Search Speed and Theory Discovery: A Model for Innovation

Arnaldo Camuffo Alfonso Gambardella Andrea Pignataro *

September 19, 2024

Abstract

We develop a model for analyzing firm innovation strategies, distinguishing between the exploitation of established firm-level paradigms (which we call "theories") and the exploration of new paradigms. Firms engage in a staged search process, balancing the exploitation of current theories with the discovery of new ones. Central to the model is the role of speed, which we identify as a key driver of innovation and growth. The empirical analysis of a sample of all US public firms with at least one patent between 1986 and 2015, reveals that the speed at which firms exploit innovations within their existing paradigms is positively correlated with their exploration of new technological areas, asset growth, and an increase in patent output relative to R&D investment. Importantly, speed and exploration exhibit a complementary relationship, which only becomes significant beyond a certain speed threshold. The model identifies four key drivers of innovation success: search speed within existing theories, the probability of discovering new theories, the potential to generate new paradigms, and capital assets. High search speeds boost early-stage productivity and resource accumulation, increasing the likelihood of discovering new theories and sustaining growth. Firms with strong theory generation and capital assets benefit significantly, explaining the hyper-growth seen in many high-tech companies. Conversely, slow search speeds lead to a focus on exploiting current theories, limiting new exploration and reinvestment. Policymakers should consider interventions that accelerate search-enhancing technologies, reduce capital costs, and foster theory generation across industries, promoting equitable growth and reducing disparities in innovation capacity.

JEL: L21, L26, M13, M21

Keywords: Innovation, exploration, AI, speed, firm strategies, growth

*Camuffo: Department of Management & Technology, Bocconi University and ION Management Science Lab, *arnaldo.camuffo@unibocconi.it*. Gambardella: Department of Management & Technology and ION Management Science Lab, Bocconi University, *alfonso.gambardella@unibocconi.it*. Pignataro: ION Group, *ap@iongroup.com*.

1 Introduction

Innovation is a key driver of economic growth. At the micro-level, there is extensive literature on innovation (Hall & Rosenberg , 2010; Hall & Helmers , 2024). At the macro-level, the new growth literature, stemming from Romer (1990) and Aghion & Howitt (1992), places innovation at the center of the growth processes for countries, industries, and firms.

Surprisingly, however, neither the micro nor macro literature has thoroughly explored one crucial concept: *speed*. Intuitively, the speed at which innovations are generated impacts firms, industries, and economies, raising critical questions about how and when speed benefits them, and the mechanisms through which it influences these processes. This issue is highly relevant. For instance, Rosenberg (1992) showed that technical advancements in scientific instrumentation enhanced both the opportunities for and the speed of innovation in industries. Similarly, various technologies, such as computers, have accelerated the pace of activities, including innovation search and generation. Today's artificial intelligence (AI) holds the promise of significantly increasing the speed of search and, more broadly, the generation of innovations.

This paper introduces a framework for analyzing firm innovation strategies through the lens of search processes. We combine speed with a fundamental feature of innovation search: firms either search for innovations within established theories or seek to discover new theories before searching within them. Building on Dosi's (1982) paradigm shift concept, which parallels Kuhn's process of scientific revolutions, we view theories as firm-level paradigms. In our dynamic model, firms search for innovations in a staged process, allocating a given amount of time (a proxy for resources) between searching within current theories and exploring new ones. Discovering a new theory resets the process, while

searching within a theory gradually exhausts its innovations, prompting firms to shift resources toward discovering new theories. Higher search speeds exhaust theories more rapidly, leading to earlier and larger resource allocations toward new theory discovery.

The model predicts that firm size and innovation are influenced by past growth, search speed, and the propensity for basic research. It highlights how firms of different sizes prioritize basic versus applied research, thereby shaping their growth strategies. Firms with high search speed, research propensity, and size tend to favor radical innovations. Additionally, the model identifies nonlinear effects of search speed, where speed only drives greater innovation in firms exceeding a certain threshold. These divergent growth strategies explain why some firms experience rapid growth while others remain small.

Section 2 reviews the literature. Section 3 provides motivating evidence, showing that an empirical proxy for firm-level speed correlates with exploration, asset growth, and the generation of new ideas. Sections 4 and 5 present our model and its implications, while Sections 6 and 7 offer discussion and conclusions. The Appendix includes derivations, proofs, and the construction of our measure of patent clusters.

2 Related Literature

The framework of our paper sheds light on several puzzles in innovation and economic growth. We address Christensen's (1997) paradox: despite efficient resource allocation, established firms struggle with disruptive innovations due to difficulty identifying new opportunities (selection problem) and the threat to existing businesses (cannibalization problem). Our model highlights the critical role of search speed in this context. Slow search incentivizes firms to exploit existing theories, leaving insufficient resources for exploring new opportunities. In contrast, faster search exhausts existing opportunities

within a theory, making the exploration of new frameworks optimal.

This finding aligns with Levinthal & March's (1993) concept of "myopia of learning," where firms specializing in a slow search environment within a bounded domain, experience diminishing returns but adopt this focus as an optimal strategy under the constraints of slow search. Slow search restricts firms to exploiting existing knowledge, hindering the exploration of new areas.

Bloom et al. (2020) showed that "ideas are harder to find" despite significant innovation resources, consistently with our model's predictions under slow search technology. Notably, our model suggests that it is not a lack of inherent capability that hinders new ideas but rather the optimal resource allocation towards existing trajectories when search is slow. This implies that slow search allows for continued investment within an established framework, even as the innovation rate declines, delaying the need to explore entirely new theories.

A growing body of literature explores AI's impact on innovation and growth (Furman & Seamans, 2019; Agrawal et al., 2019; McElheran et al., 2024). AI, characterized as a general-purpose technology (Cockburn et al., 2019; Trajtenberg, 2019) has the potential to significantly increase search speed (Agrawal et al., 2024). Our model predicts a critical threshold for search speed, beyond which speed and the probability of discovering new theories are complementary. AI's ability to dramatically accelerate search speed creates conditions for a paradigm shift. Additionally, Ludwig & Mullainathan (2024) argue that generative AI enhances the generation of novel representations, extending AI's influence beyond increasing search speed and further validating our model's predictions.

We contribute to the growth literature (Romer, 1990; Aghion & Howitt, 1992) by exploring the micro-foundations of knowledge as an asset and its implications for growth.

Building on Weitzman's (1979 & 1998) work on knowledge formation and recombination, our model addresses the broader concept of "theories" and the interplay between search within and across them, particularly how search speed influences this dynamic. This approach aligns with Ludwig & Mullainathan's (2024) view of theory generation as a collaborative process involving humans and AI. Our focus on search speed and its impact on discovery mechanisms illuminates a critical and previously unexplored dimension of the innovation process and its connection to economic growth.

Similar to Acemoglu & Restrepo (2018 & 2020), , we examine knowledge as an asset. However, while they focus on AI and task replacement, we emphasize the interplay between established ("old") and emerging ("new") theories driving innovation and growth. This interplay extends beyond AI's role in task creation.

Arrow (1962) noted that spillovers hinder private investment in basic research. Our framework supports this in the context of slow search speeds, where established theories retain long-term value. However, at high speeds, theory exhaustion necessitates new discoveries for competitive advantage, making basic research an endogenous driver for envisioning new frameworks. This aligns with high-tech firms' significant investments in basic research today. Our work expands on classic justifications for basic research investment (Nelson, 1959; Rosenberg, 1990). We show that at high search speeds, even non-diversified firms benefit from basic research, extending beyond internalizing spillovers (Nelson, 1959) or accessing external knowledge (Rosenberg, 1990). We identify broader conditions under which firms invest in basic research, aligning with Arora et al.'s (2021a) findings of declining basic research investment during a period of slow search speeds (1980-2015). Our model suggests that the rise of AI, potentially leading to faster search speeds, could induce increased investments in basic research, especially by larger firms seeking sustained competitive advantages.

3 Descriptive Evidence

3.1 Firm Size, Research Activities, and Innovation

The interplay between firm size, research activities, exploration strategies, and the ability to exploit technologies is important for understanding innovation dynamics and competitive advantage. Extensive literature reveals that both large and small firms innovate, though larger firms often face diminishing R&D returns, while smaller firms, despite limited resources, often achieve higher patent productivity. The balance between incremental and radical innovation has also been examined, with younger managers and competitive environments fostering more disruptive changes.

This section fills gaps in the literature by exploring the role of search speed, showing that faster speeds boost exploration and growth, particularly in top-performing firms. Additionally, it reveals a non-monotonic relationship between speed and performance, where faster firms excel, and slower firms lag in productivity and R&D intensity.

Before showing our evidence, we cast it into what we currently know about firm size and innovation. The innovation literature has explored extensively the optimal firm size for innovation, revealing that both large and small firms can be highly innovative and invest significantly in basic research (Cohen et al., 1987; Cohen & Klepper, 1996a; Cohen & Klepper, 1996b; Acs & Audretsch; 1988, 1990; Baumol, 2002). However, while R&D investment scales with firm size, R&D productivity often decreases, leading to the "size contingent" theory, which suggests that R&D returns vary by firm size (Mansfield, 1981; Rosen, 1991; Cohen & Klepper, 1996a).

The relationship between firm size and basic research is more complex (Cohen, 2010). Firms often underinvest in basic research due to its broadly shared societal benefits (Nelson, 1959 & 2006). Smaller firms, despite limited resources, often achieve higher patent

productivity relative to their size, indicating diminishing returns for larger firms (Scherer, 1965). Basic research, particularly in the context of General Purpose Technologies (Rosenberg, 1994; Bresnahan & Trajtenberg, 1995), plays a crucial role in boosting productivity and driving economic growth across industries. Larger firms tend to focus on process innovations, leveraging economies of scale (Cohen & Klepper, 1996a) and internal research to prevent competitors from exploiting their discoveries (Arora et al., 2021a).

The choice between incremental and radical innovation is another key factor, with incremental innovations building on existing technologies and radical innovations creating new technology clusters crucial for long-term growth. Firms with open corporate cultures and young managers are more likely to pursue radical innovations, reflecting a culture open to disruptive change (Acemoglu, Akcigit & Celik, 2022). In competitive environments, firms often focus on internal research to avoid benefiting rivals, while regulations can push firms toward more radical, high-risk strategies (Arora et al., 2021a). Additionally, while regulations typically reduce overall innovation by limiting incremental efforts, they can also push firms toward more radical, high-risk strategies (Aghion, Bergeaud & Van Reenen, 2023).

Small and medium-sized enterprises often excel in niche markets, achieving global leadership in specific areas through continuous innovation (Acs & Audretsch, 1990; Noteboom, 1994)). The German Mittelstand and Southern European SMEs in low-tech industries exemplify this success, thriving through incremental innovations despite resource constraints (De Massis, Audretsch, Uhlaner & Kammerlander, 2018). Different entrepreneurship models lead to varying growth rates and firm sizes (Lehmann, Schenkenhofer & Wirsching, 2019), with innovative startups poised for growth differing structurally from "optimally stable" small firms that maintain their size over time (Gimenez-Fernandez, Sandulli & Bogers, 2020).

3.2 Impact of Search Speed

We investigate whether search speed is related to a firm's size and whether search speed is related to the allocation of resources between incremental and radical innovation. We utilize the DISCERN database (Arora et al., 2021b; https://discern-project.com/), which includes comprehensive data on all U.S. public firms with at least one patent from 1980 to 2015. This database links firms by ownership structure, ensuring that patents filed under different names are correctly attributed to their ultimate owners. Additionally, DISCERN integrates patents and publications data with Compustat financial data, providing a robust dataset for our analysis. To our knowledge, this is the most accurate and comprehensive dataset that we could use for our purpose.

Since this dataset provides a long time series at the firm-level, we use it to develop a measure of speed at the firm-level and correlate it with firm-level measures of exploration and the growth of firm assets and ideas. Table 1 provides definitions and descriptions of all variables used in our study.

To measure exploration, we employ two variables. First, we create a measure of patent clusters over five-year windows, aggregating patents from all firms in the database that share similar combinations of CPC patent classifications. The generation of these clusters is detailed in Appendix A3. The *growth of the percentage of patent clusters* refers to the annual increase in the share of patent clusters where a firm holds one or more patents.[1]. Second, we use *the ratio of scientific publications to patents*. Scientific publications serve as a proxy for academic research, which, even when applied, tends to focus more on the broader understanding of problems than on practical technological applications. This ratio

[1]This measure offers several advantages. First, it is more sophisticated than simply matching patent classes, as it captures patents with similar combinations of classifications. Second, it accounts for changes in the breadth of a firm's search over time by comparing the growth of its patent clusters to broader trends and other firms.

Table 1: Description of Variables

	Description
Growth % Patent Clusters	Annual growth of the firm's share of 5-year window patent clusters (Our computations from DISCERN data. See Appendix A3)
Publications/Patents	Ratio of the annual number of firm's publication to patents (DISCERN)
Growth Assets	Growth of firm's total assets (Compustat)
Patents/R&D	Annual number of firm's patents over R&D expenditures (DISCERN & Compustat)
Inverse speed	5-year moving average (lagged one year) of ratio of firm's patents in the same subclasses of the past and patents in the previous year (Our computations from DISCERN data)
Growth Employees (lagged)	Growth of number of firm's employees (lagged one year) (Compustat)
Growth Invested Capital (lagged)	Growth of firm's invested capital (lagged one year) (Compustat)
R&D/Sales (lagged)	Ratio of annual firm's R&D expenditures over sales (lagged one year)

provides a complementary dimension of exploration, reflecting the idea that by engaging with the scientific community, firms can access new ideas, rather than a sole focus on patent applications.

We employ the growth of the firm's total assets and the growth of patents to R&D as measures of, respectively, the growth of firm resources, and the growth of new ideas. We use the firm's total assets as a broad measure of the total resources of the firm. We employ the growth of patents over R&D instead of the growth of patents to avoid that this measure be affected by the growth in the propensity to patent rather than a genuine increase in innovation. Since all our regressions below use firm fixed effects, we can assume that the propensity to increase patents to protect innovations, rather than increases in innovation, is more stable over time and captured in good part by these fixed effects.

We measure *the inverse of search speed* using a five-year moving average, lagged by one year, of the ratio between annual patents, excluding those in CPC subclasses where the firm has never patented, and patents from the previous year. The numerator represents the number of patents, serving as a proxy for the firm's ideas, in the same domain as the previous year. We posit that faster firms generate more patents in a previous year compared to a subsequent year. The five-year moving average allows us to identify firms that consistently generate more patents later than earlier in already established domains, indicating a slower pace. Faster firms, in contrast, would produce the bulk of these patents in earlier years.[2].

Addiitonally, we calculated this variable using patents in the same CPC classes or sections, again finding little variation in the results. However, subclasses fit well our analysis as they represent well-defined technologies and activities, while more aggregate classes are more heterogeneous. For example, class C07 refers to Organic Chemistry, while subclass C07K refers to Peptides; similarly, class G06 is related to Computing, Calculating, and Counting, while subclass G06F pertains specifically to Electric Digital Data Processing. All control variables used in our regressions are detailed in Table 1 and are derived from Compustat data matched at the ultimate owner level through the DISCERN database.[3]

Table 2 provides descriptive statistics for the variables that we use in our analyses.

Table 3. shows that *inverse speed* is negatively correlated (and significantly so) with Growth % Patent Clusters, Publications/Patents, Growth Assets, and Patents/R&D. Thus, speed increases exploration, both growth of share of patent clusters and publications over patents, growth of assets and the growth of the patent/R&D ratio as a proxy for the growth of ideas.

[2]We also tested three- and four-year moving averages with no significant change in results.
[3]We also tested alternative controls from Compustat, which produced consistent results.

Table 2: Descriptive Statistics

	mean	sd	min	perc 1%	perc 50%	perc 99%	max	N
Growth % Patent Clusters	0.074	0.643	-1.000	-1.000	-0.016	2.195	24.231	12879
Publications/Patents	0.818	2.305	0.000	0.000	0.172	10.696	53.500	13300
Growth Assets	0.096	0.421	-0.989	-0.499	0.047	1.595	14.860	13763
Patents/R&D	0.215	3.382	-1.000	-1.000	-0.038	4.150	253.553	13672
Inverse speed	1.089	0.446	0.057	0.473	1.022	2.377	9.842	13763
Growth Employees (lagged)	0.047	0.928	-1.000	-0.497	0.013	0.882	83.000	13763
Growth Invested Capital (lagged)	0.131	4.659	-65.603	-1.204	0.049	2.688	516.741	13763
R&D/Sales (lagged)	1.960	42.186	0.000	0.001	0.052	12.167	2900.629	13763

Since our regressions in the next section employ different observations depending on the dependent variables, we report descriptive statistics for the number of observations of the dependent variables, and the sample with the largest set of observations (dependent variable = growth of assets) for the independent variables.

Table 3: Regression Results

	Growth % Patent Clusters	Publications/ Patents	Growth Assets	Growth Patents/R&D
Inverse speed	-0.199***	-0.127**	-0.038***	-0.400***
	(0.000)	(0.021)	(0.000)	(0.002)
Growth Employees (lagged)	0.005	0.006	-0.001	-0.022
	(0.272)	(0.291)	(0.691)	(0.403)
Growth Invested Capital (lagged)	0.000	-0.002	-0.007***	-0.002
	(0.579)	(0.359)	(0.000)	(0.561)
R&D/Sales (lagged)	-0.000***	-0.001	0.001	-0.000
	(0.000)	(0.499)	(0.214)	(0.858)
Constant	0.294***	0.958***	0.137***	0.651***
	(0.000)	(0.000)	(0.000)	(0.000)
Firm fixed effects	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes
Clustered errors	firms & SIC	firms & SIC	firms & SIC	firms & SIC
Observations	12,879	13,300	13,763	13,672
R-squared	0.104	0.634	0.202	0.144
log likelihood	-11876	-23294	-6084	-34995

Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

Unbalanced firm-year panel 1986-2015. (We start from 1986 to account for the one-lagged 5-year moving average *Inverse speed*, which uses data down to 1980.) All US public firms with at least one patent in 1986-2015. All regressions include a dummy variable to account for acquisitions.

All regressions include firm- and time-fixed effects (and control for acquisitions). This is important in that the firm fixed effects imply that we estimate within-firm effects – that is, the effects of speed are firm-specific effect that occur at different moments in time. In other words, when firms experience higher speed, they explore more, grow their assets and produce a higher growth of ideas. Interestingly, when we remove the firm fixed-effects, the impact of speed is more noisy, suggesting that the firm fixed effects control for important heterogeneity across firms. Also, our 1986-2015 panel comprises a sufficient number of years to exploit enough variation in the time series to account for within-firm effects.

Intuitively, the relation between speed and firm performance is non-monotonic because a low speed may not be strong enough to reallocate resources for exploration, even if speed increases. We then investigate whether the effect of speed becomes significant when it exceeds a threshold. We created three pairs of dummy variables representing low-speed and high-speed firms. These variables take the value 1 for firms whose average Inverse Speed over the study period falls below or above the 25th, 50th, and 75th percentiles, respectively. For instance, the low-speed-25% / high-speed-25% pair distinguishes between the slowest 75% of firms and the fastest 25%. The higher percentiles (50% and 75%) correspond to a smaller group of slower firms.

We ran the regressions for our four dependent variables differentiating between low and high *Inverse speed* categories. Table 4 presents the results focusing on the key independent variables. For comparison, we include the average impact of Inverse speed estimated in Table 3.

Table 4 suggests that the relationship is non-monotonic. The threshold appears to be around the 75th percentile, where the effect of speed is most pronounced. For all four dependent variables, the impact of *Inverse speed* on the slowest 25% of firms is statistically insignificant, whereas it is robust and significantly higher for the fastest 25%. This

Table 4: Non-Monotonicity of Speed

	Speed Threshold: 25%	Speed Threshold 50%	Speed Threshold 75%
Growth % Patent Clusters	Average (Table 3):	-0.199***	
		(0.000)	
Inverse speed (low)	-0.176***	-0.148***	-0.123
	(0.000)	(0.001)	(0.144)
Inverse speed (high)	-0.274***	-0.244***	-0.219***
	(0.000)	(0.000)	(0.000)
Publications/Patents	Average (Table 3):	-0.127**	
		(0.021)	
Inverse speed (low)	-0.114	0.049	0.193
	(0.108)	(0.446)	(0.366)
Inverse speed (high)	-0.165**	-0.278***	-0.207***
	(0.037)	(0.004)	(0.008)
Growth Assets	Average (Table 3):	-0.038***	
		(0.000)	
Inverse speed (low)	-0.039***	-0.053***	-0.019
	(0.001)	(0.000)	(0.225)
Inverse speed (high)	-0.035*	-0.025*	-0.043***
	(0.060)	(0.053)	(0.001)
Growth Patents/R&D Assets	Average (Table 3):	-0.400***	
		(0.002)	
Inverse speed (low)	-0.404***	-0.265***	-0.269
	(0.000)	(0.001)	(0.104)
Inverse speed (high)	-0.388	-0.512***	-0.432***
	(0.187)	(0.000)	(0.000)

Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

Unbalanced firm-year panel (1986-2015). All US public firms with at least one patent in 1986-2015. *Low speed firms* = firms with minimum *Inverse* γ over 1986-2015 higher than 25th, 50th, or 75th percentile of the average minimum *Inverse* γ across firms in the whole sample. All regressions use the same controls, fixed effects and clusters of Table 3.

suggests that a considerable proportion of slower firms do not benefit from increased speed. However, for faster firms, higher speed correlates with growth in key variables, indicating that speed amplifies growth rates and, consequently, widens performance gaps. Additionally, the ratio of publications to patents suggests that faster firms invest more heavily in scientific research relative to patenting. At lower thresholds (25th and 50th percentiles), the results are noisier, likely because these thresholds include firms above the speed threshold predicted by our model. The average effects estimated in Table 3 consistently fall between the effects observed for slow and fast firms, further supporting the non-monotonicity effect.

Finally, Table 5 reveals that slower firms tend to be smaller, less productive (as indicated by lower Sales/Employees ratios), and have a higher ratio of assets to employees. They are also less R&D-intensive, indicating that they may be less dynamic overall.

4 Model

4.1 Set-Up

We conceptualize the innovation process as a search within state spaces, where firms progressively deplete opportunities. Building on Camuffo et al. (2023), a state space represents a problem's attributes where each X_i, $i = 1, 2, ...n$ comprises a set of attributes with associated probabilities. Decision-makers build theories by establishing links among attributes (causal, logical). Exploring these relationships allows for opportunity discovery, but this process exhibits diminishing returns as readily exploitable opportunities are exhausted. Discontinuous change, like discovering new attributes or spaces, triggers a restart of the search within a new domain, leading to fresh opportunities.

Firms search for new knowledge in a staged approach. Let $t_i \in (0, T)$ be the time allocated

Table 5: Characteristics of slow firms (bottom 25%)

	Dependent variable: Dummy for slowest 25% of firms
Number of Employees	-0.006***
	(0.000)
Sales/Employees	-0.000***
	(0.000)
Assets/Employees	0.000**
	(0.027)
R&D/Sales	-0.000
	(0.182)
Constant	0.776***
	(0.000)
Observations	4,887
R-squared	0.098
log likelihood	-2645

Robust pval in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Cross-section using average values of companies over 1986-2015. All US public firms with at least one patent in 1986-2015. Linear probability model: characteristics of firms in the bottom 25% in terms of speed. Regression includes year and SIC dummies.

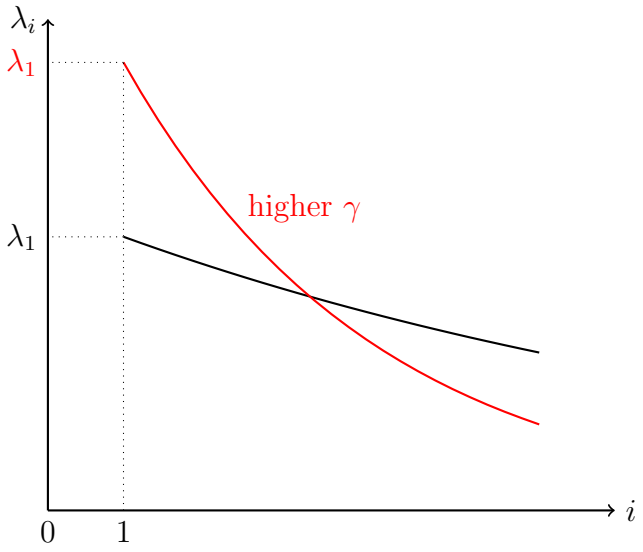
in stage $i = 1, 2, ...\infty$ to search for new ideas in the current theory, and T the time endowment in each stage. This time endowment can differ because firms can use more resources (e.g. more researchers) or more productive researchers (e.g. because of more qualified researchers or better technologies).

During each stage, they invest time t_i to uncover new ideas within the current theory and they dedicate the remaining time $(T - t_i)$ to discover new theories.

Let $\lambda_i t_i^{\frac{1}{2}}$ be the expected arrival rate of ideas in current theory in stage i, with $\lambda_i = \lambda_1 e^{-\gamma(i-1)}$. Also, let $\lambda_1 = \lambda(1 - e^{-\gamma})$ so that $\sum_{i=1}^{\infty} \lambda_i = \lambda$. The parameter λ is the expected total number of ideas of the theory.

The arrival rate of new ideas then depends on a measure of the theory ideas generation potential (λ), the efficiency of search parameter $\gamma > 0$ that represents the rate of decay of ideas across stages in the current theory, and the resources t_i allocated to the search in the current theory. A higher γ tilts the curve λ_i across stages. It raises λ_i for smaller i and lowers it for larger i, as shown by Figure 1.

Figure 1: Efficiency of Search within Theories (γ)



Thus, a more efficient search raises the productivity of search in the earlier stages of the current theory and lowers it in later stages. This depends crucially on the fact that the space defined by a given theory is bounded: if firms discover more ideas earlier (because the technology of search is more efficient), they are bound to discover fewer ideas later because they have already discovered a larger fraction of ideas within the theory.

The firm's dynamic program for knowledge accumulation K is

$$K = \sum_{i=1}^{j-1} e^{-\delta i} \lambda_i T^{\frac{1}{2}} + \sum_{i=j}^{\infty} e^{-\delta i} e^{-\pi(i-j)} \left[\lambda_i t_i^{\frac{1}{2}} + \pi \left(1 + g_i\right) K \right] \tag{1}$$

where δ is a discount factor.

Firms wait $j-1$ stages before exploring a new theory. Till then $t_i = T$. At stage j they invest $t_j \leqslant T$ in the current theory and start exploring a new theory. If they did not find a new theory in stage j, which occurs with probability $e^{-\pi}$, firms invest t_{j+1} in the current theory, and look for a new theory, which they find with probability π. If they find the new theory, they start exploiting it in the next stage, and earn K. In the generic stage $i > j$, they have not found a new theory in the previous $i-j$ stages with probability $e^{-\pi(i-j)}$, exploit the current theory by investing t_i, and find a new theory with probability π.

The term $e^{-\pi(i-j)}$ is the probability of no occurrences in $i-j$ independent Poisson processes. We approximate the probability of at least one occurrence, which is $1-e^{-\pi}$, with π. *De facto* we are assuming that π is small enough that $1-e^{-\pi} \approx \pi$ – that is, in the time span of one stage firms obtain at most one occurrence. This assumption is harmless, and basically amounts to say that the time span of each stage is sufficiently short.

The term g_i is the growth rate of knowledge in stage i, which we approximate by

$$1 + g_i = T - t_i \tag{2}$$

Thus, firms spend their exploration time to identify richer theories (i.e., with higher potential for ideas generation λ). A lower t_i, which raises $T - t_i$ increases $(1 + g_i)\lambda$, which is the expected total number of ideas of the new theory. However, the new theory is feasible only with probability π.

This set-up relies on some simplifying assumptions. We assume that firms cannot find new theories exogenously. If they do not spend time to look for them, that is $T - t_i = 0$, they do not find theories with a positive expected number of ideas. Following Cohen & Levinthal (1989) and Rosenberg (1990), they must make at least some minimal investments in absorptive capacity to understand external scientific knowledge. Also, the probability π of finding new feasible theories does not depend on the search for new theories in previous stages. Similarly, the expected number of ideas of the new theory, which depends on $1 + g_i$, only depends on the effort in the current stage and not on the efforts in the previous stages. These extensions complicate the analysis without offering new insights.

4.2 Solving the Dynamic Program

In Section A1 of the Appendix we solve the dynamic program (1) by choosing the sequence of time allocations t_i, $i \geqslant 1$. The first order condition implies that in stage i firms allocate t_i till the additional unit of t_i allocated to further exploration in the current theory is equal to what firms expect to obtain if they allocate that unit of time to the search of a new theory.

This yields the following optimal t_i for $i > j$, as a function of the initial condition t_j, $j \geqslant 1$

$$t_i^{\frac{1}{2}} = e^{-\gamma(i-j)} t_j^{\frac{1}{2}} \tag{3}$$

To determine the initial condition t_j, we assume that the optimal $t_1 \geqslant T$, which means

that either the optimal $t_1 = T$, and t_1 satisfies first order condition (A6) in Section A1 of the Appendix, or the optimal $t_1 > T$. Since T is a resource constraint, t_1 is bounded and equal to T. The optimal t_i in (3) implies that, for $i > 1$, t_i declines with i. Therefore, $\exists t_j$, $j \geqslant 1$, such that $t_j \geqslant T$ and $t_{j+1} < T$. As shown in Section A1 of the Appendix, we then solve the first order condition for t_j by approximating the optimal $t_j = T$. This amounts to solving for the optimal j. In turn, this implies that for $i \leqslant j$, $t_i = T$, and for $i > j$, t_i is determined by (3).[4]

We obtain the following first order condition

$$\frac{e^{-\gamma(j-1)}}{2\pi T} - \frac{e^{-\delta}}{1 - e^{-(\delta+\gamma)}}(1 + \Omega) = 0 \tag{4}$$

where

$$\Omega \equiv \frac{e^{-(\delta+\gamma)(j-1)}}{2}\left(1 - e^{-(\delta+\gamma)}\right)(F - G) \tag{5a}$$

$$F \equiv \frac{1}{1 - e^{-(\delta+\pi)}} + \frac{1}{1 - e^{-(\delta+2\gamma+\pi)}} \tag{5b}$$

$$G \equiv \frac{2}{1 - e^{-(\delta+\gamma)}} \tag{5c}$$

Lemma 1. *The optimal j declines with π and T*

Proof. See Appendix. ∎

Section A1 of the Appendix shows that the optimal K is

$$K = \frac{\lambda_1 e^{-\gamma(j-1)}}{2\pi T^{\frac{1}{2}}} \tag{6}$$

[4]In our model, the allocation of T between the current and new theory, the assumption that $t_1 \geqslant T$, and the decline of the optimal t_i with i, imply that the optimal j is determined by the fact that t_j falls below T. If we could only allocate T alternatively to the current or new theory, the optimal j would be determined by equating the marginal value of allocating T to one or the other theory.

By replacing $\frac{e^{-\gamma(j-1)}}{2\pi T} = \frac{e^{-\delta}}{1-e^{-(\delta+\gamma)}}(1+\Omega)$ from (4) obtain

$$K = \tilde{K}(1+\Omega) \tag{7}$$

where $\tilde{K} \equiv \frac{\lambda_1 T^{\frac{1}{2}} e^{-\delta}}{1-e^{-(\delta+\gamma)}}$ is the optimal K in the benchmark case in which there is no search for new theories ($\pi = 0$ and $j \to \infty$).

5 Implications

5.1 Search speed and discovery rate

Search speed γ and discovery rate π depend on general conditions of the economy and idiosyncratic characteristics of firms. We posit that there is a distribution of values γ or π across firms, and common factors that raise or lower the values of γ or π of all firms. For instance all firms benefit from new opportunities that raise speed (like AI) or by scientific or technological advances that open new domains. We assume for simplicity that the idiosyncratic components of γ and π are "endowments" of the firms. Firms can be faster or slower or more or less able to or interested in exploring new domains. However, firms choose t_i. Thus, the allocation of resources between current and new theories is the strategic choice that firms make given these parameters. A lower t_i reflects a higher $T - t_i$ – that is, firms invest to a greater extent in the growth of new theories.

Theorem 1. *The discovery of new theories is limited by the speed of search:* $\exists \gamma' > 0 : \gamma < \gamma' \implies \pi > 0$ *is higher than optimal.*

Proof. See Appendix. ■

This theorem states that given the endowment π of the firm, there is a threshold γ' such

that for smaller γ, the endowment of the firm is suboptimal. In other words, the firm has a higher ability to discover new theories than it is ideal. The intuition is that, for a low search speed γ, there are enough discoveries to be made in the current theory in later stages, which lowers the value of a higher ability to discover new theories. It is natural to assume that firms endowed with too high a π either lower it, or, if not, they produce suboptimal knowledge because they switch "too early" to new theories leaving ideas from the old theory "on the table". Conversely, a sufficiently high search speed makes any given ability π fully valuable. Early switch are beneficial because the old theory exhausts rapidly and a greater ability to discover a new theory pays off. Firms reallocate resources t_i away from the current theory to the new theory, and j declines. Theorem 1 also provides an explanation for Christensen's (1997) paradox whereby incumbent firms overlook new opportunities. It shows that this behavior is driven by a low search speed.

Corollary 1. *Economy-wide shocks that increase the speed of search γ raise the share of firms that invest in the discovery of new theories, and widen the knowledge gaps among firms.*

Proof. See Appendix. ■

This corollary develops the implications of Theorem 1 further. If γ increases for all firms, due to economy-wide shocks, some firms with $\gamma < \gamma'$ will now feature $\gamma \geqslant \gamma'$, but not the other way around. Thus, before the increase in γ, the π of these firms was higher than optimal, and they either worked with a lower π, or they produced less knowledge because they put too much effort $T - t_i$ on the discovery of new theories. Now they work with π equal or higher than optimal, making the focus on the search for new theories (and the lower allocation of resources to t_i) more productive. Since no firm with $\gamma \geqslant \gamma'$ before the shock now switches to $\gamma < \gamma'$, there will only be firms that raise the productivity of higher π. Moreover, Theorem 1 shows that γ and π are complementary, which implies

that the increase in γ raises the productivity gap between firms with different π.

5.2 Resources T and Growth of Ideas

In this section we show that the endogenous choice of T reinforces the effects discussed in the previous section. Higher γ and π raise T, and the combined effect of these three variables raises the growth of new ideas. As discussed in the previous section, low speed implies that firms with low or high π do not differ systematically because low speed dampens the potential of higher π. In this section we show that high speed raises the value of π, and they both raise T and a higher growth of ideas. This widens gaps across firms further.

To keep matters simple, let $V(K,T)$ be the ultimate value that the firm looks at, where K is our knowledge asset that depends on T, and the argument T represents an alternative allocation of resources such that a higher T increases K and V indirectly; the direct effect of T is negative because it represents an alternative allocations that generate value for the firm. We assume that V is concave in T, and thus there is an optimal allocation of T.

One interpretation of the direct effect of T is that not only does value depends on knowledge but also on the implementation of knowledge to generate products, services or other goods that agents use. To the extent that resources T are rivalrous, this generates a trade-off between production of knowledge and implementation. Another interpretation is that T can be allocated alternatively to activities other than the growth of knowledge of the firm. A classical example is "leisure" or more generally reallocation of the firms' proceeds to shareholders in the form of dividends rather than investment. In what follows, we keep these interpretations open. They all have similar implications for our main points.

Theorem 2. *Increases in speed γ raise investments T and the growth rate of new ideas*

of firms with higher π.

Proof. See Appendix. ■

This theorem aligns with Bloom et al. (2020), who note that new ideas are harder to find because they require an increasingly large scale of resources. In our model, a low search speed prevents resources (T) from accumulating fast enough to discover new theories, suggesting that it is an antecedent to lack of resources. New technologies that increase search speed can enable firms to discover new ideas at a higher rate.

Also, in our model, only if $\gamma > \gamma'$, the endowment π of a firm is not "too high". Thus, a low speed does not trigger the mechanisms of Theorem 2. In this case, firms lower π and increase the allocation of resources t_i on the current theory. Conversely, if $\gamma > \gamma'$, they respond to the higher γ by fully exploiting their higher ability to discover new theories, lowering allocations to t_i and increasing T, so that $T - t_i$ increases, and the endogenous T raises further the richness of the newly explored domains and growth of new ideas.

6 Discussion

Our model offers various insights into key issues discussed in the literature and the current dynamics of firms, industries, and technologies.

First, we provide an explanation for Christensen's (1997) paradox, which posits that established firms often miss new opportunities. Our model shows that in a slow search environment, it is optimal for firms to focus on current theories and postpone the search for new ones. Conversely, in a fast search environment, Christensen's conditions do not apply, and limited investment in new theories is sub-optimal.

Since at least Arrow (1962), it has been understood that firms are unlikely to invest

in basic research due to its lack of immediate application and the spillovers it generates. Arrow concluded that basic research must be publicly funded and considered the presence of individuals and institutions that enjoy conducting research and combining it with teaching a "lucky accident." Nelson (1959) argued that large, diversified firms primarily conduct basic research because they can internalize its broad spillovers. Rosenberg (1990) added that firms invest in basic research with their own funds to develop the absorptive capacity needed to understand public scientific research.

Our model makes the novel prediction that in a high search speed environment, firms will find it optimal to invest in basic research to discover new theories, a possibility not found in the existing literature. Although the model does not account for spillovers, which could reduce firms' incentives to invest in basic research, the influence of high search speed is likely to counteract this effect. Firms may focus on generating intellectual property from their basic research, potentially hindering their engagement with academia, as noted by Rosenberg.

This prediction aligns with Yue (2023), the first large-scale analysis showing a positive effect of corporate involvement on scientific progress. Focusing on AI researchers affiliated with both universities and firms, Yue (2023) reveals a significant increase in citations for papers with corporate involvement, suggesting this leads to more groundbreaking research, albeit aligned with corporate interests. In contrast to Arora et al. (2021a) , who noted a decline in U.S. firms' basic research investments from 1980-2015 due to low-speed sectors, Yue (2023) highlights the growing importance of basic research in the high-speed AI sector.

Third, our discussion aligns with Ludwig & Mullainathan (2024). They show in the context of judicial decisions that hypothesis generation, traditionally reliant on human intuition, can be systematically enhanced using AI. In our model innovations in AI such as

Large Language Models can accelerate the search process by enhancing the productivity of generating new theories.

Fourth, the ability to discover new theories can be a significant source of comparative advantage in a high-speed environment, as a firm's value (V) increases directly with the probability (π) of discovering new theories.[5] Firms investing more in basic research or enhancing their ability to increase π will enjoy significant competitive advantages. In equilibrium, firms with higher π create higher value (V), which in turn increases the resources (T) they reinvest in research. This creates a virtuous cycle explaining the high growth potential of firms investing in new theories and the hyper-growth of many high-tech firms today, especially compared to other sectors.

External shocks that increase search speed can trigger these virtuous cycles, where faster search speeds lead to higher discovery rates, which increase reinvestment resources, thereby widening gaps among firms, industries, or even countries. As shown by Corollary 1 to Theorem 1, a higher initial stock of ideas (λ) or a lower opportunity cost of capital (T) can widen these gaps. To reduce disparities or leverage new growth cycles, external interventions may be needed to create the right conditions, triggering investments in high search speed (γ), higher discovery rates (π), and greater resource accumulation (T) for reinvestment. Conversely, in a low-speed environment, the value of the firm does not benefit from a high probability (π) of discovering new theories, making differences in discovery ability irrelevant.

[5] Conversely, in a low-speed environment, the value of the firm does not benefit from a high probability π of discovering new theories, making differences in discovery theory irrelevant.

7 Conclusions

In this paper, we introduce a novel framework for understanding innovation by focusing on search speed—how quickly firms explore and exhaust ideas within a technological paradigm. This concept underscores the balance firms must strike between exploiting existing knowledge and exploring new paradigms, a crucial but often underexplored aspect of innovation. Through a multi-stage search model, we shed light on how firms optimally allocate resources between searching within an existing technological framework and seeking out new theoretical paradigms. This approach provides insights into the strategic trade-offs firms face in driving innovation forward.

A key finding is the identification of a critical threshold effect. When the efficiency of searching within the existing paradigm is low (low γ), firms may become trapped in a cycle of focusing on the existing paradigm. This is because the weak diminishing returns associated with the search within the existing paradigm make investment in exploring new theoretical frameworks unattractive, even with a high probability of discovering new paradigms (high π).

However, the model also shows that the optimal search strategy shifts dramatically when the speed of search increases beyond a threshold. In such scenarios, firms can effectively explore both avenues. A higher search speed within the existing paradigm (high γ) motivates reallocating resources from later stages of the search within the existing technological paradigm to earlier stages of exploration of new theories with greater innovation potential. The greater availability of resources at later stages then opens up opportunities to explore new theories. The probability of discovering new paradigms (π) becomes a critical success factor.

A higher γ can then lead to a "bifurcated world" in which the complementarity between

new theories frontier π and continuous advance in search frontier γ create exponential differential performance.

The model links future search resources to the value created by past innovations highlighting a critical interplay between search efficiency and the potential value of new theoretical paradigms. When search within the existing paradigm is slow, the ability to discover new theories has minimal impact on firm performance. However, when search is fast, differences in the ability to discover new theories become a major source of performance variation between firms. This underscores the importance of not only efficient search within the existing paradigm but also the ability to explore entirely new theoretical frameworks for sustained innovation success. Furthermore, speed of adoption and access to frontier technology in search γ or new theories π can become important geopolitical strategic variables (cognitive supply chain dependency) and access to the frontier can in itself become a key geopolitical decision (for example open sourcing γ but keeping proprietary π).

Furthermore, by emphasizing the critical roles of search speed and theory generation potential, the model provides insights into how policymakers can foster innovation across industries. Policies that enhance search speeds among diverse firms and industries can promote equitable growth and reduce performance gaps. For example, reducing capital costs allows firms to direct more resources toward exploring new knowledge domains. Additionally, policies that incentivize basic research and support knowledge-sharing institutions can significantly enhance the generation of new theories, thereby enriching the pool of ideas available to firms.

There are several limitations to the model. Our current measure of search speed, based on firm-level patent data, may not fully capture its complexity. Future research should incorporate more flexible, time-varying parameters to better reflect the dynamic nature of innovation cycles across industries. Additionally, the model does not account for competi-

tion and spillover effects, which significantly influence innovation strategies. By integrating these factors, we can provide a more accurate representation of how firms navigate external pressures and shared knowledge environments. Finally, extending the model to include inter-firm interactions would offer deeper insights into the broader innovation ecosystem, revealing how competition and collaboration shape the speed and direction of innovation.

References

Acemoglu, D., & Restrepo, P. (2018). The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6), 1488-1542.

Acemoglu, D., & Restrepo, P. (2020). Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Econometrica*, 88(3), 1305-1348.

Acemoglu, D., Akcigit, U., & Celik, M. A. (2022). Radical and Incremental Innovation: The Roles of Firms, Managers, and Innovators. *American Economic Journal: Macroeconomics*, 14(3), 199-249.

Acs Z.J., Audretsch, & D.B. (1988) Innovation in Large and Small Firms: An Empirical Analysis. *American Economic Review*, 78(4), 678–690.

Acs Z.J., Audretsch, D.B. (1990) *Innovation and Small Firms*. MIT Press, Cambridge, MA.

Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2), 323-351.

Aghion, P., Bergeaud, A., & Van Reenen, J. (2023). The Impact of Regulation on Innovation. *American Economic Review*, 113(11), 2894-2936.

Agrawal, A., Gans, J., & Goldfarb, A. (2021). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, Chicago.

Agrawal, A., McHale, J., & Oettl, A. (2024). Artificial Intelligence and Scientific Discovery: A Model of Prioritized Search. *Research Policy*, 53(5), 104989.

Arora, A., Belenzon, S., & Sheer, L. (2021a). Knowledge Spillovers and Corporate Investment in Scientific Research. *American Economic Review*, 111(1), 871-898.

Arora, A., Belenzon, S., & Sheer, L. (2021b). Matching Patents to Compustat Firms, 1980–2015: Dynamic Reassignment, Name Changes, and Ownership Structures. *Research Policy*, 50(5), 104271.

Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, a conference of the Universities-National Bureau Committee for Economic Research and the Social Science Research Council, 609-626. Princeton University Press, Princeton, NJ.

Baumol, W.J. (2002) *The Free-Market Innovation Machine. Analyzing the Growth Miracle of Capitalism.* Princeton university press, Princeton, NJ).

Bloom, N., Jones, C., Van Reenen, J. & Webb, M. (2020). Are Ideas Getting Harder to Find?. *American Economic Review*, 110(4), 1104-44.

Bresnahan, T. F., & Trajtenberg, M. (1995). General Purpose Technologies: 'Engines of Growth'? *Journal of Econometrics*, 65(1), 83-108.

Camuffo, A., Gambardella, A. & Pignataro, A. (2023). Theory-Driven Strategic Management Decisions. *CEPR DP 17664v4.*

Christensen, C. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Harvard Business School Press, Boston MA

Cohen, W. M., Levin, R. C., & Mowery, D. C. (1987). Firm Size and R&D Intensity: A Re-examination. *Journal of Industrial Economics*, 35(4), 543-565.

Cohen W.M., & Klepper, S. (1996a) A Reprise of Size and R & D. *Economic Journal*, 106(437), 925–951.

Cohen, W. M., & Klepper, S. (1996b). Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D. *The Review of Economics and Statistics*, 78(2), 232-243.

Cohen, W.M. (2010) Fifty Years of Empirical Studies of Innovative Activity and Performance. Hall B. & Rosenberg N. (eds.) *Handbook of the Economics of Innovation*, Elsevier, Amsterdam, 129–213.

Cockburn, I.M., Henderson, R. & Stern, S. (2019). The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis. In Agrawal, A., Gans, J., & Goldfarb, A. (eds.). *The Economics of Artificial Intelligence: An Agenda* University of Chicago Press, Chicago.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R&D. *The Economic Journal*, 99(397), 569-596.

De Massis, A., Audretsch, D., Uhlaner, L., & Kammerlander, N. (2018). Innovation

with Limited Resources: Management Lessons from the German Mittelstand. *Journal of Product Innovation Management*, 35(1), 125-146.

Dosi, G. (1982). Technological Paradigms and Technological Trajectories: A Suggested Interpretation of the Determinants and Directions of Technical Change. *Research Policy*, 11(3), 147-162.

Furman, J., & Seamans, R. (2019). AI and the Economy. *Innovation Policy and the Economy*, 19, 161-191. National Bureau of Economic Research, Cambridge.

Gimenez-Fernandez, E. M., Sandulli, F. D., & Bogers, M. (2020). Unpacking Liabilities of Newness and Smallness in Innovative Start-ups: Investigating the Differences in Innovation Performance between New and Older Small Firms. *Research Policy*, 49(10), 104049.

Hall, B.H. & Rosenberg, N. (2010). *Handbook of the Economics of Innovation*. Elsevier, Amsterdam.

Hall, B.H. & Helmers, C. (2024). *The Economics of Innovation and Intellectual Property*. Oxford University Press, Oxford.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago IL.

Lehmann, E. E., Schenkenhofer, J., & Wirsching, K. (2019). Hidden Champions and Unicorns: A Question of the Context of Human Capital Investment. *Small Business Economics*, *52*, 359-374.

Levinthal, D. A., & March, J. G. (1993). The Myopia of Learning. *Strategic Management Journal*, 14(S2), 95–112.

Ludwig, J. & Mullainathan, S. (2024). Machine Learning as a Tool for Hypothesis Generation. *Quarterly Journal of Economics* 139(2), 751-827

Mansfield, E. (1981) Composition of R and D Expenditures: Relationship to Size of Firm, Concentration, and Innovative Output. *Review of Economic Statistics*, 63(4), 610–615.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. J. (2024). AI Adoption in America: Who, What, and Where. *Journal of Economics and Management Strategy*, 33(2), 375-415.

Nelson, R. R. (1959). The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, 67(3), 297-306.

Nelson, R. R. (2006). Reflections on "The simple economics of basic scientific research": Looking back and looking forward. *Industrial and Corporate Change*, 15(6), 903-917.

Nooteboom, B. (1994). Innovation and Diffusion in Small Firms: Theory and Evidence. *Small Business Economics*, *6*, 327-347.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), S71-S102.

Rosen, R. J. (1991). Research and development with Asymmetric Firm Sizes. *The RAND Journal of Economics*, 22(3), 411-429.

Rosenberg, N. (1990). Why Do Firms Do Basic Research (With Their Own Money)? *Research Policy*, 19(2), 165-174.

Rosenberg, N. (1992). Scientific Instrumentation and University Research. *Research Policy*, 21(4), 381-390.

Rosenberg, N. (1994). *Exploring the Black Box: Technology, Economics, and History.* Cambridge University Press, Cambridge.

Scherer, F. M. (1965). Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. *American Economic Review*, 55(5), 1097-1125.

Trajtenberg, M. (2019). Artificial Intelligence as the Next GPT: A Political-Economy Perspective. In Agrawal, A., Gans, J., & Goldfarb (eds.). *The Economics of Artificial Intelligence: An Agenda* National Bureau of Economic Research, Cambridge.

Weitzman, M. L. (1979). Optimal Search for the Best Alternative. *Econometrica*, 47(3), 641-654.

Weitzman, M. L. (1998). Recombinant Growth. *Quarterly Journal of Economics*, 113(2), 331-360.

Yue, D. (2023). I, Google: Estimating the Impact of Corporate Involvment in AI Research. *working paper*, Harvard Business School.

Appendix

A1. Solving the Dynamic Program

In order to solve the dynamic program (1), first solve it for K to obtain

$$K = \frac{k}{1-\Theta} \tag{A1}$$

where

$$k \equiv e^{-\delta}\lambda_1 T^{\frac{1}{2}} \sum_{i=1}^{j-1} e^{-(\delta+\gamma)(i-1)} + e^{-\delta}e^{\pi(j-1)}\lambda_1 \sum_{i=j}^{\infty} e^{-(\delta+\gamma+\pi)(i-1)} t_i^{\frac{1}{2}} \tag{A2a}$$

$$\Theta \equiv e^{-\delta}\pi e^{\pi(j-1)} \sum_{i=j}^{\infty} e^{-(\delta+\pi)(i-1)} \left(T - t_i\right) \tag{A2b}$$

Using (A1), (A2a), and (A2b), firms choose t_i to $\max_{t_i} K$, which yields the *foc*

$$\lambda_1 e^{-\gamma(i-1)} \frac{1}{2} t_i^{-\frac{1}{2}} \left(1 - \Theta\right) - k\pi = 0 \tag{A3}$$

The ratios between each term of (A3) with the equivalent terms of the *foc* of t_j yields foc (3) in the text, that is

$$t_i^{\frac{1}{2}} = e^{-\gamma(i-j)} t_j^{\frac{1}{2}} \tag{A4}$$

Replace t_i in (A2a) and (A2b). Then, k and Θ will depend only on t_j, that is

$$k \equiv \frac{e^{-\delta}\lambda_1 T^{\frac{1}{2}} \left(1 - e^{-(\delta+\gamma)(j-1)}\right)}{1 - e^{-(\delta+\gamma)}} + \frac{\lambda_1 e^{-(\delta+\gamma)j} e^{\gamma} t_j^{\frac{1}{2}}}{1 - e^{-(\delta+2\gamma+\pi)}} \tag{A5a}$$

$$\Theta \equiv e^{-\delta j}\pi \left[\frac{T}{1 - e^{-(\delta+\pi)}} + \frac{t_j}{1 - e^{-(\delta+2\gamma+\pi)}}\right] \tag{A5b}$$

The *foc* for t_j is (A3) with t_j in lieu of t_i, which yields

$$\lambda_1 e^{-\gamma(j-1)} \frac{1}{2} t_j^{-\frac{1}{2}} \left(1 - \Theta\right) - k\pi = 0 \tag{A6}$$

where k and Θ are now (A5a) and (A5b), which are only functions of t_j.

After replacing k and Θ from (A5a) and (A5b), the *foc* (A6) for t_j becomes

$$\begin{aligned} &\lambda_1 e^{-\gamma(j-1)} \frac{1}{2}\, t_j^{-\frac{1}{2}} \left(1 - \frac{e^{-\delta j}\pi T}{1 - e^{-(\delta+\pi)}} + \frac{e^{-\delta j}\pi t_j}{1 - e^{-(\delta+2\gamma+\pi)}}\right) \\ &- \lambda_1 e^{-\delta}\pi T^{\frac{1}{2}} \frac{1 - e^{-(\delta+\gamma)(j-1)}}{1 - e^{-(\delta+\gamma)}} - \lambda_1 e^{\gamma}\pi \frac{e^{-(\delta+\gamma)j} t_j^{\frac{1}{2}}}{1 - e^{-(\delta+2\gamma+\pi)}} = 0 \end{aligned} \tag{A7}$$

As discussed in the text, replace $t_j = T$, multiply all terms by $t_j^{-\frac{1}{2}} e^{-\gamma}$, and rearrange terms to obtain (4) in the text.

To obtain the optimal knowledge K, use (A5a) and (A5b) to obtain (A1) from (A3). Replacing $t_j = T$, obtain (6) and (7) in the text.

A2. Proofs

Proof of Lemma 1. It is easy to see that when $j \to \infty$, the left-hand-side of (4) is negative. The assumption $t_1 \geqslant T$ implies that for $j = 1$, this expression is positive. These two conditions ensure that there is at least one finite j such that the *foc* (4) is satisfied. To show the second condition, derive the unconstrained optimal expression for t_1 by optimizing the initial condition t_j, with $j = 1$, using only the second summation sign in (1). The *foc* for t_1 becomes (A7) without the terms involving T. Deriving the optimal t_1, and setting $t_1 \geqslant T$ produces the condition $e^{-\delta} \pi T F \leqslant 1$, which is the same condition that makes the left-hand-side of (4) evaluated at $j = 1$ positive.

We also show that we can make the fair assumption that the left-hand-side of (4) declines monotonically with j. The sign of this differential with respect to j is the same as the sign of $-\dfrac{\gamma}{\delta + \gamma} + e^{-\delta j} \pi T (F - G)$. Using (5b) and (5c), $\gamma = 0, \pi > 0 \implies F - G < 0$, which makes this expression negative. As γ increases, $F - G$ increases and eventually becomes positive. However, $\pi (F - G)$ increases with π, and therefore for π sufficiently small, the sign of the left-hand-side of (4) negative even for $\gamma > 0$. The sign of $\dfrac{\partial(\pi F)}{\partial \pi} > 0$ depends on the sign of the sum of two terms $1 - e^{-(a+\pi)}(1 + \pi)$, with a equal to δ or $\delta + 2\gamma$. These expressions increase with π, and when $\pi = 0$ they are positive. This also makes it more likely that a small π satisfies the condition above for $t_1 \geqslant T$.

To check the sign of the differential of the left-hand-side of (4) with respect to π, multiply

both terms of this expression by π. The sign of this differential is the opposite of the sign of $\dfrac{\partial(\pi F)}{\partial \pi}$. Thus, the optimal j declines with π. The differential of the left-hand-side of (4) with respect to T is clearly negative. The optimal j decreases with T. ■

Proof of Theorem 1. Differentiate (5a) with respect to π. Using subscripts to denote derivatives, obtain

$$\Omega_{\pi} = -\left(\delta + \gamma\right) j_{\pi} \left(F - G\right) + F_{\pi} = 0 \tag{A8}$$

where, according to Lemma 1, $j_{\pi} < 0$. Set $\pi > 0$, and assume that second order effects of j_{π} are negligible. If $\gamma = 0$, from (5b) and (5c), it is not difficult to see that $F - G < 0$ and $F_{\pi} < 0$, which implies $\Omega_{\pi} < 0$. If γ increases, $\delta + \gamma$, $F - G$, and F_{π} increase, and for $\gamma \to \infty$, $F - G$ and F_{π} are finite, which ensures that for γ sufficiently large, $\Omega_{\pi} > 0$.■

Proof of Corollary 1. Theorem 1 showed that $\Omega_{\pi\gamma} > 0$. Thus, an increase in γ of all firms can switch Ω_{π} from negative to positive, but not the other way around. The endowment π of each firm can then switch from higher to equal or lower than optimal, but not vice versa. Since $\Omega_{\pi\gamma} > 0$, the higher γ widens the gap among firms with different π. ■

Proof of Theorem 2. Using subscripts to denote derivatives, the *foc* of $V(K,T)$ with respect to T is $V_K K_T - V_T = 0$. Using (7), and the expression for $\tilde{K}$, obtain $K_T = \tilde{K}_T\left(1 + \Omega\right) + \tilde{K}\Omega_T$. From Lemma 1, $j_T < 0 \implies \Omega_T > 0$, and it is easy to see that $\tilde{K}_T > 0$. If second order effects of j_T are negligible, $K_{T\gamma} = \tilde{K}_{T\gamma}\left(1 + \Omega\right) + \tilde{K}_T\left(1 + \Omega_{\gamma}\right) + \tilde{K}_{\gamma}\Omega_T$. This expression is positive because it is easy to see that $\tilde{K}_{T\gamma}, \tilde{K}_T, \tilde{K}_{\gamma} > 0$, and it is fair to assume $\Omega_{\gamma} > -1$. If second order effects of Ω_T are negligible, $K_{T\gamma\pi} = \tilde{K}_{T\gamma}\left(1 + \Omega_{\pi}\right) + \tilde{K}_T\left(1 + \Omega_{\gamma\pi}\right)$. From Theorem 1, $\Omega_{\gamma\pi} > 0$, and $\gamma > \gamma' \implies \Omega_{\pi} > 0$. Thus, when $\gamma > \gamma'$, increases in γ raise the productivity of resources T of firms with higher π. When $\gamma < \gamma'$, then the firm's

π is higher than optimal. These firms lower π which dampens the complementarity.

The growth rate of new ideas depends on whether, in (1), the firm is at the stage $i \leqslant j$ or $i > j$. In the former case, this growth rate is $\dfrac{\lambda_{i+1} t_{i+1}^{\frac{1}{2}}}{\lambda_i t_i^{\frac{1}{2}}}$. Using the expressions for λ_i and setting $t_i = T$, this yields $e^{-2\gamma}$. In the latter case, the growth rate of ideas is

$$\frac{e^{-\pi} \lambda_{i+1} t_{i+1}^{\frac{1}{2}} + \pi \left(T - t_i\right) \lambda_1}{\lambda_i t_i^{\frac{1}{2}}}$$

Replacing the optimal t_i and $t_j = T$, obtain

$$e^{-\pi - 2\gamma} + \frac{\pi \left[1 - e^{-2\gamma(i-j)}\right] T^{\frac{1}{2}}}{e^{-\gamma(2i-j-1)}} \tag{A9}$$

Because, from Lemma 1, $j_\pi < 0$, increases in π raise the second term of (A9). Neglecting second order effects, higher T increases the positive effect of higher π on this term, making it more likely that this expression is positive in spite of the negative effect of π on the first term.

The overall growth rate of ideas is a weighted average between $e^{-2\gamma}$ and (A9), where the weights account for the frequency with which, at any moment in time, the firm is on the path $i \leqslant j$ or $i > j$. If (A9) is larger than $e^{-2\gamma}$, which is more likely when γ and π, and therefore T, are high, high γ, π, and therefore T, raise the growth rate of new ideas. When $\gamma < \gamma'$, π is higher than optimal. The complementarity does not arise, dampening the growth of new ideas. ■

A3. Construction of the Patent Clusters

Selection of Machine Learning Algorithm for Patent Clustering

The goal of the algorithm is to cluster patents issued in the US between 1980 and 2015

based on the CPC classification. Patents can generally be classified in one or multiple CPC categories, which leads to specific combinations of CPC classifications. Patents issued with the same or similar combinations are assumed to belong to the same field of research. When a company issues a patent with an identical or similar combination of CPC classifications, it can be assumed that the company is continuing the research in the same field. When a company issues a patent with a new combination of CPC classifications, it can be assumed that the company is exploring a new domain.

We approached this classification problem by comparing and analyzing different clustering machine learning algorithms on a dataset that only contains the patents' application IDs and the CPC classification up to the subclass level. The CPC classifications are codes for defined fields of patents and represent categorical data. The data is pre-processed by encoding the categorical CPC classifications for the use of machine learning algorithms. A CPC classification will be assigned to the respective CPC section, CPC class, and CPC subclass. Thus, patents having multiple classifications can have positive values in multiple columns. For instance, a patent having two CPC classifications A08A and A08B, would lead to a count of 2 in the A CPC section, 2 in the A08 CPC class, and 1 in each CPC subclass. This assignment was chosen based on preliminary tests with subsets of the data and assigning a CPC classification not only to one subclass but to different levels. The results led to more stringent clustering with more patents in the miscellaneous cluster and more clusters in comparison to the clustering with the CPC subclass only. The binary encoding of CPC section, class, and subclass resulted in even more miscellaneous patents and approximately the same number of clusters. Verifying the clusters manually, the non-binary encoded data including CPC section, CPC class, and CPC subclass resulted in the most consistent clusters.

Aiming to cluster patents based on the combination of CPC classifications, the number of

clusters should be determined by the patent combinations themselves and should not be manually imposed. Moreover, the number of clusters should be treated as a result rather than as input. There are different ways to determine the number of clusters, such as the elbow method or the silhouette method. For subsamples of the patent data, the elbow method resulted in a deficient number of clusters, which is too coarse for the analysis. The silhouette method resulted in suggested parameters that yield too many and too small clusters. For the purpose of our analysis, the clusters should not be represented by too few patents since it would not capture domains but the intricacies of the patent combinations. Consequently, clustering algorithms requiring the number of clusters as input, such as K-Means, are subsequently not further considered.

Additionally, the clustering algorithm has to be suitable for categorical data which is in our case encoded in a binary way. Various algorithms were tested on a data subset but the clusters were in most cases either combined in less than five clusters or more than 1000 clusters with the majority of clusters containing only one patent. For the purpose of our analysis, we would like to have meaningful clusters containing multiple similar patents, with no cluster containing a considerable fraction of the overall data set. We decided to test the following three clustering algorithms in more detail: DBSCAN, HDBSCAN, and Agglomerative Clustering.

First, DBSCAN evaluates the proximity of points and classifies points with a sufficient number of other points within its perimeter as core points and points with an insufficient number of points within its perimeter as noise. A cluster is created by iteratively grouping core points that are density-connected and treats points out of the proximity of core points as noise (Ester, Kriegel, Sander & Xu, 1996). The advantages of DBSCAN are that clusters can have arbitrary shapes and sizes, the number of clusters is automatically determined, and the algorithm is robust to outliers and noise. The relevant disadvantage

for this dataset is that the algorithm may experience issues with clusters of similar density, and its scalability is limited for high dimensional datasets (Lytvynenko, Lurie, Krejci, Voronenko, Savina & Taif, 2019), which is the case with the categorical patent data.

Secondly, HDBSCAN is an extension of DBSCAN that includes single-linkage hierarchical clustering in addition to the DBSCAN mechanism. Thus, HDBSCAN has the same advantages as DBSCAN and is more scalable in large dimensional data. The disadvantage is that the algorithm is more computationally intense than DBSCAN.

Thirdly, Agglomerative Clustering iteratively merges the two closest clusters until a stopping criterion is met. The algorithm applies a linkage method to define cluster proximity and a distance metric. The advantage of Agglomerative Clustering is the flexible clustering of complex shapes, which determines the number of clusters itself and provides a hierarchy of clusters that can be useful for visualization and interpretation. The disadvantages of Agglomerative Clustering are the sensitivity to noise and outliers, the computational intensity, and the scalability is challenging for high dimensional datasets (Zepeda-Mendoza & Resendis-Antonio, 2013). For the purpose of our research, we prefer to have more precise clustering with noise classified in a miscellaneous category rather than having the noise included in the clusters, which can lead to deficient analysis. The clustering algorithm of DBSCAN and HDBSCAN both yield good results in separating the noise. Moreover, HDBSCAN tends to perform well for patent data in a high dimensional domain in contrast to DBSCAN, and we accept computational intensity as a tradeoff for accuracy. Thus, HDBSCAN is the suitable clustering algorithm to use for our analysis.

Regarding the inputs of HDBSCAN, we set the minimum cluster size to 10 to ensure sufficient size for clusters and not too numerous clusters. Additionally, the default metric Euclidean is applied since it delivered better results for our research considering the number of clusters, the type of data, and the computation time. The other parameters

are left on default.

Selection of Algorithm for Cluster Matching

Due to the computational requirements, the clustering was executed on a yearly basis. To match the clusters over time, the cosine-similarity algorithm was used to match the cluster ids from one year to the cluster ids of the next year while all patents in the miscellaneous cluster were excluded and were not assigned a consistent cluster id. This enables us to connect the clusters over multiple years and give a more precise measure for analysis purposes. The cosine-similarity defines the similarity of two clusters through the normalized dot product of the CPC classification features of clusters. The most similar clusters are matched with one another. Attention should be paid to the varying number of clusters per year, which leads to a small mismatch for some clusters and patents. Thus, about 2500 out of 1.32 million patents are not assigned to a consistent cluster because of this mismatch in the number of clusters per year. This problem only occurs in the first years but is nonexistent in later years since the number of consistent clusters grows over time and covers more patents.

Subsequently, the yearly clusters are connected through the matched cluster ids, based on the first year 1980. The first set of consistent clusters is given by the clusters in 1980 and clusters in later years with the same or similar patterns of CPC classifications are assigned to these consistent clusters. Since the CPC patterns evolve over time, new clusters not seen in prior years create a new consistent cluster id and are assigned to this id for subsequent years. For yearly clustering, this leads to a total of 1042 clusters.

Due to the yearly clusters changing and few observations per cluster, the consistent clusters work well for consistent clusters existing only for few years, but some clusters lasting for a long time, especially clusters lasting over 35 years, tend to have changes in the pat-

tern, leading to a different CPC pattern for the patents in the last year in comparison to the patents in the first year assigned to the same constant cluster id. Thus, consistent clusters lasting many years have "breaks" in CPC patterns that capture different CPC patterns over time.

The clustering for the 5-year moving window allows for a matching of clusters based on the overlap of the years in the clustering window. Thus, a patent generally was allocated to five clusters, except for patents at the beginning or the end of the sample. Thus, connecting the cluster ids of the individual clustering leads to sound consistent cluster ids. The consistent cluster ids lead to precise results and no observed breaks but leads to some large clusters in which many firms operate and conduct research. Examples for extensive clusters are G06F (Electrical Digital Data Processing) and H04L (Transmission of Digital Information).

Comparison of Algorithms

	DBSCAN	**HDBSCAN**	**Agglomerative Clustering**
Inputs	eps=0.5, min_samples=10	min_cluster_size=10, metric='euclidean'	n_clusters=None, distance_threshold=10, compute_full_tree=True
Miscellaneous category for noise	Yes, cluster -1	Yes, cluster -1	No
Subsample (n = 16,182)			
Number of clusters	252	297	220
10 largest clusters	-1 7559 10 417 9 345 7 174 52 168 45 162 106 157 20 146 110 137 65 129	-1 5180 259 417 263 346 257 186 139 174 124 168 260 162 172 154 122 146 289 137	159 417 110 345 6 290 33 259 57 223 56 220 37 211 29 181 5 174 7 173
10 smallest clusters	232 10 100 10 230 10 198 10 90 10 200 10 202 10 42 10 220 10 78 10	100 11 77 11 17 11 73 11 147 11 115 11 151 11 168 11 35 11 185 10	163 18 124 17 101 17 67 17 103 17 170 16 213 15 197 9 143 6 207 4

Comparison of Encoding

	Only Subclass	**All classifications – Binary Encoding**	**All classifications – Non-Binary Encoding**
CPC Classifications included	CPC subclass	CPC section, CPC class, CPC subclass	CPC section, CPC class, CPC subclass
Encoding	Binary	Binary	Non-Binary
Subsample (n = 16,182) with HDBSCAN(min_cluster_size=10, metric='euclidean')			
Number of rows of dataset	578 columns	710 columns	710 columns
Number of clusters	265	298	297
10 largest clusters	-1 4837 227 452 251 369 231 215 178 186 98 181 214 177 194 176 154 175 260 174	-1 5848 295 417 293 345 149 192 238 170 199 162 188 158 178 147 292 138 159 130	-1 5180 259 417 263 346 257 186 139 174 124 168 260 162 172 154 122 146 289 137
10 smallest clusters	196 11 254 11 60 11 2 11 0 11 72 10 151 10 31 10 28 10 38 10	14 11 261 10 231 10 272 10 116 10 42 10 53 10 96 10 74 10 165 10	100 11 77 11 17 11 73 11 147 11 115 11 151 11 168 11 35 11 185 10

Consistent Cluster ID 5-year Window

	Consistent cluster ID	Number of patents
Largest 10 clusters	14490.0	108935
	14482.0	39509
	14523.0	38702
	14466.0	18844
	14572.0	15337
	14375.0	14295
	14376.0	13998
	14245.0	12604
	14576.0	12049
	14223.0	11906
Smallest 10 clusters	14497.0	31
	14123.0	31
	13792.0	30
	3807.0	30
	14251.0	30
	13595.0	29
	8170.0	27
	14193.0	26
	14398.0	20
	14111.0	19

Number of consistent clusters: 1023

References

Ester, M., Kriegel, H.P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96).*, https://www.dbs.ifi.lmu.de/Publikationen/Papers/KDD-96.final.frame.pdf.

Lytvynenko, V., Lurie, I., Krejci, J., Voronenko, M., Savina, N., & Taif, M. A. (2019). Two Step Density-Based Object-Inductive Clustering Algorithm. *Workshop Proceedings of the 8th International Conference on "Mathematics. Information Technologies. Education," MoMLeT&DS-2019*, https://ceur-ws.org/Vol-2386/paper10.pdf.

Zepeda-Mendoza, M.L., & Resendis-Antonio, O. (2013). Hierarchical Agglomerative Clustering. *Encyclopedia of Systems Biology.*, Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_1371.