness is more readily observable).

To ensure that our results are not purely driven by inventors with a foreign country of origin citing patents issued by the same foreign country, we find in unreported robustness analysis that our results are very similar if we exclude any cases where a corporate R&D team includes an inventor with a foreign country of origin and that team then cites a foreign patent issued by the same foreign country (i.e. a team with an inventor with Japanese cultural origin cites a Japanese-issued patent), implying that our team-based findings are quite distinct from the results in Kerr (2008).

Overall, we provide some of the first large-scale micro-level evidence on how inventor teams gather and process information during the knowledge integration part of the innovation production process. Our results showing the greater willingness of more culturally diverse teams to incorporate new information in the form of more unconventional, risky knowledge sources into their research pursuits can help to explain the novel distributional effects uncovered in our study. Specifically, the greater reliance on more diverse, less proven input knowledge sources seems to lead culturally diverse teams to produce more exploratory patents and realize higher variance in their innovation outcomes, making them relatively more likely to end up in the tails of the patenting outcome distribution.

5.2 Heterogenous ability to integrate individual team member perspectives

Aside from gathering information and potential ideas from various sources for the team to consider as part of the 'knowledge integration' process, another crucial aspect of the innovation production process is for the team to attempt to synthesize each team member's proposed ideas into a cohesive technological solution. Prior studies such as Giannetti and Zhao (2019) suggest that more diverse teams in other business-related contexts tend to experience greater difficulties in reconciling the viewpoints of various team members in an expeditious manner. These challenges faced by more diverse teams can stem from more frequent disparities in starting perspectives (Horwitz and Horwitz, 2007) or difficulties in effectively communicating opinions with team members from heterogenous backgrounds and efficiently incorporating those ideas into the final product (Van den Steen, 2010). As a result, we hypothesize that more diverse teams may take longer to evaluate and synthesize different co-inventor perspectives into a new patentable innovation, even among teams with similarly sized initial information sets and similar innovative search focus.

In an ideal setting, we would like to test this hypothesis by computing the length of time that elapses from the date of initial team formation to patent application date (for more successful R&D endeavors) or project abandonment date (for failed R&D searches) on a project-by-project basis. Unfortunately, a limitation of the patent dataset is that we do not observe how long it took for an inventor team to develop their first patent together. However, we can gauge how quickly a given team can formulate and execute on *subsequent* patentable projects. Therefore, we examine the number of patents that a team produces in the next one to two years after their first patent application together. We conjecture that less culturally diverse teams, due to greater similarities in their innate information processing and communication tendencies, will be able to produce tangible R&D output in a shorter time frame after the team's first initial patent together, compared to more culturally diverse teams.

To more formally assess the possibility that teams with varying cultural value composition may exhibit heterogenous abilities to integrate various team member perspectives, we use the same initial sample of inventor teams as in Section 4.1.1 and run the following OLS regression:

$$\begin{aligned} &Team\ patent\ output\ within\ y\ years_i = \alpha + \beta_1 Team\ cultural\ similarity_i \\ &\quad +\beta_2 Scope\ of\ team's\ existing\ knowledge\ base\ across\ all\ information\ sources_i \\ &\quad +\beta_3 Degree\ of\ focus\ on\ exploitative\ innovation\ by\ team's first\ joint\ patent_i \\ &\quad +\gamma X_{i,t} + Firm \times Year\ FEs + \varepsilon_{i,t,j} \end{aligned} \tag{6}$$

where the dependent variable is the log of one plus the number of patents produced by the team within $y = \{1, 2\}$ years after the team's first patent together, subscript i denotes the identity of the inventor team, t denotes the first year that the team patents together and j denotes the firm at which the team is employed. We use the same set of controls as in Section 4.1.1, but we also include the *Scope of team's existing knowledge base across all information sources* and the *Degree of focus on exploitative innovation by the team's first joint patent together* (measured as the total backward citations to all knowledge sources including U.S. patents, foreign patents, and other references by the team's first joint patent). We use this latter measure as a proxy for the general type of R&D projects that this team is working on together, where we expect that teams that are more focused on exploitative (explorative) innovations will generally take less (more) time to co-invent subsequent patents. We use Firm × Year fixed effects in all our regressions and cluster standard errors at the team level.

In Table 7, we find evidence consistent with the theory that, due to the greater time needed to reconcile heterogenous individual inventor viewpoints and synthesize a more diverse set of input information sources, more culturally diverse teams are less able to produce subsequent patentable innovations in shorter time frames compared to less culturally diverse teams. In columns (5) and (6) of Table 7, we find that more culturally diverse teams produce significantly fewer patents in the one- or two-year period after the team's first patent together relative to less culturally diverse teams, respectively. In other words, even conditional on two inventor teams working on patents with a similar level of exploratory focus and having a similarly sized existing knowledge base, more culturally diverse inventor teams seem to take significantly longer in general to develop follow-up innovations together.[57] While we acknowledge the data limitations in this analysis, this result nevertheless provides further suggestive evidence that more culturally diverse teams face greater challenges in synthesizing varying perspectives in a timely manner compared to more homogenous teams.

Overall, the results in this section are consistent with the theory that while more culturally diverse teams may produce a lower number of patents overall due to difficulties in integrating more disparate problem-solving and communication approaches into a unified technological framework, the successful combination of these more disparate perspectives in the team innovation production process can provide more culturally diverse teams with a greater chance of discovering more impactful, radical inventions.

[57] As expected, the strong positive coefficient on the *Degree of focus on exploitative innovation by the team's first joint patent* implies that a team that initially produces a more exploitative patent together is much more likely to quickly develop a subsequent patent together, consistent with the idea that more exploitative innovations are faster to produce.

VI. ADDITIONAL ROBUSTNESS ANALYSIS

In this section, we conduct additional empirical analyses to assess the robustness of our previously reported selection and treatment effects results from Sections III and IV, respectively.

6.1 Accounting for seniority and experience in collaboration decisions

A potential concern about our selection analysis is that each inventor may not have the ability to freely choose their collaborators within the firm. For example, senior managers may dictate which inventors are assigned to work on pre-identified projects. Our first observation is that extensive survey evidence (e.g., Marvel, Griffin, Hebda, and Vojak, 2007; Corsino, Giuri, and Torrisi, 2019) shows that employee inventors, particularly those working within larger corporate R&D units, highly value and often have the freedom to select research areas and team composition as well as allocate their time across R&D projects. Furthermore, especially in our sample of large public firms that operate in dynamic innovative industries, it seems unlikely that firms and their employees could feasibly foresee and dictate all potential co-inventors and future projects on which teams may work (Zhang, 2022).

Nevertheless, we empirically mitigate the above concern by focusing on a subsample of more senior and/or more accomplished inventors because they have relatively greater choice in their co-inventor selection decisions. Specifically, we restrict our sample to only include the collaboration choices of the following set of inventors who are choosing between potential co-inventors: (a) prolific or 'star' inventors (i.e., 'Top 10% Inventors'); (b) those with 10+ years of total inventor experience to date and (c) those who have worked at their current employer for 10+ years. The results are presented in Table IA.3 which re-estimates equation (1) using only the collaboration choices of more experienced and/or more prolific corporate inventors. We find that even highly skilled and/or experienced inventors are 22%–27% more likely to choose to collaborate with colleagues who share similar cultural values compared with otherwise comparably skilled candidates working in the same corporate division and office location, but who have a more divergent set of cultural beliefs.

6.2 Accounting for alternative diversity-based explanations

In this sub-section, we re-run our treatment effects tests to examine whether diversity in other inherited traits of inventors can explain our unique distributional results for innovation outcomes with respect to team cultural value diversity.

6.2.1 Team diversity in immigrant status

One dimension of 'diversity' previously studied in the prior literature is the role of first-generation immigrant inventors in facilitating two-way knowledge spillovers with native inventors and the positive consequences of these knowledge spillovers for patented innovation (see e.g., Bernstein et al., 2022; Moser, Pasar, and San, 2023). As such, we investigate whether *Team immigrant similarity* can at least partially explain our main results.

To construct *Team immigrant similarity*, we first identify the immigrant status of each inventor in the focal team by utilizing the first five digits of social security numbers of the matched inventors in Infutor's Consumer History Plus database. Following Bernstein et al. (2022), we identify inventors who are first-generation immigrants as individuals who are more than 20 years old when assigned a social security number (SSN). To estimate when inventors received their SSN, we decode SSN provided in Infutor using https://www.ssn-verify.com, which provides information on the issue state, the first year and last year issued for each set of five-digit social security number. These first-generation immigrant inventors account for about 16% of total inventors. *Team immigrant similarity* is then calculated as one minus the Blau (1977) immigrant diversity measure, which equals to $2 \times (1 - [(\text{Percent of team that is first-generation immigrant})^2 + (\text{Percent of team that is U.S. native})^2])$. Higher values of *Team immigrant similarity* signify less team immigrant diversity.

Interestingly, we find that, once endogenous matching and selection effects are accounted for in our unique within-firm and heterogeneous treatment effects specification with Dead co-inventor fixed effects as well as controlling for an extensive set of acquired and inherited characteristic variables, *Team immigrant similarity* does *not* appear to have a significant causal impact on future team innovation output as shown in Table 8. At the same time, we find that the key coefficients on *Team cultural similarity* remain largely unchanged compared to Table 6.

Our economically significant and more nuanced results uncovering both positive and negative impacts of *Team cultural similarity* on corporate innovation highlight how our study of the effect of diversity in inventors' cultural values is distinct, both theoretically and empirically, from prior studies that compare first-generation immigrant inventors to native inventors. On a conceptual level, unlike more coarse categorical inventor attributes such as an inventor's first-generation immigrant vs. non-immigrant status, we measure "deeper-level" constructs such as cultural values that are directly related to an individual's values and beliefs, which are more likely to have a stronger impact on the creative process. Moreover, we have testable predictions regarding the role of cultural diversity on the innovation process that are grounded in theory specific to culture. On an empirical level, compared to prior studies such as Bernstein et al. (2022) that focus on *individuals* working across different institutional environments, we study initial *team* formation and compare the innovative performance of *teams* of corporate inventors working *within the same for-profit company* at the same point in time, who have similar access to physical and financial resources (by being at the same firm) and who face a similar technological/competitive landscape (by working at the same time). Importantly, consistent with the relatively low correlation between *Team cultural similarity* and *Team immigrant similarity* (approximately 0.10), we show that the impact of *Team cultural similarity* on team innovation production is unaffected by the inclusion of *Team immigrant similarity* in our empirical analysis.

6.2.2 Team diversity in other inherited characteristics (gender and age)

To further investigate the possibility that other inherited traits can account for our results, we examine whether diversity in co-inventors' gender or age can explain our unique distributional results regarding team cultural diversity.

To construct *Team gender similarity*, we first identify the gender of each inventor (male or female) in the focal team using three steps. First, we use Infutor's Consumer History Plus database, which contains detailed demographic information including gender. We match inventors in the patent database to individuals in the Infutor database according to Step 1 of Appendix IA.1. Second, for unmatched inventors, we use the dataset developed by the USPTO Office of the Chief Economist, as reported in Toole, DeGrazia, Myers, Breschi, Ferrucci, Lissoni, Miguelez, Sterzi, and Tarasconi (2019). Toole et al. (2019) develop a process that uses the (primarily first) name of the inventor and various name–gender dictionaries to attribute the gender of inventors listed on U.S. patents from 1976 to 2016. Third, for the remaining unmatched inventors, we identify their gender using genderize.io and forebears.io, supplemented by manual checks using online sources such as LinkedIn.[58] *Team gender similarity* is then calculated as one minus the Blau (1977) gender diversity measure, where the Blau gender diversity measure = 2 × (1 – [(Percent of team that is female)^2 + (Percent of team that is male)^2]). Higher values of *Team gender similarity* signify less team gender diversity.

To construct *Team age similarity*, we first identify the age of each inventor in the focal team using two sources. First, we use Infutor's Consumer History Plus database, which contains detailed demographic information including the year of birth. Second, for inventors with missing age after utilizing the Infutor database, we use matched inventor age information from the dataset created by Kaltenberg, Jaffe, and Lachman (2023) that is available on Harvard Dataverse. The authors search for age information for U.S. residing inventors from publicly available online web directories, such as Radaris, Spokeo, and Beenverified, based on name and location information from patents. Using these two sources, we obtain age information for 90% of inventors in our sample. For the remaining

[58] Both genderize.io and forebears.io predict gender based on first names. We fill in gender information for inventors with previously missing values when both genderize.io and forebears.io agree. When they disagree or only one source is available, we manually check using online sources such as LinkedIn and Wikipedia.

10% of inventors with missing age information, we interpolate their age based on other inventors with non-missing age from the same technology class and patenting year. The *Team age similarity* measure is then calculated as –1 multiplied by the standard deviation of age among the R&D team members. This measure is transformed so that its values are bounded between zero and one, where higher values indicate higher age similarity or lower age diversity among team members.

Importantly, as shown in Table 8, we find that our previous findings regarding *Team cultural diversity* are largely unaffected by the inclusion of additional team diversity measures based on other inherited inventor traits. In fact, we find that team-level diversity in these other inherited traits (i.e., *Team gender similarity* and *Team age similarity*) appear to have minimal systematic relationship with future team performance. This is consistent with the theory that, unlike diversity in "deeper-level" constructs such as cultural values that are directly related to an individual's values and beliefs, diversity in other inherited traits such as gender or age is less likely to capture fundamental differences in the information processing and communication approaches of individual inventors that in turn can meaningfully affect overall team output (Stahl et al., 2010).

6.3 Additional robustness tests

In this section, we implement several additional tests to verify the robustness of our previously reported results to potential alternative explanations.

6.3.1 Accounting for non-culture related changes at treated inventor teams post-death

One potential issue with our use of co-inventor deaths as an exogenous source of variation in team cultural diversity is that we implicitly assume that these changes in team cultural diversity are not systematically correlated with other contemporaneous changes in team characteristics such as team experience, ability, and technological focus. While we already include an extensive set of time-varying control variables, along with several layers of fixed effects, in all our main treatment effect regressions, we nevertheless augment our analysis in Sections 4.2.3 and 4.2.4 to interact all post-death changes in control variables, computed as the post-death period average minus the pre-death period average, with the post-death and treated dummies. This enables us to further isolate the incremental impact of changes in team cultural diversity on team patenting.

For brevity, we show in Table IA.6 that our results are very similar to those reported in Table 6, which is our most stringent specification of heterogeneous treatment effects around co-inventor deaths. In other words, including a full set of interacted controls in our treatment regressions has minimal impact on our key team culture measures. As discussed earlier, this is likely due to the fact that our team culture measures exhibit very low correlation (i.e., less than 0.05 in absolute terms) with other team characteristics capturing the amount and diversity of a team's acquired professional experiences and ability. This further supports the notion that the death-induced shocks to team cultural diversity that we rely on for identification are not meaningfully correlated with any other simultaneous changes in the experience, specialization, or ability of treated teams in the post-treatment period.

6.3.2 Representativeness of (patenting) inventor team sample

A potential concern with our preceding analysis is that we rely on a corporate R&D team successfully obtaining at least one granted patent during their time working together to identify the team's existence in our linked patent-inventor-team-firm database. Although this is the same general limitation faced by all prior innovation studies that rely on patent databases to identify inventor collaborations (e.g., Jaravel et al., 2018; Li and Wang, 2023), and it does not appear that this issue would significantly bias the results of our quasi-natural experiment involving co-inventor deaths in any particular direction,[59] we nevertheless undertake a novel robustness analysis utilizing scientific publications data.

[59] This is because our initial quasi-natural experiment sample is conditioned on having patented together at least once *before* the focal co-inventor's death exogenously changes the team's cultural composition. Thus, all treated inventor

As discussed in Arora et al., (2021), research conducted by corporate scientists is typically disclosed in scientific publications, even if that research does not ultimately lead to the development of a new patent. This allows us to assess whether additional intra-firm R&D collaborations, which are not captured by existing patent datasets, could potentially introduce any systematic bias to our main results. Specifically, we conduct an extensive manual search in the Web of Science Scientific Publications database for any published journal articles, working papers, and conference proceedings co-authored by a focal deceased corporate inventor within three years of their death. This search enables us to identify teams consisting solely of corporate R&D researchers who have collaborated on published (potentially patentable) firm R&D projects, even though they never developed a granted patent and are thus not observable in our linked patent datasets.[60]

We make two key observations. First, only less than 4% of the deceased treated inventors in our sample are associated with a corporate R&D-focused team that is actively seeking to develop publishable (and potentially patentable) research output around the time of their death, but these teams do not appear in our existing linked databases. This suggests that our reliance on patent databases to identify relevant inventor teams is unlikely to systematically overlook or exclude a certain segment of actually formed corporate R&D teams that could bias our main conclusions. Second, in unreported Kolmogorov-Smirnov tests, we confirm that the distribution of team cultural similarity for these "scientific publication but no patent" corporate R&D teams is not significantly different from the treated inventor teams in our sample involving the same deceased co-inventor. Therefore, while we

---

teams, irrespective of whether their team's shared culture scores exogenously increase or decrease post-death and/or whether the team produces a patent in the post-turnover period or not, remain in the entire treatment effects sample.

[60] To do this, we first collate all papers in the *Web of Science* database where one of the author's affiliations is a U.S. publicly listed firm or one of its subsidiaries. We then use a fuzzy name matching algorithm based on author surname, first initial, and middle initial to identify potential company-affiliated scientific publications co-authored by one of our deceased inventors working at the focal U.S. public firm in the 3 years surrounding their death. We then manually verify the full names and employer affiliation of all authors on the article co-authored by one of our focal deceased inventors using office/email addresses (including whether they match a corporate inventor ID in our existing dataset).

acknowledge that we may not observe all inventor teams ever formed at U.S. public firms, it appears unlikely that we are omitting a significant number of teams of a particular cultural diversity profile from our multi-faceted research design that would substantially alter our main research findings.[61]

6.3.3 Potential replacement of deceased co-inventors

Another potential concern with our research design is whether the treated teams in our sample systematically add new team members after their colleague's death, and if so, who are these newly added members. We conduct several additional (unreported) tests to address this concern. First, we observe a high level of 'stickiness' in inventor team formation, with less than 5% of the treated teams in our sample observed to add a new member in the five years following their colleague's death. Second, we verify that the (observable) characteristics of the treated teams that choose to add a new co-inventor are not significantly different from treated teams that do not add a new co-inventor. Moreover, we observe no significant relationship between changes in treated team cultural similarity and future replacement choices. Finally, our results remain qualitatively unchanged if we exclude any patents developed by these newly re-constituted inventor teams. Taken together, these findings suggest that the potential replacement decisions are unlikely to systematically bias our main results.

VII. CONCLUSION

Using information on all U.S.-based corporate inventors employed at U.S. publicly listed firms over a 40-year period, we conduct a large-scale study to examine how the alignment in cultural values of individual inventors affects their desire to work together in a corporate R&D team setting and how shared cultural values among R&D team members impact the innovation production process and the resulting innovative output.

[61] Relatedly, while it is possible that a team of inventors could move from working together at a publicly listed U.S. firm to a private company during their respective careers, this is a relatively rare phenomenon in our sample and is highly unlikely to meaningfully affect our main conclusions. For example, we do not observe any instances of a treated surviving inventor team subsequently moving to patent together at a private company in our main treatment sample.

First, we show that, even among groups of comparably skilled and experienced R&D co-workers, inventors who share similar cultural values are much more likely to collaborate on new projects. Second, using exogenous shocks to team composition arising from co-inventor deaths, we find that less culturally diverse teams produce a higher overall quantity of patents that tend to exploit existing technologies while more culturally diverse teams produce more risky, exploratory patents with a greater potential for high impact innovations. In other words, our combined empirical results imply that the outputs of less culturally diverse teams tend to fall in the middle of the patenting outcome distribution, while the outputs of more culturally diverse teams tend to fall in the tails of the patenting outcome distribution.

Furthermore, we explore the internal mechanisms of the knowledge integration process, presenting novel micro-level evidence on *how* cultural diversity impacts the nature of team innovation outputs. Specifically, we empirically trace the differing impacts of cultural diversity on innovation to differences in the range of input information sources used during the knowledge integration process and disparities in the efficiency of synthesizing the perspectives of heterogenous team members.

Overall, our results have important implications for the implementation of policies designed to promote corporate innovation in R&D intensive yet diverse workplace environments. For example, the strong homophily biases in inventor collaboration that we document point to the importance of exploration-focused firms enacting policies to incentivize existing employees to work with a more inherently diverse set of R&D team members. Our study also suggests that cultural diversity does not have a uniformly positive or negative relationship with team productivity such that firms may have a different and possibly time-varying optimal mix of diverse vs. homogenous inventor teams. We hope our paper motivates the study of important follow-on questions such as how firms can better nurture more diverse inventor teams and how local labor supply diversity can affect firm and regional economic development.

REFERENCES

Adams, R., and D. Ferreira. 2009. Women in the Boardroom and their Impact on Governance and Performance. *Journal of Financial Economics* 94: 291–309.

Aghion, P., Van Reenen, J., and L. Zingales. 2013. Innovation and Institutional Ownership. *American Economic Review* 103(1): 277–304.

Ahern, K., and A. Dittmar. 2012. The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. *Quarterly Journal of Economics* 127(1): 137–197.

Ahern, K., Daminelli, D., and C. Fracassi. 2015. Lost in Translation? The Effect of Cultural Values on Mergers Around the World. *Journal of Financial Economics* 117, 165-189.

Ahmad, M.F., de Bodt, E., and J. Harford. 2021. International Trade and the Propagation of Merger Waves. *Review of Financial Studies* 34: 4876-4925.

Aihounton, G., and A. Henningsen. 2021. Units of Measurement and the Inverse Hyperbolic Sine Transformation. *Econometrics Journal* 24(2): 334–351.

Arora, A., Belenzon, S., and L. Sheer. 2021. Knowledge Spillovers and Corporate Investment in Scientific Research. *American Economic Review* 111(3): 871–898.

Azoulay, P., Graff Zivin, J., and J. Wang. 2010. Superstar Extinction. *Quarterly Journal of Economics* 125(2): 549–589.

Balsmeier, B., Fleming, L., and G. Manso. 2017. Independent Boards and Innovation. *Journal of Financial Economics* 123: 536–557.

Baumol, W., Schilling, M., and E. Wolff. 2009. The Superstar Inventors and Entrepreneurs: How Were They Educated? *Journal of Economics & Management Strategy* 18(3): 711–728.

Bena, J. and K. Li. 2014. Corporate Innovations and Mergers and Acquisitions. *Journal of Finance* 69(5): 1923–1960.

Bernile, G., Bhagwat, V., and S. Yonker. 2018. Board Diversity, Firm Risk and Corporate Policies. *Journal of Financial Economics* 127: 588–612.

Bernstein, S., Diamond, R., Jiranaphawiboon, A., McQuade, T., and B. Pousada. 2022. The Contribution of High-Skilled Immigrants to Innovation in the United States. Working Paper, NBER.

Beugelsdijk, S., Maseland, R., and A. van Hoorn. 2015. Are Scores on Hofstede's Dimensions of National Culture Stable Over Time? A Cohort Analysis. *Global Strategy Journal* 5: 223–240.

Bisin, A., and T. Verdier. 2011. Economics of Cultural Transmission and Socialization. In: Benhabib, J., Bisin, A., Jackson, M. (Eds.), Handbook of Social Economics, vol. 1A. North-Holland.

Blau, P.M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York, NY: The Free Press.

Brochet, F., Miller, G., Naranjo, P., and G. Yu. 2019. Managers' Cultural Background and Disclosure Attributes. *The Accounting Review* 94(3): 57–86.

Calder-Wang, S., Gompers, P., and K. Huang. 2023. Diversity and Performance in Entrepreneurial Teams. Working Paper, Harvard Business School.

Chui, A.C.W., Titman, S., and K.C. Wei. 2010. Individualism and Momentum Around the World. *Journal of Finance* 65 (1): 361–392.

Corsino, M., Giuri, P., and S. Torrisi. 2019. Technology Spin-offs: Teamwork, Autonomy and the Exploitation of Business Opportunities. *Journal of Technology Transfer* 44(5): 1603–1637.

D'Acunto, F., Tate, G., and L. Yang. 2021. Entrepreneurial Teams: Diversity of Skills and Early-Stage Growth. Working Paper, University of Maryland.

Donaldson, G., and J. W. Lorsch. 1983. *Decision Making at the Top: The Shaping of Strategic Direction*. Basic Books, New York.

Donaldson, J., Malenko, N., and G. Piacentino. 2020. Deadlock on the Board. *Review of Financial Studies* 33(10): 4445–4488.

Doran, K., Gelber, A., and A. Isen. 2022. The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries. *Journal of Political Economy* 130(10): 2501–2533.

Dyck, A., Morse, A., and L. Zingales. 2010. Who Blows the Whistle on Corporate Fraud? *Journal of Finance* 65(6): 2213–2253.

Farre-Mensa, J., Hedge, D., and A. Ljungqvist. 2020. What is a Patent Worth? Evidence from the U.S. Patent "Lottery". *Journal of Finance* 75(2): 639–682.

Fernández, R., and A. Fogli. 2009. Culture: An Empirical Investigation of Beliefs, Work, and Fertility, *American Economic Journal: Macroeconomics* 1: 146–177.

Fernández, R. 2011. Does Culture Matter? In: Benhabib, J., Bisin, A., Jackson, M. O. (Eds.), *Handbook of Social Economics*, vol. 1A. North-Holland, pp. 481–510.

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

Gao, H., and W. Zhang. 2017. Employment Nondiscrimination Acts and Corporate Innovation. *Management Science* 63(9): 2982–2999.

Garlappi, L., Giammarino, R., and A. Lazrak. 2017. Ambiguity and the Corporation: Group Disagreement and Underinvestment. *Journal of Financial Economics* 125(3): 417–433.

Gattari, P. 2005. Determining Inventorship for U.S. Patent Applications. *Intellectual Property and Technology Law Journal* 17(5): 16–18.

Gera, S. 2013. Virtual Teams versus Face to Face Teams: A Review of the Literature. *Journal of Business and Management* 11(2): 1–4.

Giannetti, M., and M. Zhao. 2019. Board Ancestral Diversity and Firm-Performance Volatility. *Journal of Financial & Quantitative Analysis* 54(3): 1117–1155.

Gompers, P., Murharlyamov, V., and Y. Xuan. 2016. The Cost of Friendship. *Journal of Financial Economics* 119(3): 626–644.

Gorton, G., and A. Zentefis. 2023. Corporate Culture as a Theory of the Firm. Working Paper, NBER.

Griffin, D., Li, K., and T. Xu. 2021. Board Gender Diversity and Corporate Innovation: International Evidence. *Journal of Financial and Quantitative Analysis* 56(1): 123–154.

Gruber, J. 1994. Incidence of Mandated Maternity Benefits. *American Economic Review* 84: 622-641

Guiso, L., Sapienza, P., and L. Zingales. 2004. The Role of Social Capital in Financial Development. *American Economic Review* 94(3): 526–556.

Guiso, L., Sapienza, P., and L. Zingales. 2006. Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20, 23-48.

Guiso, L., Sapienza, P., and L. Zingales. 2015. The Value of Corporate Culture. *Journal of Financial Economics* 117(1): 60–76.

Hall, B., Jaffe, A., and M. Trajtenberg. 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Working Paper, NBER 8498.

Harhoff, D., Narin, F., Scherer, F., and K. Vopel. 1999. Citation Frequency and the Value of Patented Inventions. *The Review of Economics and Statistics* 81(3): 511–515.

Hirshleifer, D., Hsu, P.H., and D. Li. 2013. Innovative Efficiency and Stock Returns. *Journal of Financial Economics* 107(3): 632–654.

Hirshleifer, D., Hsu, P.H., and D. Li. 2018. Innovative Originality, Profitability, and Stock Returns. *Review of Financial Studies* 31(7): 2553–2605.

Hofstede, G. 1980. *Culture's Consequences: International Differences in Work-Related Values*. London: Sage Publications.

Hofstede, G., Hofstede, G.J., and M. Minkov. 2010. *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival*. 2nd Edition, McGraw-Hill, London.

Horwitz, S., and I. Horwitz. 2007. The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography. *Journal of Management* 33: 987–1015.

Ishii, J., and Y. Xuan. 2014. Acquirer-Target Social Ties and Merger Outcomes. *Journal of Financial Economics* 112: 344–363.

Islam, E., and J. Zein. 2020. Inventor CEOs. *Journal of Financial Economics* 135(2): 505–527.

Jaffe, A. 1989. Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. *Research Policy* 18: 87–97.

Jaravel, X., Petkova, N., and A. Bell. 2018. Team-Specific Capital and Innovation. *American Economic Review* 108(4–5): 1034–1073.

Jones, B. 2009. The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation getting harder? *Review of Economic Studies* 76(1): 283–317.

Kalmijn, M. 1998. Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology* 24: 395–421.

Kaltenberg, M., Jaffe, A., and M. Lachman. 2023. Invention and the Life Course: Age Differences in Patenting. *Research Policy* 52(1): 104629.

Kang, J., Kim, S., and S. Oh. 2022. Does Board Demographic Diversity Enhance Cognitive Diversity and Monitoring? *The Accounting Review* 97(6): 385–415.

Karolyi, A. 2016. The Gravity of Culture for Finance. *Journal of Corporate Finance* 41: 610-625.

Kerr, W.R. 2008. Ethnic Scientific Communities and International Technology Diffusion. *Review of Economics and Statistics* 90(3): 518–537.

Kim, D., and L. Starks. 2016. Gender Diversity on Corporate Boards: Do Women Contribute Unique Skills? *American Economic Review* 106: 267–271.

Kogan, L., Papanikolaou, D., Seru, A., and N. Stoffman. 2017. Technological Innovation, Resource Allocation and Growth. *Quarterly Journal of Economics* 132(2): 665–712.

Li, K., Mai, F., Shen, R., and X. Yan. 2021. Measuring Corporate Culture Using Machine Learning. *Review of Financial Studies* 34: 3265–3315.

Li, K., and J. Wang. 2023. Inter-Firm Inventor Collaboration & Path-Breaking Innovation: Evidence from Inventor Teams Post-Merger. *Journal of Financial & Quantitative Analysis* 58: 1144–1171.

Liu, X. 2016. Corruption Culture and Corporate Misconduct. *Journal of Financial Economics* 122(2): 307–327.

Lucas, R., and B. Moll. 2014. Knowledge Growth and the Allocation of Time. *Journal of Political Economy* 122(1): 1–51.

Ma, S. 2020. The Life Cycle of Corporate Venture Capital. *Review of Financial Studies* 33(1): 358–394.

Marvel, M., Griffin, A., Hedba, J., and B. Vojak. 2007. Examining the Technical Corporate Entrepreneurs' Motivation. *Entrepreneurship Theory and Practice* 31(5): 753–768.

Merritt, A. 2000. Culture in the Cockpit: Do Hofstede's Dimensions Replicate? *Journal of Cross-Cultural Psychology*, 31(3): 283–301.

Moser, P., Parsa, S., and S. San. 2023. Immigration, Science, and Invention: Evidence from the Quota Acts. Working Paper, New York University.

Mukherjee, S., Romero, D., Jones, B. and B. Uzzi. 2017. The Nearly Universal Link between the Age of Past Knowledge and Tomorrow's Breakthroughs in Science. *Science Advances* 3(4): e1601315.

Nguyen, D., Hagendorff, J., and A. Eshraghi. 2018. Does CEO Cultural Heritage Affect Performance Under Competitive Pressure? *Review of Financial Studies* 31: 97–141.

Ostergaard, C., Timmermans, B., and K. Kristinsson. 2011. Does a Different View create something new? The Effect of Employee Diversity on Innovation. *Research Policy* 40(3): 500–509.

Pagnini, D., and P. Morgan. 1990. Intermarriage and Social Distance among US Immigrants at the Turn of the Century. *American Journal of Sociology* 96: 405–32.

Paikeday, T., Qosja, N., Lim, S., and J. Flock. 2023. A Global Look at the Chief Diversity Officer Landscape. *Russell Reynolds Associates*.

Pan, Y., Siegel, S., and T. Wang. 2017. Corporate Risk Culture. *Journal of Financial and Quantitative Analysis* 52(6): 2327–2367.

Pan, Y., Siegel, S., and T. Wang. 2020. The Cultural Origin of CEOs' Attitudes Toward Uncertainty: Evidence from Corporate Acquisitions. *Review of Financial Studies* 33(7): 2977-3030.

Roberts, M, and T. Whited. 2013. Endogeneity in Empirical Corporate Finance. Chapter 7 in *Handbook on the Economics of Finance* 2: 493–572. Amsterdam: Elsevier.

Romer, P. 1990. Endogenous Technological Change. *Journal of Political Economy* 98: 71–102.

Schein, E. H. 1985. *Organizational Culture and Leadership*. Jossey-Bass, San Francisco.

Schumpeter, J. 1934 (reprinted 1962). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Cambridge, MA: Harvard University Press.

Schumpeter, J. 1939. *Business Cycles*. New York, NY: McGraw-Hill.

Schumpeter, J. 1942. *Capitalism, Socialism and Democracy*. New York: Harper.

Schwartz, S. 1992. Universals in the Content and Structure of Values: Theory and Empirical Tests in 20 Countries. *Advances in Experimental Social Psychology* 25: 1–65.

Seru, A. 2014. Firm Boundary Matters: Evidence from Conglomerates and R&D activity. *Journal of Financial Economics* 111: 381–405.

Shane, S. 1995. Uncertainty Avoidance and the Preference for Innovation Championing Roles. *Journal of International Business Studies* 26, 47–68.

Stahl, G. K., Maznevski, M. L., Voigt, A., and K. Jonsen. 2010. Unraveling the Effects of Cultural Diversity in Teams: A Meta-analysis of Research on Multicultural Work Groups. *Journal of International Business Studies* 41: 690–709.

Tian, X. 2011. The Causes and Consequences of Venture Capital Stage Financing. *Journal of Financial Economics* 101(1): 132–159.

Toole, A., DeGrazia, C., Myers, A., Breschi, S., Ferrucci, E., Lissoni, F., Miguelez, E., Sterzi, V., & G. Tarasconi. 2019. Profile of Women Inventors on U.S. Patents. *USPTO Data Highlights* 2: 1-18.

Van den Steen, E. 2010. Culture Clash: The Costs and Benefits of Homogeneity. *Management Science* 56(10): 1718–1738.

Van Wanrooy, B., Bewley, H., Bryson, A., Forth, J., Freeth, S., Stokes, L., and S. Wood. 2011. *Workplace Employment Relations Study: First Findings*.

Verhoeven, D. 2023. Do Broader Teams Lead to More Innovation? Working Paper, KU Leuven.

Wang, J., Cheng, G., Chen, T., and K. Leung. 2019. Team Creativity/Innovation in Culturally Diverse Teams: A Meta-analysis. *Journal of Organizational Behavior* 40(6): 693–708.

Webber, R. 2007. Using Names to Segment Customers by Cultural, Ethnic, or Religious Origin. *Direct, Data and Digital Marketing Practice* 8: 226–242.

Xiao, T., Makhija, M., and S. Karim. 2022. A Knowledge Recombination Perspective of Innovation: Review and New Research Directions. *Journal of Management* 48(6): 1724–1777.

Zhang, J. 2022. Why Do Innovative Firms Sell Patents? An Empirical Analysis of the Causes and Consequences of Secondary Market Patent Transactions. Working Paper, Boston College.

Zingales, L. 2000. In Search of New Foundations. *Journal of Finance* 55: 1623–1653.

Table 1: Summary statistics

This table reports summary statistics for the entire sample of individual inventors and inventor teams working at U.S. publicly listed firms from 1981 to 2016. Panel A presents descriptive statistics, computed across the entire sample period, for individual U.S.-based inventors working at U.S. publicly listed firms. Panel B outlines descriptive statistics for inventor teams consisting of U.S.-based inventors employed at U.S. publicly listed firms. *Size of inventor team per patent* equals the number of unique inventors listed on a particular patent. *Number of teams per inventor* equals the number of unique teams of which the focal inventor is a team member over the course of the entire sample period. *Distinct co-inventors per inventor* equals the number of unique co-inventors that the focal inventor works with over the course of the entire sample period. Panel C reports pairwise characteristics for all first-time collaborations between pairs of U.S.-based inventors where that first-time collaboration occurs at a U.S. publicly listed firm between 1981 and 2016. Appendix A outlines the definition of all the variables listed.

Panel A: Individual inventor characteristics

	Mean	Median	Std. dev.
Total patents	7.43	3.00	16.79
Years of inventor experience	19.59	18.00	10.90
Individual technical knowledge base	2.47	2.00	2.81
Top 10% inventor	0.19	0.00	0.39
Average backward cites per patent	18.24	9.50	41.64
Average forward cites per patent	16.90	7.50	34.61

Panel B: Inventor team characteristics

	Mean	Median	Std. dev.
Size of inventor team per patent	3.31	3.00	1.57
Number of teams per inventor	3.19	1.00	5.61
Distinct co-inventors per inventor	6.71	4.00	8.85

Panel C: Pairwise characteristics of newly formed co-inventor pairs (at time of first collaboration)

	Mean	Median	Std. dev.
Co-inventor cultural similarity	0.62	0.60	0.22
Co-inventor geographic distance (miles)	224.57	15.31	551.62
Co-inventor difference in avg. forward cites to date per patent	1.11	0.60	2.19
Both top 10% inventors	0.07	0.00	0.26
Co-inventor tech proximity	0.21	0.00	0.35
Co-inventor difference in avg. backward cites to date per patent	0.93	0.58	1.57
Co-inventor difference in years of inventor experience to date	7.01	5.00	6.96
Both have 5+ years of tenure at focal firm	0.18	0.00	0.38

Table 2: Determinants of choice of co-inventor – Main results

This table reports the results of conditional logit models that estimate the factors affecting the choice of co-inventor (see equation (1) in Section 3.1 for further details). The dependent variable is equal to one for all new co-inventor pairwise relationships formed at a U.S. publicly listed firm during the sample period and zero for counterfactual pairs that comprise the comparison/control group. Column (1) uses other inventors working at the same firm in the same year as the new, actually formed co-inventor pair to form the counterfactual control pairs. Column (2) uses other inventors working at the same firm, in the same year and at the same office to form the counterfactual control pairs. Column (3) uses other inventors working in the same division/subsidiary of a firm at the same time to form the counterfactual control pairs. Column (4) uses other inventors working in the same division/subsidiary of a firm, in the same year and at the same office to form the counterfactual control pairs (see Sections 3.2 and 3.3 for further details). Appendix A provides definitions for all independent variables. All regression specifications include group fixed effects that comprise the focal new realized pairing of co-inventors and its (up to) two counterfactual control pairs of potential (but ultimately unchosen) co-inventors. Robust standard errors (clustered at the group level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Counter-factual group	Same firm Same year	Same firm Same year Same office	Same division Same year	Same division Same year Same office
	(1)	(2)	(3)	(4)
Co-inventor cultural similarity	0.278*** (0.005)	0.253*** (0.005)	0.256*** (0.005)	0.270*** (0.006)
Co-inventor geographic distance	-0.403*** (0.001)	-0.125*** (0.001)	-0.358*** (0.001)	-0.110*** (0.001)
Co-inventor difference in average forward cites to date per patent	-0.003*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)
Both top 10% inventors	4.119** (1.788)	1.555*** (0.204)	3.166*** (0.975)	1.466*** (0.189)
Co-inventor tech proximity	1.776*** (0.006)	1.455*** (0.005)	1.586*** (0.006)	1.375*** (0.005)
Co-inventor difference in average backward cites to date per patent	-0.019*** (0.001)	-0.018*** (0.001)	-0.017*** (0.001)	-0.016*** (0.001)
Co-inventor difference in years of experience to date	-0.092*** (0.002)	-0.092*** (0.002)	-0.087*** (0.002)	-0.088*** (0.002)
Both have 5+ years of tenure at focal firm	-0.477*** (0.004)	-0.384*** (0.004)	-0.447*** (0.004)	-0.373*** (0.004)
Group fixed effects?	Yes	Yes	Yes	Yes
No. of observations (mil)	5,265,365	4,835,252	4,812,022	4,399,770
No. of actual pairs (mil)	1,821,150	1,752,127	1,682,644	1,608,614
No. of counter-factual pairs (mil)	3,444,215	3,083,125	3,129,378	2,791,156
Pseudo R^2	0.24	0.05	0.19	0.04

Table 3: Baseline relationship between inventor team diversity and team innovation output

This table reports how diversity in an inventor team's cultural values (namely *Team cultural similarity*) affects the type of innovation output produced by the inventor team, using the ordinary least squares (OLS) regression specification outlined in equation (2) in Section 4.1.1. We measure the quantity of team innovative output each year as *Ln(1+Total patents)*. The average quality of team innovative output is measured as *Ln(1+Average forward citations per patent)*. We measure a team's propensity for producing high impact innovations as *Ln(1+Top 10% cited patents)*. A team is designated as being more (less) focused on exploitative (explorative) innovation if they have a higher average number of backward citations per patent, *Ln(1+Average backward citations per patent)*, and a higher average number of claims per patent, *Ln(1+Average claims per patent)*. Total net innovation output is measured as *Ln(1+Total citation-weighted patents)*. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team size and team geographic diversity). See Appendix A for definitions of all variables used in this analysis. All regression specifications include Firm × Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
Team cultural similarity	0.005**	-0.003*	-0.002**	0.004**	0.003**	-0.000
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	3,385,004	3,385,004	3,385,004	3,385,004	3,385,004	3,385,004
Adjusted R^2	0.19	0.15	0.07	0.16	0.22	0.13

Table 4: Evidence supporting the validity of experimental design using co-inventor deaths

This table reports summary statistics that compare the inventor teams that form the basis of our difference-in-difference tests involving premature co-inventor deaths. Panels A and B focus on comparing inventor teams that suffer a co-inventor death ('treated' teams) with inventor teams active in the same firm at the same point in time that do not experience the loss of a team member ('control' teams) (see Section 4.2.1 for additional details for the construction of control teams). Panel A compares the average output of treated and control firms in the 5 years up to and including the year of the focal co-inventor death while Panel B compares the average pre-treatment characteristics of treated and control teams at the time of the focal co-inventor's death. In contrast, Panels C and D focus on comparing treated inventor teams whose cultural similarity increases post their focal co-inventor's death with treated inventor teams whose cultural similarity does not increase post their focal co-inventor's death. Panel C compares the average output of these treated teams in the 5 years up to and including the year of the focal co-inventor's death while Panel D compares the average pre-treatment characteristics of these treated teams at the time of the focal co-inventor's death. Finally, Panel E compares the observable characteristics of individual deceased inventors at the time of their death for treated teams whose cultural similarity does increase post the focal co-inventor's death compared to the characteristics of focal deceased inventors at treated teams whose cultural similarity does not increase after the focal co-inventor's death. See Appendix A for definitions of all dependent and independent variables listed. *, ** and *** indicate that the difference in means is statistically significant at the 10%, 5% and 1% level respectively.

Panel A: Average annual output in the 5-year period prior to the focal co-inventor's death

	Treated team Mean (1)	Control team Mean (2)	Difference (1) – (2)
Total patents	0.27	0.26	0.01
Average forward cites per patent	0.30	0.29	0.01
Top 10% cited patents	0.04	0.04	0.00
Average backward cites per patent	0.36	0.35	0.01
Average claims per patent	0.26	0.26	0.00
Total citation-weighted patents	0.40	0.38	0.02

Panel B: Team characteristics at the time of the focal co-inventor's death

	Treated team Mean (1)	Control team Mean (2)	Difference (1) – (2)
Team cultural similarity	0.58	0.59	-0.01
Team geographic diversity	260.55	280.72	-20.18
Scope of team technical knowledge	6.96	7.01	-0.05
Team average of individual technical knowledge base	2.82	2.84	-0.02
Team technical knowledge overlap	0.63	0.64	-0.01
Team average total number of patents in last 5 years	6.81	6.87	-0.06
Team average forward cites to date per patent	1.17	1.16	0.01
Team average inventor experience to date	8.48	8.35	0.13
Team average backward cites to date per patent	1.22	1.26	-0.05

Panel C: Average annual output in the 5-year period prior to the focal co-inventor's death

	Treated team (Cultural sim. Does increase) Mean (1)	Treated team (Cultural sim. Does not increase) Mean (2)	Difference (1) – (2)
Total patents	0.30	0.32	-0.02
Average forward cites per patent	0.34	0.34	0.00
Top 10% cited patents	0.06	0.06	0.00
Average backward cites per patent	0.36	0.41	-0.05
Average claims per patent	0.27	0.28	-0.01
Total citation-weighted patents	0.48	0.50	-0.02

Panel D: Team characteristics at the time of the focal co-inventor's death

	Treated team (Cultural sim. Does increase) Mean (1)	Treated team (Cultural sim. Does not increase) Mean (2)	Difference (1) – (2)
Team cultural similarity	0.56	0.57	-0.01
Team geographic diversity	269.64	258.42	11.21
Scope of team technical knowledge	6.85	7.14	-0.29
Team average of individual technical knowledge base	2.87	2.80	0.07
Team technical knowledge overlap	0.64	0.64	0.01
Team average total number of patents in last 5 years	6.61	7.08	0.47
Team average forward cites to date per patent	1.24	1.19	0.05
Team average inventor experience to date	8.90	8.57	0.33
Team average backward cites to date per patent	1.24	1.25	-0.01
Percentage of unique tech class knowledge lost due to co-inventor death	0.07	0.06	0.01
Change in team technical knowledge overlap due to co-inventor death	-0.04	-0.04	0.00

Panel E: Individual deceased inventor characteristics at the time of their death

	Treated team (Cultural sim. Does increase) Mean (1)	Treated team (Cultural sim. Does not increase) Mean (2)	Difference (1) – (2)
Age at time of death (years)	48.28	47.80	0.48
Individual technical knowledge base	3.00	2.94	0.06
Total patents in 5 years pre-death	7.36	7.47	-0.11
Average forward cites per patent	1.32	1.27	0.05
Top 10% cited patents	0.47	0.46	0.01
Years of inventor experience	9.13	8.98	0.15
Average backward cites per patent	1.27	1.30	-0.03
Average claims per patent	1.07	1.09	-0.01

Table 5: Innovation output around exogenous co-inventor turnover – Triple difference-in-difference tests

This table reports the change in innovative output for treated teams around the death of a co-inventor relative to control teams as defined in equation (3) in Section 4.2.1. *After*$_{i,t}$ is an indicator variable equal to one for all years after the focal co-inventor's death and zero otherwise. *Treated team*$_i$ is an indicator variable equal to one for all teams that suffer the death of a team member and zero otherwise. *Team cultural similarity change*$_i$ is a continuous variable equal to the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. We measure the quantity of team innovative output each year as *Ln(1+Total patents)*. The average quality of team innovative output is measured as *Ln(1+Average forward citations per patent)*. We measure a team's propensity for producing high impact innovations as *Ln(1+Top 10% cited patents)*. A team is designated as being more (less) focused on exploitative (explorative) innovation if they have higher *Average backward citations per patent* and a higher *Average number of claims per patent*. Total net innovation output is measured as *Ln(1+Total citation-weighted patents)*. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team geographic diversity). See Appendix A for definitions of all variables used in this analysis. All regression specifications include Team and Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
After$_{i,t}$	-0.213***	-0.173***	-0.031***	-0.218***	-0.204***	-0.193***
	(0.003)	(0.004)	(0.002)	(0.004)	(0.003)	(0.004)
After$_{i,t}$ × Treated team$_i$	0.006*	0.007	0.000	0.016*	0.007**	0.007
	(0.003)	(0.005)	(0.001)	(0.009)	(0.003)	(0.004)
After$_{i,t}$ × Treated team$_i$ ×	0.055**	-0.016	-0.031***	0.048**	0.047**	-0.011
Team cultural similarity change$_i$	(0.026)	(0.026)	(0.011)	(0.024)	(0.020)	(0.032)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	99,830	99,830	99,830	99,830	99,830	99,830
Adjusted R^2	0.13	0.09	0.07	0.11	0.10	0.10

Table 6: Heterogeneous treatment effects around co-inventor deaths (treated teams only)

This table reports the change in innovative output for treated teams only around the death of a co-inventor (see equation (4) in Section 4.2.4 for further details). $After_{i,t}$ is an indicator variable equal to one for all years after the focal co-inventor's death and zero otherwise. *Team cultural similarity change*$_i$ is a continuous variable equal to the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. We measure the quantity of team innovative output each year as *Ln(1+Total patents)*. The average quality of team innovative output is measured as *Ln(1+Average forward citations per patent)*. We measure a team's propensity for producing high impact innovations as *Ln(1+Top 10% cited patents)*. A team is designated as being more (less) focused on exploitative (explorative) innovation if they have higher *Average backward citations per patent* and a higher *Average number of claims per patent*. Total net innovation output is measured as *Ln(1+Total citation-weighted patents)*. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team geographic diversity). See Appendix A for definitions of all variables used in this analysis. All regression specifications include Individual dead co-inventor fixed effects, Team fixed effects and Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
$After_{i,t}$	-0.266***	-0.222***	-0.054***	-0.261***	-0.233***	-0.265***
	(0.006)	(0.006)	(0.003)	(0.006)	(0.004)	(0.008)
$After_{i,t}$ × Team cultural similarity change$_i$	0.064**	-0.005	-0.026**	0.059**	0.056***	-0.001
	(0.027)	(0.026)	(0.012)	(0.028)	(0.020)	(0.034)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Dead co-inventor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	57,392	57,392	57,392	57,392	57,392	57,392
Adjusted R^2	0.14	0.09	0.06	0.11	0.09	0.11

Table 7: Mechanism tests

This table reports on how diversity in an inventor team's cultural origins (namely *Team cultural similarity*) affects the type of input information utilized by inventor teams in their innovation production activities as well the number of subsequent patents produced by an inventor team in the initial few years after the team's first patent together (see equation (5) in Section 5.1 and equation (6) in Section 5.2 for further details). The dependent variables in columns (1) to (4) are *Team average new cite ratio*, *Percentage of back cites to 'new' foreign patents*, *Percentage of back cites to 'new' other references, and Percentage of back cites to 'new' mature U.S. patents*, respectively. The dependent variables in columns (5) and (6) are the natural log of one plus the number of patents produced by a team in the one year or two year period after the team's first patent together, respectively. *Size of team's total existing knowledge base across all information sources* is the natural log of one plus the number of U.S. patents, foreign patents, and other references in the team's existing knowledge base. *Degree of focus on exploitative innovation by team's first joint patent* equals the natural log of one plus the number of total backward citations to U.S. patents, foreign patents, and other references by the focal team's first joint patent together. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team size and team geographic diversity). See Appendix A for definitions of all variables used in this analysis. All regression specifications include Firm × Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of intermediate mechanism-related outcome		Type of input information resources used by an inventor team			Time until team's next subsequent patent	
Dependent variable	Team average new cite ratio	% of back cites to 'new' foreign patents	% of back cites to 'new' other references	% of back cites to 'new' mature U.S. patents	Ln(1+Num patents within 1 year)	Ln(1+Num patents within 2 years)
	(1)	(2)	(3)	(4)	(5)	(6)
Team cultural similarity	-0.010***	-0.005***	-0.023***	0.016***	0.007***	0.008***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Scope of team's existing knowledge base across all information sources	-0.059***	-0.000	-0.016***	-0.028***	0.043***	0.087***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Degree of focus on exploitative innovation by team's first joint patent					0.014***	0.014***
					(0.000)	(0.001)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of team-level observations	635,171	635,171	635,171	635,171	635,171	635,171
Adjusted R^2	0.28	0.15	0.23	0.25	0.09	0.12

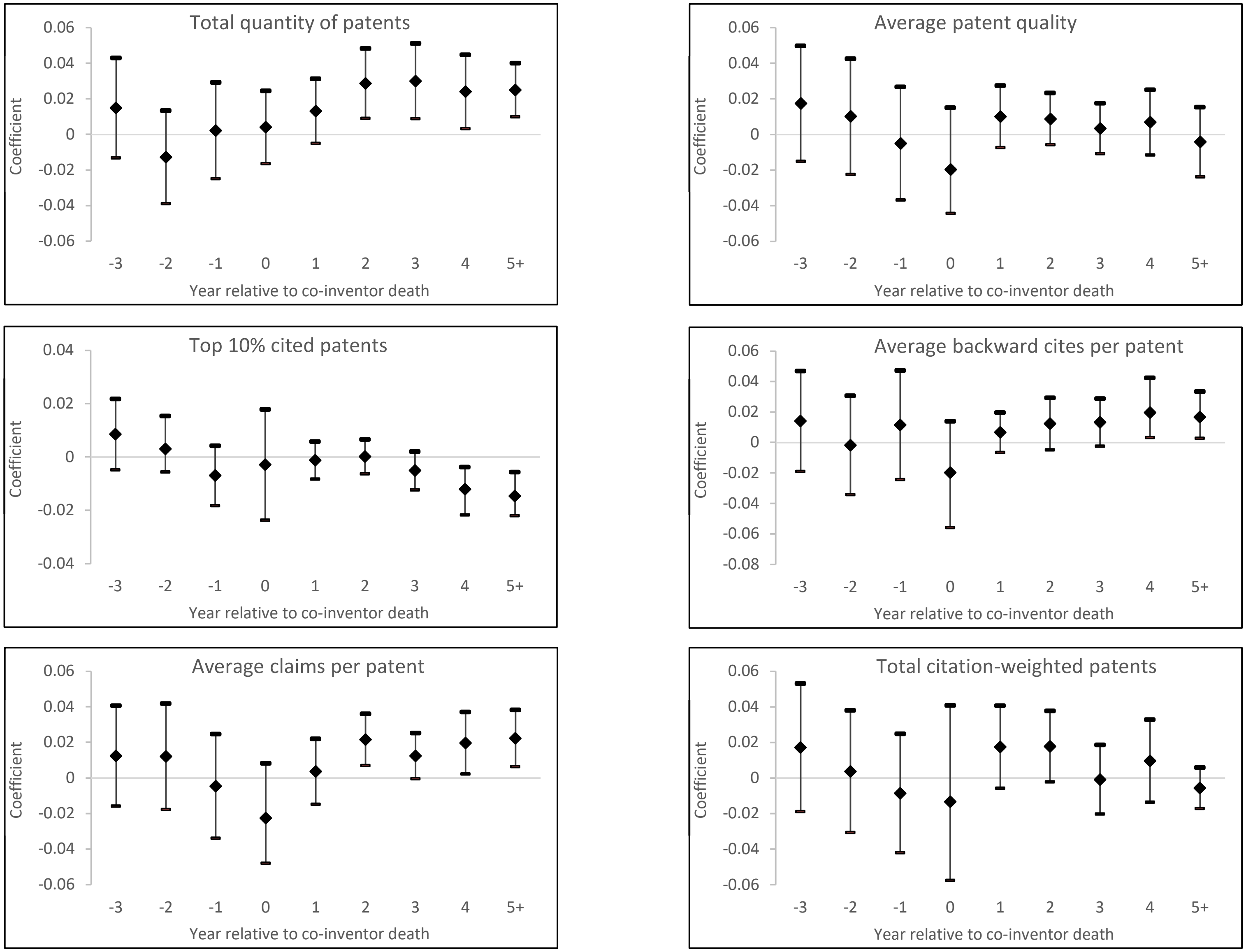
Table 8: Heterogeneous treatment effects around co-inventor deaths (treated teams only) – Inclusion of additional diversity controls

This table reports the change in innovative output for treated teams only around the death of a co-inventor (see equation (4) in Section 4.2.4 for further details). *After$_{i,t}$* is an indicator equal to one for all years after the focal co-inventor's death and zero otherwise. *Team cultural similarity change$_i$* equals the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. The innovation outcome measures are: (1) *Ln(1+Total patents)*; (2) *Ln(1+Average forward citations per patent)*; (3) *Ln(1+Top 10% cited patents)*; (4) *Ln(1+Average backward citations per patent)*; (5) *Ln(1+ Average number of claims per patent)* and (6) *Ln(1+Total citation-weighted patents)*. *Team immigrant similarity change$_i$* is a continuous variable equal to the surviving team's immigrant similarity post the focal co-inventor's death minus the pre-treatment team's immigrant similarity. *Team gender similarity change$_i$* is a continuous variable equal to the surviving team's gender similarity post the focal co-inventor's death minus the pre-treatment team's gender similarity. *Team age similarity change$_i$* is a continuous variable equal to the surviving team's age similarity post the focal co-inventor's death minus the pre-treatment team's age similarity. As outlined in Section 2.3 and Appendix A, we include controls for the size and diversity of a team's knowledge base, the amount and diversity of team members' skills, and the amount and diversity of team members' experience, as well as other controls (team geographic diversity). All regressions include Individual dead co-inventor fixed effects, Team fixed effects and Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
After$_{i,t}$	-0.267***	-0.223***	-0.054***	-0.262***	-0.234***	-0.266***
	(0.006)	(0.006)	(0.003)	(0.006)	(0.004)	(0.008)
After$_{i,t}$ × Team cultural similarity change$_i$	0.061**	-0.004	-0.026**	0.057**	0.054***	-0.001
	(0.027)	(0.026)	(0.012)	(0.029)	(0.020)	(0.035)
After$_{i,t}$ × Team immigrant similarity change$_i$	0.014	0.001	0.000	0.024	0.010	0.006
	(0.014)	(0.014)	(0.008)	(0.018)	(0.009)	(0.019)
After$_{i,t}$ × Team gender similarity change$_i$	0.008	0.009	-0.005	0.012	0.013	0.002
	(0.019)	(0.014)	(0.008)	(0.013)	(0.014)	(0.024)
After$_{i,t}$ × Team age similarity change$_i$	0.043	0.040*	0.002	0.058**	0.052***	0.047
	(0.027)	(0.021)	(0.011)	(0.026)	(0.019)	(0.029)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Dead co-inventor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Team and Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	57,392	57,392	57,392	57,392	57,392	57,392
Adjusted R^2	0.14	0.09	0.06	0.11	0.09	0.11

Figure 1: Assessment of parallel trends assumption for triple differences-in-differences analysis

This figure plots the coefficients (and 95% confidence intervals) on the triple interaction term *{Year relative to death$_{i,t}$}* × *Treated team$_i$* × *Team cultural similarity change$_i$* from 3 years prior to 5+ years post the focal co-inventor death (estimated based on equation (3) as described in Section 4.2.1).



Appendix A: Variable definitions

Variable	**Description**
Panel A: Individual inventor and inventor team innovation output	
Total patents	The number of patents filed (and subsequently granted) in a given year.
Average forward cites per patent	The average of the scaled forward citations of each patent applied for (and subsequently granted) in that year. Scaled forward citations is calculated as the number of citations that a patent receives divided by the average number of citations made to patents applied for in the same year and CPC technology sub-class.
Top 10% cited patents	A patent is classed as a top 10% cited patent if it receives citations from other patents that place the focal patent in the top decile of patents in the same application year and CPC technology sub-class.
Average backward cites per patent	The average of the scaled backward citations of each patent applied for (and subsequently granted) in that year. Scaled backward cites equals the number of citations made by the focal patent divided by the average backward citations made by patents applied for in the same year and CPC sub-class.
Average claims per patent	The average of the number of scaled claims made by each patent applied for (and subsequently granted) in that year. Scaled claims equals the number of claims that the focal patent makes divided by the average number of claims made by patents applied for in the same year and CPC technology sub-class.
Total citation-weighted patents	The total sum of each granted patent (if any) multiplied by that patent's scaled forward citations, following Kogan et al. (2017).
Panel B: Individual inventor and inventor team characteristics	
Total patents to date	A count of the number of patents granted to an inventor over their career up to the focal year.
Years of inventor experience to date	The number of years between the application date of the first (subsequently granted) patent that an inventor ever applies for and the current focal year.
Individual technical knowledge base	A count of the number of distinct CPC technology sub-classes that an inventor patents in over the course of their career up to the focal year.
Top 10% inventor	A dummy variable equal to one for inventors whose total number of patents developed to date place them in the top decile (10%) of all inventors in the USPTO universe and zero otherwise.
Average backward cites per patent	Equals the mean of the scaled backward citations made across all the patents developed over an inventor's career up to the focal year.
Average forward cites per patent	Equals the mean of the scaled forward citations received up until (and including) year $t-1$ across all the patents developed over an inventor's career to date.
Panel C: Pairwise inventor characteristics for selection tests	
Co-inventor cultural similarity	Calculated as –1 multiplied by the Euclidean pairwise distance between two inventors' cultural values based on Hofstede's framework, where pairwise distance is calculated using the following formula: $Pairwise\ distance_{ij} = [(IDV_{Inventor\ i} - IDV_{Inventor\ j})^2 + (PDI_{Inventor\ i} - PDI_{Inventor\ j})^2 + (UAI_{Inventor\ i} - UAI_{Inventor\ j})^2 + (MAS_{Inventor\ i} - MAS_{Inventor\ j})^2 + (LTO_{Inventor\ i} - LTO_{Inventor\ j})^2 + (IVR_{Inventori} - IVR_{Inventor\ j})^2]^{\frac{1}{2}}$ This measure is transformed so that its values are bound between zero and one, where higher values indicate a higher degree of cultural similarity between inventor i and inventor j. See Section 2.3.2 and Internet Appendices IA.1 and IA.2 for more details.

Variable	Description
Panel C: Pairwise inventor characteristics for selection tests (cont.)	
Co-inventor geographic distance	Natural log of the geodetic distance in miles between the two inventors' then current locations (where inventor locations are identified from patent filings).
Co-inventor difference in average forward cites to date per patent	Equals the absolute difference between each inventor's average forward cites to date per patent. Average forward cites to date per patent equals the mean of the scaled forward citations received across all the inventor's patents to date.
Both top 10% inventors	A dummy variable equal to one when both inventors classify as top 10% inventors.
Co-inventor technological proximity	Cosine similarity measure of the degree of overlap in CPC patent technology classes between patents that are applied for (and subsequently granted) to inventor 1 up to and including year $t-1$ and the similarly defined patents of inventor 2 up to and including year $t-1$. This tech proximity measure is bound between 0 and 1.
Co-inventor difference in average backward cites per patent	Equals the absolute difference between each inventor's average backward cites per patent. Average backward cites per patent equals the mean of the scaled back cites received across all the inventor's patents to date.
Co-inventor difference in years of inventor experience to date	The absolute difference between each inventor's years of inventor experience to date. Years of inventor experience equals the number of years between the application date of the first (granted) patent that an inventor ever applies for and the current year.
Both have 5+ years of tenure at the focal firm	A dummy variable equal to one when both inventors have a tenure of over 5 years at the focal U.S. public firm, and zero otherwise. The tenure of an inventor at a firm is the number of years between the application date of the first (granted) patent that an inventor applies for while working at the focal employer and the current year.
Panel D: Inventor team control variables for treatment and mechanisms tests	
Team cultural similarity	To measure cultural similarity within a team of N inventors, we compute the average of pairwise distances in all pairs of team members using the formula: $$-1 \times \frac{\sum_{i,j} Pairwise\ distance_{ij}}{\frac{N(N-1)}{2}} \quad \forall i < j$$ where $Pairwise\ distance_{ij}$ is the distance between each pair of inventors (i, j) as defined above. The measure of team cultural similarity is transformed so that its values are bound between zero and one, where higher values indicate higher cultural similarity or less cultural diversity. See Section 2.3.2 and Internet Appendices IA.1 and IA.2 for further details.
Controls for the size and the diversity of the inventor team's initial knowledge base	
Scope of team technical knowledge	Count of distinct CPC sub-classes that at least one team member has patented in prior to year t (or number of unique tech classes covered by team's collective prior patents)
Team technical knowledge overlap	The overall team-level average of each pairwise *Co-inventor technological proximity* value (based on each pair of inventor's patents to date).
Team average of individual technical knowledge bases	Average of each team member's tech class experience to date (based on the number of distinct CPC sub-classes that an inventor patents in over their career to date).
Team diversity in individual technical knowledge bases	The standard deviation in the technology class experience/individual technical knowledge bases of each team member to date.
Controls for the amount and diversity of an inventor team's skill	
Team average total number of patents to date	The average of each team member's total number of (eventually granted) patents filed to date.
Team diversity in total number of patents to date	The standard deviation in the total number of patents developed by each team member to date.

Variable	Description
Panel D: Inventor team control variables for treatment and mechanisms tests (cont.)	
Controls for the amount and diversity of an inventor team's skill (cont.)	
Team average forward citations to date per patent	The average of each team member's average forward citations to date per patent.
Team diversity in average forward citations to date	The standard deviation in the average forward citations to date per patent of each team member.
Controls for the amount and diversity of an inventor team's experience	
Team average inventor experience to date	The average of each team member's inventor experience to date.
Team diversity in inventor experience to date	The standard deviation in the inventor experience acquired by each team member to date.
Team average backward citations per patent	The average of each team member's average backward citations per patent.
Team diversity in average backward citations	The standard deviation in the average backward citations per patent of each team member.
Other controls	
Team size	The number of inventors in the focal inventor team.
Team geographic diversity	The average of the pairwise co-inventor geographic distances between all members of the inventor team.
Team's total existing knowledge base in year t	The combination of every U.S. patent, foreign patent, or other reference material that has been previously cited or created by at least one team member prior to year t.
Scope of team's existing knowledge base across all information sources	The natural log of one plus the team's total existing knowledge base in year t.
Degree of focus on exploitative innovation by team's first joint patent	The natural log of one plus the number of total backward citations made by the team's first joint patent together to U.S. patents, foreign patents, *and* other references.
Panel E: Intermediate mechanism-related outcome variables	
Team average new cite ratio	The average 'new cite ratio' across all patents ever developed by an inventor team, where each patent's new cite ratio equals the total number of backward cites to 'new knowledge' (i.e., cites to U.S. patents, foreign patents, or 'other references' that are outside of the team's 'existing knowledge base') divided by the total number of backward citations made to U.S. patents, foreign patents and 'other references'.
Percentage (%) of back cites to 'new' foreign patents	The team-level average of the ratio of all backward citations on team patents to 'new' foreign patents (i.e., cites to foreign patents that are outside of the team's existing knowledge base) divided by the total number of backward citations to any source.
Percentage (%) of back cites to 'new' other references	The team-level average of the ratio of all backward citations on team patents to 'new' other references (i.e., cites to other references that are outside of the team's existing knowledge base) divided by the total number of backward citations to any source.
Percentage (%) of back cites to 'new' mature U.S. patents	The team-level average of the ratio of all backward citations on team patents to 'new' mature U.S. patents (i.e., cites to U.S. patents that are aged 10 years or older that are outside of the team's existing knowledge base) divided by the total number of backward citations to any source.
Number of team patents within 1 year/(2 years) after first patent together	The natural logarithm of one plus the number of (eventually granted) patents applied for by the focal team within the one-year/(two-year) period after the application date of the focal team's first patent together.

The Internet Appendix of

"Shared Culture and Technological Innovation: Evidence from Corporate R&D Teams"

Appendix IA.1: Procedure to Trace Inventors' Family Trees

To determine an inventor's inherited cultural values, we identify all relatives from each inventor's family tree using the following eight-step procedure.

Step 1. We match inventors in the patent database to individuals in the Infutor database as follows:

i. Matches are performed based on first name, middle name, last name, gender, city and state of residence. We require that all cities and states associated with the inventor are matched with residence locations in Infutor.

ii. To facilitate more accurate matches, we correct for spelling errors in names and standardize city names using USPS standard city names prior to matching.

iii. For female inventors, we also use alternative last names (i.e., maiden names) provided by Infutor for matching.

iv. If multiple matches are found for an inventor, we include an additional restriction that the patent application year is between the starting and ending address years in Infutor while allowing plus or minus one year.

v. If no matches are found from Step 1(i), we relax some conditions including:

   1. Instead of requiring non-missing middle names, we allow for matches with non-conflicting middle names (i.e., one middle name is missing or both middle names are missing).

   2. Instead of requiring a match on all cities and states in which an inventor is located at the time of patenting, we require matching of at least one set of city and state between the patent data and the Infutor data.

3. Instead of requiring perfect matching of first and last names, we accommodate variations of first names such as nick names and allow for small errors between names (e.g., one letter difference).

vi. We uniquely matched 1.1 million inventors using Infutor data, which represents 73% of the 1.5 million U.S.-based inventors (including those that work in both private and public companies) that patented between 1981 and 2016.

Step 2. Using the Infutor database, we find older relatives of the matched 1.1 million inventors from Step 1 as follows:

i. Older relatives are defined as household members who were born before the inventors and who were residing at the same household at the same time sharing the same last names as the inventors. We require same last names to prevent matching with unrelated individuals such as roommates.

ii. We use inventors' last names and alternative last names (i.e., maiden names) to capture relatives on both the father's side and the mother's side.[62]

iii. We perform three iterations of matching to obtain multiple generations of relatives.[63] This step produces 8.3 million matched relatives.

Step 3. To supplement the list of relatives from Infutor and to establish whether a deceased inventor satisfies our treatment sample inclusion requirements detailed in Section 4.2.1, we manually search for 5,647 deceased inventors in the LexisNexis Public Records

[62] For example, if Mary Smith changed her name to Mary Johnson after marriage, then when we track Mary Smith's residences going back in history, we would be able to capture Mary Smith's relatives using household members sharing the same last name "Smith".

[63] Unlike Census data, it is not possible to perfectly identify parents, grandparents, and great grandparents of inventors in Infutor without information on relationship between household members (e.g., relatives can be older siblings, uncles, and aunts living in the same household). Whenever possible, we use age and information on parents from BUNMD to identify direct relatives and drop other relatives such as siblings, uncles, and aunts.

database. LexisNexis provides demographic information and identifies potential relatives based on household members at individuals' current and past residences. We first manually search for these inventors in LexisNexis using their names and locations (inventor city and state from the patent database) and use provided information on employment history, business associates, education, and email addresses to identify the correct LexisNexis record. For each inventor and the associated list of potential relatives, we save their demographic information including names, alternative names, year and month of birth, and the first five digits of social security number.

Step 4. We match inventors and their relatives from Steps 1 to 3 to the Berkeley Unified Numident Mortality Database (BUNMD) based on first name, middle name, last name, and the first five digits of their social security number. BUNMD covers more than 49 million death records from 1988 to 2005 and contains information that Infutor does not have such as individuals' place of birth, names of parents, and race. After matching inventors and their relatives from Infutor to BUNMD, we obtain 37,567 unique individuals who were born in a foreign country. These foreign-born persons are saved and further processed in Step 8.

Step 5. We drop the matched inventors and relatives (from Steps 1 to 3) who were born after 1940 or identified as foreign-born from Step 4, which yields 1.2 million individuals. Around 45% of the matched inventors from Step 1 have at least one relative who was born in 1940 or earlier. For the 1.2 million matched inventors and relatives who were born in 1940 or earlier, we match them to the full-count 1940 U.S. Census database as follows:

i. For around 20% of the sample with BUNMD information, we match these individuals to Census records based on first name, middle name, last name, gender, birth year, birthplace, race, and names of parents. For each individual, we keep the best match.

ii. For the remaining sample without BUNMD information, we match these individuals to Census records based on first name, middle name, last name, gender, birth year, and SSN issue state. We match SSN issue state from Infutor to state of birth or state of residence in the 1940 Census.[64] For each individual, we keep the best match.

iii. Of the 1.2 million inventors and relatives, 855,941 (71%) of them are matched to the 1940 Census database. For these individuals, we find all members of their households and only keep the individuals themselves, their parents, and grandparents on both the father's side and the mother's side. We identify parents and grandparents based on relationship information provided in the 1940 Census. This process yields 4.14 million unique Census household members.

iv. We do not need to match a 1940 Census household member with earlier Census records if the member was foreign-born or both of his or her parents were foreign-born. This applies to 414,417 unique 1940 Census household members, which we save and further process in Step 8. For the remaining 3.7 million 1940 Census household members, we keep tracing the ancestry origin for them using the 1930 Census database.

Step 6. We match the 3.7 million 1940 Census household members from Step 5 to the 1930 U.S. Census database using last name, middle name, first name, gender, race, birthplace, birth year, mother's birthplace, and father's birthplace. For the individuals matched to the 1930

[64] For reference, using BUNDMD data, we find that the SSN issue state is the same as the state of birth in 78% of cases.

Census database, we find all members of their households and only keep the individuals themselves, their parents, and grandparents on both the father's side and the mother's side. We identify parents and grandparents based on relationship information provided in the 1930 Census. Like the previous step, we save household members who were foreign-born or both of their parents were foreign-born in a separate file, and keep tracing the ancestry origin for the remaining household members using the 1920 census dataset.

Step 7. We repeat Step 6 for 1920, 1910, 1900, 1880, 1870, 1860, and 1850 full count Census datasets. After each iteration of matching is completed, we save the foreign-born matched household members and keep the remaining U.S.-born household members for the next round of matching.

Step 8. We collect all the foreign-born household members from the matched 1850 to 1940 Census datasets from Steps 5 to 7, which yields 2.9 million individuals who were foreign-born, or both of their parents were foreign-born. These individuals are combined with the 37,567 foreign-born individuals identified using BUNMD from Step 4. Using the country of birth information on these foreign-born individuals from inventors' family trees, we define cultural values for each inventor as follows:

i. If the inventor was born in a foreign country, we define the country of birth as the inventor's country of origin and use the six-dimension Hofstede cultural values in that country as the inventor's inherited cultural values.

ii. If the inventor was not born in a foreign country and the inventor has at least one matched foreign-born relative, then for each of the six Hofstede cultural dimensions, we take the average values across all the foreign-born relatives of the inventor.[65] We use these averaged Hofstede cultural values as the inventor's inherited cultural values.

iii. Using the method above, we obtain inherited cultural values for 322,687 unique inventors, which represents around 30% of the matched 1.1 million inventors from Step 1 and around 22% of 1.5 million U.S. inventors that patented between 1981 and 2016.

[65] For robustness, we also used a weighted average measure that gives a higher weight to relatives closer to the inventor in the family tree. We obtain very similar results using this alternative measure.

Appendix IA.2: Procedure to Determine Inventors' Countries of Origin and Inherited Cultural Values

We use a novel, multi-layered approach to identify inventors' countries of origin and determine their inherited cultural values.

Step 1. We use Infutor's Consumer History Plus (CRD4) database, LexisNexis Public Records, Berkeley Unified Numident Mortality Database (BUNMD), and the restricted full count U.S. Census records from 1850 to 1940 from the Minnesota Population Center to track down relatives from each inventor's family tree to determine the inventor's inherited cultural values based on exact countries of ancestry. Using the procedure described in Appendix IA.1, we obtain inherited cultural values based on all relatives' exact country of ancestry for 322,687 unique inventors.

Step 2. To supplement precisely determined countries of ancestry from Step 1, we manually search for inventors and determine their countries of origin for the 4,000 most productive inventors who are not matched from step 1. For these inventors, we search by hand using LinkedIn and other biographical resources such as the Marquis Who's Who database, corporate web profiles and Wikipedia. Using available information on country of birth, parents' country of birth, location of their high school or bachelor granting institution, languages spoken, and pictures, we are able to determine the country of origin for 2,013 inventors.

Step 3. For the inventors that are matched to the Infutor database from Step 1 of the procedure described in Appendix IA.1 that do not yet have inherited cultural values, we identify inventors who are first-generation immigrants as individuals who are more than 20 years old when assigned a social security number (SSN) following Bernstein et al. (2022).

i. To estimate when inventors received their SSN, we decode SSN provided in Infutor using https://www.ssn-verify.com, which provides information on the issue state, the first year and last year issued for each set of five-digit social security number. These first-generation immigrant inventors account for about 16% of total inventors.

ii. Since most of these inventors (and their relatives) will not be in the 1850 to 1940 U.S. Census records, we determine their country of origin based on their surnames using several public and commercial sources that are more suitable for more recent immigrants:[66] Specifically, we use four sources: Nationalize.io (https://nationalize.io), Forebears (https://forebears.io), NamePrism (https://www.name-prism.com), and Dictionary of American Family Names from Oxford Reference. For each inventor, we find the most likely country of origin according to each of the four sources based on the inventor's surname and a match is recorded if the most likely country of origin appears in two or more sources.[67]

Step 4. For the remaining inventors who are most likely U.S.-born second or higher generation immigrants, we identify their countries of ancestry based on their surnames using the 1850 to 1940 U.S. Census records, following the methodology of Liu (2016). For female inventors, we use their maiden names provided by Infutor for matching to maximize accuracy whenever possible. We use U.S. Census records for these inventors because

[66] According to the Pew Research Center (2015), immigrants that entered the U.S. during the latest wave of immigration post-1965 mostly come from Latin America and Asia countries. The top three source countries are Mexico, China, and India.

[67] For robustness, we use both first names and last names for matching, and we also use weighted percentage of countries instead of the most likely country of origin and obtain similar results.

they likely have relatives in the 1850 to 1940 U.S. Census. However, we are not able to track their family trees directly using the procedure described in Appendix IA.1 because we cannot find a relative for them who was born in 1940 or earlier using Infutor or LexisNexis, which has better coverage for individuals born in more recent decades. Essentially, instead of tracking down direct relatives from inventors' family trees as in Appendix IA.1, we find all possible relatives for these inventors as individuals who share the same surnames using U.S. Census records as follows:

i. We use U.S. Census records from 1850 to 1940 to identify each inventor's country of origin based on all possible relatives sharing the same surname. Specifically, we restrict the dataset to first- and second-generation immigrants whose country of birth or father's country of birth is outside of the United States. We then link each unique surname from the Census records to its most frequently associated country of birth or father's country of birth. For instance, the surname "Wong" is linked to China because 97.2% of immigrants with the same surname are from China. To further verify these matches, we also use the surname-ancestry country matching list from a commercial database, Origins Info Ltd., which is a well-known commercial vendor of name classification services. Origins Info processed the list of surnames using its proprietary database constructed based on sources such as the American Dictionary of Family names and international telephone directories. The accuracy of Origins Info's matching has been validated in prior studies (Webber, 2007).

ii. To create the final matching list, we do the following.

    1. First, we record matches where the most frequently associated country of birth from census records is the same country of origin identified by Origin Info.

2. Second, we keep surnames for which the most frequently associated country of birth appears in more than 75% of the census records.
3. Third, for surnames with different census and Origin Info country of origin, we hand-check their country of origin using sources such as dictionaries and ancestry.com, which provides a distribution of U.S. immigrants based on port entry records. We record a match if at least two sources agree on the country of origin for a surname.
4. Fourth, for the remaining unmatched surnames, we hand-check their country of origin using sources such as dictionaries and ancestry.com for 3,000 of the most common surnames. The procedure generates a list of over 1.5 million unique surnames and their associated country of origin. Finally, we merge the surname data with remaining inventors from the patent database.

Step 5. For inventors who are not matched in Step 1, we define the inventor's inherited cultural values as the six-dimension Hofstede cultural values in the country of origin as determined by Steps 2 to 4. Combined with the 322,687 inventors from Step 1, we obtain inherited cultural values for about 98% of 1.5 million U.S.-based inventors (including those working in both private and public companies) that patented between 1981 and 2016.

Appendix IA.3: Survey-based validation of culture measures

To further validate the underlying assumption of our culture measures, we empirically test the relation between the Hofstede cultural values in an individual's country of ancestry and the individual's personal beliefs using individual responses to relevant survey questions from the 1972 to 2016 sample of the U.S. General Social Survey. The survey, which is frequently used in the economics literature, asks respondents in the U.S. their background information and their opinions on a wide array of topics. To identify ancestry background, we use responses to the question: "from what countries or part of the world did your ancestors come?" Based on these responses, we construct individual-level Hofstede cultural measures analogous to the main cultural measures used in the empirical analysis for inventors. We then conduct individual-level regressions where the dependent variables are respondents' responses to survey questions relating to each of the six Hofstede cultural dimensions and the independent variables are the Hofstede cultural measures in the respondents' reported countries of ancestry and a set of individual-level controls.

The survey questions that we select for each Hofstede cultural dimension are as follows. For the individualism dimension, the selected question asks respondents to pick what quality that they think is most important to teach their children to prepare them for adulthood, where 'to think for himself or herself' is one of the possible qualities. For the power distance dimension, the selected question asks whether people should obey the law without exception or follow their consciences, which is closely related to the concept that people in high PDI societies obey authority without questioning it. For the uncertainty avoidance dimension, we select questions related to people's preference for a job that has no danger of being fired, which is related with the idea that people in high UAI societies prefer certainty. For the masculinity dimension, we select questions related to people's views on traditional gender roles and gender inequality. For the long-term orientation dimension, we infer time orientation based on smoking status since smoking generates short-term pleasure but has known long-term negative health consequences. For the indulgence dimension, we select questions related to exercising restraint (which is the opposite of indulgence) such as the importance of having self-control and working hard.

In column (1) of Table IA.1, we focus on the individualism dimension and find that U.S. respondents from ancestry countries with higher individualism scores think it is more important to teach children to think for themselves to prepare them for life as an adult. This result is in line with the notion that people in more individualist societies are expected to take care of only themselves and their immediate families. In columns (2) to (6) of Table IA.1, we examine the other five Hofstede dimensions in a similar way and find that the Hofstede cultural values in a respondent's country of ancestry are strongly related to the respondent's individual views on related issues.

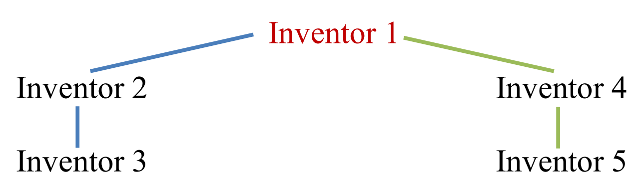
Overall, these results provide empirical support for the underlying premise of our measures of inherited cultural values that inventors' cultural heritage, as measured by the prevalent preferences in their country of ancestry, significantly impacts their values and preferences. This supports the idea that inventors sharing similar cultural backgrounds tend to have more similar values compared to inventors from different cultural backgrounds. In addition, since more than 90% of the respondents are second or higher generation immigrants, these empirical results also lend support to the idea that immigrant cultures are persistent across generations. The notion that immigrants bring their cultural beliefs to the new host country and pass down these beliefs to their descendants resulting in the persistence of cultural beliefs across generations has been empirically demonstrated by many studies in the economics literature (e.g., Guiso, Sapienza, and Zingales, 2006; Fernández and Fogli, 2009; Fernández, 2011).[68]

[68] For more evidence on the persistent intergenerational transmission of cultural beliefs, see Bisin and Verdier (2011).

Appendix IA.4: Additional explanation of how our difference-in-differences methodology controls for the *relative* contribution of each deceased inventor to each of their specific teams

As explained in Section 4.2, our empirical procedure specifies that every comparison of inventor teams in our treatment tests involves comparing two inventor teams that operate *within the same firm* at the *same point in time*. In conjunction with Team and Year fixed effects, this setup ensures that our treatment test results cannot be explained by factors across firms or time-invariant team and year effects.

Importantly, our inclusion of various (time-varying) control variables that capture the amount and diversity of knowledge, skills, and experiences of each inventor team also allows us to *explicitly control for the relative contribution of the deceased co-inventor to each of their treated teams*. This includes cases where the focal deceased inventor is the same person for two sets of treated teams in our most restrictive treatment effects analysis implemented in Section 4.2.4. As illustrated in the example below, this identification strategy focuses on comparing sets of treated teams who suffer the *loss of the exact same co-inventor* (i.e., Inventor 1 in the illustrative example below) while working at the same firm (through the inclusion of Dead co-inventor fixed effects):



Treated Team A (*pre-death*)
= Inventor 1
+ Inventor 2
+ Inventor 3

Treated Team B (*pre-death*)
= Inventor 1
+ Inventor 4
+ Inventor 5

Treated Team A (*post-death*)
= Inventor 2
+ Inventor 3

Treated Team B (*post-death*)
= Inventor 4
+ Inventor 5

For example, assume that the treated teams A and B are both focused on developing building/ construction-related technologies for the same corporate employer. Assume that Team A has one engineer (Inventor 1) and two architects (Inventor 2 and Inventor 3) while Team B has two engineers (Inventor 1 and Inventor 4) and one architect (Inventor 5). The two teams share the same engineer (Inventor 1), who passes away unexpectedly.

Inventor 1 (Engineer)

Inventor 2 (Architect) | Inventor 4 (Engineer)

Inventor 3 (Architect) | Inventor 5 (Architect)

Treated Team A (*pre-death*)
= Inventor 1 (Engineer)
+ Inventor 2 (Architect)
+ Inventor 3 (Architect)

Treated Team B (*pre-death*)
= Inventor 1 (Engineer)
+ Inventor 4 (Engineer)
+ Inventor 5 (Architect)

Treated Team A (*post-death*)
= Inventor 2 (Architect)
+ Inventor 3 (Architect)

Treated Team B (*post-death*)
= Inventor 4 (Engineer)
+ Inventor 5 (Architect)

Inventor 1's death will arguably have a larger effect on Team A than Team B because Inventor 1 was the only engineer on Team A and thus less replaceable. In other words, the deceased Inventor 1 may have a higher team-specific value-add to Team A than Team B. As such, it is possible that Team B will subsequently outperform Team A while it is also possible that Team B just so happens to be the team that experiences an increase in team cultural diversity (this possibility is hereafter referred to as the "alternative hypothesis").

Before continuing, it is important to note that our extensive univariate comparisons of observable characteristics between our two sets of treated teams in Panels D and E of Table 4 indicate that our heterogenous treatment effects sample is quite well balanced with no systematic correlation between changes in team cultural value diversity and changes in the team's base of knowledge, skills, and experiences. For example, we show in Panel D of Table 4 that the

Percentage of unique tech class knowledge lost due to the focal co-inventor's death and the *Change in team technical knowledge overlap due to co-inventor death* is insignificantly different between treated teams whose cultural value similarity increases post-death vs. treated teams whose cultural similarity does not increase post-death. The comparability of our treated teams, in terms of team patenting output, is further confirmed by Panel C of Table 4 and unreported parallel trends tests.

Nevertheless, a critical benefit of our heterogenous treatment effects specification with dead co-inventor fixed effects is that the relatively higher contribution of Inventor 1 to Team A as opposed to Team B in our hypothetical example above will naturally be accounted for as part of the change in our multitude of control variables that capture the amount and diversity of knowledge, skills, and experiences possessed by the focal inventor team in each year. This can be seen in the continuation of our stylized example:

Treated Team A (*pre*-Inventor 1 death)				**Treated Team B (*pre*-Inventor 1 death)**			
Control variable	Inventor background	Tech classes patented in	Scope of team knowledge base	Control variable	Inventor background	Tech classes patented in	Scope of team knowledge base
Inventor 1	Engineer	K,L,M		Inventor 1	Engineer	K,L,M	
Inventor 2	Architect	X,Y,Z		Inventor 4	Engineer	K,L,M	
Inventor 3	Architect	X,Y,Z		Inventor 5	Architect	X,Y,Z	
Team A control value (pre-death)			**6**	**Team B control value (pre-death)**			**6**
Treated Team A (*post*-Inventor 1 death)				**Treated Team B (*post*-Inventor 1 death)**			
Control variable	Inventor background	Tech classes patented in	Scope of team knowledge base	Control variable	Inventor background	Tech classes patented in	Scope of team knowledge base
Inventor 2	Architect	X,Y,Z		Inventor 4	Engineer	K,L,M	
Inventor 3	Architect	X,Y,Z		Inventor 5	Architect	X,Y,Z	
Team A control value (post-death)			**3**	**Team B control value (post-death)**			**6**
Change in Team A control value			**–3 (–50%)**	**Change in Team B control value**			**0 (0%)**

Note: Team A (or B) control value refers to the value of the control variable, *Scope of team knowledge base*, which is the count of unique technology classes that any team member has previously patented in up to the focal analysis year. Change in Team control value = Team control value (post-death) – Team control value (pre-death).

For example, if Inventor 1's contribution of unique technical knowledge is relatively higher for Team A than Team B, then the decline in the control variable *Scope of team technical knowledge* (which counts the number of unique technology classes that any team member has previously patented in) will be relatively greater for Team A than Team B. This same logic can be applied for our full set of extensive, time-varying control variables capturing the amount and diversity of each inventor team's knowledge, skills, and experiences.

Since we also include Team FEs in our within-firm diff-in-diff specification to control for time-invariant components of a team's knowledge, skills, and experiences, we explicitly control for both the time-varying *and* the time-invariant components of knowledge, skills, and experiences provided by the same surviving inventors in both the pre- and post-death period. This leaves the *relative* contribution (or loss) of the deceased inventor's knowledge, skill, and experience between the pre-death period and the post-death period as the primary driver of changes in team-level control variables. As such, our coefficient estimates for the effect of team cultural similarity on team output are already conditional on the relative importance of the dead inventor's knowledge, skill, and experience to each treated team of which they were a member.

Therefore, our treatment effects specifications can control for the possibility that the same co-inventor's death can impact different teams differently due to a deceased co-inventor's team-specific value-add. This in turn allows us to generate credibly causal estimates of the incremental impact of team cultural diversity on team innovation output.

Finally, in the hypothetical scenario outlined above, one would expect that Team B (assuming for illustrative purposes that the team's cultural similarity decreases post Inventor 1's death) would clearly and uniformly outperform Team A (assuming for illustrative purposes that the team's cultural similarity increases post Inventor 1's death) in the post-treatment period. However, we find that overall (citation-weighted) patent output is *not* significantly different in the post-treatment period between the two sets of treated teams but that only the type and distribution of team innovation outcomes significantly change. In other words, treated teams that become more culturally homogenous produce a higher overall quantity of patents that are more exploitative in nature while treated teams that become more culturally diverse produce more risky, exploratory patents with a greater potential for high-impact innovations. As such, our unique distributional results concerning both the positive and negative effects of team cultural value diversity on team innovation output do not appear to be logically consistent with the predictions of the alternative hypothesis. Specifically, in the hypothetical scenario above, our main empirical results would indicate that Team A outperforms Team B in terms of the quantity of patents produced while the overall (citation-weighted) patent outputs of Teams A and B do not significantly differ, which are inconsistent with the prediction of the alternative hypothesis that Team B would uniformly outperform Team A across all innovation output metrics in the post-treatment period.

Appendix IA.5: Procedure to Identify Treated and Control Teams in Diff-in-Diff Tests

The first step in implementing our triple difference-in-differences empirical strategy involves identifying active inventor teams at U.S. publicly listed firms that experience the 'premature' death of one of their team members (otherwise referred to as the "treated teams"). We begin with USPTO data and reports that directly identify inventors who died around the date of patent application.[69] We then supplement and match this information with Infutor's Consumer History Plus (CRD4) database, LexisNexis Public Records, and the Fold3 Social Security Death Index to identify those deceased inventors that:

a) Died between 1981 and 2011 (ensures we have at least 5 years of pre- and post-death data);
b) Are employed at a U.S. publicly listed firm at the time of their death; and
c) Are no older than 60 years of age at the time of their death.[70]

Next, to isolate 'active' inventor teams that are likely to be genuinely impacted by the death of their colleague, we identify all teams of 3+ inventors[71] that collaborated with the deceased inventor (at the same firm) on a patent that was applied for within 3 years of the focal inventor's death.[72] Finally, we require that we have data on each team member's cultural heritage, location, and patenting history. These filters ultimately result in the identification of 2,637 treated teams that were actively collaborating at a U.S. public firm around the date of a co-worker's premature death.

---

[69] In the inventor fields published on the USPTO website, a recently deceased inventor will have a 'deceased' or 'late' label affixed to their name and/or have their legal representative noted on the patent application. Separately, the USPTO publishes records of petitions related to deceased inventors and their patent applications. A key advantage of relying directly on this USPTO data is that it more precisely identifies the deceased inventor's name, location, and employer at time of death. This helps facilitate matching a deceased inventor's patenting history to their personal characteristics.
[70] Following Jaravel et al. (2018), we define a 'premature' co-inventor death as an inventor that was 60 years old or younger at the date of their passing to reduce the likelihood that the death is due to a long-standing health condition.
[71] In our main analysis, we exclude original two-person inventor teams (comprising the deceased inventor plus one more surviving inventor) because it is not feasible to calculate meaningful "team-based" diversity measures in the post-treatment period when the surviving treated "team" only comprises one living inventor.
[72] Our empirical results are very similar if we use a 5-year 'active' cut-off threshold instead of 3 years.

The second step in our empirical strategy is to use the following procedure to match each treated team with a corresponding counterfactual control team that does not experience a co-inventor death. First, we identify all teams of inventors working at the same firm as the treated team and keep those teams that have the same number of inventors as the treated team prior to the focal co-inventor's death. Second, we ensure that no members of the potential control team have an existing collaboration with any treated team member in our sample.[73] Third, we require that the potential control team is actively working together around the time of the focal co-inventor's death. Specifically, the control team must have successfully patented together at least once in the 3 years leading up to the focal co-inventor's death. Fourth, we require that the potential control team has developed the same number of (eventually granted) patent applications to date as the treated team at the time of the focal co-inventor's death to ensure that the control team has similar team-specific human capital and productivity as the treated team. Finally, we specify that the chosen control team has the closest proximity to the treated team in terms of cultural value diversity. We then use this control team's characteristics and patenting activity as the counterfactual for how the relevant treated inventor team would have performed if they did not suffer the loss of their collaborator. After implementing this procedure, we have a final sample for our triple diff-in-diff analysis of 2,441 treated teams and 2,441 counterfactual controls.[74]

To illustrate our approach, take as an example from our data three inventors (Mr. A, Mr. B, and Mr C.) who are working at the technology firm Z. This team applies for three (eventually granted) patents before the year 2005. We then observe that Mr. A dies before his 60th birthday in

---

[73] This ensures that control teams are not subject to any direct spill-over effects arising from a colleague's death.

[74] We are not able to find a suitable counterfactual control team for approximately 7% of our sample. This principally occurs in smaller U.S. public firms with a more limited pool of inventors that are unaffiliated with inventors comprising the treated teams. However, in unreported tests, the unmatched treated teams do not appear to be significantly different (at least on observable characteristics) from those that do find a matching control team. Note also that *all* treated teams are included in our most restrictive heterogeneous treatment effects specification outlined in Section 4.2.4.

the year 2005. We would proceed to form a pair of treated and counterfactual control teams related to this co-inventor death as follows. First, we define the post-death treated team as the combination of Mr. B. and Mr. C. and calculate all post-death team-level output and characteristics based on these two surviving individuals. Second, we find a counterfactual control team working at the same firm Z with no connection to any of the treated inventors based on the following criteria:

(1) The control team must have the same number of team members in the pre-treatment period as the treated inventor team (in this case, three individuals);

(2) Control team must be actively collaborating around the time of the co-inventor's death (in this case, the control team must have patented together at least once between 2002 to 2005);

(3) The control team must have generated three (eventually granted) patent applications by the time of the focal co-inventor's death (in this case, 2005); and

(4) The chosen control team is the one closest to the treated team's pre-death cultural values.

Table IA.1: General Social Survey

The table reports results based on individual-level regressions using the U.S. General Social Survey from 1972 to 2016, where the number of observations is the number of respondents for each survey question. THINKSELF asks "If you had to choose, which thing on this list would you pick as the most important for a child to learn to prepare him or her for life?" The answer to this question is reversed so higher values indicate it is more important "to think for himself or herself." OBEYLAW asks "In general, would you say that people should obey the law without exception, or are there exceptional occasions on which people should follow their consciences even if it means breaking the law?" The answer to this question is reversed so higher values indicate people should obey the law without exception rather than follow their consciences. JOBSEC asks "would you please look at this card and tell me which one thing on this list you would most prefer in a job?". The answer to this question is reversed so higher values indicate more importance placed on "no danger being fired". MRMOM/NO GENDEREQ equals one (zero otherwise) if the respondent agrees/strongly agrees with the statement that "it is not good if the man stays at home and cares for the children and the woman goes out to work" or answers "probably should not/definitely should not" to the question that "do you think it should or should not be the government's responsibility to promote equality between men and women?" NO SMOKE equals one (zero otherwise) if the respondent answered no to the question "do you smoke?". NO SELF-CONTROL/WORKHARD equals one (zero otherwise) if the respondent does not think it is important for a child to learn to have self-control or work hard. The main independent variables are the six Hofstede cultural dimensions in the respondent's reported country of ancestry, which are normalized to be between 0 and 1. Ln(Age) is the natural logarithm of the respondent's age at the time of the interview. Male equals one (zero otherwise) if the respondent is male. Ln(Education) is the natural logarithm of the number of years of formal education the respondent has completed. Ln(Income) is the natural logarithm of the respondent's family income converted to 1986 dollars. Married equals one (zero otherwise) if the respondent is married at the time of the interview. White and Black are indicators for the respondent's reported race, where the omitted category is other races. Employed equals one (zero otherwise) if the respondent holds a full-time or a part-time job. Intercepts are included and not reported. Survey year fixed effects are included. Robust standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

VARIABLES	THINKSELF	OBEYLAW	JOBSEC	MRMOM/ NO GENDEREQ	NO SMOKE	NO SELF-CONTROL/ WORKHARD
Predicted Signs	+	+	+	+	+	+
Individualism	0.209***					
	(0.043)					
Power distance		0.138***				
		(0.043)				
Uncertainty avoidance			0.189***			
			(0.045)			
Masculinity				0.162***		
				(0.052)		
Long-term orientation					0.051**	
					(0.020)	
Indulgence						0.033**
						(0.014)
Ln(Age)	0.060**	0.109***	0.095***	0.136***	0.134***	0.065***
	(0.026)	(0.021)	(0.026)	(0.023)	(0.012)	(0.008)
Male	-0.217***	-0.078***	0.106***	0.007	-0.060***	-0.035***
	(0.018)	(0.014)	(0.019)	(0.016)	(0.009)	(0.006)
Ln(Education)	0.933***	-0.296***	-0.795***	-0.177***	0.171***	0.001
	(0.042)	(0.033)	(0.042)	(0.042)	(0.018)	(0.012)
Ln(Income)	0.084***	-0.016*	-0.116***	-0.025**	0.022***	-0.026***
	(0.011)	(0.009)	(0.012)	(0.010)	(0.006)	(0.004)
Married	-0.080***	0.025	-0.013	0.030*	0.034***	0.009
	(0.019)	(0.016)	(0.021)	(0.018)	(0.010)	(0.007)
White	0.409***	-0.115***	-0.114*	0.026	-0.085***	0.045***
	(0.049)	(0.037)	(0.060)	(0.034)	(0.028)	(0.015)
Black	0.338***	-0.043	0.088	-0.006	-0.088***	0.046***
	(0.051)	(0.039)	(0.066)	(0.040)	(0.031)	(0.017)
Employed	0.072***	-0.056***	0.021	-0.025	-0.044***	0.009
	(0.021)	(0.018)	(0.022)	(0.019)	(0.010)	(0.007)
No. of observations	19,295	4,505	16,058	2,127	11,885	24,830
Adjusted R^2	0.076	0.074	0.072	0.054	0.036	0.106

Table IA.2: Drivers of co-inventor choice – Subsample of inventors with different countries of origin

This table reports the results of conditional logit models that estimate the factors affecting the choice of co-inventor in the subsample where pairs of inventors are *not* from the same country of origin, using alternative specifications for defining counterfactual control pairs. The dependent variable is equal to one for all new co-inventor pairwise relationships formed at a U.S. publicly listed firm during the sample period and zero for counterfactual pairs that comprise the comparison/control group. Column (1) uses other inventors working at the same firm in the same year as the new actually formed co-inventor pair to form the counterfactual control pairs. Column (2) uses other inventors working at the same firm, in the same year and at the same office to form the counterfactual control pairs. Column (3) uses other inventors working in the same division/subsidiary of a firm at the same time to form the counterfactual control pairs. Column (4) uses other inventors working in the same division/subsidiary of a firm, in the same year and at the same office to form the counterfactual control pairs (see Sections 3.2 and 3.3 for further details). Appendix A provides definitions for all independent variables. All regression specifications include group fixed effects. Robust standard errors (clustered at the group level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Counter-factual group	Same firm Same year (1)	Same firm Same year Same office (2)	Same division Same year (3)	Same division Same year Same office (4)
Co-inventor cultural similarity	0.140*** (0.008)	0.122*** (0.008)	0.127*** (0.008)	0.128*** (0.009)
Co-inventor geographic distance	-0.404*** (0.001)	-0.126*** (0.001)	-0.360*** (0.001)	-0.112*** (0.001)
Co-inventor difference in average forward cites to date per patent	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)
Both top 10% inventors	7.728 (5.937)	1.552*** (0.246)	4.007*** (1.573)	1.461*** (0.230)
Co-inventor tech proximity	1.796*** (0.006)	1.488*** (0.006)	1.608*** (0.006)	1.408*** (0.006)
Co-inventor difference in average backward cites to date per patent	-0.020*** (0.001)	-0.019*** (0.001)	-0.018*** (0.001)	-0.016*** (0.001)
Co-inventor difference in years of experience to date	-0.090*** (0.002)	-0.090*** (0.002)	0.085*** (0.002)	-0.087*** (0.002)
Both have 5+ years of tenure at focal firm	-0.475*** (0.005)	-0.384*** (0.004)	-0.445*** (0.005)	-0.370*** (0.005)
Group fixed effects?	Yes	Yes	Yes	Yes
No. of observations (mil)	4,187,693	3,829,631	3,820,326	3,480,247
No. of actual pairs (mil)	1,535,832	1,457,308	1,414,260	1,334,757
No. of counter-factual pairs (mil)	2,651,861	2,372,323	2,406,066	2,145,490
Pseudo R^2	0.24	0.04	0.20	0.04

Table IA.3: Determinants of choice of co-inventor – More experienced/prolific inventors only

This table reports the results of conditional logit models that estimate the factors affecting the choice of co-inventor. The dependent variable is equal to one for all new co-inventor pairwise relationships formed at a U.S. publicly listed firm during the sample period and zero for counterfactual pairs that comprise the comparison/ control group. Column (1) considers only the collaboration choices of more prolific inventors, defined as inventors whose total number of patents developed to date places them in the top decile (10%) of all inventors in the USPTO universe (i.e., a Top 10% inventor). Column (2) considers only the collaboration choices of inventors with more than 10 years of professional inventor experience up to and including their tenure at their current focal firm. Column (3) considers only the collaboration choices of inventors who have been working at their current focal U.S. public firm for 10 or more years (see Section 6.1 for further details). Appendix A provides definitions for all independent variables. All regression specifications include group fixed effects. Robust standard errors (clustered at the group level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Focal treated inventor collaboration choice group	Top 10% inventor only (1)	10+ years of professional inventor experience (2)	10+ years of tenure at current employer (3)
Co-inventor cultural similarity	0.241*** (0.010)	0.196*** (0.010)	0.152*** (0.012)
Co-inventor geographic distance	-0.110*** (0.002)	-0.112*** (0.002)	-0.105*** (0.002)
Co-inventor diff. in avg. forward cites to date per patent	-0.008*** (0.001)	-0.008*** (0.001)	-0.014*** (0.002)
Both top 10% inventors	1.476*** (0.200)	1.329*** (0.249)	1.202*** (0.341)
Co-inventor tech proximity	1.630*** (0.008)	1.528*** (0.008)	1.624*** (0.009)
Co-inventor diff. in avg. backward cites to date per patent	-0.023*** (0.002)	-0.017*** (0.002)	-0.023*** (0.002)
Co-inventor diff. in years of inventor experience to date	-0.087*** (0.003)	0.006 (0.004)	-0.043*** (0.004)
Both have 5+ years of tenure at focal firm	-0.511*** (0.006)	-0.397*** (0.006)	-0.457*** (0.006)
Group fixed effects?	Yes	Yes	Yes
No. of observations	1,568,093	1,683,936	1,213,197
No. of actual pairs	744,807	792,238	571,324
No. of counter-factual pairs	823,286	891,698	641,873
Pseudo R^2	0.07	0.05	0.07

Table IA.4: Innovation output around exogenous co-inventor turnover – Triple diff-in-diff tests (with Firm × Year fixed effects)

This table reports the change in innovative output for treated teams around the death of a co-inventor relative to control teams as defined in equation (3) in Section 4.2.1 (but also with Firm × Year fixed effects). $After_{i,t}$ is an indicator variable equal to one for all years after the focal co-inventor's death and zero otherwise. *Treated team*$_i$ is an indicator variable equal to one for all teams that suffer the death of a team member and zero otherwise. *Team cultural similarity change*$_i$ is a continuous variable equal to the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. We measure the quantity of team innovative output each year as *Ln(1+Total patents)*. The average quality of team innovative output is measured as *Ln(1+Average forward citations per patent)*. We measure a team's propensity for producing high impact innovations as *Ln(1+Top 10% cited patents)*. A team is designated as being more (less) focused on exploitative (explorative) innovation if they have higher *Average backward citations per patent* and a higher *Average number of claims per patent*. Total net innovation output is measured as *Ln(1+Total citation-weighted patents)*. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team geographic diversity). See Appendix A for all variable definitions. All regression specifications include both Team and Firm × Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
$After_{i,t}$	-0.210***	-0.162***	-0.025***	-0.212***	-0.198***	-0.177***
	(0.004)	(0.004)	(0.002)	(0.005)	(0.004)	(0.005)
$After_{i,t}$ × Treated $team_i$	0.006**	0.006*	0.000	0.015***	0.005	0.009
	(0.003)	(0.003)	(0.001)	(0.004)	(0.003)	(0.006)
$After_{i,t}$ × Treated $team_i$ ×	0.058**	-0.011	-0.029**	0.053*	0.047**	-0.007
Team cultural similarity $change_i$	(0.029)	(0.029)	(0.012)	(0.028)	(0.022)	(0.036)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	99,830	99,830	99,830	99,830	99,830	99,830
Adjusted R^2	0.14	0.12	0.07	0.14	0.12	0.13

Table IA.5: Heterogeneous treatment effects around co-inventor deaths (treated teams only) (inverse hyperbolic sine transformation)

This table reports the change in innovative output for treated teams only around the death of a co-inventor (see equation (4) in Section 4.2.4) but using the inverse hyperbolic sine transformation (*Asinh*) for all patent-related variables rather than the (started) natural log transformation used in Table 6. $After_{i,t}$ is an indicator variable equal to one for all years after the focal co-inventor's death and zero otherwise. *Team cultural similarity change*$_i$ is a continuous variable equal to the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. We measure the quantity of team innovative output each year as *Asinh(1+Total patents)*. The average quality of team innovative output is measured as *Asinh(Average forward citations per patent)*. We measure a team's propensity for producing high impact innovations as *Asinh(Top 10% cited patents)*. A team is designated as being more (less) focused on exploitative (explorative) innovation if they have a higher average number of backward citations per patent, *Asinh(Average backward citations per patent)*, and a higher average number of claims per patent, *Asinh(Average claims per patent)*. Total net innovation output is measured as *Asinh(Total citation-weighted patents)*. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team geographic diversity). See Appendix A for definitions of all variables used in this analysis. All regressions include Individual dead co-inventor fixed effects, Team fixed effects and Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Asinh(Total patents)	Asinh(Avg forward cites per patent)	Asinh(Top 10% cited patents)	Asinh(Avg back cites per patent)	Asinh(Avg claims per patent)	Asinh(Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
$After_{i,t}$	-0.342***	-0.283***	-0.069***	-0.333***	-0.299***	-0.336***
	(0.007)	(0.008)	(0.004)	(0.008)	(0.006)	(0.010)
$After_{i,t}$ × Team cultural similarity change$_i$	0.083**	-0.008	-0.033**	0.073**	0.073***	0.002
	(0.035)	(0.033)	(0.015)	(0.036)	(0.026)	(0.043)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Dead co-inventor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	57,392	57,392	57,392	57,392	57,392	57,392
Adjusted R^2	0.14	0.09	0.06	0.11	0.09	0.11

Table IA.6: Heterogeneous treatment effects around co-inventor deaths (treated teams only) (inclusion of fully interacted controls)

This table reports the change in innovative output for treated teams only around the death of a co-inventor (see equation (4) in Section 4.2.4 for further details). *After$_{i,t}$* is an indicator variable equal to one for all years after the focal co-inventor's death and zero otherwise. *Team cultural similarity change$_i$* is a continuous variable equal to the surviving team's cultural similarity post the focal co-inventor's death minus the pre-death/pre-treatment team's cultural similarity. The dependent variables in our analysis are *Ln(1+Total patents)*, *Ln(1+Average forward citations per patent)*, *Ln(1+Top 10% cited patents)*, *Ln(1+Average backward citations per patent)*, *Ln(1+Average claims per patent)* and *Ln(1+Total citation-weighted patents)*, respectively. As outlined in Section 2.3, we include several sets of control variables for the size (and diversity) of a team's knowledge base, the amount (and diversity) of team members' skills, and the amount (and diversity) of experience accumulated by team members, as well as other controls (team geographic diversity). In this table, we also include as controls the interaction of the post-death change in the average value of each control variable (computed as the post-death period average minus the pre-death period average) with the post-death indicator, *After$_{i,t}$*. See Appendix A for all variable definitions used in this analysis. All regressions include Individual dead co-inventor fixed effects, Team fixed effects and Year fixed effects. Robust standard errors (clustered at the team level) are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Category of innovation outcome	Quantity of innovation	Average quality of innovation	Likelihood of high impact innovation	Relative focus on exploitative innovation		Total net innovation output
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Top 10% cited patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Cite weighted patents)
	(1)	(2)	(3)	(4)	(5)	(6)
After$_{i,t}$	-0.278***	-0.242***	-0.053***	-0.323***	-0.265***	-0.270***
	(0.013)	(0.014)	(0.007)	(0.017)	(0.011)	(0.016)
After$_{i,t}$ × Team cultural similarity change$_i$	0.055**	-0.012	-0.028**	0.053**	0.048**	-0.010
	(0.027)	(0.025)	(0.012)	(0.026)	(0.019)	(0.033)
Controls for size & diversity of team knowledge base	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team skill	Yes	Yes	Yes	Yes	Yes	Yes
Controls for amount & diversity of team experience	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
After$_{i,t}$ × Change in team knowledge base controls	Yes	Yes	Yes	Yes	Yes	Yes
After$_{i,t}$ × Change in team skill controls	Yes	Yes	Yes	Yes	Yes	Yes
After$_{i,t}$ × Change in team experience controls	Yes	Yes	Yes	Yes	Yes	Yes
After$_{i,t}$ × Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Dead co-inventor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	57,392	57,392	57,392	57,392	57,392	57,392
Adjusted R^2	0.15	0.10	0.07	0.12	0.11	0.12