Received: 17 February 2023 | Revised: 23 September 2024 | Accepted: 14 November 2024

DOI: 10.1002/smj.3685

RESEARCH ARTICLE

SMS | Strategic Management Journal

WILEY

Free range startups? Market scope, academic founders, and the role of general knowledge in AI

Shinjinee Chattopadhyay[1] | **Florence Honoré**[2] | **Shinjae Won**[3]

[1]Gies College of Business, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

[2]Wisconsin School of Business, University of Wisconsin-Madison, Madison, Wisconsin, USA

[3]School of Labor and Employment Relations, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

Correspondence
Shinjinee Chattopadhyay, Gies College of Business, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA.
Email: schattop@illinois.edu

Abstract

Research Summary: High-tech startups develop technologies, the market applicability of which can vary widely, enabling startups to target a range of market segments. Using a question-driven approach to contrast startups with and without academic founders, we investigate the difference in the market applicability between the two groups on a sample of 988 startups in the artificial intelligence (AI) field. Our findings reveal that academics' pursuit of basic research drives the creation of general knowledge, which in turn leads to wider market applicability. With fewer requirements for complementary downstream assets in the AI ecosystem, academics can more easily translate their general ideas to market applications and locate downstream in the value chain. Our findings highlight the role of problem-formulation and -solving in startups and of academic startups within AI.

Managerial Summary: Using a sample of 988 startups in the Artificial Intelligence field, we find that startups with at least one academic on their founding team are associated with a higher number of verticals (potential market segments for the technology the startups developed) compared to startups without any academics. Teams with academic founders produce more general publications and patents than others, which drives the association with more verticals. Academics formulate and solve more general problems relative to non-academics, leading to the creation of more general products that are applicable to a broader range of verticals. With fewer requirements for complementary downstream assets in the AI ecosystem, academics can more easily translate their general ideas to market applications and locate downstream in the value chain.

KEYWORDS

academic entrepreneur, entrepreneurial team, generality, innovation, market opportunities, problem-solving perspective

1 | INTRODUCTION

The potential market scope of a startup is shaped by the technology created within its bounds (Kim et al., 2023; Moeen & Agarwal, 2017; Shermon & Moeen, 2022) and is an important determinant of its performance and survival (Gruber, 2010; Gruber et al., 2012; Nerkar & Shane, 2003). Entrepreneurs face a tradeoff between developing technologies that apply to fewer or more markets.[1] While technologies applicable to fewer markets lower overhead costs and may be more accessible for startups that typically cannot leverage economies of scope (Giustiziero et al., 2023; Levinthal & Wu, 2010), being constrained to fewer markets increases risk and may necessitate frequent pivoting as startups mature (Kirtley & O'Mahony, 2023; McDonald & Gao, 2019). Developing broad-scoped technologies applicable to multiple markets, by contrast, allows them to potentially navigate demand uncertainty in single markets (Penrose, 1959) and obtain higher sales (Gruber et al., 2008) but also requires a wider range of complementary assets and the existence of broad demand in downstream markets (Conti et al., 2019).

The artificial intelligence (AI) ecosystem[2] provides a context where some of the traditional barriers to entry are relaxed and can serve as a setting for scholars seeking new theoretical understanding into factors driving startups' creation of technologies and their market applicability. In contrast to traditional high-tech industries such as lasers or biotech, where firms need substantial assets to enter multiple downstream markets, the largely digital nature of the AI-based technology (Giustiziero et al., 2023) creates different considerations for multiple market entry. The need for critical assets—such as substantial computing power and large data sets—is

[1]Throughout the paper, we use the term *market applicability* to refer to the number of market segments the startups could potentially enter based on the technologies they have developed; in other words, we refer to the startups' potential market scope.

[2]AI ecosystem refers to the technologies, tools, stakeholders, and resources that contribute to the development, deployment, and advancement of AI (Jacobides et al., 2021). We provide more detail about this context and the period we study in Section 1 of the Appendix.

concentrated upstream (Jacobides et al., 2021). Hence, there is a lower need for startups to possess specialized assets to enter multiple downstream markets, and these lower barriers can potentially create new incentives for startups to develop technologies with wide market scope.

What are the implications of AI's structural features for how different kinds of entrepreneurs define and pursue markets? Academic entrepreneurs,[3] in particular, may be uniquely positioned due to their domain-specific expertise and their foundational contributions to the field of AI (Gofman & Jin, 2024; Jacobides et al., 2021; Wooldridge, 2020). Their typical weakness, moreover, —the lack of access to complementary assets such as operational knowledge, distribution networks and industry connections (e.g., Agarwal & Shah, 2014; Park et al., 2024)—is less critical in AI as the lower barriers to entry to downstream markets may benefit them more relative to other entrepreneurs. Their nonacademic counterparts, by contrast, may have other strengths, such as their industry-specific knowledge and operational and managerial know-how, that allow them to benefit from lowered barriers to entry within AI and offer products to many existing markets (e.g., Gruber, 2010). Given these important differences between academics' and nonacademics' pre-entry knowledge and capabilities, we test whether and how academic startups (i.e. founded by at least one academic) differ from nonacademic startups in the breadth of potential market applicability for the technologies they develop.

Examining 988 new ventures within the field of AI, we find that academic startups are associated with 12% higher breadth of market applicability on average than nonacademic startups. This is a surprising finding given long-standing evidence on academics' tendency to develop specialized expertise within fewer domains (Leahey, 2006, 2007) and their lack of operational and market knowledge in different domains (e.g., Vohora et al., 2004). To explain this unexpected finding, we conduct a question-driven inquiry and shed light on a process through which entrepreneurs' pre-entry knowledge is translated to their startups' market applicability.

We draw upon different streams of literature to identify two channels of influence on entrepreneurs' market applicability of their technologies. The first channel pertains to academics' and nonacademics' different approaches toward commercialization of their technologies. Academics tend to license their technologies early on (Agarwal, 2006; Jensen & Thursby, 2001) and may participate in market for technologies, thereby reaching a broader set of markets through this channel than nonacademics (Arora et al., 2001; Conti et al., 2019; Gambardella & Giarratana, 2013). The second channel of influence relates to different approaches toward knowledge creation. Academics are driven by their pursuit of "knowledge for the sake of knowledge," and are guided by scientific principles in this pursuit (Agarwal & Ohyama, 2013; Kim et al., 2023; Stephan, 1996; Stokes, 1997), while nonacademics are driven by commercial priorities (Sauermann & Stephan, 2013). These two differing approaches to knowledge creation may create differences in the kinds of technologies they develop, leading to differences in their startups' potential market scope.

Exploring the first plausible channel of influence, we find that academic startups do not license or sell their technologies to a significantly greater extent than nonacademic startups. Although past literature has found that in other industries academic entrepreneurs are more likely to be upstream than nonacademics, 86% of academic startups in AI position their business downstream. Further, location within the ecosystem does not explain differences in the breadth of market applicability across academic and nonacademic new ventures.

[3]An academic founder is defined as a founder who was working at a university as a professor, research associate, research scientist or post doc at the time of founding (e.g., Roche et al., 2020; Shane, 2004).

Examining the second channel of influence, we find that academic startups conduct more basic research relative to applied research than nonacademics, and produce more general (widely cited across domains) publications and patents than nonacademic startups. Importantly, we find that basic research leads to the creation of general knowledge, and the generality of the knowledge produced explains the association between academic startups and market applicability. Through several interviews with entrepreneurs, academics, and industry-employees, we discover that academics pursue and attempt to solve general problems[4] and develop more general solutions. Nonacademic founders, by contrast, tend to pursue more narrow-scoped problems driven by specific customers' demands and needs, leading to more targeted applicability of resulting products to fewer market segments.

Our paper uncovers a process through which founders' pre-entry knowledge translates into market applications, by showing plausible evidence on the links between founders' choice of problems (i.e., idea selection), the knowledge they create, and the potential market scope of the firm. Starting from their immersion in basic research, academic entrepreneurs follow a "science first" approach toward knowledge creation within their startups, enabling them to formulate more general problems and create more general solutions. With fewer requirements for complementary downstream assets in the AI ecosystem, academics can more easily translate their general ideas to market applications with a wide market scope. In contrast, nonacademics adopt a more demand-pull approach, where they focus on problems and solutions that are targeted toward meeting a more specific need, leading to a narrower market scope.

Our paper offers three main contributions. First, we show how academics translate their general ideas to market applications, thereby potentially reaching a wider market scope in the AI ecosystem. In contrast to other industries, where academic startups tend to locate upstream, within AI, they cluster downstream. The structural features of the AI ecosystem allow academics to leverage their distinct problem-formulation approach, thereby positioning their startups differently from what they have traditionally done in other high-tech industries. Second, our findings speak to the emerging literature stream on enabling technologies (Conti et al., 2019; Gambardella et al., 2021; Gambardella & Giarratana, 2013) by providing a snapshot of early entrepreneurial activity occurring in an emerging digital ecosystem. This study may set the stage for future research on the antecedents of such technologies and the development of their entrepreneurial ecosystem. Third, our paper draws attention to a dimension of entrepreneurial backgrounds that has not been explored within extant literature—their problem-formulation and solving approach (Nickerson et al., 2012; Nickerson & Zenger, 2004). While Shane (2000) suggests entrepreneurship arises from discovering opportunities through new technologies, our evidence presents a complementary perspective: some entrepreneurs develop technologies first, then explore market opportunities, yet still aim for broad market potential. This micro-foundational view can complement prior research on academic entrepreneurship and offers future opportunities for theorization.

In the following sections, we begin with the motivation for studying the market applicability of startups' technology in AI and review the importance of academic entrepreneurs in this context. We present our main results and identify two relevant channels that may lead to our findings. We then unpack the most plausible path these entrepreneurs pursue through empirical

[4]A *general problem* refers to a problem that is stripped of specific details and context. A general problem formulation focuses on the underlying structure or principles that define it, and allows the problem to be framed in a way that the solution can be applied to various domains or scenarios (see Section 5.2 for an example).

tests and interviews. We discuss our contributions to multiple literature streams, present the limitations of our study, future research directions, and conclude.

2 | THEORETICAL MOTIVATION IN AI

2.1 | Market applicability and AI

It is a priori unclear whether it is advantageous for startups to have a wide or narrow potential market scope. Technologies specialized for single markets can lower costs for firms, capture upsides from positive network effects, and avoid potential diseconomies of scope resulting from seeking applicability within multiple market segments (Giustiziero et al., 2023; Levinthal & Wu, 2010) but such technologies potentially limit flexibility. Broad applicability helps mitigate uncertainty and increase sales potential (Gruber et al., 2010), though it requires more complementary assets and the existence of broad demand in the downstream markets (Conti et al., 2019).

In the AI ecosystem, startups face fewer constraints compared to non-digital industries, allowing them to occupy different positions within the industry value chain. The digital nature of AI technology means fewer specialized downstream assets are needed to enter multiple markets, unlike sectors like biotechnology or manufacturing, where co-specialized equipment and physical infrastructure are essential. This renders downstream markets more accessible to new entrants in AI compared to traditional industries.

The context of AI offers a snapshot of an emerging digital ecosystem that is a hotbed of significant entrepreneurial activity—a context that is largely unexplored. AI draws from science in many different areas and can potentially impact many disparate domains, from finance, or health, human resources to the life sciences (Zhang et al., 2021) and attract entrepreneurs from different disciplines. Navigating this emerging ecosystem and establishing new ventures may entail going beyond understanding market needs, as the science behind AI requires advanced knowledge and expertise that is not readily available to all kinds of entrepreneurs (Gofman & Jin, 2024; Jacobides et al., 2021).

These features distinguish the AI ecosystem as a context that has the potential to challenge common assumptions in the innovation and entrepreneurship literature and lay a fertile ground for researchers to gain new insights into how startups' technologies guide their market applicability.

2.2 | Entrepreneurs and the AI ecosystem

Entrepreneurs' pre-entry experiences shape the outcomes of the ventures they found (e.g., Helfat & Lieberman, 2002; Beckman & Burton, 2008; Bercovitz & Feldman, 2008; Chen et al., 2012; Honoré, 2022). Within emerging industries in particular, academic entrepreneurs have traditionally represented a key actor by making foundational contributions and shaping the early stages of many industries (e.g., Agarwal & Shah, 2014; Moeen & Agarwal, 2017; Roche, 2023; Roche et al., 2020). Academics are a distinct type of entrepreneur, having little in common in their pre-entry experiences with other kinds of entrepreneurs, whether they originate from the industry, or are user entrepreneurs (Agarwal & Shah, 2014; Park et al., 2024). Academics' typical motivation is oriented toward their pursuit of knowledge to achieve scientific recognition (Agarwal & Ohyama, 2013; Merton, 1957; Roach & Sauermann, 2015; Stephan, 1996), which often persists after founding their own ventures (Haeussler & Colyvas, 2011).

In the context of AI, academics have driven and continue to contribute to many advancements (Gofman & Jin, 2024; Wooldridge, 2020). They more easily meet the required expertise threshold compared to other kinds of entrepreneurs and may be able to leverage this expertise in their startups. By contrast, their nonacademic counterparts might be in a better position to leverage a combination of market knowledge and practical technical advancement in their startups. This comparison can provide insights on what kinds of founders' pre-entry experience and knowledge shape the market applicability of startups in the emerging field of AI.

3 | DATA AND METHODOLOGY

3.1 | Sample selection strategy

To examine this question, we utilized data from various sources, including Crunchbase, LinkedIn, and PitchBook.[5] From Crunchbase, we collected the entire list of startups founded post-2002 in the AI field that are headquartered in the United States by using the keyword *AI*. Among the US-based teams, we found 5449 such startups.[6]

Since Crunchbase contains the founders' LinkedIn addresses, we used these addresses to collect their career history data from public LinkedIn profiles. We complemented the missing addresses with a manual search on LinkedIn based on the names of the founders as well as the organization name and the founded year (e.g., Reese et al., 2021) and supplemented with a web search when needed. To retain each startup in our dataset, we had to find the career history data for all founding members; this was not always possible. We were able to collect the career history data for 66.6% of associated founders, which resulted in 2235 founding teams (41% of the initial list of 5449 teams) with full employment and educational history data for all members. From this data, we captured their employment and educational background. Finally, we turned to PitchBook to get funding data and information on market applicability. Those not found in the PitchBook database had to be dropped from our sample. We were eventually left with 1696 startups. Less than half of them were created by solo founders ($n = 708$). We excluded the solo-founder startups since prior literature established that the founding conditions of firms and founders' objectives may differ for startups led by solo founders vs. teams (e.g., Agarwal et al., 2016; Delmar & Shane, 2006; Lee et al., 2024) and because we found fundamental differences between solo and non-solo founders in our data and context. We detailed

[5]We collected Crunchbase and LinkedIn data during the Summer and Fall 2020 and PitchBook during Fall 2021. Our choice of datasets was guided by established practices among scholars of strategic management. Crunchbase has served as a trusted source for identifying representative samples of new ventures and their founders (e.g., Conti & Roche, 2021; Contigiani, 2023). LinkedIn has comprehensive data on individuals' careers which scholars draw from (e.g., Ge et al., 2016). PitchBook collects detailed information on private firms and funding deals with professional investors and is gaining traction in studies on entrepreneurial finance and innovation for growth-oriented firms (e.g., Ewens et al., 2022; Hallen et al., 2023).

[6]There were 79 companies were founded in or prior to year 2002. We removed these companies since we focus on the field of AI as a nascent field and these very small number of older companies may be different in terms of their product or technology even though they had the "AI" in the Crunchbase keyword.

descriptive statistics and our rationale based on interviews as well as past literature in Section 3 of the Appendix. We ended up with a final sample of 988 founding teams with venture-level data as well as career history data for all founding members. To learn about the history and the businesses of a few representative startups refer to Section 4 in the Appendix.

3.2 | Variables

3.2.1 | Dependent variable

Breadth of market applicability

To measure market applicability, we obtained the number of "Verticals" that PitchBook assigned to each new venture based on the applicability of their technology. According to PitchBook, "[A vertical] describes a group of companies that focus on a shared niche or specialized market spanning multiple industries."[7] This definition is in line with prior scholarly work. For instance, Mosakowski (1991) explains as follows: "Vertical markets refer to specific industrial niches for computer products. The breadth of these niches varies considerably, ranging from large niches such as the health-care industry to narrow niches such as the automobile repair industry (p. 119)." Compared to industry sectors, verticals often signify applicability across various industries. For example, within Fintech, one of the verticals, one can find a product that bridges commercial lending applications, insurance products, and financial platforms. PitchBook's list of verticals as a measure of market applicability has gained legitimacy and traction outside of academia as well. National Science Foundation (NSF) refers to the same list of verticals provided by PitchBook within the report prepared by the National Science Board, titled "*Science and Engineering Indicators*."[8] The prominent investment financial services company Morningstar, which is worth approximately $12 billion in market cap,[9] recently introduced indices based on PitchBook's verticals measure, demonstrating that firms are staking future revenue on PitchBook's measure of industry verticals.[10] We provide more information on how PitchBook assigns verticals to firms as well as the complete list of verticals within our sample in Section 5 in the Appendix.

3.2.2 | Independent variable

Having an academic entrepreneur

To identify academic entrepreneurs, we followed the classification procedure by Fuller and Rothaermel (2012) and examined the job titles that individuals held at universities. We coded this dummy variable as 1 if the team had any founder who was working at a university as a

[7]Source: https://pitchbook.com/what-are-industry-verticals; Last accessed on February 6, 2023.
[8]PitchBook verticals can be found within the Technical appendix in NSB-2022-6, Technical Appendix (https://ncses.nsf.gov›pubs›nsb20204-tabs08-063). Last accessed on February 2, 2023.
[9]As of December, 2023.
[10]Source: https://www.prnewswire.com/news-releases/morningstar-and-PitchBook-introduce-new-indexes-to-provide-investors-with-deeper-insight-into-emerging-venture-capital-companies-301989283.html. Last accessed on December 16, 2023.

professor, research associate, research scientist or postdoc[11] at the time of founding and 0 otherwise (e.g., Agarwal & Shah, 2014; Roche et al., 2020; Shane, 2004).[12]

3.2.3 | Control variables

A set of control variables captures the team characteristics: team size, team's average years of experience[13] (the current year minus the start year of the first job), team's average number of employers, whether the team has any nonacademic PhD-holder (who was not working at a university at the time of firm founding, i.e., any PhD holders not qualified as an AE according to the paper's definition),[14] whether the team has any female, whether any founder graduated from a top 10 university, any serial founder (has previous founding experience) or any foreign founder (has an undergraduate degree from outside of United States). We also controlled for startup characteristics measured when we collected data on market applicability (Fall 2021): startup age, startup location (East Coast, West Coast, and the rest), and the number of X (formerly known as Twitter) followers as a proxy for prominence, as certain new ventures receive more spotlight whether it is from their initial success or factors unrelated to performance. For example, teams that include academic entrepreneurs may attract more media attention and have a higher number of verticals being assigned.

4 | MAIN FINDING: ASSOCIATION BETWEEN ACADEMIC STARTUPS AND MARKET APPLICABILITY

4.1 | Descriptive statistics

Table 1 presents the descriptive statistics. In our sample, 83 out of 988 teams (8.4% of teams) have at least one academic entrepreneur. The number of verticals ranges from 0 to 9, with a mean of 3.11 (see Figure A in the Appendix for the distribution of verticals). The average team size is 2.53, and the average firm age is 4.79 years. About one-fifth of new ventures have a female founder and 72% of new ventures have a founder with previous founding experience. The average number of years of work experience across all founders of a team is 18.59.

We also checked for observable differences across teams with an academic entrepreneur and without. The results from the *t* tests are reported in Tables A1 and A2 of the Appendix. We see that the teams with academics are more likely to include a female founder or a foreign

[11]Job titles of postdoc positions include those such as postdoctoral associate, postdoctoral fellow, postdoctoral research associate, postdoctoral research fellow, postdoctoral researcher, and postdoctoral scholar. We also included positions such as research scientist, research associate or data scientist in universities if the individual holds a PhD degree and works in research for a university. We do not include adjunct positions in our definition.

[12]When we restrict the definition to just faculty members (which is a more conservative measure) at the time of founding the firm (6% of our sample), we found that all our results were consistent.

[13]We also examined prior industry in marketing and management, and their interactions with Has an AE but did not find any association with our outcome variable.

[14]Since our independent variable *Has an Academic Entrepreneur* is highly collinear with this variable, we replaced this dummy with zero when the team included an academic entrepreneur. Hence, this variable captures the teams that includes a PhD holder (and is not an academic entrepreneur) and is mutually exclusive from our independent variable *Has an Academic Entrepreneur*.

TABLE 1 Descriptive statistics.

		Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11
1	Number of verticals	3.11	1.42	1.00										
2	Has an academic entrepreneur	0.08	0.28	0.05	1.00									
3	Team size	2.53	0.79	0.04	0.06	1.00								
4	Firm age	4.79	2.39	0.17	0.00	0.06	1.00							
5	Has a female founder	0.23	0.42	−0.03	0.06	0.11	−0.07	1.00						
6	Has a PhD holder (nonacademic)	0.3	0.46	0.08	−0.19	0.12	0.01	0.08	1.00					
7	Has a serial founder	0.72	0.45	0.02	−0.12	0.15	−0.02	−0.02	−0.04	1.00				
8	Average number of employers	5.94	2.25	−0.04	−0.09	0.00	−0.15	0.04	0.01	0.30	1.00			
9	Average experience	18.59	6.71	0.00	0.00	0.00	0.29	0.00	0.03	0.15	0.41	1.00		
10	Has a foreign founder	0.45	0.5	0.04	0.06	0.35	−0.01	0.06	0.08	0.07	−0.05	−0.03	1.00	
11	Has a founder from top 10 inst.	0.1	0.3	0.01	0.11	0.13	−0.01	0.05	0.06	0.02	0.02	0.00	0.06	1.00
12	Log number of X followers	4.17	3.19	0.10	0.02	0.04	0.08	0.01	0.02	0.07	0.01	0.05	−0.03	−0.03

founder, but less likely to include a serial founder. Also, the average number of past employers was lower for teams with academics. We observe a similar pattern when we estimate the likelihood of having an academic entrepreneur as a function of these observables, which we report in Table B1 of the Appendix. Since the teams with and without academics differed on these dimensions, we controlled for these variables in all our models. To further understand which factors determine the likelihood of including an academic entrepreneur, we also used Least Absolute Shrinkage and Selection Operator, following the approach used by Roche (2023) in her work on academic entrepreneurs. We found that most variables were important predictors of the inclusion of academics in teams. This procedure further validates the use of the control variables to address the selection bias (Belloni et al., 2014; Conti & Guzman, 2019).

4.2 | Main findings

We find that teams with at least one academic entrepreneur are associated with higher market applicability. Table 2 shows the results of our OLS regression. In Model 1, we see that teams with an academic entrepreneur have 0.38 more verticals (12% higher than the mean) compared to those without ($\beta = .38$, $p = .02$).[15] In the following sections, we investigate and test explanations for the observed relationship. We have summarized our approach along with our findings in Table 5.

5 | TWO CHANNELS OF INFLUENCE: COMMERCIALIZATION AND KNOWLEDGE CREATION APPROACHES

Based on the distinct features of academics' exploration of knowledge and their activities that set them apart from nonacademics, we draw on prior literature and argumentative reasoning to examine why academic startups are associated with a higher breadth of market applicability. Two literature streams are relevant: the first pertains to startups' approach to commercialization while the second is focused on their approach toward knowledge creation. In the following sections, we examine each of these channels.[16]

5.1 | Theoretical background: Commercialization approach

Academic entrepreneurs often lack downstream complementary assets, such as operational knowledge that are required to create products for downstream markets (Agarwal & Shah, 2014; Park et al., 2024). Their application-agnostic focus toward knowledge creation often renders them lacking in know-how on technology commercialization, markets or management (Vohora et al., 2004). Rather than entering downstream markets with products, they tend to

[15]We checked whether academic startups target different verticals based than nonacademic startups. Table C1 of the Appendix, indicates that the majority of the top 10 verticals (7 out of 10) were the same for both team types, suggesting that academic and nonacademic startups did not target different market segments. We also acknowledge that verticals may vary in the breadth; our results may be discounting the effect of entrepreneurs focusing on a few wide niches and can possibly make our results more conservative.

[16]Table A1 in the Appendix present the descriptive statistics of the variables we collected to investigate each channel.

TABLE 2 The effect of having an academic entrepreneur on market applicability.

	(1) Market applicability
Has an AE	0.38
	[.02]
Team size	−0.02
	[.78]
Firm age	0.11
	[.00]
Has a female founder	−0.10
	[.33]
Has a non-AE PhD	0.30
	[.00]
Has a serial founder	0.13
	[.24]
Average number of past employers	0.00
	[.86]
Average experience	−0.01
	[.06]
Has a foreign founder	0.09
	[.33]
Has a founder from top 10 inst.	−0.01
	[.97]
Log number of X followers	0.04
	[.01]
Constant	2.47
	[.00]
Observations	988
R-squared	.06

Note: *p*-Values in brackets. All models control for location dummies.

participate in the upstream markets for technology (Arora et al., 2001) to either license, lease or sell these technologies to downstream agents (Agrawal, 2006; Jensen & Thursby, 2001; Serrano, 2007, 2010). Since firms specializing upstream can provide technologies to multiple industries and domains (Conti et al., 2019), this body of research would predict that academic startups should be associated with a higher breadth of market applicability. Nonacademics' focus on user-driven or market-ready applications (Shermon & Moeen, 2022), on the other hand, may enable them to hone in on fewer relevant but promising market segments, leading to lower breadth of market applicability. To test whether the position within the industry value chain explains our finding, we conducted tests using three different sources: (1) patent sales, (2) licensing records, and (3) business descriptions.

5.1.1 | Examination of patent sales and licensing

Firms that specialize upstream within the industry value chain are more likely to license or sell their patents (Gambardella et al., 2021; Serrano, 2010; Teece, 1986). To examine this, we constructed a variable *Number of Patents Sold*. We collected data on 1912 patents produced by all new ventures in our sample from the USPTO database.[17] We then examined the patent assignment records from this database to determine whether these patents were sold. In cases where the new venture was acquired by another firm, we only tracked the patents produced by the team before the acquisition. Patents are assigned to the firm holding the patent and so, a change in the assignment reflects a change in ownership of the patent: "Reassignment is the legal transfer of rights, title and interest in a patent, thereby a proxy for activity in the market for technology" (Serrano, 2010). When the assignee of a patent was changed from our focal venture to another entity this indicated that the patent was sold, and we coded it as such.[18]

We found that 29% of our patents were sold. However, we did not find that academic startups sold more patents (Model 1 in Table 3). In fact, academic startups sold fewer (approximately one less on average) patents compared to others ($\beta = -.79$, $p = .05$). We also found that the number of patents sold was not associated with market applicability (Model 2 in Table 3).

Next, we examined whether teams with academics were more likely to license their patents than those without. We utilized tools provided by a private company (KTMine) to examine the number of patents developed by all firms within our sample that were licensed to public firms. We did not identify any patents from our sample as being licensed.[19]

5.1.2 | Examination of business descriptions

Second, we examined all business descriptions of the startups that we obtained from PitchBook to understand the location of the startup in their industry value chain. Firms that specialize upstream may produce technologies that buyers incorporate into their products while those that are downstream will produce end products. To test this idea, we generated a variable *Upstream*. We designated a startup to be upstream if the main activity of the business relates to the development of AI algorithms, models, data and technologies to sell to other firms. Downstream startups implement AI technologies into end-user products. We used fine-tuned ChatGPT to categorize startups' location in the value chain. We have detailed the process of using this method in Section 6 in the Appendix.

We found that 6.4% of all startups are selling technologies used in other products, indicating that they are upstream in the industry value chain. Within the sample of academic

[17]We manually searched for patent data using the venture name, founder name and the location of headquarter on PatentsView as well as Google Patents. We were able to identify 1912 patents across our founders in 264 ventures.

[18]We disregarded instances where the new assignee was a financial institution since this indicates that the patent was assigned to the financial institution as collateral for a loan, or collateral to secure financing, rather than a genuine transfer of the technology to be used by the new owner of the patent (Hochberg et al., 2018).

[19]Academic startups might license their technology to other startups, but we are unable to observe such agreements. Since nonacademic startups do not demonstrate higher market applicability, even if they potentially benefit from licensed academic knowledge, this possibility of knowledge transfer actually makes our empirical tests more conservative.

TABLE 3 Channel 1: Upstream location and market applicability.

	(1) Number of patents sold	(2) Market applicability	(3) Upstream	(4) Market applicability
Has an AE	−0.79	0.35	0.06	0.39
	[.05]	[.04]	[.03]	[.02]
Number of patents sold		−0.02		
		[.21]		
Total number of patents	0.47	0.02		
	[.00]	[.04]		
Upstream location				−0.18
				[.34]
Team size	0.23	−0.02	0.01	−0.01
	[.12]	[.77]	[.20]	[.81]
Firm age	−0.03	0.11	0.00	0.11
	[.51]	[.00]	[.17]	[.00]
Has a female founder	0.08	−0.10	−0.03	−0.11
	[.76]	[.36]	[.07]	[.30]
Has a non-AE PhD	−0.14	0.28	0.07	0.31
	[.56]	[.01]	[.00]	[.00]
Has a serial founder	0.35	0.14	−0.02	0.12
	[.17]	[.20]	[.33]	[.25]
Average number of past employers	0.07	0.01	0.00	0.00
	[.21]	[.75]	[.93]	[.86]
Average experience	−0.03	−0.02	−0.00	−0.02
	[.17]	[.04]	[.05]	[.06]
Has a foreign founder	−0.23	0.09	−0.01	0.09
	[.32]	[.35]	[.38]	[.34]
Has a founder from top 10 inst.	−0.19	−0.01	0.05	0.00
	[.60]	[.96]	[.05]	[.99]
Log number of X followers	−0.10	0.03	0.00	0.04
	[.00]	[.01]	[.68]	[.01]
Constant	−0.25	2.50	0.11	2.49
	[.65]	[.00]	[.01]	[.00]
Observations	988	988	988	988
R-squared	.5	.06	.04	.06

Note: *p*-Values in brackets. All models control for location dummies.

startups, 14% are in upstream locations. Using a linear probability model, we found that, relative to nonacademics, academic startups were 6% more likely to be located upstream (Model 3 in Table 3), but locating upstream was not associated with the breadth of market applicability (Model 4 in Table 3). The coefficient of *Has an academic entrepreneur* remained positively associated with market applicability when we control for the startup position in the value chain.

We further examined the current websites of academic startups.[20] We carefully checked all active websites (71 out of our 83 academic startups have a functioning website). We found that six new ventures are specializing upstream by developing technology that other firms can build on to create their own products, while the remaining 65 firms are developing standalone products that they bring to market directly. Interestingly, five of the six companies that participate in the market for technologies *also* commercialize applications, thereby entering both upstream and downstream.[21]

These three pieces of evidence indicate that there are no discernible differences in the patent sales, patent-licensing rates, or location within the value chain between academic and nonacademic startups that explain the differences in the startups' market applicability.

Therefore, in the next section, we investigate the second channel through which academic entrepreneurs can generate wider market applicability.

5.2 | Theoretical background: Knowledge creation approach

We turn to the second feature that sets academics apart from nonacademics: their knowledge creation process. Academics' key distinguishing feature is their quest for knowledge for the sake of knowledge or research for the sake of furthering the bounds of knowledge rather than seeking a specific application (Agarwal & Ohyama, 2013; Bieniawski, 1994; Nelson, 1959; Rosenberg & Nelson, 1994; Stokes, 1997). This quest is associated with basic research. Indeed, Tijssen (2010, p. 1) states that *"the traditional understanding of 'basic science' centers on knowledge creation within science itself, 'applied science' derives its meaning and relevance from external applications where knowledge and skills are transferred into and further utilized within environments outside the realm of science."* Similarly, the NSF views basic research as: "*Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view.*"[22]

When academics found startups, this approach is likely to inform and shape the technologies they develop, and the markets these technologies apply to. After all, academic entrepreneurs draw upon scientific principles to develop their ideas (e.g., Kim et al., 2023; Roche et al., 2020; Sauermann & Stephan, 2013). We raise the possibility that academics' pursuit of

[20]We examined the academic startups websites in Fall 2023 while we used the Fall 2021 downloads from PitchBook to analyze the business descriptions.

[21]Three of them are working on bringing machine vision to their customers' robots (Numericcal, Neurala, and Voxel51 with an open-source model). 6D.ai rented its platform and tools to developers in the augmented reality domain until its acquisition by Niantic Labs. Deepgram developed a speech recognition and transcription application, sells the features to software companies and makes its API available to developers. Finally, Lightmatter solely occupies an upstream position developing photonic chips to speed AI computation and deep learning.

[22]Source: "At the Foundation of the Foundation: Basic Research at the NSF" https://www.nsf.gov/mps/advisory/mpsac_white_papers/mpsac_basic_research_white_paper.pdf (Last accessed on September 14, 2024).

basic knowledge enables the creation of general knowledge—that is, knowledge that is applicable to a wider variety of domains (Valentini, 2012). Past research has shown that the ability to map and understand the landscape of scientific problems and solutions is key to knowledge production (Fleming & Sorenson, 2004). Firms that rely on scientific knowledge produce general inventions that are applicable across multiple domains (Novelli, 2015). Drawing upon these mechanisms, we contend that academics, with their focus on extending the knowledge frontier through their focus on basic research, may be creating general knowledge within their startups by adopting a general problem formulation and solving approach.

A *general problem formulation* refers to a formulation that is stripped of specific details and context, focusing instead on the underlying structure or principles that define the problem. A general formulation allows the problem to be framed in a way that the solution can be applied to various domains or scenarios. Take, for example, the problem of a robot solving a maze. If the maze is a priori known, the robot can be programmed to solve that exact maze, and the problem would be formulated as "solve *that* maze." The solution would be specific to that maze's particular turns, measurements, and the dead ends within it. The robot could alternately be preprogrammed to "solve *any* maze." The latter is the generalized version of the problem, and the solution would apply to any maze that the robot is given. This formulation and solution requires an understanding of the scientific principles underlying the problem. Academics' scientific training and focus on basic research may plausibly enable them to identify general problems within their field of research, and their scientific knowledge may help their ventures invent general solutions. In a context with lower barriers to entry to downstream markets such as AI, this general knowledge can be translated into technologies with applicability to multiple markets.

We examine whether AI founders conduct basic research, leading to general knowledge, and whether these two aspects of their endeavors might drive broader market applicability.

5.2.1 | Examining the role of basic research

First, we tested whether the relatively higher emphasis on basic research by founders influences the breadth of market applicability. To conduct this test, we gathered all the publications in which our sample founders were listed as authors from the SCOPUS database.[23] In total, we identified 19,291 conference proceedings and journal publications until December 2021. We then examined the abstracts of publications to classify them as basic or applied research by following the well-established definitions of basic and applied research as used in the literature on science (e.g., Bieniawski, 1994; Nelson, 1959; Rosenberg & Nelson, 1994; Stokes, 1997; Tijssen, 2010).[24] Our entire co-author team first manually coded

[23]To identify matches on SCOPUS, we utilized the names of the founders, their current and previous affiliations, and their PhD institution (if applicable). Among our 988 founding teams, 354 had a founder who had at least one paper published before entry and 409 pre- and post-startup incorporation. We retained those publications that had an abstract in SCOPUS, and at least two citations; our final sample consisted of 16,400 publications.

[24]An alternative approach is to consider what is considered a "natural science": mathematics, biology, chemistry, and physics. We identified these fields using the ASJC codes from Scopus, specifically the first two-digit codes (26 for mathematics, 13 for biology, 16 for chemistry, and 31 for physics). When we included the number of publications in each of these fields as independent variables, we found similar results indicating that basic research does not fully account for our main effect. Results are available upon request.

10% of the publications (2000 in total; randomly selected).[25] We then used fine-tuned ChatGPT to categorize rest of the abstracts according to the same principles we used to manually code our abstracts (Chattopadhyay et al., 2024). Comparing ChatGPT's categorization of the 2000 abstracts to ours, we achieved a matching rate of 78%, a level that we deemed sufficiently high, given that a correlation of 0.7 is commonly employed as a benchmark for interrater reliability (Taber, 2018). We explain our method in more detail in the Appendix Section 6.

After using ChatGPT to code the rest of the abstracts, around two-thirds of the abstracts were coded as applied and one-third as basic. We created a variable *Proportion of basic research* by dividing the number of pre-entry basic publications by the total number of pre-entry publications for each startups. On average, academic startups had 19 basic publications, which is more than twice the number of basic publications by nonacademic PhD-led startups (with an average of 7.6). Most teams without academics or PhD-holders did not have any basic publication (with an average of 0.1). Then, 27% of academic startups' publications, 17% of nonacademic PhD-led startups' publications, and 1% of publications of teams without either academics or PhD holders are basic, while the rest are applied. It is important to note that all startups tend to bring in more applied research than basic research within this context.

We first examined the effect of the number of pre-entry publications. In Model 1 of Table 4, we observed that the magnitude of the coefficient on *Has an academic entrepreneur* was reduced by 3%. This suggests that a team's number of publications explains part of the association between academics and market applicability. To explore how much of this association is explained by basic research, we added the *Proportion of basic research* in Model 2. We did not find the proportion of basic publications to be associated with market applicability.

5.2.2 | Examining the role of generality

Next, we examined whether academics produce more general knowledge than nonacademics. We used two knowledge-based measures to test for generality: (1) the founders' publications and (2) the startups' patents. To construct *Average generality of publications*, we quantified the generality of each publication by examining the forward citations each publication received. To measure generality, we considered the journal categories of the articles that cited the focal publications using the journal categorization system called All Science Journal Classifications[26] from SCOPUS. We calculated generality by counting the number of fields that cite the focal publications. We then aggregated the generality measure to the team level by taking the average of team members' generality.

Using the patent classes that cite the focal patent, we similarly constructed *Average generality of patents*.

[25]As a robustness check, we conducted the same set of analysis using hand-coded abstracts only. We found qualitatively similar result that basic research does not explain our finding that academic startups have higher market applicability. Results are available upon request.

[26]According to SCOPUS, "the classification is based on the aims and scope of the title, and on the content it publishes." A complete list of categories are listed here: https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/ (Last accessed on November 15, 2023). For the purpose of our analysis, we used the *Subject Area Classifications* which is one level below *Subject Area*.

TABLE 4 Channel 2: Knowledge and market applicability.

	Publications				Patents			Full model
	Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Market applicability	Market applicability	Average generality of pubs	Market applicability	Average generality of patents	Market applicability	Market applicability	Market applicability
Has an AE	0.37	0.35	1.45	0.3	0.55	0.36	0.33	0.26
	[.06]	[.09]	[.00]	[.14]	[.00]	[.03]	[.05]	[.21]
Proportion of basic research		0.34	3.24					0.22
		[.17]	[.00]					[.40]
Log number of research publications	0.00	−0.02	0.81	−0.04				−0.05
	[.93]	[.67]	[.00]	[.38]				[.28]
Average generality of pubs				0.04				0.04
				[.04]				[.10]
Number of patents					0.12	0.01	0.00	0.00
					[.00]	[.10]	[.57]	[.62]
Average generality of patents							0.05	0.05
							[.07]	[.08]
Team size	−0.02	−0.02	0.00	−0.02	0.12	−0.02	−0.03	−0.03
	[.78]	[.74]	[.98]	[.76]	[.07]	[.72]	[.64]	[.62]
Firm age	0.11	0.11	0.08	0.11	0.02	0.11	0.11	0.10
	[.00]	[.00]	[.01]	[.00]	[.39]	[.00]	[.00]	[.00]

TABLE 4 (Continued)

	Publications				Patents			Full model
	Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	**Market applicability**	**Market applicability**	**Average generality of pubs**	**Market applicability**	**Average generality of patents**	**Market applicability**	**Market applicability**	**Market applicability**
Has a female founder	−0.1	−0.10	−0.10	−0.10	−0.14	−0.10	−0.09	−0.09
	[.33]	[.34]	[.51]	[.36]	[.24]	[.36]	[.39]	[.42]
Has a non-AE PhD	0.29	0.27	1.16	0.24	0.18	0.28	0.27	0.21
	[.01]	[.02]	[.00]	[.05]	[.11]	[.01]	[.01]	[.07]
Has a serial founder	0.13	0.13	−0.05	0.13	0.12	0.13	0.12	0.12
	[.23]	[.24]	[.77]	[.23]	[.30]	[.22]	[.25]	[.24]
Average number of past employers	0.00	0.00	0.08	0.00	−0.04	0.01	0.01	0.00
	[.86]	[.88]	[.03]	[.98]	[.14]	[.79]	[.73]	[.84]
Average experience	−0.01	−0.01	−0.03	−0.01	0.01	−0.02	−0.02	−0.01
	[.06]	[.07]	[.00]	[.10]	[.24]	[.05]	[.04]	[.07]
Has a foreign founder	0.09	0.10	−0.14	0.1	−0.27	0.09	0.11	0.12
	[.33]	[.31]	[.34]	[.29]	[.01]	[.33]	[.26]	[.22]
Has a founder from top 10 inst.	−0.01	0.00	0.47	−0.03	−0.14	0.00	0.00	−0.01
	[.96]	[.99]	[.04]	[.87]	[.40]	[.98]	[.98]	[.95]

TABLE 4 (Continued)

	Publications				Patents			Full model
	Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Market applicability	Market applicability	Average generality of pubs	Market applicability	Average generality of patents	Market applicability	Market applicability	Market applicability
Log number of X followers	0.04	0.04	0.00	0.04	0.03	0.04	0.04	0.03
	[.01]	[.01]	[.85]	[.01]	[.07]	[.01]	[.01]	[.01]
Constant	2.47	2.47	0.11	2.46	0.08	2.50	2.5	2.50
	[.00]	[.00]	[.74]	[.00]	[.77]	[.00]	[.00]	[.00]
Observations	988	988	988	988	988	988	988	988
R-squared	.06	.06	.50	.06	.28	.06	.06	.07

Note: *p*-values in brackets. All models control for location dummies. For Models 2, 3, and 8, we used the proportion of basic science rather than a count measure due to the high collinearity between the number of publications and the number of basic publications. We confirmed that the overall VIF for each model fell below the standard cutoff of 5. For Model 3, we also checked for the quadratic effect of *the proportion of basic research* but found the inflection point to be close to 1.

TABLE 5 Summary of tests for two channels.

Step		Relevant variables	Findings
Main inquiry	Relationship between academic startups and the breadth of market applicability.	Academic entrepreneurs Number of verticals • From Pitchbook	We found a positive relationship: academic startups generate more market applicability than nonacademic startups, 12% more. (Table 2 Model 1)
Commercialization approach	Academics position them as sellers on the market for technologies, through which they can enhance market applicability.	Number of patent sales • Using patent assignment records from USPTO database Location in the value chain (upstream vs. downstream) • Using business descriptions from Pitchbook	None of the startups in our sample licenses their technology to public firms. Some startups sell their patents. We observed no significant differences between academic and nonacademic startups. Most startups pursue downstream applications—only 6.4% of all startups are located upstream based on their business description. Being upstream does not relate to the number of verticals. (Table 4, Models 1–4)
Knowledge creation approach	Academic entrepreneurs pursue more basic research, which might lead to higher market applicability.	Number of publications • From SCOPUS database Number of basic (vs. applied) publications • Examined the abstracts of publications to classify them as basic or applied research	Academics publish more, but the number of publications does not correlate with higher market applicability (Table 3 Model 1) Academics pursue more basic research than all other entrepreneurs (among the entrepreneurs who publish: 27% of academic research is basic vs. 7% within nonacademic research). However, conducting basic research does not relate to generating more market applicability. (Table 3 Model 2)
	Academics' pursuit of general problems helps their startups attain broader market applicability than nonacademic startups.	Average generality of patents • Average number of patent classes that cite the focal patents Average generality of publications	Academic startups' founders publish more general publications (cited across numerous fields) and obtain more general patents (cited across numerous classes). The startups' general outcomes explain the number of market segments

TABLE 5 (Continued)

Step		Relevant variables	Findings
		• Average number of fields that cite the focal publications	they pursue. (Table 3, Models 4–8) Interviews further revealed that the extensive research that academics conducted to solve general problems is used in their startups. Nonacademics focus on concrete problems from their own experience in their personal life or in their work. They tend to solve these problems by combining technical solutions at hand.
Conclusion	Academics startups generate more market applicability than nonacademic startups because their academic founders solve general problems that can be turned into solutions applicable to more market segments. In contrast, nonacademic startups solve concrete problems that relate to fewer market segments.		

First, we examined the role of generality using publication data. In Model 3 of Table 4, we found a positive association between academic startups and generality. We found that teams with an academic entrepreneur were on average cited across categories about twice as much as others ($\beta = 1.70$, $p = .00$; average generality = 1.90). Further, we found that the proportion of basic research is associated with higher average generality ($\beta = 3.24$, p = .00). This suggests that basic science leads to the creation of general knowledge. We then tested whether the generality of publications explains any of the association between academics and the breadth of market applicability. In Model 4, receiving a citation from one additional field is associated with approximately a 1.3% increase over the average in the number of market segments ($\beta = .04$, $p = .04$). When the generality of publications is introduced in our main regression, having an academic entrepreneur is no longer associated with the breadth of market applicability. This suggests that the generality of knowledge plausibly explains the association and is an important path through which academics can pursue more market applicability.

Next, we conducted the same set of tests using the generality of patents held by startups. In Model 6, we find that academic startups have, on average, 0.55 more classes citing the focal patents than nonacademic startups ($\beta = .55$, $p = .00$), which amounts to a difference of 73% over the mean (average generality = 0.75). In Model 6, a 5% reduction in the coefficient size compared to Model 1 in Table 2 suggests that the number of patents partially explains the association. In Model 7, we see that receiving a citation from one additional field is associated with a 2% increase over the average in the breadth of market applicability ($\beta = .05$, $p = .07$). When we added the average generality of patents, the coefficient size on *Has an academic entrepreneur* in this model further decreased by 13% compared to Model 1 in Table 2. This suggests that the generality of patents partially accounts for the positive association between academic startups and the applicability of the startup technologies to a wide range of market segments. In Model 8, we estimated a full model controlling for basic knowledge, general publications, and general patents. In the full model, we see that the generality of publications and the generality of

TABLE 6 Size of network and complexity and the effect of academic entrepreneur and market applicability.

	(1) Log number of collaborators (mean)	(2) Market applicability	(3) Complexity (mean)	(4) Market applicability
Has an AE	0.03	0.37	0.05	0.36
	[.75]	[.06]	[.33]	[.03]
Log number of collaborators (mean)		0.03		
		[.65]		
Log number of research publications	1.05	−0.03		
	[.00]	[.74]		
Number of patents			0.02	0.01
			[.00]	[.08]
Complexity (mean)				−0.07
				[.55]
Team size	0.02	−0.02	0.03	−0.02
	[.46]	[.77]	[.14]	[.74]
Firm age	0.00	0.11	0.02	0.11
	[.77]	[.00]	[.00]	[.00]
Has a female founder	0.02	−0.1	−0.07	−0.1
	[.67]	[.33]	[.02]	[.34]
Has a non-AE PhD	0.00	0.29	0.03	0.28
	[1.00]	[.01]	[.34]	[.00]
Has a serial founder	−0.03	0.13	−0.06	0.13
	[.58]	[.23]	[.05]	[.24]
Average number of past employers	0.05	0.00	0.00	0.01
	[.00]	[.91]	[.57]	[.80]
Average experience	−0.01	−0.01	0.00	−0.02
	[.00]	[.07]	[.11]	[.05]
Has a foreign founder	−0.06	0.10	0.12	0.10
	[.23]	[.32]	[.00]	[.29]
Has a founder from top 10 inst.	0.10	−0.01	−0.02	−0.01
	[.21]	[.95]	[.69]	[.97]
Log number of X followers	0.00	0.04	0.01	0.04
	[.57]	[.01]	[.13]	[.01]
Constant	0.07	2.46	−0.10	2.50
	[.55]	[.00]	[.14]	[.00]

TABLE 6 (Continued)

	(1) Log number of collaborators (mean)	(2) Market applicability	(3) Complexity (mean)	(4) Market applicability
Observations	988	988	988	988
R-squared	.81	.06	.16	.06

Note: *p*-values in brackets. All models control for location dummies.

patents are both positively associated with the breadth of market applicability, while basic knowledge is not.

Our evidence thus far suggests that generality of knowledge is a key driver of the association of academics with wider market applicability. To fully ascertain that there are no other driving factors, we investigated alternative mechanisms that may contribute to this association.[27]

First, we investigated the association between academic founders' collaborator network and the outcome variable. Academics have been shown to have extensive co-author networks (Murray, 2004) which may provide them with insights that lead to wider market applicability. To test this explanation, we examined the number of collaborators from their pre-venture publications, using the publication data that we obtained from SCOPUS. We counted the number of collaborators of each founder, and then calculated the average value of this number for each team. Findings are reported in Table 6. We did not find evidence that teams with an academic entrepreneur had more collaborators compared to others (Model 1) or that the average number of collaborators and market applicability were positively associated (Model 2), suggesting that network effects are not driving the effect.[28]

Next, we explored the possibility that technology developed by academic startups has greater interdependencies (complexity) across subtechnologies, leading to broader market applicability. We assessed the complexity of each patent using the method introduced by Ganco (2013) that uses the co-occurrence of subclasses. We calculated the interdependence for each subclass and the complexity measure for each of these patents.[29] We then aggregated this measure to the team level by taking the average value.[30] As Model 3 in Table 6 shows, we did not find a relationship between teams with academic entrepreneurs and complexity using an average measure among the team. We also did not find an association between complexity and the breadth of market applicability (Model 4).

Taken together, our evidence shows that the generality of academic startups' knowledge and inventions explains the positive association between academic startups and market applicability.

[27]We thank the reviewer team for suggesting these two tests.

[28]We also checked the maximum and minimum of the number of collaborators within teams and found similar results.

[29]We calculated complexity following the approach in Ganco (2013), complexity K for patent l is calculated as:
$= \frac{1}{count\ of\ subclass\ of\ patent\ l} \sum K_i$
where $K_i = \sum_{j \in l_{-i}} \frac{count\ of\ patents\ in\ subclasses\ i\ and\ j}{count\ of\ patents\ in\ subclass\ i}$, where j belongs to all subclasses except i.

[30]We found robust results when taking the maximum or minimum value of the complexity measure for each team.

6 | ORIGINS OF GENERALITY: PROBLEM FORMULATING AND SOLVING IN AI

What is the process through which academics' focus on scientific principles can lead to the creation of general knowledge and how does this differ from nonacademics? To further examine the origins of the differences between academics and nonacademics in terms of generality and market applicability we turned to interviews of individuals in AI.[31]

6.1 | Academics′ perspective: Choosing "big" problems and putting "science first"

The differences in how academics and nonacademics develop their respective technologies within their startups lead to differences in their market applicability. From interviews, we uncovered *two observations*: (1) academics rely on scientific principles to pursue more general problems, and they produce classes of solutions to their problems rather than ad hoc solutions, and (2) nonacademics identify market pain points or concrete problems that they solve with existing tools, they thus serve a clearly identified market in their startups. We first share observations that illustrate how academics identify and solve problems during their time in academia.

Several academics iterated that within the academe, such as in computer science and engineering, the generality of the problem drives the formulation and selection of the problem. Academics who focus on general (i.e., broadly applicable) questions are likely to fetch more recognition since their work is likely to be relevant to many more domains. Their ability to ask and solve general questions originates from their reliance on scientific principles and the pursuit of fundamental or basic research. This is revealed in one of our conversations with an academic entrepreneur, who is a faculty of computer engineering (Faculty 2):

> "It comes down to what kind of knowledge are you trying to advance. Is it a very specific kind of question, which applies to a very specific industry or a domain? Or are you trying to answer a question, which would be more broadly applicable? More broadly means it's more foundational. Because it's broad other things can be built on top of that foundation. And I think, when you're trying to do more foundational work then it'll be more influential basically, and in a university, you gain reputation by being more influential."

He further explained that he chooses not to formulate very narrow (specific) problems because often peers within other departments within the university do not have the expertise to appreciate and reward the solution of very specific problems, but do understand more general problems and solutions:

> "And if you want to be super specific, it's hard because you don't have the domain knowledge necessary to be super specific in the university, that requires, you know, 30 years of understanding of—say how car batteries are made- and we don't have that kind of expertise in it. We don't have an automotive department. So, you

[31]Table D1 in the Appendix present our interviewees' background and occupation.

> have to tackle more of the foundational questions that are more general and understood by more departments."

An example comes from another computer science faculty (Faculty 1) who described how his research group spent over a decade chipping away at the problem of indoor localization. They attempted to make global positioning systems, which then typically worked outside buildings, also work efficiently inside buildings. He described how his group attempted to solve it from several different angles, such as using radio frequency, Bluetooth, acoustic audio signals, fusing with motion sensors, and finally Wi-Fi. He explained it this way:

> "You start with the big problem that you want to solve. Break those problems into what pieces need to be solved and you go and attack the pieces that have not been solved in the past."

The lack of constraints and lower opportunity costs further bolster the pursuit of "big" problems (Conti & Roche, 2021; Perkmann et al., 2019). Beginning the knowledge creation process without prioritizing any specific applications may remove constraints in problem formulation, potentially expanding the breadth of applicability of the solutions. Academics' formulation of the problem itself is, therefore, only limited by the boundaries of their abilities and that of science. Indeed, Faculty 4 said that:

> "… it is the way it should be. Right, like one can argue about like, say, autonomous cars, right? One can argue either way like some people believe that there's no need for cars to be autonomous, and in some ways you can agree with them. But a lot of academics will believe that cars need to be autonomous because they can be."

One way that academics identify research questions is to delve into *why* the technology works. When solving a problem, they adopt a similar approach. For instance, several faculty members stated that they develop algorithms based on solid mathematical foundations that are generalizable to classes of problems, rather than one-off solutions that address one specific problem. Faculty 2 noted the difference between academia and industry in that regard:

> "So just to give an example, we are developing some algorithms for doing some verification. And then we realize that we, the algorithms people in academia, always relied on models. Someone writes down a mathematical model. But then we talk to industry and the reaction was …. [shrugs]. Well, they often don't have models, right? Or they have a partial model. So, they formulated this problem where you have a partial model and the rest of it is just some code for which you would not be able to write down the model."

Based on these interviews, our first observation is that academics pursue "big" or general research problems that they solve with a general approach. If they translate their research findings into startups, first and foremost, they do so based on scientific principles and basic research. This can be regarded as a "science push" approach to entrepreneurship. Academics' overall process—from generating and pursuing ideas, to developing technologies—results from more general problem formulation and problem solving. The knowledge that is created and the ultimate products developed are thus more likely to be widely applicable once commercialized.

6.2 | Nonacademics' problem-solving perspective: Concrete problems and markets first

We next turned to nonacademic entrepreneurs. In various conversations, entrepreneurs with prior experience in the industry explained that they started their entrepreneurial endeavors by focusing on the features of the market that demonstrated a well-defined "pain point" or a need that they believed they could fulfill. The technology they developed within the startup targeted that specific market need.

One entrepreneur described how they spent hours talking to farmers to understand the problems they faced within agribusiness. These conversations highlighted the fragmentation and information asymmetry within the agricultural market. Farmers were not connected to each other, markets, or suppliers efficiently and comprehensively; this led them to spend large amounts of time and effort tracking information and coordinating through middlemen. These observations inspired the entrepreneur to develop an app that connects farmers to various parties and provides them with information pertaining to all aspects of their value chain. The entrepreneur's idea was born from the identification of a market need; she judged that the potential demand and market size were appropriately large, and her application was therefore targeted toward solving that specific need.

Another entrepreneur was a doctoral student working for his advisor's startup that was focused on edge computing. However, he decided to quit because customers were mainly interested in current solutions that could solve problems that had manifested and were immediate.

> "You are so far advanced in the field that the problems that you're trying to solve oftentimes aren't actually problems for regular people or they're not yet right. [...] We would go in and pitch this idea, and you know it's like well network traffic isn't that bad yet, and you could kind of say: Well, it's going to be so let's start working on this now, right? But that was often a problem for industry people [not wanting to anticipate]."

He then formed a different startup, drawing upon an extensive discussion with his wife about a practical problem in her job. He and his spouse then jointly developed a solution.

Faculty 1 who does consulting in the industry and Employee 2 also mentioned how investors and executives sometimes constrain the scope of their research. Employee 2 said:

> "[W]hen engineers are proposing their ideas on solving part of the system, they need to face the questioning from an executive team. That's what I'm trying to say, okay? So they're in the same room and they're kind of the engineering team is trying to propose something. The executives are modifying it, that's a lot of what they're going to do."

The approach to solving problems within the industry differs from that within academia and will influence employees-turned-nonacademic entrepreneurs. The nonacademics start with opportunities that run on heuristics based on their past experiences.

> "So the first part is going to be like heuristic based [...], but this is one component of the business. And the second component is providing agriculture news and government subsidies, and schemes relevant to your region."

The employee added that she focuses on what works and what improves operational processes.

These conversations led us to our second observation, that those working in the industry, facing time and profitability constraints, identify a concrete problem and rely on heuristics to generate a solution that works, paying limited attention to the proof or the reason *why* a solution works (Ott et al., 2017).[32] The nonacademic approach tends to be targeted toward solving a problem at a time rather than a *class* of problems. Nonacademic entrepreneurs are more likely to anchor on a specific market "pain point" or concrete problems and develop practical solutions, which limit the market applicability of their solutions.

7 | SUMMARY OF FINDINGS

We presented two channels through which academic startups potentially achieve higher market applicability. We find that most startups are positioned downstream in the industry value chain, and the difference between academic and nonacademic startups in terms of market applicability does not play out through this channel. Instead, we find the knowledge creation channel to be much more promising. Our results show that academics' basic research leads to the creation of more general knowledge through more general publications and patents, which can potentially be applicable to the a wider scope of markets. Our interviews reveal a stark contrast between academics' "science first" approach and nonacademics' "markets first" approach, which is also reflected in academics' tendency to formulate and solve general problems, and nonacademics' focus more on specific customers' demands and needs, leading to more targeted applicability of resulting products. The lower barriers to entry within AI likely create a lower "translation distance" of technologies to market within AI, allowing academics' general technologies to be associated with higher market applicability.

8 | SUPPLEMENTARY ANALYSES

8.1 | Problem formulation and solving and market applicability among nonacademic PhDs

In this section, we delve further into data to investigate the origins of academics' problem formulation approach. We reason that as academics train their doctoral students in their mold (Roche, 2023) and the academic journey starts with the attainment of a PhD degree, doctoral training itself may be a source of differences between academics and nonacademics in their approach to problem formulation and solving. We expect that doctoral training will leave imprints on individuals' mindset (McEvily et al., 2012). If doctoral students subsequently become faculty, they continue to engage in this problem formulation and solving approach, while those who transition into industry jobs are likely to also adopt the norms and practices of the industry, due to the influence of local contexts (Bercovitz & Feldman, 2006; Dokko et al., 2009).

[32]While some larger technology companies do focus on long-term R&D and focus on developing algorithms rather than heuristic-based approaches, on average, nonacademics are focused on finding an immediate solution to a problem without caring *why* that solution works, and as a result, the solutions have inherent limited applicability.

First, we compare teams with at least one nonacademic PhD holder with those without any PhDs. In Model 1 in Table 2, we observed that founding teams with at least one nonacademic PhD holder is associated with a wider market applicability relative to teams without any PhD holders ($\beta = .30$, $p = .00$). We next examine variations *within* nonacademic PhD holders. We expect the association between PhD holders who worked in the industry and market applicability to attenuate with the amount of time spent in the industry.[33] Table E1 in the Appendix shows that the association between having a nonacademic PhD holder and the market applicability decreases with the time the PhD holder spent working in the industry. We did not find a linear effect of industry experience within this sample (Model 1) and therefore we used a spline specification to capture any nonlinearity following the approach in prior literature (Won & Bidwell, 2023).[34] We present a variety of analyses with thresholds set at 3, 5, 10, and 15 years. The results are presented in Models 2–5. Across all models, we do not see an effect of industry experience below the threshold. In Model 3, once the total number of years of experience exceeds 5 years the association between time in the industry and the market applicability becomes negative ($\beta = -.03$, $p = .07$). The effect is especially strong, and the estimation becomes more accurate once the threshold is set to 15 years ($\beta = -.07$, $p = .00$). These results suggest that a moderate number of years in the industry after receiving a PhD does not change the relationship between founders' PhD and the market applicability. However, for those who have spent at least 5 years within the industry after obtaining their PhD, each additional year (after the 5 years) that they spend within the industry is negatively associated with the market applicability.

8.2 | Breadth of market applicability and new venture outcomes

Having explored the inception of academics' problem formulation and solving approach we next analyze its potential consequences for startups. To examine whether wider breadth of market applicability matters for the quality of ventures created, we investigated the association between market applicability and funding or exit outcomes. Note that we found valuation data for only a subsample of teams (67%). In our sample, only two teams had experienced IPOs. The most common liquidity event, therefore, is acquisition by other firms.

We used an OLS regression to estimate the relationship between the breadth of market applicability and the valuation of the venture and a linear probability model to estimate the likelihood of having a liquidity event. The results are reported in Table F1 of the Appendix. We find that the number of verticals is positively associated with both measures. Having one more vertical is associated with a 14% increase in valuation (Model 1). Since having an academic is associated with 0.38 more verticals (Model 1 in Table 2), this effect is equivalent to a 5.3% increase in terms of valuation. We also found that the breadth of market applicability is positively associated ($\beta = .01$, $p = .11$) with the likelihood of having a liquidity event, where one additional vertical is associated with a 1 percentage point increase in the chance of getting acquired (Model 2).

---

[33]PhD holders who found their new venture immediately after graduation will not experience attenuation.

[34]The spline splits industry experience into two separate variables: one variable measures the effect of experience between zero and the threshold (or the "knot") and the second variable measures the effect of experience above the threshold. The estimated coefficients on these variables capture the effect of experience on the number of market opportunities *within* that range of experience.

9 | DISCUSSION AND CONCLUSION

9.1 | Key contributions

This study is motivated by the need to understand how and why academic and nonacademic startups differ in their market applicability in the AI ecosystem, where structural features differ from other high-tech industries. The AI context flips the typical constraints startups face when considering markets and thus offers a new type of entrepreneurial landscape. Our paper broadens the conventional perspectives within the innovation and entrepreneurship literature by showing that when the need for downstream complementary assets is limited, academic startups can produce technologies with large market applicability while being located downstream. While prior work has shown that only firms with well-developed complementary assets tend to pursue downstream product markets (Conti et al., 2019), our evidence indicates academics are able to leverage their problem-formulation approach and the structural features of the ecosystem to deviate from this pattern.

Our qualitative inquiry brings to light the actual process of knowledge creation relying on the founders' problem-formulation and solving approach and offers new opportunities for theorization on the micro-foundations of human capital development and innovation (Chattopadhyay, Won, & Bidwell, 2024; Heydari et al., 2024). Being able to abstract away from the bones of the problem to see its essence is not a ubiquitous skill and is not always inherent to founders. General problem-formulation results from high construal thinking. If entrepreneurs' approach relies on prior academic training, they are more likely to think in terms of high levels of abstraction, that is, high construal levels (Park & Baer, 2022; Trope & Liberman, 2010) as opposed to concrete narrow terms. Focusing on the *why* allows high-construal thinkers to view problems in terms of their essential characteristics; they view problems as belonging to different *classes*, where different problems are categorized in terms of their common traits and characteristics. They would then aim to develop solutions that ideally apply to entire classes of problems rather than individual problems. Founders' general problem-formulation therefore likely lead to broader potential market applicability.

Our interviews further indicate that academics do not go "searching" for market opportunities, suggesting that technological discoveries find the markets. Our findings, as well as interviews, thus relate to the debate between science first and demand-pull (e.g., Di Stefano et al., 2012), where academic entrepreneurs choose the former and nonacademics the latter. In a context where technological advancements progress quickly, we show that both types of entrepreneurships can coexist in the same ecosystem.

Our supplementary results provide evidence that doctoral training is a plausible means through which individuals acquire a distinct cognitive framing. This finding suggests that prior literature on academic entrepreneurship, which has largely focused on the lack of complementary assets like operational know-how and marketing knowledge and the role played by Technology Transfer Offices in providing these services (e.g., Bercovitz & Feldman, 2006, 2008; Clarysse et al., 2011), may have underestimated the role played by doctoral training in shaping new ventures' outcomes.

We moreover document new findings that provide insights into the ecosystem of the field of AI (Jacobides et al., 2021; Zhang et al., 2021). In our snapshot, startups offer AI-based technological advancements for downstream markets locating in many industries. In the future, startups may utilize and further expand the enabling features of the technology across industries (Gambardella et al., 2021) while still relying on large incumbents' upstream assets. Our paper sets the ground for more exploration of the role of different types of entrepreneurs in developing technologies transversal to many markets and industries (Agarwal et al., 2024).

In terms of methods, we confirm empirically that the field of AI is suitable for combining publications and patent data to develop insights (e.g., Miric et al., 2023). We also show how researchers can leverage a generative AI tool to automate the classification of textual data, in our case, basic versus applied publications and upstream versus downstream businesses (e.g., Chattopadhyay, et al., 2024).

Generalizing beyond our context, our paper suggests that there are different paths entrepreneurs can take in the technological entrepreneurship realms when considering market opportunities. One prevalent view of entrepreneurship is the discovery of opportunity, which is guided by the heterogeneity among individuals (not all individuals recognize the same opportunity), what people know, and the distribution of information (Shane, 2000). While some kinds of entrepreneurship can certainly be viewed through this lens, our paper offers a complementary perspective where market opportunities are shaped by the technologies developed within the startup.

9.2 | Limitations and future research

Our findings show that greater scope of market applicability is associated with higher valuation. This suggests that despite the limitations of academic entrepreneurs, and the fact that they constitute a small fraction of all entrepreneurs in AI, the academic approach to entrepreneurship, which is rooted in a distinct problem-formulation and problem-solving approach, may offer some benefits that were previously overlooked. At the same time, these findings must be contextualized in the nascency of the field of AI and the early-stage nature of the startups themselves. Given that AI is a transversal and relatively new technology, venture capitalists are likely to view wide market segment breadth as a positive signal of the firm's initial potential market size, but their assessment may evolve with time. Higher breadth may signal a lack of focus as the startup ages. Is greater market applicability a double-edged sword that can quickly become an encumbrance to startups' growth and profits? Will academic startups be able to profit from the enabling technology they are creating (Gambardella et al., 2021; Novelli, 2015)? Or is excessive generality a reason that academic startups have lower survivability than others (Park et al., 2024)? Indeed, our conversations with entrepreneurs indicated that their funding VCs encouraged the ventures to pivot and focus when the founders were preparing to seek late-stage funding for their startups (Series E and beyond) rather than remain overly general. A longitudinal study on the association of verticals with firms' outcomes can provide us with newer and better developed insights and help answer these questions. Other contexts centered on other enabling technologies or simply other technological industries may offer an interesting comparison.

Our study focused on startups with two or more founders, and within our sample, the majority of academic startups included nonacademics as well. We examined whether backgrounds of nonacademic co-founders influence the main finding and found that it did not. Additional examination of only the sample of solo-founder startups revealed that solo academics' ventures were not associated with greater market applicability compared to other solo teams. As we detailed in the Appendix, ventures founded solely by academics may differ from those in teams due to differences in motivation or selection issues (e.g., choice of consulting activity). The team composition and the nuances between solo and team founders can offer promising areas for future research, especially in settings with more variation in team compositions spanning different technological fields.

The empirical results within this article can only be interpreted as correlational and not causal as there may be selection effects that our analyses do not disentangle. Individuals who select into academia may be qualitatively different in ability and preferences from those who do not; teams that include academics may be qualitatively different from those who do not. Nevertheless, we believe that our tests and interviews document an important channel through which founders influence their startup's market applicability conditional on selection into academia and subsequent selection into entrepreneurship in teams.

10 | CONCLUSION

This article documents a positive relationship between academic startups and the startups' breadth of market applicability in the field of AI. Our findings reveal that academics' use of scientific principles and their pursuit of basic research drives the creation of general knowledge, which in turn leads to wider potential market scope. With fewer requirements for complementary downstream assets in the AI ecosystem which allows startups to locate downstream, academics can more easily translate their general ideas to wider market applicability. Our interviews revealed that academics tend to identify general problems and generate solutions that can apply to a class of problems. Nonacademic interviewees suggested a market-oriented approach, leading to the use of heuristics to solve issues at hand in a narrow set of possible applications. Our findings unpack the channel through which academic entrepreneurs can broaden their firms' potential market scope in the field of AI.

ACKNOWLEDGMENTS

All authors contributed equally, and their names are in alphabetical order. The authors thank Rajshree Agarwal and two reviewers for their valuable guidance and comments. Shinjinee Chattopadhyay thanks Sayan Mitra and friends in UIUC ECE for many hours of enjoyable conversations on academics and their research. The authors are grateful for the research assistance from Yu Han, Elton Li, Yon Jin Suh, and Ji Young Lee. The authors also benefitted greatly from comments and suggestions from Russ Coff, Martin Ganco, Danny Kim, Elena Kulchina, Mahka Moeen, Michael Roach, Deepak Somaya, David Tan, Tiantian Yang, Dan Oslzewski, the WiSER group, and participants in HEC Strategy Seminar, the 2022 Wharton Technology Conference, 2022 & 2024 Academy of Management Conferences, 2022 SMS Conference, 2023 Madrid Work and Organizations Conference, 2023 People and Organizations Conference, 2024 ISA Conference, as well as seminar participants at the University of Wisconsin-Madison and KAIST. The authors thank the Bright Data initiative for giving access to LinkedIn data.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Shinjinee Chattopadhyay https://orcid.org/0000-0002-5326-1728
Florence Honoré https://orcid.org/0000-0003-1313-5177
Shinjae Won https://orcid.org/0000-0001-9914-5936

REFERENCES

Agarwal, R., Campbell, B. A., Franco, A. M., & Ganco, M. (2016). What do I take with me? The mediating effect of spin-out team size and tenure on the founder–firm performance relationship. *Academy of Management Journal*, *59*(3), 1060–1087.

Agarwal, R., Moeen, M., Guerra, M., & Aversa, P. (2024). *Creating new industries: The generative role of Heterogenous actors*. University of Maryland Working Paper.

Agarwal, R., & Ohyama, A. (2013). Industry or academia, basic or applied? Career choices and earnings trajectories of scientists. *Management Science*, *59*(4), 950–970.

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

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.

Ahsan, M. M., Rahaman, M. S., & Anjum, N. (2023). From ChatGPT-3 to GPT-4: A Significant Leap in AI-Driven NLP Tools. Saidur and Anjum, Nishath, From ChatGPT-3 to GPT-4: A Significant Leap in AI-Driven NLP Tools (March 27, 2023).

Arora, A., Fosfuri, A., & Gambardella, A. (2001). Markets for technology and their implications for corporate strategy. *Industrial and Corporate Change*, *10*(2), 419–451.

Atoji Keene, C. (2013). This robot doesn't replace workers, it works with them—The Boston Globe published on May 19th, 2013.

Beckman, C. M., & Burton, M. D. (2008). Founding the future: Path dependence in the evolution of top management teams from founding to IPO. *Organization Science*, *19*(1), 3–24.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, *81*(2), 608–650.

Bercovitz, J., & Feldman, M. (2006). Entrepreneurial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development. *Journal of Technology Transfer*, *31*(1), 175–188.

Bercovitz, J., & Feldman, M. (2008). Academic entrepreneurs: Organizational change at the individual level. *Organization Science*, *19*(1), 69–89.

Bieniawski, Z. T. (1994). Designing a Design Degree in Engineering. In *International design engineering technical conferences and computers and information in engineering conference* (Vol. 12822, pp. 303–313). American Society of Mechanical Engineers.

Chattopadhyay, S., Han, Y., Honoré, F., & Won, S. (2024). Getting a 25th hour: Fine-tuning pre-trained LLMs to classify text data for strategy research. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4996983

Chattopadhyay, S., Won, S., & Bidwell, M. (2024). *Dynamic demands versus durable skills: Career dilemmas in pharmaceutical R&D*. University of Illinois Working Paper.

Chen, P. L., Williams, C., & Agarwal, R. (2012). Growing pains: Pre-entry experience and the challenge of transition to incumbency. *Strategic Management Journal*, *33*(3), 252–276.

Clarysse, B., Tartari, V., & Salter, A. (2011). The impact of entrepreneurial capacity, experience and organizational support on academic entrepreneurship. *Research Policy*, *40*(8), 1084–1093.

Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The economics of artificial intelligence: An agenda* (pp. 115–146). University of Chicago Press.

Conti, A., & Guzman, J. (2019). What is the US comparative advantage in entrepreneurship? Evidence from Israeli migration to the United States. Preprint, submitted April 29. https://doi.org/10.2139/ssrn.3376131

Conti, A., & Roche, M. P. (2021). Lowering the bar? External conditions, opportunity costs, and high-tech start-up outcomes. *Organization Science*, *32*(4), 965–986.

Conti, R., Gambardella, A., & Novelli, E. (2019). Specializing in generality: Firm strategies when intermediate markets work. *Organization Science*, *30*(1), 126–150.

Contigiani, A. (2023). Experimentation and appropriability in early-stage ventures: Evidence from the US software industry. *Strategic Management Journal*, *44*, 2128–2174.

Delmar, F., & Shane, S. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, *4*(3), 215–247.

Di Stefano, G., Gambardella, A., & Verona, G. (2012). Technology push and demand pull perspectives in innovation studies: Current findings and future research directions. *Research Policy*, *41*(8), 1283–1295.

Dokko, G., Wilk, S. L., & Rothbard, N. P. (2009). Unpacking prior experience: How career history affects job performance. *Organization Science*, *20*(1), 51–68.

Duggal, S. (2023). Retrieved from https://www.entrepreneur.com/en-in/news-and-trends/this-ai-driven-service-desk-auto-resolves-employee-needs/446852

Durrell, K. (2021). Retrieved from https://www.nutritioninsight.com/news/mixfit-co-founder-talks-b2b-expansion-dsm-partnership-and-new-professional-platform-for-personalized-nutrition.html

Ewens, M., Gorbenko, A., & Korteweg, A. (2022). Venture capital contracts. *Journal of Financial Economics*, *143*(1), 131–158.

Feldman, A. (2020). New venture Founded By Cognitive Scientist Gary Marcus And Roboticist Rodney Brooks Raises $15 Million To Make Building Smarter Robots Easier—Published on forbes.com.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, *25*(8–9), 909–928.

Fuller, A. W., & Rothaermel, F. T. (2012). When stars shine: The effects of faculty founders on new technology ventures. *Strategic Entrepreneurship Journal*, *6*(3), 220–235.

Gambardella, A., & Giarratana, M. S. (2013). General technological capabilities, product market fragmentation, and markets for technology. *Research Policy*, *42*(2), 315–325.

Gambardella, A., Heaton, S., Novelli, E., & Teece, D. J. (2021). Profiting from enabling technologies? *Strategy Science*, *6*(1), 75–90.

Ganco, M. (2013). Cutting the Gordian knot: The effect of knowledge complexity on employee mobility and entrepreneurship. *Strategic Management Journal*, *34*(6), 666–686.

Ge, C., Huang, K. W., & Png, I. P. (2016). Engineer/scientist careers: Patents, online profiles, and misclassification bias. *Strategic Management Journal*, *37*(1), 232–253.

Giustiziero, G., Kretschmer, T., Somaya, D., & Wu, B. (2023). Hyperspecialization and hyperscaling: A resource-based theory of the digital firm. *Strategic Management Journal*, *44*(6), 1391–1424.

Gofman, M., & Jin, Z. (2024). Artificial intelligence, education, and entrepreneurship. *Journal of Finance*, *79*(1), 631–677.

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, *52*(1), 104653.

Gruber, M. (2010). Exploring the origins of organizational paths: Empirical evidence from newly founded firms. *Journal of Management*, *36*(5), 1143–1167.

Gruber, M., Heinemann, F., Brettel, M., & Hungeling, S. (2010). Configurations of resources and capabilities and their performance implications: An exploratory study on technology ventures. *Strategic Management Journal*, *31*(12), 1337–1356.

Gruber, M., MacMillan, I. C., & Thompson, J. D. (2008). Look before you leap: Market opportunity identification in emerging technology firms. *Management Science*, *54*(9), 1652–1665.

Gruber, M., MacMillan, I. C., & Thompson, J. D. (2012). From minds to markets: How human capital endowments shape market opportunity identification of technology start-ups. *Journal of Management*, *38*(5), 1421–1449.

Haeussler, C., & Colyvas, J. A. (2011). Breaking the ivory tower: Academic entrepreneurship in the life sciences in UK and Germany. *Research Policy*, *40*(1), 41–54.

Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are seed accelerators status springboards for startups? Or sand traps? *Strategic Management Journal*, *44*, 2060–2096.

Helfat, C. E. and Lieberman, M. B., (2002). The birth of capabilities: market entry and the importance of pre-history. *Industrial and Corporate Change*, *11*(4), 725–760.

Heydari, B., Chattopadhyay, S., Padhee, S., & Karim, S. (2024). Core or Periphery: Examining where to allocate heterogenous innovators through the impact on firms' innovation. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5028175

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., Senior, S., Vanhoucke, V., Nguyen, P., Sainath, T. N., & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97.

Hochberg, Y. V., Serrano, C. J., & Ziedonis, R. H. (2018). Patent collateral, investor commitment, and the market for venture lending. *Journal of Financial Economics*, *130*(1), 74–94.

Honoré, F. (2022). Joining forces: How can founding members' prior experience variety and shared experience increase startup survival? *Academy of Management Journal*, *65*(1), 248–272.

Jacobides, M. G., Brusoni, S., & Candelon, F. (2021). The evolutionary dynamics of the artificial intelligence ecosystem. *Strategy Science*, *6*(4), 412–435.

Jensen, R., & Thursby, M. (2001). Proofs and prototypes for sale: The licensing of university inventions. *American Economic Review*, *91*(1), 240–259.

Kim, S., Agarwal, R., & Goldfarb, B. (2023). Creating radical technologies and competencies: Revisiting Interorganizational dynamics in the nascent bionic prosthetic industry. Available at SSRN 4024471.

Kirtley, J., & O'Mahony, S. (2023). What is a pivot? Explaining when and how entrepreneurial firms decide to make strategic change and pivot. *Strategic Management Journal*, *44*(1), 197–230.

Krizhevsky, A., Sutskever, I., & Hinto, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Landi, H. (2021). Retrieved from https://www.fiercehealthcare.com/tech/fresh-off-its-win-cms-challenge-closedloop-ai-nabs-34m-to-build-out-explainable-ai

Leahey, E. (2006). Gender differences in productivity: Research specialization as a missing link. *Gender & Society*, *20*(6), 754–780.

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American Sociological Review*, *72*(4), 533–561.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. In *The handbook of brain theory and neural networks* (Vol. 3361, p. 1995). The MIT Press.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Lee, H., Shah, S. K., & Agarwal, R. (2024). Spinning an entrepreneurial career: Motivation, attribution, and the development of organizational capabilities. *Strategic Management Journal*. *45*(3), 463–506.

Lee, J. (2023). Retrieved from https://www.reuters.com/technology/ai-boom-triples-valuation-lightmatter-us-startup-using-light-computing-2023-05-31/

Levinthal, D. A., & Wu, B. (2010). Opportunity costs and non-scale free capabilities: Profit maximization, corporate scope, and profit margins. *Strategic Management Journal*, *31*(7), 780–801.

Margolis, J. (2017). *Stuck in the shallow end, updated edition: Education, race, and computing*. MIT Press.

McDonald, R., & Gao, C. (2019). Pivoting isn't enough? Managing strategic reorientation in new ventures. *Organization Science*, *30*(6), 1289–1318.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine*, *27*(4), 12.

McEvily, B., Jaffee, J., & Tortoriello, M. (2012). Not all bridging ties are equal: Network imprinting and firm growth in the Nashville legal industry, 1933–1978. *Organization Science*, *23*(2), 547–563.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, *22*(6), 635–659.

Miric, M., Jia, N., & Huang, K. G. (2023). Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents. *Strategic Management Journal*, *44*(2), 491–519.

Moeen, M., & Agarwal, R. (2017). Incubation of an industry: Heterogeneous knowledge bases and modes of value capture. *Strategic Management Journal*, *38*(3), 566–587.

Mosakowski, E. (1991). Organizational boundaries and economic performance: An empirical study of entrepreneurial computer firms. *Strategic Management Journal*, *12*(2), 115–133.

Murray, F. (2004). The role of academic inventors in entrepreneurial firms: Sharing the laboratory life. *Research Policy*, *33*(4), 643–659.

Nelson, R. R. (1959). The simple economics of basic scientific research. *Journal of Political Economy*, *67*(3), 297–306.

Nerkar, A., & Shane, S. (2003). When do start-ups that exploit patented academic knowledge survive? *International Journal of Industrial Organization*, *21*(9), 1391–1410.

Nickerson, J., Yen, C. J., & Mahoney, J. T. (2012). *Exploring the problem-finding and problem-solving approach for designing organizations. Academy of Management Perspectives*, *26*(1), 52–72.

Nickerson, J. A., & Zenger, T. R. (2004). A knowledge-based theory of the firm—The problem-solving perspective. *Organization Science*, *15*(6), 617–632.

Novelli, E. (2015). An examination of the antecedents and implications of patent scope. *Research Policy*, *44*(2), 493–507.

Ott, T. E., Eisenhardt, K. M., & Bingham, C. B. (2017). Strategy formation in entrepreneurial settings: Past insights and future directions. *Strategic Entrepreneurship Journal*, *11*(3), 306–325.

Park, C. H., & Baer, M. (2022). Getting to the root of things: The role of epistemic motivation and construal levels in strategic problem formulation. *Strategy Science*, *7*, 284–299.

Park, U. D., Rha, Y., Shah, S. K., & 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.

Perkmann, M., McKelvey, M., & Phillips, N. (2019). Protecting scientists from Gordon Gekko: How organizations use hybrid spaces to engage with multiple institutional logics. *Organization Science*, *30*(2), 298–318.

Press, J. (2020). 12 AI milestones: 4. MYCIN, An expert system for infectious disease therapy (forbes.com).

Reese, D., Rieger, V., & Engelen, A. (2021). Should competencies be broadly shared in new ventures' founding teams? *Strategic Entrepreneurship Journal*, *15*(4), 568–589.

Roach, M., & 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. P. (2023). Academic entrepreneurship: Entrepreneurial advisors and their advisees' outcomes. *Organization Science*, *34*(2), 959–986.

Roche, M. P., Conti, A., & Rothaermel, F. T. (2020). Different founders, different venture outcomes: A comparative analysis of academic and non-academic startups. *Research Policy*, *49*(10), 104062.

Rosenberg, N., & Nelson, R. R. (1994). American universities and technical advance in industry. *Research Policy*, *23*(3), 323–348.

Sauermann, H., & Stephan, P. (2013). Conflicting logics? A multidimensional view of industrial and academic science. *Organization Science*, *24*(3), 889–909.

Serrano, C. J. (2007). Technology transfer and the business cycle: Evidence from patent reassignments.

Serrano, C. J. (2010). The dynamics of the transfer and renewal of patents. *RAND Journal of Economics*, *41*(4), 686–708.

Shane, S. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science*, *11*(4), 448–469.

Shane, S. A. (2004). *Academic entrepreneurship: University spinoffs and wealth creation*. Edward Elgar Publishing.

Sharma, S. (2023). Retrieved from https://venturebeat.com/data-infrastructure/virtualitics-raises-37m-for-ai-driven-3d-data-exploration/

Shermon, A., & Moeen, M. (2022). Zooming in or zooming out: Entrants' product portfolios in the nascent drone industry. *Strategic Management Journal*, *43*(11), 2217–2252.

Shortliffe, E. (Ed.). (2012). *Computer-based medical consultations: MYCIN* (Vol. 2). Elsevier.

Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature*, *34*(3), 1199–1235.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Brookings Institution Press.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*, 1273–1296.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285–305.

Tijssen, R. J. (2010). Discarding the 'basic science/applied science dichotomy: A knowledge utilization triangle classification system of research journals. *Journal of the American Society for Information Science and Technology*, *61*(9), 1842–1852.

Trajtenberg, M. (2018). Artificial intelligence as the next GPT: A political-economy perspective. In *The economics of artificial intelligence: An agenda* (pp. 175–186). University of Chicago Press.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463.

Valentini, G. (2012). Measuring the effect of M&a on patenting quantity and quality. *Strategic Management Journal*, *33*(3), 336–346.

Vohora, A., Wright, M., & Lockett, A. (2004). Critical junctures in the development of university high-tech spin-out companies. *Research Policy*, *33*(1), 147–175.

Wiggers, K. (2021). Retrieved from https://venturebeat.com/ai/ambi-robotics-which-develops-robot-grasping-systems-raises-28m/

Won, S., & Bidwell, M. (2023). Finding the right path to the top: How past interorganizational moves impact executive selection outcomes. *Strategic Management Journal*, *44*, 2341–2376.

Wooldridge, M. (2020). *A brief history of artificial intelligence*. Flatiron Books New York.

Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., Lyons, T., Manyika, J., Niebles, J. C., Sellitto, M., Shoham, Y., Clark, J., & Perrault, R. (2021). The AI index 2021 annual report, AI Index Steering Committee, Human Centered AI Institute, Stanford University, Stanford, CA, March 2021.

How to cite this article: Chattopadhyay, S., Honoré, F., & Won, S. (2025). Free range startups? Market scope, academic founders, and the role of general knowledge in AI. *Strategic Management Journal*, *46*(4), 1027–1079. https://doi.org/10.1002/smj.3685

APPENDIX

SECTION 1. CONTEXT: THE ROLE OF ACADEMICS IN THE DEVELOPMENT OF AI

AI is widely regarded as a general-purpose technology or enabling technology, with broad applicability across many different markets (Cockburn et al., 2018; Goldfarb et al., 2023; Trajtenberg, 2018). AI has applications in numerous fields such as robotics, finance, and medicine, and industries such as manufacturing, banking, and healthcare, and is also changing horizontal functional areas such as HR and marketing, which are relevant to all industries (Zhang et al., 2021).

Academics have played an important role in AI since its inception. AI was coined as a scientific field in 1955 when John McCarthy, a professor in mathematics turned computer scientist, started a summer school, gathering other academics interested in this fledging field (McCarthy et al., 1955). These summer schools offered a promising launching pad for the field as McCarthy himself, and many of the attendants created the first AI labs in their respective institutions—Stanford, MIT, and Carnegie Melon University (Wooldridge, 2020). In addition, academic researchers from cognitive science, logic, economics, and mathematics developed the first theoretical models used in AI (Wooldridge, 2020).

Within academia, AI as a field rose to fame several times during the 20th century but faced difficulties in sustaining momentum in research until the second decade of the 21st century. Nevertheless, prominent academics have led the development of significant technologies over the past five decades, which now form the basis of the AI-based ecosystem. Machine learning, based on neural networks, is one of the most significant developments within AI was first developed by academics (such as LeCun & Bengio, 1995; LeCun et al., 1998). Academics were also the first to explain how to expand it to implement deep learning (e.g., Hinton et al., 2012; Krizhevsky et al., 2012).

Throughout the evolution of AI, several academics went on to develop applications or to start their ventures. For instance, the MYCIN system co-developed in Stanford AI lab and medical school in the 1970s could infer what infectious diseases patients were most likely to suffer from by asking questions to physicians (Press, 2020; Shortliffe, 2012). More recently, Professor Rodney Brooks from MIT became one of the most famous academic entrepreneurs in robotics as the founder of iRobot and two other ventures (Atoji Keene, 2013; Feldman, 2020). Academics' long-standing contributions to the field are likely to shape the ways they approach knowledge-creation within their startups and the applicability of their knowledge to different market segments. Note that our data collection was completed before the diffusion of the generative AI technology.

SECTION 2. FIGURES AND TABLES

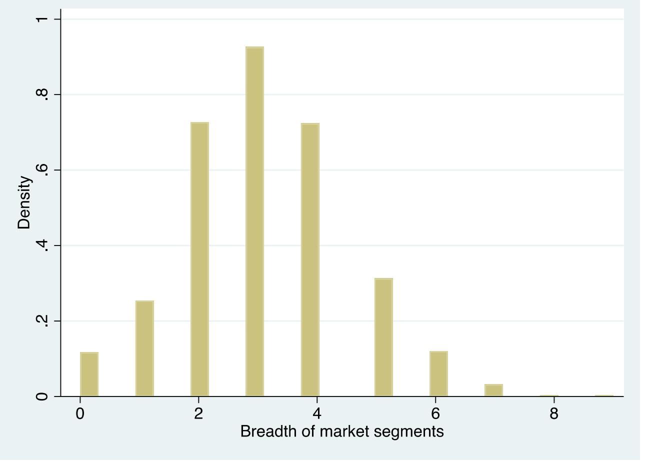

FIGURE A Market applicability (verticals).

TABLE A1 Comparison of teams with and without AE.

	Teams with AE	Teams without AE	*p*-Value from two-tailed *t* test
Team size	2.67	2.51	0.08
Firm age	4.82	4.78	0.89
Has a female founder	0.31	0.22	0.06
Average number of employers	5.31	6.00	0.01
Average experience	18.60	18.67	0.91
Has a serial founder	0.54	0.74	0.00
Has a foreign founder	0.55	0.44	0.04
Has a founder from top 10 inst.	0.20	0.10	0.00
Number of patents sold	0.54	0.59	0.93
Upstream	0.14	0.06	0.00
Avg. num of pubs	95.83	8.69	0.00
Avg. num of pre-venture pubs	58.90	6.28	0.00
Avg. num of pre-venture pubs about basic research	19.25	2.58	0.00
Proportion of pre-venture basic pubs	0.27	0.07	0.00
Avg. publication generality	5.14	1.60	0.00
Avg. num of patents	3.56	1.79	0.03
Avg. patent generality	1.40	0.70	0.00

TABLE A2 Comparison of teams with and without PhD holder among teams without AE.

	Teams without PhD holder	Teams with PhD holder	*p*-Value from two-tailed *t* test
Team size	2.44	2.66	0.00
Firm age	4.77	4.81	0.82
Has a female founder	0.19	0.28	0.00
Average number of employers	5.99	6.01	0.89
Average experience	18.40	18.94	0.25
Has a serial founder	0.76	0.70	0.05
Has a foreign founder	0.40	0.51	0.00
Has a founder from top 10 inst.	0.07	0.13	0.00
Number of patents sold	0.35	1.06	0.04
Upstream	0.04	0.10	0.00
Avg. num of pubs	0.62	24.48	0.00
Avg. num of pre-venture pubs	0.29	18.01	0.00
Avg. num of pre-venture pubs about basic research	0.07	7.51	0.00
Proportion of pre-venture basic pubs	0.01	0.17	0.00
Avg. publication generality	0.66	3.44	0.00
Avg. num of patents	1.15	3.04	0.00
Avg. patent generality	0.57	0.95	0.00

TABLE B1 Estimating the probability of having an AE.

	(1) Has an AE
Team size	0.01
	[.25]
Firm age	0.00
	[.44]
Has a female founder	0.03
	[.14]
Has a serial founder	−0.07
	[.00]
Average number of past employers	−0.01
	[.03]
Average experience	0.00
	[.08]
Has a foreign founder	0.03
	[.16]
Has a founder from top 10 inst.	0.09
	[.00]
Constant	0.10
	[.02]
Observations	988
R-squared	.04

Note: *p*-Values in brackets.

TABLE C1 Top 10 most common verticals.

	Teams with AE		Teams without AE	
		%		%
1	Artificial intelligence and machine learning	25	Artificial intelligence and machine learning	23.07
2	TMT	15.38	TMT	18.07
3	Big Data	12.5	Big Data	12.23
4	SaaS	6.25	SaaS	12.09
5	HealthTech	4.81	Mobile	4.09
6	Mobile	3.85	Marketing Tech	3.15
7	Robotics and Drones	3.37	HealthTech	2.31
8	Industrials	2.88	FinTech	2.1
9	Mobility Tech	2.4	Internet of Things	1.61
10	Autonomous cars	2.4	Robotics and Drones	1.5

TABLE D1 List of interviewees.

ID	Current profession	Has startup	Works for startup	Has ever worked for startup
Faculty 1	ECE professor and consulting	No	No	No
Faculty 2	ECE professor	Yes	Yes	No
Empl1	Employee at a large Tech Company	No	No	Yes
FTE1	First time entrepreneur	Yes	Yes	No
StartupEmp1	Employee at an early-stage startup	No	Yes	Yes
Faculty 3	Professor	No	Yes	No
Faculty 4	Professor	Yes	Yes	No
PhD Entrepreneur	Entrepreneur	Yes	Yes	No
PhD dropout	Entrepreneur	Yes	Yes	Yes

SECTION 3. SOLO FOUNDERS

In this article, we focused on new ventures founded by two or more founders based on extant research that ventures founded by a single individual may be qualitatively different. Upon speaking with faculty in engineering schools, we found that while some solo academics—faculty members and postdocs—may indeed be pursuing business ventures, many others tend to incorporate for other reasons, such as consulting for the Department of Defense and other government funding agencies or for promoting their research to obtain corporate funding. Prior literature also mentions that the founding conditions of firms, and founders' objectives may be different for startups led by solo founders versus teams (e.g., Agarwal et al., 2016; Delmar & Shane, 2006; Lee et al., 2024). Therefore, solo founders, especially academic solo founders in our setting may hold different objectives and motivations than those who partner with others and may introduce significant bias and noise in our estimates.

We found that the solo-founded ventures are indeed different from our sample of ventures with two or more founders. First, the representation of academic founders was much lower; there were only 11 teams with academic entrepreneur (1.7%, vs. 8.4% among two or more founder teams), and 111 teams with PhD-holding founder (15.68%, vs. 38% among two or more founder teams). Second, the publication rate was also less than half compared to new ventures with two or more founders. At the individual level, about 16% had at least one publication among solo founders, which is much lower than startups with two or more founders where 25% of the teams had at least one publication. Third, we examined the patenting rate. We found that 136 teams (19%) had at least one patent. This is a significantly lower patenting rate compared to our sample of teams with two or more founders (27%). Note that within the sample of solo founders, we did not find a difference between academic team versus others, in terms of whether the patent or not (18% for academics and 19% for others; $p = .93$ [two-tailed t test]), the average number of patents for those with patents (7.5 for academics and 8.7 for others, $p = .92$ [two-tailed t test]) as well as the average generality of patents (2.4 for academics and 1.8 for others, $p = .71$ [two-tailed t test]). This is possibly related to the above observation that the representation of academic founders is much lower among solo founders. Finally, and not surprisingly

TABLE E1 Subsample analyses among nonacademic PhDs: Evidence of attenuation.

	(1) Market applicability	Threshold: 3 (2) Market applicability	Threshold: 5 (3) Market applicability	Threshold: 10 (4) Market applicability	Threshold: 15 (5) Market applicability
Years of industry experience for non-AE PhDs	−0.02				
	[.22]				
Years of industry experience for non-AE PhDs (<=threshold)		0.19	0.11	0.05	0.02
		[.14]	[.14]	[.12]	[.31]
Years of industry experience for non-AE PhDs (>threshold)		−0.03	−0.03	−0.05	−0.07
		[.11]	[.07]	[.02]	[.00]
Team size	−0.03	−0.01	−0.01	0.01	0.01
	[.76]	[.87]	[.94]	[.92]	[.91]
Firm age	0.05	0.06	0.06	0.06	0.06
	[.17]	[.15]	[.14]	[.13]	[.14]
Has a female founder	−0.34	−0.36	−0.35	−0.36	−0.37
	[.05]	[.04]	[.04]	[.04]	[.03]
Has a serial founder	−0.04	−0.01	0.00	−0.01	−0.03
	[.84]	[.95]	[.99]	[.96]	[.85]
Average number of past employers	0.02	0.01	0.01	0.01	0.01
	[.61]	[.77]	[.78]	[.88]	[.86]
Average experience	0.01	0.01	0.01	0.01	0.01
	[.77]	[.70]	[.69]	[.77]	[.81]
Has a foreign founder	0.22	0.20	0.20	0.22	0.22
	[.19]	[.23]	[.23]	[.19]	[.20]
Has a founder from top 10 inst.	0.13	0.14	0.15	0.14	0.15
	[.60]	[.56]	[.53]	[.55]	[.52]
Log number of X followers	0.04	0.04	0.04	0.03	0.03
	[.08]	[.10]	[.12]	[.17]	[.21]
Constant	3.08	2.55	2.53	2.58	2.75
	[.00]	[.00]	[.00]	[.00]	[.00]
Observations	299	299	299	299	299
R-squared	.06	.07	.07	.08	.08

Note: *p*-Values in brackets. All models control for location dummies.

TABLE F1 Number of market opportunities and new venture outcomes.

	(1) Log(valuation)	(2) Liquidity event
Market applicability	0.14	0.01
	[.01]	[.11]
Has an AE	−0.06	−0.04
	[.82]	[.24]
Team size	0.21	0.02
	[.03]	[.14]
Firm age	0.10	0.03
	[.00]	[.00]
Has a female founder	−0.59	−0.01
	[.00]	[.65]
Has a non-AE PhD	0.35	0.01
	[.03]	[.77]
Has a serial founder	−0.04	−0.01
	[.82]	[.56]
Average number of past employers	−0.11	0.00
	[.00]	[.62]
Average experience	0.05	0.00
	[.00]	[.33]
Has a foreign founder	0.01	−0.02
	[.94]	[.37]
Has a founder from top 10 inst.	0.30	0.01
	[.23]	[.72]
Log number of X followers	0.16	−0.04
	[.00]	[.00]
Constant	0.42	0.01
	[.31]	[.90]
Observations	661	988
R-squared	.20	.21

Note: The main liquidity events captured from our data were acquisitions by other companies, which were experienced by 103 startups from our sample; we also observed two cases of IPOs. In total, 105 (10.25%) experienced liquidity events. *p*-Values in brackets. All models control for location dummies.

given above differences, solo-founded teams had on average lower number of verticals (2.6) compared to the two or more founder teams (3).

For all these reasons, we did not include the solo founders in our sample, but we found that the main finding that academic startups have higher market applicability was robust to combining the solo founders with our sample albeit having a weaker effect of having an academic entrepreneur compared to the effect size in our sample of two or more founders. Results are reported in Table G1. Note that the size of the coefficient of having an AE founder in Table G1

TABLE G1 The effect of having an academic entrepreneur on the number of market segments when including solo founders.

		Publications				Patents			Full model
		Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Market applicability	Market applicability	Market applicability	Average generality of pubs	Market applicability	Market applicability	Average generality of patents	Market applicability	Market applicability
Has an AE	0.27	0.29	0.28	1.45	0.24	0.55	0.26	0.23	0.21
	[.07]	[.10]	[.10]	[.00]	[.18]	[.00]	[.09]	[.14]	[.24]
Proportion of basic research			0.25	0.04	3.24				0.14
			[.22]	[.00]					[.52]
Log number of research publications		−0.01	−0.03	−0.03	0.81	−0.04			−0.06
		[.85]	[.51]	[.00]	[.33]				[.23]
Average generality of pubs					0.03	0.03			0.03
					[.09]				[.10]
Number of patents							0.09	0.01	0.00
						[.00]	[.05]	[.52]	[.55]
Average generality of patents									0.06
								[.01]	[.01]
Team size	0.12	0.12	0.11	0.00	0.12	0.22	0.12	0.11	0.10
	[.01]	[.01]	[.02]	[.98]	[.01]	[.00]	[.01]	[.02]	[.02]
Firm age	0.02	0.02	0.02	0.08	0.02	0.03	0.02	0.02	0.01
	[.12]	[.12]	[.11]	[.01]	[.14]	[.03]	[.15]	[.19]	[.23]

TABLE G1 (Continued)

	Publications					Patents			Full model
		Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**Market applicability**	**Market applicability**	**Market applicability**	**Average generality of pubs**	**Market applicability**	**Market applicability**	**Average generality of patents**	**Market applicability**	**Market applicability**
Has a female founder	−0.11	−0.11	−0.1	−0.1	−0.11	−0.07	−0.10	−0.10	−0.09
	[.23]	[.23]	[.28]	[.51]	[.24]	[.39]	[.27]	[.29]	[.30]
Has a non-AE PhD	0.33	0.34	0.22	1.16	0.30	0.22	0.32	0.31	0.29
	[.00]	[.00]	[.00]	[.00]	[.00]	[.01]	[.00]	[.00]	[.00]
Has a serial founder	0.04	0.04	0.05	−0.05	0.04	0.06	0.05	0.04	0.04
	[.57]	[.57]	[.55]	[.77]	[.59]	[.46]	[.55]	[.58]	[.62]
Average number of past employers	−0.03	−0.03	−0.03	0.08	−0.03	−0.03	−0.03	−0.03	−0.03
	[.05]	[.05]	[.06]	[.03]	[.05]	[.07]	[.06]	[.08]	[.07]
Average experience	0.00	0.00	0.01	−0.03	0.01	0.01	0.00	0.00	0.00
	[.36]	[.35]	[.32]	[.00]	[.31]	[.02]	[.41]	[.49]	[.42]
Has a foreign founder	0.11	0.11	0.11	−0.14	0.12	−0.25	0.11	0.13	0.14
	[.24]	[.23]	[.26]	[.34]	[.21]	[.01]	[.23]	[.18]	[.16]

TABLE G1 (Continued)

		Publications				Patents			Full model
		Number of publications	Basic research	General research					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Market applicability	Market applicability	Market applicability	Average generality of pubs	Market applicability	Market applicability	Average generality of patents	Market applicability	Market applicability
Has a founder from top 10 inst.	0.02	0.02	0.01	0.47	0.01	−0.14	0.02	0.03	0.02
	[.91]	[.91]	[.92]	[.04]	[.97]	[.34]	[.90]	[.86]	[.87]
Log number of X followers	0.06	0.06	0.06	0.00	0.06	0.03	0.06	0.05	0.05
	[.00]	[.00]	[.00]	[.85]	[.00]	[.01]	[.00]	[.00]	[.00]
Constant	2.30	2.30	2.32	0.11	2.30	−0.28	2.31	2.33	2.33
	[.00]	[.00]	[.00]	[.74]	[.00]	[.04]	[.00]	[.00]	[.00]
Observations	1696	1696	1696	1696	1696	1696	1696	1696	1696
R-squared	.05	.05	.05	.53	.06	.26	.06	.06	.06

Note: *p*-Values in brackets. All models additionally control for location dummies.

(solo and non-solo combined) is reduced compared to Table 2 (only non-solo). Given the reasons that we explained above, the generalized knowledge of solo academic founders may not have been effectively leveraged in their ventures toward higher market applicability.

SECTION 4. DESCRIPTIONS OF STARTUPS IN THE FIELD OF AI

As AI is a broad transversal field and not a specific industry, we described the characteristics of the AI new ventures based on their founder's background. We selected companies based on the number of publications, the availability of information, and the representativeness of different industries.

1. Academic startups co-founded by computer scientists

Founders	Founding history and products	Milestones by the end of 2023 (PitchBook)
Ken Goldberg (UC Berkeley Professor in robotics) and two of his postdocs, Jeffrey Mahler (PhD) and Stephen McKinley (PhD)	**Ambi Robotics'** dexterous robots originate from the three founders' extensive research in robotics. The three founders achieved a breakthrough in their research when they developed an operating system for robots so that the robots can learn the best way to grab objects based on camera images. They then decided to create a startup to sell dexterous robots to firms. They improved their operating system, now called AmbiOS, so that the robots can sort all kinds of objects and keep learning. While there can be many applications of this product in any line of work that requires sorting, the company obtained their first contracts in the logistics and parcel sorting industry (Wiggers, 2021)	Contract with Pitney Bowes $34.7 million raised 58 employees
Ciro Donalek (PhD and previously Computational Staff Scientists at CalTech), Scott Davidoff (PhD), Michael Amori (MS & MBA) and George Djorgovski (Professor of Astronomy and Data Science, and Director, Center for Data-Driven Discovery at Caltech)	**Virtualitics** distinguishes itself from other data visualization software by relying on AI to analyze the data in an intelligent way. The underlying algorithm uncovers relationships within the data, and spells out insights that the users may not have noticed on their own. Further, Virtualitics offers a 3D-interactive tool. The goal of the company is to facilitate the analysis of data in an interactive virtual-reality environment. Currently, the company works with the Department of Defense and offers a module integration to existing data warehouse software. The company raised an additional 37 million dollars in August 2023 and wants to pursue all types of businesses needed to analyze data to develop predictions (Sharma, 2023).	Contract with the Defense Department $73.79 million dollars raised 97 employees

Founders	Founding history and products	Milestones by the end of 2023 (PitchBook)
Nicholas Harris (PhD and postdoc at MIT in electrical engineering and computer science), Darius Bunandar (PhD), Thomas Graham (MBA)	**Lightmatter** is at the edge of photonic computing (relying on photons instead of electrons). The startup started creating chips that work with photons rather than electrons to decrease the chips consumption of power. The company's goal is developing chips and servers that can host the largest neural networks and to compete with Nvidia and other established chip producers. "We use light to link computer chips together and we also use light to do calculations for deep learning," said Harris (Lee, 2023). Such inventions would allow the AI revolution to take off even faster.	$420.28 million raised from venture capitalists and corporate venture funds 150 employees

Note: We selected Ambi Robotics and Virtualitics as their founders have the most publications among academics (#1 and #4, respectively). Lightmatter's founders are the postdocs with the most publications.

2. Academic startups co-founded with scientists in other fields

Founders	Founding history and products	Milestones by the end of 2023 (PitchBook)
Michael Abramoff Professor at the University of Iowa in ophthalmologist (MD-PhD) and Krassen Dimitrov (PhD)	**Digital Diagnostics**' lead founder is an ophthalmology professor who fully embraced the AI vision for the medical world. He imagined an AI-driven diagnostic platform for a series of pathologies. Given his specialty, the first product consists of an FDA-cleared product that detects diabetic retinopathy without a physician's input. The new venture is developing two applications in dermatology.	Generating revenues with the sales of retinopathy detection devices $130 million raised 113 employees
Omar Farha (PhD, Professor of Chemistry at Northwestern University), Christopher Wilmer (PhD, Associate Professor of Chemical Engineering at the University of Pittsburgh) and Benjamin Hernandez (MBA, JD, and BS)	Two of **NuMat**'s co-founders are world-renowned experts in metal–organic frameworks (MOFs). The MOF structure and molecules affect the realization of chemical reactions in a series of processes from gas separation to biochemistry reactions. Extensive research is conducted to find the right structure and molecules for each application. Numat created a "fully integrated platform combining physics-based modeling, proven engineering and real-world application testing to achieve high-volume manufacturing of MOFs" (Numat website, Our Process). Current users include firms in semiconductor and energy and the Department of Defense for chemical and bacterial warfare.	Running platform and two commercialized products $47.4 million raised 58 employees
Reza Zanjani (BS and MBA) and Marco Iotti (PhD, Professor in Food System at the University of Modena and Reggio Emilia)	**Mixfit**, founded by a Professor in Food systems and his co-founder, developed an application that recommends the intake of a personalized nutritional drink based on their past meals, activity level, and UV exposure among other environmental factors (Durrell, 2021).	Partnership followed by an acquisition by DSM in 2018 11 employees

Note: we selected these three startups to show the heterogeneity of academic startups whose founders' research came out of AI.

3. Nonacademic startups

Founders	Founding history and products	Milestones by the end of 2023 (PitchBook)
Andrew Eye (BS, serial entrepreneur) and Dave DeCaprio (BS)	**Closedloop.ai** developed an AI platform based on machine learning and healthcare machine learning libraries to predict the risk of health adverse events. They hope to increase the use of data science by providers and healthcare institutions to reduce the number of wrong diagnoses and to make the healthcare industry more efficient. The motivation to create this startup came from the personal experience of one of the founders. The founders' daughter received the wrong diagnosis and was on a waiting list for an organ transplant although a medical solution existed (Landi, 2021).	Won several innovation competitions in healthcare $48.14 million raised 81 employees
Mark Chung (MS), Jonathan Chu (MS) and Thomas Chung (BS)	**Verdigris** created a platform to optimize energy management in industrial and commercial buildings. One of the co-founders, an engineer by training had the idea after receiving a very expensive electric bill after returning home from his holidays. The electricity provider could not provide the reason for the peak in electrical consumption. He decided to monitor every appliance in his house with meters he bought in a retail store. He found the source of the problem (excessive energy used by broken rotor bars in the pool pump) and realized that similar issues might occur on a much larger scale for industrial and commercial buildings (Margolis, 2017). With his co-founders, he created a suite of sensors that track all energy consumption and reports the data to a platform. This startup capitalizes on the Internet of Things trend and energy monitoring. The company also provides a developer hub and shares its API. They offer solutions for real estate, hospitality, data centers and manufacturing facilities	$42.33 million raised 38 employees
Saurahb Kumar (BS and MBA) and Udaya Bashskar Reddy (BS and MBA)	**Rezolve.ai** (previously actionable science) is a good example of a company that uses chatbots to offer IT and HR service desks based on AI. Their solutions are integrated into Microsoft's Teams and incorporate OpenAI chatGPT to address around 65% of the desks' tickets (Duggal, 2023). The founders found a smart way to incorporate existing tools that both deliver and improve their solutions to customers.	$12.25 million raised 99 employees

Note: As we have a large number of nonacademic, non-PhD startups, we used different sample selection method. We first investigated Closedloop.ai whose founders published the most in this group (20 publications). We randomly selected Verdigris. We selected Rezolve.ai to represent the numerous startups that offer support functions.

SECTION 5. PITCHBOOK METHODOLOGY TO DETERMINE VERTICALS

We present definitions, conceptualization and operationalization of the term "verticals" from PitchBook.

1. From PitchBook online material

PitchBook defines verticals as "a niche market space." PitchBook developed verticals as a tool for investors (such as venture funds and angel inventors) to refine their search on the platform:

> *"Verticals provide a specially crafted and curated perspective on a niche market space. They are designed to slice across industries such that a single vertical may be comprised of companies that span multiple industries. Unlike industry codes, not every company is tagged with a vertical. A company is tagged to as many relevant verticals as possible, giving each vertical equal weight. We do not designate a primary vertical, as we do with industries. This article explains our methodology to ensure that the verticals on our platform are tagged in a way you can trust."* (Retrieved from PitchBook website on December 12, 2023).

PitchBook researchers associate verticals with firms based on the keywords they assign to each firm. The researcher team writes a company description and list of keywords for each company. According to PitchBook:

> *"Every company in PitchBook's database is accompanied by two types of descriptive textual information: a company description and a series of keywords and phrases. This information is crafted and curated by PitchBook researchers. Keyword data is used primarily to help with discovery and with assigning companies to verticals. To help the company appear in the correct search results, PitchBook researchers add keywords that contain unique variations, appropriate synonyms, and descriptions. Each company has approximately 3–10 keywords added separately from the description."* (Retrieved from PitchBook website on December 12, 2023).

Based on their explanation above, PitchBook researchers run an algorithm that assigns the keywords to a vertical. This matching is sometimes validated by a researcher.

"Verticals are constructed by using a series of keywords and phrases that help define a given space. The keywords that are linked to a vertical are run against the millions of companies in PitchBook's database, so companies with matching keywords attached to their profile may be automatically tagged to a vertical or require additional researcher validation to confirm the relationship. Companies are only tagged to verticals based on their keywords, not the words in their company description." (Retrieved from PitchBook website on December 12, 2023).

List of verticals from our sample.

3D Printing
AdTech
Advanced Manufacturing
AgTech
Artificial Intelligence & Machine Learning
AudioTech
Augmented Reality
Autonomous cars
B2B Payments
Beauty
Big Data
Cannabis
CleanTech
Climate Tech
CloudTech & DevOps
Construction Technology
Cryptocurrency/Blockchain
Cybersecurity
Digital Health
E-Commerce
EdTech
Esports
FemTech
FinTech
FoodTech
Gaming
HR Tech
HealthTech
Impact Investing
Industrials
Infrastructure
InsurTech
Internet of Things
LOHAS & Wellness
Legal Tech
Life Sciences
Manufacturing
Marketing Tech
Mobile
Mobility Tech
Mortgage Tech
Nanotechnology
Oil & Gas
Oncology
Real Estate Technology
Restaurant Technology
Robotics and Drones
SaaS
Space Technology
Supply Chain Tech
TMT
Virtual Reality
Wearables & Quantified Self
eSports

SECTION 6. CHATGPT CLASSIFICATION PROCESS

Section 6.1. Basic versus applied research

1. Manual coding

Our sample of entrepreneurs produced a body of work of 19,291 publications up to the end of 2021, among which 17,331 received more than two citations until August 2023. We retained this sample of publications and citations for our analyses. First, each of the three co-authors individually coded the same 100 abstracts, classifying them as basic or applied research based on how prior literature has defined them (Bieniawski, 1994; Nelson, 1959; Rosenberg & Nelson, 1994; Stokes, 1997; Tijssen, 2010). The co-author team then met extensively over a period of 4 weeks to resolve differences in the coding across the team. This entailed conducting our own research into the broad subject areas represented in our sample, and extensive discussions to understand better what kinds of publications would be considered basic or applied. For instance, in the healthcare domain (sometimes intersecting with life science), any identification of new genes or molecules is considered basic research. At the other end of the spectrum, clinical trials on humans constitute applied research. Similarly, in physical sciences, research on optimizing industrial processes was classified as applied. In computer science research (including in physical sciences), research testing algorithms on data were coded as applied while the mathematical development of an algorithm was coded as basic. Once we were confident that our understanding of basic vs. applied science accurately matched the nomenclature prevalent within each field, we then divided the 2000 abstracts such that each abstract was coded by two co-authors. This represented more than 10% of the total universe of the publications in our sample.

1. OpenAI ChatGPT fine-tuning—Adopted solution

Given the large volume of data and the significant time it took to manually read and understand each abstract, we refrained from manually coding all 19,291 publications. Instead, we concluded that using generative AI fine-tuning was the best option for automated coding given the ability of large language models to handle the volume of data and variety of topics covered by the abstracts.

We hired an RA who had experience using Open AI ChatGPT and programming in Python. The RA and the coauthor team developed a fine-tuned program in Python using the API of Open AI ChatGPT 3.5.[35] To fine-tune the model, we carefully prepared the abstracts from which the program learns and the prompt that provides the instruction and definitions to the program.[36] Based on the publications' Scopus subject areas, most of our abstracts belonged to healthcare, life sciences, physical sciences, and a smaller share in social sciences. We made sure that each subject area was proportionally represented in the learning sets as well as the variety within each subject area. We hand-selected 70 abstracts from the first 100 we all coded earlier: 20 in the health subject area, 20 in the life science one, 20 in the physical science one, and 10 in the social science one.

Fine-tuning ChatGPT consisted of two parts: learning and validating. We first made ChatGPT learn on 35 abstracts following the aforementioned proportions by subject area. We then validated the learning on another 35 abstracts. After we fine-tuned the model, we checked the accuracy of the model by running the model on 2000 abstracts that we coded manually. We obtained an average matching rate of 78% (since we had two human coders for each of the 2000 abstracts, this is the mean of the matching rate between ChatGPT and coder A, and that of ChatGPT and coder B). We deemed this matching rate to be sufficiently high, given that a correlation of 0.7 is commonly employed as a benchmark for interrater reliability (Taber, 2018). Hence, we applied the program to the rest (17,291) of the abstracts. Then, 33% were coded as basic and 67% as applied. In the table below, we report the proportion of basic research within each subject area: healthcare, life sciences, physical sciences, and social sciences. Since AI is inherently more applied, we believe that we are seeing a higher occurrence of applied research in our sample than other fields such as life sciences.

Proportion of basic publications across categories.

	Healthcare	Life sciences	Physical sciences	Social sciences
Proportion of basic research	0.14	0.43	0.34	0.31

[35]Upon testing the performance on a subset of abstracts using ChatGPT 4, we concluded that the ability to classify abstracts is comparable to that of ChatGPT 3.5. This is in line with observation that ChatGPT 4 performance is superior in tasks that requires more intense computing (e.g., Ahsan et al., 2023). We used ChatGPT 3.5 as the cost to run the model on all 19,291 publication abstracts was significantly lower than ChatGPT 4, while giving us a similar accuracy rate.

[36]We provided the prompt as follows: *"You are an expert in Science. Technology, Engineering, and Mathematics. Classify abstracts as \\\"applied\\\" and \\\"not applied\\\" research. An abstract should be applied when the research directly applies to an industrial or business application in the real-world context. Not applied should be research that is basic. Basic research is experimental, empirical or theoretical work undertaken primarily to acquire new knowledge or method or technique or test. If a new method or technique or knowledge or test is provided, it is basic. It is about understanding underlying mechanisms, phenomena and properties without immediate industrial or business application. If the abstract*

SECTION 6.2. UPSTREAM VERSUS DOWNSTREAM

We developed a fine-tuning program in openAI ChatGPT to code the location of the company within the AI ecosystem (upstream vs. downstream) based on the company descriptions from PitchBook.

We first manually coded 100 descriptions. Then, we developed a fine-tuning program. We carefully selected 40 descriptions: 20 descriptions for learning and 20 for validation. Given the limited number of upstream firms, we could only provide 25% of upstream description for these two steps. We wrote a prompt with the request to classify the abstract and the definitions of upstream versus downstream based on the definition adopted by the participants in the field of AI.[37] We then tested ChatGPT classification against ours. The average accuracy was 94% (which means that in 94% of the cases, GPT and the human coder had given the same classification to the firm). We ran the model on all company descriptions. We obtained 92% of businesses classified as downstream and 8% as upstream.

has a mix of applied and basic, it should be regarded as applied. Based on this principle, classify the following abstracts.\"\n"."

[37]We wrote this prompt:

"You are a research analysis assistant for AI business analysis. Based on this description: classify these AI businesses as being upstream or downstream. The definitions are as follows. Being upstream means that the main activity of the business relates to the development of AI algorithms, models, and technologies. Other upstream tasks would involve data collection, algorithm design, and training models. The class for upstream is U. Being downstream means implementing AI technologies into real-world applications or systems. The AI technology would be put to use. The class for downstream is D. If mobile applications are included, then the business would be downstream. State the classification in the first line of your response."