Knowledge Cycles and Corporate Investment

M. Cecilia Bustamante* Julien Cujean† Laurent Frésard‡

December 22, 2022

Abstract

We examine how firms invest along their knowledge cycles. If investment is only a means to accumulate capital, cycles are irrelevant for the investment-value relation, with the two declining together over cycles. But we argue that investment also creates knowledge—serendipitously—disconnecting it from value. Investment is high early *and* late in the cycle, its relation with value spikes before new cycles start and declines thereafter. We uncover this specific pattern in the data, identifying new cycles using sharp changes in patents' citations to prior technologies. Cycles' length has tripled in recent years, coinciding with concurrent changes in the investment-value relation.

We thank Luke Taylor (EFA discussant) for suggesting the Intel example. We are grateful to Daniel Andrei, Serguey Braguinsky, Nicolas Crouzet (Cavalcade discussant), Francois Derrien, Howard Kung, Jerry Hoberg, Johan Hombert, Julien Hugonnier, Nobuhiro Kiyotaki, Albert Kyle, Mark Lowenstein, Song Ma, William Mann (ASU discussant), Rich Matthews, Erwan Morellec, Lukas Schmid (WFA discussant), Philip Valta, Laura Veldkamp, Yufang Wu (WEFI discussant), Alexei Zhdanov (FOM discussant) and Jake Zhao (MFA discussant) for their feedback, and to seminar participants at the 4Nations Cup at HEC Paris, the University of Southern Denmark, University of Bern, Bocconi University, the USC Finance Organizations and Markets (FOM) Conference 2019, the ASU Sonoran Winter Finance Conference 2020, the WFA Annual Meetings 2020, the Federal Reserve Board, Universite Paris Dauphine, Queen Mary University, the Econometric Society World Congress 2020, the European Finance Association Annual meeting 2020, the Women Assistant Professors in Finance Conference at Stern, the SFS Cavalcade conference 2021, the MFA Annual Meetings 2021, the Society for Economic Dynamics annual meeting 2021, and the Workshop on Entrepreneurial Finance and Innovation 2022. We thank Tina Oreski for excellent research assistance.

*University of Maryland, R. H. Smith School of Business. Contact: 4423 Van Munching Hall, College Park, MD 20742, USA; +1 (301) 405 7934; mcbustam@rhsmith.umd.edu

†University of Bern, Institute for Financial Management, and CEPR. Contact: Office 214, Engehaldenstrasse 4, CH - 3012 Bern; +41 (0) 31 631 3853; julien.cujean@unibe.ch

‡Universita della Svizzera italiana (USI Lugano), Swiss Finance Institute, and CEPR. Contact: Via Giuseppe Buffi 13, CH - 6900 Lugano; +41 (058) 666 4491; laurent.fresard@usi.ch

1 Introduction

The creation of technological knowledge within firms follows cycles. The case of *Intel* is a classic example (Casadesus-Masanell, Yoffie, and Mattu, 2014). From the early 1970s to 1982 *Intel* developed RAM chips technology, creating unique technological knowledge of substantial value for its shareholders. Competition from Japanese producers eventually eroded profits *Intel* made on this knowledge, precipitating the decision to switch technology and start a new cycle featuring a central processors technology, that began in 1988 and declined in the early 2000s. Interestingly, the value of *Intel* closely followed these cycles, peaking early in each cycle when superior knowledge conferred a leading market position, and declining afterwards when the returns to knowledge from the existing technology plunged.

In this paper, we study how firms should invest along their knowledge cycles. In particular, if firms' value varies over knowledge cycles, as the example of *Intel* indicates, at which stage of the cycle should firms invest most? Do investment and value move together along cycles, as neoclassical models predict, or in opposite directions? What determines the exploration of new technologies, and what is the resulting length of knowledge cycles? These questions appear particularly relevant in a time of rapid technological changes (e.g., clean technologies) in which many companies must decide whether and how to adjust their technological knowledge; yet, they have received surprisingly little attention in the literature.

Our main argument is that the answers to these questions depend on the role of investment in the process of knowledge creation. We first argue that from a theoretical perspective the relevance of knowledge cycles for investment is questionable. We make this point in a frictionless neoclassical model in which cycles arise and are determined entirely by firm value. In this context *if investment is only a means to accumulate capital*, value (marginal q) dictates entirely how the firm invests. Hence, remarkably, knowledge cycles are *irrelevant* for the relation between investment and value, with the two declining hand in hand over cycles. However, borrowing a long-standing idea in economics we propose that investment is also a means to create knowledge, a way to learn about and to experiment with a technology.[1] This notion, and more broadly the idea that knowledge creation is a by-product of

[1] Braguinsky, Ohyama, Okazaki, and Syverson (2021) show empirically that expanding firms' productive capital refines their knowledge about the quality of their current technology, in support of the idea that investment creates knowledge. This idea is also used in Grossman, Kihlstrom, and Mirman (1977), or more recently by Farboodi, Mihet, Philippon, and Veldkamp (2019).

economic activity (Arrow, 1962), remains unexplored in the neoclassical literature. Yet, we show that once investment becomes a means to learn knowledge cycles become central for understanding firms' investment decisions.

We develop a model that retains the main themes of neoclassical investment models, but considers investment as a form of *experimentation* that creates knowledge, and which allows firms to engage in *exploration* (to adopt new technologies). We study a risk-neutral firm that operates a given technology whose quality is unobservable. The firm combines this technology with capital to produce output and generate profits. Knowledge is the firm's confidence, in a statistical sense, about the quality of its existing technology. More knowledge increases the firm's productivity (Jovanovic and Nyarko, 1996). The firm accumulates knowledge by observing productivity realizations (learning-by-doing) and, because investment is also a means to create knowledge, by actively experimenting with its technology through investment (learning-by-investing). Importantly and distinct from capital expansion, the accumulation of knowledge is genuinely noisy. Thus unlike neoclassical models, the uncertainty inherent to experimentation outcomes (e.g., Callander (2011)) affects the firm's investment decision.

Knowledge cycles arise in the model because the firm cannot fully appropriate the returns on its knowledge. Due to workers' mobility, social interactions, imitation or competitive pressures, the firm's own technological knowledge "dissipates" to others, and thus the dollar benefits of knowledge on an existing technology are *bounded* (Young, 1993). In other words, operating the same technology forever is not optimal, and the exploration of new technologies is necessary to generate future revenues. To the extent that knowledge is only partially transferable across technologies (Jovanovic and Nyarko, 1996), the adoption of new technologies leads to partial knowledge resets, and generates knowledge cycles.

The model delivers four main insights. In contrast to neoclassical predictions, *(1)* firm's investment is high both early *and* late in a knowledge cycle, *(2)* the relation between investment and value shifts around knowledge resets, and, most importantly, *(3)* in a specific way—it spikes late in the cycle and declines early in the cycle. In addition, whereas knowledge resets often follow pre-determined schedules in the literature (e.g., Berk, Green, and Naik (2004) or Pastor and Veronesi (2009)), knowledge cycles in our model are endogenous: *(4)* they are entirely determined by firm value, much as in the *Intel* example, they are surprisingly short, and they expand when knowledge is easier to protect.

Two separate intuitions suggest that investment should be high at the beginning of a knowledge cycle. At early stages a new technology has highest option value. In addition, early stages are associated with high uncertainty and thus high learning benefits from investing. Perhaps it seems that investment should be low late in the cycle, when both value and uncertainty are low. Yet, this intuition omits that when investment creates knowledge the possibility to explore a new technology is a put option on experimentation. Although risk-neutral the firm develops an attitude towards the uncertainty of experiments, as their outcomes affect value only after it has made its investment decision. Owing to the concavity of firm's revenues in productivity, early in the cycle firm value is concave in knowledge—experimentation is a risk the firm would prefer to avoid. Later in the cycle firm value becomes convex in knowledge since learning benefits are bounded. The firm changes attitude, and starts gambling on new technologies by investing actively when the end of a cycle is in sight.

The point is that firm's attitude towards noisy experiments creates a disconnect between investment and value. The extent of this disconnect, and thus the relation between investment and value, itself varies along the knowledge cycle in a specific way. Early on, when uncertainty about the new technology is high, the firm prefers to avoid experimentation and invests *prudently* when in fact value is highest; this weakens the relation between investment and value. A cycle ends when exploration is optimal, and this occurs when the marginal value of knowledge on the current technology approaches zero. Since more knowledge harms value late in the cycle due to accelerated revenue erosion, the marginal value of capital and that of knowledge together go up, exactly when the firm gambles on new technologies by raising investment. Hence, the relation between value and investment is stronger late in the cycle. This pattern is specific to the disconnect experimentation creates; it does not arise when investment in capital and knowledge are separate decisions nor can it be explained by traditional frictions (e.g., fixed adjustment costs), and it is flat when investment only serves to expand capital.

We exploit this novel prediction to demonstrate the relevance of knowledge cycles for firms' investment in the data. Guided by the theory, we test whether the relation between investment and value shifts around knowledge resets as the model predicts. To implement this test we develop a measure of knowledge resets using information from the patents of a large sample of U.S. publicly-listed firms between 1976 and 2017. We first define all patents that a firm cites in its own patents (in recent years) as its "knowledge base": these citations capture the technological knowledge the firm has accumulated over time (Ma, 2021).

Aggregating cited patents across NBER technology classes, we then identify major annual changes in a firm's knowledge base through cosine similarity between the firm's current and lagged distributions of its technological knowledge across classes. A "knowledge reset" corresponds to a year in which the distributions of cited technologies suddenly becomes dissimilar.

To map the measurement of investment and value to their definition in the model, we follow Peters and Taylor (2017) and consider "total investment" in physical and intangible capital and "total Q," the ratio of firm value to physical and intangible capital. We then examine the relation between the two in event time around knowledge resets. Confirming that there exists a disconnect between investment and value over knowledge cycles, we first find a clear break in the relation around resets. Furthermore and most notably, supporting the idea that the disconnect results from investment creating knowledge, we uncover a significant spike in the investment$-Q$ relation immediately prior to the reset (late in the cycle) and a steady decline (of about 40%) following it (early in the cycle).

There exists many theories for a disconnect between investment and value but these do not easily explain this specific pattern, and their potential connection with knowledge resets is unclear. Competition and the "rise of intangibles," two forces deemed responsible for a disconnect in recent years (Gutierrez and Philippon, 2018), are building blocks of our benchmark and so do not explain the pattern, which we also confirm empirically. Another established origin of a disconnect is financing frictions (Gomes, 2001; Hennessy, Levy, and Whited, 2007), but we find they do not explain the pattern, reminiscent of the result of Gutierrez and Philippon (2018). Another possibility is that investment in physical capital and intangibles are separate decisions, but we uncover empirically that the two follow similar patterns around resets. Fixed adjustment costs also create a disconnect (Caballero and Leahy, 1996) but we find theoretically they imply the opposite pattern. Nor is this pattern explained by potential changes in ownership structure around resets, nor by mechanical changes in patenting activity, nor by acquisition of firms with distinct technologies.

We also find that knowledge cycles (the time between two resets) are short, with a median of 5 to 7 years. They are shorter in more competitive markets in which knowledge likely dissipates faster and thus induce firms to explore new technologies sooner. Notably, we observe that knowledge resets have become less frequent over time, with cycle length almost tripling between 1980 and 2017 (from about 5 years to 13 years). Though not our focus, longer cycles could originate from the increased difficulty to generate new ideas (Bloom, Jones, van Reenen, and Webb, 2017) or declining competition (Gutierrez and Philippon, 2018). Since our model predicts that the investment$-Q$ relation is dictated by knowledge cycles, it is interesting to contrast their lengthening to the evolution of this relation. Gutierrez and Philippon (2017) and Alexander and Eberly (2018) document a decline in investment in the last twenty years but no decline in Q, which our sample confirms. The relationship between investment and Q being strongest around resets, its weakening over time could be plausibly tied to less frequent resets and changes in firms' process of knowledge creation.

This paper adds to a growing literature on the role of intangible capital in firms' investment and valuation. Recent models do not consider knowledge cycles and often incorporate intangible capital as a production factor that accumulates similarly to physical capital, and under perfect information (e.g., Eisfeldt and Papanikolau (2013), Peters and Taylor (2017), Crouzet and Eberly (2018), Crouzet and Eberly (2020)). Closer to this paper, Andrei, Mann, and Moyen (2018) develop a model in which technological innovation occurs randomly and firms learn about it passively. In contrast, we consider that investment is a means to create knowledge, and we study the implications of this idea for understanding the firms' investment over their knowledge cycles. Our results extend the existing literature by highlighting that investment creates knowledge and thus directly affects firms' stock of intangible capital.

We also contribute to the literature in finance that explicitly models the process of knowledge creation within firms (e.g., Berk et al. (2004), Bergemann and Hege (2005), Pastor and Veronesi (2009), Manso (2011), or Manso, Balsmeier, and Fleming (2019)). Whereas we share the view that experimentation and exploration dictate the evolution of firms' technological knowledge, we associate experimentation with investment and we focus on the implications for the relation between investment and value over knowledge cycles.

This paper also belongs to the literature on experimentation, and is closest technically to Moscarini and Smith (2001).[2] Similar to their decision maker, our firm decides how much to experiment (to invest) and when to take a payoff-relevant action (permanent stop in their case, and exploration in ours). We also find that experimentation intensity is high when the action is about to be taken. Their decision maker minimizes the present value of experimentation costs which delays high-intensity experimentation. Our firm experiments

[2] For instance, Grossman et al. (1977), Weitzman (1979), Jovanovic and Rob (1990), Rob (1991), Aghion, Bolton, Harris, and Jullien (1991), Jovanovic and Nyarko (1996), Keller and Rady (1999), Callander (2011).

more late in the cycle because exploration induces gambling on new technologies. Another difference is in the formulation and scope. They focus on sequential analysis (Wald, 1947), whereas we study investment in a neoclassical context.

2 Model setup

2.1 Main assumptions

We consider a firm that combines capital, K_t, with some technology to produce a unique non-storable output, Y_t, at time t according to the production function:

$$Y_t = A_{n,t} K_t^{\alpha}, \tag{1}$$

where $A_{n,t}$ denotes the productivity of technology indexed by $n \in \mathbb{N}$ and $\alpha \in (0,1)$ denotes returns to scale on capital. Productivity, $A_{n,t}$, of technology n follows the law of motion:

$$\mathrm{d}A_{n,t}/A_{n,t} = \tau_A^{1/2} \frac{M_n}{\Omega_{n,t}^{1/2}} \mathrm{d}t + \mathrm{d}B_t, \tag{2}$$

where M_n determines the growth rate of the productivity of the current technology and thus captures its *quality*, and B is a Brownian shock.[3] Throughout the analysis, we use the concept of technology in a broad sense as a "technology" may refer to production techniques, but also management practices (e.g., Bloom, Sadun, and van Reenen (2016)), or organization design (e.g., Prescott and Visscher (1980)). We assume that technologies only differ with respect to their respective quality.

The quality, M_n, of each technology n is unobservable, and learning about it is a noisy process. Although the firm does not observe the quality of its current technology, it observes realizations of its productivity. Each incremental realization of $\mathrm{d}A/A$ reveals noisy information about M_n, allowing the firm to revise its estimate of quality.[4] Following Jovanovic and Nyarko (1996) learning improves productivity, which we model by scaling quality, M_n, by

[3] The literature (e.g., Acemoglu (2009)) typically assumes that technological innovation modifies the level of productivity (i.e., A), as opposed to its growth rate (M). Adding this feature to the model is possible but is strictly equivalent to reducing obsolescence costs, given the way we model the evolution of capital below. Hence, we characterize technologies only in terms of their productivity growth.

[4] Formally, we define how the firm sets its own estimate about M in Section 2.2.

the conditional standard error, $\Omega_n^{1/2}$, of the firm's estimate of quality (learning reduces Ω_n in a way to be described). To simplify interpretation, we normalize the conditional variance of productivity realizations in (2) to one so that τ_A measures the informativeness of the productivity signal (about M_n) in time-invariant units of uncertainty (e.g., Kyle, Obizhaeva, and Wang (2018)).

The firm decides optimally how much to invest, and which technology to operate. At any time t, *exploration* refers to the firm's decision to abandon its current technology, n, and switch to a new one, $n+1$. Consistent with previous work incorporating exploration into neoclassical settings (see, e.g., Abel and Eberly (2012)), abandoning an existing technology to explore is costly: a fraction ω of capital K becomes obsolete. In contrast with these studies, however, in this model the link between qualities across technologies is informational. We follow Jovanovic and Nyarko (1996) and assume that switching to a new technology reduces knowledge temporarily, and that once a technology has been abandoned it can never be "recalled." In the main analysis, for simplicity, we focus on the extreme case in which each technology resets knowledge completely, or equivalently the firm "learns and forgets" (e.g., Benkard (2000)); we relax this assumption in Section 4.4, allowing knowledge to be partially transferable across technologies. Upon selecting technology $n+1$, its quality is randomly drawn as per:

$$M_{n+1} \sim \mathscr{N}(0, \tau_M^{-1}), \quad \text{with } M_n \perp M_{n+1}, \quad \forall n \in \mathbb{N},$$

so that the technology index n can be discarded, and the parameter τ_M captures prior precision regarding the quality of the technology.[5]

The firm has incentives to explore new technologies because the benefits of knowledge are *bounded* (Young, 1993). A large literature indicates that, because technological knowledge is non-rival and largely embodied in (inalienable) workers, firms cannot fully appropriate its returns. Proprietary knowledge thus typically dissipates over time (e.g., Romer (1986); Lucas (1988)). Knowledge dissipation results, for instance, from workers moving across firms (e.g., Stoyanov and Zubanov (2012)), social interactions (e.g., Glaeser, Kallal, Scheinkman, and Shleifer (1992)), imitation (e.g., Lieberman and Asaba (2006)), or illegal actions (corporate

[5]The assumption that each technology resets knowledge completely, along with the specification of the signals in Eqs. (2)–(6) below, jointly imply that the parameter τ_M plays no role in the analysis.

spying or thefts).[6] To incorporate this aspect we assume that the firm faces an isoelastic demand curve for the good (or service) it produces, $P_t = (Y_t N_t)^{-\eta}$, with N an exogenous variable that captures the dissipation of the firm's knowledge in reduced form and $\eta \in (0,1)$ the price elasticity of demand. The firm's revenues are increasing in the productivity of the current technology, A, but are decreasing in the intensity of knowledge dissipation, N:

$$\Pi\left(A_t, K_t, N_t\right) \equiv P_t Y_t = A_t^{1-\eta} K_t^{\alpha(1-\eta)} N_t^{-\eta}. \tag{3}$$

Knowledge dissipation erodes revenues, making exploration necessary for firms to regenerate future revenues.[7] Upon exploration the firm faces an informational trade-off: it gives away a known technology with low expected profitability, in exchange for a new technology with higher expected return but of unknown quality.

The remaining corporate decision is that of investment, which in our model is coupled with *experimentation*. As is customary in the neoclassical literature, the firm accumulates capital K in (1) through corporate investment. Let $I_t \geq 0$ represent investment in capital at time t and $i_t = I_t/K_t$ be the corresponding investment rate. We assume that capital depreciates at a constant rate, δ. It follows that the net change in the stock of capital is:

$$\mathrm{d}K_t = (i_t - \delta)K_t \mathrm{d}t - \omega K_{t-} \mathbf{1}_{t=\nu}, \tag{4}$$

where, throughout the paper, ν denotes dates at which the firm decides to abandon the current technology and to explore an unknown technology. Absent exploration, capital accumulation in (4) is identical to that in the neoclassical literature. In addition, investment entails adjustment costs, $\gamma_t(i)$, for purchasing and installing new capital, which for analytical convenience we assume proportional to operating revenues, Π_t (e.g., Cooper (2006); Hackbarth and Johnson (2015)):

$$\gamma_t \equiv \underbrace{\left(i + \gamma/2(i-\delta)^2\right)}_{\equiv \widehat{\gamma}(i)} \cdot \Pi_t, \tag{5}$$

[6] The firm may limit the dissipation of its knowledge by using patents (or trademarks). Intellectual property tools are however limited in their scope because certain types of knowledge cannot be precisely codified, are costly to give away through public filings (e.g., trade secrets), or because patents may fail to adequately protect property rights (e.g., Lanjouw and Schankerman (2001)).

[7] We specify the law of motion for the process N in Section 2.2.

so that the price of capital is taken to be proportional to Π/K, and the adjustment cost function is convex (e.g., Abel and Eberly (1994)). In Section 5.4.2 we consider an extension with fixed costs. We assume investment is strictly positive ($I > 0$) and, thus, irreversible.[8]

The novelty of the investment decision in this model compared to previous neoclassical studies is that investment is not only a means to expand capital, but also a way to create knowledge. Knowledge is a serendipitous by-product of economic activity (e.g., Arrow (1962); Grossman et al. (1977); Farboodi et al. (2019)). For instance, Braguinsky et al. (2021) show that expanding the set of machines used for production enlarges the productivity of machines and refines the firm's information about its current technology.[9] Experimentation thus takes the form of an expansion of capital on which the current technology is applied. Formally, investment is a form of experiment that provides the firm with an informative signal flow, $\mathrm{d}S$, about the quality of the existing technology, M:

$$\mathrm{d}S_t = \frac{\tau_S^{1/2} i_t^{1/2}}{\Omega_t^{1/2}} M \mathrm{d}t + \mathrm{d}B_{S,t}, \tag{6}$$

where τ_S is a parameter that, together with investment, $\tau_S i$, captures how informative experimentation is. The more the firm invests relative to its stock of capital, the more it learns about M. Setting τ_S to zero further allows us to shut down experimentation (see Section 3). Importantly, the last term in (6) introduces an independent Brownian, $B_S \perp B$, in signal realizations, which captures a genuine feature of experimentation: when the firm invests to create knowledge, the outcome of its experiments is uncertain (e.g., Callander (2011)). Similar to the productivity signal, informativeness of S is measured in time-invariant units of uncertainty. Since investment augments both capital and knowledge, it can be viewed as total investment in capital and knowledge (Peters and Taylor, 2017); in Section 5.4.2 we also allow the firm to make separate investment decisions in capital and in the signal flow, S.

2.2 Knowledge creation within the firm

These assumptions imply that the firm collects information about the quality of its existing technology both through productivity realizations (passive learning, as per (2)) and the signal

[8] Formally, this implies that adjustment costs are asymmetric, and that disinvestment is infinitely costly.

[9] As noted by Arrow (1962): "Each new machine produced and put into use is capable of changing the environment in which production takes place so that learning is taking place with continually new stimuli."

flow investment creates (active learning, as per (6)). These two signals together represent all information that is available to the firm:

$$\mathscr{F}_t = \sigma\left((A_u, S_u) : u \leq t\right).$$

Based on this information, the firm computes an estimate, $\widehat{M}_t = \mathbb{E}[M|\mathscr{F}_t]$, of quality. The conditional error variance of this estimate is $\Omega_t = \mathbb{V}[M|\mathscr{F}_t]$ (introduced in (6)). Standard arguments imply that the firm gradually updates these statistics according to:

$$\begin{aligned}
\mathrm{d}\widehat{M}_t &= \Omega_t^{1/2}\tau_A^{1/2}\mathrm{d}\widehat{B}_t + \Omega_t^{1/2}\tau_S^{1/2}i_t^{1/2}\mathrm{d}\widehat{B}_{S,t} - \widehat{M}_{t-}\mathbf{1}_{t=\nu}\\
\mathrm{d}\Omega_t &= -\Omega_t\left(\tau_A + \tau_S i_t\right)\mathrm{d}t + (\tau_M^{-1} - \Omega_{t-})\mathbf{1}_{t=\nu},
\end{aligned}$$

where $\widehat{B}$ and $\widehat{B}_S$ are two independent Brownian processes under the firm's probability measure.[10] The firm updates its estimates about the quality of its current technology continuously, by observing realized productivity and by experimenting with it, and resets its estimate and precision to priors (0 and τ_M, respectively) whenever it decides to trigger exploration.

We specify the dynamics of knowledge dissipation, N, that erodes revenues based on the ratio of these two numbers, $Z \equiv \widehat{M}/\Omega^{1/2}$, which will be shown to define the relevant notion of "knowledge" in the model. Since (the logarithm of) revenues in (3) are increasing linearly in Z but decreasing linearly in $\log(N)$, we let the dissipation intensity evolve according to:

$$\mathrm{d}N_t/N_t = \phi Z_t^2\mathbf{1}_{Z_t\geq 0}\mathrm{d}t, \tag{7}$$

with $\phi > 0$.[11] In words, we assume that knowledge dissipates faster as it accumulates, and this dissipation only occurs for "good" technologies, $Z > 0$. When a technology is likely of poor quality, the firm has strong incentives to explore. When the firm discovers a good technology instead, it does not let knowledge grow unboundedly by sticking to the same technology forever, as dissipation eventually erodes its revenues to zero (e.g., Jovanovic and MacDonald (1994)).

We conclude by formulating the firm's problem. The firm maximizes the expected value

[10] These dynamics follow directly from Theorem 12.7 in Lipster and Shiryaev (2001).

[11] We assume that the growth rate of $\log(N)$ is quadratic in Z, because the resulting affine-quadratic framework remains tractable. However, any *strictly* convex increase in N will keep profits bounded and produce qualitatively similar results, except in terms of the asymmetry between good and bad technologies.

of its future profits (that is, revenues net of costs) by optimally choosing its technology and its investment. Assuming the firm is risk neutral and discounts its profits at a constant rate, r, its value, $V(\cdot)$, then equals:

$$\max_{\{\nu_n\}, i_{t+s} \geq 0} \mathbb{E}\left[\int_0^\infty e^{-rs} \left(\Pi \left(A_{t+s}, K_{t+s}, N_{t+s} \right) - \gamma_{t+s} \right) \mathrm{d}s \middle| \mathscr{F}_t \right]. \tag{8}$$

In Appendix B we show that our specification of adjustment costs in (5) implies that value is homogeneous in revenues; learning improving productivity as in (2) and our specification of knowledge dissipation in (7) further mean value is homogeneous of degree zero in $(\widehat{M}, \Omega^{1/2})$:

$$V\left(N_t, A_t, K_t, \widehat{M}_t, \Omega_t\right) \equiv \Pi\left(A_t, K_t, N_t\right) \cdot v\left(\widehat{M}_t / \Omega_t^{1/2}\right) \equiv \Pi\left(A_t, K_t, N_t\right) \cdot v\left(Z_t\right), \tag{9}$$

where $v(\cdot)$ denotes the intensive value of the firm.

We can now be specific about what "knowledge" means in this model. The firm's estimate, $\widehat{M}$, and its standard error, $\Omega^{1/2}$, do not matter separately; they only matter as a $t-$statistic, Z_t, which summarizes all the information the firm has at date t about the technology it operates to decide whether to experiment or to explore:

$$\mathrm{d}Z_t = \underbrace{\tau_A/2Z_t\mathrm{d}t + \tau_A^{1/2}\mathrm{d}\widehat{B}_t}_{\text{passive learning}} + \underbrace{\tau_S i_t/2Z_t\mathrm{d}t + \tau_S^{1/2} i_t^{1/2}\mathrm{d}\widehat{B}_{S,t}}_{\text{learning by experimenting}} \underbrace{-Z_{t-}\mathbf{1}_{t=\nu}}_{\text{knowledge reset}}.$$

Positive (negative) values of Z indicate that the firm is confident, in a statistical sense, that the quality of its current technology is high (low). On average, knowledge moves away from prior beliefs, $Z_0 = 0$, accumulating in continuation of the current estimate, Z_t, at the speed with which passive and active signals reveal information ($\tau_A/2$ and $\tau_S i_t/2$, respectively). Exploring a new technology resets firm's knowledge to $Z_\nu = 0$, leading to a *knowledge reset.*

3 On the relevance of knowledge cycles: a benchmark

To highlight the role of investment as a form of experimentation, we first study the benchmark case in which investment is simply a means to accumulate capital. Even absent experimentation, the firm retains the ability to explore new technologies. Thus knowledge cycles already arise in this benchmark case, which is regularly overlooked in the neoclassical literature.

Similarly, whereas knowledge and competition in the form of knowledge dissipation are not standard elements in this literature they are building blocks of this benchmark. We explain how these features matter for theoretical conclusions whenever relevant.

As a first step, we define the marginal product of capital (hereafter q).

Definition 1. *The **marginal product of capital** q is the ratio of value of an additional unit of capital to profits, ignoring the effect of investment on knowledge:*

$$q(Z) \equiv \frac{1}{\Pi} \cdot K \cdot V_K. \tag{10}$$

A (cosmetic) difference of Eq. (10) relative to the standard definition of q in the neoclassical literature is on the relation between firm value and q. In neoclassical models, firm value is often assumed to be homogeneous in K so that q is proportional to V/K (average q). Our specification of adjustment costs, however, considers that firm value is homogeneous in profits, Π, (see Eq. (9)) so that q is proportional to V/Π (intensive firm value v):

$$q(Z) = \alpha(1-\eta)v(Z) \propto v(Z).$$

The benchmark model delivers the seminal insight in Hayashi (1982) that q is a sufficient statistic for investment. To show this, and solve for the firm's optimal investment policy, suppose for the moment that it is optimal for the firm to stick to its current technology. In this case the (intensive) value of the firm, v, associated with Eq. (8) satisfies the HJB equation:

$$\begin{aligned} &v(Z)\left(r + \eta Z^2 \phi \mathbf{1}_{Z>0} - (\eta - 1)\left(\alpha\delta + \frac{\eta}{2} - \sqrt{\tau_A} Z\right)\right) \\ &= \psi(q(Z)) + \frac{1}{2}\tau_A v''(Z) + \sqrt{\tau_A} v'(Z)\left(-2\eta + \frac{\sqrt{\tau_A} Z}{2} + 1\right), \end{aligned} \tag{11}$$

where we have defined ψ to be the maximand:

$$\psi(q) \equiv \max_{i \geq 0} \{qi - \widehat{\gamma}(i)\} \equiv (q-1)(2\gamma\delta + q - 1)/2\gamma. \tag{12}$$

The firm's optimal investment policy is given by:

$$i_t \equiv i(q) = \max\{\delta + (q-1)/\gamma, 0\},$$

thus delivering the traditional insight that q is a sufficient statistic for investment. In the benchmark experiments are uninformative and Z evolves through passive learning and resets:

$$\mathrm{d}Z_t = \tau_A/2 Z_t \mathrm{d}t + \tau_A^{1/2}\mathrm{d}\widehat{B}_t - Z_{t-}\mathbf{1}_{t=\nu}.$$

Since investment only affects value through capital accumulation, the marginal product of capital satisfies $q(Z) \equiv \frac{1}{\Pi}\frac{\mathrm{d}}{\mathrm{d}i}\frac{\mathrm{d}V}{\mathrm{d}t}$, and thus summarizes entirely how investment affects value.

Not only does q determine investment, but it is also a sufficient statistic for when the firm decides to explore a new technology. The solution assumes so far that it is optimal for the firm not to switch technology. When the firm decides to explore a new technology, it resets its knowledge to priors and incurs a lump-sum cost of obsolescence, ωK, on its existing capital stock. Since q in Eq. (10) is proportional to (intensive) firm value, q upon exploration is:

$$q^{\star} \equiv (1-\omega)^{\alpha(1-\eta)} q\left(0\right), \tag{13}$$

so that the firm switches technology whenever $q \leq q^{\star}$. Because exploration resets knowledge, of which q is a function, determining when exploration is optimal implies solving for q in Eq. (11); this requires some conjecture on the dependence of firm value on knowledge.

Intuitively, q peaks around knowledge resets, $Z \approx 0$, following exploration. As Z grows negative the firm is confident that its technology is poor, and its value must decrease as Z further declines; as the firm instead becomes confident that its technology is of high quality, knowledge dissipation intensifies and its value must decline as Z further rises. We conjecture that q has a single peak, and later use parameter values under which this conjecture is verified numerically. Thus, q is high when a technology has high option value and knowledge dissipation is low ($Z \approx 0$), and this occurs following exploration. Conversely, q is low when option value is low and dissipation is high, and this triggers exploration.

Under this conjecture, there must exist two "trigger levels" of knowledge Z, say $\underline{a} < \overline{a}$, at which $q(Z) = q^{\star}$ is satisfied and the firm decides to explore. Thus, $\mathscr{A} = (\underline{a}, \overline{a})$ denotes the region in which the firm does not explore. The times $\{\nu_k\}_{k=0}^{\infty}$ in Eq. (8) at which the firm explores correspond to hitting times at which knowledge exits $\mathscr{A}$:

$$\nu_n = \inf\{t \geq \nu_{n-1} : Z_t \notin \mathscr{A}\}, \ \ \forall n \in \mathbb{N}. \tag{14}$$

Every time the firm triggers exploration, it ends a cycle and begins a new one. We refer to the period in between exploration triggers as a "knowledge cycle".

Definition 2. *A* ***knowledge cycle*** *starts when the firm resets its knowledge to 0 by exploring a new technology ($Z_0 = 0$), and ends after a period of (random) length as per Eq.* (14) *when the firm abandons the technology to explore a new one ($Z \notin \mathscr{A}$).*

We compute firm value piecewise over the two regions $[\underline{a}, 0]$ and $(0, \overline{a}]$. Eq. (11) does not have an explicit solution; we proceed numerically, imposing the two boundary conditions implied by Eq. (13), $v(\underline{a}) = v(\overline{a}) = (1-\omega)^{\alpha(1-\eta)} v(0)$. We obtain another boundary condition by "piecing together" firm value across the two regions, $v(0_-) = v(0_+)$. Customary "smooth-pasting" conditions then provide the remaining conditions that determine the location of the optimal exploration thresholds.[12] Figure 1 illustrates q as a function of the stock of knowledge. This illustration relies on parameter values, which we discuss in Appendix A.2.

Recall that absent experimentation knowledge is irrelevant for investment—only q matters. Now, since q itself depends on knowledge, there exists an indirect relation between investment and knowledge. Figure 1 confirms that q is high (low) early (late) in the cycle. Thus, investment is highest following recent exploration ($Z \approx 0$), and weakest when q is low and triggers exploration ($Z \approx \underline{a}$ or $Z \approx \overline{a}$). However, because investment and q move exactly together their relation (investment$-q$ sensitivity) is flat (at $1/\gamma$) throughout the cycle. In other words, although knowledge cycles are present they do not affect the relation between investment and value.

[12] Smooth-pasting conditions satisfy:

$$\lim_{Z \downarrow \underline{a}} v'(Z) = \lim_{Z \uparrow \overline{a}} v'(Z) = 0, \ \text{ and } \ \lim_{Z \uparrow 0} v'(Z) = \lim_{Z \downarrow 0} v'(Z).$$

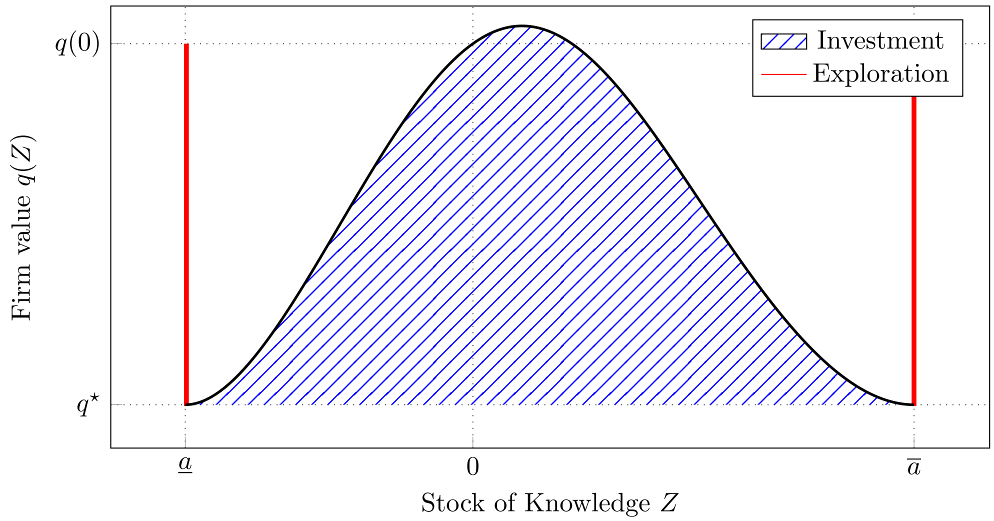

Figure 1: Marginal product of capital as a function of knowledge in the benchmark model. The figure plots q as a function of knowledge, Z. Smooth-pasting for exploration requires that q lands flat at the edges of the domain defined by the exploration thresholds $\underline{a}$ and $\overline{a}$. Parameter values are defined in Appendix A.2.

4 Revisiting knowledge cycles with experimentation

We now study the case in which the firm invests as a means to experiment and to create knowledge. The solution method in this case is similar to that of the benchmark. We conjecture that there exists a region $\mathscr{A}$, outside of which the firm decides to explore, and inside which it experiments; we solve Eq. (11) imposing identical boundary conditions. The only difference is that the maximand in Eq. (12) now not only involves q but also the *expected marginal benefit of knowledge*, c:

$$\psi(q,c) = \max_{i\geq 0} \left\{(c+q)i - \widehat{\gamma}(i)\right\}, \tag{15}$$

which we define next.

Definition 3. *The* ***expected marginal benefit of knowledge*** *is the ratio of expected value of an additional unit of knowledge to profits, ignoring the (traditional) effect of investment on capital:*

$$c(Z) \equiv \frac{1}{\Pi} \cdot \tau_S \cdot \Omega \cdot \left(-V_\Omega + \frac{1}{2} \cdot V_{\widehat{M}\widehat{M}}\right). \tag{16}$$

The expected marginal benefit of knowledge is a new variable that explains investment, in addition to the marginal product of capital, q, and which affects, to a large extent, its relation to value. From Eq. (15) the first-order condition for optimal investment, i, is:

$$\underbrace{\widehat{\gamma}'(i(Z)) = q(Z)}_{\text{neoclassical trade-off}} + \underbrace{c(Z)}_{\substack{\text{expected marginal} \\ \text{benefit of knowledge}}} . \tag{17}$$

The neoclassical trade-off, as present in the benchmark, equalizes the marginal value of capital, q, to marginal costs. When investment also creates knowledge, marginal costs equal the sum of q and the expected marginal benefit of knowledge. Hence, q is no longer a sufficient statistic for investment, and this disconnect means that the firm now follows a knowledge-contingent investment plan, $i(Z)$, which q does not subsume.

The effect of investment on knowledge, unlike traditional capital, is uncertain. For this reason, note first that c represents the *average* marginal benefit of knowledge. Formally, we can rewrite c (over a period of length $\mathrm{d}t$) as:

$$c(Z)\mathrm{d}t = \frac{1}{\Pi}\frac{\mathrm{d}}{\mathrm{d}i}\mathbb{E}\left[V_M \mathrm{d}\widehat{M} + V_\Omega \mathrm{d}\Omega + \frac{1}{2}V_{\widehat{M}\widehat{M}}\mathrm{d}\langle\widehat{M}\rangle\right] = \frac{1}{\Pi}\frac{\mathrm{d}}{\mathrm{d}i}\left(V_\Omega \mathrm{d}\Omega + \frac{1}{2}V_{\widehat{M}\widehat{M}}\mathrm{d}\langle\widehat{M}\rangle\right).$$

The first equality averages the value of an additional unit of knowledge over the firm's revisions in quality estimates, $\mathrm{d}\widehat{M}$, associated with each possible experimentation outcome. From the firm's perspective $\widehat{M}$ is a martingale (because quality is fixed for a given technology) and these revisions are pure noise and thus have mean zero, the second equality. However, and most importantly, this noise *does* matter through its instantaneous variance, $\mathrm{d}\langle\widehat{M}\rangle$.

Because experimentation noise matters for evaluating the expected marginal benefit of knowledge, this benefit is a bundle of two alternative forces. To see this, note that Eq. (9) implies $V_\Omega = -Z/2v'\Pi\Omega^{-1}$ and $V_{MM} = v''\Pi\Omega^{-1}$ and we can rewrite Eq. (16) as:

$$c(Z) \equiv \underbrace{\frac{\tau_S}{2}v'(Z)Z}_{\text{knowledge } q} + \underbrace{\frac{\tau_S}{2}v''(Z)}_{\substack{\text{attitude towards} \\ \text{noisy experimentation}}} . \tag{18}$$

Just as q measures the marginal value of capital, the first component in (18) measures the marginal value of reducing the conditional error variance, Ω, of the quality estimate $\widehat{M}$ and can be thought of as "knowledge q." In a Gaussian setup, the effect of experimentation on the conditional variance is known, much as that on capital accumulation. Yet a distinction with capital is that knowledge accumulation is noisy. The second component in Eq. (18), which is central to our argument, captures the firm's attitude towards this noise. The firm's attitude towards experimentation noise and its knowledge q evolve differently over the knowledge cycle, as we now demonstrate.

4.1 Investment and firm's attitude towards noisy experimentation

Our main argument is that when investment creates knowledge it is high at the beginning *and* at the end of the knowledge cycle. Intuitively, early stages of a cycle are associated with high q, the intuition the benchmark delivers. And if, in addition, investment is a means to learn, early stages are associated with high uncertainty and thus high learning benefits. It would seem that investment should be low late in the cycle when both q and uncertainty are low. But this conjecture ignores that when investment creates knowledge the possibility to explore new technologies is a put option on experimentation. Exploration induces firms to gamble on new technologies by investing actively when the end of a cycle is in sight.

The economic mechanism underlying our argument operates as follows. Focus first on knowledge q, the first component of the expected marginal benefit of knowledge in Eq. (18), and use it to define "total q" as:

$$q_{\text{tot}}(Z) \equiv q(Z) + \frac{\tau_S}{2} v'(Z)Z, \tag{19}$$

that is, traditional q plus knowledge q. Total q has the notable advantage that it can be reasonably constructed from the data (Peters and Taylor, 2017). Figure 2 shows how traditional q (solid line) and total q (dashed red line) differ over the knowledge cycle, separately for good ($Z > 0$, left panel) and bad ($Z < 0$, right panel) technologies. Note that in the benchmark model they are identical; with experimentation, they may only coincide at the beginning of a knowledge cycle (because $Z \equiv 0$) and at the end of it (due to smooth-pasting, $v' \equiv 0$).

Knowledge q works hand in hand with traditional q (defined earlier), except late in the knowledge cycle. Note first that experimentation does not affect the shape of traditional q (relative to that in Figure 1 in the benchmark): with or without experimentation, q is

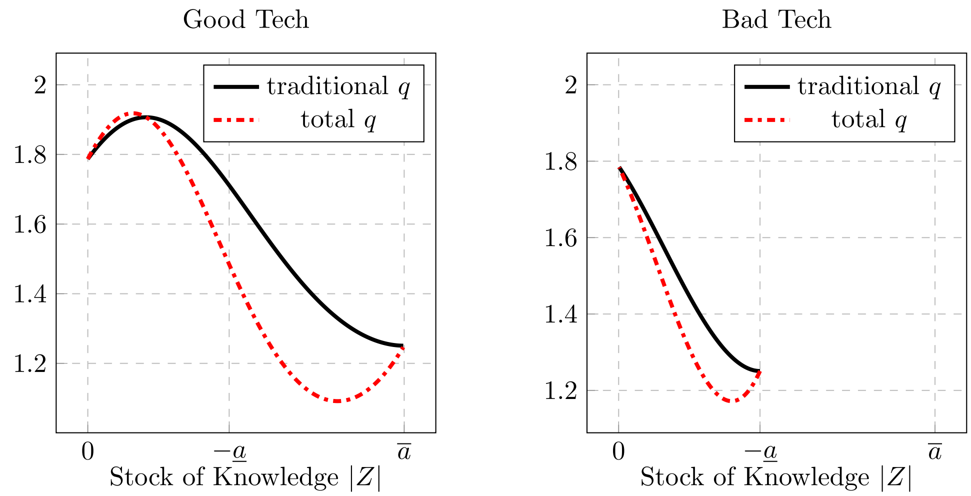

Figure 2: Traditional q and total q. This figure plots traditional q (solid black line) and total q (dashed red line) as functions of knowledge, Z, separately for good (left panel) and bad (right panel) technologies. Parameter values correspond to the baseline calibration, defined in Appendix A.2.

high early in the cycle and low late in the cycle. Since learning benefits are high early in the cycle and decline thereafter, knowledge q pushes total q above traditional q early in the cycle and causes it to decline faster after, thus reinforcing the neoclassical channel. However, for exploration to be optimal improvement in knowledge on the current technology must be eventually valueless (knowledge q is 0), causing total q to rise late in the cycle. Nevertheless, both total and traditional q may only imply low investment late in the cycle.

This intuition is incomplete, as noisy experimentation only affects firm value *ex-post*, that is, only after the firm has made its investment decision. Hence, the curvature of firm value, as measured by the second component, v'', in Eq. (18), matters for the firm's decision of ramping investment up or down. In particular, Eqs. (17), (18) and (19) together imply that total q does not constitute a sufficient statistic for investment:

$$i(Z) = \delta + \frac{1}{\gamma}\Big(q_{\text{tot}}(Z) + \underbrace{\tau_S/2v''(Z)}_{\substack{\text{attitude towards}\\ \text{noisy experimentation}}} - 1\Big), \tag{20}$$

as it ignores the firm's attitude towards noisy experimentation, an attitude which we illustrate in Figure 3.

Early in the cycle firm value is concave, $v'' < 0$, owing to the concavity of revenues in productivity, and experimentation is a risk the firm would prefer to avoid. That is, keeping capital fixed expected value post investment is below its known, pre-investment value. Halfway through the cycle, however, knowledge dissipation accelerates and firm value becomes convex, $v'' > 0$. As the firm becomes confident about the quality of the existing technology, it develops a gambling attitude towards noisy experiments and ramps up on investment. The possibility of ramping knowledge down through exploration is a put option on experimentation, inducing the firm to gamble on new technologies by raising investment late in the cycle.

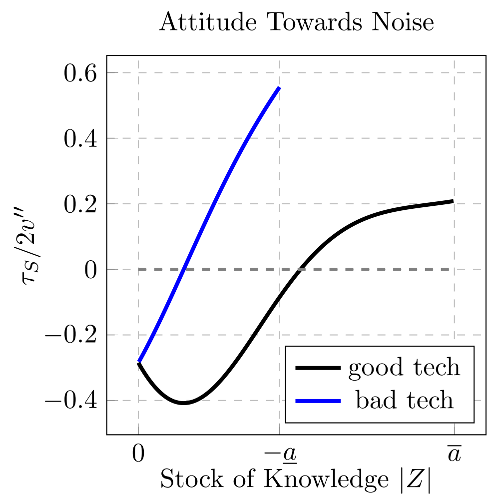

Figure 3: Attitude towards noisy experimentation over the knowledge cycle. This figure plots the second component of the expected marginal benefit of knowledge in Eq. (18), $\tau_S/2v''$, as a function of knowledge, separately for good (black line) and bad (blue line) technologies. Parameter values correspond to the baseline calibration, defined in Appendix A.2.

To examine how the firm's attitude towards noisy experiments affects investment over the cycle, Figure 4 contrasts the relation between investment and knowledge in the benchmark (solid line, a scaled version of Figure 1) to that with experimentation (dashed red line), separately for good (left panel) and bad (right panel) technologies. Focus first on the benchmark. Since neoclassical theory works in this case, investment is proportional to q: it is highest following recent exploration, and lowest late in the cycle. Furthermore, when the technology is likely good investment rises early in the cycle (q goes up), and then declines.

Activating experimentation (dashed red line) causes investment to rise late in the cycle. Early in the cycle, though total q dictates the shape of investment, the firm's reluctance towards experimentation noise flattens considerably the relation between investment and knowledge. As knowledge accumulates along the cycle the firm changes attitude and starts experimenting actively as a way to gamble on new technologies despite low total q. Figure 4 shows that this gambling attitude dominates late in the cycle and causes investment to rise. This effect is particularly important for bad technologies, and turns the neoclassical relation on its head: investment is decreasing in knowledge in the neoclassical case, but increasing in knowledge when investment is a means to experiment. For good technologies instead, we observe that investment is high not only early in the cycle—as in the benchmark model—but also late on, as exploration becomes more likely.

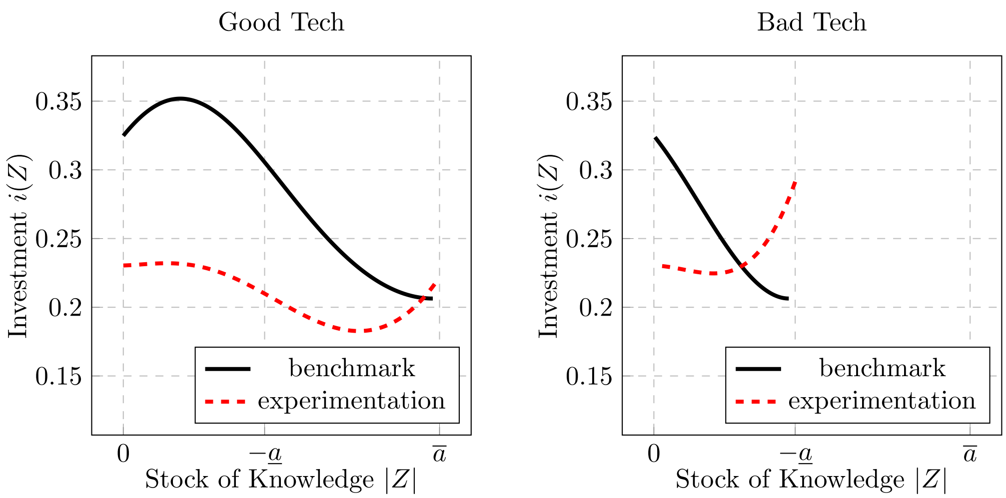

Figure 4: Knowledge-contingent investment with and without experimentation. This figure plots investment as a function of knowledge, Z, in the benchmark model of Section 3 (solid black) and in the model with experimentation of Section 4.4 (red dashed). The panel plots investment under good technologies; the right panel plots investment under bad technologies. Parameter values correspond to the baseline calibration, defined in Appendix A.2.

Finally, Figure 5 illustrates how key parameters of the model affect the relation between investment and knowledge. Making experiments more precise (a higher τ_S) may give full force to the firm's attitude towards noisy experiments, exacerbating the late rise in investment and reducing early investment (left panel). Another important parameter of the model is knowledge dissipation. When knowledge is easier to protect (lower ϕ), the investment pattern over the cycle is qualitatively similar, but the investment level is substantially higher.

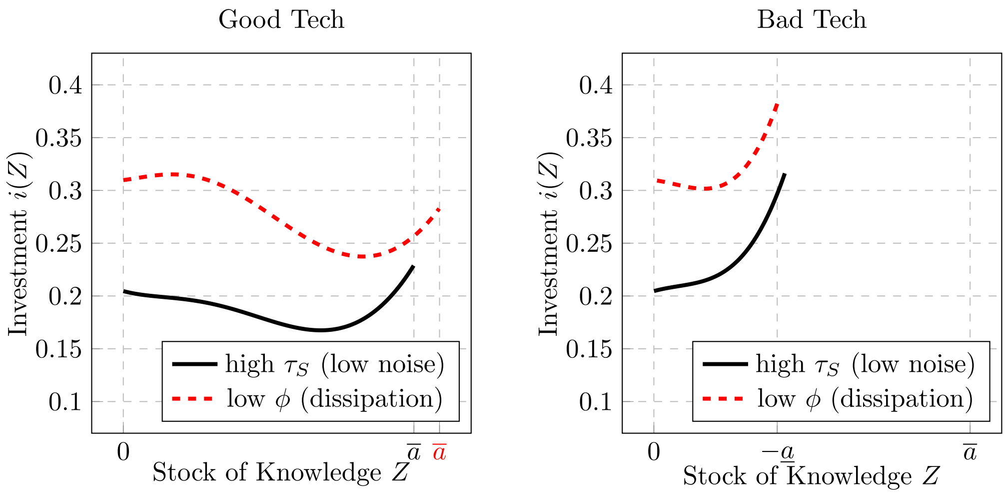

Figure 5: Investment and knowledge frictions. This figure plots investment as a function of knowledge, Z, with experimentation for alternative parameter values, separately for good (left panel) and bad (right panel) technologies. The blue line corresponds to the case of low experimentation noise ($\tau_S = 0.45 > 0.3$). The red line corresponds to the case of low knowledge dissipation ($\phi = 0.3 < 0.35$). Parameter values are defined in Appendix A.2.

4.2 Relation between investment and q over the knowledge cycle

Firm's attitude towards experimentation noise creates a disconnect between investment and total q, which causes the relation between the two to vary along the cycle in a specific way. In particular, experimentation causes total q and investment to rise late in the cycle but weakens their relation early on. Knowledge resets identify a shift in investment$-q$ sensitivity, with a spike before the reset and a drop upon and following it. Although there are alternative ways of creating a disconnect between q and investment, in Section 5.4.2 we show that these are unlikely to explain this specific pattern.

We approach the relation between investment and q in a way that can be replicated in empirical tests. From the firm's perspective, the relevant statistic for investment is the sum of traditional q and the expected marginal benefit of knowledge. Alternatively, this sum can be expressed as total q and the firm's attitude towards noise, as seen in Eq. (20). The problem is that, unlike traditional q, the expected marginal benefit of knowledge cannot be directly constructed from the data. Empirical proxies for total q do exist (Peters and Taylor, 2017), but measuring attitude towards experimentation noise seems difficult (e.g., it requires averaging value over all possible experimentation outcomes before the firm invests). Yet we know from the model that this attitude matters most and changes in opposite directions around knowledge resets. How to measure knowledge resets is itself an empirical challenge, which we tackle in Section 5. For the present section suppose these resets are available to the empiricist. Therefore, to assess the relevance of firm's attitude towards noise we estimate investment$-q$ regressions, which are common in empirical studies, but examining specifically how the relation evolves around knowledge resets.

We simulate daily time series of total q and investment, which are long enough ($8,000$ years) to cover many different knowledge cycles ($\approx 2,460$ cycles). Because we use patent data to measure knowledge resets in the data, we note that it is unlikely that bad technologies are patented, and we remove all cycles that revealed a technology to be poor, which amount to about half of the sample. Every time the firm explores a new technology ($Z \notin \mathscr{A}$), we flag the exploration date (ν). To map the data frequency in later tests we aggregate investment and total q annually:

$$I_t = \int_t^{t+1} i_s \mathrm{d}s \quad \text{and} \quad Q_t = \int_t^{t+1} q_{\mathrm{tot},s} \mathrm{d}s.$$

We keep track of whether a knowledge reset occurred over the year using the dummy variable:

$$D_t^{Reset} = \mathbf{1}_{\nu \in [t,t+1)}.$$

To examine how the relation between investment and total q evolves around resets, we track this relation 5 years preceding and following the reset by estimating the specification:

$$I_t = \alpha + \beta Q_t + \sum_{k=-5}^{+5} \gamma_k D_{t+k}^{Reset} Q_t + \sum_{k=-5}^{+5} \delta_k D_{t+k}^{Reset} + \epsilon_t, \tag{21}$$

where $k \in \{-5, ..., 5\}$ denotes the number of years preceding or following the reset year, t. The coefficient β captures the average level of investment$-q$ sensitivity; each coefficient γ_k captures variation in sensitivity relative to this level k years before or after the reset. Figure 6 reports the sensitivity, $\beta + \gamma_k$, in event time k around the reset ($k = 0$).

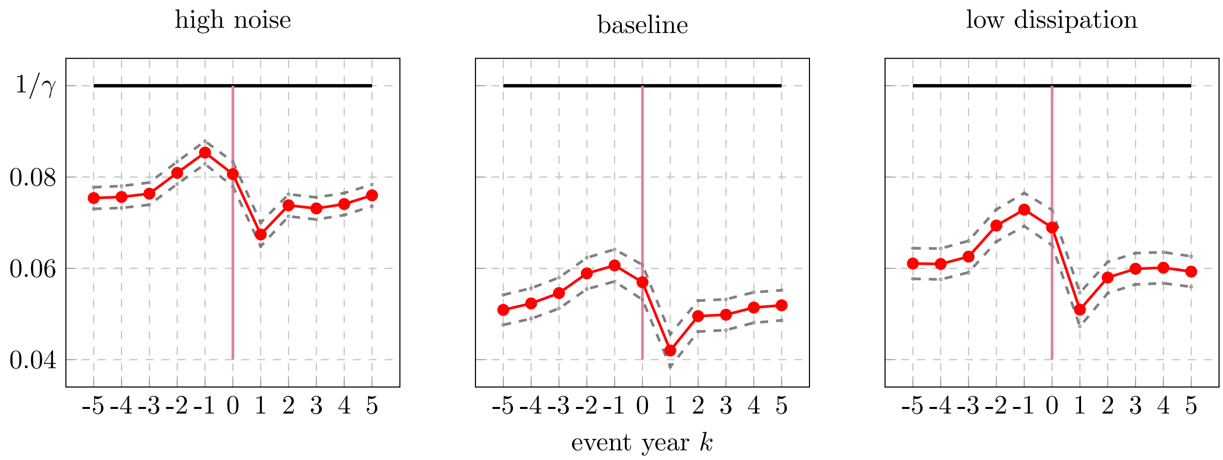

Figure 6: Investment$-q$ sensitivity in event time around a knowledge reset. Each panel plots the sensitivity of investment to total Q within the five years before and after the reset event, $k = 0$. We plot the sensitivity in the benchmark (flat black line) and with experimentation (red line). The middle panel corresponds to the baseline calibration (Appendix A.2), the left panel considers high experimentation noise ($\tau_S = 0.25 < 0.3$), and the right panel considers low knowledge dissipation ($\phi = 0.3 < 0.35$). Dashed gray lines correspond to the confidence intervals at the 95% level.

Recall that benchmark investment$-q$ sensitivity is constant and equal to $1/\gamma$ throughout the cycle (the flat black line in Figure 6). Note further that this prediction based on total q would carry over to experimentation if the firm's attitude towards noisy experiments did not matter. However, when this attitude matters, total q is not a sufficient statistic for investment—and the relation between the two must exhibit variation around resets, as is clearly apparent from the red lines in Figure 6.

Variation in the strength of the investment$-q$ relation follows a specific pattern around resets—the relation is weaker early in the cycle and stronger late in the cycle. The firm's attitude towards noise weakens the relation between investment and total q early in the cycle ($v'' < 0$). When uncertainty about the current technology is high, the firm prefers to avoid experimentation risk and invests prudently when in fact total q is highest. Late in the cycle, the firm gambles instead on new technologies by raising investment, when simultaneously knowledge q causes total q to rise. This pattern is robust to alternative values of key model parameters: it is more pronounced either when experiments are less informative or when knowledge is easier to protect (see left and right panels in Figure 6).

4.3 What is the length of a knowledge cycle?

Let $\tau_x = \inf\{t > \nu. : Z_t = x\}$ be the first time knowledge hits level x after switching to a new technology. The probability that the next exploration time for good technologies occurs before or at time $t \geq 0$ immediately following exploration (i.e., $Z_0 = 0$) is:

$$\mathbb{P}\left(\tau_{\overline{a}} \leq t | Z_0 = 0, \tau_{\overline{a}} < \tau_{\underline{a}}\right) \equiv \underbrace{\mathbb{E}\left(\mathbf{1}_{\tau_{\overline{a}} \leq t} \mathbf{1}_{\tau_{\overline{a}} < \tau_{\underline{a}}} | Z_0 = 0\right)}_{\equiv G(t;0)} / \underbrace{\mathbb{E}\left(\mathbf{1}_{\tau_{\overline{a}} < \tau_{\underline{a}}} | Z_0 = 0\right)}_{\equiv g(0)}.$$

For an arbitrary starting point $Z \in \mathscr{A}$ within the experimentation region, standard arguments show that $G(t, Z)$ solves the Kolmogorov Backward Equation (KBE):

$$G_t = \frac{1}{2}(\tau_A + \tau_S i(Z))(G_Z Z + G_{ZZ}),$$

with initial condition $G(0, Z) = 0$ and boundary conditions $G(t, \underline{a}) = 0$ and $G(t, \overline{a}) = 1$.[13] The distribution of a cycle length is $\frac{\mathrm{d}}{\mathrm{d}t} G(t; 0)/g(0)$ and is plotted in Figure 7 with and without experimentation (left panel) and with experimentation only but when knowledge is more difficult to protect (right panel).

All curves in Figure 7 suggest that knowledge cycles are short, with a median of approximately 6 years. Casual intuition may suggest otherwise. For instance, the cycles associated with *Intel*'s knowledge about RAM chips or central processor technology lasted around ten years. Thus, a median of 6 years seems substantially "shorter." However, we will show empirically in Section 5 that knowledge cycles are consistently short in the data.

Experimentation leads to shorter knowledge cycles (the left panel of Figure 7). Note that we have chosen parameter values in the benchmark model so that the benchmark exploration thresholds coincide with those under experimentation. Hence, the shortening of the cycle is a unique consequence of the learning benefits of active experimentation. Because experimentation allows the firm to learn faster about the quality of its technology, it shortens the average and median length of the cycle. Considering the substantial effect of experimentation on investment relative to neoclassical predictions, however, this effect is relatively small; this is because experimentation significantly reduces the initial level of investment due to the firm's reluctance to take experimentation risk early in the cycle.

[13] Similarly, $g(Z)$ solves $0 = (\tau_A + \tau_S i(Z))(g_Z Z + g_{ZZ})$, with boundary conditions $g(\overline{a}) = 1$ and $g(\underline{a}) = 0$.

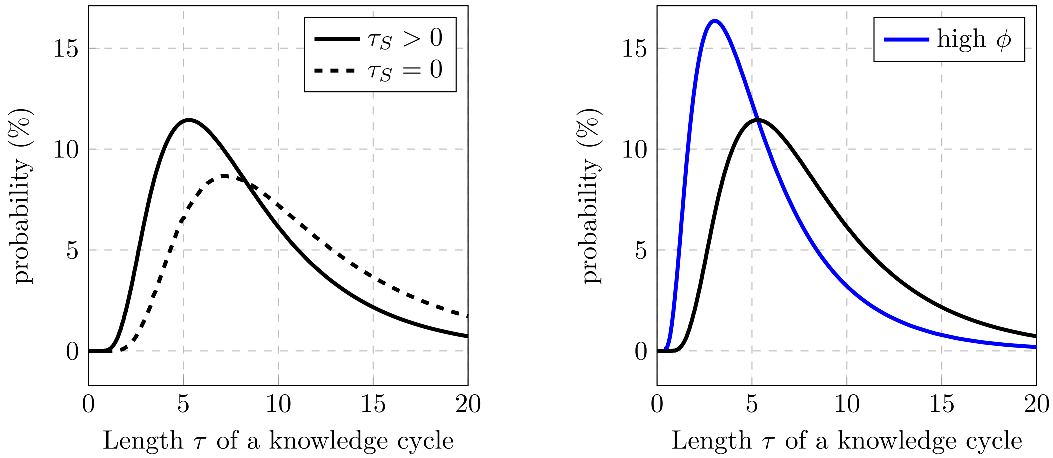

Figure 7: Distribution of the length of a knowledge cycle for good technologies. The left-hand panel plots the distribution of the length of the cycle with (solid line) and without (dashed line) experimentation. The right-hand panel does comparative statics, with the black curve denoting the case with experimentation, and the blue curve corresponding to the case with experimentation and high knowledge dissipation ($\phi = 0.8 > 0.35$). Parameters are reported in Appendix A.2

It is intuitive that knowledge dissipation, for instance in the form of competitive pressures, should reduce the length of knowledge cycles. In the case of RAM chips, competition from Japan precipitated to a large extent the end of *Intel*'s knowledge cycle. Alternatively, when property rights are harder to protect the firm's incentive to explore a new technology are higher. The right panel of Figure 7 shows that high knowledge dissipation (the blue line) reduces the length of the knowledge cycle, a prediction that we confirm later in the data.

4.4 Model extension: partially transferable knowledge

It is likely that knowledge acquired about a technology is in fact partially transferable to another. Thus, we now relax our baseline assumption that knowledge is fully reset upon exploration. In the spirit of Jovanovic and Nyarko (1996), we assume that upon exploring a new technology (which occurs at time ν) knowledge is reset to:

$$Z_\nu = \lambda \cdot Z_{\nu-}, \tag{22}$$

where $\lambda \in [0, 1)$. When $\lambda \equiv 0$, knowledge is specific to a technology and we recover the model of Section 2. When $\lambda > 0$, knowledge is partially transferable across technologies.

The case $\lambda \equiv 1$ in which knowledge is freely transferable across technologies does not lead to a stationary solution and must be discarded.

The solution method is identical to the baseline, except for the boundary conditions that apply when exploration is optimal. Value-matching conditions are now given by:

$$\lim_{Z \downarrow \underline{a}} v(Z) = (1-\omega)^{\alpha(1-\eta)} v(\lambda \cdot \underline{a}) \quad \text{and} \quad \lim_{Z \uparrow \overline{a}} v(Z) = (1-\omega)^{\alpha(1-\eta)} v(\lambda \cdot \overline{a}).$$

Smooth-pasting conditions are modified similarly.[14] Figure 8 compares investment under partially transferable knowledge ($\lambda > 0$, dashed lines) to that when knowledge is technology-specific ($\lambda = 0$, solid lines) separately for good and bad technologies.

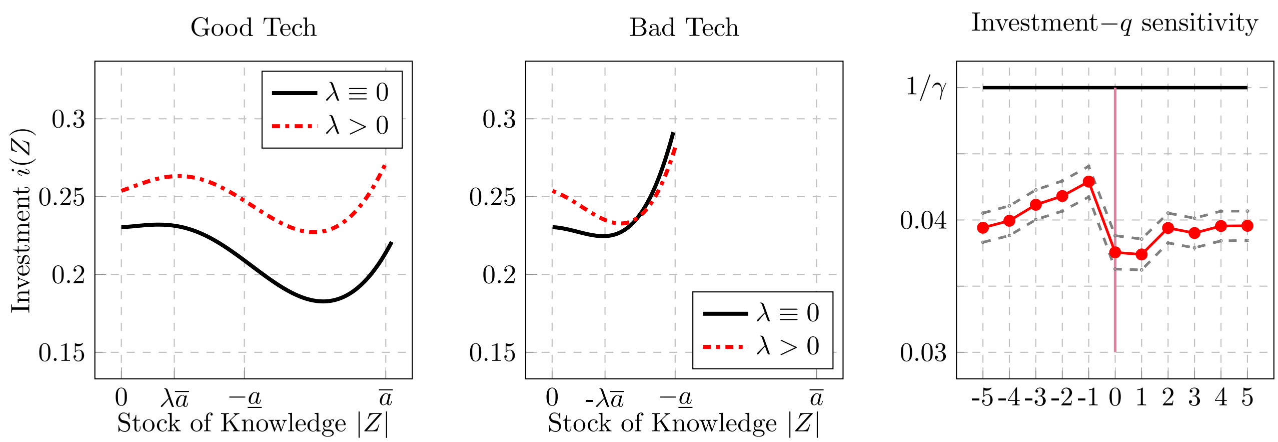

Figure 8: Investment and its sensitivity to total q when knowledge is partially transferable across technologies. The left-hand and center panels plot investment as a function of knowledge, Z, with partially transferable ($\lambda = 0.2$, dashed lines) and with technology-specific ($\lambda = 0$, solid lines) knowledge for good (left panel) and bad (center panel) technologies. The right-hand panel plots investment-q sensitivity according to the simulation procedure of Section 4.2 5 years before and after a reset event, $k = 0$, but with $\lambda = 0.2$. Parameter values are defined in Appendix A.2.

The main insights of the model remain, although the asymmetry between good and bad technologies becomes stronger. When the firm uncovers a good technology, it is now more likely to uncover another good one in the next cycle. Thus, the firm has greater incentives to invest when it is able to partially reuse its existing knowledge. The same argument implies that investment is lower for bad technologies relative to its level under technology-specific knowledge, except early in the cycle when partial transferability of knowledge raises total q.

[14] $\lim_{Z \downarrow \underline{a}} v'(Z) = \lambda \cdot (1-\omega)^{\alpha(1-\eta)} v'(\lambda \cdot \underline{a})$ and $\lim_{Z \downarrow \overline{a}} v'(Z) = \lambda \cdot (1-\omega)^{\alpha(1-\eta)} v'(\lambda \cdot \overline{a})$.

Around knowledge resets, the relation between knowledge and investment and the evolution of investment$-q$ sensitivity (right panel) remain qualitatively unchanged.

5 Empirical tests

5.1 Measuring knowledge resets

To test the model's main predictions, we measure knowledge resets using firms' patents. We obtain patent data from the United States Patent and Trademark Office (USPTO) PatentView platform. We consider all utility patents applications granted between 1976 and 2017 and link assignees to public firms in Compustat using the bridge file provided by NBER until 2006, and that extended by Stoffman, Woeppel, and Yavuz (2021) for later years. Central to our measurement of knowledge resets, the data contains all the citations each patent makes to prior patents, that we consider as the prior art that each patent is building upon. We only consider firms that were granted at least one patent over the 1976-2017 period, and require available data on total investment and total Q, as defined by Peters and Taylor (2017). Firms enter the sample when they apply for their first patent over that period and leave three years after their last patent grant. The sample comprises 1,442,813 distinct patent applications granted to 5,428 firms, or 54,250 firm-year observations. On average firms are granted 26 new patents per year (the median is 2), and each patent makes on average 15 citations to prior patents (with a median of 7).

We define firm f technological "knowledge base" at time t as the set of all the patents cited by the patents applied for by f between $t-5$ and t.[15] Following Ma (2021), this set captures the underlying technological knowledge that f has accumulated between $t-5$ and t. On average, firms' knowledge base is composed of 148 (median of 11) distinct patents cited from 133 patents (median of 11). To track the evolution of f's knowledge base, we aggregate the cited patents across the 38 distinct NBER patent technology sub-classes. Specifically, we define the vector v_t^f where each of its 38 elements represents the share of f's knowledge base in year t corresponding to each class. Intuitively, v_t^f represents the distribution of knowledge accumulated by f in year t across distinct technology areas. The innovation of some firms

[15] We consider patent application year instead of grant year to better reflect the true timing of firms' innovation. While we consider a period of 5 years to define the knowledge base, we obtain similar results if we consider longer periods of 10 or 15 years.

builds on patents from a few technology areas whereas other firms have a more dispersed knowledge base.

Then, we measure annual changes in f's' knowledge base from the cosine similarity between v_t^f and v_{t-1}^f according to:

$$\Delta v_t^f = \frac{v_t^f}{\|v_t^f\|} \cdot \frac{v_{t-1}^f}{\|v_{t-1}^f\|},$$

where the operator "$\cdot$" denotes the scalar product, and "$\|v\|$" denotes vector v's Euclidean norm. By definition, Δv_t^f is bounded in $[0, 1]$. Higher values of Δv_t^f reflect greater similarity between v_t^f and v_{t-1}^f, and indicate stability in the knowledge base. In contrast, lower values of Δv_t^f indicate instability. Knowledge base is instable when the technology developed by f in year t builds on patents from different technological classes compared to what was used previously. Hence, $\Delta v_t^f = 0$ indicates that new technology in year t builds on technologies that were not used until year $t - 1$. Table 1 (Panel A) shows that, perhaps unsurprisingly, firms' knowledge base is remarkably stable since the average value of Δv_t^f is 0.96 and the median is 0.99. The sample standard deviation of Δv_t^f is 0.10, indicating that firms' innovations typically build on a stable distribution of technologies.

To identify knowledge "resets", we flag years in which Δv_t^f is abnormally low, using the following binary variable(s):

$$Reset_t^f(\theta) = \begin{cases} 1 & \text{if } \Delta v_t^f < \overline{\Delta v^f} - \theta \times \sigma(\Delta v^f) \\ 0 & \text{otherwise,} \end{cases}$$

where $\overline{\Delta v^f}$ and $\sigma(\Delta v^f)$ are the time-series average and standard deviation of Δv_t for firm f, and $\theta \in (1, 1.5, 2, 2.5)$. A reset occurs when Δv_t^f is θ standard deviations below its average. Table 1 indicates that, across all firms and years, knowledge resets are infrequent as their occurence, the average of $Reset(\theta)$, ranges between 3.8% ($\theta = 2.5$) and 10.8% ($\theta = 1$). Although infrequent, the change in knowledge base during resets is sizable, as Δv_t^f ranges between 0.67 and 0.77, implying roughly a 30% change in the knowledge base.

To build intuition, we detail the knowledge resets of two firms in our sample.[16] The first

[16]Selected among the 25 largest firms with a least 10 patents forming their knowledge base and experiencing large resets during the sample period ($\theta = 2.5$)

Variable	mean	sd	p25	p50	p75	N
		Panel A: firm-year variables				
Δv	0.964	0.107	0.982	0.997	0.999	54,250
Reset(1.0)	0.108	0.310	0	0	0	54,250
Reset(1.5)	0.081	0.273	0	0	0	54,250
Reset(2.0)	0.058	0.234	0	0	0	54,250
Reset(2.5)	0.039	0.193	0	0	0	54,250
Δv|Reset(1.0)	0.779	0.229	0.678	0.859	0.961	5,827
Δv|Reset(1.5)	0.745	0.245	0.596	0.826	0.950	4,353
Δv|Reset(2.0)	0.708	0.258	0.535	0.780	0.930	3,107
Δv|Reset(2.5)	0.682	0.264	0.500	0.751	0.911	2,076
Total Q	1.194	1.949	0.223	0.615	1.345	54,250
Total Investment	0.236	0.185	0.124	0.183	0.278	54,250
Physical investment	0.066	0.080	0.022	0.042	0.079	54,250
Intangible Investment	0.169	0.143	0.076	0.130	0.213	54,250
		Panel B: Cycle variables				
$\tau_{data}(1.0)$	6.404	5.348	3	5	8	1,885
$\tau_{data}(1.5)$	7.526	6.429	3	6	9	1,095
$\tau_{data}(2.0)$	8.387	7.273	3	6	11	568
$\tau_{data}(2.5)$	9.674	8.298	3	7	13	227

Table 1: Summary Statistics. This table presents the summary statistics of all variables used in the empirical analysis. Panel A reports statistics for firm-year variables. Panel B reports statistics for cycle variables. The sample period is from 1976 to 2017.

is the business-to-business software company CA Technologies that experienced a knowlegde reset in 1997, and the second is the distributor of motor fuels Sunoco that experienced a reset in 1985. Table 2 displays the elements (larger than 5% for brevity) of both firms' vectors v for years t (the reset year), $t-1$, and $t-2$. Panel A reveals that CA's reset (Δv dropping from 0.99 in 1996 to 0.39 in 1997) is driven by an abrupt increase in its share of citations of patents in "computer hardware and software", "communication", and "peripherals", and "information storage", and a simulateneous decrease in citations from "material processing and handling", "coating", and "agriculture, food, and textiles". These changes reflect the company's strategy in the mid nineties to improve compatibility with products from other vendors, e.g., Hewlett-Packard or Apple Computer.[17] Panel B indicates that the knowledge reset of Sunoco (Δv dropping from 0.98 in 1984 to 0.73 in 1985) stems from its increased reliance on technologies related to "organic compounds", and a decreased

[17] See CA Technologies' history at: https://en.wikipedia.org/wiki/CA_Technologies.

NBER Sub-category	v_t	v_{t-1}	v_{t-2}
Panel A: CA Technologies (reset in t=1997)			
Agriculture, Food, Textiles	0.01	0.04	0.05
Coating	0.02	0.12	0.07
Miscellaneous	0.02	0.08	0.07
Communications	**0.05**	**0.00**	**0.00**
Computer Hardware & Software	**0.43**	**0.10**	**0.11**
Computer Peripherials	**0.05**	**0.00**	**0.00**
Information Storage	**0.14**	**0.00**	**0.00**
Electronic business methods and software	0.09	0.00	0.00
Mat. Proc & Handling	0.11	0.53	0.55
Δv	0.39	0.99	
Panel B: Sunoco (reset in t=1985)			
Gas	0.05	0.09	0.08
Organic Compounds	**0.28**	**0.03**	**0.01**
Resins	0.02	0.01	0.06
Miscellaneous	0.38	0.53	0.48
Drugs	0.07	0.04	0.03
Measuring & Testing	0.00	0.06	0.10
Mat. Proc & Handling	0.05	0.01	0.01
Heating	0.07	0.13	0.11
Δv	0.73	0.98	

Table 2: Examples of knowledge resets.

reliance on technologies related to "measuring and testing", and "heating", coinciding with its launch of the Sunoco ULTRA 94 in 1983, the market's highest octane unleaded gasoline.[18]

5.2 What is the length of knowledge cycles in the data?

We define the length of a cycle as the number of years between two resets ($\tau_{data(\theta)}$) for firms that have completed at least one cycle (two observable resets). This restricted sample comprises between 227 ($\theta = 2.5$) and 1,885 ($\theta = 1$) cycles for 190 and 1,242 distinct firms, respectively. Figure 9 displays the distributions of τ_{data} for the four values of the threshold θ. Mirroring the distribution in the model (see Figure 7), the empirical distributions are asymmetric and right skewed. Panel B of Table 1 indicates that the average cycle length varies between 6.40 and 9.67 years, and the median is between 5 and 7 years. Figure 10 reveals notable heterogeneity across broad (one-digit SIC) industries and confirms the prevalence

[18]See Sunoco's history at: https://en.wikipedia.org/wiki/Sunoco.

of skewed cycles (averages are larger than medians). The longest cycles appear in mining with an average close to 10 years, and the shortest in retail with less than 6 years. Although existing research provides no point of comparison that we are aware of, knowledge cycles appear rather short in our sample.

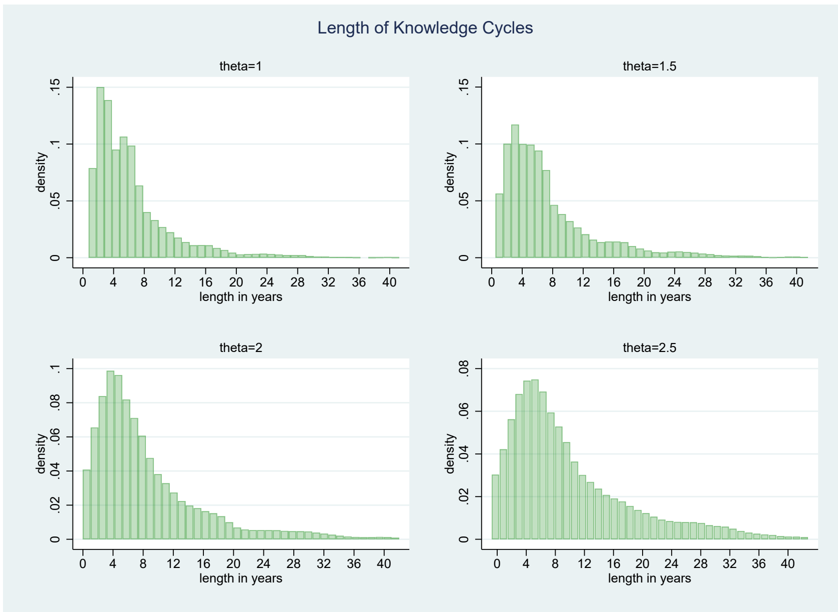

Figure 9: Empirical length of knowledge cycles. This figure plots the distribution of the length of knowledge cycles, where a cycle is given by the number of years between two knowledge resets. We consider resets based on four distinct values of θ.

Figure 11 displays length distributions separately for firms operating in competitive and concentrated industries, measured using the Herfindhal index (HHI) computed at the 3-digit SIC level below and above the sample median based on firms in Compustat. We use HHI as a proxy for the intensity of knowledge dissipation (ϕ), assuming lower dissipation in more concentrated industries (high HHI). In line with the model's prediction, Figure 11 reveals longer cycles for firms facing lower knowledge dissipation, with an average length of 8.37 years in concentrated industries compared to 6.69 years in competitive industries (the difference is statistically significant).[19] Figure 11 reveals that the intensity of knowledge dissipation is linked to the length of knowledge cycles, which reinforces the economic plausibility of our

[19]The model also predicts that the distribution of knowledge cycles should depend on the noise association with experimentation, τ_S. Unfortunately, we could not find a plausible empirical proxy for this parameter.

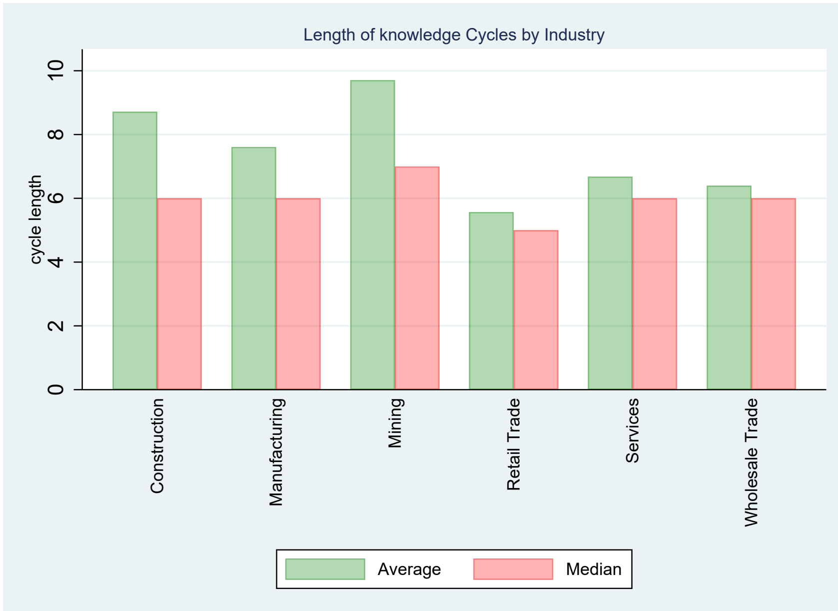

Figure 10: Empirical length of knowledge cycles across industries. This figure plots the average and median length of knowledge cycles across one-digit SIC industries, where a cycle is given by the number of years between two knowledge resets. We consider resets based on $\theta = 1.5$.

measurement of knowledge resets and cycles.

5.3 Investment and total q around knowledge resets

We examine the relation between investment and total q around knowledge resets (see Figure 6) using "total" investment from Peters and Taylor (2017) to capture the optimal investment rate $i(Z)$. Similar to $i(Z)$, total investment is the sum of capital (capex) and intangible (R&D+0.3×SG&A) expenditures divided by (lagged) total capital. Total capital is the sum of property, plant, and equipment, intangible assets on the balance sheet, research and organizational capital (accumulated R&D and SG&A spending). We use Peters and Taylor (2017)'s total Q as a proxy for the model's total q. Total Q is defined as firm value divided by total capital, which we view as a reasonable proxy for the sum of the model's traditional and knowledge q. Using simulated data, in untabulated tests, we obtain a correlation coefficient of 0.96 between the model's total q and that in Peters and Taylor (2017), and verify that

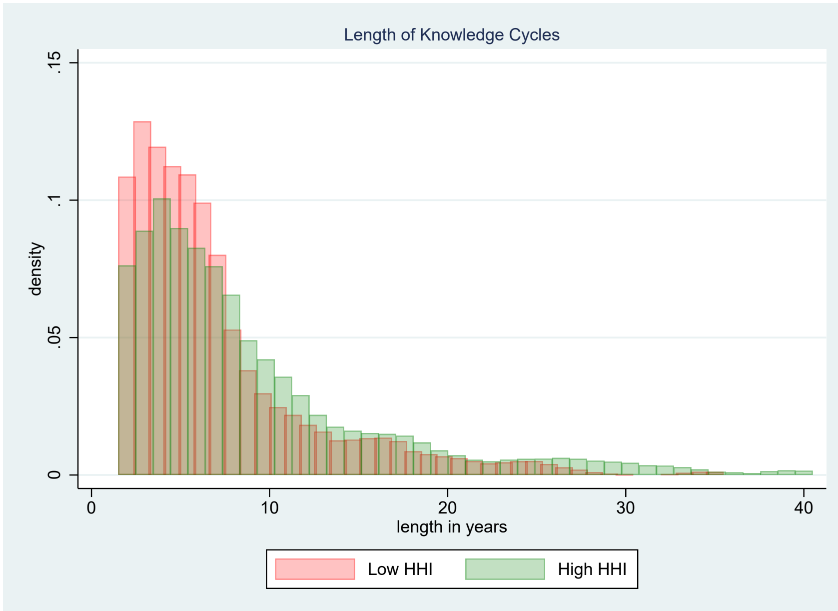

Figure 11: Empirical length of knowledge cycles by HHI. This figure plots the distribution of the length of knowledge cycles across firms in concentrated and competitive industries. Firms are assigned to groups based on whether the HHI of their industry is above or below the sample median. A cycle is given by the number of years between two knowledge resets. We consider resets based on $\theta = 1.5$.

the evolution of investment$-q$ sensitivity is virtually unchanged.[20]

Table 1 presents summary statistics. Average total Q is 1.280, with a median of 0.636. Total investment is 0.235 on average, with a median of 0.183. The average (median) physical investment is 0.065 (0.042) and the average (median) intangible investment is 0.168 (0.130). Overall, these statistics are close to those reported by Peters and Taylor (2017) on a larger sample that also includes non-patenting firms (see their Table 1). Firms in our sample invest more in intangible than in physical assets, which is expected as we focus on firms with patents.

To track the relation between total investment and total Q around resets we estimate the empirical counterpart to Eq. (21) in simulations:

$$I_{f,t} = \alpha_f + \eta_t + \beta Q_{f,t-1} + \sum_{k=-5}^{+5} \pi_k Q_{f,t-1} \times D_{t+k}^{Reset} + \sum_{k=-5}^{+5} \delta_k D_{t+k}^{Reset} + \varepsilon_{f,t}, \tag{23}$$

[20] Formally, using the notation of our model, the variable $\frac{vK}{K+Z}$ boils down to Peters and Taylor (2017)'s definition of total Q—up to a constant.

where $I_{f,t}$ is total investment of firm f in year t, and $Q_{f,t-1}$ is its lagged total Q. To closely map the model, we include firm fixed effects α_f to isolate the within-firm variation of investment and Q over resets. In addition, year fixed effects η_t absorb common variations across firms (e.g., investment or valuation booms or busts common to all firms in specific years). We cluster the standard errors at the firm level. The sample includes firms that experience, or not, a reset at some point. The indicators D^{Reset}_{t+k} identify years around resets, from 5 years before to 5 years after.[21] Hence, the coefficients π_k's on the interactions between $Q_{f,t-1}$ and D^{Reset}_{t+k} track the within-firm evolution of the investment$-Q$ sensitivity around resets. Because β captures average sensitivity outside periods surrounding resets or for control firms not experiencing resets, investment-Q sensitivity around resets is given by $\hat{\beta} + \hat{\pi}_k$ for $k \in [-5, +5]$, which Figure 12 plots (together with their 95% confidence bounds) for resets corresponding to $\theta = 1.5$.

Figure 12: Investment-Q sensitivity around knowledge resets. This figure plots the sensitivity of investment to Q ($\hat{\beta} + \hat{\gamma}_k$ for $k \in [-5, +5]$) in event-time around resets, with resets corresponding to $\theta = 1.5$.

Figure 12 reveals substantial variation in the relation between total investment and total Q around resets, which first confirms a disconnect between value and investment around

[21] Because firms could have several resets within short periods, this specification accounts for the fact that some years could be simultaneously before and after a reset.

resets. Second, and most importantly, this disconnect evolves in a way that closely mirrors the relation that experimentation implies. Specifically, resets mark a clear break in this relation, with investment-Q sensitivity increasing and spiking before the reset (at $k = -1$), and dropping following the reset (for $k \geq 0$). The evolution of investment-Q sensitivity is both economically and statistically significant. It rises to 0.05 just before resets, and reaches to 0.03 five years into the new cycle, implying a drop of 40% that is significant at 1%. Reassuringly, the magnitude of investment-Q sensitivities is in line with Peters and Taylor (2017). In their sample, sensitivity is 0.049 (see their Table 2), whereas it is 0.042 in ours.

The specific pattern of Figure 12 is quite puzzling if one ignores that investment creates knowledge. In particular, a strong (weak) relation between value and investment at the end (start) of a cycle is not easily explained by traditional arguments. Of course, there exists various ways of rationalizing a disconnect (e.g., financing fricitions) but we argue and show next that few imply the specific pattern of Figure 12. In particular, traditional origins of a disconnect are either not at play in the data around resets or imply patterns that are theoretically inconsistent with that of Figure 12. We take these counterfactuals as indicative that the economic force driving the disconnect in our model—firms' attitude towards noisy experiments—is at play in the data.

5.4 Are alternative channels at play?

In principle, the timing of knowledge resets could coincide with changes in firms' environment, unrelated to the value of knowledge or to firms' attitude towards experimentation noise, but related to their investment and Q. We consider two categories of alternative explanations, those that are mechanical and those that are based on other economic theories.

5.4.1 Mechanical channels

Our measure of knowledge resets could be affected by changes in acquisition or innovation strategies related to investment and Q. For instance, resets may arise following the purchase of firms with distinct patented technologies, or large discontinuities in patenting (e.g., the rising importance of secrecy). We track investment-Q sensitivity around resets seperately for firms that *(i)* were acquirers in the three years preceeding resets (or not), or *(ii)* experienced increase (or decrease) in the size of their patent portfolios at the time of resets.[22] Figure 13 shows that the dynamics of investment-Q sensitivity around resets ($\hat{\beta} + \hat{\pi}_k$ for $\theta = 1.5$) are similar for both groups, labeled as "Low" and "High." Thus our conclusions are unlikely due to changes in firms' acquisition or patenting activities around resets.

Figure 13: Investment-Q sensitivity around knowledge resets: robustness I. This figure plots the sensitivity of investment to Q ($\hat{\beta} + \hat{\gamma}_k$ for $k \in [-5, +5]$) in event-time around resets across groups. Firms are assigned to groups based on whether they acquired other firms in the last three years (left panel), or whether the change in the size of their patent portfolio is above of below the sample median. Resets correspond to $\theta = 1.5$.

5.4.2 Alternative economic theories

The relevance of knowledge resets for the relation between investment and value must arise from a disconnect between the two. Following Gutierrez and Philippon (2017) we focus on three origins of a potential disconnect: *(i)* financial frictions, *(ii)* competition, and *(iii)* institutional ownership. To this list we add *(iv)* the possibility that, in contrast to our standing assumption, investment in capital and intangibles are separate decisions, and *(v)* fixed adjustment costs, which are a known theoretical origin of the disconnect.

[22] Specifically, we re-estimate specification (23) but interact all variables with two binary variables identifying the two groups in which firms are assigned at the time of resets.

Financing frictions, competition and ownership. We capture financing frictions using the ratio of cash and marketable securities to assets (e.g., Fazzari, Hubbard, Petersen, Blinder, and Poterba (1988)). Furthermore, competition is another potential origin of the disconnect in the literature. We measure competition using industries' Herfindhal index computed at the 3-digit SIC level from Compustat. Finally, we measure ownership using the fraction of shares held by institutional investors. For each variable, we assign firms experiencing resets into two groups based on whether they are below or above the median taken across all resets, and track the investment-Q sensitivity seperately across groups. Figure 14 reveals no significant difference across groups, suggesting that the dynamics of investment-Q sensitivity around resets is unrelated to these channels.

At this stage a clarification regarding the effect of competition in the model is perhaps appropriate. In the model we capture competitive pressures by assuming that they exacerbate knowledge dissipation, which we parametrize with ϕ. Thus competition in this form is a building block of the model and cannot explain the relevance of knowledge resets in this context. Although competition affects knowledge cycles to a large extent (see Figure 7), absent a disconnect between investment and value it does not affect the relation between the two. And under the disconnect experimentation creates, competition merely shifts the overall level of the relation leaving its specific pattern unaffected (see the right panel of Figure 6). We conclude that the data supports this intuition.

As an alternative robustness exercise, we include all these variables along with those of Section 5.4.1 (whether a firm is an acquiror, the number of patents in its portfolio, cash, HHI and institutional ownership) in Eq. (23) as controls, both in levels and interacted with the indicators D_k^{Reset}. The lower-right panel of Figure 14 indicates that the inclusion of these variables does not affect the dynamics of the investment-Q sensitivity around resets (despite significant coefficients for some of them, not displayed for brevity).

Capital and knowledge as separate investment decisions. The key assumption we make is that knowledge is a by-product of investment in capital. Consider instead what happens in the model if investment in capital and knowledge are in fact separate decisions. In this case the firm solves:

$$\max_{\{\nu_n\},i_{K,t+s}\geq 0,i_{Z,t+s}\geq 0} \mathbb{E}\left[\int_0^\infty e^{-rs}\left(\Pi\left(A_{t+s},K_{t+s},N_{t+s}\right)-\gamma_{t+s}\right)\mathrm{d}s\,\middle|\,\mathscr{F}_t\right] \tag{24}$$

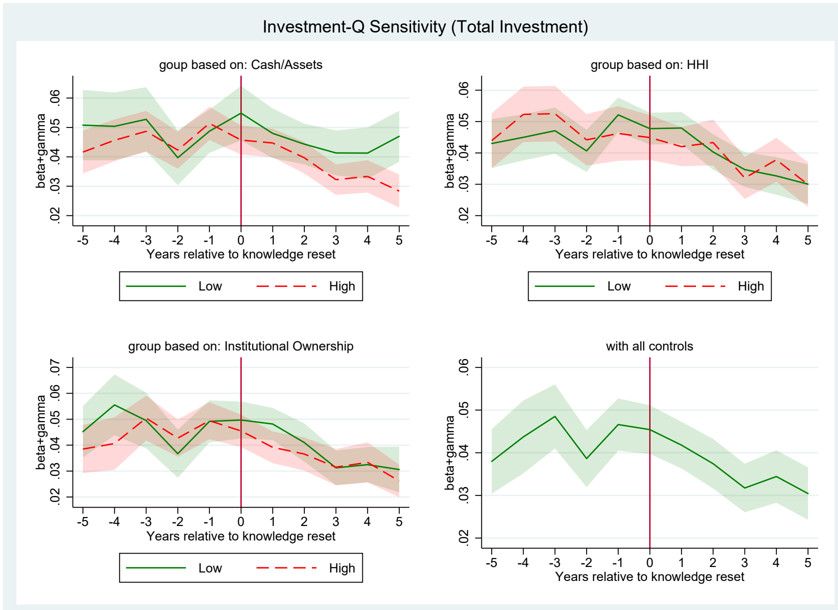

Figure 14: Investment-Q sensitivity around knowledge resets: robustness II. This figure plots the sensitivity of investment to Q ($\hat{\beta} + \hat{\gamma}_k$ for $k \in [-5, +5]$) in event-time around resets across groups. Firms are assigned to groups based on whether their cash-to-asset ratio (upper-left panel), industry HHI (upper-right panel), or institutional ownership (bottom-left panel) is above of below the sample median. In the bottom-right panel, we display the sensitivity of investment to Q after controlling for whether a firm is an acquiror, the number of patent in its portfolio, cash, HHI and institutional ownership). Resets correspond to $\theta = 1.5$.

$$\begin{aligned} \text{s.t. } \mathrm{d}Z_t &= \tau_A/2Z_t\mathrm{d}t + \tau_A^{1/2}\mathrm{d}\widehat{B}_t + \tau_S i_{Z,t}/2Z_t\mathrm{d}t + \tau_S^{1/2} i_{Z,t}^{1/2}\mathrm{d}\widehat{B}_{S,t} - Z_{t-}\mathbf{1}_{t=\nu} \\ \mathrm{d}K_t &= (i_{K,t} - \delta)K_t\mathrm{d}t - \omega K_{t-}\mathbf{1}_{t=\nu}, \end{aligned}$$

where the dynamics of A and N remain unchanged and adjustment costs are modified to:

$$\gamma_t \equiv \left(i_K + p_Z i_Z + \gamma/2(i_K - \delta)^2 + \gamma_Z/2i_Z^2\right) \cdot \Pi_t,$$

to reflect that knowledge has adjustment costs of its own (and unrelated to δ). We relegate the solution method to the appendix and plot the two separate investment policies in Figure 15. The pattern for value is similar to that in the baseline model and thus not reported.

Unsurprisingly, given that value is a sufficient statistic for investment in capital, its pattern (left panel) is very similar to that of Section 3. Furthermore, we know from Eq. (18) that the benefits of knowledge, c, are a bundle of two opposite forces, knowledge q and

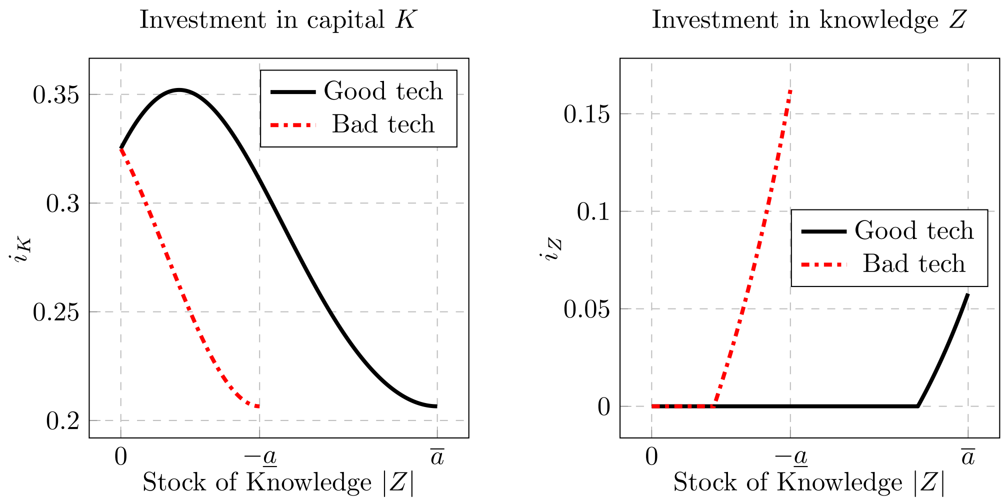

Figure 15: Investment in capital and knowledge when optimized separately. This figure plots investment in capital K (left panel) and in knowledge Z (right panel) as functions of knowledge, Z, separately for good (solid black) and bad (dashed red) technologies. Parameter values correspond to the baseline calibration, defined in Appendix A.2 with $\gamma_Z \equiv \gamma$ and $p_Z \equiv 0$.

attitude towards noise. Early in the cycle knowledge q is positive but investment is risky, which dominates, $c < 0$, so that the firm chooses to be "knowledge inactive." Late in the cycle knowledge q is negative but investment becomes a form of gambling, which dominates, $c > 0$, and triggers investment in knowledge. The point is that separating the two forms of investment creates inaction regions for investment in knowledge and thus a disconnect between investment and value. Therefore, we now reiterate the simulations of Section 4.2 and plot the sensitivity of the two kinds of investment to total q in Figure 16.

Two patterns are clear. Sensitivity of investment in knowledge to total q is weak and negative, which is due to intangible investment occuring at the end of the cycle when total q is slowly declining. In contrast, value and investment in capital, and thus their relation, follow similar patterns to those that prevail when investment creates knowledge. We conclude that the two patterns of sensitivity across the two kinds of investment differ when investment does not create knowledge. To verify whether this is the case empirically, we reiterate these regressions in the data. Figure 17 indicates that the patterns associated with distinct investment decisions are not supported in the data, with sensitivity of investment in capital knowledge following similar patterns instead. This finding suggests that the two decisions are likely entangled, and should be modeled together. Our standing assumption that investment

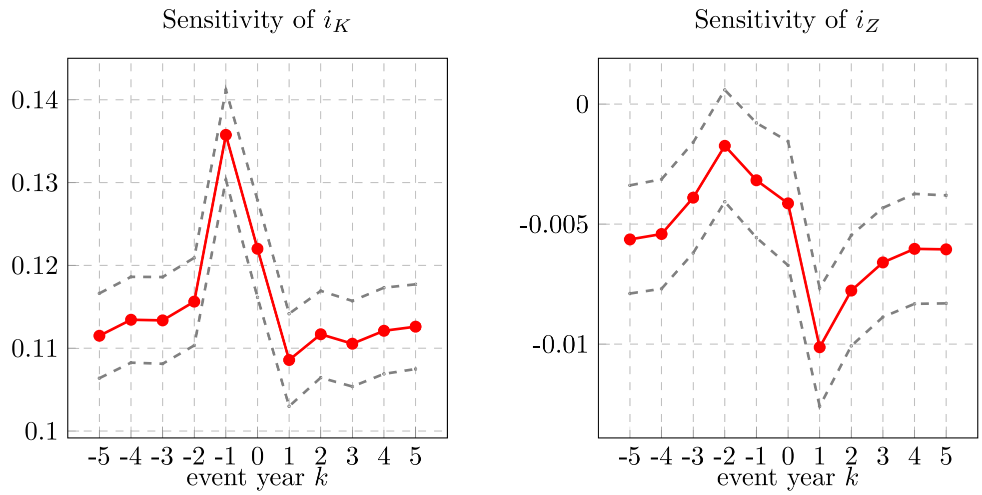

Figure 16: Sensitivity of investment in capital and knowledge to total q when optimized separately. This figure plots sensitivity of investment in capital (left panel) and investment in knowledge (right panel) to total q in event time in a 5-years window around knowledge resets, $k = 0$; Parameter values correspond to the baseline calibration, defined in Appendix A.2 with $\gamma_Z \equiv \gamma$ and $p_Z \equiv 0$.

generates knowledge is not only consistent with this evidence, but has also been validated empirically in a different context (Braguinsky et al., 2021).

Fixed adjustment costs. Another well-known cause of inaction are fixed adjustment costs (Caballero and Leahy, 1996), in which case investment is no longer driven by marginal q. To examine whether the resulting disconnect can explain the main fact we document, we now consider the benchmark framework of Section 3 (that is without experimentation, $\tau_S \equiv 0$) and modify the structure of adjustment costs to:

$$\gamma_t \equiv \left(i + F + \gamma/2(i - \delta)^2\right) \cdot \Pi_t, \tag{25}$$

where $F > 0$ is a fixed cost that is paid per unit of time, dt (Abel and Eberly, 1994). We relegate methodological details to the appendix and plot the optimal investment policy in Figure 18. The pattern for value is similar to that in the baseline model and thus not reported.

As desired, the left panel shows that fixed costs create an inaction region and thus a disconnect between value and investment. Early in the cycle when q is highest the firm follows an investment strategy that is similar to that of the benchmark. However, halfway

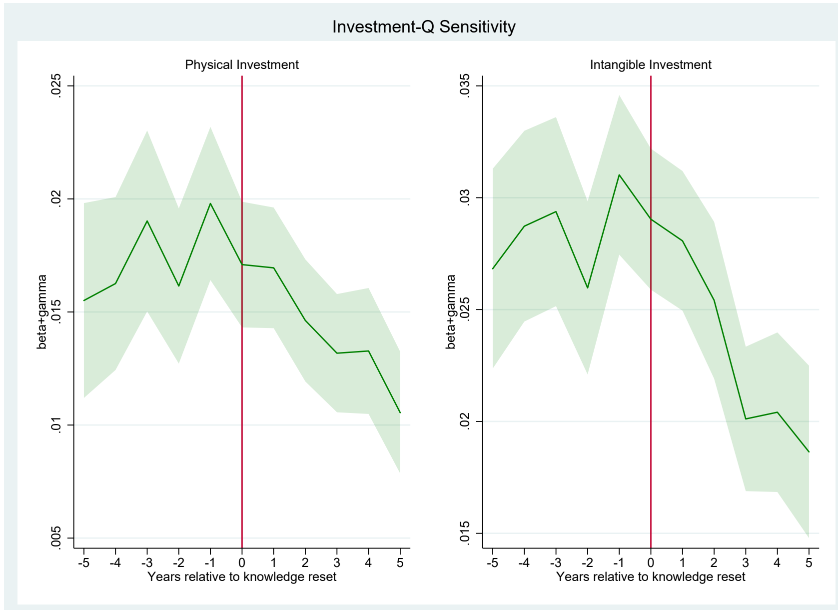

Figure 17: Investment-Q sensitivity around knowledge resets. This figure plots the sensitivity of physical (left panel) and intangible (right panel) investment to Q ($\hat{\beta} + \hat{\gamma}_k$ for $k \in [-5, +5]$) in event-time around resets, with resets corresponding to $\theta = 1.5$.

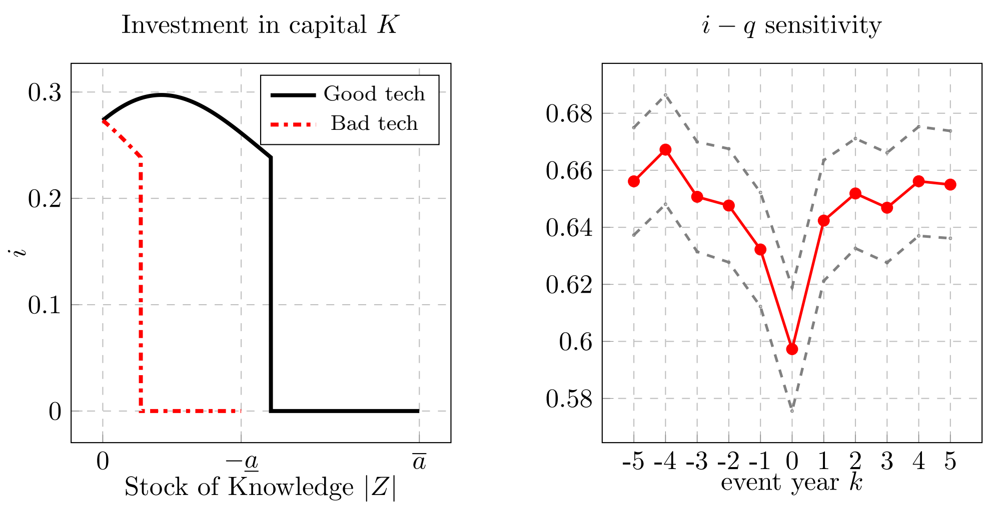

Figure 18: Investment and investment$-q$ sensitivity in the presence of fixed costs. The left panel plots investment in capital K as a function of knowledge, Z, separately for good (solid black) and bad (dashed red) technologies. The right panel plots its sensitivity to marginal q in event time in a 5-years window around knowledge resets, $k = 0$; dashed gray lines correspond to the confidence intervals at the 95% level. Parameter values correspond to the baseline calibration, defined in Appendix A.2 with $F \equiv 0.1$.

through the cycle q becomes too low to justify paying the fixed costs and investment ceases altogether. The right panel reiterates the simulations of Section 4.2 and plots the sensitivity of investment to q. The resulting disconnect creates a pattern in sensitivity that is opposite to that we uncover in the data. Specifically, sensitivity weakens ahead of the reset and strengthens following it. Intuitively, fixed costs deter investment late in the cycle and thus cannot explain a strong relation between investment and value preceding resets.

Other robustness checks. We subject our results to other checks unreported for brevity. The dynamics of investment-Q sensitivity holds if we *(i)* focus only on firms experiencing resets at some point, *(ii)* consider other values for the threshold θ, *(iii)* eliminate years before 1980 to limit the potential truncation in reset measurement, *(iv)* include the interaction between year and industry (3-digit SIC) fixed effects to absorb time-varying industry shocks (e.g., unobserved common technological changes), and *(v)* add firms' size (log of assets) and cash-flows as additional controls, mimicking typical investment specifications (e.g., Peters and Taylor (2017)). The post-reset decline in investment-Q sensitivity also remains in a difference-in-differences specification aggregating all D_k^{Reset} indicators into one capturing the post period $k > 0$ for firms experiencing resets, accounting for EIV in Q using the estimator of Erickson and Whited (2000).

6 Concluding Remarks

The relation between investment and Q drops when firms explore new technologies and reset their technological knowledge. This pattern is novel and consistent with a model in which investment is a means to accumulate capital, as is customary in the literature, but also a means to create knowledge. Investment as a means to experiment is high early in the cycle (after adopting a new technology) but also late in the cycle (just before exploring new technologies). The relation between investment and value is weaker early in the cycle and stronger late in the cycle. The model generates endogenous knowledge cycles that are short, and expand when knowledge is easier to protect. We find empirical support for these predictions, confirming that knowledge cycles are an important determinant of investment.

Knowledge resets have become less prevalent over time, as illustrated in Figure 19, re-

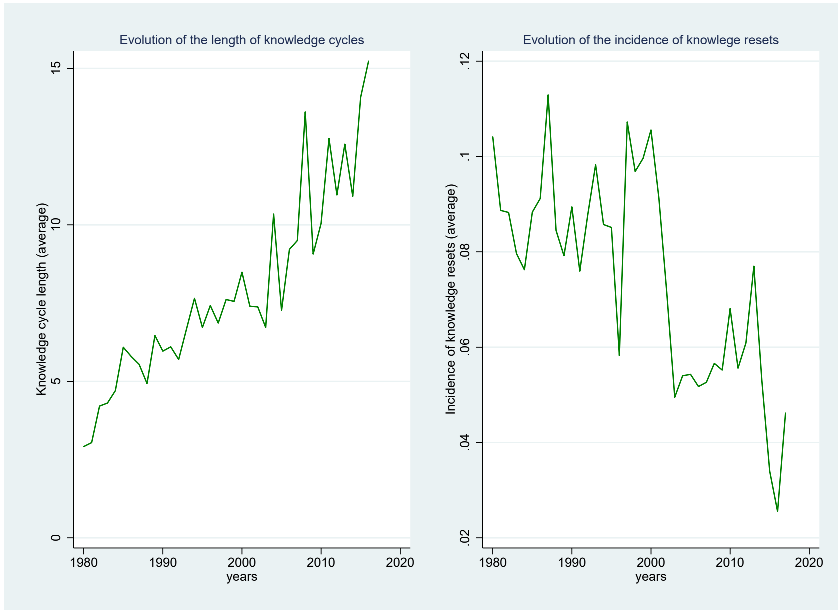

Figure 19: Evolution of cycles length and reset incidence. This figure plots the average length of knowledge cycles by year (left panel) and the average incidence of knowledge resets by year. Resets correspond to $\theta = 1.5$.

sulting in cycles whose length almost tripled between 1980 and 2017.[23] The lengthening of cycles (which we believe we are first to document) could have various origins, for instance, the increased difficulty to generate new ideas (Bloom et al., 2017) or declining product market competition (Gutierrez and Philippon, 2018). Since investment$-Q$ sensitivity is linked to knowledge cycles, it is interesting to contrast its evolution to that of knowledge cycles. Gutierrez and Philippon (2017) and Alexander and Eberly (2018) document a decline in investment in the last twenty years, despite no decline in Q, which we confirm in our sample. Because the relation between investment and Q is strongest around resets (see Figure 12), its recent weakening could be connected to less frequent resets and the underlying changes in the process of knowledge creation within firms. We leave these questions for future research.

[23]We obtain similar evolutions if we consider other values for θ, and if we control for changes in the composition of the sample by removing firm fixed effects.

References

Abel, A. and J. Eberly (1994). A unified model of investment under uncertainty. *American Economic Review 84*(5), 1369–84.

Abel, A. and J. Eberly (2011). How q and cash flow affect investment without frictions: An analytic explanation. *Review of Economic Studies 78*(4), 1179–1200.

Abel, A. and J. Eberly (2012). Investment, valuation and growth options. *Quarterly Journal of Finance* (2), 5–30.

Acemoglu, D. (2009). *Introduction to modern economic growth.* Princeton University Press.

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991). Optimal learning by experimentation. *The Review of Economic Studies 58*(4), 621–654.

Alexander, L. and J. Eberly (2018). Investment hollowing out. *IMF Economic Review* (66), 5–30.

Andrei, D., W. Mann, and N. Moyen (2018). Why did the q theory of investment start working? *The Journal of Financial Economics*.

Arrow, K. (1962). The economic implications of learning by doing. *Review of Economic Studies 29*(3), 155–173.

Benkard, L. (2000, September). Learning and forgetting: The dynamics of aircraft production. *American Economic Review 90*(4), 1034–1054.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *The RAND Journal of Economics 36*(4), 719–752.

Berk, J., R. Green, and V. Naik (2004, 01). Valuation and return dynamics of new ventures. *The Review of Financial Studies 17*(1), 1–35.

Bloom, N., C. Jones, J. van Reenen, and M. Webb (2017). Are ideas getting harder to get? NBER Working Paper 23782.

Bloom, N., R. Sadun, and J. van Reenen (2016). Management as a technology? *Working Paper*.

Braguinsky, S., A. Ohyama, T. Okazaki, and C. Syverson (2021). Product innovation, product diversification, and firm growth: Evidence from japan's early industrialization.

Caballero, R. and J. Leahy (1996). Fixed costs: The demise of marginal q. *Unpublished Manuscript*.

Callander, S. (2011, October). Searching and learning by trial and error. *American Economic Review 101*(6), 2277–2308.

Casadesus-Masanell, R., D. Yoffie, and S. Mattu (2014). Intel corporation: 1968-2003. Harvard Business School Case 703-427, November 2002. (Revised February 2010.).

Cooper, I. (2006). Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *The Journal of Finance 61*(1), 139–170.

Crouzet, N. and J. Eberly (2018, May). Intangibles, investment, and efficiency. *AEA Papers and Proceedings 108*, 426–31.

Crouzet, N. and J. Eberly (2020). Rents and intangible capital: A q + framework.

De Loecker, J., J. Eeckhout, and G. Unger (2019). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, forthcoming.

Eisfeldt, A. and D. Papanikolau (2013). Organization capital and the cross-section of expected returns. *Journal of Finance 68*, 1365–1406.

Erickson, T. and T. Whited (2000). Measurement error and the relationship between investmentand ¡em¿q¡/em¿. *Journal of Political Economy 108*(5), 1027–1057.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp (2019, May). Big data and firm dynamics. *AEA Papers and Proceedings 109*, 38–42.

Fazzari, S. M., R. G. Hubbard, B. C. Petersen, A. S. Blinder, and J. M. Poterba (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity 1988*(1), 141–206.

Foster, L., J. Haltinwanger, and C. Krizan (2001). *Aggregate Productivity Growth: Lessons from Microeconomic Evidence*, Volume 1. University of Chicago Press, Chicago IL.

Glaeser, E., H. Kallal, J. Scheinkman, and A. Shleifer (1992). Growth in cities. *Journal of Political Economy* (100), 1126–1152.

Gomes, J. F. (2001, December). Financing investment. *American Economic Review 91*(5), 1263–1285.

Grossman, S., R. Kihlstrom, and L. Mirman (1977). A bayesian approach to the production of information and learning by doing. *The Review of Economic Studies 44*(3), 533–547.

Gutierrez, G. and T. Philippon (2017). Investmentless growth: An empirical investigation. *Brookings Papers on Economic Activity*, 89–169.

Gutierrez, G. and T. Philippon (2018). Declining competition and investment in the us. *Working Paper*.

Hackbarth, D. and T. Johnson (2015). Real options and risk dynamics. *Review of Economic Studies 82*(4), 1449–1482.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica 50*(1), 213–224.

Hennessy, C. A., A. Levy, and T. M. Whited (2007). Testing q theory with financing frictions. *Journal of Financial Economics 83*(3), 691–717.

Jovanovic, B. and G. MacDonald (1994). Competitive diffusion. *Journal of Political Economy* (102), 24–52.

Jovanovic, B. and Y. Nyarko (1996, November). Learning by Doing and the Choice of Technology. *Econometrica 64*(6), 1299–1310.

Jovanovic, B. and R. Rob (1990). Long waves and short waves: Growth through intensive and extensive search. *Econometrica 58*(6), 1391–1409.

Keller, G. and S. Rady (1999). Optimal experimentation in a changing environment. *Review of Economic Studies 66*(3), 475–507.

Kyle, A., A. Obizhaeva, and Y. Wang (2018). Smooth trading with overconfidence and market power. *Review of Economic Studies 85*(1), 611–662.

Lanjouw, J. and M. Schankerman (2001). Characteristics of patent litigation: A window on competition. *Rand Journal of Economics* (32), 129–151.

Lieberman, M. and S. Asaba (2006). Why do firms initate each other? *Academy of Management Review* (31), 366–385.

Lipster, R. and A. Shiryaev (2001). *Statistics of Random Processes II.* Springer Verlag, New York.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics* (22), 3–42.

Ma, S. (2021). Obsolete Firms. *Working Paper*.

Manso, G. (2011). Motivating innovation. *The Journal of Finance 66*(5), 1823–1860.

Manso, G., B. Balsmeier, and L. Fleming (2019). Heterogenous innovation and the antifragile economy. *Working Paper*.

Moscarini, G. and L. Smith (2001). The optimal level of experimentation. *Econometrica 69*(6), 1629–1644.

Nikolov, B. and T. Whited (2014). Agency conflicts and cash estimates from a structural model. *Journal of Finance 69*(5), 1883–1921.

Pastor, L. and P. Veronesi (2009). Technological revolutions and stock prices. *The American Economic Review 99*(4), 1451–1483.

Peters, R. and L. Taylor (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics 100*, 251–272.

Prescott, E. and M. Visscher (1980). Organization capital. *Journal of Political Economy* (88), 446–461.

Rob, R. (1991). Learning and capacity expansion under demand uncertainty. *The Review of Economic Studies 58*(4), 655–675.

Romer, P. (1986). Increasing returns and long run growth. *Journal of Political Economy* (94), 1002–1037.

Stoffman, N., M. Woeppel, and D. Yavuz (2021). Small innovators: No risk, no return. *Working Paper*.

Stoyanov, A. and N. Zubanov (2012). Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics* (4), 168–98.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica 15*(4), 279–313.

Weitzman, M. (1979). Optimal search for the best alternative. *Econometrica 47*(3), 641–54.

Young, A. (1993). Invention and bounded learning by doing. *Journal of Political Economy 101*(3), 443–472.

A Model Calibration

A.1 Baseline Calibration

We first define parameter values for our baseline calibration, and summarize them in Table 3. Many parameters are standard, and thus calibrated following common practices in the literature. Consistent with Nikolov and Whited (2014), we set the discount rate to $r \equiv 5\%$, the capital depreciation rate to $\delta \equiv 13\%$, and the production function is such that $\alpha \equiv 0.8$. The profit function exhibits significant decreasing returns to scale, consistent with the rise of market power documented in Gutierrez and Philippon (2017). We set the inverse of the elasticity of demand to $\eta \equiv 0.5$, to capture the average markup of 1.5 in recent years as estimated in De Loecker, Eeckhout, and Unger (2019). Importantly, conjecturing that firm value is concave when experimentation is optimal, the"expected returns on knowledge" is endogenously lower than 0.5 (by Jensen's inequality); accordingly, we choose returns on capital, $\alpha(1-\eta) = 0.4 < 0.5$. The curvature of profits with respect to capital is therefore on the lower bound as discussed in Abel and Eberly (2011). Last, because firms' variable adjustment costs in our setting are proportional to profits, we calibrate firms' variable adjustment costs to $\gamma \equiv 5$ so that the investment rate is consistent with the summary statistics observed in our sample (see Table 1) and in Table 1 of Peters and Taylor (2017).

Table 3: Baseline parameter values in our full model. The table summarizes standard parameters in neoclassical models (part 1), and others specific to our model (part 2).

Symbol	Definition	Value
1. standard parameters		
r	risk-free rate	5%
δ	depreciation rate	13%
α	returns to scale on capital	0.8
η	inverse of price elasticity of demand	0.5
γ	variable adjustment costs	5
2. other parameters		
τ_A	informativeness of aggregate productivity	0.6
τ_S	informativeness of experiments	0.3
ω	obsolescence costs	59%
ϕ	knowledge dissipation	0.4

Other parameters determining the behavior of exploration and experimentation in our model are τ_A, τ_S, ω and ϕ. In our baseline calibration, we set the informativeness of experimentation to $\tau_S \equiv 0.3$, and the informativeness of the productivity signal to $\tau_A \equiv 0.6$, to match the observed variation in firm-level growth rates of output (i.e., Ma (2021)). Similarly, we ensure that exploration is infrequent by setting sufficiently high obsolescence costs ω. The stock of capital becomes obsolete so that firm value decreases by $1-(1-\omega)^{\alpha(1-\eta)} \approx 30\%$ upon exploration. We thus assume significant creative destruction upon breakthroughs in exploration, consistent with the evidence in Foster, Haltinwanger, and Krizan (2001). Finally, the dissipation of knowledge occurring upon experimentation determines how desirable a technology is. We set $\phi \equiv 0.35$, a value that triggers sufficient asymmetry in the exploration thresholds, so that the firm is more likely to experiment with an existing technology when the firm is confident that the technology is good ($Z > 0$).

A.2 Alternative Parametrizations

In Figure 7, we illustrate the properties of the model with respect to the expected length of the knowledge cycle by considering the same parameters as in Table 3, with the exception of $\tau_A = 0.3$.

In the extension with partially transferable knowledge, we illustrate the properties of the model numerically by considering the same parameters as in Table 3, in addition to imposing $\lambda = 0.2$.

In the extension with fixed costs, we consider all standard parameters as stated in the first panel of Table 3, as well as $\omega = 0.59$, which we also use in the baseline calibration. Considering that our assumptions on adjustment costs, knowledge dissipation and the signals extracted from experimentation in this extension are different from the baseline model, we modify the remaining parameters so that $\gamma = 0.5$, $\kappa = 0.095$, $\tau_A = 0.7$, $\tau_S = 0.2$, and $\phi = 2$.

B Homogeneity of the value function

In this appendix we show that the value function of the firm takes the form in Eq. (9). We can write the problem of the firm as:

$$\begin{aligned}
V\left(N_t, A_t, K_t, \widehat{M}_t, \Omega_t\right) \equiv \max_{\{\nu_n\}, i_{t+s}} \mathbb{E}\left[\int_0^\infty e^{-rs}\left(\Pi\left(A_{t+s}, K_{t+s}, N_{t+s}\right)\left(1-\widehat{\gamma}_{t+s}\right)\right) \mathrm{d}s \,\middle|\, \mathscr{F}_t\right] \\
\text{s.t. } \mathrm{d}N_t &= \phi N_t \left(\widehat{M}_t/\Omega_t^{1/2}\right)^2 \mathbf{1}_{\widehat{M}_t/\Omega_t^{1/2} \geq 0} \mathrm{d}t \\
\mathrm{d}A_t &= A_t(\tau_A^{1/2}\widehat{M}_t/\Omega_t^{1/2}\mathrm{d}t + \mathrm{d}\widehat{B}_t) \\
\mathrm{d}K_t &= (i_t - \delta)K_t \mathrm{d}t - \omega K_{t-}\mathbf{1}_{t=\nu} \\
\mathrm{d}\widehat{M}_t &= \Omega_t^{1/2}\tau_A^{1/2}\mathrm{d}\widehat{B}_t + \Omega_t^{1/2}\tau_S^{1/2}i_t^{1/2}\mathrm{d}\widehat{B}_{S,t} - \widehat{M}_{t-}\mathbf{1}_{t=\nu} \\
\mathrm{d}\Omega_t &= -\Omega_t\left(\tau_A + \tau_S i_t\right)\mathrm{d}t + (\tau_M^{-1} - \Omega_{t-})\mathbf{1}_{t=\nu}.
\end{aligned}$$

Consider the expectation in the firm's objective. Integrating over the dynamics of A, N and K and substituting we obtain:

$$\mathbb{E}\left[\int_0^\infty e^{-rs}\left(\Pi\left(A_{t+s}, K_{t+s}, N_{t+s}\right)\left(1-\widehat{\gamma}_{t+s}\right)\right) \mathrm{d}s \,\middle|\, \mathscr{F}_t\right]$$

$$= \Pi(A_t, K_t, N_t)\mathbb{E}\left[\int_0^\infty \frac{(1-\omega)^{\alpha(1-\eta)k_s} e^{(1-\eta)(\widehat{B}_{t+s}-\widehat{B}_t)}(1-\widehat{\gamma}_{t+s})}{e^{\int_t^{t+s}\left((1-\eta)\left(\tau_A^{1/2}\widehat{M}_u/\Omega_u^{1/2}+\alpha(i_u-\delta)-1/2\right)-\eta\phi\left(\widehat{M}_u/\Omega_u^{1/2}\right)^2 \mathbf{1}_{\widehat{M}_u/\Omega_u^{1/2}>0}-r\right)\mathrm{d}u}} \mathrm{d}s \,\middle|\, \mathscr{F}_t\right],$$

where k_s denotes the number of times exploration was triggered over the interval $[t, t+s]$. Now introduce the change from the firm's probability measure, $\widehat{\mathbb{P}}$, to another probability measure, $\widetilde{\mathbb{P}}$, according to:

$$\left.\frac{\widetilde{\mathbb{P}}}{\widehat{\mathbb{P}}}\right|_{\mathscr{F}_t} = e^{-1/2(1-\eta)^2 t + (1-\eta)\widehat{B}_t},$$

and define the function:

$$F\left(\widehat{M}_t, \Omega_t\right) \equiv \widetilde{\mathbb{E}}\left[\int_0^\infty (1-\widehat{\gamma}_{t+s})(1-\omega)^{\alpha(1-\eta)k_s} e^{\int_t^{t+s}\left((1-\eta)\left(\tau_A^{1/2}\frac{\widehat{M}_u}{\Omega_u^{1/2}}+\alpha(i_u-\delta)-\eta/2\right)-\eta\phi\left(\frac{\widehat{M}_u}{\Omega_u^{1/2}}\right)^2 \mathbf{1}_{\frac{\widehat{M}_u}{\Omega_u^{1/2}}>0}-r\right)\mathrm{d}u} \mathrm{d}s \,\middle|\, \mathscr{F}_t\right].$$

We can now rewrite the firm's value function as:

$$\begin{aligned}
V\left(N_t, A_t, K_t, \widehat{M}_t, \Omega_t\right) &= \Pi\left(A_t, K_t, N_t\right) \max_{\{\nu_n\}, i_{t+s}} F\left(\widehat{M}_t, \Omega_t\right) \\
\mathrm{d}\widehat{M}_t &= \Omega_t^{1/2}(1-\eta)\tau_A^{1/2}\mathrm{d}t + \Omega_t^{1/2}\tau_A^{1/2}\mathrm{d}\widetilde{B}_t + \Omega_t^{1/2}\tau_S^{1/2} i_t^{1/2}\mathrm{d}\widehat{B}_{S,t} - \widehat{M}_{t-}\mathbf{1}_{t=\nu} \\
\mathrm{d}\Omega_t &= -\Omega_t\left(\tau_A + \tau_S i_t\right)\mathrm{d}t + (\tau_M^{-1} - \Omega_{t-})\mathbf{1}_{t=\nu},
\end{aligned}$$

where $\widetilde{B}$ is a $\widetilde{\mathbb{P}}-$Brownian motion. The firm's problem is to choose a stochastic process $(i_u)_{t\geq t}$ and stopping times $\{\nu_k\}_{k\in\mathbb{N}}$ given the initial conditions $(\widehat{M}_t, \Omega_t)$ at date t.

Let (suppressing time indices for simplicity)

$$J\left(\widehat{M}, \Omega\right) \equiv \max_{\{\nu_n\}, i\geq 0} F\left(\widehat{M}, \Omega\right),$$

and let $\mathscr{A}$ denote the region in which sticking to the current technology is optimal. We want to show, first within the inaction region, that the value function has the property that:

$$J(\lambda^{1/2}\widehat{M}, \lambda\Omega) = J(\widehat{M}, \Omega), \tag{26}$$

for an arbitrary constant, $\lambda^{1/2} \in \mathbb{R}$, (the value function is homogenous of degree 0 in $(\widehat{M}, \Omega^{1/2})$).

The customary martingale argument implies J satisfies within the inaction region, $(\widehat{M}, \Omega) \in \mathscr{A}$:

$$0 = 1 + \left((1-\eta) \left(\tau_A^{1/2} \frac{\widehat{M}}{\Omega^{1/2}} - \alpha\delta - \eta/2 \right) - \eta\phi \left(\frac{\widehat{M}}{\Omega^{1/2}} \right)^2 \mathbf{1}_{\frac{\widehat{M}}{\Omega^{1/2}} > 0} - r \right) J \tag{27}$$
$$+ J_{\widehat{M}} \Omega^{1/2} (1-\eta) \tau_A^{1/2} + \tau_A \Omega (J_{\widehat{M}\widehat{M}} - J_\Omega) + \psi(\widehat{M}, \Omega),$$

where ψ denotes the maximand:

$$\psi(\widehat{M}, \Omega) \equiv \max_{i \geq 0} \left\{ \alpha(1-\eta) J i + \tau_S (J_{\widehat{M}, \widehat{M}} - J_\Omega) \Omega i - \widehat{\gamma}(i) \right\}.$$

From the first-order condition the optimal investment policy satisfies:

$$i(\widehat{M}, \Omega) = \begin{cases} 0 & \text{if } \alpha(1-\eta) J + \tau_S (J_{\widehat{M}, \widehat{M}} - J_\Omega) \Omega \leq 1 \\ \frac{1}{\gamma} (\alpha(1-\eta) J + \tau_S (J_{\widehat{M}, \widehat{M}} - J_\Omega) \Omega - 1) & \text{otherwise} \end{cases}.$$

Differentiating the conjectured identity in Eq. (26) twice w.r.t. $\widehat{M}$ and once w.r.t. Ω shows that the same level of investment is optimal for $(\widehat{M}, \Omega)$ and $(\lambda^{1/2} \widehat{M}, \lambda\Omega)$:

$$i(\lambda^{1/2} \widehat{M}, \lambda\Omega) = i(\widehat{M}, \Omega),$$

and the same property thus applies to the maximand. Repeating the exercise on the PDE in Eq. (27) shows that the conjectured homogeneity holds within the region $\mathscr{A}$. Now when adopting a new technology capital is depleted by ω and knowledge is reset to priors, so that for $(\widehat{M}, \Omega) \in \partial\mathscr{A}$:

$$J(\widehat{M}, \Omega) = (1-\omega)^{\alpha(1-\eta)} J(0, \tau_M^{-1}),$$

and thus the conjectured homogeneity holds upon exploration as well. Note that homogeneity upon resets follows from our baseline assumption that knowledge is reset to priors. In Section 4.4 we allow for partial resets. For homogeneity to be preserved we need that $M_\nu \propto M_{\nu-}$ and $\Omega_\nu \propto \Omega_{\nu-}$, which underlines the assumption we make in Eq. (22). Finally, letting $J(\widehat{M}/\Omega^{1/2}, 1) \equiv v(\widehat{M}/\Omega^{1/2})$ we can write:

$$V\left(N_t, A_t, K_t, \widehat{M}_t, \Omega_t\right) = \Pi\left(A_t, K_t, N_t\right) v(\widehat{M}/\Omega^{1/2}).$$

C Extensions

In this appendix we describe the solution method for the model extensions presented in Section 5.

C.1 Separate investment decisions in capital and knowledge

We start by writing the HJB equation associated with the problem in Eq. (24). This HJB takes the form in Eq. (11) with the maximand replaced by:

$$\psi(q,c) = \max_{i_K \geq 0, i_Z \geq 0} \{qi_K + ci_Z - \widehat{\gamma}(i_K, i_Z)\}.$$

The first-order conditions for optimal investment in capital and knowledge are:

$$\begin{aligned} i_K(q) &= \max\{\delta + (q-1)/\gamma, 0\} \\ i_Z(c) &= \max\{(c - p_Z)/\gamma_Z, 0\}. \end{aligned}$$

As in the benchmark case of Section 3 q is a sufficient statistic for investment in capital, and mirroring this solution the expected marginal benefit of knowledge, c, is a sufficient statistic for investment in knowledge. Yet we know from the main analysis that q and c behave differently as functions of Z. In particular, whereas marginal q is guaranteed to be positive, we know that the sign of c depends on those of $v'Z$ and v''. From the main analysis we have seen that early in the cycle $v'Z > 0$ and $v'' < 0$ with the latter dominating $c < 0$, and late in the cycle $v'Z < 0$ and $v'' > 0$ with the latter dominating $c > 0$. Therefore, the constraint that $i_Z \geq 0$ triggers inaction early in the cycle (inaction with respect to investment in knowledge only). Under the conjecture that firm value is single peaked in knowledge, there must exist two "trigger levels" of knowledge Z, $\underline{a} < \underline{b} < 0 < \overline{b} < \overline{a}$, at which:

$$c(\underline{b}) = c(\overline{b}) = 0$$

is satisfied and beyond which the firm decides to invest in knowledge. Thus $\mathscr{B} \equiv (\underline{b}, \overline{b})$ with $\mathscr{B} \subset \mathscr{A}$ denotes the region within which the firm does not invest in knowledge. We solve firm value piecewise over the four associated regions ($[\underline{a}, \underline{b}]$, $[\underline{b}, 0]$, $(0, \overline{b}]$ and $[\overline{b}, \overline{a}]$), imposing the boundary conditions in the main text and adding the smooth-pasting condition associated with each of the two b thresholds:

$$\lim_{Z \uparrow b} v'(Z) = \lim_{Z \downarrow b} v'(Z).$$

C.2 Fixed adjustment costs

The HJB equation is identical to that in Eq. (11) with the maximand in Eq. (12) but in which adjustment costs are replaced by Eq. (25). Optimal investment is as in the benchmark, entirely driven by q, but the fixed costs now triggers inaction in some regions. As q declines beyond a certain level it can no longer justify paying the fixed cost and the firm no longer invests. Given the conjectured shape of firm value there must exist two "trigger levels" of knowledge Z, $\underline{a} < \underline{b} < 0 < \overline{b} < \overline{a}$, at which marginal q is so that:

$$\psi(q(\underline{b})) = \psi(q(\overline{b})) = 0,$$

and this value of marginal q in turn corresponds to:

$$q(\underline{b}) = q(\overline{b}) = \frac{1 - \delta\gamma + \sqrt{\gamma(2F + \delta^2\gamma)}}{\alpha(1-\eta)}.$$

In contrast to Section C.1 in which inaction takes place at the center in a neighborhood of $Z \approx 0$, with fixed adjustment costs inaction takes place for large values of Z (when q is low). Thus $\underline{\mathscr{B}} \equiv (\underline{a}, \underline{b})$ and $\overline{\mathscr{B}} \equiv (\overline{a}, \overline{b})$ denote the two regions within which the firm is inactive. We solve firm value piecewise over the four associated regions ($[\underline{a}, \underline{b}]$, $[\underline{b}, 0]$, $(0, \overline{b}]$ and $[\overline{b}, \overline{a}]$), imposing the boundary conditions in the main text and adding the smooth-pasting condition associated with each of the two b thresholds:

$$\lim_{Z \uparrow b} v'(Z) = \lim_{Z \downarrow b} v'(Z).$$