Hidden Costs of Star Inventors in Organizational Innovation:
Knowledge Homogenization and Obsolescence

February 2024

Jaeho Kim
The Wharton School
kjaeho@wharton.upenn.edu

David H. Hsu
The Wharton School
dhsu@wharton.upenn.edu

***Abstract***: We explore a potential unintended consequence of research and development (R&D) teams involving star inventors. While the innovation process often requires experimentation and even questioning conventional wisdom, collaborators of star inventors may instead exhibit deferential behavior to stars and follow stars' proven knowledge to mitigate uncertainty in the innovation process. This may lead to homogenized team knowledge, and this convergent rather than divergent exploration process toward stars' knowledge may be ultimately detrimental to organizational innovation due to knowledge obsolescence. Within the context of startup inventors in the evolving fabless semiconductor industry, which experienced technological shifts during our period of study (1975-2020), we provide supportive empirical evidence based on a difference-in-differences design. We conclude with managerial implications of this research.

***Managerial summary***: While the conventional wisdom is that stars can benefit organizations through both their direct knowledge and their spillover effect on co-workers, we investigate a potential downside. We find that junior research and development (R&D) team members, when staffed with star inventors, subsequently exhibit knowledge investment behavior convergent with that of stars. This process may be antithetical to organizational environments best suited for innovation. We find evidence consistent with these ideas in the setting of startups in the fabless semiconductor industry. The results have implications for designing corporate R&D teams.

***Keywords***: star inventors; knowledge convergence; innovation performance; fabless semiconductor industry; technology startups.

## 1. INTRODUCTION

Scientific stars are those who have shown exceptional performance in inventions and scientific discoveries, reflecting their unique intellectual assets (Azoulay et al., 2010; Zucker & Darby, 1996). The traditional view in the literature is that these stars benefit organizations via their own contributions as well as through their positive idea spillovers to others in the organization. As a result, researchers have examined the organizational conditions under which stars maintain their productivity, particularly as they move across organizations (e.g., Groysberg et al., 2008; Huckman & Pisano, 2006). However, prior studies predominantly document positive knowledge spillovers between stars and others in the academic context and pay less attention to the *collaborators* of stars, particularly in the organizational enterprise setting (Chen & Garg, 2018; Oettl, 2012).

We contend that star inventors in firms may lead to negative knowledge spillover effects on their non-star collaborators. The negativity comes from star inventors homogenizing non-star inventors' knowledge toward that of stars after their collaboration. Homogenized knowledge between inventors may be antithetical to innovation, especially in rapidly moving environments (Teodoridis et al., 2019) and within emerging enterprises, undermining the quality of non-stars' subsequent innovation after interacting with stars in the same team. Hence, we suggest that the traditional account of positive star collaborator effects (e.g., Azoulay et al., 2010) may be more nuanced within an organizational (not academic institution) context.

To test these hypotheses, we examine fabless startups in the semiconductor industry. These firms specialize in semiconductor chip designs without having capital-intensive manufacturing facilities, thus lowering the entry barrier for startups. In this setting, intangible human capital knowledge assets (rather than manufacturing capability) are a key resource. The industry also experienced substantial technological changes, moving from traditional chip designs (e.g., central processing units, CPUs) to new applications (e.g., customizable chips for the Internet of Things).

Based on 372 startups, 18,621 inventors, and innovation outcomes from 1975 to 2020, we find supporting evidence for our arguments. After interacting with stars in the team, non-star inventors' knowledge becomes about 10% more similar to stars' knowledge (doubled for novice inventors). Non-star inventors experiencing knowledge homogenization during their team interaction experience 38% lower innovation quality on average as measured by citation counts (with a further 18 percentage point decline if the star had obsolete knowledge).

Our findings have a number of implications. When designing R&D teams, firms should be cautious about heavily imbalanced team composition. Attention to team behaviors based on composition is important, akin to considering individual experience or expertise in forming teams. In addition, it is important to factor in the likely change in the technology environment, compared to knowledge bases that are or will be possessed by inventing teams. For the academic literature, we emphasize the consideration of team and organizational boundaries that may modify inferences in generalizing from prior research for which such boundaries are not as salient – as is the case in comparing our results to the academic institution context in which teams are largely self-assembled and an "invisible college" mentality makes organizational boundaries arguably less important.

## 2. BACKGROUND, LITERATURE, AND HYPOTHESES

### 2.1. Background

Based on the observation that typically, a small percentage of individuals account for an outsize share of a field's overall output (e.g., Lotka, 1926), the literature has examined star researchers. For example, Zucker et al. (1998) find that at the birth of the biotechnology industry, new ventures are much more likely to geographically co-locate proximately to the most productive scientists (i.e., stars). This suggests that before specialized knowledge is codified, accessing leading scientists' tacit knowledge is particularly important for commercial enterprises utilizing new technology. Management scholars also examine the degree to which star employees operate at high levels, even across organizational contexts, with the finding that organizational and team conditions shape star performance (e.g., Groysberg et al., 2008; Huckman & Pisano, 2006).

There is also increasing interest in how stars influence their collaborators, which is the antecedent to our study. Azoulay et al. (2010), studying life science stars in academia (not in firms), find that they bring unique research ideas to their collaborative relationships, which is irreplaceable when the stars unexpectedly perish. With attention to how stars might impact *organizational* innovation more generally, Kehoe & Tzabbar (2015) find that biotechnology R&D teams involving star inventors may limit the emergence of their collaborators as innovation leaders (operationalized as the number of subsequent patents that do not involve the star). Left unaddressed is a more fulsome analysis of collaborators' behaviors and innovation outcomes (e.g., Oettl, 2012).

A final relevant strand of prior work increasingly recognizes that the way teams are composed within organizations more generally can shape innovation outcomes. For example, even holding constant individual technical experience, alternative team compositions can shape different trajectories of organizational innovation (Aggarwal et al., 2020; Chang, 2023). This literature does not consider the R&D team behavior and outcomes associated with star-collaborator team compositions, however, which we do here.

### 2.2. Knowledge homogenization between star and non-star inventors

When star and non-star inventors interact in an inventing team under knowledge asymmetry, the flow of knowledge influence is likely unidirectional, from stars to non-stars. A key premise of prior research on team collaboration is that high-performing employees are more likely to influence low-performing employees, such as "star-centric" interactions (Chen & Garg, 2018). The low performers have much to learn from the high performers, and the high performers tend to have stronger authority in team decision-making, particularly in technology-intensive environments (Astley & Sachdeva, 1984; Burke et al., 2007).

In addition to the authority-based explanation of knowledge direction, uncertainty in innovation tasks plays a critical role in shaping such unidirectional knowledge influence. The underlying uncertainty in the innovation process may drive stars to over-rely on exploiting their previously successful knowledge trajectories (Audia & Goncalo, 2007; Staats et al., 2018), establishing traditional knowledge within the existing dominant design (Suarez & Utterback, 1995). Although non-stars face lower switching costs to new technologies compared to stars (Jovanovic & Nyarko, 1996), non-stars are tempted to follow the proven knowledge of stars rather than engage in unchartered exploration. Thus, both stars and non-stars tend to focus on utilizing stars' traditional knowledge domain during their collaboration, making their knowledge bases more similar after the team interaction. This motivates our first hypothesis:

**Hypothesis 1 (H1):** *After interacting with star inventors in a team, non-star inventors' knowledge becomes homogenized toward stars' traditional knowledge domain*.

### 2.3. Non-stars' innovation performance after knowledge homogenization

We further argue the homogenized knowledge of non-stars toward stars' traditional knowledge undermines the quality of non-stars' subsequent innovation. Stars' traditional knowledge faces the issue of knowledge obsolescence over time. The knowledge obsolescence problem has been studied at the level of organization (Leonard-Barton, 1992; Sørensen & Stuart, 2000) and invention (Capaldo et al., 2017; Jain, 2016), and we apply the implications to individual inventors equipped with a specific set of knowledge domains. Although stars' traditional knowledge contributed to their stardom, such knowledge tends to lose its power and impact, especially under rapid technological advances.

Also, non-stars suffer from knowledge misfit between their existing non-traditional knowledge and the absorbed knowledge from stars. Non-stars' different knowledge backgrounds may prevent them from making the best use of traditional knowledge, due to inventor knowledge dissimilarity (Carnabuci & Operti, 2013; Dougherty, 1992). The mere absorption of knowledge driven by the unidirectional influence may not lead to a complete integration of knowledge due to the absence of mutual contributions and proactive discussions (Gardner et al., 2012; Vestal & Danneels, 2023). Thus, we make a theoretical prediction that non-stars homogenized toward stars' traditional knowledge are associated with diminished innovation quality (see Appendix 1 for an analytical model based on these concepts), motivating the second hypothesis:

**Hypothesis 2 (H2):** *After interacting with star inventors in a team, non-star inventors' innovation quality is diminished.*

## 3. DATA AND METHODS

### 3.1. Empirical setting

The empirical setting is fabless startups in the United States (US). Fabless is an industry segment in the semiconductor industry where firms specialize in semiconductor chip design by delegating the production of chips to manufacturing-focused firms, called "foundries" (Macher et al., 1998).[1] The fabless startup context is a good one for testing our hypotheses for two key reasons. First, fabless startups are knowledge-based ventures where the interaction of ideas between inventors plays a critical role. The lower entry costs based on specialization in chip design enable fabless

[1] Fabless firms emerged around the mid-1980 when the design of integrated circuits became standardized across firms through Design Rule Checking (DRC), and the introduction of Electronic Design Automation (EDA) enables firms to store their chip design blueprints electronically and transfer them to other firms (Kapoor, 2013).

startups to focus on fostering their innovative ideas. In addition, knowledge exchange and influence between stars and non-stars become more direct and intense due to the limited number and size of inventing teams in fabless startups compared to incumbent counterparts.

Second, the fabless segment experienced substantial technological changes. In the early generation, fabless startups designed traditional integrated circuits and developed programmable logic technology (e.g., CPU chip design), much like their incumbent counterparts (Kapoor, 2013). However, fabless startups in the late generation innovate for different types of chip designs, used for diverse end-products beyond the traditional semiconductor market, such as Internet of Things-(IoT) or artificial intelligence (AI)-embedded devices (Flynn et al., 2017). Recent technological advances made after 2000, such as the reconfigurable computing features in Field-Programmable Gate Arrays (FPGAs), are key enablers of flexible and unique chip designs, customizable for various products (Compton & Hauck, 2002).

**3.2. Data source and sample**

We use data from Pitchbook to identify fabless startups among venture capital-backed firms in the US. We first filter by keywords associated with fabless startups: "fabless," "semiconductor," "chip," and "circuit." We then manually investigate whether the descriptions correctly refer to the characteristics of fabless firms that specialize in chip design, which results in 638 venture-backed fabless startups (see Appendix 2 for details about fabless identification).

We use US Patent and Trademark Office (USPTO) data to track inventors' innovation outcomes records, as patenting is customary in the semiconductor industry (Hall and Ziedonis, 2001). We first fuzzy match the Pitchbook fabless startups to the USPTO assignees by their names and manually verify the potential matches (Wasi & Flaaen, 2015). For the matched firms, we further retrieve their granted patents in the US to identify co-patenting teams and associated inventors. We collect the inventors' patenting history before and after joining fabless firms. This process results in 372 startups, 18,621 inventors, 27,727 inventing teams, and innovation outcomes between 1975 and 2020 (see Appendix 2 for details about sample selection and matching).

**3.3. Dependent variables**

As the outcome variable to test H1, we use *Knowledge Similarity* by investigating whether an inventor's knowledge becomes similar to a star inventor's knowledge in the team after their interaction. For all inventors in a team, we calculate the cosine similarity in each inventor dyad before and after the team interaction by using the two inventors' patent applications (which were eventually granted) in each year (see Appendix 2 for details about variable construction). We use the Cooperative Patent Classification (CPC) codes in patents to characterize each inventor's knowledge domain in vector space.

The outcome variables to test H2 are *Scaled Citation*, *Generality*, and *Originality* to see how the quality of non-stars' inventions changes after interacting with star inventors (see Appendix 2 for details about pre- vs. post-team windows). Forward citation count is a proxy of value and novelty of innovation (Trajtenberg, 1990), and we use *Scaled Citation* that adjusts the raw citation count by considering the year and category fixed effects to control for patent citation and technological category trends (Trajtenberg et al., 1997). We further examine two additional measures of innovation quality. *Generality* measures the diversity of patent classes in future inventions that build upon the focal patent, and *Originality* captures the diversity of patent classes the focal patent relies on. Both are constructed using patent class Herfindahl indices (Hall et al., 2001; Trajtenberg et al., 1997). As the three innovation quality measures are at the patent level, we average them to construct inventor-year level observations.

**3.4. Independent variables**

In testing H1 where the unit of analysis is at the team inventor dyad-year level, *Post Team Interaction* is a binary time indicator of whether an observation is after the team interaction (defined as two years before the team patent application), and *Star in Dyad* is a binary treatment indicator of whether a dyad includes a star. Consistent with prior literature (e.g., Kehoe and Tzabbar, 2015), stars are defined by the top five percent inventors among all fabless inventors in each fabless joining year cohort, and the ranking is determined by the number of patents applied *before* joining the fabless firm, weighted by those patents' citation count.[2]

The independent variables to test H2 are constructed at the inventor-year level. *Star in Team* is a treatment indicator of whether a team includes a star. For the mechanism test, we further differentiate stars by the time gap between stars' emergence year (i.e., fabless joining year) and the team interaction year.[3] Compared to the team interaction year, stars on average emerged four years earlier (min: -21, max: 0). For teams having a larger (smaller or equal) year gap than the average, we assign a value of one to *Past Star in Team* (*Contemporary Star in Team*).

Table 1 shows summary statistics and the correlations of the variables. Note that there are two different datasets: Panel A shows the team inventor dyad-year-level dataset to test H1, whereas Panel B shows the inventor-year-level dataset to test H2. Star inventors are excluded in Panel B because H2 investigates the innovation performance changes of non-star inventors.[4]

### 3.5. Estimation strategy

We adopt a staggered difference-in-differences (DID) design to test H1 and H2 as follows:

---

[2] In addition to this main treatment variable, we also include additional treatment variables for robustness checks: *Star-Star Dyad*, *Star-Middle Dyad*, *Star-Novice Dyad*, and *Novice in Dyad*. Novice inventors are those who have two or fewer pre-fabless patents (about 66% of the inventors), and middle inventors are those who are neither star nor novice inventors.
[3] If there is more than one star in a team, we take an average of their emergence years.
[4] See Appendix 2 for details about sample restrictions.

$$\text{H1: } Knowledge\ Similarity_{i,t} = \beta_0 + \beta_1 Post\ Team\ Interaction_{i,t} + \beta_2 Post\ Team\ Interaction_{i,t} \times Star\ in\ Dyad_{i,t} + \gamma_f + \delta_m + \theta_i + \mu_o + \sigma_y + \varepsilon_{i,t}$$

$$\text{H2: } Innovation\ Quality_{j,t} = \beta_0 + \beta_1 Post\ Team\ Interaction_{j,t} + \beta_2 Post\ Team\ Interaction_{j,t} \times Star\ in\ Team_{j,t} + \gamma_f + \delta_m + \rho_j + \mu_o + \sigma_y + \varepsilon_{j,t}$$

where $i$: inventor dyad, $j$: inventor, $t$: time before and after team interaction, $\gamma_f$: firm fixed effects, $\delta_m$: team fixed effects, $\theta_i$: dyad fixed effects, $\rho_j$: inventor fixed effects, $\mu_o$: team formation year fixed effects, $\sigma_y$: year fixed effects, $\varepsilon$: the error term. Standard errors are clustered at the dyad level in H1 and the inventor level in H2.[5] For both H1 and H2, $\beta_2$ is the key estimator of interest.

Although the DID design addresses some empirical confounds, it may still suffer from selection issues because firms do not randomly assign their inventors to teams, especially their star inventors. We use multiple fixed effects in the model specification to mitigate this issue (see Appendix 3 for discussions about endogenous selection). We cannot leverage inventors' time-varying covariates as they are mostly correlated with either innovation performance or team selection, violating DID design prerequisites (Freedman et al., 2023).

In addition, we apply coarsened exact matching (CEM) in H2 where the performance measure is the dependent variable in an effort to further mitigate remaining selection issues that may threaten the parallel pre-trend assumption (Bessen et al., 2023; Burford et al., 2022). Based on inventors' pre-fabless patent count and years of patenting experience, we one-to-one match our treated sample with the most comparable set of counterfactual inventors.[6]

## 4. RESULTS

[5] The estimation for H1 is weighted by the minimum value of each inventor's patent count in each dyad-year observation, which prevents a dyad with too few patents from overinfluencing the estimates. The coefficients of *Star in Dyad* and *Star in Team* are omitted due to the dyad and team fixed effects, similar to a two-way fixed effects setting.
[6] The matching is conducted at the team-inventor level because an inventor can be associated with more than one team. The team and inventor fixed effects address any different dynamics in multiple teams.

### 4.1. Main results

Table 2 shows supporting evidence for H1. Columns (1)-(3) show the estimates with different fixed effects, and the estimates of *Post Team Interaction* × *Star in Dyad* are statistically significant in all three columns. On average, an inventor interacting with a star in a team tends to have about 10% more similar knowledge to the star after the interaction, compared to another inventor who does not interact with a star in a team. Columns (4)-(6) test how the knowledge homogenization effect differs by the focal inventor type. If the focal inventor is also a star, there is no knowledge homogenization, whereas the homogenization is strongest if the focal inventor is a novice inventor (16% more similar). One might believe novice inventors are malleable and prone to be affected by *any* (more senior) inventors, but Columns (7)-(8) show novice inventors do not show knowledge homogenization when exposed to other non-star inventors.

The results in Panel A in Table 3 support H2. Columns (1)-(2) use *Scaled Citation* as the innovation quality measure. Column (1) shows the estimate in the full sample, while the estimate in Column (2) reflects the CEM-adjusted sample. The key estimates of *Post Team Interaction* × *Star in Team* are statistically significant in both columns. On average, a non-star inventor who interacts with a star inventor in a team is associated with a 38% reduction in subsequent patent quality as compared to another non-star inventor who does not interact with a star inventor in a team. We validate the parallel pre-trends between the treatment and control groups in the temporal DID estimates and find a significant drop in *Scaled Citation* appears in t+7, linked with the fact that *Knowledge Similarity* takes about six years to increase (see Figure A1 in Appendix 3).

The other specifications in Panel A use alternative innovation quality measures (following the same specification structure as above). Columns (3)-(4) use *Generality* as the quality measure, while Columns (5)-(6) use *Originality* as the quality measure. The former group of results suggests that non-stars' previous exposure to stars is correlated with a reduced span of technologies building on the focal innovation (i.e., *Generality*), while the latter group of results suggests no change in *Originality* associated with star exposure.

**4.2. Testing mechanisms for innovation quality**

The primary driver of the main results in H2 is the obsolescence of stars' knowledge under the rapid pace of technological changes in the fabless industry. To validate it, we first explore the technological trends in fabless. Panel A in Figure 2 shows the distribution of fabless startups by founding years and their association with emerging technologies at the time of founding.[7] The number of new fabless startup foundings peaks around 2000 and decreases gradually after that, while the percentage of fabless firms targeting emerging technologies surges after 2000.[8] These patterns suggest a rapid pace of technological change among fabless ventures in the sample.

We further investigate whether stars' knowledge becomes obsolete over time with industry technological changes. To do so, we compare knowledge that star versus novice inventors bring to their fabless enterprises (their pre-fabless knowledge). For stars who joined fabless before 2000, we find the top 10 most common knowledge domains in their pre-fabless patents. We also identify the top 10 most common knowledge domains of novice inventors who joined fabless after 2010 to enhance the comparison (see Table A1 in Appendix 3 for the lists). For each knowledge domain, we track its citation count trend in the entire patent space. The yearly raw citation count of each knowledge domain is divided by the number of all citations made in each year, yielding an adjusted citation index, which we use to control for time-based citation trends. Panel B in Figure 2 shows

[7] We define a fabless startup's founding as associated with emerging technologies if its industry vertical is related to machine learning, virtual reality, autonomous cars, big data, cryptocurrency, digital health, IoT, robots, drones, or wearables.
[8] The line indicating the percentage of fabless startups' associated with emerging technologies is created by locally weighted scatterplot smoothing.

the adjusted citation index of knowledge domains, comparing stars to novice inventors. Stars' knowledge domains initially have much higher citation indices, but novices' knowledge domains catch up and eventually exhibit higher indices in later periods. This suggests the stars' knowledge tends to lose value over time, with the opposite pattern for novice inventors' knowledge.

To formally test the role of knowledge obsolescence in undermining non-stars' innovation quality, we include additional variables that consider the time gap between stars' emergence and team interaction (i.e., *Past Star in Team* and *Contemporary Star in Team*). Panel B in Table 3 shows that the major part of non-stars' reduced innovation quality is from past stars (56% reduction in *Scaled Citation* and 16% reduction in *Generality*). Interaction with past stars even reduces *Originality* (10% reduction), which does not show significant results when all stars are considered. These additional results validate the main findings and corroborate our theoretical expectations.

## 5. DISCUSSION AND CONCLUSION

We present evidence that star inventors in an inventing team homogenize non-star inventors' knowledge, and the knowledge homogenization undermines the quality of non-stars' subsequent innovation. These findings suggest that stars may not always provide positive knowledge spillovers to non-stars; instead, the relationship may depend on inventor behavior, intra- versus inter-organizational boundaries of exposure, as well as the pace of technological change in a given industrial setting relative to stars' knowledge recency and relevance (e.g., Teodoridis et al., 2019).

This research contributes to the literature on team-level innovation and knowledge sharing. Research on team-level innovation tends to focus on either team members' performance differences or knowledge domain differences. Instead, we explore the behavior and innovation consequences of intra-organizational R&D teams composed of star and non-star inventors. Our knowledge homogenization results run counter to the conventional wisdom of positive knowledge spillovers, which have mainly been explored in the domain of inter-organizational academic teams.

While this research provides novel findings and contributions, it is important to acknowledge its limitations (while providing opportunities for future research). First, our ability to observe teams is restricted to those with at least one patent. The omission of teams that fail to achieve innovation outcomes could pose a challenge, particularly if the interaction between star and non-star inventors plays a role in such failures. Future research utilizing unselected R&D team data and outcomes would be most welcome. Second, we lack information about the intermediate knowledge development processes within teams. Unraveling those steps could reveal interesting contingencies or effective interventions to stem the homogenization effects. Finally, our estimates should not be interpreted as causal, since our empirical approach lacks exogenous variation.

More broadly, future research may not only consider team composition effects in light of our results but also investigate how established levers of innovation may interact with the homogenization effects we highlight. For example, at the organizational level, a tolerance for failure (Manso, 2011) and an associated culture of exploration and experimentation has been shown to promote innovation. Can firms with such a culture staff R&D teams with stars and non-stars without suffering detrimental homogenization effects? Similarly, knowledge recombination has long been discussed as a positive influence on innovation (e.g., Fleming, 2001). Future work in this area may investigate innovation outcomes associated with staffing teams with researchers of different accomplishment levels who themselves also have knowledge and experience in disparate domains. These types of studies might shed light on the relative importance of different levers of innovation. Clearly, there is much work ahead to more fully understand this area, but our hope is that this work initiates that conversation.

**REFERENCES**


Aggarwal VA, Hsu DH, Wu A. 2020. Organizing knowledge production teams within firms for innovation. *Strategy Science* **5**(1): 1–16.

Astley WG, Sachdeva PS. 1984. Structural sources of intraorganizational power: A theoretical synthesis. *Academy of Management Review* **9**(1): 104–113.

Audia PG, Goncalo JA. 2007. Past success and creativity over time: A study of inventors in the hard disk drive industry. *Management Science* **53**(1): 1–15.

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

Bessen J, Goos M, Salomons A, Van Den Berge W. 2023. What happens to workers at firms that automate? *Review of Economics and Statistics*: 1–45.

Burford N, Shipilov A V, Furr NR. 2022. How ecosystem structure affects firm performance in response to a negative shock to interdependencies. *Strategic Management Journal* **43**(1): 30–57.

Burke MA, Fournier GM, Prasad K. 2007. The diffusion of a medical innovation: Is success in the stars? *Southern Economic Journal* **73**(3): 588–603.

Capaldo A, Lavie D, Petruzzelli AM. 2017. Knowledge maturity and the scientific value of innovations: The roles of knowledge distance and adoption. *Journal of Management* **43**(2): 503–533.

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.

Chang MH. 2023. Cascading innovation: R&D team design and performance implications of mobility. *Strategic Management Journal* **44**(5): 1218–1253.

Chen JS, Garg P. 2018. Dancing with the stars: Benefits of a star employee's temporary absence for organizational performance. *Strategic Management Journal* **39**(5): 1239–1267.

Compton K, Hauck S. 2002. Reconfigurable computing: A survey of systems and software. *ACM Computing Surveys* **34**(2): 171–210.

Dougherty D. 1992. Interpretive barriers to successful product innovation in large firms. *Organization Science* **3**(2): 179–202.

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

Flynn D, Myers J, Toh S. 2017. Design methodologies for IoT systems on a chip. In *Enabling the Internet of Things*, Alioto M (ed). Springer: 271–286.

Freedman SM, Hollingsworth A, Simon KI, Wing C, Yozwiak M. 2023. Designing difference in difference studies with staggered treatment adoption: Key concepts and practical guidelines. *NBER Working Paper*.

Gardner HK, Gino F, Staats BR. 2012. Dynamically integrating knowledge in teams: Transforming resources into performance. *Academy of Management Journal* **55**(4): 998–1022.

Groysberg B, Lee L-E, Nanda A. 2008. Can they take it with them? The portability of star knowledge workers' performance. *Management Science* **54**(7): 1213–1230.

Hall BH, Jaffe AB, Trajtenberg M. 2001. The NBER patent citations data file: lessons, insights, and methodological tools. *NBER Working Paper 8498*.

Huckman RS, Pisano GP. 2006. The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science* **52**(4): 473–488.
Jain A. 2016. Learning by hiring and change to organizational knowledge: Countering obsolescence as organizations age. *Strategic Management Journal* **37**(8): 1667–1687.
Jovanovic B, Nyarko Y. 1996. Learning by doing and the choice of technology. *Econometrica* **64**(6): 1299–1310.
Kapoor R. 2013. Persistence of integration in the face of specialization: How firms navigated the winds of disintegration and shaped the architecture of the semiconductor industry. *Organization Science* **24**(4): 1195–1213.
Kehoe RR, Tzabbar D. 2015. Lighting the way or stealing the shine? An examination of the duality in star scientists' effects on firm innovative performance. *Strategic Management Journal* **36**(5): 709–727.
Leonard-Barton D. 1992. Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal* **13**(S1): 111–125.
Lotka AJ. 1926. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* **16**(12): 317–323.
Macher JT, Mowery DC, Hodges DA. 1998. Reversal of fortune? The recovery of the US semiconductor industry. *California Management Review* **41**(1): 107–136.
Manso G. 2011. Motivating innovation. *Journal of Finance* **66**(5): 1823–1860.
Oettl A. 2012. Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science* **58**(6): 1122–1140.
Sørensen JB, Stuart TE. 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly* **45**(1): 81–112.
Staats BR, KC DS, Gino F. 2018. Maintaining beliefs in the face of negative news: The moderating role of experience. *Management Science* **64**(2): 804–824.
Suarez FF, Utterback JM. 1995. Dominant designs and the survival of firms. *Strategic Management Journal* **16**(6): 415–430.
Teodoridis F, Bikard M, Vakili K. 2019. Creativity at the knowledge frontier: The impact of specialization in fast-and slow-paced domains. *Administrative Science Quarterly* **64**(4): 894–927.
Trajtenberg M. 1990. A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics* **21**(1): 172–187.
Trajtenberg M, Henderson R, Jaffe AB. 1997. University versus corporate patents: a window on the basicness of invention. *Economics of Innovation and New Technology* **5**(1): 19–50.
Vestal A, Danneels E. 2023. Unlocking the inventive potential of knowledge distance in teams: How intrateam network configurations provide a key. *Organization Science*.
Wasi N, Flaaen A. 2015. Record linkage using Stata: preprocessing, linking, and reviewing utilities. *Stata Journal* **15**(3): 672–697.
Zucker LG, Darby MR. 1996. Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences* **93**(23): 12709–12716.
Zucker LG, Darby MR, Brewer MB. 1998. Intellectual capital and the birth of US biotechnology enterprises. *American Economic Review* **88**(1): 290–306.

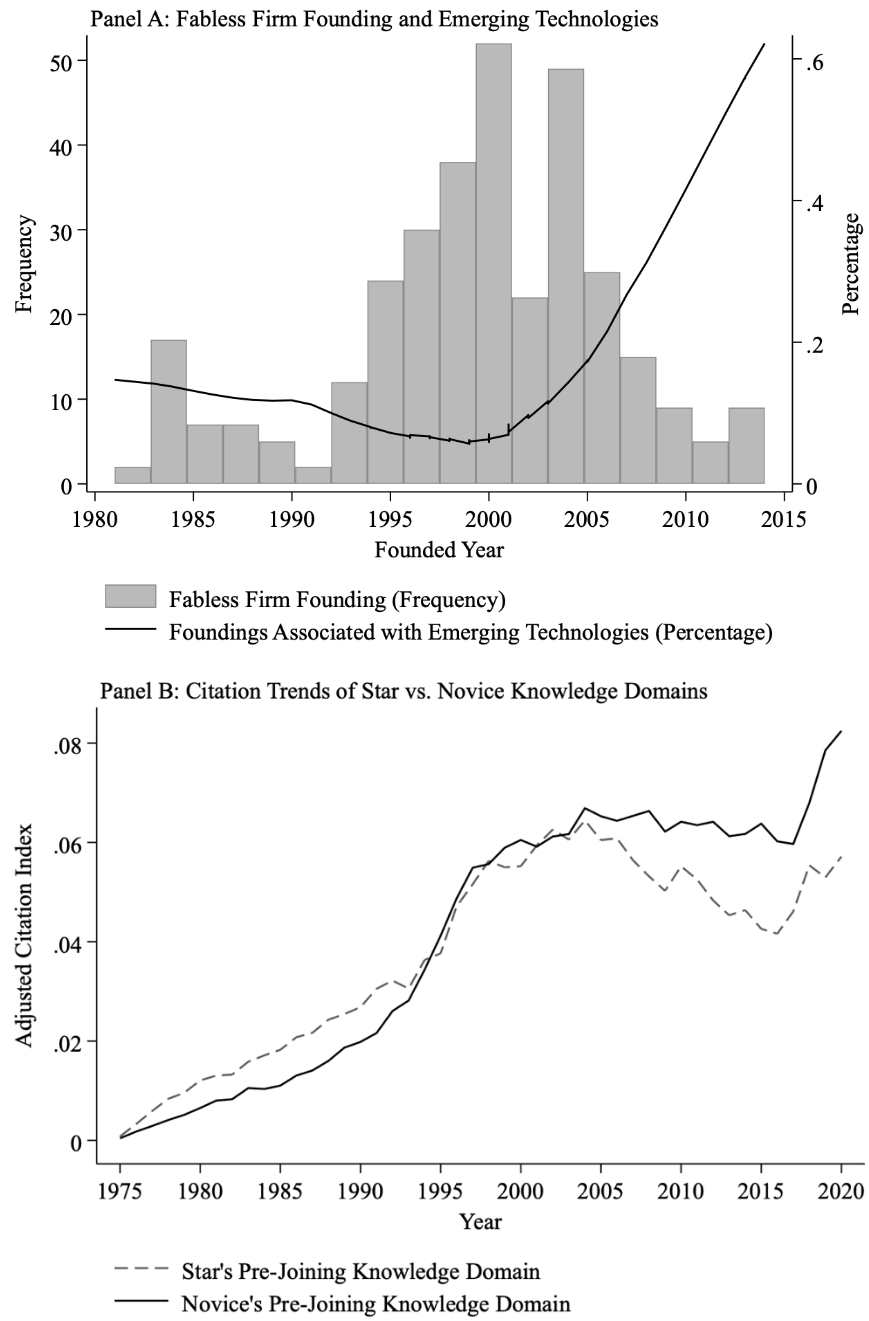


**FIGURE 1: Trends of Fabless Firm Founding and Knowledge Domains.** Panel A shows fabless startups' founding years (gray bars) and the percentage of fabless startups targeting emerging technologies at their founding (black line). Emerging technologies include machine learning, virtual reality, autonomous cars, big data, cryptocurrency, digital health, Internet of Things (IoT), robots, drones, and wearables. Panel B shows the citation trends of the top 10 most common knowledge domains (see Appendix Table A1) that star and novice inventors have before joining fabless firms. Stars who entered before 2000 and novice inventors who joined fabless firms after 2010 are included. The adjusted citation index is the citation count that each set of knowledge domains receives divided by the total citations made in each year.

**TABLE 1: Summary Statistics and Correlations.** Panel A shows the observations for testing Hypothesis 1 where the unit of analysis is at the team inventor dyad-year level. Panel B shows the observations for testing Hypotheses 2 where the unit of analysis is at the firm-inventor-year level. Panel B excludes star inventors to investigate the innovation outcome of non-star inventors. Panel A includes teams with co-patenting years up to 2015 (this cutoff provides plenty of post-team time to elapse, allowing us to avoid data incompleteness issues). The time period cutoff is shortened in Panel B because patent citation data suffers more from the truncation problem in recent years, particularly after 2020. To alleviate the imbalance problem, we decrease the cutoff to 2012 in Panel B.

Panel A: Inventor dyad-year sample (at firm-team-inventor dyad-year level; including star inventors)

| | N | Mean | SD | Min | Max | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *Knowledge Similarity* | 2,034,353 | 0.334 | 0.438 | 0.000 | 1.000 | | | | | | | |
| (2) *Post Team Interaction* | 2,034,353 | 0.679 | 0.467 | 0.000 | 1.000 | 0.178 | | | | | | |
| (3) *Star in Dyad* | 2,034,353 | 0.153 | 0.360 | 0.000 | 1.000 | 0.021 | -0.035 | | | | | |
| (4) *Star-Star Dyad* | 2,034,353 | 0.013 | 0.115 | 0.000 | 1.000 | 0.040 | -0.006 | 0.243 | | | | |
| (5) *Star-Middle Dyad* | 2,034,353 | 0.077 | 0.266 | 0.000 | 1.000 | 0.030 | -0.031 | 0.725 | -0.021 | | | |
| (6) *Star-Novice Dyad* | 2,034,353 | 0.063 | 0.243 | 0.000 | 1.000 | -0.020 | -0.016 | 0.595 | -0.017 | -0.052 | | |
| (7) *Novice in Dyad* | 2,034,353 | 0.665 | 0.472 | 0.000 | 1.000 | -0.035 | 0.072 | -0.218 | -0.121 | -0.362 | 0.143 | |

Panel B: Inventor-year sample (at firm-team-inventor-year level; excluding star inventors)

| | N | Mean | SD | Min | Max | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *Scaled Citation* | 1,032,695 | 0.587 | 1.326 | 0.000 | 47.107 | | | | | | | |
| (2) *Generality* | 1,032,695 | 0.235 | 0.280 | 0.000 | 0.933 | 0.495 | | | | | | |
| (3) *Originality* | 1,032,695 | 0.277 | 0.291 | 0.000 | 0.942 | 0.368 | 0.706 | | | | | |
| (4) *Post Team Interaction* | 1,032,695 | 0.747 | 0.435 | 0.000 | 1.000 | -0.109 | -0.115 | -0.026 | | | | |
| (5) *Star in Team* | 1,032,695 | 0.124 | 0.329 | 0.000 | 1.000 | 0.068 | 0.077 | 0.082 | -0.049 | | | |
| (6) *Past Star in Team* | 1,032,695 | 0.046 | 0.209 | 0.000 | 1.000 | 0.019 | 0.034 | 0.042 | -0.048 | 0.593 | | |
| (7) *Contemporary Star in Team* | 1,032,695 | 0.078 | 0.268 | 0.000 | 1.000 | 0.067 | 0.066 | 0.066 | -0.021 | 0.742 | -0.100 | |

**TABLE 2: Knowledge Similarity between Stars and Non-Stars After Team Interaction.** Fixed-effect regressions at the firm-team-inventor dyad-year level. The dependent variable is *Knowledge Similarity*. The "Dyads with Stars" sample includes dyads having at least one star inventor, whereas the "Dyads without Stars" sample includes the other dyads. Robust standard errors are clustered by inventor dyad and shown in parentheses, and p-values are in brackets.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| *Post Team Interaction* | 0.291 | 0.269 | 0.247 | 0.253 | 0.249 | 0.248 | 0.205 | 0.258 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.006) | (0.003) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| *Star in Dyad* | -0.058 | -0.027 | | | | | | |
| | (0.004) | (0.004) | | | | | | |
| | [0.000] | [0.000] | | | | | | |
| *Post Team Interaction × Star in Dyad* | 0.021 | 0.029 | 0.028 | | | | | |
| | (0.005) | (0.004) | (0.004) | | | | | |
| | [0.000] | [0.000] | [0.000] | | | | | |
| *Post Team Interaction × Star-Star Dyad* | | | | -0.128 | | | | |
| | | | | (0.019) | | | | |
| | | | | [0.000] | | | | |
| *Post Team Interaction × Star-Mid Dyad* | | | | | 0.028 | | | |
| | | | | | (0.005) | | | |
| | | | | | [0.000] | | | |
| *Post Team Interaction × Star-Novice Dyad* | | | | | | 0.055 | | |
| | | | | | | (0.006) | | |
| | | | | | | [0.000] | | |
| *Post Team Interaction × Novice in Dyad* | | | | | | | 0.046 | -0.005 |
| | | | | | | | (0.008) | (0.003) |
| | | | | | | | [0.000] | [0.087] |
| Sample | Full | Full | Full | Full | Full | Full | Dyads with Stars | Dyads without Stars |
| Firm FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Team FE | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Dyad FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Team Formation Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-squared | 0.122 | 0.316 | 0.461 | 0.461 | 0.461 | 0.461 | 0.415 | 0.465 |
| Firm Count | 372 | 372 | 372 | 372 | 372 | 372 | 150 | 368 |
| Inventor Count | 18,621 | 18,621 | 18,621 | 18,621 | 18,621 | 18,621 | 3,859 | 17,883 |
| Team Count | 27,727 | 27,727 | 27,727 | 27,727 | 27,727 | 27,727 | 4,132 | 26,062 |
| Team-Inventor Count | 87,745 | 87,745 | 87,745 | 87,745 | 87,745 | 87,745 | 14,280 | 81,436 |
| Dyad Count | 122,240 | 122,240 | 122,240 | 122,240 | 122,240 | 122,240 | 11,489 | 110,751 |
| Observations | 2,034,353 | 2,034,353 | 2,034,353 | 2,034,353 | 2,034,353 | 2,034,353 | 220,258 | 1,814,095 |

**TABLE 3: Non-Stars' Innovation Performance After Team Interaction with Stars.** Fixed-effect regressions at the firm-team-inventor-year level. The "CEM" sample includes the one-to-one matched observations between the treated and control groups by coarsened exact matching at the team-inventor level. Robust standard errors clustered by inventor are shown in parentheses, and p-values are in brackets.

| | Panel A: Main Results | | | | | | Panel B: Mechanism Tests | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DV: | Scaled Citation | | Generality | | Originality | | Scaled Citation | | Generality | | Originality | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *Post Team Interaction* | 0.174 | 0.311 | 0.075 | 0.081 | 0.081 | 0.084 | 0.266 | 0.170 | 0.081 | 0.069 | 0.089 | 0.077 |
| | (0.015) | (0.037) | (0.003) | (0.005) | (0.003) | (0.005) | (0.035) | (0.025) | (0.004) | (0.004) | (0.005) | (0.004) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| *Post Team Interaction* | -0.245 | -0.255 | -0.016 | -0.017 | -0.003 | -0.004 | | | | | | |
| *× Star in Team* | (0.076) | (0.074) | (0.007) | (0.008) | (0.007) | (0.007) | | | | | | |
| | [0.001] | [0.001] | [0.035] | [0.026] | [0.691] | [0.617] | | | | | | |
| *Post Team Interaction* | | | | | | | -0.371 | | -0.040 | | -0.029 | |
| *× Past Star in Team* | | | | | | | (0.092) | | (0.009) | | (0.010) | |
| | | | | | | | [0.000] | | [0.000] | | [0.003] | |
| *Post Team Interaction* | | | | | | | | -0.017 | | 0.010 | | 0.019 |
| *× Contemporary Star in Team* | | | | | | | | (0.080) | | (0.009) | | (0.008) |
| | | | | | | | | [0.834] | | [0.228] | | [0.022] |
| Mean of DV | 0.587 | 0.666 | 0.235 | 0.246 | 0.277 | 0.290 | 0.666 | 0.666 | 0.246 | 0.246 | 0.290 | 0.290 |
| Sample | Full | CEM | Full | CEM | Full | CEM | CEM | CEM | CEM | CEM | CEM | CEM |
| Firm FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Team FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Inventor FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Team Formation Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-squared | 0.213 | 0.200 | 0.310 | 0.280 | 0.308 | 0.291 | 0.201 | 0.199 | 0.281 | 0.280 | 0.291 | 0.291 |
| Firm Count | 339 | 283 | 339 | 283 | 339 | 283 | 283 | 283 | 283 | 283 | 283 | 283 |
| Inventor Count | 15,209 | 6,254 | 15,209 | 6,254 | 15,209 | 6,254 | 6,254 | 6,254 | 6,254 | 6,254 | 6,254 | 6,254 |
| Team Count | 22,483 | 9,800 | 22,483 | 9,800 | 22,483 | 9,800 | 9,800 | 9,800 | 9,800 | 9,800 | 9,800 | 9,800 |
| Team-Inventor Count | 66,532 | 15,634 | 66,532 | 15,634 | 66,532 | 15,634 | 15,634 | 15,634 | 15,634 | 15,634 | 15,634 | 15,634 |
| Observations | 1,032,695 | 251,286 | 1,032,695 | 251,286 | 1,032,695 | 251,286 | 251,286 | 251,286 | 251,286 | 251,286 | 251,286 | 251,286 |

**Appendix 1: Analytical Model**

We present a simple analytical model to illustrate how non-stars' knowledge homogenized toward traditional knowledge domains (as would be the case with emulating star inventors' knowledge) becomes suboptimal. We define non-star inventors' innovation quality as a function of (a) knowledge inputs between traditional versus non-traditional knowledge and (b) time in the industry represents the maturity of each knowledge domain. We make the following assumptions: (1) both knowledge domain and time in the industry have a positive relationship with performance, with decreasing returns, and (2) traditional knowledge emerges earlier than non-traditional knowledge. In this framework, we show that under the performance maximization condition, non-stars' overreliance on traditional knowledge becomes suboptimal if (a) traditional knowledge obsolescence is severe or (b) knowledge misfit between the two knowledge inputs is severe.

We define non-stars' innovation performance function as follows[9]:

$$P = \underbrace{\frac{\alpha}{m}\ln(tx)}_{Traditional} + \underbrace{\beta\ln\left(t(N-x)\right)}_{Non-traditional}$$

where $P$: innovation performance (i.e., quality), $x$: the amount of traditional knowledge in the total knowledge mix $N$ ($1 \leq x \leq N$), $t$: time in the industry ($t > 0$), $\alpha(\beta)$: productivity coefficient of traditional (non-traditional) knowledge ($\alpha, \beta \geq 1$), $m$: discount factor of $\alpha$ due to knowledge misfit in traditional knowledge against non-traditional knowledge ($m > 1$).

The first-order condition for performance maximization by knowledge input is the following[10]:

[9] We choose a natural logarithmic function because (a) it satisfies the assumption that knowledge input and time have a positive relationship with innovation performance, with decreasing returns, and (b) both the knowledge input and time in the industry are non-zero positive values allowing differentiability in the given domain. The implications from this model can be applied to any functional forms following these conditions without loss of generality.

[10] Checking the second-order condition is unnecessary because we assume an increasing function with non-zero positive inputs.

$$
\begin{aligned}
&\frac{\partial P}{\partial x} = \frac{\alpha t}{m(tx)} - \frac{\beta t}{t(N-x)} = 0 \\
&\Rightarrow \alpha t(N-x) = \beta m(tx) \\
&\Rightarrow x^* = \frac{\alpha N t - \beta m}{\beta m t + \alpha t}
\end{aligned}
$$

Under this maximization condition, we further investigate when the optimal knowledge mix represents a greater portion of traditional knowledge:

$$
\begin{aligned}
&x^* > \frac{N}{2} \\
&\Leftrightarrow \frac{\alpha N t - \beta m}{\beta m t + \alpha t} > \frac{N}{2} \\
&\Leftrightarrow \alpha N t > \beta(mNt + 2m) \\
&\Leftrightarrow \alpha > \underbrace{\left(\frac{mNt + 2m}{Nt}\right)}_{Friction} \beta
\end{aligned}
$$

This inequality suggests that the optional condition holds if traditional knowledge has a higher productivity coefficient compared to non-traditional knowledge in the absence of the friction term (i.e., $\alpha > \beta$). The friction term provides the key condition under which the inequality does not hold (i.e., non-stars' overreliance on traditional knowledge leads to suboptimal performance). The inequality is violated if $t$ or $m$ is sufficiently high. When sufficient technological advances are made in both types of knowledge (i.e., $t$ is high), overreliance on traditional knowledge becomes suboptimal due to amplified decreasing returns in traditional knowledge. The optimal condition is also not met if the knowledge misfit problem is severe (i.e., $m$ is high) because it dampens the effective use of traditional knowledge. Thus, these conditions explain why the problems of knowledge obsolescence and misfit may lead to undermined innovation quality of non-stars homogenized toward traditional knowledge.

## Appendix 2: Further Data Details

### A2.1. Pitchbook data

Among 113,882 unique venture-backed firms founded until 2021, the first fabless screening based on the four keywords leaves 1,950 firms. For the screened-in firms having at least one keyword match in their business descriptions, we manually investigate whether the descriptions correctly refer to the characteristics of fabless firms that specialize in chip design. Common false positives are suppliers or facilitators of semiconductor-related materials or products, e.g., "provider of semiconductor packaging technologies," "manufacturer of electronic components using nitride semiconductor materials," and "provider of optical devices for semiconductor-based lamps." These firms should not be considered fabless startups focusing on chip design as their business is closer to the production of chips or semiconductor-related products. This manual verification process results in 638 venture-backed fabless startups.

### A2.2. USPTO data

We only include granted utility patents because those patents pass the minimum quality threshold determined by the USPTO, contributing to technological innovation. Regarding the timing of patenting outcomes, we use the application years of the granted patents (not granted years) to measure the date of knowledge creation accurately, as is customary in the literature.

For the matching between the Pitchbook and USPTO datasets, we use Stata modules to (a) preprocess raw names to address inconsistencies in firm names in the two data sources (i.e., "stnd_compname"), (b) conduct probabilistic linking to calculate match scores based on the proximity between the preprocessed name components (i.e., "reclink2"), and (c) manually review potential matches that pass a certain threshold when they are not perfect matches (i.e., "clrevmatch"). We set the threshold to a 0.95 match score out of 1. The fuzzy matching results in

482 correct matches between the two data sources. There are 37 firms in which the earliest team patent application year precedes the founding year in the Pitchbook data (mostly one year ahead of the founded year). For our purposes of investigating inventor interactions, we replace those founding years with the earliest team inventing year in the USPTO data. For the matched firms, we only include inventors who have been associated with at least one inventing team in the sample based on co-patenting. Team size is limited to 10, excluding 144 out of 27,727 teams (0.005%).

**A2.3. Variables**

For each inventor dyad in H1, we first consider up to 10 years before and after the team interaction. Within this window, the earliest pre-team year is when at least one inventor starts having a patenting outcome (if it is later than t-10), and the latest post-team year is when at least one inventor continues issuing a patent (if it is earlier than t+10). We fill in missing observations with zeros within the final pre- and post-team year window.

To calculate *Knowledge Similarity*, we use level-2 CPC codes (e.g., H01: electric elements), constructing a vector space with dimension = 129. Using more granular levels could denote inventors' knowledge domains more precisely (as used in Section 4.2), but the dimensionality becomes too high, making the similarity between inventors nearly zero in most cases. For example, level-3 CPC codes would entail constructing a vector space with dimension = 670.

For each inventor to test H2, we first consider up to 10 years before and after the team interaction. Within this window, the minimum pre-team year is when a focal inventor starts issuing a patent (if it is later than t-10), whereas the maximum post-team year is five years after the last patent is issued (if it is earlier than t+10). We consider the five-year buffer in the post-team period because it is more likely that the inventor could not achieve inventions after the final patent application (not because the inventor completely retires from any patenting work). We fill in missing observations with zeros within the final pre- and post-team year window.

### A2.4. Sample restrictions

The dyad-year-level dataset for H1 (Panel A in Table 1) includes teams with co-patenting years up to 2015. Given that the most recent and “complete” patent data is around 2022, this cutoff provides sufficient post-team observations for the teams patented until 2015. The cutoff needs to be shortened in the inventor-year-level dataset (i.e., Panel B in Table 1) because patent citation data suffers more from the truncation problem in recent years, particularly after 2020. To alleviate the imbalance problem, we decrease the cutoff to 2012 in Panel B.

## Appendix 3: Additional Empirical Analyses and Discussions

### A3.1. Endogenous selection in inventing teams

We discuss two major selection issues that may confound our main results in H2 and explain how our fixed effects can mitigate the biases. The first selection is from the firm side. There could be firms with or without star inventors in the first place. If this firm selection is significant, our counterfactual inventors (non-stars who do not interact with stars) may have lower innovation performance because they are in firms with limited resources or capabilities. This scenario implies downward biases in our estimates because the observed innovation quality loss of non-star inventors in the treatment group is compensated by these inventors' relatively higher performance (or increasing trends) compared to the counterfactual inventors. Firm fixed effects in our model address these fundamental differences between firms with or without star inventors.

The second selection is from the team side. Firms may assign their star inventors disproportionately to the inventing teams. One could argue that firms strategically match their star inventors and high-performing non-star inventors in the same team with the hope that the non-stars learn from the stars and keep improving their innovation performance. In this case, this team selection also causes an underestimation problem due to the same reasoning explained above. To alleviate this issue, we include team fixed effects that rule out the underlying differences between teams with or without stars.

As we discussed in Section 3.5, we further apply CEM to resolve remaining confounders, and the CEM-adjusted results are consistent with the underestimation scenario. Column (2) in Table 3 shows the estimated coefficient becomes larger in magnitude with smaller standard errors compared to Column (1), meaning having a more comparable set of treatment versus control groups may reduce the underestimation problem. We find the same results in Column (3) and

Column (4) where the dependent variable is *Generality*. Therefore, although our efforts may not fully overcome the endogeneity problem due to the lack of exogenous variation, we believe our estimates provide reasonable lower bounds of the true effects.

**A3.2. Parallel trends in the temporal DID estimates**

The core assumption of the DID setup is the parallel trend between the treatment and control groups. To validate this assumption, we plot the temporal DID estimates of our main models (Column (8) in Table 2 for H1 and Column (2) in Table 3 for H2). Years between t-4 and t+10 are included in the temporal regression to ensure sufficient observations at each time point. Figure A1 visualizes the temporal DID estimates for *Knowledge Similarity* (Panel A) and *Scaled Citation* (i.e., Panel B). Dots represent point estimates, and capped lines denote 95% confidence intervals. In Panel B, we further compare the confidence intervals from the full sample (i.e., shaded area) and CEM sample (i.e., dots and capped lines). There is no pre-trend in both cases, and our F-tests do not reject the null hypothesis that estimates of t-1 to t-4 are jointly zero (p-value: 0.529 for *Knowledge Similarity*; p-value: 0.346 for *Scaled Citation*).

**A3.3. Knowledge domains of star versus novice inventors**

To identify inventors' knowledge domains, we use level-3 CPC codes (e.g., H01L21). Using level-3 codes helps us differentiate knowledge domains between star versus novice inventors more clearly, as compared to level-2 codes used in calculating *Knowledge Similarity*. We exclude indexing schemes in the ranking because these codes' definitions are too broad to provide specific information about different knowledge domains. Table A1 shows the top 10 knowledge domains of stars (joined before 2000) and novice inventors (joined after 2010). It shows the two groups of inventors had very different knowledge domains before they joined fabless firms.

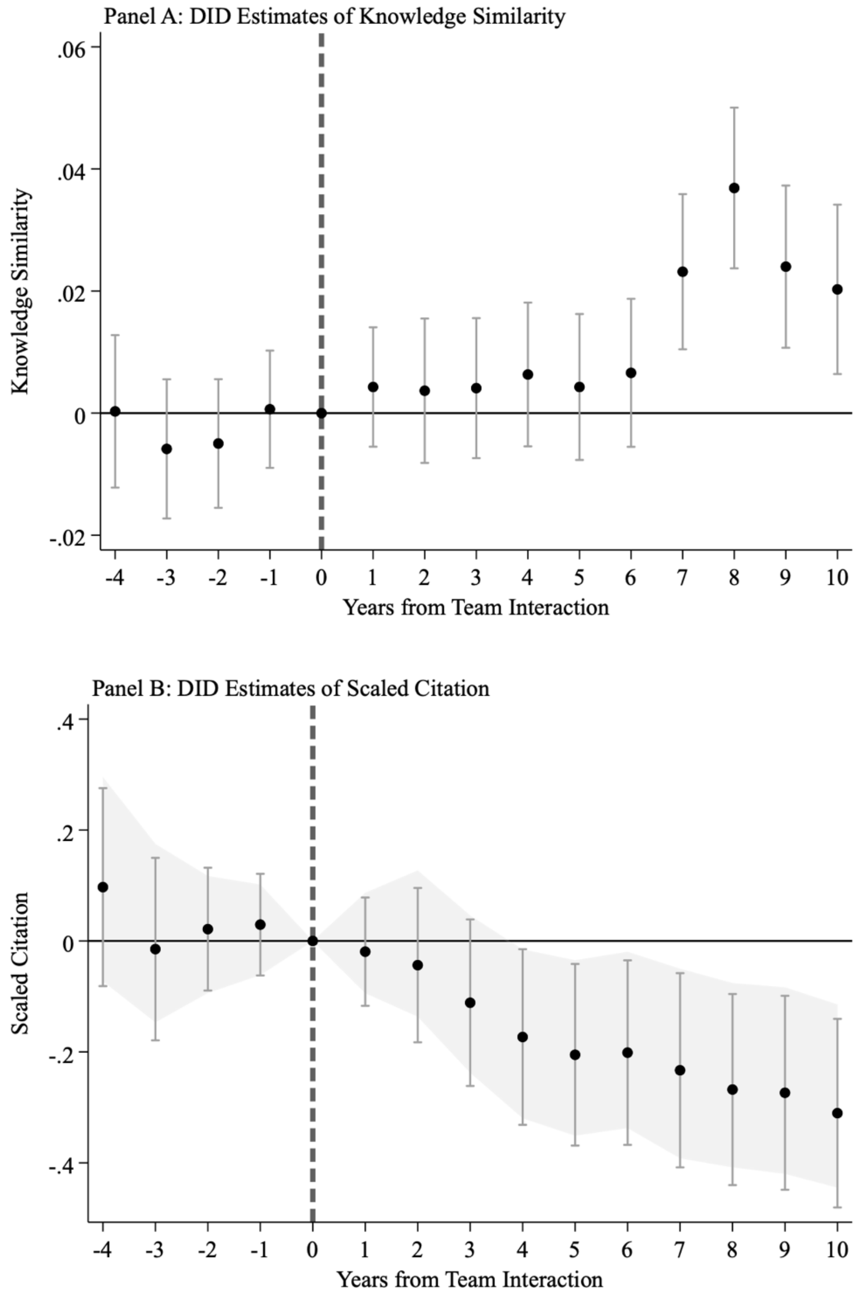


**FIGURE A1: Temporal Effects of Team Interaction with Stars on Knowledge Homogenization and Innovation Performance.** Panel A shows the year-by-year DID estimates of team interaction with stars on knowledge similarity between inventors measured by *Knowledge Similarity* (Hypothesis 1). Black dots indicate the estimates, and gray capped lines represent 95% confidence intervals. Panel B shows the DID estimates of team interaction with stars on non-star inventors' innovation performance measured by *Scaled Citation* (Hypothesis 2). The shaded area in Panel B shows confidence intervals in the full sample without matching, whereas the dots and lines represent the estimations based on the matched sample.

**TABLE A1: List of Knowledge Domains by Star and Novice Inventors.**

| | Star Inventors (joined fabless before 2000) | | Novice Inventors (joined fabless after 2010) | |
|---|---|---|---|---|
| Rank | CPC Code | Description | CPC Code | Description |
| 1 | H01L21 | Processes or apparatus adapted for the manufacture or treatment of semiconductor or solid state devices or of parts thereof | G06F3 | Input arrangements for transferring data to be processed into a form capable of being handled by the computer; Output arrangements for transferring data from processing unit to output unit, e.g. interface arrangements |
| 2 | H01L29 | Semiconductor devices adapted for rectifying, amplifying, oscillating or switching, or capacitors or resistors with at least one potential-jump barrier or surface barrier | G06F9 | Arrangements for program control, e.g. control units |
| 3 | H01L27 | Devices consisting of a plurality of semiconductor or other solid-state components formed in or on a common substrate | G06F30 | Computer-aided design [CAD] |
| 4 | H01L23 | Details of semiconductor or other solid state devices | G06F13 | Interconnection of, or transfer of information or other signals between, memories, input/output devices or central processing units |
| 5 | H03K19 | Logic circuits, i.e. having at least two inputs acting on one output; Inverting circuits | G06F2119 | Details relating to the type or aim of the analysis or the optimisation |
| 6 | Y10S148 | Metal treatment | G06F2203 | Interaction techniques based on graphical user interfaces [GUI] |
| 7 | Y10S438 | Semiconductor device manufacturing: process | G06F8 | Arrangements for software engineering |
| 8 | G11C11 | Digital stores characterised by the use of particular electric or magnetic storage elements; Storage elements therefor | H01L23 | Details of semiconductor or other solid state devices |
| 9 | H03K3 | Circuits for generating electric pulses; Monostable, bistable or multistable circuits | H01L24 | Arrangements for connecting or disconnecting semiconductor or solid-state bodies; Methods or apparatus related thereto |
| 10 | G11C7 | Arrangements for writing information into, or reading information out from, a digital store | G06F3 | Input arrangements for transferring data to be processed into a form capable of being handled by the computer; Output arrangements for transferring data from processing unit to output unit, e.g. interface arrangements |