Working Paper 24-021

Bringing Science to Market: Knowledge Foundations, Inventor-Founders, and Performance

Justine Boudou
Maria Roche

Harvard Business School

Bringing Science to Market: Knowledge Foundations, Inventor-Founders, and Performance

Justine Boudou
Harvard Business School

Maria Roche
Harvard Business School

Working Paper 24-021

Funding for this research was provided in part by Harvard Business School.

Bringing Science to Market: Knowledge Foundations, Inventor-Founders, and Performance*

Justine Boudou
Harvard Business School

Maria Roche
Harvard Business School

February 17, 2025

Abstract

In this paper, we examine how a startup's knowledge foundations – embedded in its core technology – influence its performance in the exit market. Using a dataset of 1,006 biomedicine startups founded between 2005 and 2015, we focus on two key factors: (1) the degree of scientific specialization in the startup's core technology and (2) whether the technology's inventor is also the startup's founder. Counterintuitively, we find that greater scientific specialization in a startup's knowledge-base correlates with poorer exit market outcomes. Additional analyses suggest that this stems from such startups relying on narrower, less integrable technologies heavily dependent on tacit knowledge, which can hinder engagement with external stakeholders. However, the presence of an inventor-founder—an individual who invents the core technology and establishes the startup—moderates this relationship. When an inventor-founder is involved, the negative relationship with knowledge specialization is almost entirely mitigated. This suggests that inventor-founders may enhance the strategic value of specialized knowledge by making it more accessible to key stakeholders while also reinforcing its defensibility, counterbalancing its associated challenges. Interestingly, we also find that startups with an inventor-founder but without a specialized knowledge-base, or vice versa, perform worse on the exit market. These findings underscore the contingent value of knowledge-based resources in entrepreneurial contexts, emphasizing the importance of aligning knowledge characteristics with the founder to optimize firm outcomes. Our research highlights the nuanced relationship between knowledge specialization, founder roles, and startup performance, contributing to a deeper understanding of how knowledge-based resources shape firm success.

Keywords: *Firm Performance*, *Exits*, *Knowledge Foundations*, *Inventor-Founder*, *Specialized Scientific Knowledge*

*The authors thank Annamaria Conti for data access. We thank Tim Holzner and Richie Zitomer for excellent research assistance. Many thanks to participants at the Barcelona School of Economics Entrepreneurship Working Group, DRUID and the Productivity Seminar. We are indebted to Ashish Arora, George Chondrakis, Annamaria Conti, Shane Greenstein, Jacqueline Lane, Brian Silverman, and Keyvan Vakili for feedback on early drafts. We are grateful for the valuable insights provided by several technology transfer offices and entrepreneurial faculty. We gratefully acknowledge funding from the Harvard Business School Division of Research and Faculty Development. All errors and omissions are our own. Contact: jboudou@hbs.edu; mroche@hbs.edu

1 Introduction

Possessing and leveraging unique resources has long been considered a critical component in achieving competitive advantage in the marketplace (Barney, 1991). In particular, firms that are able to harness unique, complex and tacit knowledge, especially difficult to replicate or transfer, often benefit from a sustained advantage over competitors (Grant, 1996). Despite the considerable body of research that examines the role of knowledge in firms' creation (Agarwal and Shah, 2014; Botelho et al., 2021; Chattopadhyay et al., 2024; Ganco, 2013; Grant, 1996; Park et al., 2024; Shane and Venkataraman, 2000), little is known about how building a new venture around specialized scientific knowledge may shape long-term performance outcomes. This is important to understand because the impact of unique resources in entrepreneurship may differ from that in established firms, as startups' success critically depends on their ability to engage with and appeal to external stakeholders. Moreover, the literature remains largely silent on the distinction between the individual who develops an invention and the individual who seeks to commercialize it – specifically, whether the inventor of a technology is also the founder of the startup that aims to bring it to market – which may influence the tacitness and transferability of the knowledge in question. This distinction is particularly salient in the context of startups, where the nature of the knowledge-base and the identity of the individual holding it can be more easily disentangled than in established firms, providing a unique lens to study how these elements jointly shape firm performance.

In this paper, we ask (1) how the extent to which a technology is built upon specialized scientific knowledge – advanced, in-depth, individual-specific expertise in a particular scientific field – influences subsequent startup's performance outcomes, and (2) how this relationship varies depending on whether the inventor of the technology is also the founder of the startup. Specifically, we examine the impact of two key decisions founders make when starting a company.[1] First, they can choose to build a company around an invention that heavily relies on the specialized scientific knowledge of its inventor or draws more broadly from other scientific sources of knowledge, thereby shaping the characteristics of their startup's knowledge-base. Second, they can decide whether to commercialize their own invention or that of another inventor. In doing so, we shed light on the relationship between specialized scientific knowledge and subsequent firm performance

[1] For the purpose of this paper, we are agnostic to the order.

outcomes and how the presence of an "inventor-founder" – which we define as an individual who both creates a novel invention and establishes a venture to commercialize that invention – moderates this relationship.

Generally, we may expect that a high degree of scientific specialization within a venture's knowledge-base would foster better performance outcomes by providing unique knowledge and technical expertise, thus, creating critical "isolating mechanisms" (Rumelt et al. 1984, p. 568) necessary to achieve and sustain competitive advantage. However, in the context of new venture formation, it is possible that these advantages – many of which have been identified in the context of established firms – may be offset by specific challenges, shifting the net relationship into the red. Some of these frictions may stem from external stakeholders' – such as investors and acquirers – struggles to fully understand, evaluate, or integrate the specialized and tacit knowledge embedded in technologies that heavily rely on the scientific expertise of their inventors (Balachandran, 2024; Eckhardt and Shane, 2003; Makri et al., 2010; Polanyi, 2009; Puranam et al., 2009; Zaheer et al., 2010). Given such stakeholders' importance for achieving performance milestones, the capacity to adapt and effectively engage with external parties may be more critical for startups' success than having a unique knowledge-base. Therefore, it is feasible that the relationship between the extent to which a technology builds on specialized scientific knowledge and subsequent startup performance varies depending on the relative impact of the advantages and challenges that knowledge specialization brings.[2]

Beyond the characteristics of the knowledge-base itself, the individual who seeks to bring the invention to market may be a further crucial determining factor for long-run success. In particular, when the inventor of the core technology is also the founder of the startup, the knowledge may become more accessible and less portable. Indeed, the inventor-founder's direct involvement may provide privileged access to deeper expertise and tacit components of the knowledge, ensuring it can be fully leveraged (Aghasi et al., 2022). At the same time, this knowledge becomes more closely tied to the founder, enhancing its value as a unique, hard-to-replicate resource (Barney, 1991; Wernerfelt, 1984). Conversely, if the founder of the firm is not the inventor – such as when the invention originates from an external inventor or an employee – the knowledge may be more flexible and expandable but also harder to access and protect.

[2]See Table A1 for a summary of the arguments.

From this, we propose that introducing a key factor, namely, whether the founder is also the inventor of the startup's underlying technology – an inventor-founder[3] – is critical in understanding the relationship between the degree of scientific specialization of the knowledge-base and performance outcomes. Taken together, we propose a framework that distinguishes between the degree of scientific specialization of a startup's technology and the individual who aims to commercialize this technology. We argue that whether the inventor of a startup's core technology is also in charge of commercializing it or not might moderate the relationship between the scientific specialization of the knowledge-base and startup performance on the exit market, making the joint consideration of these two factors crucial to fully understand their relationship with startup performance.

To empirically assess these questions, we construct a unique dataset using Crunchbase and the Reliance on Science (RoS) dataset (Marx and Fuegi, 2020), which we combine with a variety of other sources such as LinkedIn and Bloomberg. Our final sample consists of 1,006 startups in the bio-medicine sector founded between 2005 and 2015.[4] We choose this sector as it is particularly well suited for our study given the tight link between the product-market and scientific research and the relatively high propensity to patent. This is critical in our approach since it enables us to proxy the initial knowledge-base of a startup, which should at least be partially codified in a new venture's granted patents. To quantify the extent to which startups are built around specialized scientific knowledge, we introduce a novel measure based on patents' citations to the scientific work of their inventor. This measure reflects the degree to which a startup's technology is directly tied to the scientific expertise of the inventor. We then investigate how this measure is associated with performance outcomes on the exit market.

Our analyses reveal notable findings regarding the impact of scientific specialization on startup performance. Surprisingly, we uncover that startups with a higher reliance on specialized scientific knowledge are less likely to be successful on the exit market. Specifically, a 10 percentage-point increase in the extent to which a startup's core technology builds on the inventor's own scientific work is associated with a 4.5% decrease in the likelihood of getting acquired or going public, representing

[3]For simplicity, we use the terms 'inventor' and 'founder'. However, we recognize that an invention may have multiple inventors and a startup could have several founders. Similarly, we use the word 'inventor-founder' but it is possible that multiple inventors are also founders, or that only one of the inventors is a founder. These nuances do not conceptually alter our discussion.

[4]Our dataset only contains startups that *own* their patents. In particular, we exclude firms building on licensed technologies to ensure our sample is comparable.

a 26% decrease with respect to the average exit rate. This result is consistent with the idea that specialized scientific knowledge may introduce frictions in engaging with external stakeholders that outweigh the benefits of possessing a unique, hard-to-replicate resource, ultimately hampering performance on the exit market.

In order to provide a clearer understanding of the potential mechanisms behind these results and the types of frictions that may emerge, we consider further characteristics of a startup's core knowledge foundations as observed in their patents. Here, our evidence suggests that a startup whose knowledge-base relies more on specialized scientific knowledge also builds on research that is narrower and less cited by other patents.[5] In addition, we find that these startups' patents have less claims and are associated with a smaller team of inventors. This is consistent with the idea that inventions relying more on specialized scientific knowledge are narrower, harder to integrate into broader technological solutions, and more reliant on tacit knowledge. This might limit founders' ability to develop solutions that resonate with market needs and negatively impact their startup's overall performance on the exit market.

Next, we examine how the relationship between specialized scientific knowledge foundations and new venture performance varies based on whether the founder of the startup is also the inventor of the technology. To do so, we track and measure the extent to which a startup's patents emanate from the founder, or from other inventors. We find that on average, having an inventor-founder is negatively associated with performance on the exit market and only weakly mediates the relationship between specialized scientific knowledge and performance. However, our examination of the interaction coefficients indicates that the inventor-founder role appears to moderate the relationship between specialized scientific knowledge and exit outcomes. Specifically, we find that when the startup's knowledge-base is highly scientifically specialized *and* the inventor is a founder, our baseline negative relationship is almost entirely mitigated (though the combined impact remains negative).

Overall, we interpret our results such that the inventor-founder's intimate understanding of the specialized scientific knowledge may help ensure this knowledge can be more effectively leveraged, and integrated. At the same time, the presence of the inventor-founder reduces the mobility of the specialized scientific knowledge, ensuring it remains uniquely tied to the venture and can be protected. Hence, acquirers and public investors gain access to an otherwise hard-to-utilize resource,

[5] We proxy for research breadth by using the number of unique 'concepts' a paper entails.

unlocking its value while reducing its mobility. In essence, the combination of specialized scientific knowledge (the 'horse') and the inventor-founder (the 'jockey') seems to transform what could otherwise be a severe liability into a potential strategic asset (Conti et al., 2024).

Taken together, this paper makes at least three contributions. First, we shed light on the complexities of knowledge-based advantages in new firm formation, paying particular attention to the role of specialized scientific knowledge and how it may influence critical performance outcomes. Second, we introduce a new measure that allows us to quantify how much a startup's knowledge-base relies on the specialized scientific knowledge of the inventor behind its core technology. To that end, we analyze startups' patents and their citations to scientific literature to characterize a startup's knowledge-base and specifically link them to their degree of reliance on their inventors' scientific work. Third, we distinguish between the roles of inventor and founder, providing empirical evidence suggesting that this distinction is critical in determining the role of the portability of knowledge for achieving performance milestones. Specifically, by studying whether the involvement of the inventor as the founder can exacerbate or mitigate the difficulties of achieving a successful exit event, we introduce crucial inventor-founder interactions.

From a managerial and strategic perspective, our results offer important insights for startups and their stakeholders. For founders, the findings underscore the need to carefully consider how their venture's reliance on highly specialized knowledge may hinder their ability to successfully navigate their performance on the exit market (Makri et al., 2010; Polidoro and Yang, 2021; Puranam et al., 2009; Stuart et al., 1999). To mitigate these challenges, potential strategies might involve efforts to codify specialized knowledge, making it easier to transfer and understand by external stakeholders (Nonaka and Von Krogh, 2009; Zander and Kogut, 1995). This may involve a more targeted communication strategy, a role that technology transfer offices (TTO) can potentially fulfill (Shah and Pahnke, 2014). Our findings also suggest that the identity of the knowledge-holder is a critical factor to consider (Campbell et al., 2012). In particular, leveraging one's own invention does not always seem to confer an advantage. Rather, our findings suggest that when it comes to specialized scientific knowledge, it helps alleviate concerns from potential acquirers and public investors only when the invention is deeply rooted in the inventor's own scientific knowledge. Importantly, our findings suggest that it is crucial for founders to ensure that their focus on advancing specialized knowledge aligns with their intended exit strategy and the priorities of key stakeholders they aim to engage, recognizing that different stakeholders may be needed at various stages of the venture's life. By balancing specialized expertise with market-driven considerations, startups potentially enhance the relevance and scalability of their technologies, making them more appealing to investors, acquirers, and the broader market.

For strategy scholars and practitioners, our results suggest that while specialized knowledge can create competitive differentiation (Barney, 1991; Conner and Prahalad, 1996; Grant, 1996; Leahey, 2007; Teece, 1986), it may also introduce significant barriers to growth, integration, and alignment with external stakeholders, especially in the context of new firm formation, thus requiring a more nuanced approach to leveraging knowledge-based advantages in ventures. Highly specialized scientific knowledge on its own does not appear to be a panacea for achieving competitive advantage. Finally, while our study cannot causally establish the role of specialized knowledge or an inventor-founder in startup success because clear unobservable factors are at play due to selection mechanisms – e.g., founders who commercialize a technology that is not their own may be more actively aware of specific market needs, whereas those leveraging specialized knowledge may face reduced technological uncertainty but lack market awareness – our results provide important insights into which types of startups are more or less likely to succeed and some of the key factors influencing this outcome.

2 Background and Hypotheses

A notable body of work has made fundamental strides in understanding the impact of capability differentials in shaping both the decision to become an entrepreneur and the entry mode into entrepreneurship (Agarwal et al., 2004; Agarwal and Shah, 2014; Åstebro et al., 2011; Hoang and Gimeno, 2010; Lee et al., 2024; Roach, 2017; Roach and Sauermann, 2015; Shah et al., 2019; Sørensen, 2007; Stenard and Sauermann, 2016). However, a gap remains in understanding how specific types of initial knowledge endowments influence subsequent startup performance, particularly when it comes to exit outcomes such as acquisitions and IPOs. While the literature has extensively documented the broad distinction between utilizing and not utilizing scientific knowledge, less is known about the ways in which the *nature* of scientific knowledge used can impact entrepreneurial outcomes.

Throughout this paper, we think of startups as a vehicle for commercializing a core technology. We conceptualize a startup's knowledge-base as the body of scientific and technical knowledge that informs its core technology. In this section, we discuss how variations in the extent to which startups' knowledge-base relies on *specialized scientific knowledge* may influence their performance outcomes. We thereby conceptualize specialized scientific knowledge as advanced, in-depth, individual-specific expertise in a particular scientific field, developed through focused research and study. Furthermore, we examine how this relationship is influenced by whether the inventor of the core technology is also the founder of the startup. We do so because the individual who embodies the knowledge at the core of the startup is critical for determining how transferable, scalable, and understandable that knowledge is to external stakeholders (Becker, 1964). Figure A1 summarizes these main conceptual elements.

By shifting our focus from a binary perspective to exploring the *degree* to which a startup's knowledge-base relies on specialized scientific expertise — and whether that expertise emanates from the inventor-founder — we provide a more nuanced understanding of how specialized resources, in our context specialized scientific knowledge, contributes to firm success. More insight into these nuances is critical for founders, and investors as adjustments regarding the knowledge-base and founder-team[6] are feasible to make and small changes could have far-reaching implications.

2.1 Knowledge Endowments and Firm Performance

The long established resource-based view of the firm perceives the firm as a unique bundle of idiosyncratic resources, with distinctive competencies and heterogeneous capabilities that drive competitive advantage (Barney, 1991; Penrose, 1959; Wernerfelt, 1984). Resources play a central role in shaping not only established firms', but also new ventures' strategic outcomes. Indeed, entrepreneurial opportunities are presumed to exist primarily because of differences in beliefs about the relative value of resources (Schumpeter, 1912; Shane, 2000). In particular, new ventures are often formed to commercialize a specific technology, where initial resource endowments, such as social capital, technical knowledge, or intellectual property, can serve as critical foundations for long-term performance (Dencker et al., 2009; Shane and Stuart, 2002). Among those, knowledge stands out as one of the most strategically important assets a firm can possess as it often involves a tacit component making, e.g., imitation by competitors more difficult (Agarwal and Shah, 2014; Grant, 1996; McEvily and Chakravarthy, 2002; Spender and Grant, 1996). In fact, in a recent

[6] Note that there may be team dynamics that play a critical role in achieving certain performance milestones. For the purpose of argument, we will not go into detail on these in this paper.

meta-analysis, Bergh et al. (2024) find that knowledge resources are the most important strategic components for strong stock market, financial and growth performance.

Within this broad category, scientific knowledge represents a particularly valuable subset, and several prominent examples highlight the crucial role of scientific knowledge in driving innovations that enhance firm performance. For instance, the groundbreaking work by Jennifer Doudna and Emmanuelle Charpentier on CRISPR-Cas9 gene-editing technology revolutionized biotechnology and led to the success of firms like Editas Medicine and CRISPR Therapeutics. Similarly, the development of the PageRank algorithm by Larry Page and Sergey Brin, based on scientific principles from computer science and graph theory, led to the creation of Google's PageRank algorithm which transformed search engine technology and propelled Google to become a leading tech giant. These examples are far from being anecdotal, and the literature finds consistent evidence of higher social and private values derived from leveraging scientific knowledge (Baruffaldi and Poege, 2024). For instance, high-technology firms that foster a research-oriented culture by supporting independent research, attending conferences, or encouraging publications, tend to perform better than those that do not (Henderson and Cockburn, 1994). Furthermore, patents that incorporate scientific content have been shown to generate more follow-on citations (Ahmadpoor and Jones, 2017; Fleming and Sorenson, 2004; Sorenson and Fleming, 2004). This is partly because the norms of science, such as openness and publication, accelerate the diffusion of knowledge (Sorenson and Fleming, 2004), and because scientific research aids inventors in identifying better combinations of knowledge components (Fleming and Sorenson, 2004). Moreover, scientific knowledge enhances absorptive capacity, enabling firms to better assimilate and utilize new information (Cohen and Levinthal, 1990). Krieger et al. (2024) also show that patents which build on scientific knowledge have higher private value, potentially due to their greater novelty. Additionally, patents that rely on science are more likely to be traded, as their scientific grounding is believed to facilitate codification, reduce search costs for buyers, and improve the evaluation and integration of inventions (Arora et al., 2022).

While the distinction between scientific and non-scientific knowledge endowments is well-documented, the nature of scientific knowledge used in an invention can vary significantly, potentially bearing important consequences. To examine this further, we turn our focus to the degree of scientific specialization of an invention.

2.2 The Double-Edged Sword of Specialized Knowledge in Entrepreneurship

We conceptualize the *specialized* scientific knowledge of an individual as the scientific knowledge and expertise they have accumulated through their *own* peer reviewed published research in a specific scientific domain. When individuals engage in scientific research, they delve into novel and unexplored aspects of a field, making them the expert of a specific topic. In doing so, they ultimately develop a deep understanding of a particular area of science (Jones, 2009). This understanding is based on rigorously tested and peer-reviewed work derived from systematic research and empirical evidence, typically emanating from academia (Dasgupta and David, 1994).

An invention can vary in the degree to which it leverages the scientific expertise of its inventor, thereby shaping the characteristics of the startup's knowledge-base. Inventions that heavily rely on the inventor's specialized scientific knowledge draw extensively on their prior research. Any product or technology commercialized by startups based on that invention will be directly informed by the inventor's scientific work, making the startup's knowledge-base highly scientifically specialized. In contrast, inventions that minimally leverage the inventor's specialized scientific knowledge may incorporate some insights from their research but will also rely on scientific knowledge from other sources. As a result, products or technologies developed by startups will be less tightly connected to the inventor's research. In this case, we will consider the startup's knowledge-base to be less scientifically specialized.

At first, inventions that rely on their inventor's specialized scientific knowledge may be expected to have significant advantages, particularly in terms of innovation potential, technical expertise, and competitive differentiation. Indeed, specialized knowledge embedded within the invention might provide deep insights into a specific domain, making the technology more effective in addressing complex technical challenges and driving innovation (Kaplan and Vakili, 2015; Leahey, 2007), especially in faster-paced environments (Teodoridis et al., 2019). Moreover, the direct application of this specialized knowledge in the technology development process can lead to more efficient problem-solving, as the technology reflects a high level of familiarity with the intricacies and challenges of the domain (Nagle, 2018). In addition, a strong foundation of specialized knowledge may signal technical credibility and expertise to external stakeholders, such as investors or strategic partners, increasing the attractiveness of the startup for collaboration and funding opportunities (Stuart, 2000). Inventions built on specialized scientific knowledge may also be better equipped to navigate potential risks and roadblocks throughout the development process, as the inventor's deep expertise enables more accurate anticipation and assessment of challenges (Fleming, 2001; Roche et al., 2020).

Importantly, firms that are able to harness rare and valuable resources, such as specialized knowledge, have been suggested to be better able to build a sustainable competitive advantage (Barney, 1991). In particular, specialized scientific knowledge could serve as a unique and hard-to-replicate resource because it relies more heavily on the inventor's tacit knowledge, thereby allowing the startup to differentiate itself from competitors who lack access to this level of expertise (Collis and Montgomery, 2008; Teece, 1986). This advantage may be further solidified by the intellectual property (IP) associated with specialized scientific knowledge, as patents rooted in deep expertise tend to be more defensible and harder to replicate (Krieger et al., 2024; Teece, 1986). Additionally, specialized knowledge may help foster collaborations with institutions or firms seeking complementary expertise, which strengthens the startup's position in niche markets and provides access to critical resources (Zucker et al., 2002).

In the context of entrepreneurship, inventors' specialized knowledge may grant them the ability to "recognize the value of new, external information, assimilate it, and apply it to commercial ends [...]" (Cohen and Levinthal 1990, p.128). By providing more intimate knowledge about the scientific roots of a technology, the more extensive use of scientific specialized knowledge may make inventors better equipped to identify when and why a technology might succeed or fail (Kacperczyk and Younkin, 2017), potentially providing the startup that builds on their invention with a competitive advantage in terms of technical expertise, innovation potential and opportunity recognition. As a consequence, ventures whose knowledge-base relies more extensively on the scientific work of their technology's inventors may be better positioned to develop groundbreaking and unique innovations and capture market opportunities that are inaccessible to others.

However, the literature also highlights that specialization can be a double-edged sword (Peteraf, 1993). For example, Nagle and Teodoridis (2020) show that specialized researchers find it more difficult to integrate new external knowledge compared to those with broader expertise. Additionally, reliance on familiar knowledge may lead to diminishing returns and organizational rigidity (Dosi, 1988; Katila and Ahuja, 2002; Tripsas and Gavetti, 2017). This could limit a firm's ability to grow and adapt, ultimately affecting its broader success and long-term trajectory.

Moreover, since a venture success is often measured by exit events such as acquisitions or IPOs (Roche et al., 2020), this further suggests that a venture's success is critically dependent on its ability to engage with and appeal to external stakeholders (Garg et al., 2025; Gompers and Lerner, 2001; Stuart and Sorenson, 2007). Provided documented differences in institutional logics (Sauermann and Stephan, 2013), this introduces important nuances to our understanding of how specialized scientific knowledge influences venture outcomes. Given asymmetry of information and divergence in interests, the engagement of ventures with external stakeholders is already delicate to begin with (Jensen and Meckling, 2019; Junkunc and Eckhardt, 2009; Leland and Pyle, 1977). This may be exacerbated for ventures whose technology relies more extensively on specialized scientific knowledge. While scientific knowledge is often publicly available in the form of publications, inventions rooted in the inventor's own specialized expertise may depend on deeper, tacit elements that are not fully codified or readily understandable by stakeholders on the exit market (Polanyi, 2009; Zaheer et al., 2010). This is because simply disclosing scientific knowledge does not imply an understanding of the deep underlying mechanisms of the knowledge itself, which typically takes years to decades of work to obtain (Roche et al., 2020). By contrast, when inventors build on the research of others, they are more likely to leverage 'visible' or codified components of prior work, which are easier to communicate and transfer. Specialized scientific knowledge might also tend to correlate with smaller teams of inventors or even solo inventors (Jones, 2009), which limits opportunities for knowledge sharing, codification, and diffusion. While this reliance on tacit, individually held knowledge may make the invention more rare and unique, it might also make it difficult for external stakeholders to fully understand or assess the venture's value. This might increase the risk of adverse selection and moral hazard, ultimately hindering engagement of potential investors and acquirers. Moreover, this might also create communication and collaboration challenges between the venture and external stakeholders, potentially limiting access to crucial funding and partnership opportunities (Polidoro and Yang, 2021; Stuart et al., 1999).

Beyond appeal and engagement with external stakeholders, ventures that rely heavily on specialized scientific knowledge may face additional challenges related to the scalability and growth of their technology. First, inventions that draw heavily on the inventor's specialized scientific knowledge may be based more on familiarity rather than market potential (Klepper, 2001; Shane and Venkataraman, 2000). Indeed, specialized scientific knowledge is often deeply embedded in a specific theoretical or methodological framework, making it more difficult to apply beyond its original research context (Kogut and Zander, 1992). Combined with its more tacit elements, this may limit its adaptability, as it may not readily lend itself to modular adaptation or integration into complementary technological systems.

Second, inventions that draw heavily on the inventor's specialized scientific knowledge may result in a narrower knowledge-base for the startup, characterized by a less diverse set of scientific ideas or concepts from which the technology draws. This might restrict the number of ways the technology can be applied, making it less adaptable across different markets and limiting opportunities for technological recombination (Dosi, 1988; Katila and Ahuja, 2002). A narrower knowledge-base may also limit a venture's exit prospects because their technology may lack sufficient scientific and technical complementarities with others' existing operations, making integration more challenging and reducing the appeal for potential buyers or users (Makri et al., 2010). The higher reliance of a startup's knowledge-base on specialized scientific knowledge may also lead to post-acquisition challenges, as the acquiring firm faces difficulties in effectively integrating the knowledge assets of the target firm into its own operations, and face coordination challenges (Jain and Mitchell, 2022), especially in the case of technology's acquisitions (Puranam et al., 2009). This might make these ventures less attractive as an acquisition target or a business opportunity. The lack of diversity in expertise and perspectives can further inhibit the firm's ability to respond to evolving market demands and new technological developments, which is crucial for long-term growth.

Overall, the literature, especially the stream building on arguments for more established firms, suggests a positive relationship between specialized scientific knowledge and performance outcomes. Regarding startups, similar advantages may prevail on the exit market because it can provide them with a unique competitive advantage. However, this specialization can also introduce challenges, particularly in engaging with external stakeholders that are critical for startups' success on the exit market. This trade-off suggests that *the relationship between specialized scientific knowledge and startup performance depends on whether its advantages outweigh its challenges.* Appendix Table A1 summarizes the core theoretical arguments and the relationship with exit outcomes.

2.3 The moderating role of the inventor-founder

In this section, we focus on the *individual* who aims to commercialize the invention – the founder of the startup. While some startups are established by individuals who are also the original inventors of the core technology, others rely on inventions developed by individuals who are not part of the founding team, highlighting a distinction between those who develop the technology and those who bring it to market. For example, venture capitalists or experienced entrepreneurs often identify and acquire promising technologies from inventors, founding startups that harness these innovations without the inventor's direct involvement (Hellmann and Puri, 2002). Additionally, corporate spin-offs may arise when a large firm decides to commercialize internally developed innovations, appointing a leadership team distinct from the original inventors to drive the venture (Cirillo et al., 2014). In other cases, employee-inventors within existing startups or corporations may develop new technologies that are later commercialized by external entrepreneurs or other organizational members who are not directly involved in the invention process (Gambardella et al., 2015; Kim, 2022; Roach and Sauermann, 2015). Furthermore, some high-profile scientists, such as Nobel laureates, may join ventures primarily for signaling purposes, enhancing the venture's credibility without necessarily contributing to the underlying technology (Roche et al., 2020; Zucker et al., 2002).[7] These scenarios illustrate how the roles of inventor and founder can diverge, highlighting the potential importance of separating technological creation from its commercialization.

We therefore introduce a critical moderating factor: whether the inventor of the startup's core technology is also the startup's founder – an 'inventor-founder'. Although the literature on entrepreneurship has made significant strides in understanding the role of knowledge in venture creation, it has largely overlooked the critical distinction between founders who are also inventors and those who are not. This gap is particularly evident in studies that implicitly assume that the source of the venture's knowledge-base comes from the founder's own expertise, without considering variations in how knowledge is sourced or embodied within the startup (Hong et al., 2022; Roche et al., 2020). This distinction has important implications for how ventures scale, integrate knowledge, and attract external investment, yet remains underexplored in existing research – to some degree because making the direct link between person, knowledge-base and firm is tedious.

[7] Although this is a credible path, we will not focus on licensed technologies in this paper.

To illustrate this, Figure 1 differentiates ventures based on two dimensions: (1) the degree of scientific specialization of the startup's knowledge-base (i.e., whether the startup's core technology heavily relies on the inventor's own scientific work or not) and (2) whether the inventor of the core technology is also the founder of the startup. In Quadrant 1, the founder creates the venture based on their own technology that is not rooted in their own scientific work: the technology has a low degree of scientific specialization and is tied to the founder. An illustrative example would be a software engineer who develops a machine learning tool for hospital scheduling based on existing optimization techniques rather than their own academic research and starts a company to commercialize it. In Quadrant 2, the founder creates the venture based on their own technology, which is highly rooted in their own scientific work: the technology has a high degree of scientific specialization and is tied to the founder. An example for this scenario would be a biomedical researcher starting a biotech company based on their discovery of a new protein engineering technique for drug development. In Quadrant 3, the founder builds the venture from a technology developed by another inventor which is highly rooted in the inventor's scientific work: the technology has a high degree of scientific specialization and is not tied to the founder. This could be a biotech startup founded by a former pharmaceutical executive who recruits a team to commercialize a novel cancer immunotherapy, originally developed in an academic lab. Finally, in Quadrant 4, the founder develops the venture around a technology developed by another inventor which is not rooted in that inventor's own scientific work: the technology has a low degree of scientific specialization and is not tied to the founder. This could be a tech executive founding a green energy startup that develops improved solar panel coatings, based on scientific principles from materials science research that is not specifically conducted by the inventor.

Insert Figure 1 here

The presence of an inventor-founder may influence how specialized scientific knowledge is perceived, particularly by making the knowledge appear less portable. Indeed, given the intricate and tacit nature of specialized scientific knowledge, variation in whether the founder is also the inventor might indicate the extent to which this knowledge is codifiable and understandable (Agrawal, 2006). When the founder is not the inventor, it suggests that the specialized scientific knowledge embedded in the technology was transferable and understandable enough for someone else to establish a startup around it. In contrast, when the founder is the inventor, this highly specialized scientific knowledge may be perceived as less easily transferable and understandable because it is more dependent on the founder's tacit knowledge. This dynamic can amplify perceptions of dependency on the founder, as ventures with an inventor-founder may already appear more closely tied to the founder (Agarwal and Shah, 2014; Coff, 1997)

For stakeholders on the exit market, however, this reduced portability may serve as an advantage when the technology has a high degree of scientific specialization (Quadrant (2)) (Campbell et al., 2012). Contrary to investors such as VCs who often prefer to replace the founder with an outside CEO (Hellmann and Puri, 2002), acquirers and public investors may view the presence of an inventor-founder positively when a startup relies heavily on specialized scientific knowledge. First, the presence of the inventor-founder may provide acquirers and investors with essential access to a resource that is otherwise difficult to utilize effectively because of its reliance on tacit elements (Agarwal and Shah, 2014; Aghasi et al., 2022). Indeed, for acquirers, the possibility of retaining the inventor-founder post-acquisition gives them the opportunity to unlock the value of the specialized knowledge embedded in the technology (Kim, 2024), ensuring smoother integration and maximizing the technology's potential, particularly when it disrupts core business lines (Boyacıoğlu et al., 2024; Graebner, 2004). For public investors, the inventor-founder's presence ensures that the tacit knowledge and deep expertise are actively embodied within the venture. The inventor-founder's involvement may also signal strong commitment to the technology's success and reduces agency costs, providing reassurance to public investors (Jain and Kini, 1994; Nelson, 2003).

Second, the reduced mobility of the specialized scientific knowledge associated with the inventor-founder's involvement may enhance its perception as a unique, hard-to-replicate resource that offers competitive advantage by embedding the knowledge more firmly within the venture (Barney, 1991; Wernerfelt, 1984). This ensures that the knowledge cannot be easily imitated or transferred, increasing its strategic value. Thus, the interaction between specialized scientific knowledge and the inventor-founder may offer strategic advantages. The pairing of the "jockey" (in this paper, the inventor-founder) with the "horse" (in this paper, the specialized scientific knowledge) may transform what might otherwise be a liability into a strategic asset (Conti et al., 2024; Gompers and Lerner, 2001): stakeholders can leverage a highly idiosyncratic resource while simultaneously reducing its mobility, enhancing the venture's appeal and competitive positioning on the exit market. Taken together, we expect that:

The relationship between the degree to which the new venture's knowledge-base (as proxied by the invention) leverages the specialized scientific knowledge of its inventor and performance outcomes on the exit market is moderated by the founder also being the inventor. We expect that the relationship will be more positive for startups with an inventor-founder.

3 Data

3.1 Construction

We construct our dataset from the population of U.S. startups listed on Crunchbase, which provides significant information about startups' founding team, sector and financing.[8] Importantly for our analysis, Crunchbase has a broader coverage of technology startups than other sources since it also provides information on startups seeking to raise capital, regardless of whether they have successfully raised the funds, limiting potential selection and survivor bias. We then keep startups in bio-medicine (i.e., biotechnology and medical devices) because this sector is tightly linked to scientific research and has a relatively high propensity to patent. This enables us to capture the initial knowledge-base of a startup and relate it to its degree of scientific specialization. We focus on startups started after 2004 (because Crunchbase has been found to be more accurate in recent years) and before 2015 (in order to have sufficient time to observe outcomes) (Conti and Roche, 2021). Because of the incomplete coverage of founders in Crunchbase, additional information about the founding team, such as identity, educational level, graduation year and prior working experience, was derived from each startup's website, LinkedIn and Bloomberg through extensive manual searches. We retrieve information about the patents assigned to each startup using Clarivate Analytics that we complement with PatentsView and the Reliance on Science (RoS) dataset (Marx and Fuegi, 2020, 2022), which we further describe below. While our dataset accounts for patents' reassignment, so that we are able to observe patents that the startup did not apply for but is the current assignee of, one limitation of our data is that we do not capture licensing. Thus, our findings are most applicable

[8]To build this data, we closely follow Roche et al. (2020).

to technology-driven startups that rely more heavily on patent ownership rather than licensing for commercializing their technologies. We keep startups with at least one (granted) patent because we use patents to proxy for startups' knowledge-base.[9] Our final sample consists of 1,006 startups.

3.2 Specialized scientific knowledge

Our main independent variable aims to capture the extent to which a startup's core technology heavily relies on the specialized scientific knowledge of its inventor, or draws more broadly from other sources – i.e., the degree to which the startup's knowledge-base is scientifically specialized. Using the bio-medicine sector as our empirical setting is convenient because it is tightly linked to scientific research and has a relatively high propensity to patent, two features that allow us to make progress on the measurement of the independent variable we are interested in.

We capture the knowledge-base of a startup at the time of creation by considering the (granted) patents where the startup is the assignee. Patents are a good proxy for the knowledge-base of a startup because they represent formal claims of novel inventions and technological advancements that a startup aims to commercialize. These patents reflect the startup's foundational intellectual property and the areas of innovation it focuses on. We capture inventors' scientific expertise by considering their pool of scientific publications. Scientific publications are a good proxy for specialized scientific knowledge because they represent an individual's rigorous research work that has been validated as novel and significant through the peer-review process, often establishing the authors as experts in their respective fields.

We then capture the extent to which these patents build upon their inventors' own specialized scientific knowledge by considering the citations patents make to the scientific literature. When a company applies for a patent, it must list all the knowledge which it builds on, including other patents and importantly for us, scientific papers. This allows us to differentiate between citations that patents make to inventors' scientific work vs citations that patents make to other researchers' work. We then calculate the degree of scientific specialization of the startup's knowledge-base by computing the percentage of scientific citations that patents make to their inventors' scientific

[9] We keep patents applied for before startup creation or within 3 years of inception to avoid issues related to (1) capturing irrelevant or later-stage innovations not representative of the startup's initial knowledge-base, and (2) overemphasizing post-creation patenting activity that may reflect external investments or acquisitions rather than the original capabilities of the startup.

papers.

In practice, we take each patent associated with a startup and match them to the Reliance on Science (RoS) dataset (Marx and Fuegi, 2020, 2022), which provides a publicly-available set of citations from U.S. patents to scientific articles.[10] For each academic article cited by a patent, we create a self-citation dummy equal to 1 when at least one author of the academic article is matched to an inventor with a confidence score above 50.[11] This identifies instances where we can reasonably be confident that at least one inventor of the patent is citing their own academic work. We normalize this measure to account for the number of inventors, as patents with more inventors naturally have more opportunities to be associated with a self-citation. We then calculate the percentage of self-citations at the patent level by dividing the total number of self-cites by the total number of scientific citations.

$$\% \text{ self-citations} = \frac{\text{Number of scientific self-cites}}{\text{Total number of scientific cites}} \tag{1}$$

Finally, we average this measure at the firm level for startups with multiple patents. The higher the percentage of self-cites, the more a startup's knowledge-base relies on specialized scientific knowledge. Figure A2 shows an example of how we calculate this variable.

Note that our measure is inherently constructed to capture variations in inventors' specialized scientific knowledge publicly disclosed through publications. While scientific publications are arguably a good proxy for specialized scientific knowledge, it is likely that some scientific research remains private instead of being publicly disseminated, especially for inventors working in the private sector (e.g., due to trade secrecy). While we cannot rule out this possibility, we chose the bio-medical sector on purpose because it is research-driven, with individuals in this industry often coming from academic or research backgrounds where publications are a primary mean of disseminating new findings and demonstrating expertise. Most companies in this industry also allow their scientists to publish as an investment in absorptive capacity or as a way to hire the most talented individuals

[10] As detailed in previous work, patents may cite scientific research on the front-page or in the body of the text. Front-page citations are usually aimed at citing prior art while in-text citations are closer to the role played by academic citations, incorporating knowledge by reference (Bryan et al., 2020). Most research so far focused on front-page citations because they were easier to extract. Because the RoS dataset contains both types of citations, our analysis considers both front-page and in-text citations.

[11] The match is performed based on last, middle and first name (Marx and Fuegi, 2020, 2022) Results are robust to the use of more stringent thresholds, such as 75.

(Cockburn and Henderson, 1998; Stern, 2004). Yet, this means that our measure of specialized scientific knowledge might represent a lower bound. This limitation does not bias our empirical estimates as long as the extent to which the degree of scientific knowledge that remains private is not systematically correlated with startups' exit outcomes. In other words, as long as startups whose patents rely more on unpublished, proprietary knowledge do not experience systematically different exit outcomes compared to those whose patents rely more on publicly disclosed knowledge, our estimates remain unbiased. To capture differences in scientific dissemination which might correlate with our main independent variable as well as with startups' performance, we control for whether the venture is an academic startup or not (i.e., founded by at least one professor), the educational background of founders (e.g., PhD vs MD etc.) and whether founders have previous industry experience. Additionally, we report our results on the sample of academic startups only, as the presence of a professor-founder may mitigate issues of proprietary knowledge by promoting openness and dissemination norms common in Academia, allowing us to obtain a cleaner measure of specialized scientific knowledge.[12]

In subsequent analysis, we characterize the type of knowledge embodied in each scientific paper cited by a patent in more detail. To that end, we retrieve information about each publication using Dimensions AI.[13]

3.3 Summary statistics

Figure 2 presents the histogram of our main independent variable, *% self-citations*. This variable is skewed, with an average value of 1.9%. 51% of startups have patents that do not rely on their inventors' research, either because they cite scientific papers that are never written by their inventors (78%) or because they do not cite any scientific paper (22%). In both cases, the main independent variable of interest takes a value of 0, implying a relatively low scientific specialization of the startup's knowledge-base. We show in robustness that results are robust to keeping startups with at least one patent that cites science.

[12] Note that these controls are related to founders (vs inventors) because we are interested in founders' choices when creating their startup.
[13] https://www.dimensions.ai/

Insert Figure 2 here

Table 1 displays summary statistics for our sample of 1,006 firms. Startups have on average 8 patents and on average 80% of these patents cite scientific literature. 678 startups are identified as having an inventor-founder, representing 67% of the sample. 30% of the ventures in our sample are academic startups, i.e., that they have at least one professor in their founding team. The average founding-team size is 1.8 people, with 10% of ventures having at least one female founder, 60% having at least one founder with a PhD, 30% having at least one founder with a MD and 60% having at least one founder with prior industry experience. 10% of startups have a founder with prior entrepreneurship experience. 20% of the startups in our sample have a successful exit event, defined as being acquired or going public.

Insert Table 1 here

Startups in our dataset are primarily located in California (31%), Massachusetts (14%), Pennsylvania (5%) and Texas (4%).

4 Methods and Results

4.1 Empirical Strategy

The relationship between ventures' outcomes on the exit market and the extent to which they build on an invention that leverages specialized scientific knowledge is subject to the classic problem of selection: individuals choose whether to enter entrepreneurship and conditional on entering it, choose whether they predominantly rely on inventions based on specialized scientific knowledge or not. Our main empirical strategy consists of controlling for the most obvious confounding factors. While our results cannot be interpreted as definitively causal, they provide insights into which types of startups are more or less likely to succeed and some of the key factors driving these differences in outcomes. Our conversations with TTOs also suggest that inventions in our sample are likely to be comparable in terms of commercialization potential. The main difference in potential is typically detected between inventions that proceed through entrepreneurship versus other commercialization paths.[14]

To examine how the extent to which a technology is built upon specialized scientific knowledge influences subsequent startup's performance outcomes in the exit market, we run regressions of the form:

$$Y_i = \beta\% \text{ self-citations}_i + \gamma X_i + \delta_{\text{Founding year}} + \delta_{\text{State}} + \delta_{\text{Sector}} + \epsilon_i \tag{2}$$

with i indexing startups. We use robust standard errors, clustered at the startup level.

Dependent Variable: Our main outcome of interest, Success_i, is an indicator equal to 1 if the startup was acquired or went public and 0 otherwise.

Control Variables: We include controls related to the venture. $\delta_{\text{Founding year}}$ are founding year fixed-effects to control for startup age that might correlate with both outcomes on the exit market as well as the degree to which startups are based on their patents' inventors own scientific work (e.g., if the use of specialized scientific knowledge increases over time). Similarly, δ_{State} and δ_{Sector} are state and sector fixed-effects that control for state and technology trends that might be correlated with both venture outcomes and our main independent variable of interest. X_i includes a control for the log number of patents relying on scientific literature as it might influence both the calculation of our independent variable and the outcomes.[15]

X_i also includes a variety of other controls related to the founding team in order to capture factors that could influence both founders' propensity to build on specialized scientific knowledge and their venture's outcomes. We add an indicator variable equal to 1 if the startup includes a professor in the founding team, as there might be differences between academic and non-academic startups regarding the use of specialized scientific knowledge, as well as success on the exit market (Roche et al., 2020). Similarly, we include an indicator variable equal to 1 if at least one founder has some previous industry-experience and an indicator equal to 1 if at least one founder has already

[14] As a reminder, our dataset only contains startups that *own* their patents. In particular, we do not observe licensing. The decision to license is a strategic one and hinges on the expected value of the invention (Bearson and Roche, 2025). We exclude firms building on licensed technologies to ensure our sample is comparable.

[15] In practice, the correlation between these 2 variables is low and equal to 0.01.

founded one or several startups before,[16] as industry and entrepreneurial experiences might affect founders' propensity to use specialized scientific knowledge as well as their venture's success. We also include an indicator equal to 1 if at least one founder has a PhD and an indicator equal to 1 if at least one founder has a MD. These variables aim to capture differences in training and expertise that might influence the use of specialized science as well as exit outcomes. We also include a control for the number of founders at inception (hereafter referred to as 'team size'). Indeed, larger teams may have a more diverse set of knowledge, increasing the likelihood that they will draw from broader knowledge sources beyond specialized scientific expertise. Moreover, larger teams may benefit from stronger networks, potentially improving access to resources such as funding or partnerships that enhance the probability of successful exits. Because female founders may have different propensities to use specialized scientific knowledge and also face distinct challenges in securing funding or achieving successful exits, we also include an indicator equal to 1 if the founding team comprises at least one female. Finally, we also add a linear and quadratic terms for the average working experience of the founding team.[17] Together with the indicators related to education, the experience controls allow us to capture differences in founders' use of specialized scientific knowledge that might affect our independent variable and (potentially) our outcomes as well.

Moderating Variable: Our moderating variable aims to capture whether the inventor of the technology is also the founder of the startup. For each patent associated with a startup, we create an indicator equal to 1 if at least one of the inventors of the patent is also a founder of the venture. We then average this measure at the startup level and create an indicator equal to 1 if more than 50% of a startup's patents have an inventor-founder and 0 otherwise.[18] In Table A2, we observe a positive correlation between inventor-founder status and the extent to which the knowledge-base of a startup is based on specialized scientific knowledge. In particular, having an inventor-founder is associated with about a 1% increase in the use of specialized scientific knowledge in a startup's knowledge-base, representing a 38% increase with respect to the mean. This indicates that inventor-founders tend to use their own scientific work more intensively when building their ventures. This correlation is

[16] 115 startups (11% of the sample) have what we call a "serial" founder.

[17] To proxy for work experience, we manually retrieve each founder's graduation year for their highest degree obtained and measure the difference between startup founding year and this graduation year. We then take the maximum of this measure at the startup level. Results are robust to using the average level of experience across founders instead.

[18] Results are robust to using different thresholds such as 25% or 75%.

important for interpreting our interaction results, as it suggests that inventor-founders may already be predisposed to rely more on specialized scientific knowledge.

4.2 Main Results

Results are presented in Table 2. For clarity, we only keep the main coefficients of interest and present the detailed regression results in Table A3.

Insert Table 2 here

Column (1) focuses on the role of specialized scientific knowledge. Results show that startups built around an invention that relies more heavily on specialized scientific knowledge are associated with a lower likelihood of success on the exit market. More precisely, a 10p.p increase in the use of specialized scientific knowledge is associated with a 4.5% decrease in the probability of getting acquired or going public.

In Column (2), we explore the role that an inventor-founder plays on subsequent startup outcomes. Results show that when the inventor of a startup's core technology is also the founder of the startup that seeks to commercialize it, there is a negative and significant association with the likelihood of success on the exit market.

Column (3) includes both variables. The magnitude of the coefficient on specialized scientific knowledge becomes less negative but remains statistically significant, suggesting that inventor-founder status only weakly mediates the effect of scientific specialization.

Core to our conceptual framework, Column (4) examines whether the inventor-founder status might moderate the relationship between the degree of scientific specialization of the knowledge-base and startups' performance on the exit market. The main effects of scientific specialization and having an inventor-founder remain negative and significant. However, the interaction term between these two variable is positive and significant. This suggests that the negative relationship between startups' reliance on specialized scientific knowledge and their likelihood of getting acquired is driven by startups *without* an inventor-founder. The sign of the coefficient on the interaction term suggests that having an inventor-founder is a positive attribute for potential acquirers and investors on the public market when startups build around an invention that relies more heavily on specialized scientific knowledge. Column (5) further controls for the log of funding received within 5 years of inception. We find similar results, ensuring that the results in Column (4) are not confounded by differences in initial funding levels.

Overall, the results of Column (4) align with the predictions of our conceptual framework. We find that the challenges associated with specialized scientific knowledge seem to outweigh its benefits for stakeholders in the exit market. We also find that a higher degree of scientific specialization is mitigated by the presence of an inventor-founder. These findings offer a nuanced perspective on how the characteristics of a startup's knowledge-base shape performance outcomes, and how this relationship varies depending on the identity of the startup's founder.

In Table A4, we add sector $\times$ startup creation year fixed effects to control for time-varying sector-specific trends (e.g., sector-specific use of specialized scientific knowledge over time). These models confirm our results.

5 Towards finding potential explanations

5.1 Characteristics of specialized scientific knowledge

The first main finding of the previous section suggests that startups which rely more on specialized scientific knowledge perform worse on the exit market. The natural question that ensues is why? In this section, we run additional analysis to provide tentative explanations for our findings. We first highlight the micro-foundational differences that distinguish the knowledge-base of startups that rely more heavily on specialized scientific knowledge. To that end, we take each paper cited by the startup's patents and retrieve information about them using Dimensions AI. We also use information about the patents themselves. Following the conceptual framework, we explore differences in (i) research breadth, (ii) commercial appeal, (iii) patent breadth and (iv) number of inventors, to shed light on knowledge tacitness, scalability and ease of integration into technological solutions.[19]

For the purpose of our study, we conceptualize the *breadth* of a paper as the diversity of topics it entails. While it is hard to measure knowledge and in particular its breadth, we proxy for it using a feature of Dimensions AI called *concepts*. For each paper, Dimensions AI uses a machine

[19]We summarize how we empirically proxy for these variables in Table A5. Table A6 shows summary statistics for each variable.

learning algorithm that assigns concepts (i.e., topics) to each paper as well as a relevance score. We provide example of papers abstracts and their associated concepts in Appendix A. For each paper cited by the patents of a startup, we count the number of unique concepts with a score of 0.5 or above, in order to assign the most relevant topics to a given paper. We then average this measure at the startup level. Results are presented in Panel A Column (1) of Table 3 and show a significantly negative coefficient. This suggests that startups whose technology leverages more heavily specialized scientific knowledge build on research that is narrower (i.e., incorporate a less diverse range of ideas).

We then examine whether technologies that leverages more heavily specialized scientific knowledge build on research that is less easily *integrated into technological solutions*. Specialized scientific knowledge may be more individual-specific and address specific scientific challenges, rather than being driven by immediate market needs or practical applications. As a result, such research may be less aligned with the demands of industry, making it more difficult to adapt into scalable, commercially viable innovations. To empirically test this idea, we calculate the number of citations that papers received from patents (excluding patents from the focal startup) and average this measure at the startup level.[20] When a scientific paper receives a higher number of citations from patents, it might indicate that the knowledge in that paper is more easily integrated into technologies. Panel A Column (2) of Table 3 shows that a higher reliance on specialized scientific knowledge is associated with the use of papers that tend to be *less used by other patents*, suggesting that they are less easily integrated into technological solutions. In Panel A Column (3), we show that this is not simply related to the fact that these papers do more 'basic' (as opposed to 'applied') science. For each paper, we retrieved information about the Journal Commercial Impact Factor (JCIF) of the journal the paper was published in.[21] JCIF is a measure designed to capture the 'appliedness' of scientific journals by reflecting how often the research published in those journals is cited by patents (Bikard and Marx, 2020). We then average this measure at the startup level and use it as outcome. Our results are inconclusive, suggesting that the negative relationship between startups' reliance on more specialized scientific knowledge and patent citations is not simply related to the use of

[20]This measure is normalized following Perry and Reny (2016) in order to account for differences in publication year and field.

[21]This measure is normalized following Perry and Reny (2016) in order to account for differences in publication year and field.

research that would be more 'basic' in nature. In other words, the reduced integration of these papers into technological solutions is not merely a function of their fundamental or theoretical focus but likely stems from the inherent difficulty in translating highly specialized, individual-specific knowledge into widely applicable, scalable technologies.

Panel A Columns (4) and (5) examine the characteristics of the patents themselves and reveal further information. We find that an increase in the use of specialized scientific knowledge in patents is associated with a decrease in both the number of claims and the number of inventors. The decrease in the number of claims is consistent with the idea that patents relying more on specialized scientific knowledge may be narrower in scope, reflecting the more focused, individual-specific nature of their invention and a lower commercial potential. Additionally, the decrease in the number of inventors implies that these patents are more often driven by smaller teams. This might limit opportunities for knowledge sharing and codification, making the patent appear more reliant on tacit, individually-held knowledge.[22]

Panel B shows no significant relationship between *Inventor-founder* and the type of scientific knowledge used, except for the JCIF measure which is negative. This suggests that the role of the inventor-founder might not alter the fundamental characteristics of the knowledge used. Table A8 logs these outcomes. Results are directionally similar though noisier.[23]

Insert Table 3 here

5.2 The inventor-founder and the portability of knowledge

To deepen our understanding of the role played by the inventor-founder, we analyze heterogeneity in performance outcomes, focusing on funding, acquisition, and IPO separately. Each of these outcomes highlights distinct stakeholder priorities, offering insight into how the inventor-founder

[22]Table A7 runs the main regressions of interest of *Success* on *% self-citations* controlling for each of the measures previously described. While we remain careful in arguing about a well-identified mechanism, we find that the coefficient on *% self-citations* decreases, probing our hypothesis that specialized scientific knowledge might impact performance through smaller research breadth, a lower commercial appeal, smaller patent breadth and a lower number of inventors.

[23]Note that the dependent variables we explore in this section are continuous measures because they are averages taken at the startup level and are normalized at the field and publication year level. Therefore, we cannot use Poisson regressions.

status influences the portability of the specialized scientific knowledge. Table A6 shows that in our sample, 10% of the startups are acquired and 10% of them go public through an IPO. The amount of funds raised within 5 years of inception is skewed, with an average of $U.S. 14.7 million.

Funding: Investors play a critical vetting role in early-stage ventures. Their primary concern is to evaluate the startup's potential and assemble a successful leadership team to drive the necessary changes. Importantly, investors often prefer to replace the founder with an outside CEO to optimize business expansion and market strategy (Hellmann and Puri, 2002). If the presence of an inventor-founder reduces the portability of knowledge – making it appear more embedded in the individual and harder to transfer – investors may perceive these ventures as riskier investments, given the challenges of leadership transitions and broader market adaptation. To test this, we examine the likelihood of raising more than $10 million in the first five years of inception.[24] Results are presented in Panel A of Table 4. We find that the negative relationship between startups' reliance on specialized scientific knowledge and their likelihood of raising at least $U.S. 10 million within 5 years of inception is driven by startups with an inventor-founder. More precisely, Column (4) shows that the main relationship on scientific specialization is not significant on conventional levels, implying that there is likely no relationship between startups' reliance on specialized scientific knowledge and the likelihood of raising a significant amount of funds when the founder is not the inventor. However, the interaction term is negative and significant. This implies that when the founder is also the inventor of the core startup's technology, a higher reliance on specialized scientific knowledge is negatively associated with the likelihood of raising at least $U.S. 10 million. Overall, we infer from this that investors appear particularly concerned when ventures aim to commercialize an invention that is highly scientifically specialized and tightly linked to the founder, consistent with concerns over the lower portability of such knowledge.

Acquisition: Acquirers, in contrast, tend to focus on post-deal integration and scalability. The tacit and narrower nature of knowledge used in inventions that heavily rely on specialized scientific knowledge could complicate integration. However, the presence of an inventor-founder may actually be advantageous in this context: because the specialized scientific knowledge is deeply embedded in the founder, acquirers may retain this individual post-acquisition to ensure a smoother transfer of

[24] Based on conversations with investors, this is a cutoff typical for the industry when considering serious "growth"-oriented investments.

knowledge and facilitate integration (Agarwal and Shah, 2014; Aghasi et al., 2022). Table 4, Panel B, shows that while specialization is negatively associated with acquisition likelihood, the interaction term between scientific specialization and the presence of an inventor-founder is insignificant. This suggests that having an inventor-founder does not exacerbate concerns for acquirers. If anything, the interaction term is positive, potentially because retaining the inventor-founder in the post-acquisition phase helps preserve the knowledge required for integration.

IPO: The IPO process emphasizes the market potential of a standalone company rather than immediate knowledge transfer or integration. In this context, the presence of an inventor-founder may offer significant advantages when a startup's knowledge-base relies heavily on specialized scientific knowledge. The inventor-founder's involvement ensures that the deep and tacit knowledge is utilized and remains actively embodied within the venture, creating a unique and hard-to-replicate resource that provides competitive advantage. Moreover, the presence of an inventor-founder may signal strong commitment to the venture's success, reducing agency costs and enhancing public investors' confidence. Table 4, Panel C reveals a positive and significant interaction between specialization and inventor-founder, suggesting that the presence of an inventor-founder partially mitigates the challenges associated with specialized scientific knowledge.

Overall, these results suggest that the inventor-founder plays a critical role in shaping the portability and perception of knowledge. For funding, we find that the presence of an inventor-founder seems to heighten investors' concerns about ventures with a highly scientifically specialized knowledge-base. This is consistent with the idea that the presence of the inventor-founder amplifies the perceived dependence of the knowledge on a specific individual. This may make investors more cautious, as they seek ventures with knowledge that is both codifiable and easily transferable to enable leadership transitions or growth-focused restructuring. In acquisitions, specialized scientific knowledge presents some challenges, but this seems less problematic with the presence of an inventor-founder, consistent with the idea that the retention of the inventor-founder offers a pathway to mitigate these concerns. For IPOs, where immediate knowledge transfer is less of a concern, the inventor-founder dynamic paired with high scientific specialization is actually positive. This is consistent with the notion that the inventor-founder's presence ensures that the specialized scientific knowledge can be fully exploited and not easily replicated, offering the venture a stronger avenue to achieve competitive advantage.

Taken together, these results highlight the critical role of the inventor-founder in influencing the portability of specialized scientific knowledge. By examining funding, acquisition, and IPO independently, our findings reveal that the ability to effectively transfer or integrate specialized knowledge is not an inherent characteristic but is shaped by the individual embodying the knowledge. This highlights the strategic importance of the inventor-founder in navigating the challenges and opportunities of leveraging specialized scientific knowledge in entrepreneurial ventures. It also underscores the importance of aligning the characteristics of the knowledge-holder and the knowledge-base with the venture's intended exit strategy, as the interplay between these factors may significantly shape the likelihood of success across different outcomes.

Insert Table 4 here

---

6 Robustness checks

In Table A9, we run our main regressions regarding startups' performance on the sample of academic startups only. This is useful because academic startups tend to have a more consistent engagement with specialized scientific knowledge, ensuring that the nature of the knowledge-base is more uniform across ventures. By focusing on this subset, we reduce the variability introduced by non-academic startups, offering clearer insights into the relationship between specialized scientific knowledge and performance outcomes. Despite the lower sample size, we find very similar results as our main regressions. In particular, startups that build upon a technology that leverages more heavily specialized scientific knowledge have a lower likelihood of success on the exit market. Table A10 adds sector × startup creation year fixed effects and displays similar conclusions. Tables A11 to A12 explore the micro-foundational differences that distinguish the knowledge-base of academic startups that rely more or less on specialized scientific knowledge and show similar conclusions. In particular, startups whose knowledge-base relies more heavily on specialized scientific knowledge build on research that is narrower and less easily integrated into technological solutions. Their patents are also narrower and appear more reliant on tacit knowledge.

In Table A13, we reiterate our main performance regressions excluding startups that have no patent that rely on scientific literature. This ensures that when the independent variable equals 0, it now exclusively comes from firms that have patents that cite scientific literature but do not cite their investors' own scientific work, vs startups that do not cite any science at all. We find similar results. Table A14 adds sector $\times$ startup creation year fixed effects and displays similar conclusions. Tables A15 to A16 display similar conclusions regarding the micro-foundational differences of the knowledge-base of startups that rely more or less on specialized scientific knowledge.

In order to account as best as possible for the underlying research quality and potential, Table A17 shows the main performance regressions, including controls for the number of scientific citations received by papers cited in a startup's patents and the quality of the journals in which these papers were published as proxied by the Journal Impact Factor (JIF). We find similar results and, if anything, stronger magnitudes.

7 Discussion and Conclusion

Our study investigates the role of specialized scientific knowledge in shaping startup performance outcomes, with a particular focus on how this knowledge interacts with the founder's role as inventor of the startup's underlying technology. Through our analysis of 1,006 bio-medicine startups, we find that startups exhibit poorer performance outcomes when they build on a technology that leverages more heavily specialized scientific knowledge. These findings are consistent with the idea that specialized scientific knowledge, while a source of competitive differentiation, may introduce significant challenges in engaging with external stakeholders, who may struggle to fully understand or integrate such knowledge into scalable commercial ventures. Delving deeper into the mechanisms behind these results, we find that startups whose technology relies more on specialized scientific knowledge build on research that is narrower and less cited by patents. Their patents are also narrower in scope and associated with a smaller team of inventors. This suggests that inventions that rely more heavily on specialized scientific knowledge might be harder to integrate into technological solutions, of lower commercial appeal and more reliant on tacit knowledge. This is consistent with anecdotal evidence from informal interviews, where stakeholders have observed that inventors building a technology closer to their own scientific work may try to push it forward at the expense of broader considerations, which may subsequently hinder startups' ability to develop solutions that align with market needs and attract external support.

Importantly, our research highlights the critical role of the knowledge-holder in shaping how specialized scientific knowledge is perceived by external stakeholders. Our results show that the presence of an inventor-founder significantly moderates the relationship between specialized scientific knowledge and startup performance on the exit market. Specifically, having an inventor-founder seems to make the knowledge appear less portable. This might help to mitigate the challenges associated with specialized scientific knowledge by enabling acquirers and investors to effectively unlock the value of a resource that would otherwise be difficult to utilize, while simultaneously enhancing its defensibility as a unique and hard-to-replicate asset.

We contribute to a body of work examining the importance of knowledge for firm strategy by providing suggestive evidence that the benefits of specialized knowledge may be offset by challenges in the context of entrepreneurship (Conner and Prahalad, 1996; Galunic and Rodan, 1998; Grant, 1996; Wang et al., 2009). Moreover, our study highlights how the role of the inventor-founder, which has been largely overlooked in the literature, plays a critical part in influencing the 'portability' of this knowledge. Managerially, our findings suggest that founders must carefully balance the benefits of deep scientific expertise with the risks it poses in terms of external stakeholder engagement. Codifying specialized scientific knowledge, broadening the venture's technological base, and fostering partnerships that facilitate knowledge transfer may help mitigate the challenges of commercializing highly specialized scientific knowledge. Additionally, the presence of an inventor-founder, while valuable in certain contexts, critically influences firm success depending on the stage of the venture's life.

Our study is not without limitations. Although we take great care in controlling for potential confounds and show that our results are robust to different specifications, our results cannot be formally interpreted as causal. Additionally, our sample is limited to U.S.-based startups in the bio-medicine sector. While this sector plays a critical role in driving technological innovation, and its strong ties to academic research make it particularly relevant for studying knowledge specialization, it may not fully represent other industries or countries where knowledge specialization operates differently. Future studies could examine how our findings vary across other technology-intensive sectors with distinct commercialization paths. Moreover, our measure of scientific knowledge specialization primarily captures specialized scientific knowledge through scientific publications.

Future studies could explore alternative or complementary measures that capture specialized scientific knowledge through additional sources. Finally, future work may also examine features other than scientific knowledge specialization and their impact on entrepreneurial success. Given the complexity of firm creation, and persistent consequences of early decisions around the resources used as the foundation of the firm (Geroski et al., 2010), opportunities for follow-on work abound.

References

Agarwal, R., Echambadi, R., Franco, A. M. and Sarkar, M. B. (2004), 'Knowledge transfer through inheritance: Spin-out generation, development, and survival', *Academy of Management Journal* **47**(4), 501–522.

Agarwal, R. and Shah, S. K. (2014), 'Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators', *Research Policy* **43**(7), 1109–1133.

Aghasi, K., Colombo, M. G. and Rossi-Lamastra, C. (2022), 'Post-acquisition retention of target founder-CEOs: Looking beneath the surface', *Journal of Management Studies* **59**(4), 958–997.

Agrawal, A. (2006), 'Engaging the inventor: Exploring licensing strategies for university inventions and the role of latent knowledge', *Strategic Management Journal* **27**(1), 63–79.

Ahmadpoor, M. and Jones, B. F. (2017), 'The dual frontier: Patented inventions and prior scientific advance', *Science* **357**(6351), 583–587.

Arora, A., Belenzon, S. and Suh, J. (2022), 'Science and the market for technology', *Management Science* **68**(10), 7176–7201.

Åstebro, T., Chen, J. and Thompson, P. (2011), 'Stars and misfits: Self-employment and labor market frictions', *Management Science* **57**(11), 1999–2017.

Balachandran, S. (2024), 'The inside track: Entrepreneurs' corporate experience and startups' access to incumbent partners' resources', *Strategic Management Journal* **45**(6), 1117–1150.

Barney, J. (1991), 'Firm resources and sustained competitive advantage', *Journal of Management* **17**(1), 99–120.

Baruffaldi, S. and Poege, F. (2024), 'Like stars: How firms learn at scientific conferences', *Management Science* .

Becker, G. S. (1964), *Human capital: A theoretical and empirical analysis, with special reference to education*, University of Chicago Press.

Bergh, D. D., D'Oria, L., Crook, T. R. and Roccapriore, A. (2024), 'Is knowledge really the most important strategic resource? A meta-analytic review', *Strategic Management Journal* .

Bikard, M. and Marx, M. (2020), 'Bridging academia and industry: How geographic hubs connect university science and corporate technology', *Management Science* **66**(8), 3425–3443.

Botelho, T. L., Fehder, D. and Hochberg, Y. (2021), Innovation-driven entrepreneurship, Technical report, National Bureau of Economic Research.

Boyacıoğlu, B., Özdemir, M. N. and Karim, S. (2024), 'Acqui-hires: Redeployment and retention of human capital post-acquisition', *Strategic Management Journal* **45**(2), 205–237.

Bryan, K. A., Ozcan, Y. and Sampat, B. (2020), 'In-text patent citations: A user's guide', *Research Policy* **49**(4), 103946.

Campbell, B. A., Ganco, M., Franco, A. M. and Agarwal, R. (2012), 'Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance', *Strategic Management Journal* **33**(1), 65–87.

Chattopadhyay, S., Honoré, F. and Won, S. (2024), 'Free range startups? Market scope, academic founders, and the role of general knowledge in AI', *Strategic Management Journal* .

Cirillo, B., Brusoni, S. and Valentini, G. (2014), 'The rejuvenation of inventors through corporate spinouts', *Organization Science* **25**(6), 1764–1784.

Cockburn, I. M. and Henderson, R. M. (1998), 'Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery', *The Journal of Industrial Economics* **46**(2), 157–182.

Coff, R. W. (1997), 'Human assets and management dilemmas: Coping with hazards on the road to

resource-based theory', *Academy of Management Review* **22**(2), 374–402.
Cohen, W. M. and Levinthal, D. A. (1990), 'Absorptive capacity: A new perspective on learning and innovation', *Administrative Science Quarterly* **35**(1), 128–152.
Collis, D. J. and Montgomery, C. A. (2008), 'Competing on resources', *Harvard Business Review* **86**(7/8), 140.
Conner, K. R. and Prahalad, C. K. (1996), 'A resource-based theory of the firm: Knowledge versus opportunism', *Organization Science* **7**(5), 477–501.
Conti, A., Peukert, C. and Roche, M. P. (2024), 'Beefing IT up for your investor? Engagement with open source communities, innovation and startup funding: Evidence from github', *Organization Science* .
Conti, A. and Roche, M. (2021), 'Lowering the bar? External conditions, opportunity costs, and high-tech start-up outcomes', *Organization Science* **32**(4), 965–986.
Dasgupta, P. and David, P. A. (1994), 'Toward a new economics of science', *Research Policy* **23**(5), 487–521.
Dencker, J. C., Gruber, M. and Shah, S. K. (2009), 'Pre-entry knowledge, learning, and the survival of new firms', *Organization Science* **20**(3), 516–537.
Dosi, G. (1988), 'Sources, procedures, and microeconomic effects of innovation', *Journal of Economic Literature* pp. 1120–1171.
Eckhardt, J. T. and Shane, S. A. (2003), 'Opportunities and entrepreneurship', *Journal of Management* **29**(3), 333–349.
Fleming, L. (2001), 'Recombinant uncertainty in technological search', *Management Science* **47**(1), 117–132.
Fleming, L. and Sorenson, O. (2004), 'Science as a map in technological search', *Strategic Management Journal* **25**(8-9), 909–928.
Galunic, D. C. and Rodan, S. (1998), 'Resource recombinations in the firm: Knowledge structures and the potential for schumpeterian innovation', *Strategic Management Journal* **19**(12), 1193–1201.
Gambardella, A., Ganco, M. and Honoré, F. (2015), 'Using what you know: Patented knowledge in incumbent firms and employee entrepreneurship', *Organization Science* **26**(2), 456–474.
Ganco, M. (2013), 'Cutting the gordian knot: The effect of knowledge complexity on employee mobility and entrepreneurship', *Strategic Management Journal* **34**(6), 666–686.
Garg, S., Howard, M. and Pahnke, E. C. (2025), 'Directors in new technology-based ventures: An empirical inquiry', *Journal of Business Venturing* **40**(2), 106431.
Geroski, P. A., Mata, J. and Portugal, P. (2010), 'Founding conditions and the survival of new firms', *Strategic Management Journal* **31**(5), 510–529.
Gompers, P. and Lerner, J. (2001), 'The venture capital revolution', *Journal of Economic Perspectives* **15**(2), 145–168.
Graebner, M. E. (2004), 'Momentum and serendipity: How acquired leaders create value in the integration of technology firms', *Strategic Management Journal* **25**(8-9), 751–777.
Grant, R. M. (1996), 'Toward a knowledge-based theory of the firm', *Strategic Management Journal* **17**(S2), 109–122.
Hellmann, T. and Puri, M. (2002), 'Venture capital and the professionalization of start-up firms: Empirical evidence', *The Journal of Finance* **57**(1), 169–197.
Henderson, R. and Cockburn, I. (1994), 'Measuring competence? Exploring firm effects in pharmaceutical research', *Strategic Management Journal* **15**(S1), 63–84.
Hoang, H. and Gimeno, J. (2010), 'Becoming a founder: How founder role identity affects entrepreneurial transitions and persistence in founding', *Journal of Business Venturing* **25**(1), 41–53.
Hong, S., Guo, K., Zhang, H. and Hsu, D. H. (2022), 'When do ventures founded by research-oriented entrepreneurs outperform?', *Working Paper* .
Jain, A. and Mitchell, W. (2022), 'Specialization as a double-edged sword: The relationship of scientist specialization with R&D productivity and impact following collaborator change', *Strategic Management Journal* **43**(5), 986–1024.
Jain, B. A. and Kini, O. (1994), 'The post-issue operating performance of IPO firms', *The Journal of Finance* **49**(5), 1699–1726.
Jensen, M. C. and Meckling, W. H. (2019), Theory of the firm: Managerial behavior, agency costs and ownership structure, *in* 'Corporate governance', Gower, pp. 77–132.
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.
Junkunc, M. T. and Eckhardt, J. T. (2009), 'Technical specialized knowledge and secondary shares in initial

public offerings', *Management Science* **55**(10), 1670–1687.
Kacperczyk, A. and Younkin, P. (2017), 'The paradox of breadth: The tension between experience and legitimacy in the transition to entrepreneurship', *Administrative Science Quarterly* **62**(4), 731–764.
Kaplan, S. and Vakili, K. (2015), 'The double-edged sword of recombination in breakthrough innovation', *Strategic Management Journal* **36**(10), 1435–1457.
Katila, R. and Ahuja, G. (2002), 'Something old, something new: A longitudinal study of search behavior and new product introduction', *Academy of Management Journal* **45**(6), 1183–1194.
Kim, J. D. (2022), 'Startup acquisitions, relocation, and employee entrepreneurship', *Strategic Management Journal* **43**(11), 2189–2216.
Kim, J. D. (2024), 'Startup acquisitions as a hiring strategy: Turnover differences between acquired and regular hires', *Strategy Science* **9**(2), 118–134.
Klepper, S. (2001), 'Employee startups in high-tech industries', *Industrial and Corporate Change* **10**(3), 639–674.
Kogut, B. and Zander, U. (1992), 'Knowledge of the firm, combinative capabilities, and the replication of technology', *Organization Science* **3**(3), 383–397.
Krieger, J. L., Schnitzer, M. and Watzinger, M. (2024), 'Standing on the shoulders of science', *Strategic Management Journal* **45**(9), 1670–1695.
Leahey, E. (2007), 'Not by productivity alone: How visibility and specialization contribute to academic earnings', *American Sociological Review* **72**(4), 533–561.
Lee, H., Shah, S. K. and Agarwal, R. (2024), 'Spinning an entrepreneurial career: Motivation, attribution, and the development of organizational capabilities', *Strategic Management Journal* **45**(3), 463–506.
Leland, H. E. and Pyle, D. H. (1977), 'Informational asymmetries, financial structure, and financial intermediation', *The Journal of Finance* **32**(2), 371–387.
Makri, M., Hitt, M. A. and Lane, P. J. (2010), 'Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions', *Strategic Management Journal* **31**(6), 602–628.
Marx, M. and Fuegi, A. (2020), 'Reliance on science: Worldwide front-page patent citations to scientific articles', *Strategic Management Journal* **41**(9), 1572–1594.
Marx, M. and Fuegi, A. (2022), 'Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations', *Journal of Economics & Management Strategy* **31**(2), 369–392.
McEvily, S. K. and Chakravarthy, B. (2002), 'The persistence of knowledge-based advantage: an empirical test for product performance and technological knowledge', *Strategic Management Journal* **23**(4), 285–305.
Nagle, F. (2018), 'Learning by contributing: Gaining competitive advantage through contribution to crowdsourced public goods', *Organization Science* **29**(4), 569–587.
Nagle, F. and Teodoridis, F. (2020), 'Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration', *Strategic Management Journal* **41**(1), 55–85.
Nelson, T. (2003), 'The persistence of founder influence: Management, ownership, and performance effects at initial public offering', *Strategic Management Journal* **24**(8), 707–724.
Nonaka, I. and Von Krogh, G. (2009), 'Perspective—tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory', *Organization Science* **20**(3), 635–652.
Park, U. D., Rha, Y., Shah, S. K. and Chattopadhyay, S. (2024), 'An image of industry: exploring the effects of knowledge sources in the medical imaging industry', *Industrial and Corporate Change* **33**(1), 64–89.
Penrose, E. T. (1959), *The Theory of the Growth of the Firm*, Oxford University Press.
Perry, M. and Reny, P. J. (2016), 'How to count citations if you must', *American Economic Review* **106**(9), 2722–2741.
Peteraf, M. A. (1993), 'The cornerstones of competitive advantage: A resource-based view', *Strategic Management Journal* **14**(3), 179–191.
Polanyi, M. (2009), The tacit dimension, *in* 'Knowledge in organisations', Routledge, pp. 135–146.
Polidoro, F. and Yang, W. (2021), 'Corporate investment relationships and the search for innovations: An examination of startups' search shift toward incumbents', *Organization Science* **32**(4), 909–939.
Puranam, P., Singh, H. and Chaudhuri, S. (2009), 'Integrating acquired capabilities: When structural integration is (un) necessary', *Organization Science* **20**(2), 313–328.
Roach, M. (2017), 'Encouraging entrepreneurship in university labs: Research activities, research outputs, and early doctorate careers', *PLOS ONE* **12**(2).

Roach, M. and Sauermann, H. (2015), 'Founder or joiner? The role of preferences and context in shaping different entrepreneurial interests', *Management Science* **61**(9), 2160–2184.
Roche, M., Conti, A. and Rothaermel, F. T. (2020), 'Different founders, different venture outcomes: A comparative analysis of academic and non-academic startups', *Research Policy* **49**(10), 104062.
Rumelt, R. P. et al. (1984), 'Towards a strategic theory of the firm', *Competitive Strategic Management* **26**(3), 556–570.
Sauermann, H. and Stephan, P. (2013), 'Conflicting logics? A multidimensional view of industrial and academic science', *Organization Science* **24**(3), 889–909.
Schumpeter, J. A. (1912), 'Theorie der Wirtschaftlichen Entwicklung: Eine Untersuchung ueber Unternehmergewinn', *Kaptial, Kredit, Zins und den Konjunkturzyklus, Duncker&Humblot, Berlin* .
Shah, S. K., Agarwal, R. and Echambadi, R. (2019), 'Jewels in the crown: Exploring the motivations and team building processes of employee entrepreneurs', *Strategic Management Journal* **40**(9), 1417–1452.
Shah, S. K. and Pahnke, E. C. (2014), 'Parting the ivory curtain: Understanding how universities support a diverse set of startups', *The Journal of Technology Transfer* **39**, 780–792.
Shane, S. (2000), 'Prior knowledge and the discovery of entrepreneurial opportunities', *Organization Science* **11**(4), 448–469.
Shane, S. and Stuart, T. (2002), 'Organizational endowments and the performance of university start-ups', *Management Science* **48**(1), 154–170.
Shane, S. and Venkataraman, S. (2000), 'The promise of entrepreneurship as a field of research', *Academy of Management Review* **25**(1), 217–226.
Sørensen, J. B. (2007), 'Bureaucracy and entrepreneurship: Workplace effects on entrepreneurial entry', *Administrative Science Quarterly* **52**(3), 387–412.
Sorenson, O. and Fleming, L. (2004), 'Science and the diffusion of knowledge', *Research Policy* **33**(10), 1615–1634.
Spender, J.-C. and Grant, R. M. (1996), 'Knowledge and the firm: Overview', *Strategic Management Journal* **17**(S2), 5–9.
Stenard, B. S. and Sauermann, H. (2016), 'Educational mismatch, work outcomes, and entry into entrepreneurship', *Organization Science* **27**(4), 801–824.
Stern, S. (2004), 'Do scientists pay to be scientists?', *Management Science* **50**(6), 835–853.
Stuart, T. E. (2000), 'Interorganizational alliances and the performance of firms: a study of growth and innovation rates in a high-technology industry', *Strategic Management Journal* **21**(8), 791–811.
Stuart, T. E., Hoang, H. and Hybels, R. C. (1999), 'Interorganizational endorsements and the performance of entrepreneurial ventures', *Administrative Science Quarterly* **44**(2), 315–349.
Stuart, T. E. and Sorenson, O. (2007), 'Strategic networks and entrepreneurial ventures', *Strategic Entrepreneurship Journal* **1**(3-4), 211–227.
Teece, D. J. (1986), 'Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy', *Research Policy* **15**(6), 285–305.
Teodoridis, F., Bikard, M. and 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.
Tripsas, M. and Gavetti, G. (2017), 'Capabilities, cognition, and inertia: Evidence from digital imaging', *The SMS Blackwell Handbook of Organizational Capabilities* pp. 393–412.
Wang, H. C., He, J. and Mahoney, J. T. (2009), 'Firm-specific knowledge resources and competitive advantage: the roles of economic-and relationship-based employee governance mechanisms', *Strategic Management Journal* **30**(12), 1265–1285.
Wernerfelt, B. (1984), 'A resource-based view of the firm', *Strategic Management Journal* **5**(2), 171–180.
Zaheer, A., Hernandez, E. and Banerjee, S. (2010), 'Prior alliances with targets and acquisition performance in knowledge-intensive industries', *Organization Science* **21**(5), 1072–1091.
Zander, U. and Kogut, B. (1995), 'Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test', *Organization Science* **6**(1), 76–92.
Zucker, L. G., Darby, M. R. and Armstrong, J. S. (2002), 'Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology', *Management Science* **48**(1), 138–153.

Figures

Figure 1: Conceptual framework: Scientific specialization and the inventor-founder status

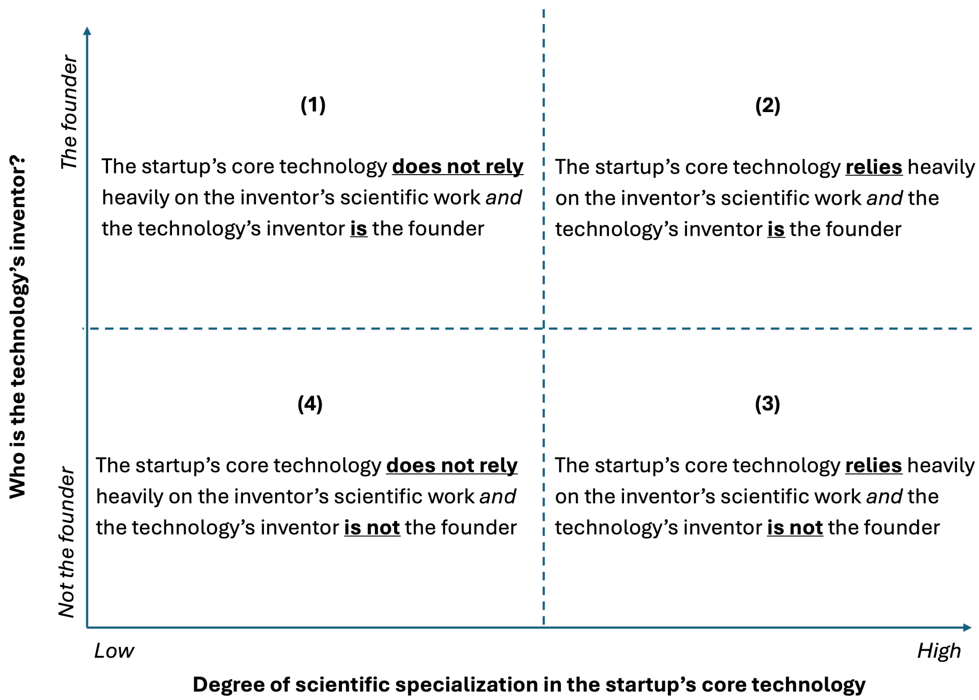
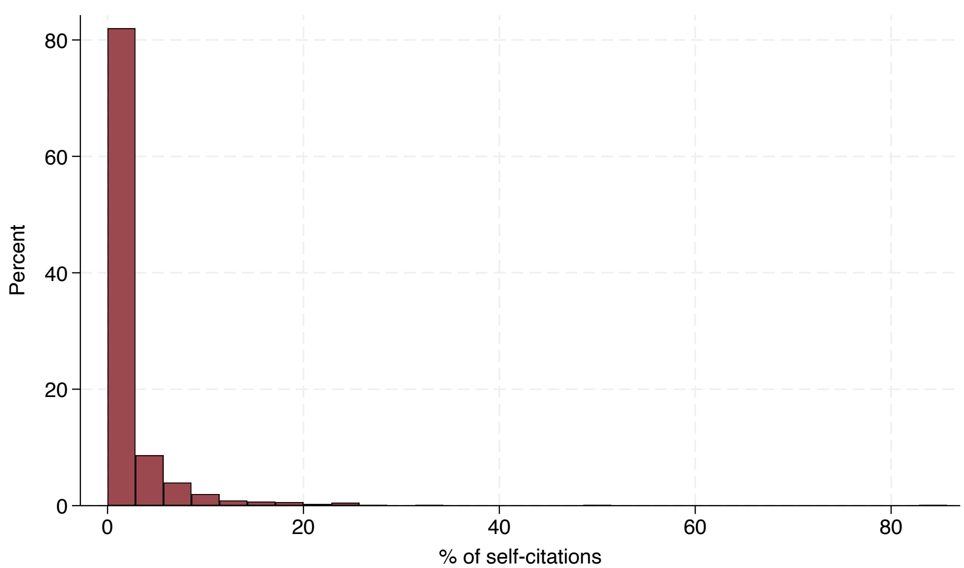


Figure 2: Histogram of the independent variable



Notes: This figure displays the distribution of our main independent variable, *% self-citations*. To calculate this variable, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. See Figure A2 for an example.

Tables

Table 1: Summary statistics, startup level

	Min	Mean	SD	Max
# of patents	1.0	7.6	15.6	341.0
# of patents citing science	0.0	6.1	10.6	151.0
Self-citations (%)	0.0	1.9	4.9	85.7
Inventor-founder	0.0	0.7	0.5	1.0
Academic startup	0.0	0.3	0.5	1.0
Founding team size	1.0	1.8	0.9	7.0
At least one female founder	0.0	0.1	0.3	1.0
At least one PhD founder	0.0	0.6	0.5	1.0
At least one MD founder	0.0	0.3	0.5	1.0
Industry experience	0.0	0.6	0.5	1.0
Entrepreneurship experience	0.0	0.1	0.3	1.0
Biotech sector	0.0	0.7	0.5	1.0
Founding year	2005.0	2008.0	2.1	2012.0
Success (0/1)	0.0	0.2	0.4	1.0
Observations	1006			

Notes: This table displays summary statistics at the startup level for our sample of 1,006 startups

Table 2: Success on the exit market - Summary

	Success (0/1)				
	(1)	(2)	(3)	(4)	(5)
% of self-citations	-0.448**		-0.383**	-1.701***	-1.896***
	(0.190)		(0.186)	(0.449)	(0.462)
Inventor-founder		-0.111***	-0.108***	-0.132***	-0.131***
		(0.034)	(0.034)	(0.037)	(0.037)
Inventor-founder=1 × % of self-citations				1.551***	1.772***
				(0.474)	(0.488)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2	0.2
Observations	739	739	739	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. All columns include controls related to the founding team and are described in section 4.1. Column (5) adds the log of funding received within 5 years of inception. Robust standard errors (in parentheses) are clustered at the startup level.

Table 3: Characteristics of specialized scientific knowledge

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-16.728*	-2.644***	2.577	-12.308*	-4.806***
	(9.542)	(0.950)	(2.828)	(7.106)	(1.208)
R-Sq	0.2	0.1	0.1	0.1	0.1
Panel B: Inventor-Founder					
Inventor-founder	-0.332	0.041	-0.064*	-1.478	-0.098
	(1.312)	(0.152)	(0.038)	(1.166)	(0.151)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.1	0.1	0.1	0.1
Dep. Var. Mean	40.9	1.9	1.0	20.2	3.3
Observations	659	662	656	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. In column (1), we calculate for each paper cited by a patent the number of unique concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, normalized to account for publication year and field and excluding citations from the focal startup's patents and average this measure at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

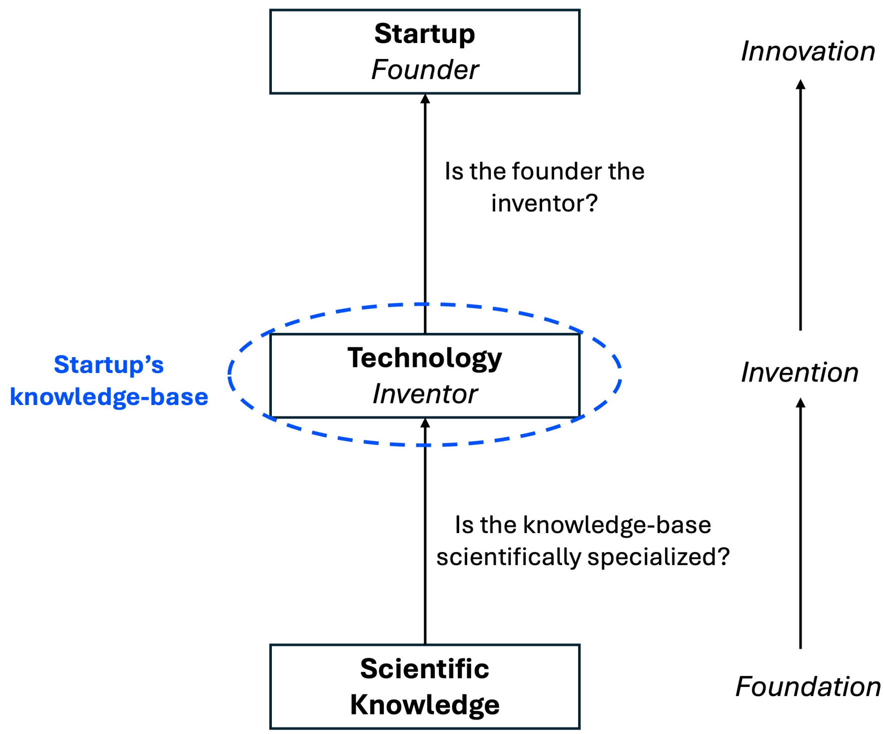
Table 4: Heterogeneity in performance outcomes - Summary

	(1)	(2)	(3)	(4)
Panel A: Funding				
% of self-citations	-0.428*		-0.347	1.098
	(0.242)		(0.230)	(0.957)
Inventor-founder		-0.137***	-0.135***	-0.109**
		(0.039)	(0.039)	(0.042)
Inventor-founder=1 × % of self-citations				-1.700*
				(0.994)
R-Sq	0.1	0.1	0.1	0.1
Dep. Var. Mean	0.3	0.3	0.3	0.3
Panel B: Acquisition				
% of self-citations	-0.241*		-0.229*	-0.751*
	(0.134)		(0.133)	(0.398)
Inventor-founder		-0.021	-0.019	-0.029
		(0.027)	(0.027)	(0.030)
Inventor-founder=1 × % of self-citations				0.614
				(0.414)
R-Sq	0.1	0.1	0.1	0.1
Dep. Var. Mean	0.1	0.1	0.1	0.1
Panel C: IPO				
% of self-citations	-0.188		-0.130	-0.801**
	(0.126)		(0.124)	(0.332)
Inventor-founder		-0.098***	-0.097***	-0.109***
		(0.027)	(0.027)	(0.029)
Inventor-founder=1 × % of self-citations				0.790**
				(0.348)
R-Sq	0.1	0.2	0.2	0.2
Dep. Var. Mean	0.1	0.1	0.1	0.1
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	739	739	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. In Panel A, the outcome *Funding* is an indicator variable equal to 1 if the startup raised more than $US 10 million within the first 5 years of inception. In Panel B, the outcome *Acquisition* is an indicator variable equal to 1 if the startup is acquired. In Panel C, the outcome *IPO* is an indicator variable equal to 1 if the startup went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

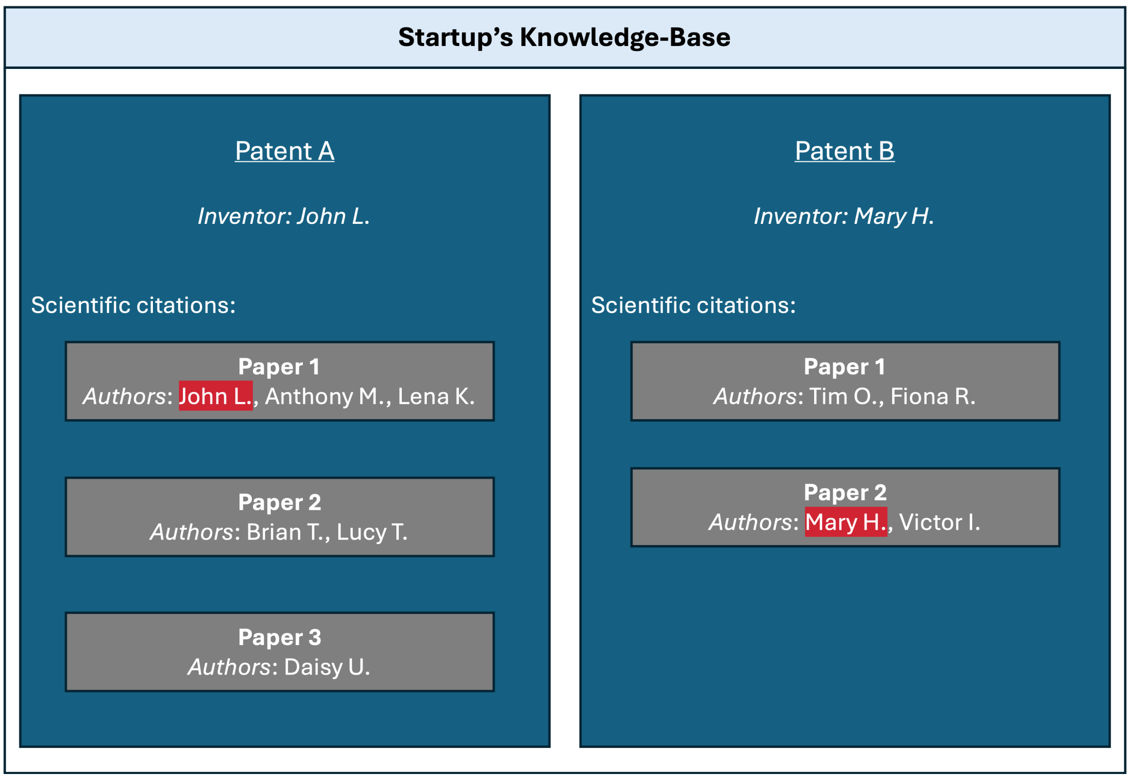
Online Appendix for: Bringing Science to Market: Knowledge Foundations, Inventor-Founders, and Performance

Figure A1: Conceptual elements



Notes: This figure displays the main conceptual elements used in this paper. We think of startups as a vehicle for commercializing a core technology. We conceptualize a startup's knowledge-base – as represented by the dashed circle – as the body of scientific and technical knowledge that informs its core technology. We define a startup's knowledge-base has being scientifically specialized if its technology heavily relies on the inventor's own scientific work. A startup has an inventor-founder if the inventor of the startup's core technology is also the founder of the startup.

Figure A2: Calculation of the degree of specialized scientific knowledge - Example



Notes: This figure shows with a simple example how we calculate the extent to which a startup's knowledge-base relies on specialized scientific knowledge. In this example, the startup's knowledge-base is made of 2 patents: patent A and patent B. Patent A cites 3 scientific articles. One of this article has patent A's inventor as an author. The share of self-citations of patent A is therefore 1/3. Patent B cites 2 scientific articles. One of this article has patent B's inventor as an author. The share of self-citations of patent B is therefore 1/2. The overall degree of reliance of specialized scientific knowledge for the startup is (1/3+1/2)/2=5/12

Table A1: Summary of some theoretical arguments about exit performance outcomes

Theoretical Arguments	*Specialized Scientific Knowledge*
Innovation potential: Deep insights into a specific domain enable inventions that address complex technical challenges effectively, fostering groundbreaking innovations (Cohen & Levinthal, 1990; Kaplan & Vakili, 2015).	+
Expertise: Signals technical expertise and credibility to external stakeholders, increasing attractiveness for funding and partnerships and allowing a better development process (Stuart, 2000; Fleming 2001).	+
Idiosyncratic resource: Specialized knowledge might leverage more tacit elements, making it a unique and hard-to-replicate resource (Teece, 1986; Collis & Montgomery 2008).	+
Engagement with stakeholders: Specialized knowledge is harder to communicate and codify because of its tacit components, creating frictions with external stakeholders (Stuart et al., 1999; Junkunc & Eckhardt, 2009).	-
Integration into technological solutions: Specialized scientific knowledge may be less adaptable and less aligned with market demand (Kogut & Zander, 1992; Klepper, 2001).	-
Scalability: Specialized scientific knowledge may result in a narrower knowledge-base, limiting technological scalability and appeal for acquirers (Puranam, 2009; Makri et al., 2010).	-

Notes: This table summarizes the main theoretical arguments underlying the relationship between specialized scientific knowledge and performance outcomes. + denotes that the relationship is expected to be positive and - denotes that the relationship is expected to be negative.

Table A2: The relationship between inventor-founder and the degree of scientific specialization

	% self-citations	
	(1)	(2)
Inventor-founder	0.009***	0.007**
	(0.003)	(0.004)
Academic startup		0.021***
		(0.005)
# of first patents (log)		-0.002
		(0.002)
Industry experience		-0.004
		(0.003)
Founding team size (log)		-0.013***
		(0.004)
At least one female founder		-0.002
		(0.004)
Entrepreneurship experience		-0.008*
		(0.004)
At least one PhD founder		0.011***
		(0.004)
At least one MD founder		-0.000
		(0.004)
Avg team exp		0.000
		(0.000)
Avg team exp squared		0.000
		(0.000)
Sector FE	Yes	Yes
Founding Year FE	Yes	Yes
State FE	Yes	Yes
R-Sq	0.1	0.1
Dep. Var. Mean	0.02	0.02
Observations	1001	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome *% self-citations* captures the degree of scientific specialization of startups' knowledge-base. It is constructed as follows: we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A3: Success on the exit market - Detailed regressions

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.448**		-0.383**	-1.701***
	(0.190)		(0.186)	(0.449)
Inventor-founder		-0.111***	-0.108***	-0.132***
		(0.034)	(0.034)	(0.037)
Inventor-founder=1 × % of self-citations				1.551***
				(0.474)
Academic startup	-0.103***	-0.105***	-0.097**	-0.105***
	(0.039)	(0.039)	(0.039)	(0.039)
# of first patents (log)	0.057***	0.053***	0.052***	0.051***
	(0.014)	(0.014)	(0.014)	(0.014)
Industry experience	-0.004	-0.003	-0.004	-0.005
	(0.030)	(0.030)	(0.030)	(0.030)
Founding team size (log)	0.063*	0.072**	0.067*	0.065*
	(0.035)	(0.035)	(0.035)	(0.035)
At least one female founder	-0.109***	-0.107***	-0.108***	-0.108***
	(0.033)	(0.032)	(0.032)	(0.032)
Entrepreneurship experience	-0.019	-0.024	-0.027	-0.024
	(0.041)	(0.040)	(0.040)	(0.040)
At least one PhD founder	0.043	0.053	0.057	0.060
	(0.039)	(0.038)	(0.039)	(0.039)
At least one MD founder	0.078**	0.081**	0.081**	0.082**
	(0.035)	(0.035)	(0.035)	(0.035)
Avg team exp	0.003	0.003	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Avg team exp squared	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2
Observations	739	739	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A4: Success on the exit market - Summary, Founding Year × Sector FE

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.457**		-0.395**	-1.717***
	(0.198)		(0.193)	(0.455)
Inventor-founder		-0.110***	-0.107***	-0.131***
		(0.034)	(0.034)	(0.037)
Inventor-founder=1 × % of self-citations				1.555***
				(0.482)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
Sector FE × Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2
Observations	739	739	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A5: Characterization of specialized scientific knowledge - Definition of variables

Variable	Definition	Empirical calculation
Concepts	This variable captures the breadth of a paper, as defined by the diversity of topics it entails	The number of unique concepts associated with a paper's abstract.
Patent Cites	This variable captures how easy the knowledge embodied in a paper can be integrated into a technological solution	Citations received by the paper from patents.
JCIF	This variable captures the 'appliedness' of the journal in which a paper is published, providing an indirect indication of whether the research tends to be more basic or applied	Journal Commercial Impact Factor of the journal where the paper is published. The JCIF of a journal in year t is calculated as the number of times articles from years $t-1$ and $t-2$ were cited by patents during year t, divided by the number of articles published during years $t-1$ and $t-2$.
Claims	This variable captures the breadth of a patent and its commercial application	The number of claims associated with a patent.
Inventors	This variable captures the degree to which the tacit knowledge embedded in a patent has had the opportunity to be shared and codified	The number of inventors associated with a patent.

Notes: This table describes the main variables we use to characterize specialized scientific knowledge.

Table A6: Summary Statistics - Mechanisms

	Min	Mean	SD	Max
A: Characteristics of the papers				
# of unique concepts	1.0	39.9	15.9	110.2
Citations from patents	0.0	1.9	2.1	18.0
JCIF	0.0	1.0	0.6	12.2
B: Characteristics of the patents				
# of claims	1.0	19.6	10.9	115.0
# inventors	1.0	3.3	1.8	16.0
C: Performance outcomes				
Acquired (0/1)	0.0	0.1	0.3	1.0
IPO (0/1)	0.0	0.1	0.3	1.0
Funds raised within 5y (USD million)	0.0	14.7	37.6	507.0
Observations	1006			

Notes: Panel A shows summary statistics for the main variables we use to characterize papers. Panel B shows summary statistics for the main variables we use to characterize patents. Panel C shows summary statistics for different performance outcomes. All variables are averaged at the startup level.

Table A7: Success on the exit market and characteristics of specialized scientific knowledge

	Success			
	(1)	(2)	(3)	(4)
% of self-citations	-0.411**	-0.405**	-0.440**	-0.370**
	(0.195)	(0.185)	(0.189)	(0.183)
# of unique concepts	0.001			
	(0.001)			
Citations from patents		0.017*		
		(0.009)		
# of claims			0.001	
			(0.001)	
# inventors				0.016**
				(0.008)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2
Observations	659	662	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. To calculate *# of unique concepts*, we calculate for each paper cited by a patent the number of unique concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. To calculate *Citations from patents*, we calculate the number of patent citations received by each paper cited by a patent, normalized to account for publication year and field and excluding citations from the focal startup's patents and average this measure at the startup level. To calculate *# of claims*, we calculate the number of claims associated with each patent and average this measure at the startup level. To calculate *# of inventors*, we calculate the number of inventors associated with each patent and average this measure at the startup level. Robust standard errors (in parentheses) are clustered at the startup level.

Table A8: Characteristics of specialized scientific knowledge - Logged outcomes

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-0.397	-0.732**	0.189	-0.249	-1.474***
	(0.327)	(0.337)	(1.131)	(0.338)	(0.322)
R-Sq	0.2	0.2	0.2	0.1	0.1
Panel B: Inventor-Founder					
Inventor-founder	0.007	-0.000	-0.108**	0.009	-0.051
	(0.044)	(0.043)	(0.046)	(0.053)	(0.045)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.1	0.1
Dep. Var. Mean	3.6	0.9	-0.1	2.9	1.1
Observations	659	662	652	739	739

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. Outcome in column (2) is calculated as log(1+X), all other outcomes are logged. In column (1), we calculate for each paper cited by a patent the unique number of concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, normalized to account for publication year and field and excluding citations from the focal startup's patents and average this measure at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A9: Success on the exit market - Summary, Academic startups only

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.365**		-0.232	-1.806**
	(0.179)		(0.166)	(0.819)
Inventor-founder		-0.125**	-0.119**	-0.142**
		(0.053)	(0.053)	(0.056)
Inventor-founder=1 $\times$ % of self-citations				1.702**
				(0.856)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.3	0.3	0.3
Dep. Var. Mean	0.1	0.1	0.1	0.1
Observations	315	315	315	315

Notes: The unit of observation is a startup. We restrict the sample to academic startups, defined as startups with at least one professor in the founding team. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A10: Success on the exit market - Summary with Founding Year × Sector FE, Academic startups only

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.333*		-0.215	-2.153**
	(0.199)		(0.186)	(0.831)
Inventor-founder		-0.118**	-0.113**	-0.141**
		(0.053)	(0.054)	(0.057)
Inventor-founder=1 × % of self-citations				2.089**
				(0.880)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
Sector FE × Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.3	0.3	0.3	0.3
Dep. Var. Mean	0	0	0	0
Observations	315	315	315	315

Notes: The unit of observation is a startup. We restrict the sample to academic startups, defined as startups with at least one professor in the founding team. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A11: Characteristics of specialized scientific knowledge - Academic startups

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-17.305*	-2.504**	3.186	-14.933*	-5.424***
	(9.975)	(1.144)	(3.255)	(8.381)	(1.451)
R-Sq	0.2	0.2	0.2	0.1	0.2
Panel B: Inventor-Founder					
Inventor-founder	0.616	-0.110	0.037	-5.872**	-0.198
	(2.272)	(0.233)	(0.067)	(2.364)	(0.256)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.1	0.1	0.2
Dep. Var. Mean	43.1	2.0	1.1	20.7	3.5
Observations	302	301	298	315	315

Notes: The unit of observation is a startup. We restrict the sample to academic startups, defined as startups with at least one professor in the founding team. The observation count shown is the count with remaining variation after fixed effects are included. In column (1), we calculate for each paper cited by a patent the unique number of concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, excluding citations from the focal startup patents. We normalize this measure following Perry and Reny (2016) to account for paper publication year and field, and average it at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A12: Characteristics of specialized scientific knowledge (logged outcomes) - Academic startups

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-0.603*	-0.784*	0.279	-0.263	-1.689***
	(0.332)	(0.408)	(1.158)	(0.339)	(0.374)
R-Sq	0.2	0.3	0.2	0.2	0.2
Panel B: Inventor-Founder					
Inventor-founder	0.052	-0.036	0.026	-0.166*	-0.028
	(0.073)	(0.066)	(0.062)	(0.088)	(0.071)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2	0.2
Dep. Var. Mean	3.7	1.0	-0.0	2.9	1.1
Observations	302	301	297	315	315

Notes: The unit of observation is a startup. We restrict the sample to academic startups, defined as startups with at least one professor in the founding team. The observation count shown is the count with remaining variation after fixed effects are included. Outcome in column (2) is calculated as log(1+X), all other outcomes are logged. In column (1), we calculate for each paper cited by a patent the unique number of concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, excluding citations from the focal startup patents. We normalize this measure following Perry and Reny (2016) to account for paper publication year and field, and average it at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A13: Success on the exit market - Summary, startups relying on science

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.433**		-0.361*	-1.697***
	(0.190)		(0.184)	(0.463)
Inventor-founder		-0.117***	-0.114***	-0.139***
		(0.036)	(0.036)	(0.039)
Inventor-founder=1 × % of self-citations				1.573***
				(0.490)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2
Observations	672	672	672	672

Notes: The unit of observation is a startup. We exclude startups whose patents do not reference any scientific literature. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A14: Success on the exit market - Summary with Founding Year × Sector FE, Startups relying on science

	Success (0/1)			
	(1)	(2)	(3)	(4)
% of self-citations	-0.462**		-0.391**	-1.716***
	(0.201)		(0.194)	(0.468)
Inventor-founder		-0.117***	-0.114***	-0.139***
		(0.036)	(0.036)	(0.039)
Inventor-founder=1 × % of self-citations				1.559***
				(0.497)
Sector FE	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes
Sector FE × Founding Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2
Observations	672	672	672	672

Notes: The unit of observation is a startup. We exclude startups whose patents do not reference any scientific literature. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A15: Characteristics of specialized scientific knowledge - Startups relying on science

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-16.728*	-2.644***	2.577	-13.982*	-4.863***
	(9.542)	(0.950)	(2.828)	(7.312)	(1.229)
R-Sq	0.2	0.1	0.1	0.1	0.1
Panel B: Inventor-Founder					
Inventor-founder	-0.332	0.041	-0.064*	-1.159	-0.164
	(1.312)	(0.152)	(0.038)	(1.177)	(0.158)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.1	0.1	0.1	0.1
Dep. Var. Mean	40.9	1.9	1.0	20.5	3.4
Observations	659	662	656	672	672

Notes: The unit of observation is a startup. We exclude startups whose patents do not reference any scientific literature. The observation count shown is the count with remaining variation after fixed effects are included. In column (1), we calculate for each paper cited by a patent the unique number of concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, normalized to account for publication year and field and excluding citations from the focal startup's patents and average this measure at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A16: Characteristics of specialized scientific knowledge (logged outcomes) - Startups relying on science

	Papers			Patents	
	(1) Concepts	(2) Patent Cites	(3) JCIF	(4) Claims	(5) Inventors
Panel A: Specialization					
% of self-citations	-0.397	-0.732**	0.189	-0.478	-1.534***
	(0.327)	(0.337)	(1.131)	(0.328)	(0.347)
R-Sq	0.2	0.2	0.2	0.1	0.1
Panel B: Inventor-Founder					
Inventor-founder	0.007	-0.000	-0.108**	0.001	-0.073
	(0.044)	(0.043)	(0.046)	(0.049)	(0.047)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.1	0.1
Dep. Var. Mean	3.6	0.9	-0.1	2.9	1.1
Observations	659	662	652	672	672

Notes: The unit of observation is a startup. We exclude startups whose patents do not reference any scientific literature. The observation count shown is the count with remaining variation after fixed effects are included. Outcome in column (2) is calculated as log(1+X), all other outcomes are logged. In column (1), we calculate for each paper cited by a patent the unique number of concepts it entails and take the maximum of this measure at the patent level. We then average this measure at the startup level. In column (2), we calculate the number of patent citations received by each paper cited by a patent, normalized to account for publication year and field and excluding citations from the focal startup's patents and average this measure at the startup level. In column (3), the outcome is the average JCIF value, normalized to account for publication year and field, of papers cited by a startup's patents. In column (4), we calculate the number of claims associated with each patent and average this measure at the startup level. In column (5), we calculate the number of inventors associated with each patent and average this measure at the startup level. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founder(s) as inventor. Robust standard errors (in parentheses) are clustered at the startup level.

Table A17: Success on the exit market controlling for scientific potential - Summary

	Success (0/1)				
	(1)	(2)	(3)	(4)	(5)
% of self-citations	-0.513***		-0.439**	-1.713***	-1.923***
	(0.188)		(0.184)	(0.441)	(0.455)
Inventor-founder		-0.117***	-0.114***	-0.138***	-0.137***
		(0.036)	(0.036)	(0.040)	(0.040)
Inventor-founder=1 × % of self-citations				1.507***	1.745***
				(0.471)	(0.487)
Log(1+Sc. Cites)	-0.075	-0.080*	-0.085*	-0.084*	-0.083*
	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)
JIF	-0.038	-0.016	-0.019	-0.013	-0.012
	(0.060)	(0.060)	(0.060)	(0.060)	(0.059)
Sector FE	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
R-Sq	0.2	0.2	0.2	0.2	0.2
Dep. Var. Mean	0.2	0.2	0.2	0.2	0.2
Observations	658	658	658	658	658

Notes: The unit of observation is a startup. The observation count shown is the count with remaining variation after fixed effects are included. The outcome is an indicator variable equal to 1 if the startup was acquired or went public via an IPO. To calculate the variable *% of self-citations*, we calculate for each patent the share of scientific articles it cites where at least one inventor of the patent is also an author of the cited article. We normalize this measure by dividing it by the number of inventors on the patent and average it at the startup level. The variable *Inventor-founder* is an indicator equal to 1 if more than half of the patents have one of the founders as inventor. All columns include controls related to the founding team and are described in section 4.1. Column (5) adds the log of funding received within 5 years of inception. The variable *Sc. Cites* represents the number of scientific citations received by papers cited by a startup's patents, normalized to account for publication year and field. The variable *JIF* represents the average JIF value of journals in which papers cited by a startup's patents are published, normalized to account for publication year and field. Robust standard errors (in parentheses) are clustered at the startup level.

A Dimensions AI and Concepts

Definition

Dimensions AI categorizes and organizes vast amounts of academic research data using machine learning techniques. One of its key features that we utilize in this paper is its ability to extract and identify "concepts" from scientific texts.

Concept Identification: Dimensions AI scans through publications' abstracts and identifies "concepts" associated with each paper. These concepts are meant to represent key themes or topics within a paper. They are derived from the text using machine learning models that recognize important keywords, phrases, and themes commonly associated with specific scientific domains.

Relevance Scoring: Each concept identified by Dimensions is assigned a relevance score, which indicates how central the concept is. While the exact scoring scheme is proprietary, it is a function of the field of study of the paper.

We present two examples to illustrate the use of concepts. Both papers were published in 1995.

Paper 1:

- Abstract: *We have synthesized a highly fluorescent (quantum yield 0.88) guanosine analog, (3-methyl-8-(2-deoxy-beta-D-ribofuranosyl) isoxanthopterin (3-Mi) in a dimethoxytrityl, phosphoramidite protected form, which can be site-specifically inserted into oligonucleotides through a 3',5'-phosphodiester linkage using an automated DNA synthesizer. Fluorescence is partially quenched within an oligonucleotide and the degree of quench is a function of the fluorophore's proximity to purines and its position in the oligonucleotide. As an example of the potential utility of this class of fluorophores, we developed a continuous assay for HIV-1 integrase 3'-processing reaction by incorporating 3-MI at the cleavage site in a double-stranded oligonucleotide identical to the U5 terminal sequence of the HIV genome. Integrase cleaves the 3'-terminal dinucleotide containing the fluorophore, resulting in an increase in fluorescence which can be monitored on a spectrofluorometer. Substitution of the fluorophore for guanosine at the cleavage site does not inhibit integrase activity. This assay is specific for the 3'-processing reaction. The change in fluorescence intensity is linear over time and proportional to the rate of the reaction. This assay demonstrates the potential utility of this new class of fluorophore for continuous monitoring of protein/DNA interactions.*
- Concepts: guanosine analog, guanosine, reaction, analogs, incorporation, HIV-1
- Breadth (as calculated with our measure): 6

Paper 2:

- Abstract: *An approach for genome analysis based on sequencing and assembly of unselected pieces of DNA from the whole chromosome has been applied to obtain the complete nucleotide sequence (1,830,137 base pairs) of the genome from the bacterium Haemophilus influenzae Rd. This approach eliminates the need for initial mapping efforts and is therefore applicable to the vast array of microbial species for which genome maps are unavailable. The H. influenzae Rd genome sequence (Genome Sequence DataBase accession number L42023) represents the only complete genome sequence from a free-living organism.*
- Concepts: Haemophilus influenzae Rd, genome sequence, complete nucleotide sequence, free-living organisms, whole genome, genome mapping, genomic analysis, nucleotide sequence, whole chromosomes,

microbial species, genome, sequence, mapping efforts, assembly, chromosome, bacterium, DNA, species, random sequence, organization, maps, analysis, efforts

- Breadth (as calculated with our measure): 23