Entrepreneurial Entry Over the Technology Lifecycle*

Innessa Colaiacovo[†]
University of Oregon
Lundquist College of Business

Daniel P. Gross[‡]
Duke University
Fuqua School of Business
and NBER

Jorge A. Guzman [§]
Columbia Business School
and NBER

March 2026

Abstract

We use data on nearly all U.S. business registrations since the early twentieth century to study entrepreneurial entry following the emergence of 860 influential technologies. We document a large and varied set of entrants, spanning technology producers, component suppliers, distributors, integrators, service providers, and more. Both the level and nature of this entrepreneurial entry evolve over time following a technology's emergence: entry rises, falls, and shifts from upstream, manufacturing-related activities to downstream activities such as aftermarket parts and service. We show that aggregate industry lifecycles are composites of overlapping sub-dynamics across the value chain. Our findings reveal empirical regularities about the scope and evolution of entrepreneurial entry around technological innovation that generalize across a large and diverse set of technologies, providing systematic evidence complementing prior case studies of ecosystems around specific innovations.

Keywords: entrepreneurship, innovation, ecosystems, industry dynamics, value chain

*We are grateful to the editor (Rajshree Agarwal) and two anonymous referees for constructive and generative comments. We also thank Ashish Arora, Tim Simcoe, Daniel Keum, Naomi Lamoreaux, John Wallis, participants at the University of Utah Strategy seminar, the DRUID, Doriot, and Strategy Research Forum annual conferences, the AOM annual meeting, the NBER Innovation Information Initiative and DAE meetings, and especially discussants Oliver Alexy and Vera Rocha, for helpful comments. All errors are our own.
[†]Address: University of Oregon Lundquist College of Business, Eugene, OR, USA; email: innessa@uoregon.edu.
[‡]Address: Duke University Fuqua School of Business, Durham, NC, USA; email: daniel.gross@duke.edu.
[§]Address: Columbia Business School, New York, NY, USA; email: jag2367@columbia.edu.

1 Introduction

A central question in strategic management is where and when new market opportunities emerge. A significant literature examines opportunities engendered by technological innovation (Schumpeter 1942, Tushman and Anderson 1986) and how firms create and capture value from new technology (Teece 1986). A long tradition of research has in particular identified entire industries born of breakthroughs, studying which firms enter, how that entry unfolds over time, how industries evolve as technologies mature, and how firms compete at each stage of the industry lifecycle (Abernathy and Utterback 1978, Gort and Klepper 1982, Suarez and Utterback 1995, Klepper 1996, Agarwal and Gort 1996, Agarwal and Tripsas 2008).

In the classical industry evolution literature, industry dynamics are typically evaluated based on the entry and exit of firms that produce a specific technology-based product. Entry, growth, shakeout, and consolidation among these producers are taken as the primary markers of opportunity and competition, and industry lifecycles are often characterized accordingly. However, more recent work complicates this picture: detailed studies of individual industries have shown that substantial entrepreneurial activity can occur outside technology production, including among firms supplying components, distributing products, integrating systems, or providing complementary services. In settings ranging from agricultural biotechnology to mobile money, solar energy, semiconductors, and bionic prosthetics, researchers have documented the importance of non-producing firms in shaping how technologies are commercialized, diffused, and used (Adner and Kapoor 2010, Moeen 2017, Wormald et al. 2021, Szerb and Furr 2025, Kim et al. 2025, among others). Conceptual syntheses have drawn on this evidence to emphasize the role of heterogeneous actors and broader ecosystems in industry emergence and growth (Moeen et al. 2020, Agarwal et al. 2025). Yet despite these advances, systematic evidence is lacking on non-producer firm entry and how it evolves over time. Whether patterns observed in a small number of intensively studied industries reflect general features of technological change, or are context-specific, remains an open question. Addressing this question has historically been difficult due to the necessity of data that span many technologies, long time horizons, and a wide range of firm types all at once.

Using data on nearly all U.S. business registrations since the turn of the twentieth century, we examine firm creation associated with 860 influential technologies. We link firms to technologies using words in firm names and classify firms by economic sector and value chain activity, allowing us to characterize how many firms enter with direct linkages to new technologies, what they do, how these patterns change as technologies mature, and what these patterns may reflect about changes in commercial opportunity over the technology lifecycle. We find that entrepreneurial entry around technologies is large, predominantly non-manufacturing, and shifts systematically from upstream (production) to downstream (aftermarket) activities as technologies mature.

To systematically study technology-enabled entrepreneurship, we develop a new dataset linking technological innovation to business creation at scale. We first use patent data to identify nearly 1,000 major technologies of the past century, using keywords from U.S. patent publications. We next use administrative business registration records from 47 states—which include all new corporations, limited partnerships, and limited liability companies—to measure over a century of firm creation, and a newly-developed procedure to classify these firms into economic sectors via words in firms' names. We link firms to innovation by identifying firms that include one of our focal technologies in their name. The logic of this approach is that the invocation of a technology in a firm's name explicitly indicates it is closely related to, or even built upon, that technology.[1] We validate this inference with a hold-out sample and by examining the 4-digit SIC industries that firms with these terms are most likely to be associated with in commercial firm directories like Dun & Bradstreet, where industries will often reveal firms' technological connections.

Our analysis is organized around three specific sets of empirical questions. First, how much entrepreneurial entry is associated with new technologies, and how is this entry distributed across different types of firms? Second, how does the composition of entrepreneurial entry change over time following a technology's emergence? Third, how do aggregate patterns of industry evolution relate to entry dynamics within different segments of the value chain?

We in turn document three findings. First, entrepreneurial entry around new technologies is substantial in scale and extends well beyond firms that manufacture the technology itself. Across the twentieth century, roughly 1-2% of all new U.S. businesses registered in a given year explicitly reference one of the technologies in our sample. For example, at their peak in the 1920s and 1980s, respectively, radio- and computer-related firms comprised roughly 1 of every 200 new businesses in the U.S. The majority of these firms are in non-manufacturing industries and exhibit characteristics consistent with small, locally oriented enterprises rather than high-growth, innovation-driven startups. This population includes firms supplying components, distributing products, installing equipment, maintaining capital, and providing specialized services.

[1] Business names are intended to be a faithful description of the business and are often direct and transparent. Prior work has highlighted how the choices entrepreneurs make at founding, including the company's name, are indicative of their identity and market orientation (Belenzon et al. 2020, McDevitt 2014).

Second, the composition of entrepreneurial entry evolves systematically over time. Early in a technology's lifecycle, entry is relatively more concentrated in upstream activities related to research, development, and manufacturing. As technologies mature, however, entrepreneurial entry increasingly shifts toward downstream activities such as distribution, parts, and service. These changes are visible across multiple measures, including sectoral classifications, value-chain activities, and specific words used in firm names, where words like *manufacturing* and *engineering* are more likely to appear in firm names early in the technology lifecycle, and words like *maintenance*, *parts*, and *service* become more common later. Importantly, this downstream migration of entry often continues even as entry into technology production declines.

Third—and perhaps most fundamentally—when entrepreneurial entry is disaggregated and lifecycles are measured for individual industry segments, it becomes apparent that the aggregate industry lifecycle is a composite of overlapping sub-industry dynamics. Classic lifecycle patterns (e.g., growth, peak, and decline) remain visible at the aggregate level, but the aggregate pattern masks substantial heterogeneity in the timing and nature of opportunity across the value chain. From this perspective, declining entry among producers does not necessarily signal the exhaustion of entrepreneurial opportunity and industry stagnation or decline; instead, opportunity on average migrates across segments as technologies age and markets evolve.

These findings contribute to several literatures. First, they extend classic work on industry evolution by broadening the unit of analysis beyond technology producers. Whereas prior studies have provided rich insights into entry and competition among producers (Klepper and Graddy 1990, Klepper 1996, Agarwal and Gort 1996), our results show that a much larger share of entrepreneurial activity occurs in complementary segments, and that this activity has distinct temporal patterns. In emphasizing gross entry, we are arguably closest to Agarwal and Gort (1996) in this regard. At the same time, our findings also complement detailed single-industry studies by providing systematic evidence (across many technologies and long time periods) of substantial and predictably evolving entry elsewhere in the value chain.[2] In doing so, we show that some findings of this literature are not idiosyncratic to their specific contexts but rather represent more general features of technology-based industry evolution, and we identify new regularities that are difficult to recognize without a large multi-industry sample of the kind we have compiled. Putting the two together, the central insight of this paper is that a focus on innovative firms overlooks a wider range of entrepreneurial

[2]Among many examples, see Ozcan and Eisenhardt (2009), Moeen and Mitchell (2020), and especially Hannah and Eisenhardt (2018), Wormald et al. (2023), Kim et al. (2025), or Szerb and Furr (2025) for focused studies of how firms have navigated ecosystems in emerging technology-based industries.

activity which emerges at different times and evolves over different horizons.

Though our principal goals are descriptive and analytical, the findings also relate to research in entrepreneurial strategy. Although entrepreneurial strategy research often frames opportunity identification as a problem of choosing among emerging technologies (e.g., Gans et al. 2021), our evidence suggests that for many entrepreneurs, opportunity may instead lie in choosing where to position within a technology's value chain, and that these opportunities vary systematically over time. While our data do not allow us to evaluate firm performance or competitive outcomes directly, they reveal a large and patterned set of entrepreneurial responses to technological change that have received relatively limited attention in strategy research. Understanding the strategies these firms employ to identify opportunities and capture value represents an important frontier for research on entrepreneurship around technological change.

Our approach nevertheless has limitations. Our analysis is restricted to a select set of technologies meeting two conditions: they were heavily developed over time, and they are observable in the patent record. We are thus unable to make statements about business creation around *all* new technologies, including those that are unsuccessful, short-lived, or not patented or patentable. Additionally, mentions of a technology in firm names incompletely measure its actual use. For example, the word "computer" does not appear in the name of prominent companies like Gateway or Microsoft. However, these are not the modal firms in the economy: far more common are firms like "Personal Computer Services Inc." or "Computer Rental Corp." (real firms in our data). These considerations mean that, even with comprehensive data on firm registrations, we can only partially characterize entry related to individual technologies.

Beyond the methodological limitations related to the sampled technologies and firms' links to them, the main constraint is a measurement challenge: we can observe firm entry but not exit, performance, or growth.[3] As a result, our findings should be interpreted as documenting patterns of entrepreneurial entry, rather than characterizing realized returns or competitive dynamics. Taken together, however, our results suggest a reinterpretation of entrepreneurial dynamics around new technologies: rather than viewing industry evolution solely through the lens of producer entry

[3] An additional limitation is intrinsic to business registration data as an indicator of entry, which specifically measures the creation of legal entities. The business registration data include all S-Corps, C-Corps, LLCs, and LPs and LLPs, including many small non-employer businesses, but will not cover unregistered sole proprietorships or informal businesses. A different population-level administrative data source is the Census Bureau's business formation statistics, which are derived from applications for employer tax identification numbers and can include sole proprietors—but these data are limited in different ways, particularly in their restriction to employers, which excludes independent operators. We consider business registration data sufficient for our aims.

and consolidation, we show that entrepreneurial opportunity often persists and evolves across the value chain as technologies mature. This perspective helps reconcile classic lifecycle patterns with recent ecosystem-based accounts and highlights the importance of considering the full range of entrepreneurial activity associated with technological change.

The paper proceeds as follows. Section 2 describes our data and methodological approach to identifying technologies, measuring firms' relationship to them, and associating firms to economic sectors or specific activities. Section 3 presents descriptive evidence on entrepreneurial entry around the technologies in our sample, illustrating variation across technologies in the types of associated new ventures and lifecycle patterns through detailed examples. Section 4 presents our core empirical analysis of how entrepreneurial entry disperses across sectors and migrates down the value chain. Section 5 concludes by discussing our findings in relation to prior literature, limitations and boundary conditions, and questions left for future work.

2 Data and Measurement

To undertake this analysis, we develop a new dataset to measure entrepreneurial activity that emerges around technological innovation. This dataset, in turn, requires identifying specific technologies, measuring firm creation, and drawing links between them.

2.1 Identifying new technologies through patents

Measuring technologies is less straightforward than might initially appear. There is no single authoritative list of technologies for us to study, in part because a technology, as commonly understood, is not unitary: "technology" often represents a bundle of innovations. Though often used in research, patent data present challenges: patents are written on inventions which are often narrow and comprise parts, devices, or processes which are embodied in products, rather than final products themselves. Patent classes are broader than any one patent alone, yet they can be hard to map to individual technologies (as commonly understood).

Instead of patents or patent classes, we use patent keywords to identify discrete technologies that each patent is associated with, in plain-language—following a growing body of work in management and other fields which takes similar approaches (e.g., Packalen and Bhattacharya 2012, Arts et al. 2021, Gross 2023, or Kalyani et al. 2024). These keywords were obtained from Google Patents public data (using BigQuery) and represent each patent's "top 10 salient terms extracted from the patent's title, abstract, claims, and description."[4] Keywords are typically technological, and sufficiently general that they appear across multiple patents and often clearly identify the technology(ies) to which a given patent relates: for example, some frequently-recurring keywords include *television*, *radio*, *semiconductor*, and *internet*. Other keywords are simply terms that co-occur in patents with these technologies, such as *receiver*, *signal*, *substrate*, or *data*.

We retrieve these keywords for all patents issued by the U.S. Patent and Trademark Office (USPTO) since 1836 (when the patent record begins) and use them to identify new technologies as follows. We first identify the first use of each keyword in the patent record and measure the total number of patents with that keyword (through 2009). We then filter to the 200 most heavily used keywords introduced each decade from the 1880s (thus capturing influential 20th century technologies such as the automobile) to the 2000s (covering modern technologies such as the internet)—drawing uniformly across decades to create a representative sample.[5]

As the above examples illustrate, not all keywords are technologies as typically conceived. We therefore manually review all keywords that are not, by our reading, obviously or unambiguously identifying of a technology in non-patent contexts, and removed these from our sample. For instance, the term "programming" appears in our list, but in addition to software engineering, it also has non-technological uses (e.g., community programming). We also drop biological or chemical entities (e.g., "alkylene"), units of measure (e.g., "pixel"), and terms with less than three characters, as well as a handful of terms related to nuclear physics (e.g., "atomic"), which we have found are often used colloquially in firm names. The end result is a set of 860 technologies. Of these, 658 technologies can be linked to firms (discussed below). These comprise the sample for our analysis (see Appendix Table B.1 for a sample list of these technologies).[6]

This approach comes with a few limitations. One is that our sample is conditioned on the total number of patents filed on a technology, which implies we will be studying technologies that with hindsight were influential. Since we do not know how selection creates differences between these technologies and others, we are not able to make statements about the population of all new technologies an entrepreneur may observe at any given moment in time, nor about entrepreneurship

[4]No formal documentation is available on Google's methodology, but the data developers shared with us over email that these terms are obtained from a procedure similar to TF-IDF (a common method for measuring words' importance in textual documents), with corrections for very common terms and bigrams. This method aims to find terms that are common in a patent, but rare across the population of patents.

[5]We use natural language processing tools to singularize plural nouns before performing this aggregation.

[6]Though we explored other refinements such as grouping seemingly-related terms (e.g., automotive and automobile), ambiguity over when to group terms—particularly when they could be considered synonyms but are not perfect substitutes—led us to err towards caution by keeping these terms independent.

around ultimately unsuccessful technologies. Additionally, we sample on patents, and will therefore miss innovations that are not patentable or patented, including many organizational innovations and most business process innovation.[7] Finally, our use of text-based measures raises the question of whether, and how, language has evolved over time. Language associated with novel ideas may develop alongside them, and initial patents may not fully capture the innovation in contemporary language. We believe this is unlikely to be problematic in our case, as most of the keywords in our sample balance specificity with genericity and entered the English lexicon when they began to grow. However, care in interpretation is nevertheless advisable.

2.2 Measuring firm creation through business registration records

We measure firm creation using state-level business registration records. Business registration is the act of creating a new limited partnership, a corporation, or—since 1993, in most of the U.S.—limited liability companies (LLCs). As such, it represents the legal founding of a company and, in most cases, an anchoring moment of entrepreneurship. This provides us a way to measure business creation over long horizons, as this paper requires, with sufficiently consistent definitions across space and over time. We explain these measures next.

While U.S. corporations have been created since at least the 1600s, it was traditionally a slow process that was tightly controlled by the state (pre-American Revolution, by England; later, by state governments). Since creating new corporations often required state legislatures to pass an act providing it a charter, there was historically a considerable difference between overall entrepreneurship and legal firm formation. The incorporation process was gradually opened up in the late 1800s, beginning with New Jersey in 1896, and quickly expanding to other states. By roughly 1915, two features characterized business incorporation across the U.S. First, all states allowed general (i.e., open) incorporation. Second, firms in every state could choose a legal jurisdiction (a sort of statutory domicile) different from their physical headquarters location. Incorporating firms in practice typically chose between two options: their local state law, which advantages new firms operating in the state, and Delaware corporate law, which is more beneficial for larger firms that engage in business across states, have a large operation, or intend to list on the stock market. In a more recent, post-1988 sample, being registered in Delaware is associated with a 23 times higher probability of reaching a high growth outcome (Guzman and Stern 2020).

[7] The U.S. Federal Court of Appeals established the patentability of business methods in a 1998 decision in the case of State Street Bank v. Signature Financial Group (149 F.3d 1368 (Fed. Cir. 1998)).

We collected data on all business registrations across 50 U.S. states through the Startup Cartography Project (Andrews et al. 2022), of which 47 included historical data.[8] Each record provides the company's name, physical location, state of registration, jurisdiction of choice, corporate form (i.e., corporation versus LLC), and registration date. These data allow us to measure all new legal entities registered in the U.S. since the late 1800s, providing us a rich view of entrepreneurship over time. The major limitations, in our view, are twofold. First, the data do not include firms' industry or business characteristics, which motivates our developing a new way to categorize businesses into sectors. Second, they specifically measure the creation of legal entities, and do not include unincorporated sole proprietors or general partnerships.

Focusing on new legal entities allows us to study an anchoring moment in firm formation that maps naturally to a definition of entrepreneurship: it is the moment a business idea gets translated into an independent organization tasked with executing on it. This is a consistent definition that our data allow us to apply over our entire sample. One implication of this approach, however, is that our measures of firm creation will be sensitive to changes over time in the propensity for incorporation over other organizational forms, including unregistered partnership and sole proprietorship.[9] To limit these concerns, we include technology fixed effects in our analysis where we can, and condition on invention half-decades (and, in robustness checks, invention year; see Appendix C) to compare technologies born into similar institutional environments.

2.2.1 Industry classification

Registration data do not include a firm's line of business. In administrative datasets such as those of the U.S. Census Bureau, firms' industry is categorized into a standardized industry classification by professional coders and proprietary recommendation algorithms, using data collected by mandatory-response census instruments. State offices that register businesses, however, do not classify new registrations into industries, nor do they consistently collect information to support a direct ex-post industry classification by administrators or researchers.

For our aims in this paper, however, we wish to know what business a firm is in. To do so, we build on recent research showing that firm names predict performance and growth potential (McDevitt 2014, Belenzon et al. 2017), and consider that firm names might also provide useful

[8] The Startup Cartography Project primarily documents entrepreneurship since 1988, but we are able to retrieve the full history of firm registrations, both active and deceased, from nearly all states.

[9] As Guinnane et al. (2007) point out, incorporation patterns have varied across U.S. history, especially in the postwar period as corporate and personal income tax rates diverged, preferencing incorporation.

information about industry. We develop a procedure that associates words in firm names with each of ten economic sectors, which we can use to classify firms to sectors. Concretely, we use Dun & Bradstreet (D&B) data, which include both firm names and Standard Industrial Classification (SIC) codes, to train an algorithm that measures the frequency with which individual words in firm names appear in 10 high-level sectors (into which the U.S. SIC classification is organized) and apply this to impute firms' sectors based on their names.[10] This approach mimics methods used by the U.S. Census Bureau to classify firms into industries, which uses a combination of firm names and one-line business descriptions (Kearney and Kornbau 2005), and by D&B itself, which at times uses firm names to impute their industry (Cramer 2017).

As we show in Appendix A, this procedure does a reliably good job of identifying firms' economic sectors in multiple validation samples. It comes with the added advantage that it supports a more flexible classification than SICS or NAICS: firms can be fractionally associated to multiple sectors and thus span sectoral boundaries, as many firms do in practice.

2.2.2 Measuring value chain activities

Following a similar logic, we associate firms to specific value-adding activities related to technological innovation using words in firm names. We focus on five categories which can be viewed as sequential steps in the value chain—research, manufacturing, sales, distribution, and service, omitting other activities that might be more nebulously defined or measured (e.g., marketing). In contrast to our sector classification, which takes a data-driven approach to fractionally associate firms to each of ten economic sectors, we connect the firms in our sample to these five value-adding activities via hand-chosen words and substrings (such as *manufact*, *store*, *service*, or *repair*), which prioritizes precision over completeness. When this procedure associates a firm to multiple activities (because, e.g., the firm's name has both "manufacturing" and "supply" in the name), we discretely measure it as engaged in its most upstream associated activity.

2.3 Linking technologies and firms

The last step in our data collection is to draw links between innovations and firm creation. To do so, we take our initial set of 860 technologies, and search for these technologies in firm names. In doing so, we measure firms which explicitly self-identify with each technology, which we assume indicates that the firm is closely related to that technology. Within technologies, we then aggregate

[10] These sectors are: Agriculture/Forestry, Construction, Finance/Insurance/Real Estate, Manufacturing, Mining, Public Administration, Retail Trade, Services, Transportation/Public Utilities, and Wholesale Trade.

registrations to count firms for all years, by year, and of various types (e.g., manufacturing firms vs. others, or Delaware vs. local jurisdiction—a signal of growth intension, per Guzman and Stern 2020). We link a total of 658 technologies to new firms in this way.

2.4 Limitations and validation

Our approach, while systematic, also has limits. There are almost certain to be false negatives: firm names will not fully measure firms' relation to new technologies, both because some names do not explicitly convey the business a firm is in and because some firms may be engaging with multiple technologies or a single technology may not be sufficiently characteristic of the firm to be in its name. There is also a risk of false positives: firms that mention a technology in their name without any substantive relationship to it. Though it seems unlikely that firms misrepresent their line of business, particularly given that names are used to signal firm types (McDevitt 2014), some technological language may be faddish and its use represents marketable naming conventions rather than substantive signals of linkages to specific technologies.

To evaluate whether the presence of a technology's keyword in a firm's name is a meaningful indicator of the firm's line of business, we examine the industries that firms in our sample are classified to in the Dun & Bradstreet data. Table 1 lists all technologies in our data with at least 500 associated firms in D&B, and the top three 4-digit SIC industries they are in (according to the D&B data). The evidence is affirming: across the table, these industries generally reflect the technology in question. The top three industries associated with "RADIO" are radio and television repair (35.7%), radio and television stores (17.1%), and radio broadcasting stations (17.1%). The top three industries associated with "COMPUTER" are computer and software stores (19.1%), computer related services (13.5%), and computer maintenance and repair (9.8%). The most closely associated with "DIESEL" are general automotive repair (60.6%), industrial machinery and equipment (10.1%), and repair services (6.1%). Other examples are similar.[11]

Our validation so far considers whether firm names can identify a general relationship to new technology. The next question is whether co-occurring words in firm names are reliable indicators of the sectors and value chain activities that firms are engaged in. Both here and in Appendix A

[11] As additional validation, Appendix Tables B.1 and B.2 show the firm registration counts for technologies with the most patents and most associated firms (respectively), the most common other words in firm names, and three randomly-chosen associated firms as examples. Both tables provide additional validating evidence: Appendix Table B.1 shows that the most commonly-occurring paired word in the names of firms associated with these technologies is in fact "TECHNOLOGY", and Appendix Table B.2 illustrates the range of businesses that can be associated with a given technology through the random examples it presents.

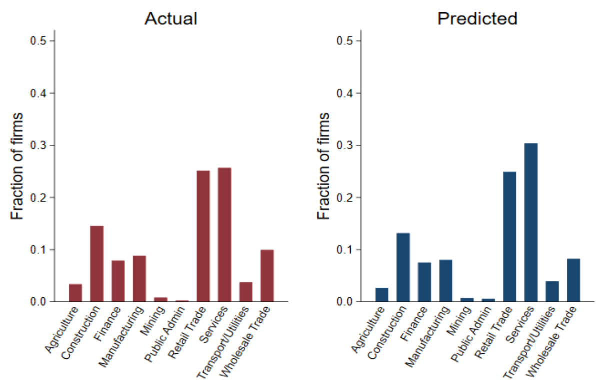
we run an extensive battery of validation tests for our sector classification method, first within a D&B hold-out sample, and then in an independent dataset of U.S. firm listings from InfoGroup USA. Figure 1 first presents the aggregate view, showing that the distribution of predicted sectors in the D&B holdout sample closely matches that of reported sectors. We then show in Appendix Tables A.3 and A.4 that at the firm level, the D&B-reported sector matches our highest-probability predicted sector roughly 70% of the time, and one of the top two predicted sectors 85% of the time; in all sectors, the predicted sector is most likely to be the actual sector. We consider these high rates, as we do not expect them to be 100%, due to the fact that many firms span industry boundaries (creating noise in classification, both by us and by D&B), as well as potential error in the source (D&B itself sometimes reclassifies establishments across SICs). We further find these patterns are stable over time and similar in an independent InfoGroup firm sample.

Table 1: SIC industries most heavily associated with specific technologies in D&B sample

	Largest SIC		2nd-largest SIC		3rd-largest SIC	
Technology	Share	Description	Share	Description	Share	Description
AIRCRAFT	16.8%	Airports, flying fields, and services	13.5%	Transportation equipment and supplies	13.1%	Automotive dealers, nec
ALTERNATOR	56.2%	Automotive repair shops, nec	12.1%	Motor vehicle supplies and new parts	10.2%	General automotive repair shops
AUDIO	21.1%	Radio, television, and electronic stores	8.9%	Radio and television repair	7.0%	Radio And Television Stores
AUTOMOBILE	22.9%	Membership organizations, nec	13.0%	New and used car dealers	10.6%	Used car dealers
AUTOMOTIVE	45.4%	General automotive repair shops	15.4%	Motor vehicle supplies and new parts	7.8%	Auto and home supply stores
BUTANE	71.6%	Liquefied petroleum gas dealers	9.2%	Petroleum products, nec	3.9%	Toys and hobby goods and supplies
CARDIAC	55.3%	Offices and clinics of medical doctors	9.1%	Medical laboratories	6.1%	Specialty outpatient clinics, nec
CATALYST	14.9%	Management consulting services	5.6%	Industrial inorganic chemicals, nec	5.1%	Business consulting, nec
COMPUTER	19.1%	Computer and software stores	13.5%	Computer related services, nec	9.8%	Computer maintenance and repair
COMPUTERIZED	12.9%	Accounting, auditing, and bookkeeping	5.9%	Data processing and preparation	4.9%	Custom computer programming services
DIESEL	60.6%	General automotive repair shops	10.1%	Industrial machinery and equipment	6.1%	Repair services, nec
ELECTRONIC	15.7%	Radio and television repair	15.4%	Electronic parts and equipment, nec	8.8%	Radio, television, and electronic stores
EXTRUSION	17.2%	Aluminum extruded products	12.7%	Plastics products, nec	11.9%	Special dies, tools, jigs, and fixtures
FIBERGLASS	12.5%	Boatbuilding and repairing	10.9%	Repair services, nec	9.0%	Special trade contractors, nec
GENETIC	12.2%	Livestock services, except veterinary	11.2%	Medical laboratories	7.7%	Commercial physical research
HELICOPTER	19.9%	Air transportation, nonscheduled	11.3%	Crop planting and protection	9.2%	Air transportation, scheduled
HVAC	79.8%	Plumbing, heating, air-conditioning	6.2%	Warm air heating and air conditioning	2.4%	Refrigeration service and repair
INTERNET	32.6%	Telephone communication, except radio	12.8%	Data processing and preparation	7.3%	Computer related services, nec
KARAOKE	69.0%	Entertainers and entertainment groups	5.7%	Radio, television, and electronic stores	3.1%	Eating places
LAMINATE	13.2%	Wood partitions and fixtures	6.9%	Special trade contractors, nec	6.4%	Lumber, plywood, and millwork
MACHINING	75.0%	Industrial machinery, nec	2.8%	Special dies, tools, jigs, and fixtures	1.7%	Repair services, nec
MAILBOX	58.6%	Business services, nec	8.3%	Direct mail advertising services	3.6%	Hardware stores
MICRO	10.6%	Computer and software stores	8.9%	Computers, peripherals, and software	6.3%	Computer related services, nec
MICROFILM	40.4%	Business services, nec	29.6%	Business Services, nec	5.5%	Office equipment
MICROWAVE	15.4%	Electronic components, nec	14.3%	Household appliance stores	13.1%	Electrical repair shops
MOTORCYCLE	46.8%	Motorcycle dealers	18.9%	Repair services, nec	9.5%	Automotive dealers, nec
MULTIMEDIA	8.7%	Data processing and preparation	8.4%	Custom computer programming services	7.7%	Motion picture and video production
ONLINE	19.1%	Telephone communication, except radio	8.8%	Data processing and preparation	6.2%	Computer related services, nec
ORTHODONTIC	54.0%	Offices and clinics of dentists	32.1%	Dental laboratories	7.3%	Dental equipment and supplies
PAGER	40.7%	Radiotelephone communication	25.9%	Miscellaneous retail stores, nec	15.9%	Electronic parts and equipment, nec
PAINTBALL	38.6%	Sporting goods and bicycle shops	25.3%	Amusement and recreation, nec	6.2%	Hobby, toy, and game shops
PHONE	25.6%	Miscellaneous retail stores, nec	9.2%	Business services, nec	8.9%	Telephone communication, except radio
PIPELINE	26.5%	Water, sewer, and utility lines	13.2%	Crude petroleum pipelines	12.7%	Natural gas transmission
PLYWOOD	27.5%	Lumber, plywood, and millwork	27.1%	Lumber and other building materials	7.0%	Hardwood veneer and plywood
POLYMER	13.4%	Plastics materials and resins	11.2%	Plastics materials and basic shapes	8.1%	Plastics products, nec
PROCESSOR	11.0%	Meat packing plants	6.3%	Scrap and waste materials	3.6%	Livestock services, except veterinary
PROPANE	75.3%	Liquefied petroleum gas dealers	6.9%	Petroleum products, nec	2.8%	Fuel oil dealers
PROSTHETIC	37.0%	Surgical appliances and supplies	30.8%	Miscellaneous retail stores, nec	18.6%	Dental laboratories
RADIO	35.7%	Radio and television repair	17.1%	Radio And Television Stores	17.1%	Radio broadcasting stations
SEMICONDUCTOR	37.6%	Semiconductors and related devices	31.3%	Electronic parts and equipment, nec	2.0%	Engineering services
SNOWMOBILE	62.9%	Automotive dealers, nec	8.1%	Repair services, nec	4.3%	Membership sports and recreation clubs
SOFTWARE	45.8%	Prepackaged software	23.2%	Custom computer programming services	8.9%	Computer and software stores
TELECOMMUNICATION	27.0%	Telephone communication, except radio	11.6%	Electrical work	9.8%	Business consulting, nec
TELEVISION	39.5%	Radio and television repair	18.6%	Radio And Television Stores	8.6%	Television broadcasting stations
TRACTOR	28.4%	Farm and garden machinery	10.0%	Repair services, nec	8.4%	Miscellaneous retail stores, nec
TYPEWRITER	47.9%	Miscellaneous retail stores, nec	27.8%	Repair services, nec	7.5%	Commercial Machines And Equipment
URETHANE	24.2%	Plastering, drywall, and insulation	20.3%	Roofing, siding, and sheetmetal work	6.0%	Special trade contractors, nec
VIDEO	43.7%	Video tape rental	9.7%	Motion picture and video production	8.1%	Radio, television, and electronic stores
VINYL	20.3%	Roofing, siding, and sheetmetal work	8.9%	Floor covering stores	7.8%	Top and body repair and paint shops
WINDSHIELD	56.2%	Automotive glass replacement shops	15.4%	Paint, glass, and wallpaper stores	11.0%	Automotive repair shops, nec
WIRELESS	37.3%	Radiotelephone communication	32.4%	Miscellaneous retail stores, nec	6.5%	Telephone communication, except radio
WIRELINE	86.3%	Oil and gas field services, nec	1.4%	Drilling oil and gas wells	1.2%	Water, sewer, and utility lines

Notes: Table lists technologies in our sample with $\geq$500 associated firms in our D&B data, and the 4-digit SIC industries most heavily associated with each. The close relationship between the technology and their firms' principal industries suggests that the presence of a technology in a firm's name generally suggests that the firm's principal line of business is related to that technology.

Figure 1: Predicted vs. actual distribution of firms across sectors in D&B testing sample



Notes: Figure shows the predicted and actual distribution of sectors in our D&B test sample, using the sector classification algorithm described in the paper.

Though these tests are validating, our measurement faces a few residual limitations. One is that for a small share of firms, we are unable to classify to sectors because the words in their names are either generic stop words or were not in our D&B training data. However, even under our most restrictive procedure, we classify 90% of firms, and this rate is stable over time. Some words included in our technologies list can also be contested, particularly non-nouns, and in robustness checks, we will further limit our analytical sample to only nouns.

3 Top Technologies and Illustrative Examples

Table 2 provides a first look at the technologies in our data by listing, for each decade from the 1900s to the 1990s, the technologies associated with the largest number of new firms in our registration data. The technologies that appear—such as wireless communications, automobiles, video, semiconductors, software, and the internet—correspond to well-known technological advances of the twentieth century.[12] For each technology, the table reports the number of associated firms and patents, the share of firms we classify to manufacturing versus other sectors, three representative firm names, and the most common co-occurring words in firm names.

Two features of the table are immediately apparent. First, these technologies are associated with substantial entrepreneurial entry. Second, for most technologies, relatively little of this entry occurs among firms engaged in technology production. With a few notable exceptions—such as semiconductors—associated firm creation spans a broad set of predominantly non-manufacturing activities, including distribution, installation, repair, and service.

[12] Computers are notably absent from this table because the term emerges in the patent data in the late nineteenth century (in reference to mechanical calculating machines), but they are included in our broader dataset.

Table 2: Top technology from each decade of the 20th century, and associated characteristics

		Total (to 2009)				Most common	
Decade	Top technology	Firms	Patents	Mfg. share	Other share	paired words	Example firms
1900	WIRELESS	28732	10565	6%	94%	COMMUNICATIONS SERVICES SOLUTIONS	VANGUARD WIRELESS, INC CYVO WIRELESS LLC WIRELESS BUILDERS, LLC
1910	AUTOMOTIVE	70954	2438	5%	95%	SERVICE SERVICES REPAIR	BNA AUTOMOTIVE GROUP, INC. GALYEON AUTOMOTIVE, LLC W.A.M. AUTOMOTIVE, INC.
1920	TELEVISION	12642	7598	4%	96%	CABLE SERVICE RADIO	GABLES TELEVISION, INC. BUSINESS TELEVISION, INC. U.S. PANTS & TELEVISION, INC.
1930	VIDEO	52958	26315	4%	96%	PRODUCTIONS AUDIO SERVICES	LEGAL VIDEO SERVICES, INC. COBB VIDEO, INC. VIDEO EMPORIUM, INC.
1940	SEMICONDUCTOR	1977	88544	48%	52%	EQUIPMENT TECHNOLOGY TECHNOLOGIES	PHILIPS SEMICONDUCTORS INC. SAMSUNG SEMICONDUCTOR, INC. IDEAL SEMICONDUCTOR, INC.
1950	BROADBAND	3643	909	23%	77%	SERVICES COMMUNICATIONS WIRELESS	N2 BROADBAND, INC. ZTARK BROADBAND CORPORATION BROADBAND PROPERTIES, LLC
1960	PAGER	574	565	8%	92%	CELLULAR PLUS REPAIR	PAGERS TO GO, INC. PAGERS PLUS INC. LET'S TALK PAGERS, INC.
1970	SOFTWARE	63181	6198	5%	95%	SOLUTIONS SYSTEMS SERVICES	MINI B SOFTWARE, INC. DENIZEN SOFTWARE, LLC BEST SOFTWARE OF CALIFORNIA
1980	ROBOTIC	1632	756	39%	61%	SYSTEMS AUTOMATION TECHNOLOGY	MAZOR ROBOTICS INC. INTERNATIONAL ROBOTICS INC. FUTURE ROBOTICS, INC.
1990	INTERNET	17268	2387	4%	96%	SERVICES SOLUTIONS MARKETING	DC INTERNET GROUP, LLC SAEROM INTERNET WORLD II INC. WORLDPORT INTERNET, INC.

Notes: Table lists the technology from each decade of the 20th century with the most firms created, alongside (i) the number of associated new firms and patents, (iii) the sectoral composition of those firms, (iv) the three most common paired words, and (v) three example firms.

Table 3 expands on the select set of examples in Table 2 with summary statistics for the full technology sample. The table shows large variation across technologies in the number of associated firms, which ranges from zero to nearly 100,000 (with a mean of ∼1,000 and large skew). We generally see a low share of firms registered in Delaware (a signal of growth intention) or which we classify to manufacturing: at the median, only 2% and 13%, respectively (average 9% and 17%).

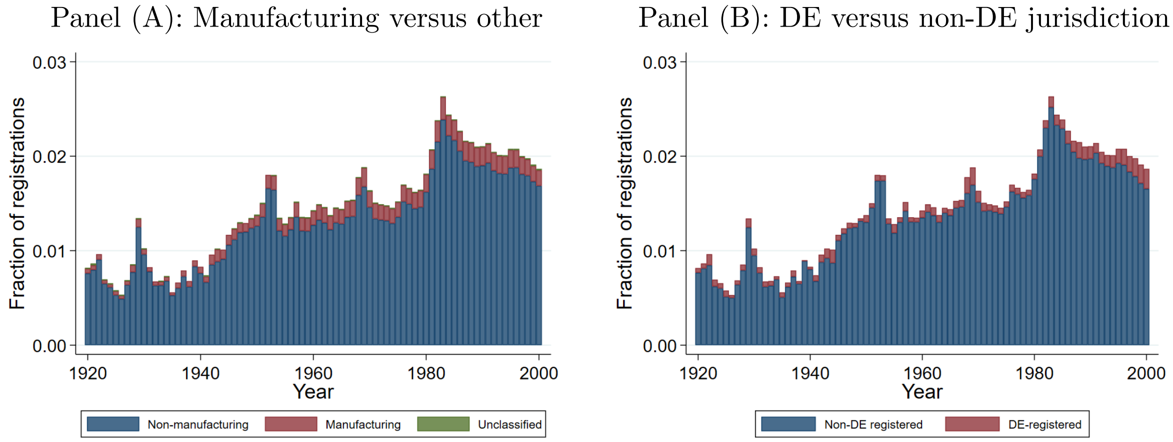
Table 3: Summary statistics at the technology level

Variable	N	Mean	P10	P25	P50	P75	P90	Max
Patents	860	2919	151	281	909	1653	4915	161347
Firms	658	1040	2	3	19	64	673	98530
Share Delaware jurisdiction	658	0.09	0.00	0.00	0.02	0.08	0.25	1.00
Share non-Delaware jurisdiction	658	0.91	0.75	0.87	0.98	1.00	1.00	1.00
Share manufacturing	625	0.17	0.04	0.06	0.13	0.19	0.33	0.91
Share non-manufacturing	625	0.83	0.67	0.76	0.87	0.92	0.96	1.00

Notes: Table shows summary statistics for technologies in our sample (each measured through 2009). The sample size in each row reflects the total number of technologies identified from patent data (N=860), the number we associate to at least one firm (N=658), and the number we associate to at least one firm that we can classify to an economic sector (N=625).

Beyond variation across technologies, we are also interested in understanding how large technology-related entry is relative to overall entrepreneurial entry in the economy. In short, we find that these firms represent a substantial share of all new U.S. firms. Figure 2 shows the share of all U.S. business registrations each year associated with the technologies in our sample by year, separately reporting firms we classify into manufacturing versus other sectors (left panel) and firms registered in Delaware versus other jurisdictions (right panel). Between 1950 and 2000, roughly 1-2% of new firms each year explicitly include one of the technologies in our sample in their name—a total which intrinsically *under*counts the target population of industry participants due to both our focused technology sample ($N = 860$), our measurement of entry via firm creation (which excludes existing firms), and our reliance on firm names (which only partially measure the phenomenon). Nearly 85% of these firms are local non-manufacturing firms. Manual inspection of the sample reinforces that rather than being fast-moving, high-growth innovators, most firms linked to new technologies have founding characteristics consistent with small, local enterprises that create and capture value by providing ancillary goods and services—such as local equipment sales, rental, installation, replacement or maintenance, and advisory services.

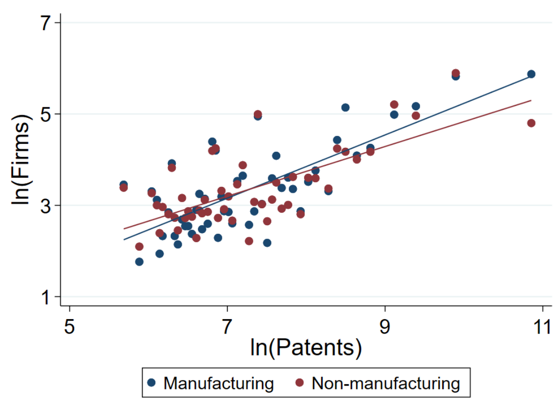
Figure 2: Focal technology-related firm share of U.S. business registrations, 1900-2000



Notes: Figure plots the share of U.S. business registrations over time related to the technologies in our sample, as measured via firm names. See text for further discussion.

More heavily-developed technologies also have more associated new firm registrations. Figure 3 presents a binned scatterplot of log firms against log patents for our sampled technologies, conditional on the decade of each technology's first patent. We find a large correlation between them: a doubling of patents is associated with 53% more entry around a given technology. This relationship is statistically and quantitatively very similar for both manufacturing and non-manufacturing firms, revealing an additional pattern: not only are more developed technologies associated with more entry, but they evidently are so for firms of every type.

Figure 3: Correlation of total entry and patenting in individual technologies



Notes: Figure shows a binned scatterplot of the number of patents (horizontal axis) and number of associated business registrations (vertical axis). Chart conditions on first patent decade.

Example technologies

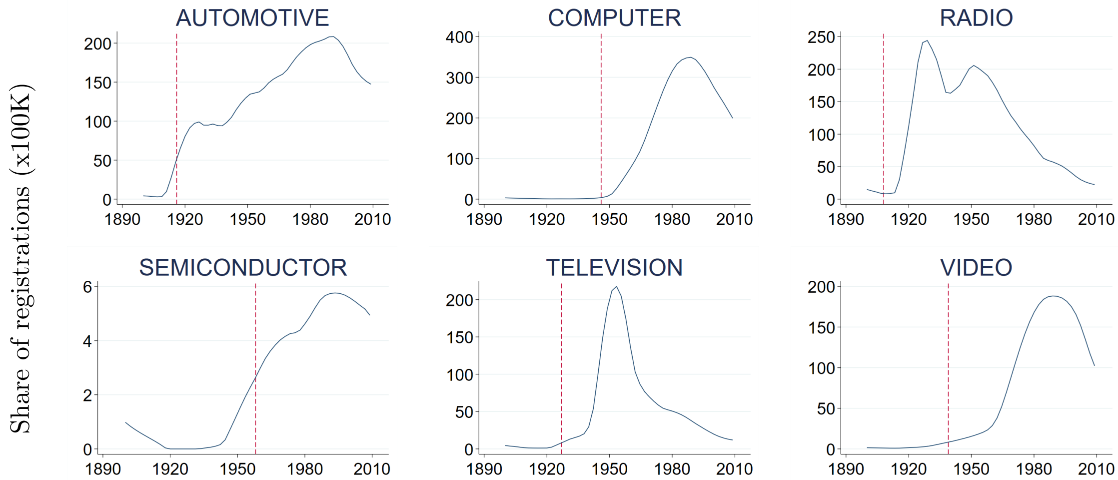
To begin to understand how lifecycle patterns are reflected in firm registration data, we chose ten example technological terms in our sample to illustrate specific dynamics and other features in the data. Our selection spans the twentieth century and covers a range of terms that vary in scale. Figure 4 charts the evolution of gross entry over time for these ten examples. For each one, we measure the annual share of all U.S. firm registrations that are associated to each technology (scaled by 100,000 for readability), and smooth sharp fluctuations for lower-count examples by calculating 5-year moving averages.[13] Vertical lines mark each technology's "emergence" year, which we define as the first year when it has at least 10 cumulative associated firms and enters a five-year period with more than 10% average annual growth in firm entry—in effect, when firm creation first begins to take off (akin to the approach taken by Kalyani et al. 2024).

In all cases, entry grows, peaks, and (in all but the most recent cases) declines, but there is also substantial heterogeneity. As Table 4 shows, some of these technologies at their peak attract an order of magnitude more entry than others. In some cases, entry is slow to build; in others, fast. Some technologies are multi-peaked, while others stimulate short but sharp bursts of firm creation. Similar heterogeneity is present in the raw data in previous cross-industry analyses of industry lifecycles (e.g., Gort and Klepper 1982 and its descendants), despite these papers' emphasis on averages and common patterns. This can be seen in Appendix Figure A.1, where we similarly plot the rate of gross entry over time for products in Klepper and Graddy (1990) which we could reliably search for in our firm registration data (some technological, some not): despite differences in both what is measured (gross vs. net entry) and sampling methods (firm registrations vs. firm directories), we find similar lifecycle patterns in these examples, whose timing aligns with Klepper and Graddy's stages of industry evolution. Importantly, in both this prior tradition and the present paper, these patterns represent overall industry dynamics and mask the heterogeneity that may exist across industry segments, geography, or other dimensions.

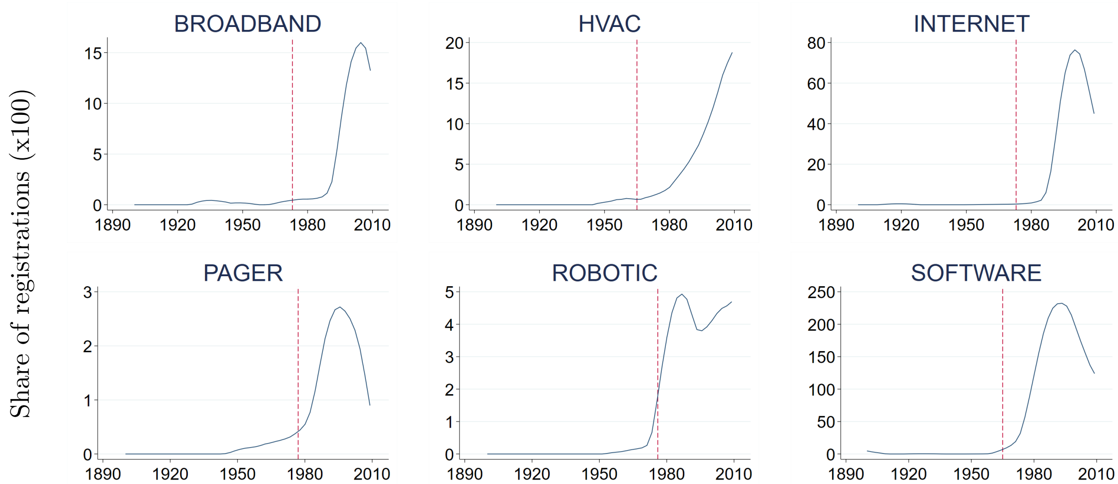
[13] We plot associated entry as a share of all firm registrations rather than in counts, which are strictly increasing over time—driven more by the growth of the U.S. population and economy than by industry dynamics.

Figure 4: Diffusion of example technologies into new firm names

Panel (A): Early technologies



Panel (B): More recent technologies



Notes: Figure illustrates entrepreneurial lifecycles for individual technologies, as measured via firm names, for technologies developed before 1950 (Panel A) and after (Panel B). Vertical axis measures the technology's share of business registrations in a given year. The dotted line in each subfigure marks the technology's emergence year, defined as the earliest year after which the technology had at least 10 cumulative firm registrations and ≥10% average growth in registrations over the next five years. Lifecycles plotted using local polynomial smoothing. See text for further details.

Table 4: Example technologies' maximum
share of business registrations (x1000)

Panel A. Early technologies (pre-1950)		Panel B. More recent technologies	
Variable	Max share	Variable	Max share
AUTOMOTIVE	2.3	BROADBAND	0.3
COMPUTER	5.8	HVAC	0.2
RADIO	4.6	INTERNET	1.7
SEMICONDUCTOR	0.1	PAGER	0.1
TELEVISION	4.7	ROBOTIC	0.1
VIDEO	5.1	SOFTWARE	2.8

Notes: Table lists ten example technologies and their peak share of new business registrations in our data.

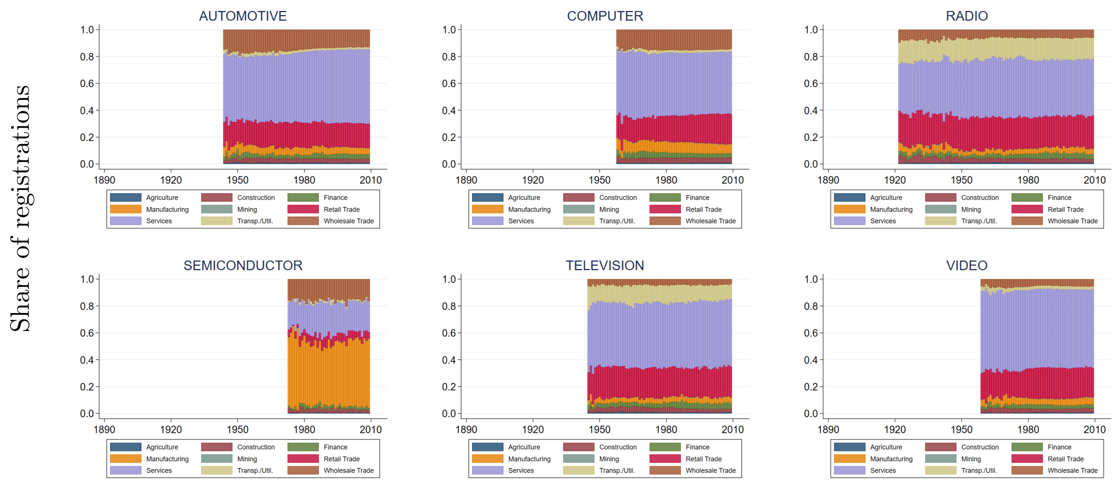
Although these lifecycle patterns resemble those documented in prior work, they describe aggregate entry and do not reveal how the composition of entrants changes as technologies mature. The fundamental departure we make in this paper is to examine the composition and dynamics of technology-enabled entrepreneurship *within* the industry lifecycle. In practice, doing so requires disaggregating overall patterns into meaningful subcategories. Throughout the remaining analysis, we will attempt to partition firms into industry segments, typically in two ways: by (i) associating firms to specific value-adding activities, based on the presence of specific words in their name, and (ii) associating firms to SIC sectors using our predictive machinery.

To illustrate the substantial heterogeneity that exists within our sample, we first focus on sectoral composition—motivated by the fact that our procedure can associate nearly every firm to these sectors and that it supports fractional associations (reflecting that in practice, many firms span boundaries). Figure 5 shows the sectoral composition of new firms in our ten example technologies.[14] Though a few technologies have a large share of entry in manufacturing (e.g., semiconductors, and to a lesser degree electronics and aircraft), others concentrate in construction (HVAC), transportation and utilities (aircraft, radio, television), or services (automotive, software, internet). The figure provides initial evidence (from a small set of examples) suggesting that these ancillary firms which create value around new technologies but do not specifically produce it have been the "dark matter" of the empirical industry evolution literature, responsible for a large share of economic activity but not previously systematically studied. Beyond providing distinct value to a different set of customers through different activities, these firms may have systematically different industry dynamics than aggregate trends would suggest or than technology-producing firms which have been the principal focus of well-known empirical studies.

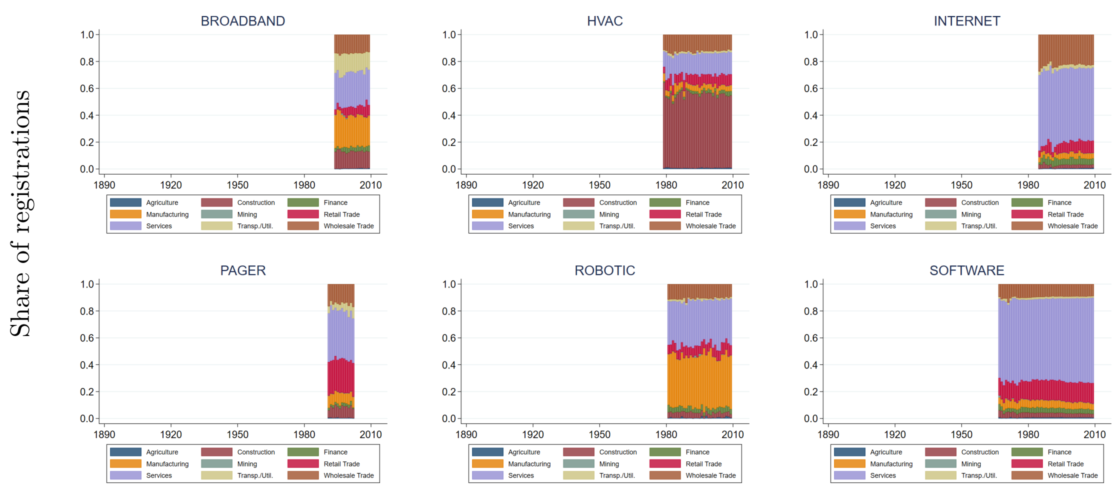
[14]In preparing these figures, we present technologies' sector shares for contiguous years with at least 10 firms, to reduce noise and aid visualization. Note that this chart is most easily read in full color.

Figure 5: Sector composition of new firms created around example technologies

Panel (A): Early technologies



Panel (B): More recent technologies



Notes: Figure shows sectoral composition of new business registrations associated with specific technologies, over time, for technologies developed before 1950 (Panel A) and after (Panel B). Each subfigure restricts to technology-years with at least 10 associated registrations, and for presentation purposes restricts to years for which a continuous time series can be constructed without gaps (i.e., the series begins in the year after which the technology always has at least 10 associated firms).

4 Industry Evolution Across the Value Chain

Our primary goal in this paper is to systematically evaluate, across a large sample, how firm creation and industrial activity change over technology lifecycles, looking within industries to examine how the set of participants evolves as new technologies mature.

To do so, we return to our full technology sample. Our analysis progresses in three steps. First, we generate summary measures of entry dispersion across activities and sectors at the technology-year level and evaluate how they change over the lifecycle. We then specifically evaluate changes over time in the relative rate of entry in these activities and sectors, testing the degree to which entry shifts down the value chain. We then synthesize this analysis into a three-phase framework for disaggregating industry lifecycles into component, overlapping evolutionary paths of narrower segments operating within them, discussing how these segments interact and how these interactions are likely to shape their scale and timing. We conclude this analysis by presenting stylized facts on these three industry phases, based on averages in the data.

4.1 The Dispersion of Entrepreneurial Entry

Our starting point is to examine the dispersion of entrepreneurial entry. Because there are many ways to measure dispersion, we take several approaches. One is to measure the top activity's (or sector's) share of firm creation. Another is to construct a Herfindahl index across activities (sectors). A third is to measure the coefficient of variation (C.V.), dividing the standard deviation of activities' (sectors') share of firm creation in a given technology by the mean (as Kalyani et al. 2024 do in studying the geographic spread of jobs related to new technologies). All three are concentration measures that, when they decline, will signal entrepreneurial entry growing more heterogeneous and diffuse, and (for the coefficient of variation) more balanced across sectors.

In Table 5 we estimate a regression of each measure, calculated at the technology-year level, on indicators for each decade of a technology's lifecycle, controlling for technology and calendar year fixed effects. Concretely, we estimate the following equation:

$$Y_{it} = \sum_{s} \beta_s \cdot \mathbb{1}(\text{Technology is } s \text{ years into lifecycle}) + \alpha_i + \delta_t + \varepsilon_{it} \qquad (1)$$

where i and t index technologies and calendar years, α_i and δ_t are fixed effects, and β_s are the parameters of interest. Columns (1) to (3) of the table examine activity-based concentration measures, and Columns (4) to (6) complementary sector-based measures.

Table 5 shows that all of these measures decline monotonically across technologies' lifecycles. Despite most technologies in our sample intrinsically being more or less closely associable to specific activities and sectors (as in Figure 5), the top activity's share of entry declines roughly 20 percentage points over the first 60 years of the lifecycle (Column 1). The Herfindahl index declines nearly 25 points (out of 100; Column 2), and the coefficient of variation declines 45 points (out of a max of

2.2; Column 3).[15] The estimated decline in concentration is more muted for sector-based measures (Columns 4 to 6), which wash out more of the variation, and stop growing after roughly 40 years (rather than 60), but they remain statistically significant. Read together, these results point to a substantial dispersing of entry as technology-based industries age.

Table 5: Dispersion of entrepreneurial entry across the technology lifecycle

	Activity			Sector		
	(1)	(2)	(3)	(4)	(5)	(6)
	Top 1	HHI	C.V.	Top 1	HHI	C.V.
11-20 years into lifecycle	-0.058	-0.059	-0.112	-0.010	-0.012	-0.027
	(0.014)	(0.016)	(0.030)	(0.006)	(0.005)	(0.013)
21-30 years into lifecycle	-0.075	-0.072	-0.143	-0.017	-0.020	-0.047
	(0.019)	(0.023)	(0.043)	(0.006)	(0.006)	(0.015)
31-40 years into lifecycle	-0.108	-0.110	-0.213	-0.025	-0.030	-0.071
	(0.025)	(0.031)	(0.058)	(0.008)	(0.007)	(0.018)
41-50 years into lifecycle	-0.135	-0.140	-0.273	-0.025	-0.030	-0.072
	(0.032)	(0.039)	(0.073)	(0.009)	(0.008)	(0.022)
51-60 years into lifecycle	-0.174	-0.183	-0.355	-0.026	-0.029	-0.076
	(0.038)	(0.046)	(0.087)	(0.011)	(0.009)	(0.025)
61+ years into lifecycle	-0.215	-0.242	-0.459	-0.025	-0.029	-0.077
	(0.048)	(0.058)	(0.111)	(0.013)	(0.010)	(0.030)
N	4348	4348	4348	12308	12308	12308
R^2	0.32	0.39	0.39	0.52	0.46	0.52
Token FEs	Y	Y	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y	Y	Y
Y mean	0.88	0.84	1.96	0.50	0.38	0.77

Notes: Table estimates changes in the dispersion of firm creation across economic sectors or the value chain over time. Dispersion measures are (i) the top activity/sector share ("Top 1"; Columns 1 and 4), the activity/sector HHI ("HHI"; Columns 2 and 5), and the coefficient of variation ("C.V."—in essence, the standard deviation of activity/sector shares; Columns 3 and 6). SEs clustered by technology in parentheses.

4.2 Migration of Entrepreneurship Down the Value Chain

As entry disperses, how does the positioning of these firms change? Table 6 uses co-occurring words and our sector and activity classifications to evaluate how economic activities undertaken by technology-related startups vary over the lifecycle. Panel (A) reports the relative frequency of the manufacturing, services, and 'other' sectors among the first 100 firms in a given technology, the 101-200th, the 201-500th, and so on. Panel (B) does so for specific value chain activities, and Panel (C) for specific co-occurring words in firm names, which we (inductively) partition into three sets: words that are relatively common early in the technology lifecycle but decrease across it, those that increase, and those that trace an inverted-U growth pattern.

[15] Because our measures span five activities, the maximum potential coefficient of variation is 2.24 (when all entry is concentrated in one activity). The minimum is 0 (when entry is equally dispersed).

Table 6: Changing nature of entrepreneurial entry across the technology lifecycle

Panel A. Economic sectors

	Number of firms into lifecycle					
	(1)	(2)	(3)	(4)	(5)	(6)
Sector	1-100	101-200	201-500	501-1000	1001-2000	2001+
Manufacturing	1.00	0.94	0.83	0.75	0.68	0.39
Sales/Distribution	1.00	1.02	1.07	1.09	1.09	1.22
Services	1.00	1.00	1.01	1.02	1.05	1.25
All other	1.00	1.02	1.06	1.09	1.10	0.84

Panel B. Value chain activities

	Number of firms into lifecycle					
	(1)	(2)	(3)	(4)	(5)	(6)
Activity	1-100	101-200	201-500	501-1000	1001-2000	2001+
Research	1.00	0.64	0.37	0.42	0.36	0.19
Manufacturing	1.00	0.92	0.73	0.73	0.51	0.27
Sales	1.00	1.14	1.39	1.88	1.68	1.27
Distribution	1.00	1.22	1.09	1.65	1.35	1.72
Service	1.00	1.36	1.69	1.80	1.68	1.99

Panel C. Specific paired words

	Number of firms into lifecycle					
	(1)	(2)	(3)	(4)	(5)	(6)
Paired word	1-100	101-200	201-500	501-1000	1001-2000	2001+
Industry	1.00	0.80	0.50	0.53	0.46	0.27
Manufacturing	1.00	0.96	0.73	0.72	0.53	0.27
Engineering	1.00	0.96	0.68	0.53	0.58	0.47
Research	1.00	0.63	0.37	0.42	0.36	0.19
Development	1.00	0.58	0.55	0.63	0.57	0.68
Product	1.00	0.65	0.64	0.50	0.52	0.36
Maintenance	1.00	0.90	1.21	0.94	1.11	1.67
Part	1.00	2.55	3.58	5.03	8.07	6.13
Service	1.00	1.30	1.57	1.59	1.55	1.82
Equipment	1.00	1.11	1.69	1.62	1.24	0.80
Supply	1.00	1.07	1.29	1.73	1.64	1.01

Notes: Table evaluates changes in the relative frequency of firm creation by economic sector (Panel A), activity (Panel B), and specific paired words in firm names (Panel C). To produce this table, we identify firms in the first 1-100, 101-200, etc. firms of a given technology's lifecycle. We then aggregate these to measure the share of all "first 100" firms which are associated with each sector or activity, as measured via firm names. To ease interpretation we then index these shares to that of the first 100 firms, which is therefore always equal to one and provides a reference point to compare against as firm creation changes over time.

We observe three patterns. First, entry related to the production and development of a technology tends to peak early in its lifecycle. This is apparent for manufacturing sector firms (Panel A), firms we associate to research and production activities (Panel B), and paired words like "manufacturing" or "engineering" (Panel C). Second, as technologies mature, downstream entry appears to grow: Table 6 shows an increasing relative frequency for service sector firms (Panel A), firms we associate to service activities (Panel B), and for terms like "maintenance", "parts", and "service" (Panel C). Third, depending on the measures, firms related to sales or distribution sometimes follow an inverted-U pattern (as seen in Panel B, particularly for sales, and in Panel C, for terms like "equipment" and "supply"). The evidence suggests that as technology matures, entry shifts from production to servicing, which adds value later in the technology lifecycle by extending the life of technologies that have already diffused. In between, entry also gravitates towards complementary activities, including in supporting adoption, like sales and distribution. Whereas prior work has emphasized a decline in entry as technological industries mature and consolidate (Klepper and Graddy 1990), we instead find a downstream *migration* of entry and industry participation—following apparent downstream shifts in the commercial opportunity set.

We econometrically evaluate these transitional dynamics in Table 7, which estimates changes in the share of firms associated to manufacturing versus services, at the technology-year level, on indicators for each decade of the technology lifecycle (Equation 1). Columns (1) and (2) of Table 7 examine activity shares, and Columns (3) and (4) sector shares. The results demonstrate a similar, monotonic shift as in Table 6, from production-oriented firms to service-oriented firms, with the decline in the former mostly offset by growth in the latter.

Table 7: Upstream to downstream industry evolution

	Activity		Sector	
	(1)	(2)	(3)	(4)
	Manuf.	Services	Manuf.	Services
11-20 years into lifecycle	-0.026	0.043	-0.010	0.009
	(0.032)	(0.044)	(0.005)	(0.005)
21-30 years into lifecycle	-0.083	0.122	-0.013	0.024
	(0.035)	(0.047)	(0.005)	(0.006)
31-40 years into lifecycle	-0.094	0.112	-0.022	0.027
	(0.035)	(0.052)	(0.006)	(0.007)
41-50 years into lifecycle	-0.122	0.166	-0.028	0.035
	(0.035)	(0.057)	(0.007)	(0.008)
51-60 years into lifecycle	-0.137	0.149	-0.028	0.038
	(0.039)	(0.064)	(0.009)	(0.010)
61+ years into lifecycle	-0.154	0.166	-0.034	0.048
	(0.041)	(0.069)	(0.010)	(0.012)
N	4348	4348	12308	12308
R^2	0.33	0.44	0.76	0.69
Token FEs	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y
Y mean	0.07	0.69	0.19	0.35

Notes: Table estimates changes in the share of firm creation across value chain activities and economic sectors, comparing manufacturing and services. SEs clustered by technology in parentheses.

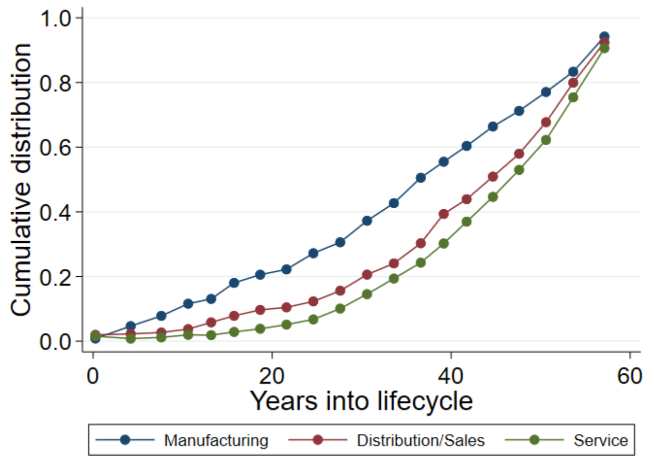
4.3 Industry Lifecycles as a Mixture of Sub-Industry Dynamics

Tables 6 and 7 show that entrepreneurial entry systematically migrates downstream as technologies mature, with production-oriented firms peaking early, followed by entry into distribution and, later, service. Together, these patterns suggest a reinterpretation of the canonical bell-shaped industry lifecycle found in prior work (and seen in, e.g., Figure 4). Rather than reflecting a single, uniform process, aggregate industry lifecycles can be understood as the composite of multiple, overlapping trajectories corresponding to distinct segments of the value chain. To illustrate, consider an industry comprised of these three segments (manufacturing, distribution, service). Suppose each exhibits its own bell-shaped entry dynamics, but with distinct timing and scale. Aggregating across them would then be liable to yield the familiar unimodal lifecycle pattern but obscure substantial heterogeneity in entrepreneurial opportunity across these segments.

Figure 6 provides direct evidence for this interpretation. The figure plots the cross-technology average cumulative distribution of firm creation over the first 60 years of the technology lifecycle, separately for manufacturing, sales and distribution, and service. The ordering of the curves exhibits clear stochastic dominance: manufacturing-related entry grows first on average, followed by distribution, and then service. This ordering is difficult to reconcile with a single-process view of industry evolution, but it follows naturally from a value-chain perspective in which different firm activities become commercially valuable at different points in time.

Taken as a whole, the evidence suggests that industry lifecycles are better described as mixtures of segment-specific dynamics rather than as a single, aggregate process. To be clear, this interpretation does not imply a new theory of industry evolution, nor does it require strong assumptions about underlying mechanisms; it instead provides a descriptive framework for organizing observed patterns of entrepreneurial entry and clarifying how aggregate lifecycle regularities can coexist with the substantial heterogeneity across the value chain we find.

Figure 6: Cumulative distribution of firm creation, by activity,
over the first 60 years of technology lifecycle



Notes: Figure shows the average cumulative distribution of firm creation in manufacturing, distribution and sales, and service (blue, red, and green, respectively) over the first 60 years of a technology's lifecycle. Firm activities identified using words in firm name. Sample restricted to technologies with at least 10 associated firms. Consistent with the theoretical arguments of (and Figure 1 from) Agarwal and Gort (1996), the figure suggests that gross entry evolves from upstream to downstream activities over the technology lifecycle: manufacturing stochastically dominates distribution, which in turn stochastically dominates service.

5 Discussion

This paper documents across 860 technologies and the near-universe of U.S. firm registrations that entrepreneurial entry around new technologies is both substantial in scale and structured in its evolution. We find that entrepreneurial activity extends well beyond the firms that produce a given technology, and that over the technology lifecycle, entry systematically migrates from upstream activities associated with manufacturing toward downstream activities such as distribution, parts, and service. Though these dynamics are intuitive in well-known cases—such as the dense networks of suppliers, dealers, and service providers surrounding mature automobile manufacturers—our contribution is to characterize these patterns systematically, across many technologies and over the lifecycle. In doing so, we connect and extend insights from detailed single-industry studies with a broader, more general account of entrepreneurial entry over time.

A first implication of our findings is to confirm that a broad constellation of entrepreneurial entrants around new technologies is not exceptional but rather a general feature of technology-driven industry formation. In doing so, this evidence links to and integrates with existing theory and case studies. For example, Moeen et al. (2020) and Agarwal et al. (2025) provide frameworks that emphasize how nascent industries emerge through entrepreneurial activities undertaken by a range of heterogeneous actors. Much of the work these papers build on showcase through in-depth studies a "hidden" web of entrepreneurial activities that extends beyond producers. This research—with work on industries as diverse as semiconductors (Adner and Kapoor 2010), agricultural biotechnology (Moeen and Agarwal 2017), mobile money (Wormald et al. 2021, 2023), solar energy (Kapoor and Furr 2015, Szerb and Furr 2025, Guerra and Agarwal 2021), and bionic prosthetics (Kim et al. 2025)—has documented the complex interplay between firms that invent and commercialize technologies and other firms engaged in complementary activities. Consistent with our findings, several of these studies have shown that early in the lifecycle, firms focus more on the creation of technological systems through integration of components, but later in the lifecycle, on creating complementary products and services and disintegration. Our evidence demonstrates that these insights are not confined to a small number of intensively studied industries. Across a large and diverse set of technologies, we find that the majority of entrepreneurial entry occurs outside manufacturing, and that firms occupying complementary positions along the value chain constitute a large share of the entrepreneurial economy associated with new technologies.

This contribution speaks directly to the industry evolution literature and its measurement traditions. Classic empirical studies relied on sources such as the Thomas Register of American Manufacturers, focusing on producers of a technology and largely abstracting away from ancillary and complementary activity (Klepper and Graddy 1990, Klepper 1996). Even later work documenting gross entry, while more closely aligned with our focus on business creation, did not differentiate systematically between producers, distributors, and service providers, nor did it explicitly consider the broader ecosystem of technology-enabled firms (e.g., Agarwal and Gort 1996). By linking firm registrations to technologies and classifying firms by sector and activity, we are able to examine gross entrepreneurial entry across the value chain rather than within a single segment. This shift in perspective reveals that the seeming decline in entrepreneurial opportunities in producer-focused data may instead reflect a reallocation of entry toward downstream activities.

Another implication concerns the temporal scope of entrepreneurial opportunity. Our results are consistent with arguments by Moeen et al. (2020) and Agarwal et al. (2025) that early uncertainties in emerging industries get resolved through the participation of different actors, but they also suggest an important extension. We find that entrepreneurial entry around new technologies continues even after the initial uncertainties surrounding an industry are likely resolved and producers consolidated. Instead, entry persists well into later stages of the lifecycle, with the composition of new firms shifting predictably toward activities that support diffusion, use, maintenance, and the extension of the technology's productive life. In this sense, the entrepreneurial ecosystem surrounding a technology continues to evolve even as the production segment matures. This downstream migration of entry is not fully captured by existing frameworks that focus primarily on incubation through early growth, and it suggests that the role of complementary firms in industry evolution extends well beyond the formative stages of industry development.

Taken together, these findings point toward a view of industry lifecycles as composites of overlapping sub-dynamics across the value chain, which are seen directly in Figure 6. Aggregate patterns of entry and decline documented in prior work masks critical heterogeneity in the timing and nature of entrepreneurial opportunity across these segments. Early in the lifecycle, entrepreneurial activity is concentrated in activities related to developing and producing the technology itself. As technologies mature, opportunities increasingly arise in complementary activities that facilitate adoption, distribution, and ongoing use. In contrast to innovative firms that pioneer new technologies, these latter firms enter later in the technology lifecycle, often in supporting roles, and at an arms-length distance from other players. From this perspective, industry evolution is not characterized solely by the rise and fall of producers, but by a shifting locus of entrepreneurial entry across activities that collectively support a new technology's economic impact.

5.1 Implications for Entrepreneurial Strategy Frameworks

Our findings also relate to how entrepreneurial opportunity is conceptualized in the strategy literature. Much existing work in this area—particularly in technology-intensive settings—frames opportunity identification as a problem of choosing among emerging technologies (Gans et al. 2021). In this view, entrepreneurs assess the technical and commercial prospects of new technologies and decide whether and how to build ventures around them. This framing fits with innovation-driven startups that seek to develop or commercialize new technologies, and it offers important insights on entry, competition, and value capture in high-tech markets.

At the same time, our evidence suggests that a technology-centric framing captures only part of the opportunity landscape associated with technological change. Across a wide range of technologies and over much of the lifecycle, we observe substantial entrepreneurial entry by firms that do not produce the focal technology, but instead occupy complementary positions along the value chain. For these firms, the relevant strategic choice may be less about which technology to pursue, and more about which activity to undertake—such as distribution, installation, maintenance, or service—and when to enter as technologies mature, standardize, and diffuse.

This reframing does not imply that activity-first strategies dominate technology-first strategies, nor does it speak to the relative performance of firms pursuing different approaches—questions that are beyond the scope of our data. Rather, it highlights that entrepreneurial opportunities associated with new technologies are more heterogeneous than is typically emphasized in strategy research. Focusing exclusively on technology-producing startups risks overlooking a large and patterned set of entrepreneurial responses to technological change.

5.2 Limitations and Boundary Conditions

As with any empirical accounting of this scope, our work is subject to limitations. We emphasize three. First, our analysis is of entrepreneurial entry, rather than performance outcomes such as profitability, survival, or growth. This focus is both feature and bug: it allows us to characterize the scale, composition, and timing of entrepreneurial responses to new technologies over long horizons, but it also restricts our ability to speak to the competitive success of these firms, or to measure net industry growth or consolidation. As a result, our findings should be interpreted as documenting the set of opportunities associated with new technologies that new firms pursue—which ostensibly reflect their expected returns—rather than realized returns.

Second, our results specifically describe patterns of entrepreneurial entry around influential, patented technologies, which is a particular category of innovation: by construction, our technology sample is drawn from the patent record and conditioned on sustained inventive activity. This allows us to study technologies that we know with hindsight were economically important, but it also limits our ability to (i) characterize entrepreneurial responses to short-lived, unsuccessful, or non-patented innovations or (ii) predict responses to emerging technologies today. The dynamics we document may differ in settings where technologies diffuse rapidly without formal intellectual property, or for innovations which are not reflected in patents or firm names.

Third, although we document systematic shifts in the composition of entry over the technology lifecycle, we remain agnostic about the mechanisms that generate these patterns. The downstream migration of entry is consistent with increasing gains from specialization that drive industry disintegration (e.g., Stigler 1951, Jacobides 2005), the emergence of standardized interfaces which support decentralization (e.g., Sanchez and Mahoney 1996, Baldwin and Clark 2000), and other theoretical mechanisms, but our data are missing critical detail that would be needed to adjudicate among these explanations. For example, we do not observe the organizational boundaries or contractual arrangements that link technology-producing firms to the broader set of entrants we study. In our view, understanding these relationships, and how they vary across technologies and institutional contexts, remains an important direction for future research.

Taken together, these limitations suggest this paper is best viewed as establishing a set of empirical regularities about entrepreneurial entry around new technologies, rather than as a complete account of industry evolution or entrepreneurial performance. By documenting the scale, timing, and heterogeneity of entry across the value chain, we aim to lay a foundation for additional work that will more directly examine firm relationships and outcomes within the broad entrepreneurial ecosystems that evidently accompany technological change.

5.3 Questions for Future Research

Our findings also point to other questions. What explains the heterogeneity across technologies in the scale and composition of entrepreneurial entry? Even among influential innovations, we see substantial variation in the sectors and activities attracting new firms. Understanding whether this heterogeneity is driven primarily by technology characteristics, institutional context, complementarities with existing industries, or historical contingencies might clarify when and where different forms of technology-enabled entrepreneurship are likely to emerge.

A second set of questions relates to entrants' strategic positioning and interactions. Although we find that entrepreneurial entry increasingly shifts toward downstream activities as technologies mature, we do not observe how these firms establish positions, compete, or coordinate with technology producers or one another. Future work might explore how producers' boundary choices, contracting practices, intellectual property (IP) management, or ecosystem strategies shape the organization of complementary activity, and conversely how non-producer firms navigate markets where core IP is controlled by other parties, entry barriers are low, and scale is limited. Addressing these questions would deepen our understanding of the evolution of technology-based industries, and of strategy for a fuller range of technology-enabled ventures.

References

Abernathy, William J and James M Utterback. 1978. “Patterns of industrial innovation,” *Technology Review*, Vol. 80, No. 7, pp. 40–47.

Adner, Ron and Rahul Kapoor. 2010. “Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations,” *Strategic Management Journal*, Vol. 31, No. 3, pp. 306–333.

Agarwal, Rajshree and Michael Gort. 1996. “The evolution of markets and entry, exit and survival of firms,” *The Review of Economics and Statistics*, pp. 489–498.

Agarwal, Rajshree, Mara Guerra, Mahka Moeen, and Paolo Aversa. 2025. “Creating new industries: The generative role of heterogeneous actors,” *Academy of Management Annals*, Vol. 19, No. 2, p. 476–521.

Agarwal, Rajshree and Mary Tripsas. 2008. “Technology and industry evolution,” *The handbook of technology and innovation management*, Vol. 1, pp. 1–55.

Andrews, Raymond J, Catherine Fazio, Jorge Guzman, Yupeng Liu, and Scott Stern. 2022. “The Startup Cartography Project: Measuring and mapping entrepreneurial ecosystems,” *Research Policy*, Vol. 51, No. 2, p. 104437.

Arts, Sam, Jianan Hou, and Juan Carlos Gomez. 2021. “Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures,” *Research Policy*, Vol. 50, No. 2, pp. 104–144.

Baldwin, Carliss Y and Kim B Clark. 2000. *Design rules: The power of modularity*. Cambridge: MIT Press.

Belenzon, Sharon, Aaron K Chatterji, and Brendan Daley. 2017. “Eponymous entrepreneurs,” *American Economic Review*, Vol. 107, No. 6, pp. 1638–55.

——— . 2020. “Choosing between growth and glory,” *Management Science*, Vol. 66, No. 5, pp. 2050–2074.

Cramer, Jonathan James. 2017. “Dun & Bradstreet enlists the help of AI to solve a decades-old problem,” Available at https://www.dnb.com/perspectives/master-data/how-dnb-improves-sic-classification-ai.html.

Gans, Joshua S, Michael Kearney, Erin L Scott, and Scott Stern. 2021. “Choosing technology: An entrepreneurial strategy approach,” *Strategy Science*, Vol. 6, No. 1, pp. 39–53.

Gort, Michael and Steven Klepper. 1982. “Time paths in the diffusion of product innovations,” *The Economic Journal*, Vol. 92, No. 367, pp. 630–653.

Gross, Daniel P. 2023. “The hidden costs of securing innovation: The manifold impacts of compulsory invention secrecy,” *Management Science*, Vol. 69, pp. 2318–2338.

Guerra, Mara and Rajshree Agarwal. 2021. “Failing to emerge: An empirical investigation of why industries do not succeed,” in *Academy of Management Proceedings*, Vol. 2021, p. 15311.

Guinnane, Timothy, Ron Harris, Naomi R Lamoreaux, and Jean-Laurent Rosenthal. 2007. “Putting the corporation in its place,” *Enterprise & Society*, Vol. 8, No. 3, pp. 687–729.

Guzman, Jorge and Scott Stern. 2020. “The state of American entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 US states, 1988-2014,” *American Economic Journal: Economic Policy*, Vol. 12, No. 4, pp. 212–43.

Hannah, Douglas P and Kathleen M Eisenhardt. 2018. “How firms navigate cooperation and competition in nascent ecosystems,” *Strategic management journal*, Vol. 39, No. 12, pp. 3163–3192.

Jacobides, Michael G. 2005. “Industry change through vertical disintegration: How and why markets emerged in mortgage banking,” *Academy of Management Journal*, Vol. 48, No. 3, pp. 465–498.

Kalyani, Aakash, Nicholas Bloom, Marcela Carvalho, Tarek Alexander Hassan, Josh Lerner, and Ahmed Tahoun. 2024. *The diffusion of new technologies*. NBER Working Paper No. 28999.

Kapoor, Rahul and Nathan R Furr. 2015. “Complementarities and competition: Unpacking the drivers of entrants’ technology choices in the solar photovoltaic industry,” *Strategic management journal*, Vol. 36, No. 3, pp. 416–436.

Kearney, Anne T and Michael E Kornbau. 2005. “An automated industry coding application for new US business establishments,” in *Proceedings of the American Statistical Association*.

Kim, Seojin, Rajshree Agarwal, and Brent Goldfarb. 2025. "Creating radical technologies and competencies: Revisiting Interorganizational dynamics in the nascent bionic prosthetic industry," *Organization Science*, Vol. 37, No. 1, p. 272–301.

Klepper, Steven. 1996. "Entry, exit, growth, and innovation over the product life cycle," *American Economic Review*, pp. 562–583.

Klepper, Steven and Elizabeth Graddy. 1990. "The evolution of new industries and the determinants of market structure," *RAND Journal of Economics*, pp. 27–44.

McDevitt, Ryan C. 2014. ""A" business by any other name: Firm name choice as a signal of firm quality," *Journal of Political Economy*, Vol. 122, No. 4, pp. 909–944.

Moeen, Mahka. 2017. "Entry into nascent industries: Disentangling a firm's capability portfolio at the time of investment versus market entry," *Strategic Management Journal*, Vol. 38, No. 10, pp. 1986–2004.

Moeen, Mahka and Rajshree Agarwal. 2017. "Incubation of an industry: Heterogeneous knowledge bases and modes of value capture," *Strategic Management Journal*, Vol. 38, No. 3, pp. 566–587.

Moeen, Mahka, Rajshree Agarwal, and Sonali K Shah. 2020. "Building industries by building knowledge: Uncertainty reduction over industry milestones," *Strategy Science*, Vol. 5, No. 3, pp. 218–244.

Moeen, Mahka and Will Mitchell. 2020. "How do pre-entrants to the industry incubation stage choose between alliances and acquisitions for technical capabilities and specialized complementary assets?" *Strategic Management Journal*, Vol. 41, No. 8, pp. 1450–1489.

Ozcan, Pinar and Kathleen M Eisenhardt. 2009. "Origin of alliance portfolios: Entrepreneurs, network strategies, and firm performance," *Academy of management journal*, Vol. 52, No. 2, pp. 246–279.

Packalen, Mikko and Jay Bhattacharya. 2012. *Words in patents: Research inputs and the value of innovativeness in invention*. NBER Working Paper No. 18494.

Sanchez, Ron and Joseph T Mahoney. 1996. "Modularity, flexibility, and knowledge management in product and organization design," *Strategic Management Journal*, Vol. 17, No. S2, pp. 63–76.

Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy*. New York: Harper & Brothers.

Stigler, George J. 1951. "The division of labor is limited by the extent of the market," *Journal of Political Economy*, Vol. 59, No. 3, pp. 185–193.

Suarez, Fernando F and James M Utterback. 1995. "Dominant designs and the survival of firms," *Strategic Management Journal*, Vol. 16, No. 6, pp. 415–430.

Szerb, Anna and Nathan Furr. 2025. "The emergence and evolution of innovation ecosystems: Evidence from the solar PV ecosystem," *Strategic Management Journal*, Vol. 46, No. 10, pp. 2605–2657.

Teece, David J. 1986. "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy," *Research Policy*, Vol. 15, No. 6, pp. 285–305.

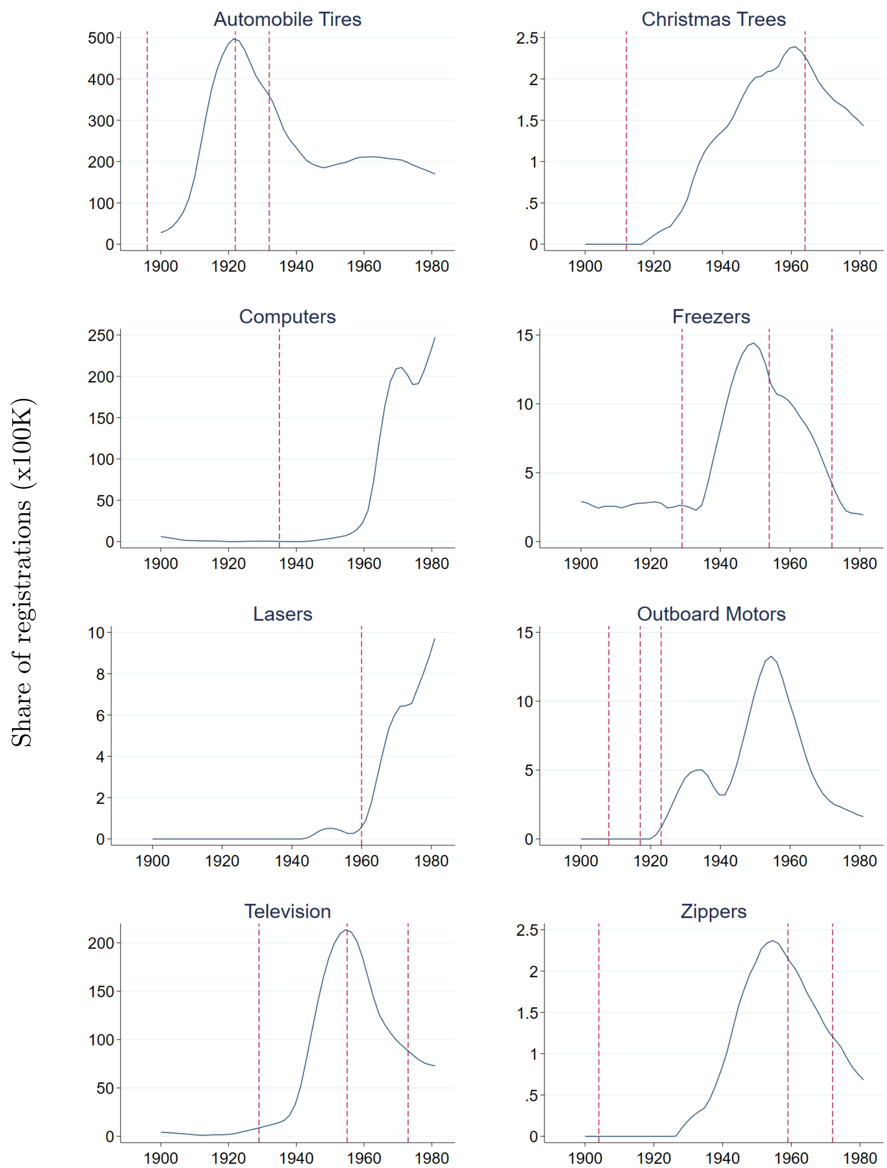
Tushman, Michael L and Philip Anderson. 1986. "Technological discontinuities and organizational environments," *Administrative Science Quarterly*, pp. 439–465.

Wormald, Audra, Rajshree Agarwal, Serguey Braguinsky, and Sonali K Shah. 2021. "David overshadows Goliath: Specializing in generality for internationalization in the global mobile money industry," *Strategic Management Journal*, Vol. 42, No. 8, pp. 1459–1489.

Wormald, Audra, Sonali K Shah, Serguey Braguinsky, and Rajshree Agarwal. 2023. "Pioneering digital platform ecosystems: The role of aligned capabilities and motives in shaping key choices and performance outcomes," *Strategic Management Journal*, Vol. 44, No. 7, pp. 1653–1697.

A Comparison to Klepper and Graddy (1990) lifecycle stages

Figure A.1: Comparison of SCP-derived gross entry lifecycle
to Klepper and Graddy (1990) product lifecycles for select products



Notes: Figure shows the smoothed rate of firm creation (across all sectors) in select products studied by Klepper and Graddy (1990)—who previously identified periods of growth, contraction (shakeout), and stability in the number of manufacturing firms active in each of 46 product categories. We link our business registration data to these products via a manually-curated set of words which we search for in firm names, and focus our analysis here on eight technologies which appear in enough firm names to be able to measure their lifecycles in our data. Each subfigure shows up to three vertical, dashed lines: the first line marks the beginning of Klepper and Graddy Stage 1 (takeoff); if present, the second line marks the start of Stage 2 (shakeout); and if present, the third line marks the start of Stage 3 (stability). Lifecycles plotted using local polynomial smoothing.

Online Appendix

Entrepreneurial Entry Over the Technology Lifecycle

Innessa Colaiacovo, Daniel P. Gross, and Jorge A. Guzman

A Sector Classification Method

In this appendix, we describe our sector classification procedure in detail, and provide supplementary validation results to those in the body of the paper. Portions of this section may reproduce passages in the paper, adding detail and context. We begin as follows.

Let i index firms, j tokens, s sectors. Each firm i has J_i tokens in its name, and there are J unique tokens in the corpus. Let τ_{js} represent a token-sector score, which will be defined below as the fraction of token j's total occurrences that are in each sector s. φ_{is} will represent a firm-sector score, which will be an average of τ_{js} across all J_i tokens in firm i's name. τ_{js} and φ_{is} will thus measure each token or firm's association with each of $S = 10$ sectors. To aid the exposition below, let n_{js} represent the number of times token j is in the name of firms in sector s, and $N_J = \sum_{s=1}^{10} n_{js}$ represent the number of firms in which token j appears overall.

A.1 Measuring token-sector scores

We first undertake several steps to prepare the D&B data for training and validating our classification procedure. We begin by cleaning firm names of stop words, special characters, tokens of length $\leq$2, and all-numeric tokens. We then split out the tokens in firm names, and reshape the firm-level data to a token-level dataset, with one observation per firm-token. Because the D&B company name field is fixed in width, words in firm names are sometimes abbreviated. We thus extend our data cleaning effort by creating an extensive, part-manual and part-automated crosswalk from tokens which appear to be abbreviations or minor misspellings to full words. A total of roughly 65,000 abbreviated tokens get crosswalked to full words through this approach. For example, "IMPRVMNT" is updated to "IMPROVEMENT", and "PIZZARIA" to "PIZZERIA". We believe these changes ultimately increase the fidelity of our data to the underlying economic reality and will improve the quality of our sector classification. A total of 65,000 unique tokens in the D&B 1990 sample are revised through these methods, out of roughly 10 times as many tokens in the data. The three most common words in this sample are "SERVICE", "CONSTRUCTION", and "SHOP"—words which are (seemingly) immediately revealing of the sector that a firm with those words is in, and illustrate the face validity of our proposed approach.

We take two approaches to further reducing (or not) the set of tokens which we will use to predict firms' economic sector. In the first variant, we deploy this set of tokens as is. In our second variant, we remove additional stop words, excluding those which are *not* an English-language word but are any of (i) a word in a U.S. state name, (ii) a word in a U.S. city name, (iii) a common given name, or (iv) a common surname.[1] We exclude these tokens because we think they are

[1]Our sample of given names consists of the union of the 1,000 most common baby names for birth years 1880 to 2009, according to the Social Security Administration, obtained from https://github.com/hadley/data-baby-names. Our sample of surnames consists of all surnames which appeared $\geq$100 times in the 2010 census, obtained from https://www.census.gov/topics/population/genealogy/data/2010_surnames.html. In addition to excluding these names, we also exclude variants of these names plus the letter 'S', to account for the sizable number of businesses named in the possessive (e.g., "MIKE'S FOOD & SPIRITS" or "DAVE'S FRESH PASTA").

both frequently occurring in firm names, due to geographic specialization or eponymy, and because they are sector-agnostic and thus may obscure our sectoral predictions. The resulting dataset has firm-tokens and their associated sector, obtained from each firm's D&B-reported SIC code.

We then measure τ_{js} as sector s's share of all uses of token j: $\tau_{js} = \frac{n_{js}}{N_J}$. Although we experimented with other approaches to scoring tokens' associations with each sector, we found that this simple approach outperformed others across many validation tests.[2] By construction, each token's sector-level scores τ_{js} will add to 1 when summed across sectors.

A.2 Measuring firm-sector scores

To get from token-sector scores to firm-sector scores (our target output), we need to aggregate across tokens in firm names—reducing the dimensionality of each firm i from $J_i \times 10$ to 10. We experimented with two approaches: straight averages and weighted averages, weighting by (the square root of) a token's total usage. This latter approach will overweight more-common words, whose sector associations we might measure more precisely. Formally, these two approaches to measuring the firm-sector score φ_{is} can be characterized as follows:

$$\textit{Approach 1:} \qquad \varphi_{is}^{uwtd} = \frac{1}{J_i} \sum_{j=1}^{J_i} \tau_{js}$$

$$\textit{Approach 2:} \qquad \varphi_{is}^{wtd} = \frac{1}{\sum_{j=1}^{J_i} \sqrt{N_j}} \sum_{j=1}^{J_i} \sqrt{N_j} \tau_{js}$$

To illustrate, suppose there are two tokens (A and B) and two sectors (1 and 2). To make this example concrete, let us specifically imagine a firm with the name "Master Plumbing", and that there are two sectors, *Construction* and *Retail Trade*. We don't ex-ante know what sector the firm is in. It might be a construction contractor. It might be a plumbing fixture store. But the words in its name give us clues—especially when we can gauge how often these words associate with each of these sectors in our training data. For this example, let us denote "Master" and "Plumbing" as tokens A and B, and *Construction* and *Retail Trade* as sectors 1 and 2.

Suppose that in our training sample, token A ("Master") is used in 225 firm names, and appears 50% of the time in sector 1 (*Construction*) and 50% in sector 2 (*Retail Trade*). Token B ("Plumbing"), on the other hand, is in 25 firm names, and appears 90% of the time in sector 1 and 10% in sector 2. Table A.1 provides the computed sector scores under each of these two approaches. In practice, we find that the simpler aggregation method (unweighted averages) better predicts the true distribution, which makes this our preferred approach.

[2] An alternative approach we explored was to square these shares, to overweight high token-sector associations. Doing so, however, reduced the rate at which we correctly predicted firms' actual sector.

Table A.1: Firm-sector scores for "Master Plumbing" (contrived example)

	Sector 1	Sector 2
Approach 1: Unweighted	$\varphi_{i1} = \frac{0.5+0.9}{2} = 0.7$	$\varphi_{i2} = \frac{0.5+0.1}{2} = 0.3$
Approach 2: Weighted	$\varphi_{i1} = \frac{\sqrt{225}\cdot 0.5+\sqrt{25}\cdot 0.9}{\sqrt{225}+\sqrt{25}} = 0.6$	$\varphi_{i2} = \frac{\sqrt{225}\cdot 0.5+\sqrt{25}\cdot 0.1}{\sqrt{225}+\sqrt{25}} = 0.4$

For a real example from our data, consider the firm "Anderson Home Appliances", whose words can be fractionally classified across each of the original 10 economic sectors. In our data, these words associate with sectors as shown in Table A.2 (where only the top five sectors are shown), leading to the firm-level scores shown in the last row of the table. Our procedure predicts *Retail Trade* and *Services* as equally likely sectors. The actual primary sector reported for this firm in the D&B data is *Services*, and the secondary sector reported in D&B is *Retail Trade*—though we might infer that Anderson Home Appliances is related to *Construction* too.

Table A.2: Firm-sector scores for "Anderson Home Appliances" (true example)

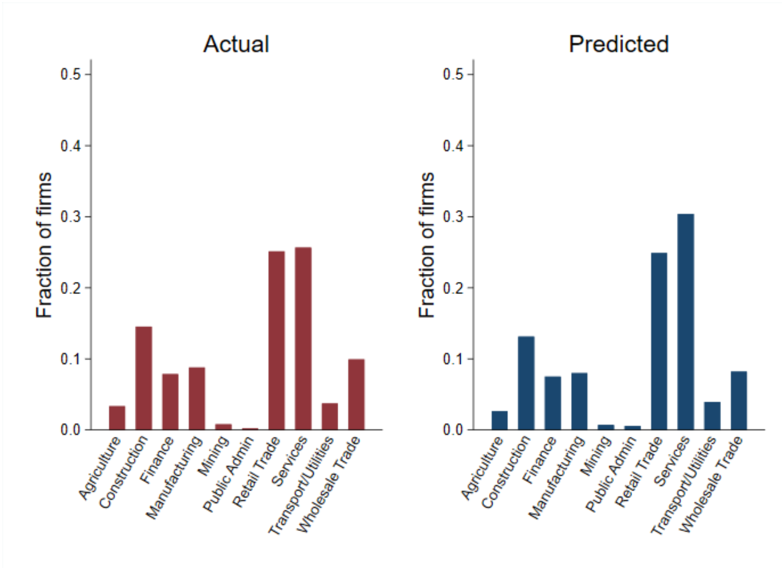
	Token-sector scores and cross-token averages				
Word	Agriculture	Construction	Fin./Ins./RE	Retail Trade	Services
Anderson	8%	17%	10%	14%	34%
Home	0%	20%	10%	17%	48%
Appliances	0%	2%	0%	71%	21%
Average	3%	13%	7%	34%	34%

A.3 Validating our procedure

We validate this sector classification procedure in several ways. We begin by validating within our D&B test sample, where we can classify firms into sectors and compare our predictions to what is reported in the data. We focus our validation evidence on our more restrictive variant, where we remove place and person names from firm names before classifying sectors.

Figure A.1 begins with a high level view, aggregating sector shares across the sample and showing the overall frequency of firms in each of our 10 sectors as observed (left panel) and as predicted (right panel). The distributions line up closely, though we slightly overpredict *Services* and slightly underpredict *Wholesale Trade*. Table A.3 provides a more nuanced view of our predictive validity, showing, for each sector (first column), how often firms predicted to be in that sector are in fact reported by D&B to be in that sector (second column), as well as how often in each such case our second-most likely prediction is the D&B-reported sector (third column). The final column shows the sum: the probability that one of the top two predicted sectors is the reported sector. Across all predicted sectors, we match D&B 70 to 90% of the time with our top one or two predictions. We consider these rates to be high, particularly as we do not expect to match D&B 100% of the time, given that many firms span boundaries between sectors.

Figure A.1: Predicted vs. actual distribution of firms across sectors in D&B testing sample



Notes: Figure shows the predicted and actual distribution of sectors in our D&B test sample, using the sector classification algorithm described in the paper.

Table A.3: Sector classification algorithm: Predictive performance, by sector

	Fraction of firms where:		
Predicted sector	Prediction correct	Runner-up prediction correct	Total
Agriculture, Forestry, & Fishing	78%	9%	87%
Construction	84%	7%	91%
Finance, Insurance, & Real Estate	82%	8%	90%
Manufacturing	67%	15%	82%
Mining	58%	13%	71%
Public Administration	81%	9%	91%
Retail Trade	73%	13%	86%
Services	66%	16%	83%
Transportation & Public Utilities	84%	6%	91%
Wholesale Trade	57%	21%	78%
Overall	71%	14%	85%

Notes: Table shows, for each sector (first column), how often firms predicted to be in that sector (i.e., whose top predicted sector is that shown in the left column) are in fact reported by D&B to be in that sector (second column), as well as how often in each such case our second-most likely prediction is the D&B-reported sector (third column).

In Table A.4, we break these patterns down even further. Here we document, for each top-predicted sector (left column), the fraction of firms with that prediction that are reported by D&B in each of our 10 sectors. The first features of this table to observe is that the mass is overwhelmingly concentrated along the diagonal (which reproduces Column 2 of Table A.3). In addition, the table also illustrates the sectors that are likely to be jointly present—or incorrectly classified, depending on interpretation. Firms which we predict to be in *Wholesale Trade* (row 10), for example, are 18% of the time actually reported as being in *Retail Trade*—illustrating that firms may sometimes be a bit of each sector, and the occasional challenging of distinguishing the two.

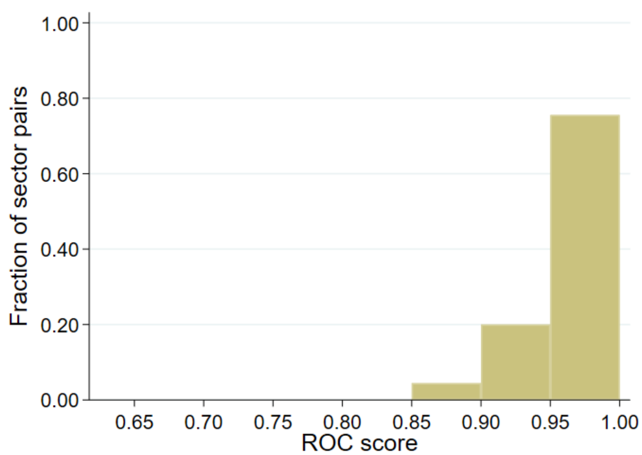
Table A.4: Sector classification algorithm performance: Predicted vs. actual sector

When top predicted sector is...		Pr(Actual sector is...)									
Sector	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Agriculture, Forestry, & Fishing	**78%**	2%	2%	1%	0%	0%	7%	5%	1%	4%
2	Construction	1%	**84%**	1%	3%	1%	0%	3%	4%	1%	3%
3	Finance, Insurance, & Real Estate	1%	4%	**82%**	1%	0%	0%	3%	7%	1%	2%
4	Manufacturing	1%	4%	1%	**67%**	0%	0%	7%	9%	1%	11%
5	Mining	1%	4%	4%	4%	**58%**	0%	9%	6%	3%	12%
6	Public Administration	1%	1%	4%	1%	0%	**81%**	2%	6%	3%	2%
7	Retail Trade	2%	2%	2%	5%	0%	0%	**73%**	9%	1%	7%
8	Services	2%	4%	4%	4%	0%	1%	11%	**66%**	2%	5%
9	Transportation & Public Utilities	1%	2%	1%	1%	0%	0%	3%	5%	**84%**	2%
10	Wholesale Trade	2%	3%	1%	8%	2%	0%	18%	7%	2%	**57%**

Notes: Table shows, for each sector, how often firms predicted to be in that sector (i.e., whose top predicted sector is that shown in the left column) are fact reported by D&B to be in each of the ten sectors. The diagonal reproduces the rates from the second column of Table A.3.

As another check, we make all unique pairs of sectors, and for each pair, we filter to firms whose reported sector is one of the two and use receiver-operating characteristic (ROC) analysis to evaluate for how many firms the more probable sector is the reported one. Figure A.2 plots the distribution of ROC scores among these pairs. The scores are near one, indicating that in each of these two-sector horseraces, we nearly always predict the D&B-reported sector.

Figure A.2: Distribution of ROC scores for prediction in sector pairs in D&B testing sample

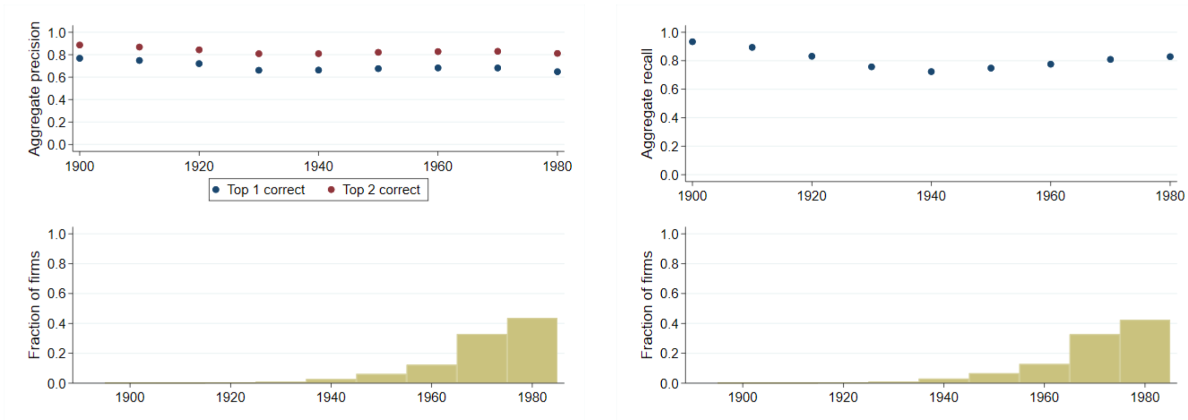


Notes: Figure shows distribution of ROC scores over sector pairs, evaluating across our test sample how often the more-predicted sector in each pair is the firm's reported sector, conditional on one of the two sectors in the pair being the true sector. With 10 sectors, the distribution in this figure is over $\sum_{s=1}^{9} s$ pairs. In the vast majority of these pairings, our algorithm predicts a firm's true sector over the alternative >95% of the time.

In additional validation tests, we examine the stability of this performance across time, for firms with different D&B-reported founding years, and we find that our procedure's performance is stable in both the rate at which its predictions match D&B-reported sectors, and in the fraction of firms that get classified to a sector (Figure A.3). We also perform validation tests against an independent third sample of firms in the Infogroup USA 2000 data extract. The performance of our procedure in this sample is similar to that in our D&B test sample, suggesting that the results above are not driven by distinctive shared features of the D&B training and test samples.

Figure A.3: Precision of sector predictions vs. rate of prediction in D&B testing sample

Panel (A): Pr(True sector is top-predicted) | Panel (B): Fraction of firms classified w/ sector



Notes: The top chart in each panel shows how precision and recall rates vary over time (across firm registration decade) in the D&B testing sample. Bottom panel shows the distribution of firms in this sample across decades. Aggregate precision and recall is a weighted average of those in the upper scatterplots, weighted by the density in the lower histograms.

A.4 Classifying the business registration sample

Our final step is to classify firms in our business registration data to economic sectors. We clean firm names by the same approach applied to D&B firms in Section A.1, up to expanding abbreviations and correcting minor misspellings, and again remove place and person names that are not English-language words. We then apply the token-sector scores calculated in Section A.1 to these firms' names, and aggregate to the firm level following Section A.2. The resulting dataset has sector scores for each registered firm, across each of ten economic sectors (*Agriculture* to *Wholesale Trade*). Recall that these scores will add to one, and can alternatively be interpreted as the probability that a firm is in a given sector, or the firm's fractional association with that sector—which allows firms to span sectoral boundaries, as many firms do in practice.

This classification nevertheless bears limitations. One is that some words in the set of firm names in the registration sample may not be present in the Dun & Bradstreet sample, and thus will be unable to be scored. Due to this, as well as the removal of stop words, some firms in the registration data may not be able to have their sector predicted. We are reassured, however, by the fact that even under our more restrictive sector classification procedure, we classify 90% of firms—and this rate is stable over time. We believe this is high enough to generate useful measures of the sectoral distribution of U.S. entrepreneurship over long periods.

B Additional Evidence on Technology-Firm Linkages: Patent and Firm Counts, Common Paired Words in Firm Names, and Example Firms

Table B.1: List of technology terms and related statistics from firm names

Technology	Patents	Firms	Year of 10th firm	Most paired word	2nd most paired word	3rd most paired word
SUBSTRATE	161347	80	1992	TECHNOLOGY	SOLUTION	LASER
SEMICONDUCTOR	88544	1977	1960	TECHNOLOGY	EQUIPMENT	SERVICE
POLYMER	85253	4184	1952	TECHNOLOGY	SYSTEM	PRODUCT
CATALYST	77798	3174	1964	SERVICE	CONSULTING	TECHNOLOGY
TRANSISTOR	76986	75	1971	DEVICE	SERVICE	ALL
SENSOR	73909	1231	1967	TECHNOLOGY	SYSTEM	CONTROL
ROTOR	62808	447	1956	WING	SERVICE	AVIATION
COMPUTER	57295	98530	1948	SERVICE	SYSTEM	SOLUTION
AMPLIFIER	48044	118	1989	CALIFORNIA	TELEVISION	CUSTOM
CAPACITOR	46707	119	1967	TECHNOLOGY	SALE	CALIFORNIA
RESISTOR	38977	50	1982	PRECISION	ELECTRONIC	GLOVER
ELECTRONIC	35422	59984	1923	SERVICE	SYSTEM	COMMUNICATION
COPOLYMER	32494	30	1984	LION	CHEMICAL	DSM
PROCESSOR	32428	2665	1937	FOOD	MEAT	GAS
TRANSDUCER	27111	114	1968	TECHNOLOGY	SYSTEM	RESEARCH
VIDEO	26315	52958	1941	PRODUCTION	AUDIO	SERVICE
DIODE	24390	65	1981	LASER	TECHNOLOGY	HOLDING
AUTOMOBILE	24072	6624	1900	CLUB	SERVICE	DEALER
VINYL	21286	4689	1954	SIDING	PRODUCT	CARPET
RADIO	20932	21350	1910	CLUB	AMATEUR	SERVICE
WAVELENGTH	19665	178	1979	COMMUNICATION	PRODUCTION	TECHNOLOGY
AIRCRAFT	18820	18145	1922	SERVICE	SALE	LEASING
SIDEWALL	18030	19	0	ROOFING	PRODUCT	SPECIALTY
PLANAR	18004	74	1983	SYSTEM	TECHNOLOGY	IPEC
THERMOPLASTIC	17216	166	1966	CAL	SYSTEM	ADVANCED
TRACTOR	16566	11879	1913	SERVICE	EQUIPMENT	TRUCK
ANALOG	15469	384	1973	DIGITAL	DESIGN	SYSTEM
WAVEGUIDE	15196	37	2001	CONSULTING	COMMUNICATION	ENGINEERING
AUDIO	14143	17062	1948	VIDEO	VISUAL	SERVICE
PHOTOSENSITIVE	13246	3	0			
POLYESTER	13150	55	1995	FILM	MITSUBISHI	NESTE
DATABASE	12350	1791	1978	SOLUTION	SYSTEM	SERVICE
PARTICULATE	11167	19	0	CONTROL	SYSTEM	TECHNOLOGY
EPOXY	11094	519	1968	SYSTEM	COATING	FLOORING
CIRCUITRY	10819	54	1987	ADVANCED	BOARD	DESIGN
WIRELESS	10565	28732	1924	COMMUNICATION	SERVICE	SOLUTION
COMPARATOR	10560	10	0	SERVICE	PRODUCT	SYSTEM
MOLECULE	10293	75	1990	DESIGNED	IMMUNO	RESEARCH
PHARMACEUTICALLY	10257	2	0	ELEGANT	FORMERLY	PHARMACEUTIC
PIEZOELECTRIC	10242	2	0	DEVICE	INVESTOR	TECHNOLOGY
FORMALDEHYDE	10189	12	0	FOAM	INSULATION	UREA
COOLANT	9818	74	1989	SERVICE	SYSTEM	MANAGEMENT
ENZYME	9717	225	1965	TECHNOLOGY	RESEARCH	PRODUCT
MICROWAVE	8833	1633	1955	SERVICE	SYSTEM	TECHNOLOGY
ULTRASONIC	8772	401	1964	CLEANING	BLIND	SERVICE
INVERTER	8620	27	1999	POWER	SERVICE	SOLAR
SERVO	8472	231	1967	SYSTEM	TECHNOLOGY	DYNAMIC
REFRACTIVE	8212	264	1982	SURGERY	CENTER	LASER
BINARY	8211	594	1975	SYSTEM	SOLUTION	TECHNOLOGY

Notes: Table lists the 50 top technologies in our sample (by number of patents) and (i) their number of associated firms in our data, (ii) the year of the tenth firm, and (iii) the top three words which co-occur in the names of firms with that technology (excluding stop words).

Table B.2: Example firms for each technology (randomly chosen)

Technology	Patents	Firms	Example 1	Example 2	Example 3
COMPUTER	57295	98530	KEPLER COMPUTER INC.	ST. CROIX COMPUTERS, LLC	COMPUTER DIRECTIONS, INC.
AUTOMOTIVE	2438	70954	GALYEON AUTOMOTIVE, LLC	W.A.M. AUTOMOTIVE, INC.	BNA AUTOMOTIVE GROUP, INC.
SOFTWARE	6198	63181	MINI B SOFTWARE, INC.	DENIZEN SOFTWARE, LLC	BEST SOFTWARE OF CALIFORNIA
ELECTRONIC	35422	59984	DM ELECTRONICS INC.	COLT ELECTRONICS CO.	GEMCO ELECTRONICS, INC.
VIDEO	26315	52958	COBB VIDEO, INC.	VIDEO EMPORIUM, INC.	LEGAL VIDEO SERVICES, INC.
WIRELESS	10565	28732	CYVO WIRELESS LLC	VANGUARD WIRELESS, INC	WIRELESS BUILDERS, LLC
RADIO	20932	21350	RADIO FREE FRISCO LLC	MUSIC SOUND RADIO, INC.	CONTINENTAL RADIO ASSOCIATION
AIRCRAFT	18820	18145	DAMAR AIRCRAFT CORPORATION	LAIRD AIRCRAFT CORPORATION	AIRCRAFT-MARINE PRODUCTS INC
INTERNET	2387	17268	DC INTERNET GROUP, LLC	WORLDPORT INTERNET, INC.	SAEROM INTERNET WORLD II INC.
AUDIO	14143	17062	MARINE AUDIO INCORPORATED	AUDIO INTERNATIONAL, INC.	SAPPHIRE AUDIO & VIDEO, LLC
TELECOMMUNICATION	2701	15645	MJM TELECOMMUNICATIONS, INC.	DB SAT TELECOMMUNICATIONS LLC	K & M TELECOMMUNICATIONS INC.
TELEVISION	7598	12642	GABLES TELEVISION, INC.	BUSINESS TELEVISION, INC.	U.S. PANTS & TELEVISION, INC.
ONLINE	319	12423	REO ONLINE INVESTORS, LLC	SALVAGE AUTOS ONLINE, INC.	SIGHTLINES ONLINE SALES, LLC
MICRO	6560	12377	MICRO FLUIDS, L.L.C.	CENTURY MICRO TEK, INC.	MICRO CONSTRUCTION, INC
TRACTOR	16566	11879	TRACTOR 6943, INC.	KELLEY TRACTOR COMPANY INC	C & W TRACTOR COMPANY, INC.
PIPELINE	5007	11485	SEVEN-A PIPELINE, LLC	JUNIOR'S PIPELINE CO., INC.	INTERCHANGE PIPELINE COMPANY
MOTORCYCLE	2465	8315	MOAK'S MOTORCYCLES, INC.	MOTORCYCLE RENTALS, INC.	NEVADA MOTORCYCLE ADVENTURES
DIESEL	1842	8132	DIESEL DIRECT LLC	DIESEL WORKS, INC.	DIESEL DISPATH INC
PHONE	2883	7345	THE ENTER-PHONE, INC.	CALIFORNIA PHONES LP-16	FRANKY THE PHONE GUY INC
MACHINING	5514	7154	J & W MACHINING, INC.	MOHAWK VALLEY MACHINING, INC.	MACON PRECISION MACHINING LLC
AUTOMOBILE	24072	6624	WESTERN AUTOMOBILE COMPANY	BIG FOUR AUTOMOBILE COMPANY	ILLINOIS AUTOMOBILE CLUB INC
MULTIMEDIA	1351	5562	CREAM MULTIMEDIA INC	MARION MULTIMEDIA INC.	STUDIO H MULTIMEDIA, INC.
VINYL	21286	4689	T & C VINYL, LLC	HANSEN?S VINYL SIDING LLC	VINYL-MASTER INTERNATIONAL
HELICOPTER	1687	4478	T A HELICOPTERS, LLC	HELICOPTER PATROL, INC.	VACA VALLEY HELICOPTERS, INC.
PROPANE	2483	4444	PROPANE CORPORATION	SUBURBAN PROPANE, L.P.	PLANTATION PROPANE, INC.
HVAC	282	4278	AIR-FLO HVAC, INC.	HVAC SOLUTIONS, LLC	AMETECH HVAC CORPORATION
POLYMER	85253	4184	HTI POLYMER, INC.	TICONA POLYMERS, INC.	PORTLAND POLYMERS, LLC
BROADBAND	909	3643	N2 BROADBAND, INC.	BROADBAND PROPERTIES, LLC	ZTARK BROADBAND CORPORATION
ORTHODONTIC	1029	3584	A & A ORTHODONTICS, INC.	JENNINGS ORTHODONTICS, PLLC	R. NIKODEM ORTHODONTIC, INC.
FIBERGLASS	951	3217	H AND H FIBERGLASS, INC.	BLADOW'S FIBERGLASS, INC.	BOSS FIBERGLASS REPAIR LLC
CATALYST	77798	3174	CATALYST PROPERTIES, LLC	CATALYST IT SERVICES, INC.	SOLUTIONS CATALYST GROUP, LLC
PROCESSOR	32428	2665	PET PROCESSORS, LLC	STEEL PROCESSORS, LLC	APPLIED PROCESSORS CORP.
PROSTHETIC	1425	2560	VIVO PROSTHETICS, INC.	CREATIVE PROSTHETICS, INC.	DALE CLARK PROSTHETICS, INC.
CARDIAC	2302	2434	CARDIAC PACEMAKERS, INC.	CARDIAC REHABILITATION, INC.	CARDIAC PATHWAYS CORPORATION
PAINTBALL	197	2428	JP PAINTBALL LLC	NYS PAINTBALL SUPPLY LLC	GLADIATOR PAINTBALL, INC.
GENETIC	482	2225	GENETIC ID, INC.	BRAND GENETICS, LLC	ULTIMATE GENETICS, LLC
COMPUTERIZED	486	2000	COMPUTERIZED SPORTS, INC.	COMPUTERIZED SERVICES, INC.	COMPUTERIZED MARKETING, INC.
SNOWMOBILE	678	1994	ALBANY SNOWMOBILE CLUB, INC.	SUNRISE LAKE SNOWMOBILE CLUB	SKYLINE SNOWMOBILE CLUB, INC.
SEMICONDUCTOR	88544	1977	IDEAL SEMICONDUCTOR, INC.	PHILIPS SEMICONDUCTORS INC.	SAMSUNG SEMICONDUCTOR, INC.
DATABASE	12350	1791	MATERIAL DATABASE CORPORATION	DATABASE DEVELOPERS GROUP LLC	DATABASE RESEARCH CORPORATION
MICROWAVE	8833	1633	CVN MICROWAVE, INC.	ASTERIA MICROWAVE LLC	MID-TEXAS MICROWAVE, INC.
ROBOTIC	756	1632	MAZOR ROBOTICS INC.	FUTURE ROBOTICS, INC.	INTERNATIONAL ROBOTICS INC.
INTEGRATOR	3184	1561	ELITE INTEGRATORS LLC	BAUD SYSTEM INTEGRATORS LLC	TECHNOLOGY INTEGRATORS INC.
ULTRASOUND	2522	1541	ABC ULTRASOUND, LLC	BUFFALO ULTRASOUND, INC.	PYRAMID ULTRASOUND, INC.
PLYWOOD	721	1422	PLYWOOD IMPORTERS, LTD.	CITY PLYWOOD CENTER, INC.	VANCOUVER PLYWOOD CO., INC.
WATERMARK	1043	1321	WATERMARK REALTY, LLC	WATERMARK SIGNAGE, INC.	THE WATERMARK GROUP, INC.
SENSOR	73909	1231	CLOUDTECH SENSORS, INC.	SENSOR TECHNOLOGIES, LLC	INDUSTRIAL SENSORS, INC.
WINDSHIELD	6383	1079	MR. WINDSHIELD, INC.	WINDSHIELD MAN, INC.	B C C WINDSHIELD REPAIR LLC
WEBSITE	265	1071	WEBSITE EXPRESS	CHEETAH WEBSITES INC.	ICON WEBSITE DESIGN LLC

Notes: Table lists the 50 top technologies in our sample (by number of firms) and provides up to three examples of firms with that word in their name. Example firms restricted to those with less than 30 characters in firm name for readability.

C Supplementary Results

The following tables reproduce Tables 5 and 7 of the paper, limiting our technology sample to only nouns. The results are quantitatively and statistically similar.

Table C.1: Dispersion of entrepreneurial activity across the technology lifecycle
robustness check: excluding non-nouns from the technology sample

	Activity			Sector		
	(1)	(2)	(3)	(4)	(5)	(6)
	Top 1	HHI	C.V.	Top 1	HHI	C.V.
11-20 years into lifecycle	-0.059	-0.063	-0.119	-0.010	-0.012	-0.027
	(0.015)	(0.017)	(0.032)	(0.006)	(0.005)	(0.014)
21-30 years into lifecycle	-0.074	-0.073	-0.145	-0.014	-0.017	-0.040
	(0.021)	(0.025)	(0.045)	(0.007)	(0.006)	(0.016)
31-40 years into lifecycle	-0.110	-0.116	-0.222	-0.023	-0.028	-0.064
	(0.026)	(0.033)	(0.062)	(0.008)	(0.007)	(0.019)
41-50 years into lifecycle	-0.135	-0.145	-0.284	-0.022	-0.026	-0.064
	(0.034)	(0.042)	(0.078)	(0.010)	(0.008)	(0.023)
51-60 years into lifecycle	-0.182	-0.194	-0.375	-0.021	-0.022	-0.060
	(0.040)	(0.049)	(0.093)	(0.011)	(0.009)	(0.027)
61+ years into lifecycle	-0.214	-0.236	-0.451	-0.021	-0.024	-0.068
	(0.051)	(0.063)	(0.119)	(0.014)	(0.010)	(0.032)
N	3771	3771	3771	10421	10421	10421
R^2	0.32	0.38	0.38	0.53	0.47	0.53
Token FEs	Y	Y	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y	Y	Y
Y mean	0.88	0.84	1.96	0.50	0.38	0.77

Notes: Table estimates changes in the dispersion of firm creation across economic sectors or the value chain over time. Dispersion measures are (i) the top activity/sector share ("Top 1"; Columns 1 and 4), the activity/sector HHI ("HHI"; Columns 2 and 5), and the coefficient of variation ("C.V."—in essence, the standard deviation of activity/sector shares; Columns 3 and 6). SEs clustered by technology in parentheses.

Table C.2: Upstream to downstream industry evolution
robustness check: excluding non-nouns from the technology sample

	Activity		Sector	
	(1)	(2)	(3)	(4)
	Manuf.	Services	Manuf.	Services
11-20 years into lifecycle	-0.031	0.047	-0.007	0.008
	(0.038)	(0.046)	(0.005)	(0.006)
21-30 years into lifecycle	-0.101	0.109	-0.011	0.026
	(0.040)	(0.051)	(0.006)	(0.007)
31-40 years into lifecycle	-0.119	0.099	-0.023	0.029
	(0.040)	(0.056)	(0.007)	(0.008)
41-50 years into lifecycle	-0.149	0.148	-0.026	0.036
	(0.040)	(0.060)	(0.007)	(0.009)
51-60 years into lifecycle	-0.168	0.136	-0.024	0.037
	(0.044)	(0.067)	(0.009)	(0.010)
61+ years into lifecycle	-0.178	0.165	-0.030	0.046
	(0.046)	(0.073)	(0.010)	(0.012)
N	3771	3771	10421	10421
R^2	0.33	0.45	0.76	0.70
Token FEs	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y
Y mean	0.07	0.70	0.18	0.35

Notes: Table estimates changes in the share of firm creation across value chain activities and economic sectors, comparing manufacturing and services. SEs clustered by technology in parentheses.