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

June 2021

ABSTRACT

Given the rising focus on workplace diversity and labor productivity in today's economy, we open the black box process of corporate innovation production by examining the most important input into the firm R&D process, namely the individual employees tasked with creating new inventions. Using information on over two million inventors employed at U.S. public firms, we examine how individual inventors' inherited traits (cultural values and gender) and acquired career experiences affect their desire to collaborate with others in a corporate R&D setting and how shared cultural values affects innovative output. We first provide novel evidence that, even amongst groups of comparably experienced inventors working in the same office, inventors who share similar cultural values are 20% more likely to work together on new research projects. Second, using exogenous shocks to inventor team composition arising from premature co-inventor deaths, we find that more culturally homogenous teams produce a higher quantity of patents that are more likely to exploit existing technologies and become moderately successful inventions. In contrast, more culturally diverse teams produce a higher share of risky, more exploratory patents that have a greater chance of becoming high impact innovations. Our results have key implications for promoting different types of innovation in corporate workplaces and the likely effectiveness of diversity hiring policies.

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

JEL Classification: G18; G31; J24; M14; M54; O31; O34; Z1

We would like to thank Rui Silva, Danielle Li, Cesare Fracassi and Francesco Lissoni (discussants), Jarrad Harford, Gerard Hoberg, Gordon Phillips, Ting Xu and Paige Ouimet as well as all seminar and conference participants at the 2021 American Finance Association (AFA) Annual Meeting, 2021 Western Finance Association (WFA) Annual Meeting, 2021 Society of Labor Economists (SOLE) Annual Meeting, 2020 Lone Star Finance Symposium, 2020 EPFL Innovation Seminar, University of Massachusetts, and Texas A&M University for their invaluable feedback. We acknowledge the generous support of the Mays Innovation Research Center and the Mays Business School. All errors are our own.

I. INTRODUCTION

Given the key role of technological innovation in driving the real economy (Stiglitz & Greenwald, 2015; Schumpeter, 1942), the rising importance of teamwork in spurring new inventions has become one of the defining features of the modern knowledge economy (Jones, 2009; Jaravel, Petkova & Bell, 2018; Lucas & Moll, 2014).[1] Concurrent with this phenomenon, there has been a conscious effort on the part of government regulators,[2] private corporations[3] and institutional investors[4] to implement policies that increase workplace diversity across several dimensions (for example cultural heritage and gender identity). In this paper, we investigate what the intersection of these two key trends, namely increased inventor teamwork and workplace diversity, means for inventor team formation and team-level labor productivity (and thus, by extension, firm innovation and economy-wide technological development). In particular, we examine what factors drive the collaboration decisions of inventors and how inventor team diversity influences team innovation output.

Despite the clear importance of these questions for both firms and policymakers, there is considerable disagreement in the literature on the relationship between workplace diversity and technological innovation. With respect to team formation, there is evidence of 'homophily', namely the tendency of individuals to prefer to work with others who share similar personal or social characteristics, in other business-related contexts (e.g. Gompers, Mukharlyamov & Xuan, 2016 (venture capitalists); Ishii & Xuan, 2014 (directors and executives); Gompers, Huang & Wang,

---

[1] The rise in the number and size of inventor teams has been attributed to a number of factors, including the benefits of co-specialization and knowledge sharing (Lucas & Moll, 2014) as well as a means for successive generations of individual inventors to more efficiently reduce their "knowledge burden" (Jones, 2009).

[2] Prohibitions on discriminatory employment practices are almost universal worldwide (e.g. *United Nations Convention No. 111 – Discrimination (Employment and Occupation*; Title VII of U.S. *Civil Rights Act 1964*). Furthermore, some regulations require firms to take affirmative actions in order to ensure that equal employment opportunity is provided to all current and prospective employees (e.g. see *U.S. Presidential Executive Order 11246* covering federal contractors. For an example of its enforcement, see *U.S. Department of Labor vs. Oracle America Inc., 2017*). In addition, some jurisdictions have enacted explicit diversity quota rules (e.g. Norway's *Public Limited Liability Companies Act (1997)* and California's *SB 826 (2018)* impose minimum female Board representation requirements).

[3] For example, approximately half of S&P 500 companies have established a dedicated Chief Diversity Officer (CDO) position or equivalent, with many of these appointments occurring over the last 3 years (Russell Reynolds Associates, 2019). In addition, many large public corporations provide annual reports that outline their diversity policies as well as publish statistics on a firm's diversity performance (e.g. Alphabet (Google), Amazon, Apple, Facebook, Microsoft).

[4] For example, two of the world's largest asset managers, Blackrock and State Street Global Advisors, have recently implemented proxy voting guidelines stating that they will vote against a firm's nominating/governance committee members if the firm does not have at least one or two women on its Board of Directors. Similarly, one of the world's largest proxy advisory firms, Glass Lewis, amended its voting guidelines in 2018 to factor Board diversity explicitly into its voting recommendations.

2017 (firm founding teams)). However, no prior study to the best of our knowledge has systematically evaluated the relative importance of inherited characteristics like cultural values vis-à-vis acquired career experiences in explaining the collaboration decisions of employee inventors in a more structured corporate R&D environment (see generally Herkenhoff, Lise, Menzio & Phillips, 2018). With respect to diversity and performance, there is much conflicting empirical evidence on whether:

a) Corporate Board diversity has a positive or negative impact on firm economic output (e.g. An, Chen, Wu & Zhang, 2020; Griffin, Li & Xu, 2020; Bernile, Bhagwat & Yonker, 2018, and Delis, Gaganis, Hasan & Pasiouras, 2017 c.f. Matsa & Miller, 2013; Ahern & Dittmar, 2012; Anderson, Reeb, Upadhyay & Zhao, 2011, and Adams & Ferreira, 2009);
b) Investor diversity enhances or impedes investment performance (e.g. Gompers et al., 2016 c.f. Cohen, Frazzini & Malloy, 2008; Hedge & Tumlinson, 2014); and
c) Employee diversity promotes or hinders firm innovation (e.g. Schubert & Tavassoli, 2020 and Ostergaard, Timmermans & Kristinsson, 2011 c.f. Doran, Gelber & Isen, 2016; Horwitz & Horwitz, 2007).

These strikingly divergent views likely stem from the formidable measurement and endogeneity issues encountered when studying the relationship between diversity and economic performance. First, measurement problems arise, for example, in the definition of the focal 'team'[5] and the associated 'team output'[6] as well as data limitations on the large-scale compilation of individual team member characteristics. This has led much of the prior literature to either rely on relatively small and somewhat subjective survey-based datasets with limited external validity[7] or focus on more coarse, publicly available firm-level data that only provides biographical details for a select group of company executives and Board members. Second, formidable endogeneity issues arise from challenges in identifying a valid counterfactual control group and difficulties in disentangling causal treatment effects from selection effects due to the endogenous matching of firms and employees and unobserved moderating factors at the industry-, regional-, firm- and/or CEO-level.

[5] For instance, when studying firm-level innovation outcomes, it is unclear whether firm management diversity, Board diversity and/or entire firm-wide diversity should be the appropriate unit of 'team-level' analysis.
[6] For example, when studying the impact of Board diversity on corporate innovation, how much responsibility should the Board receive for each patent developed by the firm?
[7] See, for example, Schubert & Tavassoli (2020) (2004-2012 biennial survey of approximately 500 Swedish firms that are concentrated in 'low-tech' manufacturing/service sectors) and Ostergaard et al. (2011) (uses a single 2006 survey of approximately 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").

In contrast to the conflicting prior literature that has focused primarily on analyzing highly endogenous firm-level relationships, we focus our analysis on understanding the formation and productivity of distinct inventor teams *within the same firm* at the *same point in time* (and even at the *same corporate division* at the *same geographic location* in certain specifications). There are several advantages to our unique corporate R&D setting. First, this setting provides us with a natural comparison sample of counterfactual inventors because we only compare inventor teams that work at the same firm (thus having similar access to physical firm resources, financing etc.) at the same time (thus facing a similar technological and competitive landscape) at the same location (thus controlling for local economic and labor market conditions) (Bernstein, McQuade & Townsend, 2019). As a result, our study can more cleanly isolate the relative importance of inherited traits (for example, cultural values) vis-à-vis acquired career experiences in driving inventor team formation and productivity. Second, we can use the detailed USPTO patent data to objectively identify the key contributors to a specific new patentable technology as well as trace the career history of millions of individual employee inventors and inventor teams. Finally, our focus on the most critical element of the corporate R&D process, namely the inventors who are ultimately responsible for creating novel technologies, allows us to identify some of underlying drivers of productivity in internal labor markets and thus corporate innovation (beyond the traditional CEO-, firm- and industry-level factors studied in the prior literature).

Our baseline sample consists of all teams of two or more inventors formed between 1981 and 2011 at U.S. publicly listed firms. Based on a sample of 1.2 million first-time collaborations between pairs of U.S. based inventors, we provide novel evidence that inherited affinity-based traits such as shared cultural values and shared gender play a critical role in observed inventor team formation. For example, two similarly experienced inventors currently working at the same firm in the same corporate office are 16%–20% more likely to collaborate with one another if they share similar cultural values. Similarly, female inventors are 14%–16% more likely to choose to collaborate with other female inventors in the same corporate office, all else being equal. These affinity-based preferences are robust across several within-firm specifications and are at least as important as a potential collaborator's prior technical experience in explaining the observed collaboration decisions of employee inventors.

Given this large-scale evidence documenting the strong preferences of firm R&D employees to collaborate with others that possess similar personal characteristics (for example, cultural values), we next examine the important question of whether these familiarity biases in co-inventor network formation enhance or impede team-level innovation productivity. In order to disentangle the selection effect from the treatment effect of inventor team diversity on team performance, we utilize a quasi-natural experiment involving premature co-inventor deaths that provides an exogenous source of variation in the cultural composition of inventor teams. Interestingly, 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 backgrounds. In particular, for treated teams that experience a decrease in team cultural diversity, these teams significantly increase their overall patent production by shifting their focus to the incremental exploitation of existing technologies, thus producing a relatively higher quantity of moderately successful inventions. In contrast, treated teams that become more culturally diverse because of their co-inventor's death subsequently tend to produce a higher share of risky, more exploratory patents that have a relatively greater chance of becoming high impact innovations. This dichotomy is consistent with the conjecture that while diversity in inherited traits can impede information sharing and the integration of disparate viewpoints (van Knippenberg & Schippers, 2007; Jehn, Northcraft & Neale, 1999; Williams & O'Reilly, 1998), a successful combination of differing perspectives amongst inventors can have a positive impact on the pursuit of technological innovation, particularly more high-risk, high-reward type inventions (Schubert & Tavassoli, 2020; Eesley, Hsu & Roberts, 2014; Fleming, 2001).

In further robustness tests, we reach similar conclusions using a heterogeneous treatment effects specification using only the treated teams subsample, where we compare the output of treated teams that suffer the *same* co-inventor death (and thus the same loss of co-inventor human capital) but experience *differential* impacts on teams' cultural value diversity. This combined evidence confirms that inventor diversity has both positive and negative consequences for innovation production with the net overall effect depending on the type of innovation pursued.

The findings of this paper contribute to several strands of the literature. First, there has been a growing body of literature documenting the importance of co-inventor networks for individual productivity (e.g. Jaravel et al., 2018; Azoulay, Graff Zivin & Wang, 2010) and the essential role of teams in innovation-related tasks (e.g. Alexander & Knippenberg, 2014; Jones, 2009; Agrawal, Kapur & McHale, 2008; De Dreu, 2006). We extend this prior work by studying how these co-inventor networks form in the first place and how team composition can directly influence the type of innovation produced. For example, we are, to the best of our knowledge, the first paper to study the relative importance of inherited affinity-based traits and professional career experiences in driving intra-firm inventor team formation and its consequent impact on team productivity.

Second, our paper augments the extensive literature that examines the relationship between various CEO-, firm-, investor- and industry-level characteristics and firm-level innovation (e.g. Custodio, Ferreira & Matos, 2019; Hirshleifer, Low & Teoh, 2012; Islam & Zein, 2020; Fitzgerald, 2020; Seru, 2014; Fang, Tian & Tice, 2014; He & Tian, 2013; Aghion, van Reenen & Zingales, 2013; Aghion, Bloom, Blundell, Griffith & Howitt, 2005). While these firm characteristics are significant drivers of corporate investment policy and firm innovative activity, the breadth and depth of our dataset allows us to probe one level deeper into the corporate R&D process by identifying the individuals directly involved in creating a new technological innovation (Bernstein et al., 2019). This in turn allows us to develop precise measures of team-specific homogeneity in terms of both inherited traits and acquired experiences. Our focus on individual inventors collaborating within a firm (which is an economically important yet relatively unexplored setting within the innovation networks literature) allows us to better understand the determinants of the collaboration choices of R&D employees and make cleaner inferences about the relationship between inventor team composition and innovation, beyond the existing factors identified at the firm level.

Third, our paper has important policy implications for both firms and government regulators with respect to workplace diversity and labor productivity. For instance, our evidence on 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. In particular, firms may need to enact proactive policies that incentivize existing employees to form a more diverse set of R&D collaborations, otherwise the strong homophily biases that we document may result in the oversupply of relatively homogenous teams. Furthermore, our empirical evidence 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 diverse teams appear to have a greater ability to produce more high impact innovations, it is important to acknowledge that more diverse teams are also relatively more likely to engage in failed research pursuits that negatively impact firm productivity. As such, our results suggest that each firm will have a different optimal mix of diverse and homogenous inventor teams to produce the desired combination of exploratory and exploitative innovations.

We organize the remainder of the paper as follows. Section II presents the data and construction of the 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 characteristics on an inventor team's innovation output. Section V discusses the key conclusions of our study.

II. DATA AND VARIABLES

2.1 Sample overview

Our analysis utilizes 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 (e.g. Fitzgerald, Balsmeier, Fleming & Manso, 2019; Atanassov & Liu, 2018; Hall & Lerner, 2010), we first collect information on U.S. patenting from three main sources: the United States Patent & Trademark Office (USPTO), PatentsView and the Berkeley-Fung Patent Database. The PatentsView database contains detailed disambiguated USPTO patent data from 1976 to 2018 and includes a patent's application and grant year, technology class, inventor names and locations, patent assignee names and locations (where the patent assignee is usually the firm or subsidiary 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 patent–Compustat assignee bridge, which spans 1976 to 2006, in the NBER Patent Database through to 2016) in conjunction with PatentsView company assignee ID numbers and our own database extensions,[8] we are able to identify all patents that are assigned to U.S. publicly listed firms between 1976 and 2018.

With respect to the patent-related data on individual inventors, we base our analysis on the PatentsView disambiguated inventor database. These files use information from the USPTO to assign each inventor a unique time-invariant ID to track each individual 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 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, 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. In the event that more than a year elapses between patent filings by the same inventor, we follow the prior literature and assume that an inventor changes employers at the midpoint between the two patent application years (see e.g. Song, Almeida & Wu, 2003; Baghai,

[8] 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).

Silva & Ye, 2019; Li & Wang, 2020).[9] 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 in order to be classified as an "active inventor" working at that firm. For example, if an inventor files a patent with Firm A in 2000 and another patent with Firm B in 2010 then we assume that this inventor is an employee of Firm A up to and including 2003 and is an employee of Firm B from 2007 onwards.

Our main sample covers the period from 1981 to 2011 in order to ensure that we have at least 5 years of inventor activity both before 1981 and after 2011. We focus exclusively on U.S. based inventors working at U.S. publicly listed firms because these publicly held firms are likely to have a sufficiently large pool of individual inventors and distinct inventor teams to facilitate the large-scale within-firm analysis that is the basis of our empirical identification strategy.

2.2 Identification of inventor teams at individual firms

Using augmented PatentsView inventor ID codes (which track all of patents developed by an individual inventor across time)[10] and 'big data' processing techniques, we identify all "teams" of 2+ inventors that ever collaborate on a patent and assign each of these teams a unique identifier. This process yields over 165 million distinct 'teams' that co-invent at least one patent during our sample period. This unique team identifier allows us to observe the patenting output, citation performance and technological specialization of each 'team' formed since 1976.

We then use this resulting combination of patent, inventor and team identifiers, along with patent assignee names, to identify all of the individual inventors and inventor teams working at each U.S. publicly listed employer at a given time and location. This enables us to understand the factors that influence the initial collaboration decisions of inventors currently employed at the same firm and to compare the innovative output of employee teams with differing levels of cultural value diversity and inventor experience.

[9] For example, if an inventor files a patent with Firm X in 2005 and then files their next patent with Firm Y in 2010, we assume that the inventor is an employee of Firm X up to and including 2007 and is an employee of Firm Y from 2008 onwards.

[10] Similar to the process described above for patent assignees, we initially use name- and location-based algorithms to identify potential duplicate inventor ID codes (i.e. an individual inventor is erroneously assigned multiple separate inventor IDs) and potential 'over-aggregating' inventor ID codes (i.e. the 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 general Google searches, to investigate potential inventor ID misclassifications and to augment the PatentsView inventor database where clear misclassifications are identified.

2.3 Variable construction

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

2.3.1 Innovation output

We employ several patent-based measures at both the individual inventor and inventor team-level to assess innovation productivity. Our first measure, *Total patents*, represents the number of patents filed (and subsequently granted) in a given year (Gao & Zhang, 2017). Our second measure, *Average forward cites per patent*, estimates the quality or impact of a patent by counting the number of citations that it receives. Scaled forward citations equals 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 Cooperative Patent Classification (CPC) technology sub-class.[11] We scale the raw citation count to account for potential variation in citation rates over time and across technologies (Bernstein, 2015) as well as to address truncation bias that results in patents granted towards the end of the sample having less time to accumulate citations (Hall, Jaffe & Trajtenberg, 2005). We then form the inventor team-year level measure by calculating the average scaled forward citations across all of the team's patents applied for in that year.

In addition to measures of patent quality based on future citations, we also assess patent quality utilizing the Kogan, Papanikolaou, Seru & Stoffman (2017) measure of the market value of patents (based on the stock price reaction to the announcement of new patent grants). We then aggregate these patent-level market values to create the average market value of patents created by an inventor team (*Average market value of patents*).

Beyond measures of the number and average quality of patents, we also seek to understand the type of innovation undertaken by inventor teams. According to March (1991), 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.

[11] The CPC is a patent classification system jointly developed by the USPTO and European Patent Office (EPO). There are approximately 640 4-digit CPC technology sub-classes that group patents together based on the similarity of their subject matter. It replaced the United States Patent Classification (USPC) system (with approximately 450 analogous 3-digit primary technology classes) in January 2013.

First, we calculate *Average backward cites per patent* 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 described above). As outlined in Balsmeier, Fleming & Manso (2017) and Lanjouw & Schankerman (2004), patents with more backward citations correlates to patenting in relatively more crowded, better known and more mature technology areas (consistent with a greater focus on exploitative innovation). Second, we compute *Average claims per patent* based on the number of claims made by a patent (using the same year and technology class scaling adjustment as for forward citations described above). As discussed in Balsmeier et al. (2017), a higher number of claims on a patent 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 the pressures for immediate and quantifiable results rises (consistent with a greater focus on exploitative innovation).

Finally, we consider where each of a team's patents fall in the patent citation distribution. In particular, given that more high-risk, high-reward explorative innovation is more likely to fall in the tail ends of the patent citation distribution (Balsmeier et al., 2017), we examine whether the patents developed by a particular team are "high impact innovations." We define "high impact innovations" as those patents that receive citations within the top 5% of patents in the same application year and CPC technology sub-class (*Top 5% cited patent*) or within the highest decile of patents in the same application year and CPC technology sub-class (*Top 10% cited patent*).

2.3.2 Inherited characteristics of individual inventors

We construct measures of inventors' inherited cultural characteristics in the following way. First, we identify the country of origin of inventors using their surnames following the methodology of Liu (2016) as described in Appendix B. This approach of using an individual's surname to identify their cultural background has been relied upon extensively in a wide variety of business disciplines.[12] Using the methodology in Liu (2016), we can link surnames to their countries of origin for close to 90% of the U.S. inventors in the patent database. While surnames have been frequently used to identify countries of origin, we acknowledge that this procedure may not yield

[12] While the use of names to classify populations into different ethnic groups has been around since the early 1900s (Rossiter, 1909), most recent efforts have been concentrated in the public health and population genetics literature (Mateos, 2007). Several recent studies in the accounting and finance literature have used surnames to identify the cultural background of inventors (Kerr & Lincoln, 2010), venture capitalists (Hegde & Tumlinson. 2014; Gompers et al. 2016), and company executives (Liu, 2016; Pan, Siegel & Wang, 2017; Brochet, Miller, Naranjo & Yu, 2019).

exact matches in all cases. For instance, individuals may assume more Anglo-Saxon surnames upon entry into the United States or women may take on their husbands' last names. However, such classification errors for a relatively small proportion of our entire sample of over two million inventors is, if anything, more likely to make our shared culture measures noisier and thus create a bias against finding any significant results.

Second, to measure the inherited cultural characteristics of inventors from a given country of origin, we use the cultural framework of Hofstede (1980) that classifies national culture into six dimensions: uncertainty avoidance, individualism, power distance, masculinity, long-term orientation, and indulgence. The first 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. This index is related to the concept of risk aversion and has been used as a measure of risk culture by prior studies (e.g., Pan, Siegel & Wang, 2017). The second cultural dimension is the individualism index (IDV), which measures the degree to which individuals are integrated into groups. The third cultural dimension is the masculinity index (MAS), which measures the degree to which masculine values such as competition and assertiveness dominate over feminine values such as cooperation and human relationships. The fourth cultural dimension is the power distance index (PDI), which measures the extent to which the less powerful members of organizations and institutions accept and expect that power is distributed unequally. The fifth cultural dimension is the long-term orientation index (LTO), which is related to the fostering of virtues oriented towards future rewards, in particular perseverance and thrift. The sixth cultural dimension is the indulgence index (IVR), which measures the extent to which people try to control their desires and impulses.

For robustness, we use two other measures to identity inherited cultural characteristics. The first measure is based on the cultural framework of Schwartz (1992) that classifies national culture into seven dimensions: harmony, embeddedness, hierarchy, mastery, affective autonomy, intellectual autonomy, and egalitarianism. The second measure is based on trust beliefs, calculated as the average response in each country of origin to the following question in the World Value Survey from 1982 to 2008: "*Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?*"[13] This measure has been used by prior

[13] Responses to the question are coded as one if "most people can be trusted" is selected and coded as zero if "you can never be too careful when dealing with others" is selected.

studies (e.g., Guiso, Sapienza & Zingales, 2006; Ahern, Daminelli & Fracassi, 2012) as a key dimension of national culture.

With respect to identifying the gender of inventors (male or female), we use the methodology and data developed by the Office of the Chief Economist at the USPTO, as reported in Toole, Myers, Breschi, Ferrucci, Lissoni, Miguelez, Sterzi & Tarasconi (2019). Toole et al. (2019) develop a 14-step methodology that utilizes the (primarily first) name of the inventor and various international name–gender dictionaries to attribute the gender of over 93% of inventors listed on granted U.S. patents from 1976 to 2016.

2.3.3 Acquired experiences of individual inventors

Using the patenting history of each individual inventor, we are able to develop several measures of the professional experience acquired by an inventor over the course of their career. First, we catalogue an individual's experience to date at both their current employer and across their entire inventor career. We calculate *years of inventor experience* as the number of years between the application date of the first (subsequently granted) patent that an inventor ever applies for and the current analysis year. Relatedly, we compute the tenure of the inventor at the focal firm as the number of years between the application date of the first (subsequently granted) patent that an inventor applies for whilst working at the focal employer and the current year.

Second, we seek to gauge how accomplished or successful an individual inventor has been relative to their peers in their inventor career to date, with the expectation that more similarly accomplished inventors are more likely to collaborate with each other. As such, we compute *average forward cites to date per patent* (defined as the mean of the scaled forward citations received across all of the inventor's patents to date) and the binary variable *Top 10% inventor* (which equals one for inventors whose total number of patents developed to date place them in the top decile of all inventors).

Third, we measure the type of innovative activity undertaken by inventors on the basis that inventors with similar technical expertise are more likely to work with one another. To this end, we compute *average backward citations per patent* (defined as the mean of the scaled backward citations received across all of the inventor's patents to date) and *technology class experience* (defined as a count of the number of distinct CPC sub-classes in which an inventor patents).

Appendix A provides a description of the other inventor-level characteristics presented in our empirical analysis.

2.3.4 Pairwise and team co-inventor characteristics

In order to undertake our empirical tests, we utilize the inherited traits and the acquired characteristics of individual inventors described above and transform them into pairwise inventor variables (for our inventor team formation, i.e. collaborator selection, tests) and team-level inventor variables (for the effect of cultural diversity on team performance, i.e. treatment effects, tests).

To measure the similarity in cultural values between a pair of inventors, we first calculate the Euclidean distance between two inventors' cultural values based on Hofstede's framework and then multiply this value by -1 using the following formula:

$$Pairwise\ similarity\ in\ inventor\ cultural\ values =$$
$$\begin{aligned} = -1 \times [&(PDI_{Inventor\ 1} - PDI_{Inventor\ 2})^2 + (IDV_{Inventor\ 1} - IDV_{Inventor\ 2})^2 \\ &+ (MAS_{Inventor\ 1} - MAS_{Inventor\ 2})^2 + (UAI_{Inventor\ 1} - UAI_{Inventor\ 2})^2 \\ &+ (LTO_{Inventor\ 1} - LTO_{Inventor\ 2})^2 + (IVR_{Inventor\ 1} - IVR_{Inventor\ 2})^2]^{\frac{1}{2}} \end{aligned}$$

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 1 and inventor 2.

To measure culture similarity within a team of N inventors, we use the following formula:

$$R\&D\ team\ cultural\ similarity = -1 \times \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(culture_{Inventor\ i} - culture_{Team\ mean})^2}$$

where the distance between each inventor's culture and the inventor's team culture is calculated as

$$\begin{aligned} &culture_{Inventor\ i} - culture_{Team\ mean} \\ &\quad = [(PDI_{Inventor\ i} - PDI_{Team\ mean})^2 + (IDV_{Inventor\ i} - IDV_{Team\ mean})^2 \\ &\quad + (MAS_{Inventor\ i} - MAS_{Team\ mean})^2 + (UAI_{Inventor\ i} - UAI_{Team\ mean})^2 \\ &\quad + (LTO_{Inventor\ i} - LTO_{Team\ mean})^2 + (IVR_{Inventor\ i} - IVR_{Team\ mean})^2]^{\frac{1}{2}} \end{aligned}$$

This measure is transformed so that its values are bounded between zero and one, where higher values indicate higher cultural similarity or less cultural diversity amongst the R&D team members.

To capture the role of shared gender in inventor team formation, we define the binary variable *Both female* as equal to one if both inventors in the inventor pair are female and zero otherwise. To measure *Team gender diversity* within a team of N inventors, we use the Blau (1977) Index of Gender Diversity, computed as follows:

$$Team\ gender\ diversity_i = 2 \times \left(1 - \sum_{s=1}^{S} (p_s)^2\right)$$

where s = male or female and p_s is the proportion of inventors in the team that are male or female respectively. As such, the value of this measure is bounded between zero and one with higher values indicating greater diversity or difference in genders within the inventor team.

To account for the role of co-inventor 'proximity' in increasing the ease of team formation, we develop several measures to capture the overlap between inventors in terms of geographic location and accumulated technical expertise. We calculate *co-inventor geographic distance* as the natural logarithm of one plus the geodetic distance (in miles) between the two inventors' then current locations. We also follow Jaffe (1989) and calculate the cosine similarity measure *co-inventor technological proximity* as the degree of overlap in technology classes between patents developed by each inventor in the focal analysis pair up to and including year *t* – 1.

To incorporate how prior inventor experience and prior patenting success may affect the choice of future R&D collaborators and inventor productivity, we transform the individual acquired experience variables described in Section 2.3.3 into the following additional pairwise inventor characteristics. Initially, we compute the continuous variables *co-inventor difference in average forward cites to date per patent*, *co-inventor difference in average backward cites per patent* and *co-inventor difference in years of inventor experience to date*. We then compute the binary dummy variables *both top 10% inventors* and *both have 5+ years of tenure at the focal firm*. Please refer to Appendix A for details of these transformations.

In our treatment effects specifications, we include a number of control variables that could influence the innovation production of inventor teams. To capture how the overall 'level' of team members' prior professional accomplishments and/or experiences may affect subsequent team productivity, we compile the team-level average of each individual team members' relevant characteristics. We thus include the following "inventor team average controls" in our regressions: *team average total number of patents to date*, *team average inventor experience to date*, *team average technology class experience to date*, *team average forward citations to date per patent* and *team average backward citations per patent*. To capture how 'diversity' in team members' prior professional accomplishments and/or experiences may affect subsequent team productivity, we compute the team-level standard deviation across each team members' relevant characteristics. We thus include the following "inventor team diversity controls" in our empirical specifications: *team gender diversity*, *team geographic diversity*, *team diversity in total number of patents to date*, *team diversity in inventor experience to date*, *team diversity in technology class experience to date*, *team diversity in average forward citations to date per patent* and *team diversity in average backward citations per patent*. Please refer to Appendix A for variable definitions.

2.4 Summary statistics

Table 1 provides the mean, median and standard deviation of the various characteristics of the U.S. based inventors working at U.S. listed firms that comprise our sample.

In Table 1 – Panel A, we present summary statistics detailing the lifetime characteristics of all individual inventors working at U.S. public firms in our sample. In Table 1 – Panel B, 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 amongst 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 Table 1 – Panel C, 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 backgrounds) across all three co-inventor cultural similarity measures described in Section 2.3.2. 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.

III. SELECTION FACTORS IN INVENTOR TEAM FORMATION

In this section, we seek to understand the relative importance of inherited affinity-based traits vis-a-vis acquired career experiences in explaining the establishment of new collaborations between inventors employed at the same firm.

3.1 Empirical methodology – ex ante selection

Following Dyck, Morse & Zingales (2010), Bena & Li (2015) and Bereskin (2018), we estimate the following conditional logit regressions using cross-sectional data:

$$\begin{aligned}
&Actual\ Pair_{ij,t} = \alpha + \beta_1 Coinventor\ cultural\ similarity_{ij,t-1} \qquad (1)\\
&+ \beta_2 Both\ female_{ij,t-1} + \ \beta_3 Coinventor\ geographic\ distance_{ij,t-1}\\
&+ \beta_4 Coinventor\ difference\ in\ average\ forward\ cites\ to\ date\ per\ patent\ _{ij,t-1}\\
&+ \beta_5 Both\ top\ 10\%\ inventors_{ij,t-1} + \ \beta_6 Coinventor\ technological\ proximity_{ij,t-1}\\
&+ \beta_7 Coinventor\ difference\ in\ average\ backward\ cites\ to\ date\ per\ patent_{ij,t-1}\\
&+ \beta_8 Coinventor\ difference\ in\ years\ of\ inventor\ experience\ to\ date_{ij,t-1}\\
&+ \beta_9 Both\ have\ 5 + years\ of\ tenure\ at\ focal\ firm_{ij,t-1} + Group\ FE + \varepsilon_{ij,t}
\end{aligned}$$

The dependent variable $Actual\ Pair_{ij,t}$ is equal to one if the inventor pair ij represents a real, first time collaboration formed at a U.S. publicly listed firm in year t (sometimes also referred to as a 'treated pair' or 'realized pair')[14] and zero otherwise.[15] For each realized inventor pairing observation, there are two associated counterfactual co-inventor pair observations (sometimes also referred to as a 'control pair') whose construction is outlined in Section 3.2 below. 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 variables of interest in this analysis are $Coinventor\ cultural\ similarity_{ij,t-1}$ and $Both\ female_{ij,t-1}$ which capture the degree to which the affinity-based characteristics of shared

[14] We focus our analysis on the very first collaboration between two individual U.S. based inventors since the decision to collaborate for the first time is unaffected by confounding factors such as experience with past collaborations and accumulated team-specific relationship capital. For clarity, we do not include cases where two inventors work together for the first time at an unlisted organization and then subsequently patent together at a U.S. publicly listed firm.

[15] Following the prior literature, we use the patent application date as an objective estimate of the time when two co-inventors commenced their new research collaboration (Jaravel et al., 2018).

cultural values and shared gender are meaningful drivers of realized co-inventor pairing (after accounting for firm and time effects as well as controls for relative inventor ability and similarity in past professional experiences). Appendix A defines all other independent variables. Following Bena & Li (2015), we use robust standard errors clustered at the 'group' level.

3.2 Counterfactual inventor pairs

In order to understand which factors significantly influence the formation of new collaborations between inventors, we utilize our unique within-firm corporate 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 set of counterfactual inventors enables us to construct counterfactual pairs of firm inventors (i.e. a set of control pairs) that, when compared against the set of realized first-time inventor collaborations, allows us to isolate the importance of affinity-based personal characteristics and other explanatory variables in driving the choice of co-inventor (see generally Gompers et al., 2016).

We commence the process of generating credible counterfactual inventor pairs by first identifying each actual pair of inventors that initiate a first-time collaboration at a U.S. publicly listed firm f in the year t (denote these co-inventors as $Inventor_{T1}$ and $Inventor_{T2}$ respectively). Next, we generate counterfactual or 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 collaborator on the following dimensions (denote these pseudo co-inventors as $Inventor_{C1}$ and $Inventor_{C2}$ respectively). First, the counterfactual co-inventor and the actual or treated co-inventor must be currently employed at the same firm f in the year t. Second, the counterfactual co-inventor must not have ever collaborated with either inventor in the treated pair prior to or during year t.[16] Third, for the firm's inventors that are available to serve as potential counterfactual control inventors after applying the first two filters, 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 co-inventors $Inventor_{C1}$ and $Inventor_{C2}$ respectively).[17] This final

[16] In our baseline analysis, we only use information that is available at the date that the first-time collaboration is formed in order to avoid the introduction of any "look ahead" bias. Nevertheless, our results are virtually identical if we instead exclude any inventors from the counterfactual control sample who ever collaborate with either inventor in the treated pair at any time during our sample period (even if this occurs after year t).

[17] 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 amongst these potential counterfactual inventors at random.

requirement helps to ensure that the counterfactual inventor has similar innovation experience and patenting productivity to the actually chosen collaborator.

As a result of implementing this procedure, our baseline pairwise dataset contains 1.297 million first time collaborations formed in the period spanning 1981 to 2011 between pairs of U.S. based inventors employed at the same U.S. publicly listed firm (i.e. treated or realized pair $Inventor_{T1}$–$Inventor_{T2}$) and 2.417 million counterfactual control pairs (i.e. pseudo pairs $Inventor_{T1}$–$Inventor_{C2}$ and $Inventor_{C1}$–$Inventor_{T2}$).[18] Once we merge in our shared culture measures and other inventor characteristics, we obtain a final pairwise dataset of 1.130 million treated pairs and 2.004 counterfactual control pairs.

3.3 Determinants of inventor collaboration decisions

Table 2 outlines the ability of various pairwise inventor characteristics to explain the observed collaboration decisions of inventors working within the same firm. We find strong evidence for homophily driven partnering decisions in internal skilled labor markets. This result applies to affinity-based pairwise traits, overlaps in technical skills and relative career experience.

With respect to affinity-based characteristics, both shared culture and gender play a key role in the decision of inventors to form new teams with each other. For example, we find that two inventors with more similar cultural origins are approximately 20% more likely to work with one another than two otherwise comparable inventors who have a one standard deviation greater distance in their cultural backgrounds. This result holds across all of the three cultural value classification schemes discussed in Section 2.3.2 and Appendix A and are uniformly significant at the 1% level.[19] Similarly, we consistently find across all specifications that, all else being equal, female inventors are about 16% more likely to choose to collaborate with other female inventors.

Beyond the affinity-based characteristics previously discussed, we unsurprisingly find that inventors that are relatively more successful (based on professional outcomes to date) are more

---

[18] In less than 1% 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 sample.

[19] When considering our findings with respect to the relationship between shared cultural values and inventor collaboration choices, it is important to note that our results are not simply driven by two inventors having the same country of origin. In Internet Appendix Table IA1, we re-run our conditional logit regressions only on the subsample of inventor pairs where the two inventors are *not* from the same country of origin (i.e. the *Same nation* dummy is equal to zero). In this subsample, we still find that the coefficient on our three shared culture measures remains strongly positive and significant. This provides further support for the notion that shared cultural value systems are a nuanced yet critical determinant of inventor collaboration choices.

likely to collaborate with one another in the future (Herkenhoff et al., 2018). In particular, we first find that inventors that exhibit a greater disparity in the number of subsequent citations made to their previous work are significantly less likely to collaborate with one another. Furthermore, we document that two inventors are much more likely to work with each other if the total number of patents that they have generated across their career to date places them in the top 10% of their profession (so-called "star" inventors). This implies that an inventor's patenting history and observable success to date has crucial signaling value for potential collaborators about that inventor's inherent ability and thus their desirability as a future research partner.

A natural follow-on question is whether the effect of cultural similarity on the collaboration choices of inventors is still present for "star" inventors who may potentially have a larger set of potential co-inventors to choose from for future research projects. Interestingly, while we observe a more muted influence of cultural value similarity on the collaboration decisions of more professionally accomplished inventors, we still find in unreported regressions that otherwise comparable "star" inventors are approximately 13% more likely to work with one another if they share similar cultural values.

Inventors also appear to exhibit a strong preference to collaborate with colleagues who have previously worked in similar technology fields. For example, greater overlaps in the technological experience of potential collaborators (as captured by the cosine similarity measure *Co-inventor technological proximity*) as well as greater shared expertise in technology fields with similar lifecycle properties (as captured by the average scaled backward citations of each inventor's prior patent portfolio) strongly increase the likelihood of a realized future collaboration.

Interestingly, we also find that the seniority of R&D personnel within a given organizational structure has a considerable impact on their future collaboration choices. Specifically, it is quite rare for relatively more senior colleagues (those with 5+ years of experience at the firm) to start collaborating on new R&D projects. Indeed, the probability that two of the firm's inventors decide to work together is reduced by over 30% if both of those inventors have tenures exceeding 5 years at their current employer. This provides suggestive evidence on the key role of organizational structure in building internal co-inventor networks.[20]

[20] Whether this observed phenomenon is a conscious organizational choice (whereby firms actively encourage senior colleagues to work with and mentor more junior colleagues) or the result of unconscious biases (for example arising due to two senior colleagues being unable or unwilling to adjust their respective approaches to innovation production in order to work together) is an open question that we leave to future research.

Finally, we find that physical proximity between inventors is a key determinant of the decision to collaborate with a colleague working in a similar R&D related field. In particular, we find that inventors are far more likely to start working with one another if they live in close geographic proximity to their potential collaborator. This is consistent with the general conclusions of the networking literature that emphasizes the importance of face-to-face interactions in promoting teamwork and network formation (e.g. Morrison-Smith & Ruiz, 2020; Gera, 2013).

Overall, our results on co-inventor partnering decisions demonstrate that individual inventors are much more likely to collaborate with others who possess similar personal characteristics and professional experiences, irrespective of whether these traits are correlated with technical skills and/or inherent ability. In particular, the role of affinity-based personal characteristics such as shared cultural values and shared gender identity in shaping observed co-inventor networks within the boundaries of firm are highly significant in both statistical and economic terms.

3.4 Robustness tests

While we claim that the use of similarly accomplished inventors working at the same company at the same point in time is a valid comparison group for the treated pairs in our sample, we nevertheless test the robustness of our previous results by applying stricter filters in the construction of counterfactual control pairs.

First, since geographic proximity is an important driver of inventor partnering decisions (that may in turn correlate with other unobserved relationship-based characteristics and/or reflect differences in local economic conditions), we specify that the potential (but ultimately not chosen) co-inventor must also be located within 50 miles of the actual treated co-inventor (see column (1) of Table 3).[21] Second, given the possibility that different divisions/subsidiaries 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 not chosen) co-inventors that work in the same division or subsidiary of the firm at the same point in time (see column (2) of Table 3).[22] Finally, our most restrictive counterfactual sample requires that both treated and

[21] We use a 50-mile radius cut-off as an estimate of the high likelihood that two co-inventors work in the same corporate office location (see generally Tian, 2010). Nevertheless, our results are qualitatively unchanged if we instead use a 25 mile or 100 mile threshold.

[22] Using both the Berkeley-Fung patent assignee database and the USPTO patent database, we identify two inventors as being part of the same division/subsidiary if there is an exact match on the name and location of the patent assignee in their most recently developed patents. For example, diversified corporations such as Tesla, Inc. will usually file

counterfactual co-inventor pairs must involve two inventors who: (a) work in the same division/ subsidiary of the same company, (b) work at the same geographic location/office *and* (c) are active R&D researchers in the firm at the same point in time (see column (3) of Table 3).

Table 3 presents the conditional logit regression results from estimating equation (1) with each of these alternative counterfactual control samples. Despite the use of these more restrictive and refined definitions of counterfactual co-inventors for comparison purposes, the results in Table 3 highlight the very important role of affinity-based characteristics in influencing the decision of two inventors to collaborate with one another. With respect to shared culture, two inventors working in the same company division at the same geographic location are approximately 16% more likely to work with each other if there is a one standard deviation increase in the similarity of their cultural background and beliefs (utilizing the Hofstede cultural value classification system).[23] Similarly, a female inventor is approximately 14% more likely to choose to collaborate with another female inventor relative to an otherwise similar male inventor working in the same corporate office. These are economically and statistically significant effects, especially when assessed in the context that we are comparing pairs of similarly experienced inventors working in the same corporate division at the same geographic location at the same point in time.

To provide a sense of the economic magnitude of our results, our estimates imply that shared cultural values are at least as important as prior technical experience in explaining co-inventor matching. As such, it is clear that affinity-based characteristics such as shared cultural values are of first order importance in explaining the observed variation in inventor collaboration decisions.

patents under the specific subsidiary or division that created the new invention (such as Tesla Motors Inc., Tesla Engineering Ltd, Tesla Laboratories LLC, Tesla Nanocoatings Inc. and Tesla Electronics Inc.).

[23] In unreported results, the coefficients on our other measures of shared cultural values (based on the Schwartz culture classification system and shared trust measure) remain positive and statistically significant at the 1% level.

IV. TREATMENT EFFECT OF INVENTOR HOMOPHILY ON INNOVATION

Thus far, we have established the important role of shared inherited traits (such as shared cultural values) between individual inventors in the decision to collaborate as part of a team in a corporate R&D setting. Given the strong evidence for homophily in the collaboration decisions of corporate inventors, a natural question to explore is whether these familiarity biases enhance or impede different types of innovative output (the 'ex post treatment effect'). The identification challenge in this context is that the selection of co-inventors is not a random process. For example, inventors may intentionally target new collaborations with individuals from relatively different personal and professional backgrounds in furtherance of the pursuit of a more risky, exploration-focused innovation search strategy. As a result, any differences in the average innovation outcomes of diverse and non-diverse inventor teams may be due to *selection* effects that arise from endogenous co-inventor matching or the direct *treatment* effect (i.e. the causal impact) of inventor team diversity on team performance. Thus, in order to disentangle these potential selection and treatment effects, we use triple difference-in-difference comparisons around exogenous co-inventor deaths to identify the causal impact on inventor team diversity on teams' subsequent innovation productivity.

4.1 Quasi-natural experiment involving co-inventor deaths

A potential issue with OLS and matching estimators is that it is still possible that there are unobserved differences in team characteristics that drive both the initial decision to collaborate and subsequent team innovation production. In order 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 change in the cultural composition of the inventor team.

4.1.1 Triple difference-in-differences approach

Relying upon the premature death methodology used in Bennedsen, Nielsen, Perez-Gonzalez & Wolfenzon (2007), Azoulay, Graff Zivin & Wang (2010), Nguyen & Nielsen (2010) and Jaravel et al. (2018), we provide causal estimates of how an inventor team's innovation production would change if there were an exogenous shift in the diversity of the cultural background of its team members. Using data from the USPTO, LexisNexis Public Records and the Fold3 Social Security

Death Index, we use 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 an inventor 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 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 inventor 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 in order 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 (such factors may include, for example, the productivity shock to team skill and experience arising from a colleague's unexpected departure).

However, a unique aspect of our empirical 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 vis-à-vis 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 the date of co-inventor death (Gruber, 1994; see Section 4.1.2 for further discussion), we can use a triple difference-in-difference regression specification to isolate the causal effect of inventor team cultural diversity on team innovation output.

We use the following triple difference-in-difference empirical setup to investigate how changes in the degree of inventor team cultural diversity impacts subsequent team performance. We estimate the following 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 vis-à-vis the difference in output between treated teams whose cultural similarity decreased after a co-inventor's death relative to their matched control teams:

$$\begin{aligned} Team\ Innovation_{i,t} &= \alpha + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Treated\ Team_i \\ &+ \beta_3 After_{i,t} \times Treated\ Team_i \times Team\ Cultural\ Similarity\ Change_i \\ &+ \gamma X_{i,t} + Team\ FEs + Year\ FEs + \varepsilon_{i,t} \end{aligned} \tag{2}$$

The dependent variable $Team\ Innovation_{i,t}$ is one of the patent-based outcome measures in each year t as described in Section 2.3.1.[24] The indicator variable $After_{i,t}$ equals one for all years after the focal co-inventor's death and zero otherwise. The indicator variable $Treat\ Team_i$ 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 treated team's surviving inventors) minus the treated team's cultural similarity immediately prior to their death (which includes the eventually deceased co-inventor).[25] As a result, $Team\ Cultural\ Similarity\ Change_i$ will be positive when the treated team's cultural 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.[26] The key coefficient of interest in this regression is β_3 which compares the relative pre- and post-death impact on team innovation performance for cases where treated inventor team cultural similarity increases vis-a-vis cases where treated inventor 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.[27]

[24] 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 similar if we narrow our focus to the 5 year period pre- and post- the focal co-inventor's death.

[25] Note that treated teams only experience a change in cultural similarity associated with the removal of the decreased inventor from the treated team while control teams do not experience any change in team cultural similarity over the test period because membership of the control team remains constant over time.

[26] In 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 obtain similar results to those presented using the continuous variable *Team Cultural Similarity Change* with similar levels of statistical significance.

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

Obviously, the death of a co-inventor can induce other important changes in team-related characteristics besides team cultural similarity (for example, a colleague's death may change the level and/or diversity of team inventor experience). As such, we explicitly control for other changes in non-culture related team variables induced by the focal co-inventor's death. The vector $X_{i,t}$ comprises various control variables including team gender diversity; team geographic diversity; the average (and standard deviation) in total patents per team member; the average (and standard deviation) of co-inventor experience; the average (and standard deviation) in the number of different technology classes that each co-inventor has previously patented in; the average (and standard deviation) of backward citations per patent developed by each co-inventor and the average and standard deviation of scaled forward citations to date for patents developed by each co-inventor (please see Appendix A for further details on the construction of each of these variables).

The first step in implementing our empirical strategy involves identifying active inventor teams at U.S. publicly listed firms who suffer 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.[28] We then match this information with LexisNexis Public Records and the Fold3 Social Security Death Index to identify those deceased inventors that are:

a) Employed at a U.S. publicly listed firm at the time of their death;
b) Died between 1981 and 2011 (in order to ensure we have at least 5 years of pre-death and post-death patent output); and
c) Are no older than 60 years of age at the time of their death.[29]

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 that collaborated with the deceased inventor (whilst all team members were employed at the focal U.S. public firm) on a patent that was applied for

[28] In the inventor fields on the patent data 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 the approximate time of death. This helps to facilitate the process of matching this deceased co-inventor's patenting history and personal characteristics.

[29] 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. We chose this threshold to reduce the likelihood that the death results from a long-standing health condition.

within three years of the focal inventor's death.[30] Finally, we require that we have data on each team member's cultural heritage, gender, location and patenting history. These filters ultimately result in the identification of 1,599 treated teams that were actively collaborating at a U.S. publicly listed firm around the date of a co-worker's premature death.

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 specify that none of the potential control team's constituents have an existing collaboration with any member of a treated team in our sample (including the deceased inventor).[31] Third, we require that the potential control team is actively working together around the time of the focal co-inventor's death (namely the control team must have successfully patented together at least once in the three 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 patenting productivity as the treated team. Finally, we specify that the chosen control team must have the closest proximity to the treated inventor team in terms of cultural value diversity. We then use this control team's characteristics and patenting activity as the assumed counterfactual for how the relevant treated inventor team would have performed if they did not suffer the loss of their collaborator. After we implement this procedure for identifying counterfactual control teams, we have a final sample of 1,460 treated teams and a corresponding 1,460 counterfactual controls.[32]

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

---

[30] 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" measures in the post-treatment period when the surviving treated "team" only comprises one living inventor. Nevertheless, we obtain qualitatively similar results when we include two-person treated teams in our diff-in-diff analysis.

[31] This is to ensure that these control teams are not subject to any direct spill over effects arising from the death of one of their R&D colleagues at the firm.

[32] We are not able to find a suitable counterfactual control team for approximately 9% 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, the unmatched treated teams do not appear to be significantly different on observable characteristics from those that do find a matching control team.

60th birthday in 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 the surviving Mr. B. and Ms. C. and calculate all post-death team-level output (e.g. total patents) and characteristics (e.g. the diversity in cultural values) based on these two individuals. Second, we find a counterfactual control inventor 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) The control team must be "actively" working together around the time of the focal co-inventor's death (in this case, we require that the control team must have patented together at least once between 2002 and 2005);

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

(4) We select as the counterfactual control team the team that is most similar in terms of cultural value diversity to the treated team (pre-death).

4.1.2 Evidence supporting identification assumptions

As discussed previously, our triple difference-in-difference 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 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 (delineated based on the change in team cultural similarity post co-inventor death).

To assess the reasonableness of the assumptions underlying our empirical strategy, we first explore whether there are significant differences in observable characteristics between treated and control teams. As illustrated 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 (quantity, quality and explorativeness of patents produced) and observable team characteristics (including the average experience and technical expertise of team members) in the period leading up to the focal co-inventor's death.

To more directly evaluate the reasonableness of the key identifying assumption underlying our triple difference-in-difference empirical strategy, we next compare the pre-treatment innovation output and characteristics of treated inventor teams that are split 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 the set of treated inventor teams whose cultural similarity increases post-death vis-à-vis the set of treated teams whose cultural similarity does not increase post-death in terms of their innovation output and observable team characteristics in the period leading up to the focal co-inventor's death. Furthermore, in Panel E of Table 4, we show that there does not appear to be any significant difference in the observable characteristics and productivity of the individual deceased inventors who exogenously depart treated teams whose cultural similarity increases post their death vis-à-vis the characteristics of deceased individuals at treated teams whose cultural similarity does not increase post their death. This is consistent with the idea that the random distribution of inventor deaths across time and individuals represents an exogenous shock to team cultural value diversity that is not systematically correlated with changes in team quality and other unobserved team characteristics.

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

4.1.3 Empirical results

Table 5 provides the results of our triple difference-in-difference research design. First, as shown in column (1) of Table 5, we find that teams that experience an increase/(decrease) in the level of team cultural value similarity (i.e. become less/(more) culturally diverse) significantly increase/(decrease) the quantity of patents produced. This is consistent with the organization behavioural theory that it is relatively easier to produce immediate, quantifiable output when team members share common perspectives. Second, we find that the quality of patents produced by more or less culturally diverse teams (as measured using average forward patent citations and the average market value of each patent in columns (2) and (3) of Table 5 respectively) does not significantly change between the pre- and post-treatment period. Third, we document in columns (4) and (5) of Table 5 that teams who become more/(less) diverse after a co-inventor's death are relatively more likely to produce more exploratory/(exploitative) patents. This is consistent with the idea that teams with greater homogeneity in their cognitive thought processes and beliefs are more likely to search in their common known areas of technological expertise for more incremental improvements while teams that are more diverse are more likely to engage in risky research pursuits outside the boundaries of existing knowledge. Finally, we find evidence to support the hypothesis that while more culturally diverse teams are more likely to engage in failed research endeavours (whether in terms of a decrease in the number of patents generated and/or an increased likelihood of producing patents that are not highly valued by the scientific or investor community), more culturally diverse teams appear to have a greater ability to produce breakthrough innovations (as measured by the number of team patents that fall in the top 5% or top 10% of the patent citation distribution).

Overall, our evidence strongly suggests that the level of diversity in an inventor team's cultural values has a first order effect on team productivity. In particular, we show that more culturally homogenous teams tend to produce a higher quantity of patents that are more likely to exploit existing technologies and become moderately successful inventions while 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 inventors with more diverse cultural backgrounds may encounter greater difficulties in successfully synthesizing different viewpoints and communication preferences (leading to less total patenting), the successful combination of these more novel and disparate perspectives can foster more exploratory and high-impact innovation (Choudhury & Kim, 2019; Uhlbach & Anckaert, 2020).

4.1.4 Alternative heterogeneous treatment effects specification (treated teams only)

A potential objection to the triple differences-in-differences with matched controls approach used previously is that it may be unreasonable to rely on any comparisons of the innovation trajectory of a treated inventor team that suffers the upheaval and numerical disadvantage associated with a co-inventor's death with the innovation trajectory of a counterfactual control team that does not experience such turmoil. As such, we employ an alternative heterogeneous treatment effects specification that only examines the treated team subsample, focusing on the comparison of innovation outcomes for treated teams that experience an increase in cultural value similarity after a co-inventor's death relative to treated teams that experience a decrease in the team's similarity of cultural values post a team member's death.

There are several advantages to this alternative heterogeneous treatment effects specification. First, all of the teams in this subsample analysis experience an exogenous shock to their personal and professional composition arising from a co-inventor's premature death. Second, we can include dead inventor fixed effects to control for a deceased individual's unique accumulated human capital and personal/professional traits. As such, this empirical test involves comparing sets of treated teams that are actively working together at the same firm and who suffer the loss of the *exact same* co-inventor but where the cultural similarity of one treated team increases after the co-inventor's death while the cultural similarity of the other treated team instead decreases. In other words, by keeping the loss of inventor-specific human capital/skill constant (because both teams experience the same individual co-inventor death), we can better isolate the unique effect that changes in inventor team cultural value diversity exert on inventor team performance.

We use the following regression specification to undertake this alternative test:

$$\begin{aligned} Team\ Innovation_{i,t} = \alpha + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Team\ Cultural\ Similarity\ Change_i \\ + \gamma X_{i,t} + Team\ FEs + Dead\ Inventor\ FEs + Year\ FEs + \varepsilon_{i,t} \quad (3) \end{aligned}$$

Where the dependant variable $Team\ Innovation_{i,t}$ is one of the patent-based 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 (based on the treated team's surviving inventors) minus the treated team's cultural similarity pre-death (which includes the eventually deceased co-inventor). As a result, $Team\ Cultural\ Similarity\ Change_i$ will be positive when the treated team's cultural 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.[33] We include team fixed effects to difference away any time-invariant team-level characteristics, dead inventor fixed effects to control for an individual deceased inventor's traits and experiences and year fixed effects to absorb any common

[33] In 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 obtain similar results to those presented using the continuous variable *Team Cultural Similarity Change* with similar levels of statistical significance.

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 used in Section 4.1.1 and are designed to control for any changes in non-culture related team characteristics over the test period.

Table 6 presents coefficient estimates from equation (3) using only treated teams that suffered the exogenous shock of losing an active co-inventor. Consistent with the results of the triple difference-in-difference with matched controls specification in Section 4.1.3, we see that inventor teams that experience an exogenous decline in team cultural diversity (i.e. the remaining inventor team is more similar in terms of cultural values than the original, pre-death inventor team) tend to produce a higher quantity of more exploitative, moderately cited patents. Conversely, we observe that inventor teams that experience an exogenous increase in the teams' cultural diversity (i.e. the remaining inventor team's cultural backgrounds are less similar than the original, pre-death team) tend to shift their patenting towards the production of risky, more explorative patents that have a greater probability of falling into the upper tails of the citation distribution. Once again, this implies a large degree of specialization in the generation of new, breakthrough technologies by more culturally diverse inventor teams.

V. CONCLUSION

In our paper, we open the "black box" process of corporate innovation production by focusing on the most important input into the firm's R&D process, namely the individual employees tasked with developing new inventions. Using information on over two million inventors employed at U.S. public firms, we investigate how individual inventors' inherited traits (particularly shared cultural values) and acquired career experiences affect their desire to collaborate with one another in a corporate R&D team setting and how shared cultural values amongst R&D team members affects innovative output. We provide novel evidence that, even amongst groups of comparably experienced inventors working for the same firm in the same local office, inventors who share similar cultural values are significantly more likely to work together on new research projects. Second, utilizing exogenous shocks to inventor team composition arising from premature co-inventor deaths, we find that more culturally homogenous teams produce a higher quantity of patents that are more likely to exploit existing technologies and become moderately successful inventions. In contrast, more culturally diverse inventor teams tend to produce a higher share of risky, more exploratory patents that have a greater chance of becoming high impact innovations.

Overall, our results have important implications for the implementation of policies designed to promote corporate innovation in R&D intensive yet culturally diverse workplace environments. For example, the strong homophily biases in inventor collaboration choices 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. In addition, our study suggests that cultural diversity has a non-monotonic relationship with team productivity such that firms may have a different (possibly time-varying) optimal mix of diverse and homogenous inventor teams in order to execute the firm's chosen innovation search strategy.

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., Bloom, N., Blundell, R., Griffith, R. and P. Howitt. 2005. Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics* 120(2): 701–728.

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

Agrawal, A., D. Kapur and J. McHale. 2008. How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data. *Journal of Urban Economics* 64(2): 258–269.

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., D. Daminelli and C. Fracassi. 2012. Lost in Translation? The Effect of Cultural Values on Mergers Around the World. *Journal of Financial Economics* 117, 165-189.

Alexander, L. and D. van Knippenberg. 2014. Teams in Pursuit of Radical Innovation: A Goal Orientation Perspective. *Academy of Management Review* 39(4): 423–438.

An, H., C. Chen, Q. Wu & T. Zhang. 2020. Corporate Innovation: Do Diverse Boards Help? *Journal of Financial and Quantitative Analysis*, forthcoming.

Anderson, R., Reeb, D., Upadhyay, A. and W. Zhao. 2011. The Economics of Director Heterogeneity. *Financial Management* 40(1): 5–38.

Atanassov, J. and Liu, X. 2018. Can Corporate Income Tax Cuts Stimulate Innovation? *Journal of Financial and Quantitative Analysis*, forthcoming.

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

Baghai, R., R. Silva and L. Ye. 2019. Teams and Bankruptcy. Working Paper, London Business School.

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

Bennedsen, M., K. M. Nielsen, F. Perez-Gonzalez and D. Wolfenzon. 2007. Inside the Family Firm: The Role of Families in Succession Decisions and Performance. *Quarterly Journal of Economics* 122(2): 647–691.

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

Bernstein, S. 2015. Does Going Public Affect Innovation? *Journal of Finance* 70(4): 1365–1403.

Bernstein, S., McQuade, T. and R. Townsend. 2019. Do Household Wealth Shocks Affect Productivity? Evidence from Innovative Workers During the Great Recession. *Journal of Finance*, forthcoming.

Blau, P. 1977. *Inequality and Heterogeneity*. New York: Free Press.

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

Choudhury, P. and D. Kim. 2019. The Ethnic Migrant Inventor Effect: Codification and Recombination of Knowledge Across Borders. *Strategic Management Journal* 40(2): 203–229.

Cohen, L., A. Frazzini and C. Malloy. 2008. The Small World of Investing: Board Connections and Mutual Fund Returns. *Journal of Political Economy* 116(5): 951 – 979.

Custodio, C., Ferreira, M. and P. Matos. 2019. Do General Managerial Skills Spur Innovation? *Management Science* 65(2): 459–476.

De Dreu, C. 2006. When Too Little or Too Much Hurts: Evidence for a Curvelinear Relationship between Task Conflict and Innovation in Teams. *Journal of Management* 32(1): 83–107.

Delis, M., C. Gaganis, I. Hasan and F. Pasiouras. 2017. The Effect of Board Directors from Countries with Different Genetic Diversity Levels on Corporate Performance. *Management Science* 63(1): 231–249.

Doran, K., A. Gelber and A. Isen. 2016. The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries. Working Paper, University of Notre Dame.

Eesley, C., Hsu, D. and Roberts, E. 2014. The Contingent Effects of Top Management Teams on Venture Performance: Aligning Founding Team Composition with Innovation Strategy and Commercialization Environment. *Strategic Management Journal* 35: 1798–1817.

Fang, X., Tian, X. and S. Tice. 2014. Does Stock Liquidity Enhance or Impede Firm Innovation? *Journal of Finance* 69(5): 2085–2125.

Fitzgerald, T., Balsmeier, B., Fleming, L. and G. Manso. 2019. Innovation Search Strategy and Predictable Returns. *Management Science*, forthcoming.

Fitzgerald, T. 2020. 'Til Death Do Us Part: The Relative Merits of Founder CEOs. Working Paper, Texas A&M Mays Business School.

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

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

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

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

Gompers, P., K. Huang and S. Wang. 2017. Homophily in Entrepreneurial Team Formation. Working Paper, Harvard Business School.

Griffin, D., K. Li and T. Xu. 2020. Board Gender Diversity and Corporate Innovation: International Evidence. *Journal of Financial and Quantitative Analysis*, forthcoming.

Gruber, J. 1994. The Incidence of Mandated Maternity Benefits. *American Economic Review* 84(3): 622–641.

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

Hall, B., Jaffe, A. and M. Trajtenberg. 2005. Market value and patent citations. RAND Journal of Economics 36: 16–38.

He, J. and X. Tian. 2013. The Dark Side of Analyst Coverage: The Case of Innovation. *Journal of Financial Economics* 109(3): 856–878.

Hedge, D. and J. Tumlinson. 2014. Does Social Proximity Enhance Business Partnerships? Theory and Evidence from Ethnicity's Role in U.S. Venture Capital. *Management Science* 60(9): 2355–2380.

Herkenhoff, K., Lise, J., Menzio, G. and G. Phillips. 2018. Production and Learning in Teams. Working Paper, University of Minnesota.

Hirshleifer, D., A. Low and S. Teoh. 2012. Are Overconfident CEOs Better Innovators? *Journal of Finance* 67(4): 1457–1498.

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

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

Jehn, K., Northcraft, G. and M. Neale. 1999. Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups. *Administrative Science Quarterly* 44: 741–763.

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.

Kerr, W. and W. Lincoln. 2010. The Supply Side of Innovation: H1-B Visa Reforms and U.S. Ethnic Invention. *Journal of Labor Economics* 28(3): 473–508.

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

Lanjouw, J. and M. Schankerman. 2004. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. *The Economic Journal Volume* 114: 441–465.

Li, K. and J. Wang. 2020. Acquiring Talent and Recombination via Mergers and Acquisitions. Working Paper, University of British Columbia Sauder School of Business.

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.

March, J. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* 2(1): 71–87.

Mateos, P. 2007. A Review of Name-Based Ethnicity Classification Methods and Their Potential in Population Studies. *Population, Space and Place* 13(4): 243–263.

Matsa, D. and A. Miller. 2013. A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics* 5(3): 136–169.

Morrison-Smith, S. and J. Ruiz. 2020. Challenges and Barriers in Virtual Teams: A Literature Review. *SN Applied Sciences* 2: 1096–1129.

Nguyen, B.D. and K.M. Nielsen. 2010. The Value of Independent Directors: Evidence of Sudden Deaths. *Journal of Financial Economics* 98(3): 550–567.

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.

Paikeday, T., H. Sachar and A. Stuart. 2019. A Leader's Guide: Finding and Keeping Your Next Chief Diversity Officer. *Russell Reynolds Associates*.

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

Rossiter, W. 1909. *A Century of Population Growth, from the First Census of the United States to the Twelfth, 1790–1900*. Washington, D.C.: Government Printing Office.

Schubert, T. and S. Tavassoli. 2020. Product Innovation and Educational Diversity in Top and Middle Management Teams. *Academy of Management Journal* 63(1): 272–294.

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.

Song, J., P. Almedia and G. Wu. 2003. Learning-By-Hiring: When is mobility more likely to facilitate inter-firm knowledge transfer? *Management Science* 49(4): 351–365.

Stiglitz, J. and B. Greenwald. 2015. *Creating a Learning Society: A New Approach to Growth, Development and Social Progress*. New York: Columbia University Press.

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

Toole, A., A. Myers, S. Breschi, E. Ferrucci, F. Lissoni, E. Miguelez, V. Sterzi and G. Tarasconi. 2019. Progress and Potential: A Profile of Women Inventors on U.S. Patents. *USPTO – Office of the Chief Economist IP Data Highlights* 2: 1–18.

Uhlbach, W. and P. Anckaert. 2020. In Search for New Knowledge: When Does Hiring Foreign R&D Workers Foster Exploration. Working Paper, Copenhagen Business School.

Van Knippenberg, D. and M. Schippers. 2007. Work Group Diversity. *Annual Review of Psychology* 58: 515–541.

Williams, K. Y. and C. A. O'Reilly. 1998. Demography and Diversity in Organizations: A Review of 40 Years of Research. *Research in Organizational Behavior* 20: 77–140.

Table 1: Summary statistics

This table reports summary statistics for the entire sample of individual inventors and inventor teams at U.S. public firms from 1981 to 2011. Panel A presents descriptive statistics for individual U.S. based inventors working at U.S. publicly listed firms. Panel B outlines descriptive statistics for inventor teams that comprise U.S. based inventors and that are first formed whilst employed at a U.S. publicly listed firm. 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. Appendix A outlines the definition of the variables listed.

Panel A: Individual lifetime inventor characteristics

	Mean	Median	Std. dev.
Female	0.12	0.00	0.31
Total lifetime patents	5.68	2.00	15.97
Years of inventor experience (lifetime)	5.72	4.00	7.14
Technology class experience (no. of classes)	2.12	1.00	3.08
Top 10% Inventor	0.18	0.00	0.41

Panel B: Inventor team characteristics

	Mean	Median	Std. dev.
Size of inventor team per patent	2.59	2.00	1.94
Number of teams per inventor (lifetime)	2.80	1.00	4.56
Distinct co-inventors per inventor (lifetime)	3.75	2.00	4.09

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

	Mean	Median	Std. dev.
Co-inventor cultural similarity – Hofstede	0.62	0.60	0.24
Co-inventor cultural similarity – Schwartz	0.67	0.66	0.23
Co-inventor cultural similarity – Trust only	0.85	0.89	0.16
Same nation	0.19	0.00	0.39
Both female	0.02	0.00	0.12
Co-inventor geographic distance (miles)	214.10	20.11	537.75
Co-inventor difference in avg. forward cites to date per patent	1.06	0.59	1.91
Both top 10% inventors	0.07	0.00	0.25
Co-inventor tech proximity	0.22	0.00	0.36
Co-inventor difference in avg. backward cites to date per patent	0.90	0.60	1.42
Co-inventor difference in years of inventor experience to date	6.60	5.00	6.65
Both have 5+ years of tenure at focal firm	0.17	0.00	0.37

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. The dependent variable is equal to one for all new, actually formed co-inventor pairwise relationships and zero for the counterfactual pairs that form the comparison/control group (based on inventors working at the same firm in the same year as the new actually formed co-inventor pair: see Section 3.2 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 (2)	Same Firm Same Year (3)
Co-inventor cultural similarity – Hofstede	0.205***		
	(0.006)		
Co-inventor cultural similarity – Schwartz		0.240***	
		(0.007)	
Co-inventor cultural similarity – Trust only			0.225***
			(0.001)
Both female	0.142***	0.146***	0.143***
	(0.014)	(0.014)	(0.014)
Co-inventor geographic distance	-0.419***	-0.419***	-0.420***
	(0.001)	(0.001)	(0.001)
Co-inventor difference in avg. forward cites to date per patent	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
Both top 10% inventors	3.097***	3.101***	3.109***
	(0.878)	(0.877)	(0.885)
Co-inventor tech proximity	1.769***	1.772***	1.769***
	(0.007)	(0.008)	(0.007)
Co-inventor difference in avg. backward cites to date per patent	-0.019***	-0.019***	-0.019***
	(0.001)	(0.001)	(0.001)
Co-inventor difference in years of inventor experience to date	-0.130***	-0.129***	-0.130***
	(0.003)	(0.003)	(0.003)
Both have 5+ years of tenure at focal firm	-0.472***	-0.472***	-0.471***
	(0.006)	(0.006)	(0.006)
Group fixed effects?	Yes	Yes	Yes
No. of observations (mil)	3.134	3.042	3.063
No. of actual pairs (mil)	1.130	1.102	1.108
No. of counter-factual pairs (mil)	2.004	1.940	1.955
Pseudo R^2	0.25	0.25	0.25

Table 3: Determinants of choice of co-inventor – Alternative counterfactual controls

This table reports the results of conditional logit models that estimate the factors affecting the choice of co-inventor, using alternative specifications for defining counterfactual control pairs. The dependent variable is equal to one for all new, actually formed co-inventor pairwise relationships and zero for the counterfactual pairs that form the comparison/control group. Column (1) uses other inventors working at the same firm, in the same year and at the same location to form the counterfactual control pairs. Column (2) uses other inventors working in the same division/subsidiary of a firm at the same time to form the counterfactual control pairs. Column (3) uses other inventors working in the same division/subsidiary of a firm, in the same year and at the same location to form the counterfactual control pairs (see Section 3.4 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 Same Location	Same Division Same Year	Same Division Same Year Same Location
	(1)	(2)	(3)
Co-inventor cultural similarity – Hofstede	0.180*** (0.006)	0.208*** (0.006)	0.176*** (0.006)
Both female	0.163*** (0.012)	0.156*** (0.014)	0.160*** (0.012)
Co-inventor geographic distance	-0.121*** (0.001)	-0.379*** (0.001)	-0.110*** (0.001)
Co-inventor difference in average forward cites to date per patent	-0.003*** (0.001)	-0.003*** (0.001)	-0.004*** (0.001)
Both top 10% inventors	1.429*** (0.247)	3.347*** (0.920)	1.359*** (0.214)
Co-inventor tech proximity	1.450*** (0.006)	1.589*** (0.007)	1.350*** (0.006)
Co-inventor difference in average backward cites to date per patent	-0.016*** (0.001)	-0.017*** (0.001)	-0.016*** (0.001)
Co-inventor difference in years of inventor experience to date	-0.123*** (0.003)	-0.128*** (0.003)	-0.118*** (0.003)
Both have 5+ years of tenure at focal firm	-0.384*** (0.005)	-0.450*** (0.006)	-0.367*** (0.005)
Group fixed effects?	Yes	Yes	Yes
No. of observations (mil)	2.869	3.090	2.818
No. of actual pairs (mil)	1.077	1.121	1.063
No. of counter-factual pairs (mil)	1.792	1.969	1.755
Pseudo R^2	0.05	0.21	0.04

Table 4: Evidence supporting 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 time that do not experience the loss of a team member ('control' teams) (see Section 4.1.1 for additional details for the construction of control teams). Panel A compares the average output of treated and control firms in the 3 years before the focal co-inventor death while Panel B compares the average pre-treatment characteristics of treated and control teams. In contrast, Panels C and D focus on comparing inventor teams whose cultural similarity increases post their co-inventor's death with inventor teams whose cultural similarity does not increase post their colleague's death. Panel C and Panel D compares these teams' 3-year average pre-treatment output and characteristics at the time of co-inventor death respectively. 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 vis-à-vis the characteristics of focal deceased inventors at treated teams whose cultural similarity does not increase after the focal co-inventor's death. Please refer to Appendix A for definitions of all outcome 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 3-year period prior to the focal co-inventor's death

	Treated team Mean	Control team Mean	Difference
Total patents	0.39	0.38	0.01
Average forward cites per patent	0.49	0.51	-0.02
Average market value of patents	4.97	4.53	0.44
Average backward cites per patent	0.54	0.51	0.03
Average claims per patent	0.38	0.39	-0.01
Top 5% cited patents	0.05	0.05	0.00

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

	Treated team Mean	Control team Mean	Difference
Team cultural similarity (Hofstede)	0.52	0.53	-0.01
Team gender diversity	0.20	0.19	0.01
Team geographic diversity	250.52	232.57	17.95
Team average total number of patents to date	13.66	13.11	0.55
Team average inventor experience to date	8.51	8.64	-0.13
Team average technology class experience to date	2.90	2.79	0.11
Team average backward citations to date per patent	1.21	1.26	-0.05
Team average forward citations to date per patent	1.15	1.21	-0.06

Panel C: Average annual output in the 3-year period prior to the focal co-inventor's death

	Treated team (Cultural sim. does increase) Mean	Treated team (Cultural sim. does not increase) Mean	Difference
Total patents	0.48	0.49	-0.01
Average forward cites per patent	0.52	0.49	0.03
Average market value of patents	4.28	4.64	-0.36
Average backward cites per patent	0.56	0.61	-0.05
Average claims per patent	0.41	0.40	0.01
Top 5% cited patents	0.06	0.05	0.01

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

	Treated team (Cultural sim. does increase) Mean	Treated team (Cultural sim. does not increase) Mean	Difference
Team cultural similarity (Hofstede)	0.48	0.50	-0.02
Team gender diversity	0.20	0.18	0.02
Team geographic diversity	294.72	267.58	27.14
Team average total number of patents to date	14.59	13.48	1.11
Team average inventor experience to date	8.84	8.52	0.32
Team average technology class experience to date	2.87	2.91	-0.04
Team average backward citations to date per patent	1.21	1.28	-0.07
Team average forward citations to date per patent	1.34	1.29	0.05

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

	Treated team (Cultural sim. does increase) Mean	Treated team (Cultural sim. does not increase) Mean	Difference
Total patents	10.13	9.84	0.29
Years of inventor experience	8.96	8.53	0.43
Technology class experience (no. of classes)	3.64	3.50	0.14
Average scaled backward citations per patent	1.37	1.50	-0.13
Average scaled forward citations per patent	1.51	1.54	-0.03

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 (2) in Section 4.1.1. $After_{i,t}$ 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 remaining 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 alternatively measured as *Ln(1+Average forward citations per patent)* and *Ln(1+Average market value of patents)*. A team is designated as being more focused on innovation exploitation 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)*. We measure the propensity for producing high impact innovations as *Ln(1+Top 5% cited patents)* and *Ln(1+Top 10% cited patents)*. Inventor team average controls comprise the team-level mean of each team members' total number of patents produced to date, inventor experience and technology class experience, average forward citations to date per patent and average backward citations per patent. Inventor team diversity controls comprise team gender diversity and team geographic diversity as well as the standard deviation (i.e. diversity) in team members' total patents to date, inventor experience and technology class experience, average forward citations to date per patent and average backward cites per patent. Please refer to Appendix A for the definitions of dependent and independent 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 or value of innovation		Relative focus on exploitative innovation		Likelihood of producing high impact innovation	
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Avg market value of patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Top 5% cited patents)	Ln(1+Top 10% cited patents)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$After_{i,t}$	-0.261***	-0.240***	-0.523***	-0.256***	-0.243***	-0.033***	-0.059***
	(0.005)	(0.006)	(0.013)	(0.007)	(0.004)	(0.002)	(0.003)
$After_{i,t}$ × Treated $Team_i$	-0.002	0.007	-0.022	-0.002	0.006*	0.002	0.000
	(0.004)	(0.004)	(0.015)	(0.005)	(0.004)	(0.002)	(0.002)
$After_{i,t}$ × Treated $Team_i$ ×	0.079**	0.029	0.066	0.095**	0.079***	-0.017**	-0.015*
Team Cultural Similarity $Change_i$	(0.033)	(0.033)	(0.062)	(0.038)	(0.025)	(0.008)	(0.008)
Inventor team average controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inventor team diversity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	63,474	63,474	63,474	63,474	63,474	63,474	63,474
Adjusted R^2	0.16	0.11	0.10	0.13	0.12	0.11	0.12

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 (3) in Section 4.1.4 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 remaining 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 alternatively measured as *Ln(1+Average forward citations per patent)* and *Ln(1+Average market value of patents)*. A team is designated as being more focused on innovation exploitation 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)*. We measure the propensity for producing high impact innovations as *Ln(1+Top 5% cited patents)* and *Ln(1+Top 10% cited patents)*. Inventor team average controls comprise the team-level mean of each team members' total number of patents produced to date, inventor experience and technology class experience, average forward citations per patent and average backward citations per patent. Inventor team diversity controls comprise team gender diversity and team geographic diversity as well as the standard deviation (i.e. diversity) in team members' total patents to date, inventor experience and technology class experience, average forward citations to date per patent and average backward citations per patent. Please refer to Appendix A for the definitions of dependent and independent variables used in this analysis. All regression specifications include Team fixed effects, Individual dead co-inventor 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 or value of innovation		Relative focus on exploitative innovation		Likelihood of producing high impact innovation	
Dependent variable	Ln(1+Total patents)	Ln(1+Avg forward cites per patent)	Ln(1+Avg market value of patents)	Ln(1+Avg back cites per patent)	Ln(1+Avg claims per patent)	Ln(1+Top 5% cited patents)	Ln(1+Top 10% cited patents)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After$_{i,t}$	-0.309***	-0.259***	-0.533***	-0.303***	-0.268***	-0.036***	-0.065***
	(0.008)	(0.008)	(0.018)	(0.009)	(0.005)	(0.003)	(0.008)
After$_{i,t}$ × Team Cultural Similarity Change$_i$	0.071**	0.020	0.044	0.086**	0.070***	-0.017*	-0.019*
	(0.033)	(0.033)	(0.061)	(0.038)	(0.024)	(0.009)	(0.011)
Inventor team average controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inventor team diversity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dead co-inventor fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	34,894	34,894	34,894	34,894	34,894	34,894	34,894
Adjusted R^2	0.16	0.11	0.10	0.13	0.11	0.12	0.11

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 patent portfolio 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 the focal patent receives divided by the average number of citations made to patents applied for in the same year and CPC technology sub-class.
Average market value of patents	The average dollar value of (subsequently granted) patents (in millions of nominal U.S. dollars) that an inventor team applies for in a given year (see methodology and data in Kogan et al., 2017).
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.
Top 5% cited patents	A patent is classed as a top 5% cited patent if it receives citations from other patents that place the focal patent in the top five percent of patents in the same application 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.
Panel B: Individual inventor and inventor team characteristics	
Female	A dummy variable equal to one if the inventor is a female and zero otherwise (see methodology and data in Toole et al., 2019).
Total lifetime patents	A count of the number of patents granted to an inventor over their entire career.
Years of inventor experience (lifetime)	The number of years between the application date of the first (subsequently granted) patent that an inventor ever applies for and the application date of the last (subsequently granted) patent filed for by an inventor.
Technology class experience (no. of classes)	A count of the number of distinct CPC technology sub-classes that an inventor patents in over the course of their career.
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 (lifetime)	Equals the mean of the scaled backward citations made across all of patents developed over an inventor's lifetime.
Average forward cites per patent (lifetime)	Equals the mean of the scaled forward citations received up until the end of the sample period across all of patents developed over an inventor's lifetime.
Size of inventor team per patent	The number of unique inventors listed on a particular patent.
Number of teams per inventor (lifetime)	The number of unique teams of which the focal inventor is a team member over the course of their entire career.
Distinct co-inventors per inventor (lifetime)	The number of unique co-inventors that the focal inventor works with over the course of their entire career.

Variable	Description
Panel C: Pairwise inventor characteristics	
Co-inventor cultural similarity – Hofstede	The pairwise inventor cultural similarity measure is calculated as –1 multiplied by the Euclidean distance between two inventors' cultural values based on Hofstede's framework using the following formula: $-1 \times [(PDI_{Inventor\ 1} - PDI_{Inventor\ 2})^2 + (IDV_{Inventor\ 1} - IDV_{Inventor\ 2})^2 + (MAS_{Inventor\ 1} - MAS_{Inventor\ 2})^2 + (UAI_{Inventor\ 1} - UAI_{Inventor\ 2})^2 + (LTO_{Inventor\ 1} - LTO_{Inventor\ 2})^2 + (IVR_{Inventor\ 1} - IVR_{Inventor\ 2})^2]^{\frac{1}{2}}$. This measure is transformed so that its values are bound between zero and one, where higher values indicate higher cultural similarity between inventor 1 and inventor 2.
Co-inventor cultural similarity – Schwartz	Measured as –1 multiplied by the Euclidean distance between two inventors based on the cultural framework of Schwartz (1992) that classifies national culture into seven dimensions: harmony, embeddedness, hierarchy, mastery, affective autonomy, intellectual autonomy, and egalitarianism. This measure is transformed so that its values are bound between zero and one, where higher values indicate higher cultural similarity between inventor 1 and inventor 2.
Co-inventor cultural similarity – Trust only	The absolute difference between inventor 1's and inventor 2's trust beliefs, where trust is calculated as the average response in each country of origin to the following question in the World Value Survey from 1982 to 2008: "*Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?*" This measure is transformed so that its values are bound between zero and one, where higher values indicate higher cultural similarity between inventor 1 and inventor 2.
Same nation	A dummy variable equal to one if both inventors have the same country of origin and zero otherwise.
Both Female	A dummy variable equal to one when both inventors are female and zero otherwise.
Co-inventor geographic distance	The natural logarithm of one plus the geodetic distance (in miles) between the two inventors' then current locations (where inventor locations are identified from current and previous 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 of the inventor's patents to date.
Both top 10% inventors	A dummy variable equal to one when both inventors classify as top 10% inventors and zero otherwise.
Co-inventor technological proximity	Cosine similarity measure following Jaffe (1989) that computes the degree of overlap in CPC patent technology classes between patents that are applied for (and subsequently granted) to inventor A up to and including year *t*-1 and the patents that are applied for (and subsequently granted) to inventor B 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 of the inventor's patents to date.
Co-inventor difference in years of inventor experience to date	Equals the absolute difference between each inventor's years of inventor experience to date. Years of inventor experience to date equals the number of years between the application date of the first (subsequently granted) patent that an inventor ever applies for and the current year.

Variable	Description
Panel C: Pairwise inventor characteristics (cont.)	
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. publicly listed firm and zero otherwise. The tenure of an inventor at a firm equals the number of years between the application date of the first (subsequently granted) patent that an inventor applies for whilst working at the focal employer and the current year.
Panel D: Inventor team average and diversity control variables	
Team cultural similarity	To measure the culture similarity within a team of N inventors, we use the following formula: $-1 \times \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(culture_{Inventor\ i} - culture_{Team\ mean})^2}$ where the distance between each inventor's culture and the inventor's team culture, $culture_{Inventor\ i} - culture_{Team\ mean}$, is calculated as $[(PDI_{Inventor\ i} - PDI_{Team\ mean})^2 + (IDV_{Inventor\ i} - IDV_{Team\ mean})^2 + (MAS_{Inventor\ i} - MAS_{Team\ mean})^2 + (UAI_{Inventor\ i} - UAI_{Team\ mean})^2 + (LTO_{Inventor\ i} - LTO_{Team\ mean})^2 + (IVR_{Inventor\ i} - IVR_{Team\ mean})^2]^{\frac{1}{2}}$. This team cultural similarity measure is transformed so that its values are bound between zero and one, where higher values indicate higher cultural similarity or less cultural diversity amongst the R&D team members.
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 average inventor experience to date	The average of each team member's inventor experience to date.
Team average technology class experience to date	The average of each team member's technology class experience to date (based on a count of the number of distinct CPC technology sub-classes that an inventor patents in over their career to date).
Team average forward citations to date per patent	The average of each team member's average forward citations to date per patent.
Team average backward citations per patent	The average of each team member's average backward citations per patent.
Team gender diversity	Within a team of N inventors, the Blau (1977) Index of Gender Diversity equals $2 \times (1 - \sum_{s=1}^{S}(p_s)^2)$ where s = male or female and p_s is the proportion of inventors in the team that are male or female respectively. The value of this measure is bound between zero and one.
Team geographic diversity	The average of the pairwise co-inventor geographic distances between all of the members of the inventor team.
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.
Team diversity in inventor experience to date	The standard deviation in the inventor experience acquired by each team member to date.
Team diversity in technology class experience to date	The standard deviation in the technology class experience of each team member to date.
Team diversity in average forward citations to date per patent	The standard deviation in the average forward citations to date per patent of each team member.
Team diversity in average backward citations per patent	The standard deviation in the average backward citations per patent of each team member.

Appendix B: Surname Matching Procedure

Inventors' countries of ancestry are identified using their surnames similar to the methodology of Lauderdale and Kestenbaum (2000). Following Liu (2016), we use two main sources to identify the country of origin of surnames in a systematic way. First, we use U.S. Census records from 1850 to 1940. These records represent the complete set of Census records available to the public in which the respondents' names are disclosed since they are no longer subject to the 72-year confidentiality rule. For several of these datasets (1880, 1920, 1930, 1940), we acquired access to 100% of the records through the Minnesota Population Center. For the other years, only 1% of the records are currently available. To identify the country of origin of surnames, 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, which yields 54 million census records. 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.

Second, we use the surname-ancestry country matching list from a commercial database. Origins Info Ltd., a well-known commercial vendor of name classification services, 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).

To create the final matching list, we do the following. 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. Second, we keep surnames for which the most frequently associated country of birth appears in more than 75% of the census records. 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. 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 the list of inventors from the patent database.

INTERNET APPENDIX

Table IA1: 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 (i.e. the Same nation dummy is equal to 0), using alternative specifications for defining counterfactual control pairs. The dependent variable is equal to one for all new, actually formed co-inventor pairwise relationships and zero for the counterfactual pairs that form 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 location 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 location to form the counterfactual control pairs (see Sections 3.2 and 3.4 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 Location (2)	Same Division Same Year (3)	Same Division Same Year Same Location (4)
Co-inventor cultural similarity – Hofstede	0.161*** (0.012)	0.139*** (0.010)	0.164*** (0.011)	0.141*** (0.010)
Both female	0.144*** (0.017)	0.170*** (0.015)	0.150*** (0.016)	0.169*** (0.015)
Co-inventor geographic distance	-0.421*** (0.001)	-0.123*** (0.002)	-0.381*** (0.001)	-0.112*** (0.002)
Co-inventor difference in average forward cites to date per patent	-0.002*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.004*** (0.001)
Both top 10% inventors	3.338*** (1.138)	1.618*** (0.318)	3.804*** (1.327)	1.382*** (0.264)
Co-inventor tech proximity	1.799*** (0.009)	1.490*** (0.008)	1.621*** (0.009)	1.391*** (0.008)
Co-inventor difference in average backward cites to date per patent	-0.020*** (0.001)	-0.018*** (0.001)	-0.018*** (0.001)	-0.016*** (0.001)
Co-inventor difference in years of experience to date	-0.131*** (0.004)	-0.128*** (0.003)	0.130*** (0.004)	-0.123*** (0.003)
Both have 5+ years of tenure at focal firm	-0.475*** (0.007)	-0.390*** (0.006)	-0.456*** (0.007)	-0.380*** (0.006)
Group fixed effects?	Yes	Yes	Yes	Yes
No. of observations (mil)	2.229	2.031	2.197	1.993
No. of actual pairs (mil)	0.865	0.811	0.856	0.800
No. of counter-factual pairs (mil)	1.364	1.220	1.341	1.193
Pseudo R^2	0.26	0.04	0.21	0.04