A Theory of Market Pioneers, Dynamic Capabilities and Industry Evolution*

Matthew Mitchell†and Andrzej Skrzypacz‡

July 19, 2013

Abstract

We analyze a model of industry evolution where the number of active submarkets is endogenously determined by pioneering innovation from incumbents and entrants. Incumbents enjoy an advantage at innovation in submarkets similar to ones in which they currently produce. We complement the existing literature - that focuses on exogenous arrival of submarkets ([21] and [27]) or the steady state of a model with constant submarkets ([22]) - by describing how competition, free entry, and the dynamic capability of incumbents drives the evolution of an industry. An important driving force comes from the demand side, as in [1]: increased competition drives down profits, which in turn makes the proportion of pioneering done by incumbent firms, with the advantaged position, rise over time. The total number of submarkets follows an S-shape, consistent with empirical studies. The shift from immature to mature submarkets can lead to a shakeout in firm numbers. Innovation shifts from pioneering to non-pioneering as the industry evolves, which is consistent with evidence on innovation and industry evolution.

*We thank three anonymous referees and Associate Editor Bruno Cassiman as well as Philippe Aghion, Serguey Braguinsky, Alex Frankel, Paul Romer and participants at seminars at the Carnegie Mellon, Duke, the Federal Reserve Bank of Minneapolis, Iowa State, and Toronto as well as at DRIUD 2012 and the 4th Duke/Texas/Northwestern IO theory conference for suggestions.

†Rotman School of Management, University of Toronto.

‡Stanford GSB.

1 Introduction

Firms are driven by a constant process of finding new profit opportunities. Sometimes those profit opportunities involve overtaking an existing firm's place at the top of one activity. Other times the profit opportunities involve pioneering new activities, for instance developing a new submarket, which we term "market pioneering." New submarkets are an important driver of industry evolution: both [21] and [27] show that, taking arrivals of new submarkets as exogenous, such a model can help explain firm and industry dynamics. On the other hand, [22] show that the steady state of a model with a constant set of submarkets can generate predictions about the cross section of firm size and innovative behavior consistent with empirical evidence. This paper provides a bridge between the two, where submarkets are generated *endogenously* through market pioneering. We show that the endogenous evolution of the industry in our model is consistent with empirical evidence.

A key driving force behind the evolution of submarkets in our model is the dynamic capabilities of incumbent firms. [12] describes dynamic capabilities as "the capacity of an organization to purposefully create, extend, or modify its resource base." The literature on dynamic capabilities, more broadly, shows that entry into new markets is driven by the past experiences of firms, and is not random.[1] In particular, firms' entry decisions are highly driven by experience in related industries. Our model, therefore, is driven by firms with dynamic capabilities that link similar submarkets. In our model, as in [22], incumbent firms at the frontier generate new innovations from a resource that comes from their current leadership position.[2] This is consistent with evidence in [15] and [13], who stress the benefit of experience in generating dynamic capabilities. In [22], every submarket is identical, and the set of submarkets is fixed, so there is no sense in which *what* a firm is doing now impacts the *sorts* of markets it might enter in the future. We focus our attention on dynamics driven by firms' abilities at entering submarkets that are similar to the ones where they have expertise.

Empirical evidence draws a tight link between market pioneers and the activities they undertook before pioneering. The message of many of these papers is broadly similar: an important input firms use in entering new submarkets is experience from *related* submarkets. Entry into the newest, most

[1] For a general discussion, see [28, 29, 12]

[2] One natural interpretation is that the resource is knowledge.

advanced submarkets is tied to participation is similarly recent submarkets. [3] To operationalize this idea, we assume that submarkets mature stochastically. As a result, at any point in time some submarkets are immature; these are, on average, relatively new submarkets, which are most likely to be technologically advanced but also unsettled. Like [22], we take the aggregate innovation technology to be constant returns; however, there are two technologies for producing innovations, with pioneering innovation fundamentally different from the sort of overtaking innovation that is the focus of [22]. Leadership in immature submarkets generates the relevant stock of dynamic capabilities used in pioneering innovation and entry into existing immature areas. Similarly, a dynamic capability of firms involved in mature submarkets aids in the generation of innovations which overtake current leaders in mature submarkets. The relative returns and the stocks of dynamic capabilities generate the relative innovation rates across the different types of submarkets. This evolution is characterized by competitive innovation, and therefore is tightly tied to the returns to innovation of different types, which we characterize.

Since the stock of incumbents is fixed at any point in time, their innovative inputs may still leave room for entry by de novo entrants (i.e. entry by firms with no operations in other submarkets) who lack the dynamic capability, and therefore operate a less efficient innovation technology. These entrants "take up the slack," and, from the standpoint of analyzing the model, pin down the return to innovation via the free entry condition whenever they are active.

We study the endogenous evolution of submarkets for an industry that starts from a tiny number of immature submarkets. During the early phase,

[3] For instance, in the hard drive industry, [6] document that entrants into new submarkets (new market diameters) disproportionately come from the most recent, high tech firms, suggesting that production in the most recent submarkets is relevant to entry into new submarkets. More generally, our model is consistent with the assumption that new submarkets are technologically similar to more recent, immature submarkets. [11] notes that R&D in a new submarket is buttressed by knowledge in similar submarkets. [26] found that entry into a new drug is tied to a firm's experience with drugs having similar characteristics. [14] use this notion to drive "technological trajectories," where a firms prior experience determines its future decisions. [5] show that, more generally, innovative firms with a greater stock of knowledge are more likely to introduce new products. We therefore assume that innovation in new and recent submarkets is related to participation in the current stock of recent submarkets, which we take as a measure of the sort of knowledge stock highlighted in [5].

since submarkets are disproportionately immature, research is focused on immature submarkets, and generates new ones at a relatively fast rate. The reason is that firms capabilities are naturally skewed toward immature submarkets during this period. As immature submarkets gradually mature, capabilities switch toward mature submarkets, leading to more investment in overtaking in mature submarkets and less in developing new areas. Eventually this slows to the point where the immature submarkets are shrinking in number, as they mature faster than they are generated. The industry moves toward a stable set of mature submarkets.

We show how the set of submarkets and the number of firms evolves over this process. The growth of submarkets is S-shaped: first it rises at an increasing rate, but eventually grows more and more slowly as it reaches a plateau. This pattern for the evolution of submarkets within an industry matches the one that [30] argues is both consistent with evidence from the tire industry and useful in replicating a variety of facts about industry dynamics.

Entry eventually peaks, and can decrease fast enough relative to exit to generate a shakeout. The force behind the shakeout is a combination of demand and supply side factors. On the demand side, increased competition is driving down profits per submarket. This can eventually drive down de novo entry without eliminating innovation by incumbents, since incumbents have a dynamic capability on the supply side. We show that entry drops rapidly (in fact discontinuously). At this point, incumbents with a weak market position are bound to fail, forced out by a combination of other firms' strong position, and bad luck. Exit of those firms is not offset by entering firms, since entering firms do not have sufficient capabilities to merit spending the cost of entry. [4]

Our model of the shakeout is related to the one that derives from [17], further applied in [19, 3]. In that framework, prices fall, eventually making entry unattractive. Here, competition makes entry difficult because only the incumbents have the requisite dynamic capability to efficiently enter under the more competitive circumstances. In other words, the model of the shakeout introduced here is a complementary story, where demand side changes

[4]In that sense our model incorporates the demand side force familiar from [1], who demonstrate that changes in demand characteristics can drive industry evolution. In deploying dynamic capabilities, firms' innovation is determined by returns that are sensitive to demand characteristics. Unlike [1], firms in our model use their dynamic capability according to the full present discounted expected value of the return to innovation, as in [22].

impact industry innovation through the incumbents' relative capability to innovate. The outcome is similar: as suggested in [18] and [20], early entrants generate a capability that helps them to survive even after entry falls. Our model is therefore broadly consistent with the evidence in [3], that new submarkets might be associated with a shift toward innovation by "leading incumbents." In our model, early entrants are more likely to hold leadership positions in mature submarkets, whereas later entrants are more likely to be in immature submarkets whose failure generates the shakeout. Moreover, the notion that new submarkets strengthen incumbents positions relative to entrants is consistent with the message of [10], who show that new submarkets for multipurpose tractors in Germany benefited incumbents with related market experience. Here we take related market experience, as is stressed in the dynamic capabilities literature, to include experience in other recently introduced submarkets.

The model allows immature submarkets to differ from mature submarkets on other dimensions besides the dynamic capability they bring. On one hand, immature submarkets may not be fully commercialized. On the other hand, early movers might enjoy benefits not attained later on. Both differences are important, since innovation levels are determined by the dynamic return to innovation, and therefore respond to current profitability as well as the expected impact of industry evolution on submarket profitability. Our model, therefore, incorporates various sorts of implications of early entry as described in [23]. We show, in fact, that in some cases the measured returns to early movers are entirely generated on the supply side by the relative cost of de novo entry. Put another way, differences in capabilities of the firms *that follow* the early movers determine the return to early moving, and not necessarily the capabilities of the early movers themselves.

In addition to generating predictions about entry and exit that are consistent with evidence on evolution, the model also generates predictions regarding the way in which the volume and composition of innovation changes over the life cycle. Pioneering innovation rises and then falls. Non pioneering research, under general conditions, is rising as the industry reaches maturity. It is natural to interpret process innovations as disproportionately non-pioneering, while product innovations represent, at least partially, pioneering of new submarkets. With that interpretation, we can compare the model's predictions on pioneering to well known evidence on product innovations over the course of an industry life cycle. This evidence was documented first by [31], and has been further discussed in papers including [4] and [16].

Innovations move from product to process innovations, with product innovations steadily falling and process innovations rising. Moreover, our model is consistent with the depiction of industry evolution driven by a changing standard product contained in [16]. One can interpret mature submarkets as variants of the "standard" product with a particular unique feature; immature submarkets are variants that are not yet accepted as a standard, and may never be. Maturity reflects a submarket's integration into the standard under that interpretation.

We show, further, that the model can lead to an industry which is innovative even in its mature phase, including the possibility that product innovation persists. The model can accommodate this fact either by assuming that mature submarkets are sufficiently product-innovation intensive, or by allowing for the possibility that mature submarkets sometimes die. This alternative leads to persistent pioneering innovation, that can be naturally interpreted as product innovation.

We focus on the lifecycle predictions of the model as the number and types of submarkets evolves to a mature industry. Our model converges to a steady state identical to [22]. They show that, for a fixed set of submarkets, a quality ladder style model can explain a variety of cross sectional facts about R&D and firm size in the stationary distribution of firms. Our model of dynamic capabilities, therefore, delivers a model of submarket arrival that is consistent with the empirical features of models of overtaking, while at the same time delivering arrivals of new submarkets in a way consistent with [21] and [27]. The key force in the evolution of the industry is the evolution of dynamic capabilities.

Outside of the literature on dynamic capabilities, the notion that firms diversify into related product areas has long been documented. [7] showed that diversifying firms chose related areas. This basic fact has both motivated a variety of models (for instance [24]) and led to a wide variety of papers studying the forces behind the phenomenon. We focus on the role that a firm's experience has in generating a capability to enter similar submarkets. Finally, as in [22], our approach to modeling innovation draws heavily on the endogenous growth literature in economics, such as [9, 2]. Our model adds endogenous variety, and in that sense is similar to a long line of growth theory papers such as [25]. Our model merges these ideas with a richer innovative structure including dynamic capabilities in order to study the industry life cycle.

2 Model

At any given time t there is an industry made up of a continuum of submarkets of measure N_t. Of the N_t submarkets, M_t have reached maturity, and I_t remain immature, so $I_t + M_t = N_t$. The industry evolves over an infinite horizon of continuous time, with future payouts discounted at the interest rate r.

In this section we take each submarket to be characterized by a profit making leader, and follower firms who earn zero profits, as in the canonical quality ladder models of [22] and [9]. In section 5.2 we show that the model is amenable to allowing several profit making firms per submarket at only the cost of notational complexity. Each leader of a mature submarket earns $\pi(M_t, I_t)$ per instant they are the leader, and each leader of an immature submarket earns $\alpha\pi(M_t, I_t)$ per instant, with $\alpha > 0$. We assume that π is continuous and decreasing in both arguments, reflecting the notion that there is elasticity of substitution between submarkets, and therefore more submarkets lead to less profits per submarket. In section 5.3 we introduce an explicit model of consumer preferences and show that it delivers this structure for profits; however, we suppress it here since all of the fundamentally new analysis does not require a specific interpretation of the origin of profits. The key here is determining how those profits translate into valuations for submarket leaders, and in turn innovation rates. This simple model of profits by submarket is analogous to assumptions in both [22] and [21]; Klette and Kortum use the term "goods" and Klepper and Thompson use "submarkets."

Industry evolution comes via innovation. There are two types of innovation: one focused on immature submarkets, which we term pioneering innovation, and innovation focused on mature submarkets. Mature submarket innovation works exactly like the quality ladder structure in [22], which is borrowed from [9]: a successful innovation into mature submarkets generates a new, higher quality version of some submarket, and therefore makes the innovator become the new submarket leader.[5]

Innovative effort in immature submarkets is the source of both improved products in those immature submarkets, as well as new designs that generate new submarkets. This pioneering research, therefore, is a form of early mov-

[5] One can take this research to be undirected or directed across submarkets. Incumbents would never want to innovate in a submarket they led in due to the Arrow replacement effect; as a result, in a symmetric equilibrium, every submarket would be researched equally, and never by its current leader, just as under undirected research.

ing, either as a first mover, or as one of the firms which enters the submarket soon after the first mover. A fraction (or, identically, probability) $1-\phi$ of innovations from research into immature submarkets generates an improvement to an existing immature submarket, resulting in a simple changing of leadership in that immature area. This matches the notion that immature areas still attract commercial competition, but perhaps in different amounts from mature submarkets. The remaining fraction of successes generate an entirely new immature submarket.[6]

Immature submarkets are fundamentally different from mature ones in two other ways, besides the profit difference ($\alpha \neq 1$) or the possibility of generating new innovations ($\phi > 0$). First, the cost of research, described below, may be different in immature areas. Second, not all immature submarkets eventually become viable, mature submarkets. Immature submarkets sometimes fail to become viable, dying at Poisson rate λ. Immature submarkets mature at Poisson rate μ, at which point they are viable and permanent.

Denoting total innovation in immature submarkets as a rate of i_t units per immature submarket, the change over time in the measure of immature submarkets is the new arrivals minus the maturing and failing submarkets:[7]

$$\dot{I} = i\phi I - \mu I - \lambda I \tag{1}$$

Here we take intensity to be symmetric across immature submarkets, since they are identical; this corresponds to a symmetric equilibrium.

The change in the measure of mature submarkets comes from maturing immature submarkets. The change over time in the measure of mature submarkets is

$$\dot{M} = \mu I \tag{2}$$

Denote the total rate of innovation in mature submarkets as m_t per mature submarket. Note that this does not change M, since these innovations are entirely generating new leaders among the mature submarkets.[8]

In order to model the dynamic capability of incumbents, we will assume that the technology for generating innovations differs across incumbents and

[6] The model could be extended to allow research in both types of submarkets to generate new submarkets. We choose to focus on the role of immature submarket research in generating new submarkets because it fits with the notion that immature submarkets are more similar to undiscovered submarkets, on average, than mature submarkets are.

[7] Here we suppress the t subscripts to streamline the presentation.

[8] Below we relax the assumption that mature submarkets continue forever.

entrants. Innovation by incumbent firms comes about via a constant returns production function. For incumbents, the arrival rate of innovations is determined by the production function

$$F_M(K_M, L_M)$$

for mature industries and

$$F_I(K_I, L_I)$$

for immature industries. There are a continuum of firms, each with a finite number of leadership positions of the two types. Here L is the number of leadership positions the firm has in a particular type of submarket, and K is all other inputs in the production of innovation (and can potentially be multidimensional). We normalize the units of measure to be in dollars, i.e. cost of one dollar per unit, and assume that $F_I(0, L_I) = 0$. In general, we can allow the production functions to differ in arbitrary ways across the two research types, but we assume they are both concave and constant returns.

The form of this production function follows [22] and is the key feature that embodies the nature of the dynamic capability that we assume. Leadership in immature and mature submarkets generates the capability needed to operate the production function for innovation in those submarkets. A firm with more of the dynamic capability conferred by L generates more innovations for a given amount of K. As a result, firms with entry into recent, immature submarkets are assumed to have a resource that generates additional innovations in immature submarkets, in keeping with the literature on entry into related areas.[9] Because of the constant returns assumption, firms will optimally choose employment of K in proportion to their dynamic capability L, as in [22].

Firms come in many sizes, corresponding to a different number of leadership positions; the distribution of firms is over the number of mature and immature submarket leadership positions it holds. A firm with twice as many leadership positions of a particular type will hire twice as much of the other inputs, and generate twice as many innovations of that type. We therefore

[9] The technology operated by incumbents need not be interpreted as generating entry into new submarkets solely by incumbents themselves; indeed, papers including [6] stress the role of spin-outs in generating entry into technologically advanced submarkets. Our model allows the incumbent innovation which benefits from the dynamic capability of the firm as being executed by employees who leave the firm. As in [6] we will assume that rents coming from such activity are captured by the parent firm.

analyze the decision of incumbents on a per-submarket basis; the same decisions apply to any incumbent, regardless of how many leadership positions it holds. We return to the distribution of firms only after everything about the equilibrium has been characterized at the submarket level.

For instance, for immature submarkets, define

$$f_I(k) = F_I(K, 1)$$

where k is interpreted as the input of K per leadership position.[10] The dynamic capability conferred by immature submarket leadership might come in different forms, sometimes favoring developing leadership positions in other immature submarkets, and other times generating a capability in developing entirely new submarkets. In that case ϕ represents the fraction of that leads to moving first in new markets, while the remaining fraction $1-\phi$ is associated with entering immature markets as an early mover.

In addition to the incumbent innovation technology there is a *de novo* entry technology, which has constant cost normalized to 1 for mature submarkets and c for immature submarkets.[11] This normalization implies that all profits and costs are expressed in terms of the cost of de novo innovation in mature markets. Because the incumbent innovation technology is concave, there is always a positive return to operating it when it is operated, i.e. there is a dynamic capability. By contrast, competition leads to zero net profits for de novo entrants' innovation technology.

In order to characterize innovation levels, we need to describe the present value of being a submarket leader. Denote by V_t the present discounted value of a leadership position at time t in a mature submarket when the current state is (M_t, I_t), and by W_t the value of a leadership position in an immature submarket. Consider a de novo entrant investing 1 unit in the mature innovation technology for dt units of time. This generates a payoff of

$$V_t dt - dt$$

while investing c for dt units of time in the immature technology generates

$$W_t dt - c dt$$

[10] We suppress the footnotes on the capital input since it always matches the one on the associated production function.

[11] Note that this need not be the only entry technology; the model could allow for exogenous entry from other sources, including spin outs.

On the other hand, the incumbents choose k to maximize

$$V_t f_M(k)dt - kdt$$

for a mature submarket, and

$$W_t f_I(k)dt - kdt$$

for an immature submarket. Denote these optimal choices by $k_M(V)$ and $k_I(W)$. For a mature submarket, for instance, the return on this investment, when a mature submarket is valued at V, is $V f_M(k_M(V)) - k_M(V)$.

An important variable is the quantity of innovation per submarket leader per instant, when de novo entrants are making zero profits, i.e. at times where $V_t = 1$ and $W_t = c$; this determines the maximum amount of innovation the incumbents will ever generate, when de novo entry occurs and the net return to such entry is exactly zero. Let $\bar{k}_I = k_I(c)$, $\bar{k}_M = k_M(1)$, $\bar{i} = f_I(\bar{k}_I)$, and $\bar{m} = f_M(\bar{k}_M)$. Total innovation is m_t and i_t, of which $\bar{m}$ and $\bar{i}$ are accounted for by incumbents; the net amount of innovation contributed by de novo entrants is $i - \bar{i}$ for immature submarkets and $m - \bar{m}$ for mature submarkets, if the total innovation rate exceeds what the incumbents offer .

When free entry binds so that $m_t > \bar{m}$ or $i_t > \bar{i}$, incumbents and entrants are both innovating, with incumbents having lower average cost by concavity of f. As a result, the incumbents have dynamic returns from investment in new submarkets that are strictly positive. This is the dynamic capability in the model. Note that an incumbent with leadership positions in more submarkets L has a greater dynamic capability, in the sense that it has more of the resource that generates this cost advantage. The value of this resource is proportional to the number of submarkets in which the firm is a leader. If the resource incumbents have is , industry wide, insufficient relative to the returns, de novo entry takes up the slack.

An *industry equilibrium* for some initial M_0, I_0 is a sequence $V_t, W_t, m_t, i_t, M_t, I_t$ such that

1. M_t and I_t satisfy (1) and (2)
2. $m_t > \bar{m}$ implies $V_t = 1$ (profit maximization for de novo mature entrants)
3. $i_t > \bar{i}$ implies $W_t = c$ (profit maximization for de novo immature entrants)
4. $m_t \leq \bar{m}$ implies $m_t = f_M(k_M(V_t))$ (profit maximization for incumbents in mature submarkets)

5. $i_t \le \bar{i}$ implies $i_t = f_I(k_I(W_t))$ (profit maximization for incumbents in immature submarkets)

6. V_t and W_t satisfy (3) and (5)(below)

The key task of the next section is to characterize the values V_t and W_t described in the final equilibrium condition for all possible combinations of M and I, which in turn determines innovation rates. Even before doing that, however, the model gives an immediate insight into the sources of returns for early movers (i.e. innovators in immature submarkets) and late movers (i.e. new leaders in mature submarkets), such as that described in [23]. In the model it is assumed that per instant profits differ between the two types of submarkets by a factor of α. However, when de novo entry in both areas is positive (i.e. $m_t > \bar{m}$ and $i_t > \bar{i}$), the relative gross return to entry in the two areas W/V is exactly c; α is irrelevant. Intuitively, the capabilities of *future* entering firms in each of the two areas, and not the current profitability of the areas, determines any measured "first mover advantage" of early versus late movers, as measured by gross return to successful entry in immature versus mature submarkets. This shows the difficulty in assessing the inherent benefit from being an early mover, as defined by the relative flow rate of rewards for early entrants (i.e. α) compared to the realized discounted returns to early moving, which depends on the endogenous response of other firms.

3 Equilibrium

3.1 Mature Submarkets: Perpetual Innovation by Incumbents

Mature submarkets behave as all submarkets do in [22], with perpetual innovation and changing leadership. This benchmark characterization is not the key prediction of the model; on the contrary, this section merely shows the sense in which the model follows the line of previous work: an industry populated with a constant set of mature submarkets would behave exactly as in [22], and deliver the same predictions about innovation that they deliver. We then build an endogenous evolution of the number and type of submarkets, including immature submarkets, in the theory of market pioneering that follows.[12]

[12] Since immature submarkets may eventually mature, we must compute the return in mature submarkets first, since it is part of the expected return in an immature submarket.

To insure that innovation in mature industries is perpetual, and therefore an industry populated by mature industries will behave like [22], we assume

Assumption 1. For all M, I, $\pi(M,I) > r + \bar{m}$

This assumption simply implies that profits are always high enough to attract de novo entrants to mature submarkets, so that they mimic the structure of [22]. [13]

Lemma 1. *For all* t, $V_t = 1$

Proof. Suppose $V_t < 1$. By the free entry condition it must be the case that $m_t \leq \bar{m}$. If the free entry condition binds again in T periods, then the payoff is

$$\begin{aligned} V_t &\geq \int_0^T e^{-(r+\bar{m})t}\pi(M_t, I_t)dt + e^{-(r+\bar{m})T} \\ &\geq \int_0^T e^{-(r+\bar{m})t} min_{0<t<T}\pi(M_t, I_t)dt + e^{-(r+\bar{m})T} \\ &> 1 \end{aligned}$$

where the first inequality is because V_t discounts using $m_t < \bar{m}$, and the last inequality is by Assumption 1. Therefore the contradiction implies $V_t = 1$. □

As a result of the lemma, we can characterize the return to mature innovation in terms of a simple Bellman equation, familiar from pricing equations in finance:

$$rV_t = \pi(M_t, I_t) - m_t V_t + (\bar{m}V_t - \bar{k}_M) + \dot{V}_t \tag{3}$$

Mature submarket leadership generates a flow payoff of π, and has a risk m of losing all value. There is also a benefit to the dynamic capability: it generates a new leadership position at rate $\bar{m}$, at a cost of $\bar{k}_M$, the net value of which is contained in the term in parenthesis. Since f is concave, this term is strictly positive: there is value in the dynamic capability. Finally, in such a valuation, one must take account of the possibility that the value of leadership in a mature submarket might change over time due to changes in M_t and I_t, which we denote $\dot{V}_t$, for the time derivative of V. The key in this development is noting that, by Lemma 1, $V_t = 1$ so the value is unchanged when M and I changes, and therefore $\dot{V}_t = 0$

Further, we can substitute $V = 1$ and compute the rate of innovation:

[13] This assumption can be weakened to only hold on the “relevant” range of M and I that is generated in equilibrium.

Proposition 1. *The rate of innovation in mature submarkets is given by*

$$m_t = \pi(M_t, I_t) + \bar{m} - \bar{k}_M - r \tag{4}$$

Note the "demand side" characterization of m: it changes as the returns to innovation change, through the impact on π. The model has both the free entry characteristics of [22] and demand side mechanics in the spirit of [1]. The Klette and Kortum characterization of innovation rates across firms is perfectly compatible with our model if M and I are constant. We will show below that, in fact, in the long-run, M and I converge to constant values, and therefore innovation in our equilibrium converges to the one in Klette and Kortum with constant innovation per submarket. We add a pioneering innovation component that generates evolution of the industry according to the free entry conditions, with "demand side" predictions about the life cycle of the industry, as in [1]. We characterize pioneering innovation next.

3.2 Immature Submarket Innovation

In this section we evaluate market pioneering given that mature submarkets generate a constant payoff of 1, according to Lemma 1. In the next section we put the pieces together and examine the model's predictions for the evolution of the stock of mature and immature submarkets implied by innovation in immature submarkets, and the eventual maturity of those submarket.

An immature submarket has discounted returns W determined by the recursion

$$rW = \alpha\pi(M, I) - i(1-\phi)W - \lambda W - \mu(W - V) + (Wf(k_I(W)) - k_I(W)) + \dot{W} \tag{5}$$

The first term is the current profits generated from leadership; the second and third terms are the expected capital loss from either an improvement which displaces the current leader or failure of the entire submarket, either of which ends that dividend payment. The fourth term is the capital gain or loss when the submarket matures, accounting for the loss of W and the gain of a mature leadership position valued at V. Note that $V = 1$ so this term simplifies further. The next term is the value of the resource that generates the dynamic capability. The final term is the time derivative of W.

We first explore how the value function is determined when there is entry by de novo firms. In that case, $W = c$. On the interior of any such region,

W is therefore constant so $\dot{W} = 0$. Therefore we can rewrite (5) as

$$rc = \alpha\pi(M, I) - i(1-\phi)c - \lambda c - \mu(c-1) + (\bar{i}c - \bar{k}_I)$$

so that

$$i = \frac{1}{(1-\phi)c}(\alpha\pi(M, I) + \mu - (\mu + \lambda + r - 1 - \bar{i})c - \bar{k}_I) \tag{6}$$

As a result, i varies continuously in the range since π is continuous.[14]

Alternatively, it could be the case that there is only innovation by incumbents, so $i = f_I(k_I(W))$. In that case (5) can be rewritten as

$$rW = \alpha\pi(M, I) + \phi iW - \lambda W - \mu(W-1) - k_I(W) + \dot{W} \tag{7}$$

To make sure this value is well-defined, we assume that

Assumption 2. $r + \lambda + \mu > \phi\bar{i}$

Assumption 2 guarantees that discounting and eventual exit from immaturity (either through death or maturity) is sufficient to keep the dynamic capability in immature industries from replicating itself so rapidly that a leadership position has infinite value, generating additional leadership positions faster than the value depreciates.

The analysis simplifies in the case where the maturation process of firms makes it more difficult for other incumbent firms to profit. It is both a natural assumption and consistent with evidence that price declines as firms mature as is evidenced in many papers including [8]. The assumption, which we maintain for the remainder of the paper, is

Assumption 3.

$$\mu\frac{d\pi(M, I)}{dM} - (\mu + \lambda)\frac{d\pi(M, I)}{dI} < 0$$

Assumption 3 ensures that an industry with immature submarkets both dying and maturing becomes more competitive, other things equal, in the sense that profits for each submarket decline, over time; the first term is the impact of the gain in mature submarkets, while the second is the impact of the decline in immature submarkets. The assumption is not essential for our

[14] Expression (6) is simplified due to our previous observation in Lemma 1 that the value to the mature leadership is constant in equilibrium. If it was not constant, (6) would contain also a V_t, but the qualitative features of our model would remain unchanged.

model's life cycle predictions, but it simplifies the analysis greatly without being inconsistent with evidence on firm dynamics. If we take profits to be a function of output, and output to be linear in the two types of firms, i.e. $\pi(M, I) \equiv \pi(M + \gamma I)$, then this simplifies to

$$\frac{\mu}{\mu + \lambda} > \gamma$$

Assumption 3 implies that profits per submarket, and as a result innovation per submarket, decrease over the lifetime of the industry.

Lemma 2. *$\dot{\pi} \leq 0$, strictly if $I > 0$.*

Proof. Since

$$d\pi/dt = \frac{d\pi(M, I)}{dM}\mu I + \frac{d\pi(M, I)}{dI}(i\phi - \mu - \lambda)I$$

And both derivatives of π are negative, it is sufficient that $\frac{d\pi(M,I)}{dM}\mu < \frac{d\pi(M,I)}{dI}(\mu + \lambda)$. □

Since profits are falling, once de novo entry is unprofitable, it remains unprofitable forever after.

Lemma 3. *If $i < \bar{k}_I$ at t, $i < \bar{k}_I$ for all $s > t$.*

Proof. If $I > 0$, the result is immediate from Lemmas 6 (in the appendix) and 2; if $I = 0$, M and I are constant and therefore the industry remains at $i < \bar{k}_I$ forever. □

Whenever $i < \bar{k}_I$, there is never any de novo entry in the future, since such entry is currently unprofitable and industry conditions are becoming more competitive under Assumption 3.

Since $i < \bar{i}$ is an absorbing state, we can construct W directly. Let $\hat{W}(M, I)$ to be the present discounted value of a firm in state M, I, assuming that only incumbents innovate forever after in the immature submarkets, i.e. as if the industry had no de novo entry technology available for immature submarkets. Since M and I are greater at every future state starting from a greater initial M or I, the resulting $\hat{W}(M, I)$ is strictly decreasing in both arguments.[15] The equilibrium value function W is, therefore, either $\hat{W}$ (if free entry does not bind) or c (if it does). In other words

$$W(M, I) = min\{c, \hat{W}(M, I)\}$$

We can describe the set of points where de novo entry ends by first describing the set of points $I = g_0(M)$ defined by

$$\hat{W}(M, g_0(M)) = c$$

For $I \geq g_0(M)$, $i \leq \bar{i}$. For $I < g_0(M)$, $i \geq \bar{i}$. Since $\hat{W}$ is decreasing in both arguments, g_0 must be decreasing.

We can therefore further characterize innovation in the range where all innovation is by incumbents. First we show formally that W declines in the region where de novo entry has ceased.

Lemma 4. *Suppose $i < \bar{i}$. Then $\dot{W} \leq 0$.*

Proof. In (7), continuity requires $\dot{W}$ be continuous since, when there is no de novo entry, continuity of W implies continuity of i. If $\dot{W} > 0$, it must eventually flatten out; at that point $\ddot{W} \leq 0$ but then by (9) $\dot{W} < 0$, a contradiction □

For incumbents, concavity of f implies that falling W leads to falling $k_I(W)$. Combined with the fact that i is decreasing when the free entry condition binds from (6), we conclude that

Proposition 2. *i is decreasing.*

Innovation in immature submarkets declines over time. We can make a further characterization: if i reaches $\bar{i}$ at some finite date T, it does not do so continuously; it *jumps* down. From (5), i must move discontinuously to keep continuity of W at the boundary between the two regions, since the slope of W jumps. The formal proof is left for the appendix.

Proposition 3. $lim_{\epsilon \downarrow 0} i_{T-\epsilon} > \bar{i}$

To understand the discontinuity in i, consider the time just before and after free entry condition binds, where W is equal to (approximately) c. Consider at any point in time expected profits over the next dt units of time.

[15] A formal statement of this would follow the same argument as Lemma 4

These expected dividends are forever strictly declining after free entry stops binding as competition gets more and more fierce. If expected dividends were roughly the same before and after the change, and continuation value went from constant (when free entry binds) to strictly declining (after), then W would *jump* down. But W is continuous; to equate the present discounted value just before and just after free entry stops binding, expected dividends must therefore jump *up* discontinuously, to offset the fact that they will *decline* from then on. This upward jump comes through the probability of losing your market leadership: only it can change discontinuously, and so it must *decline* discontinuously to make the expected dividends jump up, keeping the discounted sum of expected dividends constant.

The discontinuity result, in particular, contrasts with the result for pioneering innovation when the free entry condition starts binding; in the appendix we show that, were one to dispense with assumption 3, and at some point free entry went from not binding to binding, the evolution of i is continuous. Innovation by entrants ends suddenly, even though the model would have it begin smoothly if such a case were to arise. We take this "crash" of innovation by entrants, then, to be a characteristic of the model of dynamic capabilities and free entry applied to pioneering innovation.

This feature naturally connects the forces of the model to the shakeout: when entry goes down suddenly, there is a strong force toward contraction in firm numbers. In order to show this formally, we develop the dynamics of the model in more detail in the next section.

3.3 Life Cycle Dynamics

We are now ready to characterize the evolution of M and I. We accomplish this by studying the derivative of the two variables as a function of their current levels. We therefore are especially concerned with the set of points where $\dot{I}$ goes from positive to negative, so that the industry goes from rising immature submarkets to declining. We denote this set of combinations of I and M by $I = g(M)$. This is defined by a level of immature innovation sufficient to offset submarkets that either become mature or fail; if there are more submarkets, innovation is less attractive and therefore insufficient to maintain I. Since π is decreasing in both arguments, the greater is M, the less is I to sustain the same level of innovation. We therefore have[16]

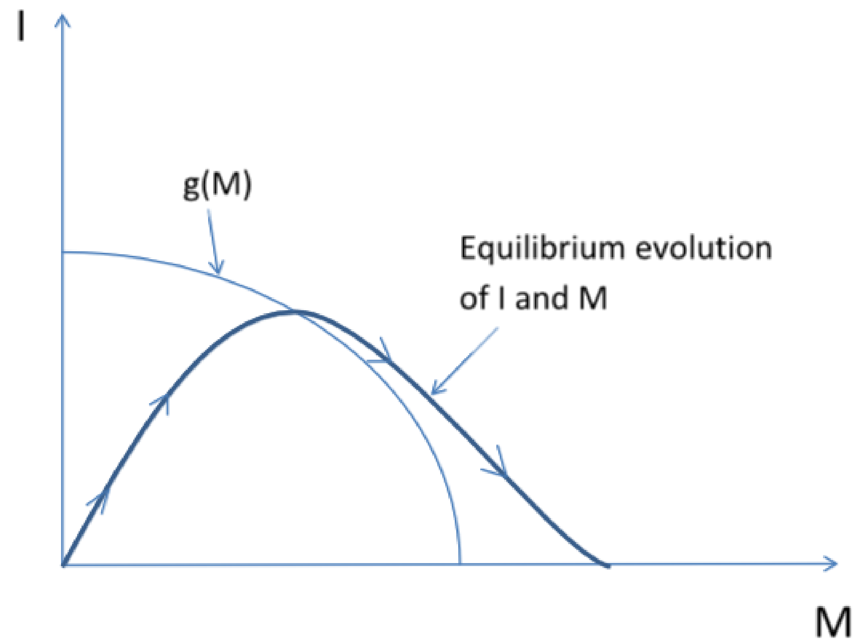
[16] Again the formal proof is in the appendix.

Proposition 4. *There exists a decreasing function $g(M)$ such that, if $I > g(M)$, $\dot{I} < 0$, and if $I < g(M)$, $\dot{I} > 0$*

Imagine an industry that begins with no submarkets. We assume that at this point there is sufficient pioneering innovation for I to grow to some small positive amount. It is a minimal assumption for the industry to grow from a small number of submarkets.

Assumption 4. $g(0) > 0$

We analyze the system using a phase diagram. Everywhere above the M axis, M is rising, since $\dot{M} > 0$ if $I > 0$. I is rising for $I < g(M)$, and falling for $I > g(M)$. The fact that I cannot be falling and M must be rising in the region where $I < g(M)$ implies that once the equilibrium path leaves that region, the equilibrium path can never reenter it. The industry therefore follows a path like the one described by the arrows:



This industry begins with increases in both mature and immature submarkets. Eventually the level of pioneering research cannot sustain the level of immature submarkets, and they are maturing or dying faster than they are being created, leading to a decline in immature submarkets. In the long run all submarkets have matured, and we have a stable set of submarkets that can be thought of as the "dominant design" as in [31]. The stocks I and M correspond to the industry stock of the dynamic capability: dynamic capabilities toward research in mature submarkets are ever growing, while dynamic capabilities in immature industries first rise, and then fall.[17]

[17] We describe in section 5.1 an extension where pioneering innovation persists forever.

4 Implications

The underlying goal of the model is to endogenize the arrival of submarkets from birth to steady state. Using the model we first evaluate its implications for the evolution of the total number of submarkets. We show that our model generates an endogenous evolution of the total number of submarkets which follows an S-shape. The S-shape has been highlighted by [30]. We then show how the S-shape can be followed by a period of a declining number of submarkets. This indicates a possible force behind the shakeout: declining immature submarkets. As this is only suggestive, we then connect the evolution of dynamic capabilities to the shakeout more directly: we show that, as the stock of dynamic capabilities in immature submarkets shrinks, it can lead to a shakeout in firms associated with immature submarkets. Finally, we describe how the underlying innovation mechanism we describe maps to facts about the quantity and type of innovation over the lifecycle.

4.1 Evolution of the Number of Submarkets: the S-shape and possible decline

A natural question in a model that endogenizes the arrival of submarkets is how their number evolves over time. [30] argues that an S-shape is a good assumption for the evolution of submarkets, and fits it to the experience of the tire industry. Moreover, [30] uses an exogenous S-shaped increase in submarkets to explain facts about the industry life cycle. In this section we show that our model generates this S-shaped evolution of submarkets endogenously.

The change in the number of submarkets over time is

$$\dot{N} = \dot{M} + \dot{I} = \phi i I - \lambda I$$

The number of submarkets changes over time as new submarkets arrive (at rate $\phi I i$) or die before reaching maturity (at rate λI); maturity itself simply changes a submarket from immature to mature. We first show that, from the point where immature submarkets are maximized to the point where total submarkets are maximized, growth in N is slowing.

Proposition 5. *From the time where $\dot{I}$ stops being positive until $\dot{N} = 0$, $\ddot{N} < 0$*

Proof. Computing $\ddot{N}$

$$\ddot{N} \;=\; \phi\dot{i}I + (\phi i - \lambda)\dot{I} \tag{8}$$

The first term is negative since i is falling by Proposition 2; the second term is negative as the product of a positive $(\phi i - \lambda)$ and negative $(\dot{I})$ terms. $\square$

This feature implies that submarkets are growing at a declining rate during the period where the number of immature submarkets is falling. On the other hand, the reverse has to be true very early in the industry's evolution:

Proposition 6. *For N near zero, $\ddot{N} > 0$*

Proof. Note that $\dot{N} = \phi iI - \lambda I$, so, since profits and therefore i is bounded, $lim_{N\downarrow 0}\dot{N} = 0$. For N small, therefore, ϕiI must be rising faster than λI, or N would become negative. This implies that $\ddot{N} > 0$. $\square$

Submarkets are rising from zero until $\dot{N} = 0$. The pattern for N is S-shaped: first at an increasing rate, and then at a decreasing rate. Note that this pattern is not a consequence of details of the curvature of the profit function; the only assumption about how π changes is Assumption 3; it is generated entirely by the evolution of submarkets via competition and dynamic capabilities.

The previous results pertain to the period where N is rising. Indeed that may be true throughout the dynamics. On the other hand total submarkets may decline, since $\dot{N} < 0$, if eventually $i < \lambda/\phi$. Since i is decreasing in π, and $i = 0$ if $\pi = 0$, it is clear that there always exists a rate of decline in π such that i falls to the point where $\dot{N} < 0$.

Declining submarket numbers near the steady state is interesting because it is related, intuitively, to the model's ability to generate a shakeout in firm numbers. The steady state of the model mimics[22]. In that model, firm numbers are proportional to the (exogenous) number of submarkets that exist. If submarkets are declining near the steady state, therefore, it seems natural that the model would generate a shakeout. We explore this possibility in the next section. The mechanism comes about through the declining number of immature submarkets, which is the driver of declining N.

4.2 Entry, Exit, and the Shakeout through Immature Submarkets

Authors including [18] and [20] have stressed that innovative early entrants to an industry tend to survive shakeouts. Firms who were early movers at the industry level are more likely, in our structure, to hold leadership positions in mature submarkets, since maturation of submarkets takes time. In this subsection we therefore study a shakeout focusing on firms that are active in the immature sector, since these firms are the most recent entrants to the industry.

4.2.1 Shakeout from the Immature Sector

De novo entry of firms into immature submarkets is

$$E^I = (i - \bar{\imath})I$$

There are two forces behind the evolution of entry. On the one hand, in the early part of the life cycle, I is rising, which increases entry. On the other hand, as i falls, the share of pioneering done by entering firms, $(i - \bar{\imath})/i$, falls. Since entry starts near zero, entry must initially rise to account for the existence of new firms; eventually, $\dot{I} = 0$ and therefore entry falls.

Exit occurs when a firm with a single submarket loses its leadership position. Therefore exit from firms in immature submarkets is

$$X_t^I = (1 - \phi)\omega_t I_t i_t$$

where ω_t is the fraction of immature submarkets led by a firm with a single leadership position. In general, out of steady state, ω_t is difficult to characterize. The discontinuity in i, however, is guaranteed to generate a point where $X^I > 0$.

Proposition 7. *Suppose that at some date t, i drops discontinuously to $\bar{\imath}$. Then $X_t^I > 0$.*

Proof. Since entry is strictly positive for some interval before the drop in i, there must be a positive fraction of firms from that set who still have only one leadership position, i.e. $\omega > 0$. Therefore $X_t^I > 0$. □

Immediately after the discontinuity, de novo entry falls to zero; therefore $X^I > 0$ implies a shakeout among firms operating in the immature sector, since firm numbers change over time by $-X^I$ after the crash in entry, where entry is zero. If $\phi\bar{i} \geq \mu + \lambda$, the shakeout is a necessity; innovation must fall below $\bar{i}$ before $\dot{I} = 0$. On the other hand, if $\phi\bar{i} < \mu + \lambda$, de novo entry may persist. It must be declining, however: beyond the point where $\dot{I} = 0$, E^I is declining since both i and I are falling. In this case the decline in i may or may not be fast enough to generate a shakeout.

Intuitively, exit is a reflection of accumulated past entry and hence changes continuously over time. In contrast, entry may drop discontinuously or very rapidly. In that case, the number of immature firms in the industry must drop. A sudden "crash" in pioneering innovation guarantees the drop is fast enough, but the fall could be sufficiently fast elsewhere. The story of the shakeout in immature firms is that rising competition eventually forces entrants without some competitive advantage out of the industry, lowering entry below exit.[18]

The results in this section focus entirely on the immature sector. This shakeout can apply to the firm numbers as a whole, however. For instance, suppose that the ratio of immature to mature submarkets is very high. This occurs if the maturation rate is very low relative to the death rate for immature submarkets; it takes a large stock of immature submarkets to generate a few successful, mature submarkets. In that case the shakeout, led by a fall in firms operating in the immature sector, will apply to firm numbers as a whole. This line of argument naturally mirrors the notion in [18] and [20] that recent entrants are most susceptible to the shakeout; if that is true, then the shakeout is most likely to occur when the industry has a large number of immature submarkets relative to mature ones, and therefore a relatively large number of young firms.

4.2.2 The Shakeout, Industry Profits, and Consolidation

The driving force in the model is changes in profits over time as competition increases. The model does not necessarily have a prediction about aggregate (or average) profitability over time, though; even though $\pi(M, I)$ is decreasing, the composition of immature and mature submarkets is evolving. For

[18] Note that, although the drop is to zero de novo entry, the model could allow for another stream of de novo entrants (perhaps a limited number with access to a favorable technology) such that entry was positive before and after the discontinuity. The key is that, at some point, a group of potential entrants goes from making zero profits (i.e. free entry for that group holds) to being unprofitable.

instance, an interesting feature of the model is that, at the peak of firm numbers where the shakeout begins, *total* industry profits can be rising, and even average profits per submarket, despite the shakeout being caused by falling profits *per submarket of a given type*. This rise in profits with contraction in firm numbers might appear to be an "industry consolidation," in the sense that fewer firms are generating more profits, but here it is not coming as a result of increased concentration at the firm level, as everything is constant returns and perfectly competitive. The profit effect comes because the composition of submarkets is changing toward mature submarkets.

Whether profits can be rising or falling depends on whether mature submarkets are more or less profitable than immature ones. The industry profit rate per submarket is

$$(M\pi + \alpha I\pi)/(M + I)$$

This is either increasing or decreasing in M/I depending on whether α is smaller than or greater than one. Therefore when $\alpha < 1$, the loss in profits over time through $\dot{\pi}$ is offset if M/I is rising. For $\dot{I}$ negative or positive but low, M/I rises. When $\alpha < 1$, then, the shakeout can look like a consolidation, in terms of profitability, when in fact it simply coincides with the contraction of the less profitable immature sector.

4.3 Innovation over the Life Cycle

Our model generates innovation rates that respond to changing profit rates. In this subsection we describe some relevant features of that evolution over the industry life cycle.

4.3.1 The Composition of Innovation over the Life Cycle

[8] document that the rise in firms is met with a rise in patenting. It must be the case that innovation and firms rise early in the life cycle in our model. Our model, however, allows us to further study the composition of innovation. Total innovation in immature submarkets is iI. Note that the rate of change of this variable is identical to the rate of change of immature entry; they differ by a constant. A constant fraction ϕ of this innovation pioneers new submarkets. As a result, market pioneering peaks before the total number of submarkets; this is consistent with the observation in [16] that "major innovations tend to reach a peak during the growth in the number of producers." In that paper, major innovations are associated with increasing

versions of the product, which is a natural interpretation of the submarkets introduced in our model.[19] In Klepper's model, the return to process innovation changes over time as scale changes. In our model, both the return (through π) and aggregate cost (through the stock of incumbents with experience and a dynamic capability) of both types of innovation can change over time.

Under the interpretation that pioneering innovation corresponds to product innovation, and mature submarkets focus on process innovation, our model is also consistent with [31], who stress that product innovation declines as the dominant design emerges. Since the change over time of pioneering innovation is proportional to $i\dot{I} + \dot{i}I$, this must turn negative before $\dot{I} = 0$. [31] also document a change from innovations that require original components, to ones that focus on adopted components and products, which fits with the notion of pioneering innovation that we use.

4.3.2 Persistently Innovative Industries

[1] stress that mature products might still be very innovative, including having many product innovations. Our model offers at least two interpretations of this fact that industries are persistently innovative. First, there is no necessity to connect product innovation exclusively to new submarkets; one could imagine new leadership positions in existing submarkets coming from either improved functionality or reduced costs.[20] Under the assumption that product innovation is ϕiI, the model replicates the rise and fall of product innovation; under the assumption that product innovation is $\phi iI + mM$, however, product innovation continues indefinitely. Both m and M are strictly positive in the long run. Although m is declining, M is rising; total mature innovation can therefore be either rising or falling. Mature innovation is

$$Mm = M\pi(M, I) + M(\bar{m} - \bar{k}_M - r)$$

The long run characterization of innovation is determined by the shape of π, i.e. the impact of competition on profits. Sufficient conditions for mature innovation to be rising in the latter part of the life cycle, where $\dot{I} \leq 0$, are $\bar{m} - \bar{k}_M - r > 0$, and that $M\pi(M, I)$ increases in M, i.e. competition between

[19] [8] also document a shift from major to minor innovations.

[20] Indeed, the quality ladder model upon which the model is based can be interpreted either as a model of product or process improvements. The details of that underlying model are described in more detail in section 5.3.

submarkets is not too fierce. The first condition can be interpreted as assuming the value of the dynamic capability in mature submarkets is sufficiently large relative to r. That $M\pi(M,I)$ is increasing can be interpreted as increases in M growing the market for mature submarkets sufficiently to offset the lost profits from increased competition. Under these conditions, pioneering innovation is falling in the latter part of the life cycle, while innovation in mature areas remains high.

Our model shares the demand-side characterization of innovation [1], and shares the flexibility that innovation can persist in the long run, or decline, depending on the shape of π. One could imagine that differences in whether mature submarket innovation is product or process would be a natural way to generate different patters of innovation ranging from the ones stressed by [31] to the ones described in [1]. Moreover, we discuss next an extension where pioneering is perpetual, which would further allow for a channel by which product innovation does not decline in the long run, even if one thinks that product innovation is largely in immature areas.

5 Extensions

5.1 Death of Mature Submarkets and Perpetual Market Pioneering

Our model is compatible with permanent pioneering if mature submarkets periodically die. Let mature submarkets be eliminated at rate δ. This alters the value of a mature submarket slightly:

$$rV = \pi(M,I) - (m+\delta)V + (\bar{m}V_t - \bar{k}_M)$$

The more substantive change comes about because of how it impacts the time derivative of M:

$$\dot{M} = \mu I - \delta M$$

Instead of M rising for any $I > 0$, now $\dot{M} = 0$ when $I/M = \delta/\mu$. Below that line, M falls. The steady state, rather than having no immature submarkets, has M where both $\dot{M} = 0$ and $\dot{I} = 0$; since the latter is defined by $g(M)$, this intersection occurs when M solves

$$g(M)/M = \delta/\mu$$

and $I = g(M) > 0$. Since $g(M)$ does not depend on δ, and $g(M)/M$ is decreasing, the steady state number of mature submarkets is decreasing in δ.

There is perpetual market pioneering in the steady state, in order to offset the death of mature submarkets. The steady state is on the $g(M)$ function (where it intersects $\dot{M} = 0$) rather than on the M axis. Since $I = M\delta/\mu$ and $\dot{I} = 0$ when $i = (\mu + \lambda)/\phi$, we can compute the steady state mature submarkets from an analogous equation to (6):

$$\frac{1}{(1-\phi)c}(\alpha\pi(M, M\delta/\mu) - \mu(c-1) - \lambda c) = (\mu + \lambda)/\phi$$

All of the earlier characterization of the shakeout near the point where de novo entry into immature submarkets crashes continues to be true. At the steady state $i = (\mu + \lambda)/\phi$. If this is smaller than $\bar{i}$, it is certain that there is a shakeout; even without it, entry is declining near the steady state, since $\dot{I} = 0$ there, which can generate a shakeout even if $i > \bar{i}$.

5.2 More than one profiting firm per submarket

Suppose that both the leader and second-leader (i.e. the most recently displaced leader) made profits in each submarket. We then have four values to define, for leaders and followers (which we denote 1 and 2, for first and second) for each type of submarket. Denoting profits of the firms by π^1 and π^2 for leaders and followers:

$$\begin{aligned}
rV^1 &= \pi^1(M,I) - m(V^1 - V^2) + \bar{m}V_t - \bar{k}_M \\
rV^2 &= \pi^2(M,I) - mV^2 + \bar{m}V_t - \bar{k}_M \\
rW^1 &= \alpha\pi^1(M,I) - i(1-\phi)(W^1 - W^2) - \mu(W^1 - V^1) - \lambda W^1 + Wf(k_I(W)) - k_I(W) + \dot{W}^1 \\
rW^2 &= \alpha\pi^2(M,I) - i(1-\phi)W^2 - \mu(W^2 - V^2) - \lambda W^2 + Wf(k_I(W)) - k_I(W) + \dot{W}^2
\end{aligned}$$

For leader firms, arrival of an innovation in their submarket knocks them down to followers; followers are eliminated. Maturation maintains the firms rank. Here we impose, as above, that de novo entry is profitable for mature industries, although that is not necessary. Moreover we could allow the dynamic capability to differ for leaders and laggards, by making f differ; here both firms maintain the capability and the value that goes with it. None of this changes the basic mechanisms of the model. The free entry conditions are

$$\begin{aligned} V^1(M,I) &\leq 1, \text{ with equality unless } m < \bar{m} \\ W^1(M,I) &\leq c, \text{ with equality unless } i < \bar{i} \end{aligned}$$

None of the qualitative features of the model are changed; the number of equations describing the equilibrium simply rise. One could extend this analogously to 3 or more profit making firms per submarket.

5.3 Consumer Preferences and Explicit Bertrand Competition within Submarkets

In this section we show how a model of consumers preferences delivers the structure for profits we study above. Suppose that, at each instant, there is a representative consumer with utility function over consumption bundles a across submarkets by

$$\int_0^M ln(a_j)dj + \gamma \int_0^I ln(a_l)dl$$

subject to a fixed budget, normalized to one, to spend on the products:[21]

$$\int_0^M p_j a_j dj + \int_0^I p_l a_l dl = 1$$

From the first order conditions for the two types of products we have that

$$p_j a_j = \gamma p_l a_l$$

Revenue for a mature submarket is

$$R(M,I) = 1/(M + \gamma I)$$

and $\gamma R(M,I)$ for an immature industry.

Price, however, is determined by competition between quality levels, as in [9, 2]. In a given moment of time in submarket j there is a set of firms J_j. Firm $n \in J_j$ can produce the good at a constant marginal cost $b_j^n \leq 1$

[21] One interpretation of the fixed budget is that the consumer has Cobb-Douglas preferences over this industry and an outside good, and therefore has constant spending on the industry's products.

per unit of quality. This allows innovations to be alternatively viewed as product innovations that raise units of quality per unit of cost, or process innovations that simply reduce cost. Firms within a submarket are ordered in a decreasing order of costs. For a given submarket i, the representative consumer consumes a_j^n units of products from firm j. This leads to d_i units from the submarket, where

$$a_i = \sum_j a_i^j$$

In equilibrium consumers will all consume the lowest cost product, denoted simply b_j, for each submarket. We assume that innovation reduces costs per quality unit by a factor $\beta > 1$. That is, if in a given submarket the lowest cost firm j has a cost b_i^j, if a new improvement is developed, it results in costs $b_i^{j+1} = b_i^j/\beta$. The first firm to operate in a submarket has cost $b_i^1 = 1/\beta$. For simplicity we assume that, in each submarket, if only one product has been invented, the consumers have an outside option that is provided competitively at marginal cost 1. One can interpret this as the next best alternative product that might substitute for submarket i.[22]

Non-lowest-cost firms price at marginal cost; to match this price, the lowest cost producer charges $p_i^j = 1/\beta^{j-1}$ for $j > 1$ and $p_i^1 = 1$. Profits for a mature industry are therefore

$$R(M, I)(1 - w/p) \quad = \quad R(M, I)(1 - 1/\beta)$$

Note that, if one wants to have immature industries have higher profits, despite having the industry more competitive as firms mature, one can make immature firms have a greater β to overcome $\gamma < 1$. In the language of section 2 where immature firms earned α times what mature firms earn,

$$\alpha = \gamma \frac{(1 - \beta_i)\beta_m}{\beta_i(1 - \beta_m)}$$

Note that as long as $\gamma < 1$, Assumption 3 can be met for suitably chosen μ and λ; meanwhile α can be either bigger or smaller than one.

[22] Alternatively, the first entrant would set price equal to the unconstrained monopoly price. This would force us to have three types of firms: mature firms, immature firms, and first-movers; nothing substantive about the model or its predictions would change.

6 Conclusions

Dynamic capabilities can be an important driver of the industry life cycle. In the model introduced here, industries evolve as the set of submarkets changes over time. Those submarkets start out immature, but some survive to maturity. Consistent with empirical evidence, we model incumbency as generating an advantage at innovation in areas related to the incumbent's position. The model generates industry life cycle dynamics that are consistent with a variety of empirical regularities, including the shakeout of firms as the industry approaches maturity, and the evolution of innovation over the industry life cycle. The model demonstrates the central role that dynamic capabilities can have in the evolution of industry.

Our model takes, as its base, the model of dynamic capabilities contained in [22]. We modify their model to take account of the fundamental feature of dynamic capabilities: that the advantage they confer applies to related product areas. We show how such a model can be used not only to make steady state predictions of the sort highlighted in [22], but also to study the non-stationary evolution from an industry's birth. The model shares the desirable steady state features of [22], while expanding it to better understand the industry life cycle and the role dynamic capabilities play in its development.

Appendix

Proof of Proposition 2

To show this, we first show the kink in W: at any such date where entry ends, W is strictly decreasing, and the the change in the slope of W is discontinuous. Denote that date by T.

Lemma 5. *$lim_{\epsilon \downarrow 0} \dot{W}_{T+\epsilon} < 0$ whenever $I > 0$*

Proof. If $lim_{t \downarrow T} \dot{W} = 0$, then W is differentiable at T and T is a local maximum of W, with $\dot{W}_T = 0$ and $\ddot{W}_T \leq 0$. But then, from (9), since $\dot{\pi}_T < 0$, $\dot{W}_T < 0$, a contradiction.

Value W is continuous at T, since it is an integral of future expected profits, and therefore cannot change suddenly. A kink in W requires a discontinuity in i, therefore, to offset the sudden change in $\dot{W}$, and therefore proposition 2 is immediate from continuity of W and (5) □

Lemma 6. *Suppose the free entry condition begins binding at T. Then i is continuous at T and $\dot{\pi} \geq 0$ for t approaching T*

Proof. Note that $\dot{W} \geq 0$ for t near T (since W is approaching its upper bound, c), and i is rising. W is clearly continuous, and W in (5) can only be varying continuously at T if $lim_{t \to T}\dot{W} = 0$ and i varies continuously at T. Differentiating (7):

$$\dot{W} = \frac{1}{r + \mu + \lambda - \phi i - k'(W)(\phi W f_I'(k) - 1)}(\alpha\dot{\pi} + \ddot{W}) \tag{9}$$

Note that, since $Wf' - 1 = 0$ by the first order condition for the choice of k, the last term $\phi Wf' - 1 < 0$, and therefore the denominator is positive since k is increasing in W. Since $\dot{W} \geq 0$ near T and $lim_{t \to T}\dot{W} = 0$, $\ddot{W} \leq 0$ as T is a local maximum of W, and therefore $\dot{\pi} \geq 0$ to make this expression positive. □

Proof of Proposition 4

Proof. First, suppose that parameters are such that $\phi\bar{i} > \mu + \lambda$. In this case, if there were de novo entry, i.e. $i > \bar{i}$, new submarket generation would be $\phi i > \phi\bar{i} > \mu + \lambda$, so the number of immature submarkets would be rising, $\dot{I} > 0$. As a result, $g(M)$ must occur where de novo entry is exactly zero and $W < c$ so that existing firms generate less than $\phi\bar{i}$ in new submarkets. In that case, $g(M)$ is defined by the set of points where $W(M, I)$ reaches a point where

$$\phi f_I(k_I(W(M, I))) = \mu + \lambda$$

so that innovation by incumbents is $i = (\mu + \lambda)/\phi$. Since W is decreasing in this region in M and I, $g(M)$ is decreasing.

On the other hand, if $\phi\bar{i} < \mu + \lambda$, the characterization of g is more complicated. Either $\dot{I} = 0$ when de novo entrants are generating new submarkets, or I goes from increasing to decreasing at precisely the point where entry crashes. In the former case, define $g_1(M)$ to be the set of points where (6) implies $i = (\mu + \lambda)/\phi$. Since π is decreasing in both arguments, g_1 is decreasing. From the prior section, g_0 is the set of points where de novo entry crashes. Clearly if de novo entry crashes when $\dot{I} > 0$, this describes the points where $\dot{I}$ changes signs; therefore define

$$g(M) = min\{g_0(M), g_1(M)\}$$

Since both g_0 and g_1 are decreasing, g is decreasing. In both cases, I is decreasing if $I > g(M)$, and increasing if $I < g(M)$. $\square$

References

[1] Ron Adner and Daniel Levinthal. Demand heterogeneity and technology evolution: Implications for product and process innovation. *Management Science*, 47(5):611–628, 2001.

[2] Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–51, 1992.

[3] Guido Buenstorf and Steven Klepper. Submarket dynamics and innovation: the case of the us tire industry. *Industrial and Corporate Change*, 19(5):1563–1587, 2010.

[4] Wesley M. Cohen and Steven Klepper. Firm size and the nature of innovation within industries: The case of process and product r&d. *The Review of Economics and Statistics*, 78(2):232–243, 1996.

[5] J.M. De Figueiredo and M.K. Kyle. Surviving the gales of creative destruction: The determinants of product turnover. *Strategic Management Journal*, 27(3):241–264, 2006.

[6] A.M. Franco and D. Filson. Spin-outs: knowledge diffusion through employee mobility. *The RAND Journal of Economics*, 37(4):841–860, 2006.

[7] Michael Gort. *Diversification and Integration in American Industry*. Princeton University Press, Princeton, NJ, 1962.

[8] Michael Gort and Steven Klepper. Time paths in the diffusion of product innovations. *The Economic Journal*, 92(367):630–653, 1982.

[9] Gene M. Grossman and Elhanan Helpman. Quality ladders in the theory of economic growth. *Review of Economic Studies*, 58:43–61, 1991.

[10] Sebastian Wilfling Guido Buenstorf, Christina Guenther. Trekkers, bulldogs and a shakeout: submarkets and pre-entry experience in the evolution of the german farm tractor industry. 2012.

[11] C.E. Helfat. Know-how and asset complementarity and dynamic capability accumulation: The case of r&d. *Strategic Management Journal*, 18(5):339–360, 1997.

[12] C.E. Helfat, S. Finkelstein, W. Mitchell, M.A. Peteraf, H. Singh, D.J. Teece, and S.G. Winter. *Dynamic Capabilities: Understanding strategic change in organizations*. Blackwell Pub., 2007.

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

[14] D.J. Kim and B. Kogut. Technological platforms and diversification. *Organization Science*, 7(3):283–301, 1996.

[15] A.A. King and C.L. Tucci. Incumbent entry into new market niches: The role of experience and managerial choice in the creation of dynamic capabilities. *Management Science*, 48(2):171–186, 2002.

[16] S. Klepper. Entry, exit, growth, and innovation over the product life cycle. *The American Economic Review*, pages 562–583, 1996.

[17] Steven Klepper. Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 86(3):562–583, 1996.

[18] Steven Klepper. Industry life cycles. *Industrial and Corporate Change*, 6(1):145–182, 1997.

[19] Steven Klepper. Firm survival and the evolution of oligopoly. *RAND Journal of Economics*, 33(1):37–61, Spring 2002.

[20] Steven Klepper and Kenneth Simons. Industry shakeouts and technological change. *International Journal of Industrial Organization*, pages 23–43, 2005.

[21] Steven Klepper and Peter Thompson. Submarkets and the evolution of market structure. *RAND Journal of Economics*, 37(4):861–886, 2006.

[22] Tor Jakob Klette and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018, 2004.

[23] M. Lieberman and D. Montgomery. First-mover advantages. *Strategic Management Journal*, 9(S1):41–58, 1988.

[24] Matthew Mitchell. The scope and organization of production: firm dynamics over the learning curve. *RAND Journal of Economics*, 31(1):180–205, Spring 2000.

[25] Paul M Romer. Growth based on increasing returns due to specialization. *American Economic Review*, 77(2):56–62, 1987.

[26] F.M. Scott Morton. Entry decisions in the generic pharmaceutical industry. *The Rand journal of economics*, pages 421–440, 1999.

[27] John Sutton. *Technology and Market Structure*. MIT Press, Cambridge, Mass., 1998.

[28] D. Teece and G. Pisano. The dynamic capabilities of firms: an introduction. *Industrial and corporate change*, 3(3):537–556, 1994.

[29] D.J. Teece. *Dynamic capabilities and strategic management: organizing for innovation and growth*. Oxford University Press, USA, 2009.

[30] Jian Tong. Explaining the shakeout process: a successive submarkets model. *Economic Journal*, 119(537):950–975, 2009.

[31] James Utterback and William Abernathy. A dynamic model of process and product innovation. *Omega*, 3(6), 1975.