Received: 25 March 2026 | Revised: 25 May 2026 | Accepted: 10 June 2026

DOI: 10.1002/smj.70108

RESEARCH ARTICLE

SMS | Strategic Management Journal

WILEY

# Ethnic integration and the value of innovation

**Stefano Breschi**[1,2] | **Luisa Gagliardi**[1,2] | **Thomas Yoon**[3]

[1]Department of Management & Technology, Bocconi University, Milan, Italy

[2]ICRIOS, Bocconi University, Milan, Italy

[3]School of Management, University of Liverpool, Liverpool, UK

**Correspondence**
Stefano Breschi, Department of Management & Technology, Bocconi University, Via Sarfatti 25, 20136, Milan, Italy.
Email: stefano.breschi@unibocconi.it

**Funding information**
NextGenerationEU, Grant/Award Numbers: ECS00000037 (MUSA), ECS00000041 (GRINS)

**Abstract**

**Research Summary:** Organizations have become increasingly ethnically diverse, and this diversity often brings together distinct technological competencies. Understanding how firms can leverage such diversity to produce valuable innovations therefore becomes essential. This study investigates how ethnic integration within inventor collaboration networks, and its interaction with ethnic diversity, influence the economic value of innovation. Using patent and co-inventor data from 890 publicly traded U.S. firms (1990–2015), we construct a network-based index of ethnic integration, defined as the ratio of cross-ethnic to within-ethnic co-invention ties, and link it to stock market reactions surrounding patent grants. We find that ethnic integration significantly enhances patent value, particularly in firms that are both ethnically diverse and engaged in technologically complex innovation. These findings are robust to an instrumental-variables strategy addressing potential endogeneity concerns. By distinguishing between workforce composition and collaborative structure, our study advances the literature on diversity and innovation and underscores the critical role of integration mechanisms in unlocking the full potential of diverse organizations.

**Managerial Summary:** Organizations have become increasingly ethnically diverse, and this diversity brings distinct technological competencies; yet many still struggle to translate diversity into measurable innovation outcomes. This study shows that how diverse talent collaborates is just as critical as who is on the team. Analyzing over 20 years of patent data from 890 U.S. public companies, we find that firms where inventors frequently collaborate across ethnic lines—what we call ethnic integration—generate significantly more valuable innovations. This effect is especially strong in firms tackling complex technologies, where integrating diverse knowledge is crucial. Importantly, diversity without integration offers limited returns. For leaders, the key takeaway is clear: investing in diverse talent is not enough. Organizations must also build collaborative cultures and structures—from inclusive leadership and cross-ethnic mentorship to team design and onboarding practices—that enable integration. In high-tech and R&D-intensive industries, integration is not optional; it is a strategic lever for innovation performance.

**KEYWORDS**

ethnic diversity, innovation, integration, knowledge complexity

## 1 | INTRODUCTION

Innovation has become increasingly complex, and firms that effectively integrate diverse internal knowledge sources often generate more novel and economically valuable inventions (Ahuja & Lampert, 2001; Carnabuci & Operti, 2013; Fleming, 2001). Central to this process are individual innovators who both generate ideas and collaborate across domains, transforming creative insights into tangible inventions (Jones, 2009). However, while heterogeneous expertise among knowledge specialists is a necessary condition for superior innovation performance, firms also require integration mechanisms that bring these disparate resources together to realize their full potential (Grant, 1996; Tortoriello et al., 2012).

In this study, we examine inventors' ethnicity as a critical dimension of both knowledge diversity and its integration within organizations. Specifically, we conceptualize ethnic background as an individual lens or cognitive framework that shapes the knowledge people accumulate, the ties they form, and the ways they deploy that knowledge in the generation of innovative ideas (Gagliardi & Novelli, 2025; Koning et al., 2020).

Over the past three decades, the ethnic composition of skilled workforces in large, R&D-intensive U.S. firms has shifted dramatically: between 1994 and 2006, the ratio of foreign-born to native science and engineering workers doubled from 0.16 to 0.32 (Sana, 2010). Ethnic background often correlates with specialization in specific technological domains—for instance, Chinese inventors are overrepresented in electrical and chemical engineering, whereas Indian inventors predominate in computer science (Kerr, 2008). These patterns reflect not only differences in educational investment but also broader cultural values and career pathways (Choudhury & Kim, 2019), as seen in the pivotal role of skilled Chinese and Indian immigrants in Silicon Valley's high-tech sector (Saxenian, 2002). Consequently, ethnically diverse firms possess substantial latent potential to integrate frontier knowledge across technological domains.

Ethnicity also shapes within-firm collaboration dynamics. Several studies have documented the strong tendency among scientists and inventors to collaborate primarily within their own ethnic groups (Almeida et al., 2015; Freeman & Huang, 2015). Consequently, ethnically diverse organizations may not realize their innovation potential if co-inventor ties remain confined to homogeneous clusters rather than spanning across ethnic boundaries. Thus, the benefits of ethnic diversity for innovation hinge critically on the degree of integration within the firm's collaborative network.

To test our argument, we construct a network-based index of ethnic integration within firms, defined by the ratio of cross-ethnic to within-ethnic co-inventor ties. Ethnically integrated organizations exhibit a network configuration where any skilled worker has an equal likelihood of interacting and collaborating with any other skilled individual in the organization, regardless of their ethnic or cultural background. At the opposite extreme, segregated firms exhibit strong homophily, with collaboration networks partitioned into nearly disjoint clusters of same-ethnicity ties. We hypothesize that, for a given level of ethnic diversity, greater ethnic integration enhances the value of a firm's innovation output. Furthermore, we propose that the effect of diversity on the economic value of inventions is contingent on the degree of integration: high integration links the latent potential of a diverse workforce to its actual realization in valuable innovations.

To investigate how ethnic integration influences the economic value of inventions, we assembled a longitudinal dataset of patenting activity for 890 publicly traded U.S. firms spanning 1990–2015. We classified each inventor's ethnicity using the `Ethnea` name classifier (Torvik & Smalheiser, 2009) and constructed firm-level co-invention networks. From these networks, we derive an *ethnic integration* index defined as the ratio of cross-ethnic to same-ethnicity collaboration ties, with higher values indicating more integrated collaborative patterns. We then link this index to patent value, measured by the cumulative three-day stock price response around the patent grant (Kogan et al., 2017). Our baseline specification controls for the level of ethnic diversity, includes firm and year fixed effects, and incorporates 2-digit SIC industry-specific linear time trends to capture evolving industry conditions.

Our empirical results show that firms with higher ethnic integration produce patents with significantly greater economic value. This relationship remains robust after addressing endogeneity via an instrumental-variables strategy that instruments the firm's integration index with expected linguistic distance among inventors in its local CBSA labor market.[1] Moreover, the positive effect of integration is concentrated among firms with higher levels of ethnic diversity. Together, these findings underscore how ethnic integration and diversity reinforce each other

[1]Core Based Statistical Areas. These areas include both Metropolitan Statistical Areas (population of 50,000+) and Micropolitan Statistical Areas (population 10,000–49,999).

and that integrating inventors from different ethnic backgrounds is essential for transforming workforce diversity's latent potential into realized innovation value.

In the final section of this study, we investigate a potential mechanism driving the relationship between ethnic integration and patent value. Building on the premise that ethnic diversity correlates with distinct technological specializations, and that integration fosters collaboration among frontier specialists, we conduct two empirical exercises. First, we map specialization profiles by ethnicity across all U.S. inventors, showing that different ethnic groups concentrate in particular technological fields where they are more likely to produce frontier innovations. Second, we assess how the value-enhancing effect of ethnic integration varies with innovation complexity. Consistent with evidence that complex inventions—those recombining disparate technological domains—yield higher economic returns (Ahuja & Lampert, 2001; Fleming, 2001; Moser & Nicholas, 2004), we find that integration's benefits are especially pronounced for firms producing these complex innovations.

The remainder of the paper is organized as follows. Section 2 reviews the literature on diversity, integration, and innovation. Section 3 describes our data and empirical methodology. Section 4 presents the main results and robustness checks. Section 5 discusses theoretical and managerial implications, acknowledges study limitations, and offers directions for future research.

## 2 | CONCEPTUAL FRAMEWORK

### 2.1 | Integration: The missing link between diversity and innovation

Scholarly interest in ethnic diversity and its impact on innovation has a longstanding tradition, yet it has also produced contrasting predictions. The value-in-diversity perspective builds on the notion of knowledge complementarity, arguing that skilled immigrants from diverse cultural and educational backgrounds bring distinct cognitive frameworks and problem-solving approaches compared to natives (Alesina et al., 2016; Cox & Blake, 1991; Dahlin et al., 2005; Gagliardi, 2015; Herring, 2009; Kerr, 2008; Lazear, 1999b; McLeod et al., 1996). Such heterogeneity enhances problem-solving, yielding higher-quality solutions and boosting the productivity of the innovation process (Crescenzi & Gagliardi, 2018; Hong & Page, 2004; Simonton, 2003). Consistent with this view, empirical research shows that ethnically diverse organizations generate more frequent and impactful innovations by recombining a broader and more specialized pool of knowledge (Fleming, 2001; Freeman & Huang, 2015; Kerr & Kerr, 2018; Nathan & Lee, 2013; Rodan & Galunic, 2004). By contrast, a substantial body of research cautions that, despite its informational advantages, ethnic diversity can increase coordination costs, foster interpersonal conflict, and exacerbate communication barriers (Herring, 2009; Williams & O'Reilly, 1998). This creates a fundamental tension between coordination efficiency and the innovation potential of ethnically diverse organizations (Reagans, 2013).

We argue that integration is the missing link reconciling these opposing effects. Integration, defined as an organizational environment where individuals are equally likely to collaborate regardless of their ethnic and cultural background, acts as the enabling factor that allows organizations to capitalize on the benefits of diversity while mitigating its costs. This perspective aligns with research showing that inclusive organizational practices are essential for translating diversity into performance, and that while ethnic diversity enhances innovation potential, its realization critically depends on the level of integration within the organization (Ely & Thomas, 2001; Herring, 2009).

Extensive evidence at the micro-level of innovative teams shows that the benefits of diversity are fully realized only when members are integrated through strong collaborative ties (Reagans & Zuckerman, 2001). In inventor teams, the probability of producing high-value innovations is higher when ethnic and non-ethnic inventors collaborate, fostering creative recombination of heterogeneous knowledge (Choudhury & Kim, 2019). Extending this insight to the organizational level, integration promotes effective collaboration across diverse ethnic and cultural backgrounds, leveraging a broader array of perspectives and expertise in the innovation process (Laursen et al., 2020).

Integrated organizations exhibit collaboration networks that are not partitioned into homophilous clusters—where individuals connect primarily with similar others—but instead are characterized by configurations in which any skilled worker is equally likely to interact with any other, irrespective of ethnic or cultural background. Such cross-ethnic ties enhance knowledge sharing and co-ordination, facilitating the exchange and recombination of diverse knowledge sets (Tortoriello et al., 2012).

Bridging ties that connect diverse individuals and groups are critical for innovation and become even more important as inventions grow in complexity (Breschi & Lissoni, 2009). The "burden of knowledge" compels inventors to specialize at the frontier of narrowly defined technological domains, limiting their individual scope (Jones, 2009). Integrative collaboration across these domain boundaries, however, fuels creativity and enables firms to execute complex, multi-field innovation projects (Fleming et al., 2007). Given that high-value innovations increasingly arise from recombining disparate technological knowledge (Ahuja & Lampert, 2001; Fleming, 2001; Moser & Nicholas, 2004), integration becomes a key driver for organizations seeking sustained innovation performance.

## 2.2 | A relational approach to integration

In this study, we adopt a relational perspective on integration by drawing on Kogut and Zander's (1996) conceptualization of the firm as a social community that coordinates specialized knowledge through shared identity and relational processes. From this vantage point, ethnic integration captures the degree to which a firm's collaborative network transcends subgroup boundaries, reducing fragmentation and segregation by fostering cross-ethnic ties. These bridging connections, which span ethnic groups, enable inclusive collaboration, linking diverse cognitive resources and facilitating the transfer and recombination of specialized knowledge (Burt, 2004; Reagans & Zuckerman, 2001).

Within this framework, diversity and integration emerge as two distinct theoretical and empirical constructs. Diversity refers to the heterogeneity of individuals within the organization along dimensions such as ethnic and cultural background, reflecting the range of knowledge and perspectives available. It denotes the mere coexistence of different identities and capabilities, but does not, in and of itself, imply interaction or collaboration. Integration, by contrast, describes the degree to which those diverse actors are connected through collaborative ties that cross identity boundaries. An integrated network is characterized by robust cross-group interactions, in which members from different ethnic and cultural backgrounds actively collaborate and share knowledge.

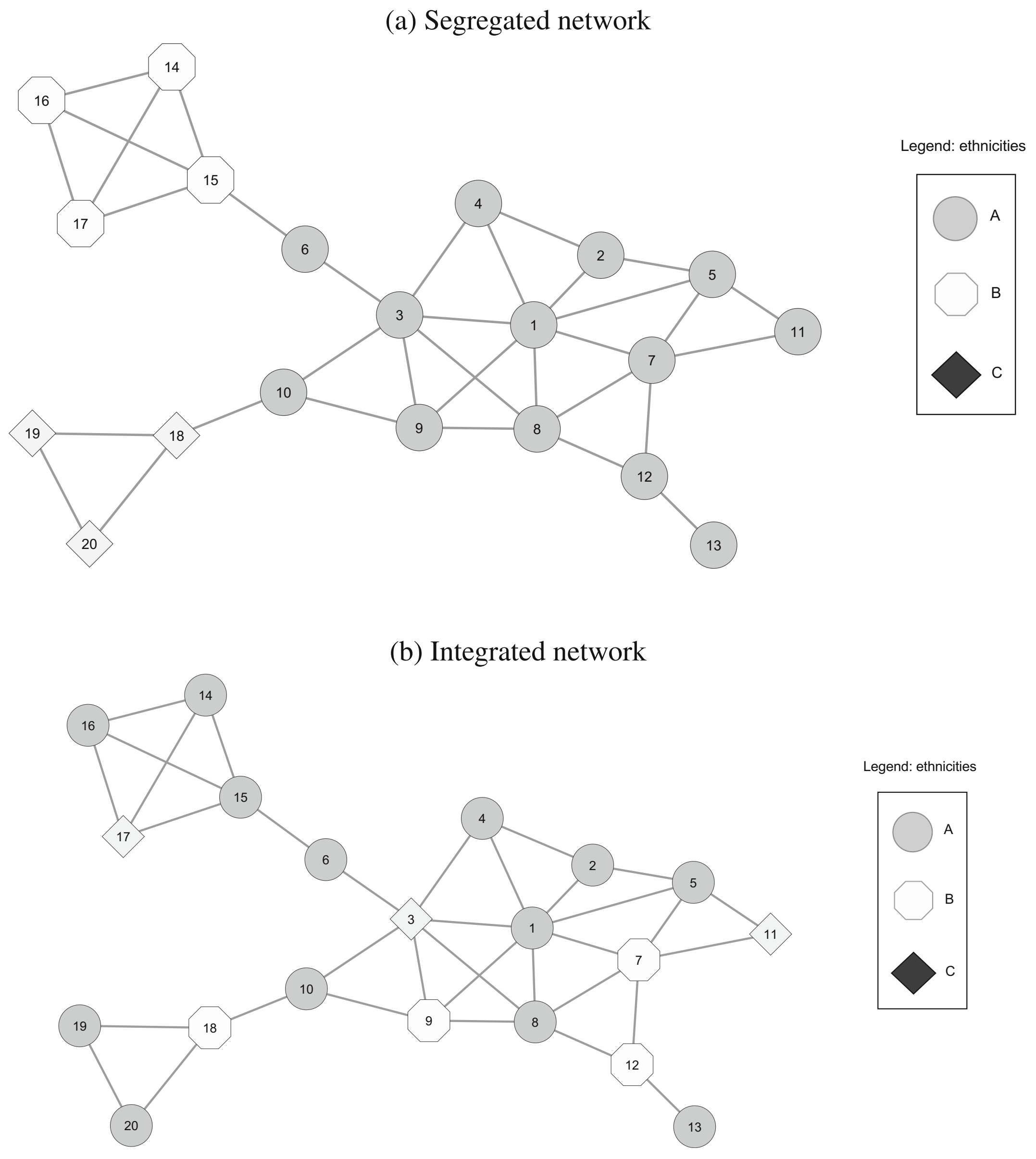


**FIGURE 1** Ethnic inventors and network integration. (a) Segregated network. (b) Integrated network. The two panels illustrate hypothetical co-inventor networks that share the same number of inventors and edges. Both networks also contain the same ethnic groups, with an equal number of inventors in each group. Panel (a) illustrates a segregated network, in which most ties occur within the same ethnic group. Panel (b) illustrates an integrated network, where ties span across ethnic groups.

To illustrate this argument, Figure 1 depicts two different configurations of the relational (collaboration) network underpinning the innovative activities of a hypothetical organization. Each node represents an individual member, with colors and shapes indicating their ethnic and cultural backgrounds. For this simplified example, the organization comprises 20 members divided into three ethnic groups: Group A (majority) with 13 members, and ethnic minorities B and C with 4 and 3 members, respectively. Although both panels reflect the same level of ethnic diversity, panel (a) and panel (b) portray distinct network configurations. The key difference lies in the positioning of members across ethnic groups. Panel (a) shows a segregated network, where members predominantly interact with others from the same ethnic group. In contrast, panel (b) illustrates an integrated network, where members from different ethnic groups freely "mix and match" regardless of cultural identity.

## 2.3 | Ethnic diversity, integration, and the value of innovation

Effective mobilization of diverse knowledge through collaboration is essential to realize the innovation value of ethnic heterogeneity. Although ethnic diversity introduces a broader spectrum of perspectives, experiences, and technical specializations, such variety yields superior innovation outcomes only when individuals actively collaborate across ethnic boundaries. We conceptualize ethnic integration as a relational structure in which co-invention ties span ethnic groups, enabling the mixing and recombination of heterogeneous knowledge to produce novel, high-impact inventions.

A fundamental premise underlying our argument is that ethnic background is systematically associated with distinct technological specializations. This association does not arise mechanically from ethnicity per se, but from the social and institutional processes that shape human capital accumulation and career trajectories across countries. First, educational systems differ across countries in their curricular priorities, methodological approaches, and access to particular scientific domains, leading migrants to enter the labor market with domain-specific competencies (Choudhury & Kim, 2019; Kerr, 2008). Second, migrant networks and diaspora communities often channel new entrants into specific industries and technological niches, reinforcing patterns of specialization over time (Saxenian, 2002). Third, homophily in professional networks, documented extensively in the innovation literature (Almeida et al., 2015; Freeman & Huang, 2015), can amplify initial differences by facilitating mentorship, sponsorship, and collaboration within ethnic communities. Through these mechanisms, ethnicity becomes correlated with clusters of technological expertise and frontier skills.

Importantly, however, ethnicity does not merely sort inventors into certain technological domains. While technological classifications capture *what* knowledge area an inventor operates in, ethnicity can also shape *how* that knowledge is accumulated, interpreted, and deployed. Cultural background influences cognitive frames, problem-solving heuristics, communication styles, and norms of interaction (Hong & Page, 2004; Lazear, 1999b). Even within the same formal technological class, inventors from different ethnic backgrounds may approach problems differently because of differences in training traditions, epistemic norms, and prior exposure to complementary domains. As a result, ethnic boundaries can correspond to variation in tacit knowledge, interpretative schemes, and social capital that are not fully observable through patent classifications alone.

These distinctions matter because bridging technological domains is not equivalent to bridging ethnic communities. Integration across technological sub-units can recombine codified knowledge domains, but integration across ethnic boundaries also requires overcoming social and relational frictions rooted in homophily, language, and identity. A large body of research shows that scientists and inventors disproportionately collaborate with same-ethnicity peers (Freeman & Huang, 2015), reflecting trust, shared background, and ease of communication. These tendencies can generate parallel, partially segregated collaboration clusters within the same firm, even when those clusters work in related technological areas. In this sense, ethnic boundaries may structure the informal architecture through which knowledge flows inside organizations, shaping who interacts with whom beyond formal task assignments.

Ethnic integration is therefore a not trivial source of heterogeneity in knowledge integration. In the absence of integrative mechanisms, organizations may display technological breadth yet remain socially segmented, with different ethnic clusters specializing in different knowledge domains but rarely collaborating. In such settings, the firm's knowledge base remains partitioned, and the potential gains from recombination are left unrealized. Integration across ethnic lines mitigates this fragmentation by fostering bridging ties that connect socially distinct communities, facilitating the transfer of tacit knowledge and enabling deeper recombination than would occur through formal cross-domain coordination alone (Reagans & Zuckerman, 2001; Tortoriello et al., 2012).

This relational perspective clarifies why ethnicity constitutes a theoretically meaningful lens rather than a redundant proxy for technological specialization. If ethnic groups are systematically embedded in distinct technological niches and social networks, then collaboration that spans ethnic boundaries simultaneously bridges (i) differentiated domain expertise and (ii) socially segmented knowledge communities. This dual bridging may be particularly consequential in complex innovation settings, where recombining frontier knowledge from multiple domains requires not only technical complementarity but also trust, mutual understanding, and iterative knowledge exchange.

By reducing homophily—where inventors cluster primarily with similar others—ethnic integration fosters cross-cultural collaboration and ensures the free flow of ideas across otherwise segmented knowledge communities. In integrated networks, diverse perspectives are not merely co-located but actively combined and refined through repeated interaction, increasing the likelihood that firms will generate innovations that are both novel and economically valuable.

Based on these arguments, we propose:

> **Hypothesis 1.** Organizations exhibiting an ethnically integrated network generate innovations of higher economic value for any given level of ethnic diversity.

Furthermore, the payoff from integration depends on the degree of workforce diversity. When diversity is low, cognitive distances among inventors are small and the marginal benefit of bridging ties is limited. As diversity rises—encompassing multiple ethnic, disciplinary, and cultural domains—the opportunity for novel recombinations grows substantially, heightening both potential rewards and co-ordination challenges. In such contexts, integration is critical for ensuring that informational richness is leveraged rather than hindered by fragmentation.

This leads to our second hypothesis:

> **Hypothesis 2.** The degree of ethnic integration positively moderates the effect of ethnic diversity on the economic value of innovations produced.

## 3 | DATA AND EMPIRICAL METHODS

### 3.1 | Sample of firms

To validate our hypotheses, we examined a sample of relatively large, innovative U.S. firms, for which the inflow of skilled immigrants has been particularly relevant in the last three decades.

Specifically, using Compustat-Capital IQ, we gathered the population of U.S. public firms listed in the period 1990–2015, omitting firms headquartered outside of the United States. It is worth noting that the period of time firms remain public varies across companies. Whereas some firms remained consistently public throughout the whole period (e.g., General Motors Co.), others (e.g., Dell Technologies Inc. and Sun Microsystems Inc.) were listed only for a subset of years during this period. For each of these firms, we collected data on their patenting activity at the United States Patent and Trademark Office (USPTO) during the years they were publicly listed in the period 1990–2015.

Recognizing that large firms sometimes file patents under varied names and that corporate subsidiaries may also engage in patenting, we exploited the DISCERN database (Arora et al., 2020). This database provides a careful mapping between USPTO patent applications and U.S. public firms from 1980 to 2015, by standardizing patent assignee names, matching subsidiaries and ultimate owners, and tracking acquisitions and corporate name changes. This yielded a pool of 4191 firms with at least one patent in this period.

To avoid introducing confounding factors, we excluded patents filed by inventors located outside the United States, thereby focusing exclusively on the *domestic* innovation activities of the selected firms. This further reduced the sample to 4077 firms.

Given the substantial heterogeneity in size and innovative intensity across firms and industries, we filtered the population of U.S. public companies to include only firms with a consistent innovation activity. Specifically, we retained only those firms with at least one patent in every year during the time in which they were listed in the period 1990–2015. This criterion ensures that we capture firms with sustained inventive activity, rather than those with sporadic or isolated patent filings. This reduced the sample to 1894 firms.

Finally, we removed companies with less than 7 years of patenting activity. This filtering criterion ensures that we have a long enough period to compute our focal independent variables, given our use of a 5-year rolling window (see below, Section 3.4.1).

After dropping singletons and firms with missing values on our covariates, our final sample consists of an unbalanced panel containing 11,608 firm-year observations for 890 firms over the period 1990–2015. Overall, the firms in our sample are responsible for 921,131 patents made by 485,987 unique inventors. This constitutes around 85.2% of all patents made by U.S. public firms, and about 40% of the overall U.S. domestic patents in the same period of time.[2]

## 3.2 | Ethnicity data

While patent data provides detailed information regarding individual inventors, it does not report either their nationality or their country of birth. Adhering to a widely-recognized methodology in the *migration and innovation* literature pioneered by Kerr (2008), we adopted a second-best approach, which consists of leveraging information on the first and last names of inventors to infer their most likely ethnic and cultural background. Specifically, we employed `Ethnea`—an ethnicity name classifier developed at the University of Illinois by Vetle Torvik and colleagues (Torvik & Agarwal, 2016; Torvik & Smalheiser, 2009). When provided with a person's name, `Ethnea` pinpoints all of the instances of authors with the same (or similar) name in `PubMed`, the leading biomedical literature database with more than 15 million

[2]Due to space constraints, we refer the interested reader to Supporting Information Appendix S1, where we report a detailed comparison between the public firms included and excluded from the sample.

abstracts, and it couples them with the respective affiliation country. Subsequently, it probabilistically maps the person's name to a set of 26 predefined ethnicities.[3] We assigned each inventor to the dominant ethnicity as predicted by the name classifier.[4] In what follows, we will consider inventors with *English* ethnicity as the majority ethnic group. In Supporting Information Appendix S1, we describe in detail our methodology for assigning ethnicity and include a comparison of ethnicity versus nationality for a subset of inventors.

## 3.3 | Dependent variable

We measured firm $i$'s innovation performance in year $t$ as the total real economic value of patents with application dates in that year:

$$Y_{i,t} = \sum_{p \in P_{i,t}} V_p, \tag{1}$$

where $P_{i,t}$ is the set of all granted patents whose application date falls in year $t$ and that belong to firm $i$. Here, $V_p$ is the value of innovation for patent $p$, deflated to 1982 dollars. To construct each patent's private value $V_p$, we draw on the data compiled by Kogan et al. (2017).[5] Specifically, for each patent $p$, they isolated the abnormal stock return around the grant date, scaled this return by the patent's ex ante probability of grant to filter out noise, and then multiplied the filtered return by the firm's market capitalization just prior to grant to obtain $V_p$. Summing these patent-level values yields $Y_{i,t}$ in real dollar terms.

A key advantage of this market-based, real-dollar valuation over citation-count metrics is that it reflects the firm-specific economic impact of each invention, independent of competitors' subsequent innovation activity. In our empirical models, we therefore include $\log(Y_{i,t})$ as our dependent variable.

## 3.4 | Independent variables

### 3.4.1 | Ethnic integration

The primary independent variable we considered is the degree of ethnic integration present within the research collaboration networks of sampled firms. To this purpose, we constructed a rolling-window network for each firm $i$ in year $t$ by drawing an undirected one-mode projection of the inventor–patent bipartite graph: two inventors are connected if they co-author at least one patent filed by firm $i$ during the interval $[t-5, t)$.

[3]The 26 ethnicities are: *English*, *German*, *Hispanic*, *Slav*, *Thai*, *Italian*, *French*, *Turkish*, *Indian*, *Chinese*, *Arab*, *Japanese*, *Nordic*, *Israeli*, *Dutch*, *Korean*, *Romanian*, *Hungarian*, *Baltic*, *African*, *Greek*, *Vietnamese*, *Indonesian*, *Caribbean*, *Mongolian*, *Polynesian*.

[4]It is worth noting that `Ethnea` also predicts dual-ethnicities (e.g., *English-Chinese*). Since including dual-ethnicities would have considerably complicated the analysis, we disregarded them and took a conservative approach. We classified as pure *English* any inventor for whom one of the dominant ethnicities is *English*.

[5]https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data (accessed July 21, 2021).

We quantify ethnic integration using the External–Internal (*EI*) index (Krackhardt & Stern, 1988):

$$EI = \frac{E - I}{E + I}, \tag{2}$$

where $E$ is the number of edges connecting inventors from *different* ethnic groups ("external" ties), and, $I$ i is the number of edges connecting inventors from the *same* ethnic group ("internal" ties). By construction, $-1 \leq EI \leq +1$. $EI = -1$ when all ties are internal (no cross-ethnic collaboration), while $EI = +1$ when all ties are external (no within-group collaboration). Lower values signal stronger segregation, and higher values indicate greater integration across ethnic groups.

Beyond its intuitive simplicity, the EI index is unaffected by isolated nodes—that is, inventors with no co-authors—unlike many alternative measures. However, it does not incorporate a null model for benchmarking observed integration against random expectations (Bojanowski & Corten, 2014). In other words, it provides no built-in "yardstick" to tell whether a given EI value reflects genuinely unusual patterns of collaboration or merely the level of mixing one would obtain by chance in a network of the same size and density. Consequently, we cannot determine whether a firm's observed integration is unusually high or low, only how it compares to the theoretical extreme values.

To build such a comparison, we proceeded as follows. For each firm–year co-invention network, we first extracted the inventor–patent bipartite graph over the preceding 5-year window. We then generated an ensemble of 1000 randomized networks by applying a degree-preserving "edge-swap" algorithm—using the `BiRewire` R package—to the bipartite matrix, thereby shuffling who works with whom while leaving each inventor's and each patent's degree unchanged. For each randomized replicate, we projected back onto the inventor layer and recomputed the EI index. This yielded a null distribution of EI values that reflects the level of integration expected by chance given the same firm's activity and network density. Finally, we compared the observed EI to this null distribution by computing the standardized $z$-score.

$$\mathrm{EI}^{s}_{i,t} = \frac{\mathrm{EI}_{i,t} - \mu_{i,t}}{\sigma_{i,t}}, \tag{3}$$

where $\mu_{i,t}$ and $\sigma_{i,t}$ are the mean and standard deviation of the 1000 randomized EI values. By construction, $\mathrm{EI}^{s}_{i,t} > 0$ indicates that the observed EI exceeds its null expectation (more cross-ethnic ties than random), whereas $\mathrm{EI}^{s}_{i,t} < 0$ denotes greater segregation than expected by chance. Hence, our focal independent variable, *Ethnic integration*, is represented by $\mathrm{EI}^{s}_{i,t}$.

### 3.4.2 | Ethnic diversity

We measure each firm's ethnic diversity over the 5-year window $[t-5, t)$ using the fractionalization index, a widely used metric in the literature (Alesina et al., 2016):

$$\text{Ethnic diversity}_{i,t} = 1 - \sum_{e=1}^{E} (s_{e,i,t})^2, \tag{4}$$

where $s_{e,i,t}$ is the share of inventors in firm $i$ belonging to ethnic group $e$ during $[t-5, t)$. This index captures the probability that two randomly selected inventors have different ethnic backgrounds: it equals 0 when all inventors share the same ethnicity and increases toward 1 as diversity rises.

### 3.4.3 | Control variables

Our empirical specification also includes a set of firm-year control variables. While our analysis is restricted to inventors and patents within the U.S., the firms in our sample frequently operate across multiple geographical regions, resulting in spatially dispersed inventor teams. Conceptually, geographic dispersion and ethnic integration are distinct dimensions of diversity. Geographic dispersion primarily captures spatial distances among inventors, affecting collaboration costs, co-ordination challenges, and localized knowledge spillovers. In contrast, ethnic integration specifically pertains to collaboration patterns that bridge culturally and technologically distinct groups, influencing the recombination of heterogeneous domain-specific knowledge.

The effect of geographic dispersion on innovation performance is a priori uncertain. Dispersed teams might face greater communication barriers and coordination difficulties, potentially hindering effective knowledge transfer and innovation (Stahl et al., 2010). Conversely, dispersed inventors might benefit from access to diverse localized knowledge pools and external spillovers, thus positively affecting innovation.

Empirically, controlling for geographic dispersion is crucial to isolate the influence of ethnic integration from potential colocation effects. To accomplish this, we employ the spatial dispersion index proposed by O'Leary and Cummings (2007), defined as follows:

$$\text{Spatial dispersion index}_{i,t} = \frac{\sum_{j-k}^{M} \text{Km}_{(j-k)} \cdot n_{i,j,t} \cdot n_{i,k,t}}{\left(N_{i,t}^2 - N_{i,t}\right)/2}, \tag{5}$$

where $\text{Km}_{(j-k)}$ denotes the distance in kilometers between states $j$ and $k$, where firm $i$ has operations, $M$ is the total number of states in which the firm is active, $n_{i,j,t}$ is the count of inventors of firm $i$ patenting in state $j$ during $[t-5, t)$, and $N_{i,t}$ is the total number of inventors from firm $i$'s patenting across all states during the same period. Including this measure helps ensure our estimates accurately reflect the effects of ethnic integration on innovation performance, rather than capturing spatial colocation effects.

In addition, all regressions control for the (logarithm of the) network size—that is, the number of inventors included in firm $i$'s network in the time window $[t-5, t-1]$, that is $log(N_{i,t})$. This variable captures the extent of human resources devoted by firms to innovation and controlling for it is essential for two reasons, one conceptual and one methodological. From a theoretical perspective, controlling for the size of the co-inventor network is necessary because network size captures critical structural characteristics such as the volume and scope of resources available for innovation within the firm.[6]

[6]Although R&D capital stock traditionally represents a key input to innovation, it overlaps conceptually and empirically with network size. Theoretically, R&D capital stock proxies the financial resources firms dedicate to innovation, while network size captures the human and relational resources available for recombination and innovation generation. Empirically, including both variables simultaneously is problematic due to their very high correlation ($r \approx 0.87$), leading to severe multicollinearity, inflated standard errors, and difficulty in separately estimating their effects. In unreported regressions, we tested robustness by including R&D capital stock computed with depreciation rates of both 15% and 30%; our main findings remain unchanged.

Moreover, larger inventor networks foster more frequent and diverse interpersonal interactions among inventors, facilitating novel recombination of knowledge, and thereby enhancing innovation outcomes (Carnabuci & Operti, 2013; Singh & Fleming, 2010). Methodologically, explicit inclusion of network size helps isolate genuine relational integration effects measured by the external-internal (EI) index, independent from scale effects arising simply from network size (Borgatti et al., 2006).

Finally, our specification includes a network structural control to capture the general level of connectivity within firms' co-invention networks. We computed the fraction of inventors belonging to the largest connected component of each firm's network:

$$\text{Largest component}_{i,t} = \frac{LC_{i,[t-5,t)}}{N_{i,[t-5,t)}}, \tag{6}$$

where $LC_{i,[t-5,t)}$ is the number of inventors in the largest connected component of firm $i$'s co-invention network during $[t-5,t)$, and $N_{i,[t-5,t)}$ is the total number of inventors in firm $i$ during the same timeframe. Including this network-level measure is theoretically justified because the connectivity and cohesion of a firm's inventor network significantly influence how efficiently knowledge diffuses within the organization and shapes inventive recombination processes (Carnabuci & Operti, 2013). Firms whose internal networks consist of small, disconnected inventor clusters are likely to experience diminished communication and limited knowledge recombination, which can constrain innovation performance. Conversely, more interconnected networks facilitate rapid knowledge diffusion, richer cross-fertilization of ideas, and higher rates of recombinant innovation.

While multiple network metrics, such as density, average path length, or clustering coefficients, could potentially reflect aspects of organizational connectivity, these were deliberately excluded due to either redundancy with the chosen metric or susceptibility to network-size biases. For instance, network density disproportionately declines as networks grow larger, providing limited additional insight beyond overall firm-size proxies. Similarly, average path length measures are less meaningful when networks are fragmented into many small components, offering less theoretical consistency and interpretability. Hence, the chosen measure strikes a balance, providing a clear, interpretable, and robust measure of organizational connectivity without the confounding effects associated with other potential network controls.

## 3.5 | Econometric approach

To test our hypotheses, we estimate the following specification using OLS:

$$\begin{aligned}\log(Y_{i,t}) = {} & \alpha + \beta_1 \text{Ethnic integration}_{i,t} + \beta_2 \text{Ethnic diversity}_{i,t} \\ & + \beta_3 \log(\text{Network size}_{i,t}) + \beta_4 \text{ Spatial dispersion index}_{i,t} \\ & + \beta_5 \text{Largest component}_{i,t} + \mu_i + \lambda_t + \delta_{s(i),t} + \varepsilon_{i,t}.\end{aligned} \tag{7}$$

In this equation, $Y_{i,t}$ denotes firm $i$'s total real economic value of patents with application dates in year $t$ (see Section 3.3 and Equation 1). The explanatory variables include Ethnic integration$_{i,t}$, reflecting the extent of ethnic integration in firm $i$'s co-invention network during the preceding 5-year window $[t-5,t)$ (see Section 3.4.1, Equation 3); Ethnic diversity$_{i,t}$, measuring the firm's ethnic diversity over the same period (see Section 3.4.2, Equation 4); and $\log(\text{Network size}_{i,t})$, the logarithm of inventors within the firm's co-invention network (see Section 3.4.3). Additionally, we include the $\text{Spatial dispersion index}_{i,t}$, quantifying inventors' geographical dispersion (see Section 3.4.3, Equation 5), and $\text{Largest component}_{i,t}$, the fraction of inventors belonging to the largest connected component (see Section 3.4.3, Equation 6). The specification incorporates firm-fixed effects $(\mu_i)$ to control for time-invariant firm characteristics, year-fixed effects $(\lambda_t)$ capturing common shocks across firms, and SIC 2-digit industry-specific linear time trends $(\delta_{s(i),t})$ to account for industry-level evolving factors.

Our specification estimates the effect of workforce integration on the economic value of patents within firms, controlling for both firm and year fixed effects, and for the endogenous evolution of industries over time. While this approach mitigates endogeneity concerns arising from omitted variables, concerns related to reverse causality may still persist. In particular, organizations that generate high-value patents may, over time, become more adept at managing an ethnically integrated workforce—potentially as a result of improved management practices. This may occur if firms endogenously design team composition by balancing the benefits of diversity against potential co-ordination and communication costs, depending on the expected value of each innovation project. As a consequence, teams responsible for high-value patents may be systematically different in composition from those associated with lower-value inventions.

To address this possibility, we develop an instrument that exploits city-wide linguistic distance among inventors residing in the same CSBA as a source of exogenous variation in the likelihood of observing integration within firms located in that city (Larkey, 1996; Ottaviano & Peri, 2005).

Operationally, the instrument is constructed by looking at the pairwise linguistic distance between language $e$ and language $k$, where $\text{Similarity}(L_e, L_k)$ is taken from Bella et al. (2021).[7] Our instrument for each firm $i$ and year $t$ is defined as

$$IV_{i,t} = \sum_{e=1}^{E} \sum_{k=1}^{E} s_e s_k d(L_e, L_k), \tag{8}$$

where $s_e$ and $s_k$ are, respectively, the shares of inventors belonging to ethnic groups $e$ and $k$, residing in the city (CBSA) where firm $i$ has its headquarters, patenting in the period $[t-6, t-2]$, but not affiliated with firm $i$, and

$$d(L_e, L_k) = 100 - \text{Similarity}(L_e, L_k). \tag{9}$$

The intuition behind the proposed instrumental variable builds on the idea that in cities characterized by high levels of linguistic distance, individuals from diverse ethnic groups face greater communication frictions and are more likely to cluster into linguistically homogenous groups (Lazear, 1999a). These linguistic barriers correlate with higher levels of urban segregation, which in turn may foster a local cultural environment less conducive to integration (Hong et al., 2023). Such an environment can shape social norms and behaviors that spill over into organizational contexts, influencing the likelihood of integration among employees of different ethnic groups in firms located in a given city. Firms in segregated urban contexts are embedded in a fragmented social fabric, which increases the likelihood of reproducing—both

[7]Data available at http://ukc.disi.unitn.it/index.php/lexsim/.

intentionally, through the assignment of individuals to teams and projects, and unintentionally, as individuals organically join teams—similar patterns of social and collaborative segregation within the organization. Importantly for the validity of our instrument, while it is plausible that linguistic distance within cities affects workforce composition and patterns of collaboration within organizations located in these areas (Lazear, 1999a), it is also likely to reflect historical settlement patterns that are uncorrelated with the current behavior and decisions of any individual firm. Finally, although plausibly exogenous, the instrument is constructed at the city level and applied to firm-level observations, meaning it varies across cities but not within them. Consequently, the effective variation is limited to the number of cities, which may reduce the precision of our instrumental variable estimates and inflate standard errors in the second stage.

## 3.6 | Descriptive statistics

Before proceeding to illustrate our results, Table 1 reports descriptive statistics and correlations for the variables included in the empirical model. Further and more detailed descriptives are reported in Supporting Information Appendix S1. Moreover, Figures 2 and 3 illustrate the evolution over time for the sample of selected companies of, respectively, the share of inventors by ethnic groups and the average values of ethnic diversity and ethnic segregation measures.

From observing Figure 2, we note the rapid growth in the share of all (non-English) ethnic inventors. Chinese and Indian inventors increased their share from around 3% in 1990 to slightly less than 10% in 2015, whereas the share of all other ethnic groups increased from about 11% to more than 16% (left $y$-axis). Correspondingly, the share of English-ethnic inventors declined from around 83% to slightly more than 63% (right $y$-axis). Clearly, the changed ethnic composition of the inventors' population is reflected in the average ethnic diversity across the firms in the sample, which went up from around 0.59 to around 1.05 (see Figure 3).

With respect to ethnic integration ($EI^s_{i,t}$, Equation 2), we observe a clear downward trend over the sample period (Figure 3). The index declines from approximately −0.5—indicating neither pronounced integration nor segregation—to around −3, reflecting substantial segregation. As the inventor pool became more ethnically diverse, firms in our sample faced increasing difficulty integrating these inventors into their technological collaboration networks.

**TABLE 1** Descriptive statistics and correlations.

| | Mean | Std. dev. | Min. | Max. | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|---|---|---|
| (1) Monetary value (log) | 4.16 | 2.54 | 0.00 | 11.73 | | | | | |
| (2) Ethnic integration | −1.52 | 3.72 | −55.46 | 19.30 | −0.36 | | | | |
| (3) Ethnic diversity | 0.35 | 0.18 | 0.00 | 0.84 | 0.07 | −0.30 | | | |
| (4) Spatial dispersion index | 0.89 | 0.57 | 0.00 | 4.16 | 0.15 | −0.13 | 0.03 | | |
| (5) Network size (log) | 4.49 | 1.47 | 0.00 | 9.57 | 0.76 | −0.50 | 0.06 | 0.26 | |
| (6) Share of largest component | 0.48 | 0.27 | 0.02 | 1.00 | −0.05 | −0.11 | 0.28 | −0.31 | −0.08 |

*Note*: This table reports descriptive statistics (mean, standard deviation, minimum, and maximum) and presents the correlation matrix for the variables used in the regressions. The dependent variable is the logarithm of the total real economic value of patents, measured in 1982-dollar millions. Please refer to Supporting Information Appendix S1 for additional descriptive analysis.

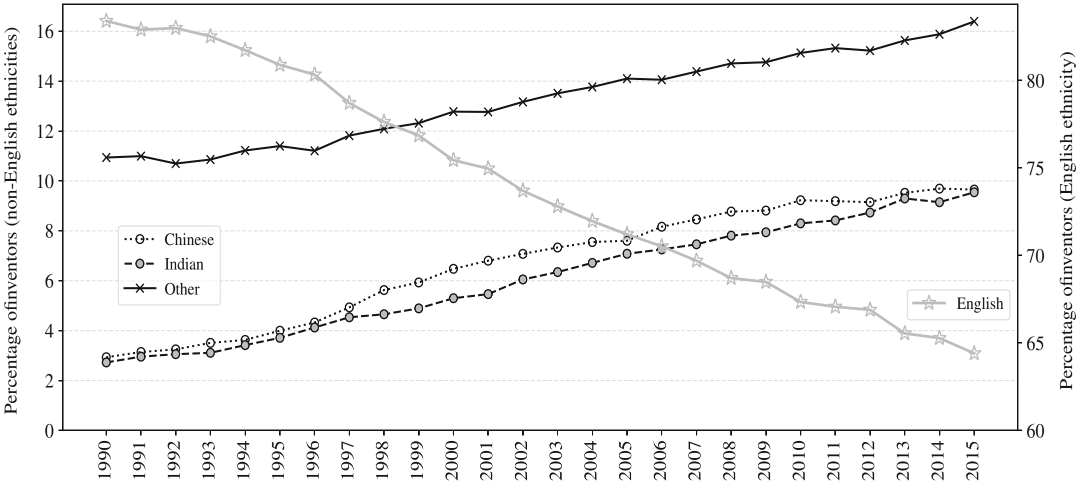


**FIGURE 2** Share of inventors by ethnicity and year. The figure illustrates the share of inventors of English ethnicity (right axis) and the share of Chinese, Indian, and other ethnic groups (left axis) over the period covered in this study, based on our sample of 890 U.S. public firms.

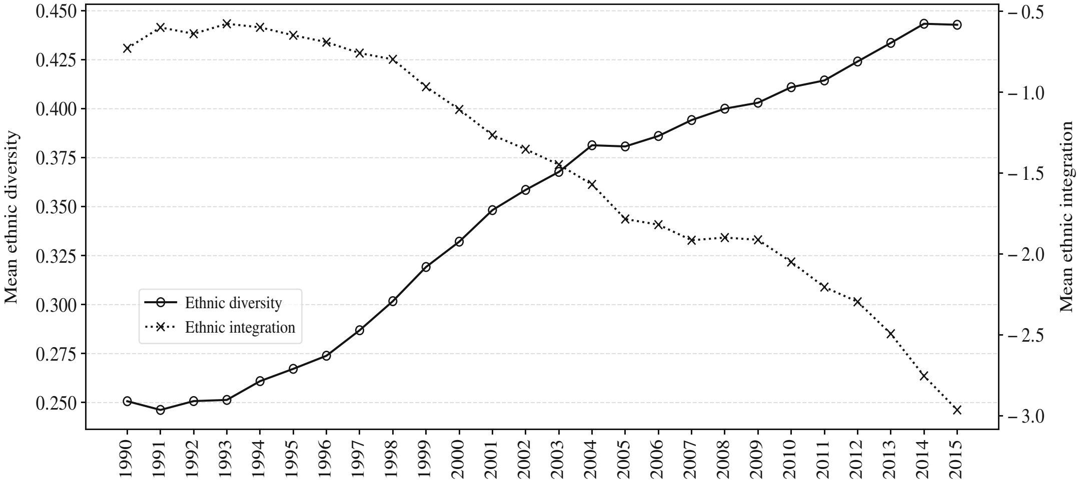


**FIGURE 3** Ethnic diversity and integration by year. The figure shows the evolution of average ethnic diversity (see Equation 4, left axis) and average ethnic integration (see Equation 3, right axis).

# 4 | RESULTS

## 4.1 | Ethnic integration, diversity, and the economic value of innovation within organizations

In Table 2, we begin by examining the correlation between ethnic integration, ethnic diversity, and the economic value of innovation, controlling for the size of the network, firm and year fixed effects, and sector specific linear trends. Both ethnic integration and diversity are

**TABLE 2** OLS estimates of ethnic integration on patent monetary value.

| | (1) | (2) | (3) |
|---|---|---|---|
| Ethnic integration | 0.014 | 0.014 | 0.016 |
| | (.001) | (.001) | (.000) |
| Ethnic diversity | 0.129 | 0.107 | 0.103 |
| | (.455) | (.535) | (.550) |
| Network size (log) | 0.699 | 0.711 | 0.705 |
| | (.000) | (.000) | (.000) |
| Spatial dispersion index | | −0.101 | −0.066 |
| | | (.022) | (.150) |
| Largest component (%) | | | 0.231 |
| | | | (.005) |
| Constant | 1.003 | 1.047 | 0.937 |
| | (.000) | (.000) | (.000) |
| Firm FE | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| SIC 2-digit × Time | Yes | Yes | Yes |
| Observations | 11,885 | 11,885 | 11,885 |
| Number of firms | 890 | 890 | 890 |
| $R^2$ | 0.833 | 0.833 | 0.833 |
| $R^2$-Within | 0.089 | 0.090 | 0.090 |

*Note*: The columns report high-dimensional fixed-effects OLS estimates of the effect of ethnic integration on the natural log of patent monetary value. Each model absorbs firm and year fixed effects, as well as interactions between 2-digit SIC industry codes and time, to net out any unobserved, industry-specific linear trends. Standard errors are robust to heteroskedasticity. *p*-values are shown in parentheses.

positively correlated with patent value, although only integration is statistically significant ($p$-value = .001). Column 2 adds a measure of spatial dispersion to capture the degree of geographical colocation among inventors, while Column 3 further controls for network connectivity by including the share of inventors in the largest connected component of the firm's co-inventor network.

In our most comprehensive specification, ethnic integration remains positively and significantly associated with patent value ($p$-value = .000). Notably, ethnic diversity is positively associated with innovation value but is statistically insignificant across specifications.[8]

In Table 3, we test the robustness of our findings against endogeneity concerns and adopt the instrument presented earlier in the paper. Column 1 reports the first stage regression where the instrument based on linguistic distance turns—as expected—to be negatively and significantly correlated with our measure of ethnic integration within organizations. Firms located in cities characterized by a high level of linguistic distance are more likely to show a lower level of ethnic integration within their network of inventors. The second stage results reported in Column 2 confirm the OLS estimates in Column 3 of Table 2. Ethnic integration remains positively

[8]Results remain consistent when using an alternative measure of ethnic diversity based on the Shannon Index; coefficient estimates are similar in magnitude and significance and are available upon request.

and statistically significantly correlated with the economic value of patents ($p$-value = .000). As expected, the coefficient increases in size, but so do standard errors: the (unreported) standard error rises from 0.004 in Column 4 of Table 2 to 0.073 in Column 2 of Table 3.

This evidence reflects both features of the instrument and the difference between local treatment effects and the broader average effects captured in the baseline specification. First, because the IV is constructed at the city level, it shifts treatment through a common, aggregated source of variation, which can attenuate the within-firm signal available to OLS and make the 2SLS estimate more imprecise. Second, 2SLS identifies a local average treatment effect (LATE) rather than the population average treatment effect (ATE), since it recovers the effect only for compliers—firms whose treatment intensity responds to the city-level instrument. These compliers are not necessarily "average" firms; they are those for which cross-ethnic collaboration is not fixed by technology or corporate policy, but can expand or contract when local conditions make communication across ethnic groups easier or harder. A natural set of marginal firms driving the IV results are those engaged in high value, complex innovations, where inventive output depends on recombining complementary expertise across people rather than isolated individual work. In these settings, cross-ethnic collaboration is both more valuable and more sensitive to coordination frictions: when local linguistic barriers make coordination harder, firms optimally segment teams; when those frictions are reduced, the same firms expand cross-ethnic teaming at the margin. Because the instrument shifts integration precisely through the feasibility and cost of coordination, the induced changes are concentrated in firms where collaboration is central to innovation and innovation value is likely higher—so the resulting LATE can plausibly exceed the average association captured by OLS. Conditional on this caveat, the results in Table 3 support Hypothesis 1: higher integration within organizations is associated with more valuable innovations.

A remaining concern is that in thin local labor markets the city-level instrument may partly reflect the inventive activity of an anchor employer, even after excluding the focal firm's own inventors from the construction of the instrument. To address this, we implement a conservative robustness check that removes firm-years headquartered in CBSAs where local patenting is highly concentrated in a single assignee. Specifically, using the universe of USPTO assignees (public and private), we compute for each CBSA-year the share of patents accounted for by the largest assignee in rolling 5-year windows, following the same moving-window procedure used to construct our main patent-based measures, and flag CBSAs in which this share exceeds 40% in any year.[9] We then re-estimate the 2SLS specification excluding all firm-years located in these "dominant-assignee" CBSAs.

Columns 3 and 4 of Table 3 report the corresponding first- and second-stage estimates. The instrument continues to exhibit strong relevance, with a large first-stage F-statistic, and the second-stage coefficient on ethnic integration remains positive and statistically significant. These results indicate that our findings are not driven by thin local markets dominated by a single firm and are consistent with the interpretation that the instrument captures broader local conditions affecting cross-ethnic collaboration rather than firm-specific shocks.[10]

[9]Results are qualitatively unchanged if we adopt a more stringent definition of dominant employers and classify CBSAs as concentrated when the largest assignee accounts for more than 25% of local patenting in the rolling 5-year window. The first- and second-stage estimates remain similar in magnitude and statistical significance.

[10]To address the related concern that the instrument may be influenced by very old, foundational firms that were central to the historical development of certain CBSAs, we also re-estimated our IV specification after progressively excluding firms founded before 1940, 1960, and 1980. Founding-year information was obtained from Jay Ritter's data (https://site.warrington.ufl.edu/ritter/files/IPOs-Age-of-Companies-Going-Public.pdf) and supplemented with manual verification where necessary. Across all specifications, the first-stage relationship remains strong and the second-stage coefficient on ethnic integration remains positive and statistically significant. These results confirm that our findings are not driven by a small set of long-established firms. Results are available upon request.

**TABLE 3** Instrumental-variables estimates of ethnic integration on patent monetary value.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | | | 1st stage | 2nd stage |
| | 1st stage | 2nd stage | (Excluding dominant-assignee CBSAs) | |
| Ethnic integration | | 0.286 (.000) | | 0.175 (.001) |
| Ethnic diversity | −2.377 (.000) | 0.841 (.002) | −2.095 (.000) | 0.463 (.040) |
| Network size (log) | −1.090 (.000) | 1.015 (.000) | −1.050 (.000) | 0.886 (.000) |
| Spatial dispersion index | 0.173 (.027) | −0.098 (.053) | 0.182 (.025) | −0.090 (.064) |
| Largest component (%) | −1.500 (.000) | 0.651 (.000) | −1.417 (.000) | 0.552 (.000) |
| Language distance | −0.140 (.000) | | −0.202 (.000) | |
| Constant | 7.697 (.000) | | 8.761 (.000) | |
| Firm FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| SIC 2-digit × Time FE | Yes | Yes | Yes | Yes |
| Observations | 11,871 | 11,871 | 10,716 | 10,716 |
| Number of firms | 888 | 888 | 808 | 808 |
| First-stage $F$ | 88.440 | | 81.952 | |

*Note*: The table reports two-stage least squares estimates of the effect of ethnic integration on the log of patent monetary value. Ethnic integration is instrumented by the average linguistic distance in the CBSA where the firm is headquartered. All regressions include firm and year fixed effects, as well as interactions between 2-digit SIC industry codes and time, to net out any unobserved, industry-specific linear trends. In the first stage, ethnic integration is regressed on the instrument and the same set of exogenous controls that appear in the second stage; the reported first-stage $F$ is the heteroskedasticity-robust $F$-statistic for the joint significance of all first-stage regressors. In the second stage, the fitted value of ethnic integration enters alongside those controls. Estimates reported in Columns 3 and 4 exclude firm-years located in CBSAs where a single assignee (not necessarily a public company in our sample) accounted for more than 40% of 5-year rolling patent counts in any year. Standard errors are robust to heteroskedasticity. $p$-values are shown in parentheses. Two firms are omitted relative to the OLS sample because their CBSAs could not be determined, preventing IV computation.

Next, we test Hypothesis 2 by splitting the sample according to whether a firm's diversity in a given year is below or above the sample median. This yields two subsamples—low- and high-diversity firms—and allows us to compare the relationship between ethnic integration and innovation value across the two groups. The results of this exercise are reported in Table 4. Columns 1–2 report results for low-diversity firms, while Columns 3–4 focus on high-diversity firms. In both subsamples, the first stage indicates that the instrument is strongly correlated with the endogenous variable in the expected direction (Columns 1 and 3). In the second stage, however, we find evidence of a positive effect of ethnic integration on innovation value only among high-diversity firms (Column 4; $p$-value = .028). For low-diversity firms, we find no statistically significant association between integration and innovation value.

Overall, results in Table 4 support Hypothesis 2, indicating that integration and diversity function as distinct yet mutually reinforcing constructs. The benefits of workforce diversity are capitalized only in the presence of effective integration, suggesting that ethnic integration is a critical enabler for converting the latent potential of diversity into innovation value.

## 4.2 | Plausible mechanisms: Ethnic-based competitive advantages and the complexity of innovation

A foundational premise of our study is that individuals' ethnic backgrounds are systematically associated with specialization in particular technological domains. These patterns arise not only from differences in access to and investment in education and training, but also from broader cultural influences, including values, aspirations, and prevailing career trajectories across ethnic groups.

If this is the case, we should observe evidence supporting two stylized facts. First, there should be systematic ethnic comparative advantages across distinct technological domains, reflecting patterns of specialization shaped by education, experience, and cultural background. Second, the benefits of integration should be especially pronounced in contexts of high innovation complexity—where the value of combining diverse sources of frontier knowledge spanning across different technological domains is greatest, and co-ordination across different expertise areas is most critical.

### 4.2.1 | Ethnic-based competitive advantages across technological domains

To provide evidence on our first prior, we collected all USPTO patents and split them into periods of 5 years (e.g., 1990–1994, 1995–1999, etc.). Then, we used the NBER technology classification that assigns each patent to six broad technological areas. For each patent, we fractionally counted inventors by ethnicity (e.g., a patent with two Indian inventors and one English inventor contributes 2/3 to Indian ethnicity and 1/3 to English ethnicity). Within each 5-year window $T$, we then computed

$$\mathrm{R}_{c,e}(T) = \frac{\mathrm{S}_{c,e}(T)}{\mathrm{S}_{e}(T)}, \tag{10}$$

where $\mathrm{S}_{c,e}(T)$ is the fraction of fractional inventor-units in technology area $c$ during period $T$ that belong to ethnicity $e$, and $\mathrm{S}_{e}(T)$ is the fraction of all fractional inventor-units in period $T$ that belong to ethnicity $e$. A ratio $\mathrm{R}_{c,e}(T)>1$ indicates that ethnic group $e$ is relatively specialized in category $c$ (i.e., appears more often in $c$ than would be expected from its overall share of inventorship), whereas $\mathrm{R}_{c,e}(T)<1$ signals under-representation.

Figure 4 shows the heatmap for 2000–2004, illustrating the values of $\mathrm{R}_{c,e}(T)$ across groups of homogeneous technologies. Indian inventors are 61% more represented in Computers & Communications than would be expected from their overall share of inventors. Similarly, Chinese inventors are 30% and 38% more represented in Electrical & Electronic and Chemical technologies, respectively, than would be expected from their overall share of inventors. The

**TABLE 4** Ethnic integration and patent monetary value in low and high diversity firms.

| | (1) Ethnic diversity < median | (2) | (3) Ethnic diversity ≥ median | (4) |
|---|---|---|---|---|
| | 1st stage | 2nd stage | 1st stage | 2nd stage |
| Ethnic integration | | 0.242 | | 0.345 |
| | | (.429) | | (.028) |
| Ethnic diversity | 1.099 | 0.364 | −5.552 | 1.134 |
| | (.011) | (.454) | (.000) | (.242) |
| Network size (log) | −0.455 | 0.833 | −1.641 | 1.202 |
| | (.000) | (.000) | (.000) | (.000) |
| Spatial dispersion index | 0.262 | −0.009 | 0.188 | −0.015 |
| | (.004) | (.934) | (.142) | (.850) |
| Fraction of inventors in largest component | −0.371 | 0.415 | −1.815 | 0.801 |
| | (.027) | (.021) | (.000) | (.010) |
| IV (Language CBSA) | −0.040 | | −0.143 | |
| | (.005) | | (.000) | |
| Constant | 1.807 | | 12.075 | |
| | (.000) | | (.000) | |
| Firm FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| SIC 2-digit × Time FE | Yes | Yes | Yes | Yes |
| Observations | 5772 | 5772 | 6008 | 6008 |
| First-stage $F$ | 18.289 | | 59.555 | |

*Note*: The table reports two-stage least squares estimates of the effect of ethnic integration on the log of patent monetary value. Ethnic integration is instrumented by the average linguistic distance in the CBSA where the firm is headquartered. All regressions include firm and year fixed effects, as well as interactions between 2-digit SIC industry codes and time, to net out any unobserved, industry-specific linear trends. Estimates in Columns 1 and 2 are for the subsample of firms with a diversity index below the median value in the sample in a given year, Columns 3 and 4 focus on firms with an above-median value of the diversity index. The estimation drops 91 singleton observations that do not contribute to identification given the absorbed fixed effects. Standard errors are robust to heteroskedasticity. $p$-values are shown in parentheses.

patterns depicted in the heatmap confirm that different ethnic groups tend to specialize in distinct technological domains, leading to variation in their capacity to develop frontier skills within those areas.

### 4.2.2 | Ethnic integration and the complexity of innovation

To test whether the role of integration changes with the complexity of innovation developed within each firm, we begin by building a global USPTO-subclass co-occurrence network using all patents filed in the 5-year window $[t-5, t)$. Each patent can be classified into multiple USPTO subclasses (e.g., subclass 252/299.62 "Liquid Crystal Compositions ..."). We take the set of all subclasses observed in that window and draw an undirected edge between any two subclasses $u$ and $v$ if at least one patent in the window is co-classified in both $u$ and $v$. The weight $w_{uv}$ on each edge is simply the total number of such co-classifications. In this way, frequently paired subclasses (which indicate a strong technological tie) receive larger weights, while rare or incidental pairings receive smaller weights. The result is a weighted graph whose nodes are all observed subclasses, whose edges represent co-classification links, and whose edge-weights count those links.

Once this global network is constructed, we apply the Leiden community-detection algorithm (Traag et al., 2019) to partition the graph into clusters. Each cluster can be interpreted as a coherent "global technology domain," since subclasses that often co-occur in the same patents end up grouped together. As output, Leiden assigns each subclass $u$ to a unique cluster $c(u)$. We treat these clusters as our reference domains when assessing firm-level complexity. (Supporting Information Appendix S1 provides full technical details and illustrative cluster maps).

Next, for each firm $i$ and year $t$, we extract the subgraph $G_{i,t}$ induced by exactly those subclasses in which firm $i$ has patented over the same 5-year window. In other words, $G_{i,t}$ contains all subclasses that appear in firm $i$'s patents filed between years $t-5$ and $t-1$, together with the co-occurrence edges (and weights) inherited from the global network. If firm $i$ focuses narrowly on a single domain, then $G_{i,t}$ will be small and densely connected within one cluster; if firm $i$ spans multiple distant domains, then $G_{i,t}$ will include subclasses from several clusters linked by cross-cluster edges.

To measure the breadth and balance of firm $i$'s portfolio, we count, for each global cluster $c$, the number $n_c$ of subclasses in $G_{i,t}$ that belong to cluster $c$. Let $N$ be the total number of subclasses in $G_{i,t}$ (so $N=\sum_c n_c$). We then compute the share of subclasses in each cluster as

$$p_c = \frac{n_c}{N}. \tag{11}$$

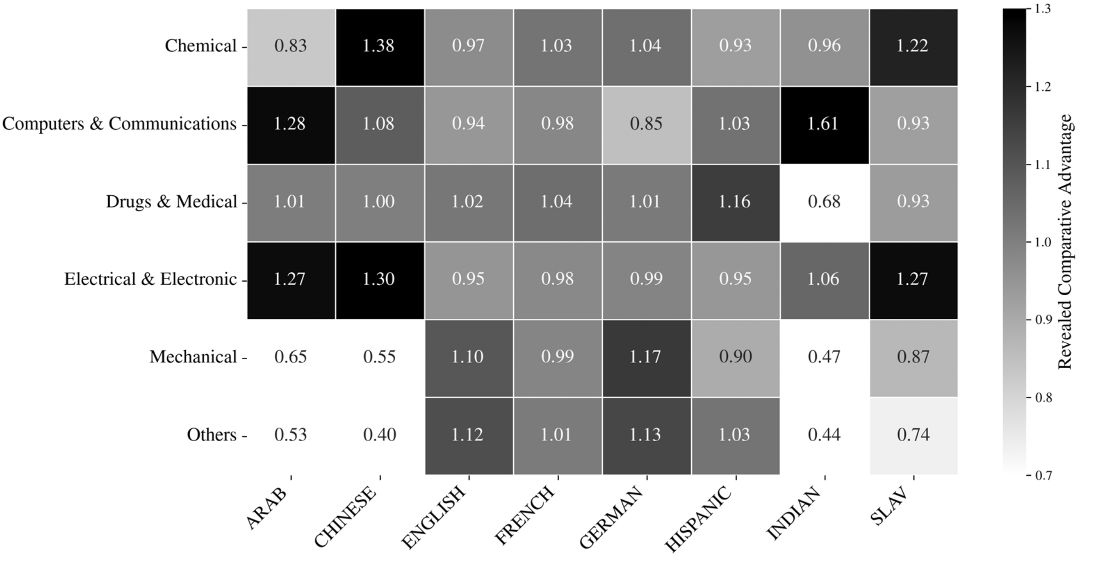


**FIGURE 4** Over- and under-representation of ethnic inventors across technological domains. This figure plots the specialization measure $R_{c,e}(T)$ (Equation 10) over 2000–2004 for those ethnic groups that account for at least 1% of U.S. inventors in that interval. Comparable charts for other 5-year windows appear in Supporting Information Appendix S1.

Finally, our complexity indicator is the Shannon entropy of these shares:

$$H_{i,t} = -\sum_{c} p_c \ln(p_c). \tag{12}$$

By construction, $H_{i,t}=0$ if all of firm $i$'s subclasses lie in a single cluster, and $H_{i,t}$ increases as the firm's subclasses spread more evenly across multiple technological domains. As a robustness check, we also computed an alternative measure that captures the share of a firm's co-occurrence edge-weight crossing different clusters. Results are qualitatively similar; details appear in Supporting Information Appendix S1.

We use this measure to rank organizations based on the complexity of the innovations they develop. Since complexity evolves over time, a given firm may move up or down in the ranking depending on the variety of technological domains it draws upon at each point in time. This approach, combined with the inclusion of firm fixed effects, allows us to capture within-firm variation in innovation complexity—rather than simply reflecting that some firms have inherently more complex knowledge bases than others.

To provide a correlational account of the role of complexity as a potential mechanism at play, we estimate our full specification—as in column (3) of Table 2—on two subsamples: firms that, at any point in time, develop innovations ranked below the median in terms of complexity (Columns 1 and 2 of Table 5), and firms whose innovations are above the median complexity level in the sample (Columns 3 and 4 of Table 5). The results indicate that the positive effect of ethnic integration on the economic value of patents is concentrated among firms engaged in more complex innovations (Column 3). In contrast, no significant effects emerge among firms with lower levels of innovation complexity (Column 1). We also find that the contribution of integration for converting the latent potential of diversity within organizations becomes key only in high complexity firms (Column 4). In this subsample the interaction term between complexity and integration turns positive and statistically significant ($p$-value: .009). No supportive evidence for Hypothesis 2 is detected for the subsample of low complexity firms (Column 2).[11]

This evidence supports the idea that the benefits of integrating diverse sources of knowledge within an organization are channeled through the complexity of the innovations it develops. Given that ethnic diversity is associated with specialization in distinct technological domains—and thus with access to frontier expertise—firms undertaking complex innovations, which require the combination of multiple technological competencies, benefit more from effective integration. In such contexts, the ability to align and synthesize these diverse capabilities within a coherent collaborative framework becomes essential for generating innovation value. Where the developed innovation is less complex and does not require the integration of distinct technological competences, the contribution arising from actual integration remains limited.

## 5 | DISCUSSION AND IMPLICATIONS

As innovation becomes increasingly complex and dependent on the integration of diverse knowledge domains, understanding how to structure and leverage collaboration has become a

[11]To avoid additionally splitting the sample by diversity in this exercise, we instead include an interaction term between diversity and integration. This allows us to test whether combining diversity with integration is particularly valuable for firms developing complex innovations, which would further reinforce our claims regarding the theorized mechanisms.

**TABLE 5** Ethnic integration and technological complexity: Split sample.

| | (1) Low complexity | (2) Low complexity | (3) High complexity | (4) High complexity |
|---|---|---|---|---|
| Ethnic integration | −0.006 | −0.032 | 0.019 | −0.017 |
| | (.571) | (.201) | (.000) | (.252) |
| Ethnic diversity | 0.433 | 0.471 | −0.261 | −0.136 |
| | (.068) | (.048) | (.363) | (.641) |
| Ethnic integration × Ethnic diversity | | 0.059 | | 0.072 |
| | | (.249) | | (.009) |
| Network size (log) | 0.514 | 0.520 | 0.938 | 0.935 |
| | (.000) | (.000) | (.000) | (.000) |
| Spatial dispersion index | −0.020 | −0.019 | −0.114 | −0.101 |
| | (.749) | (.769) | (.145) | (.201) |
| Largest component (%) | 0.080 | 0.077 | 0.166 | 0.176 |
| | (.507) | (.524) | (.184) | (.160) |
| Constant | 1.209 | 1.175 | 0.250 | 0.201 |
| | (.000) | (.000) | (.310) | (.415) |
| Firm FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| SIC 2-digit × Time | Yes | Yes | Yes | Yes |
| Observations | 5438 | 5438 | 6362 | 6362 |
| $R^2$ | 0.816 | 0.816 | 0.852 | 0.852 |
| $R^2$-Within | 0.048 | 0.048 | 0.116 | 0.117 |

*Note*: Columns 1 and 2 report high-dimensional OLS estimates of the log of firm-level patent monetary value on ethnic integration, ethnic diversity, and a common set of firm-level controls, using separate low- and high-complexity samples. Each model absorbs firm and year fixed effects, as well as interactions between 2-digit SIC industry codes and time, to net out any unobserved, industry-specific linear trends. The estimation drops 85 singleton observations that do not contribute to identification given the absorbed fixed effects. Standard errors are robust to heteroskedasticity. *p*-values are shown in parentheses.

critical concern for both researchers and practitioners. In this context—where firms face growing pressure to harness global talent and where the ethnic composition of large organizations has become markedly more diverse—the central question is no longer whether diversity matters, but how it can be effectively translated into innovation outcomes. This study contributes to a deeper understanding of how diverse organizations operate by examining the role of ethnic integration within collaborative networks as a key driver of economically valuable innovation.

Our findings show that firms with higher levels of ethnic integration—defined as the extent to which inventors collaborate across ethnic boundaries—produce patents with significantly greater economic value. This positive effect is amplified in firms that are more ethnically diverse, demonstrating that integration and diversity are mutually reinforcing. Furthermore, the benefits of integration are especially pronounced in the context of complex innovation, where success depends on recombining specialized knowledge from different domains. In such settings, the ability to bridge both cultural and technical divides becomes a prerequisite for innovation performance.

Our study adds a novel dimension to the literature on diversity and innovation by distinguishing between the composition of inventor teams and the structure of their collaborative networks. While prior work has largely focused on the presence of diversity, we show that network-level integration dynamics are essential to realizing the performance benefits of a heterogeneous workforce.

By documenting how the value of integration intensifies with the complexity of innovation, we also bridge the literature on knowledge recombination with that on organizational social structure. Our findings emphasize that the value of integration is contingent—it is not universally beneficial, but becomes increasingly important as organizations engage in more cognitively demanding and interdisciplinary innovation efforts.

From a managerial standpoint, our results underscore the value of organizational structures and cultures that enable cross-ethnic collaboration. Recruiting diverse talent is not sufficient on its own: firms also need systems and practices that translate diversity into integrative collaboration—such as cross-functional team design, inclusive leadership development, boundary-spanning roles, and interventions that reduce informal segregation in everyday work. These investments are especially salient in settings characterized by complex innovation, where success depends on effectively combining distinct specializations and perspectives; in such contexts, ethnic integration can function as a strategic asset that supports knowledge recombination and higher-value invention.

At the same time, it is important to acknowledge that collaboration-network integration can arise through multiple—and likely heterogeneous—processes across firms. In some organizations, integration is plausibly intentional, shaped through organizational design and cultural practices (e.g., deliberate staffing and team formation, cross-project rotation, incentives that reward knowledge sharing, and norms that legitimize cross-boundary collaboration). In other organizations, integration may be more emergent, reflecting local labor-market conditions, hiring pipelines, informal social ties, or path-dependent project assignments rather than an explicit strategy. These pathways carry different implications.

If integration is primarily designed, our evidence is consistent with returns to purposeful managerial choices that reduce coordination frictions and facilitate recombination of complementary expertise. However, even when integration is primarily emergent—a byproduct of how work and people flow through an organization rather than the outcome of an explicit integration strategy—its presence or absence should not be viewed as random. Instead, it often reflects the unplanned consequences of existing frictions and routines: absent deliberate countervailing practices, everyday coordination costs and social-network inertia can produce persistent silos. This perspective also helps explain why intentionally inducing integration may not replicate the same benefits. If integration tends to arise organically when there is genuine task complementarity, repeated interaction, and sufficient coordination bandwidth, then mechanically mixing people without shared objectives, stable opportunities for collaboration, or supporting routines may generate only superficial integration—potentially increasing miscommunication, conflict, or duplicated effort—without delivering deep knowledge recombination. A practical takeaway is therefore to manage the conditions under which organic integration becomes productive—by shaping project interfaces, creating sustained cross-boundary interaction (rather than one-off mixing), clarifying shared goals, and providing lightweight coordination infrastructure, rather than treating integration as a purely compositional target.

Our findings also contribute to ongoing debates in policy and management practice. In fields such as STEM talent acquisition, R&D team formation, and innovation management, diversity is often treated as an end in itself. However, our evidence suggests a more nuanced view: diversity is a necessary but not sufficient condition for enhanced innovation outcomes. Without intentional integration mechanisms, the latent potential of diversity may remain unrealized. This perspective has practical implications for how organizations design, monitor, and evaluate their diversity initiatives—not only in terms of representation but also in terms of the extent to which collaboration spans ethnic and cultural divides.

Overall, this study tries to advance a more complete understanding of when and how ethnic diversity contributes to innovation. It emphasizes that diversity without integration is a missed opportunity, particularly in knowledge-intensive and innovation-driven firms. In a geopolitical climate where diversity initiatives are increasingly contested, our findings serve as a reminder that firms should not only continue to promote workforce diversity, but also go the extra mile—by investing in organizational practices that foster the effective integration of diverse talent within their internal collaboration networks.

## 5.1 | Limitations and further research

Our study is not without limitations that also point to opportunities for future research. First, our contribution lies in estimating the value of marginal increases in realized integration, rather than the effect of any single managerial intervention aimed at raising integration within firms. Put differently, we cannot distinguish contexts in which integration is intentionally induced through organizational design or cultural practices from contexts in which it develops more accidentally and emerges from routine staffing, local labor-market conditions, or informal networks. Future research can open this black box by clarifying when firms actively engineer integration versus when it evolves organically, identifying the specific levers that move it, and establishing the conditions under which designed integration reproduces the innovation benefits documented here.

Second, by relying on patent data, we observe only whether inventors co-author; we cannot assess the quality, intensity, or depth of their interactions. Nor do we capture informal or unpatented collaborations such as mentorship, internal project work, or informal knowledge exchanges. While it is unclear whether these unobserved ties differ in their degree of ethnic integration, future work could explore complementary data sources—such as communication logs, team project records, or mentoring networks—to gain deeper insight into how cross-ethnic collaboration emerges and operates.

Third, although we specify a linear relationship between ethnic integration and innovation performance, very low or very high integration could produce diminishing returns or even exacerbate conflict, suggesting a possible curvilinear effect. Future studies could explore this nonlinearity. Moreover, our focus on cross-group ties does not fully capture minority inventors' access to structurally central network positions—roles that facilitate knowledge recombination and grant greater influence over innovation processes. Incorporating group-level brokerage measures (e.g., betweenness centrality by ethnicity) would be a valuable extension.

Fourth, organizational design—such as R&D centralization versus decentralization—may correlate with both ethnic diversity and integration. If decentralized R&D units both hire more diverse inventors and foster cross-lab collaboration, some of our integration effect could in fact reflect organizational design. Although our controls and fixed-effects strategy mitigate this concern, future research with direct measures of R&D centralization could disentangle these intertwined effects.

Fifth, our sample primarily comprises large, highly innovative firms. Accordingly, the challenges and incentives for managing ethnic diversity may differ markedly in younger or smaller organizations. On the one hand, smaller workforces and simpler structures may make it easier to foster cross-ethnic ties and integration. On the other hand, such firms often face tighter resource constraints and lack formalized diversity-management practices, rendering integration more reliant on informal networks. Future research should examine how firm size and age moderate the relationship between ethnic integration and innovation performance.

Finally, our method of assigning ethnicity based on name classification, though widely used, has limitations. Probabilistic classifiers may misidentify some individuals' backgrounds. Future research could enhance accuracy by linking patent data with additional sources such as CVs or HR records (e.g., from platforms like Revelio Labs), which would enable more robust identification of inventors' demographics and career trajectories.

*Supporting Information*: Supporting Information Appendix S1 documents the methodology behind our study. It explains how the regression sample of 890 firms was constructed and presents descriptive statistics showing that inventor pools have grown more ethnically diverse but increasingly segregated along ethnic lines in co-invention networks. It details the construction of the standardized External–Internal (E–I) integration index, which compares observed cross-ethnic collaboration ties against a degree-preserving randomized null distribution. It also describes a firm-level measure of technological complexity built from patent subclass co-occurrence networks, plus the name-based ethnicity classification method (using Ethnea) and its validation against inventors' self-reported nationalities. Finally, it reports robustness checks using a patent-weighted version of the integration index.

## ACKNOWLEDGMENTS

This work was made possible through funding from the European Union's NextGenerationEU, under the National Recovery and Resilience Plan (NRRP) Mission 4, Component 2, Investment Line 1.5 (Project: MUSA—Multilayered Urban Sustainability Action) and Mission 4 Component 2 Investment Line 1.5 (Project: GRINS—Growing Resilient, INclusive and Sustainable). Open access publishing facilitated by Universita Bocconi, as part of the Wiley - CRUI-CARE agreement.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

*Stefano Breschi* https://orcid.org/0000-0002-1642-4233
*Luisa Gagliardi* https://orcid.org/0000-0001-6266-4572

## REFERENCES

Ahuja, G., & Lampert, C. M. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, *22*(6–7), 521–543.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*(2), 101–138.

Almeida, P., Phene, A., & Li, S. (2015). The influence of ethnic community knowledge on Indian inventor innovativeness. *Organization Science*, *26*(1), 198–217.

Arora, A., Belenzon, S., & Sheer, L. (2020). *DISCERN: Duke Innovation & SCientific Enterprises Research Network*. [Data set].

Bella, G., Batsuren, K., & Giunchiglia, F. (2021). A database and visualization of the similarity of contemporary lexicons. In K. Ekštein, F. Pártl, & M. Konopík (Eds.), *Text, speech, and dialogue* (pp. 95–104). Springer International Publishing.

Bojanowski, M., & Corten, R. (2014). Measuring segregation in social networks. *Social Networks*, *39*(1), 14–32.

Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, *28*(2), 124–136.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439–468.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349–399.

Carnabuci, G., & Operti, E. (2013). Where do firms' recombinant capabilities come from? Intraorganizational networks, knowledge, and firms' ability to innovate through technological recombination. *Strategic Management Journal*, *34*(13), 1591–1613.

Choudhury, P., & Kim, D. Y. (2019). The ethnic migrant inventor effect: Codification and recombination of knowledge across borders. *Strategic Management Journal*, *40*(2), 203–229.

Cox, T. H., & Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *Academy of Management Perspectives*, *5*(3), 45–56.

Crescenzi, R., & Gagliardi, L. (2018). The innovative performance of firms in heterogeneous environments: The interplay between external knowledge and internal absorptive capacities. *Research Policy*, *47*(4), 782–795.

Dahlin, K. B., Weingart, L. R., & Hinds, P. J. (2005). Team diversity and information use. *Academy of Management Journal*, *48*(6), 1107–1123.

Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, *46*(2), 229–273.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, *47*(1), 117–132.

Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly*, *52*(3), 443–475.

Freeman, R. B., & Huang, W. (2015). Collaborating with people like me: Ethnic coauthorship within the United States. *Journal of Labor Economics*, *33*(S1), S289–S318.

Gagliardi, L. (2015). Does skilled migration foster innovative performance? Evidence from British local areas. *Papers in Regional Science*, *94*(4), 773–795.

Gagliardi, L., & Novelli, E. (2025). Female entrepreneurs targeting women: Strategic redirection under scientific decision-making. *Organization Science*, *36*(2), 718–736.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Herring, C. (2009). Does diversity pay?: Race, gender, and the business case for diversity. *American Sociological Review*, *74*(2), 208–224.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(46), 16385–16389.

Hong, S.-Y., Kim, Y., & Lee, Y. (2023). Ethnic segregation on linguistic landscapes. *Environment and Planning B: Urban Analytics and City Science*, *50*(7), 1916–1931.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.

Kerr, S. P., & Kerr, W. R. (2018). Global collaborative patents. *The Economic Journal*, *128*(612), F235–F272.

Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, *90*(3), 518–537.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665–712.

Kogut, B., & Zander, U. (1996). What firms do? Coordination, identity, and learning. *Organization Science*, *7*(5), 502–518.

Koning, R., Samila, S., & Ferguson, J.-P. (2020). Inventor gender and the direction of invention. *AEA Papers and Proceedings*, *110*, 250–254.

Krackhardt, D., & Stern, R. N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, *51*(2), 123–140.

Larkey, L. K. (1996). Toward a theory of communicative interactions in culturally diverse workgroups. *The Academy of Management Review*, *21*(2), 463–491.

Laursen, K., Leten, B., Nguyen, N. H., & Vancauteren, M. (2020). Mounting corporate innovation performance: The effects of high-skilled migrant hires and integration capacity. *Research Policy*, *49*(9), 104034.

Lazear, E. P. (1999a). Culture and language. *Journal of Political Economy*, *107*(S6), S95–S126.
Lazear, E. P. (1999b). Globalisation and the market for team-mates. *The Economic Journal*, *109*(454), 15–40.
McLeod, P. L., Lobel, S. A., & Cox, T. H. (1996). Ethnic diversity and creativity in small groups. *Small Group Research*, *27*(2), 248–263.
Moser, P., & Nicholas, T. (2004). Was electricity a general purpose technology? Evidence from historical patent data. *American Economic Review*, *94*(2), 388–394.
Nathan, M., & Lee, N. (2013). Cultural diversity, innovation, and entrepreneurship: Firm-level evidence from London. *Economic Geography*, *89*(4), 367–394.
O'Leary, M. B., & Cummings, J. N. (2007). The spatial, temporal, and configurational characteristics of geographic dispersion in teams. *MIS Quarterly: Management Information Systems*, *31*(3), 433–452.
Ottaviano, G. I., & Peri, G. (2005). Cities and cultures. *Journal of Urban Economics*, *58*(2), 304–337.
Reagans, R. (2013). Demographic diversity as network connections: Homophily and the diversity-performance debate. In Q. M. Roberson (Ed.), *The Oxford handbook of diversity and work*, (Chapter 11) (pp. 192–204). Oxford University Press.
Reagans, R., & Zuckerman, E. W. (2001). Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science*, *12*(4), 502–517.
Rodan, S., & Galunic, C. (2004). More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness. *Strategic Management Journal*, *25*(6), 541–562.
Sana, M. (2010). Immigrants and natives in U.S. science and engineering occupations, 1994–2006. *Demography*, *47*(3), 801–820.
Saxenian, A. (2002). Silicon valley's new immigrant high-growth entrepreneurs. *Economic Development Quarterly*, *16*(1), 20–31.
Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin*, *129*(4), 475–494.
Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management Science*, *56*(1), 41–56.
Stahl, G. K., Maznevski, M. L., Voigt, A., & Jonsen, K. (2010). Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups. *Journal of International Business Studies*, *41*(4), 690–709.
Tortoriello, M., Reagans, R., & McEvily, B. (2012). Bridging the knowledge gap: The influence of strong ties, network cohesion, and network range on the transfer of knowledge between organizational units. *Organization Science*, *23*(4), 1024–1039.
Torvik, V. I., & Agarwal, S. (2016). Ethnea—An instance-based ethnicity classifier based on geo-coded author names in a large-scale bibliographic database. In *International symposium on science of science*. Library of Congress.
Torvik, V. I., & Smalheiser, N. R. (2009). Author name disambiguation in MEDLINE. *ACM Transactions on Knowledge Discovery from Data*, *3*(3), 1–29.
Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233.
Williams, K., & O'Reilly, C. A. (1998). Demography and diversity: A review of 40 years of research. In B. Staw & R. Sutton (Eds.), *Research in organizational behavior* (pp. 77–140). JAI Press.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Breschi, S., Gagliardi, L., & Yoon, T. (2026). Ethnic integration and the value of innovation. *Strategic Management Journal*, 1–29. https://doi.org/10.1002/smj.70108