Making Waves: To Innovate or Be a Fast Second?

Chris Yung∗

Abstract

Internal finance leads to a stalemate in innovation games; each firm wants to free-ride on the others’ costly experimentation. When instead innovation is financed externally (e.g., with venture capital or initial public offerings), there is an endogenous cost to delay. Waiting to make risky irreversible investment conveys pessimistic information. I characterize the relative sizes of waves of leaders and followers in innovation cycles, and the endogenous, intertemporal distribution of quality as each wave builds and crashes. Finally, old waves leave an adverse selection “hangover,” such that too much early innovation can cause the market for future innovation to break down.

I. Introduction

Innovation requires costly and time-consuming experimentation. Yet it is unclear whether rational competing firms have much incentive to bear these costs (Markides and Geroski (2004)). Why not instead simply imitate competitors’ innovations and not bear the cost of the failures and false starts? Without a solution to this problem, no innovation can occur. The standard solution to this stalemate is to endow innovators with market power (arising either naturally or by legal construction) or to assume a first-mover benefit in the product market (e.g., brand loyalty).

In this paper, I demonstrate that external financing reduces the incentive to wait and learn from the innovations of others, thereby breaking this stalemate. This effect is quite distinct from, and may exist in addition to, any first-mover advantage on the product-market side. The model of sequential innovation in this paper addresses several features of innovation cycles: i) the difference in leaders’ and followers’ intrinsic quality, ii) the timing of wave propagation and crashes, iii) how this timing depends on the risk of the innovation and the amount of external finance, and iv) how future innovation is tied to the current levels of innovation.

∗Yung (corresponding author), chris.yung@virginia.edu, University of Virginia, McIntire School of Commerce, Charlottesville, VA 22904. I thank Zhaohui Chen, Michael Gallmeyer, Jarrad Harford (the editor), Andrey Malenko (the referee), Gustav Martinsson, Bruce Peterson, Roberto Pinheiro, Morten Sorensen, Luke Taylor, Bill Wilhelm, and Jaime Zender for helpful comments.
The association between innovation and external financing is consistent with casual empiricism as well as recent academic work. The 1990s research and development (R&D) wave was almost entirely driven by industries with large influxes of equity capital (Brown, Fazzari, and Peterson (2009)). Public firms that are dependent on external financing generate more numerous and higher quality patents than their private counterparts, whereas this does not hold for internally financed firms (Acharya and Xu (2016)). Consequently, a recent branch of literature gives a central role to financial market development in fostering innovation.¹

Two assumptions are key to establishing this paper’s main result. First, firms are heterogeneous. Payoffs to taking the innovative action are risky, having both a public component (observable and common to all) and a private, firm-specific component. Second, firms must obtain external financing, so that payoffs are shared by the firm and the financier. Convincing others to supply funds requires proving that one’s quality is of an acceptable level.

All firms prefer more information about the common signal, so there is a benefit of waiting and free-riding on the costly experimentation of others. Yet this incentive to delay is weaker for good firms, whose projects remain profitable over a wider range of parameters. Equilibrium security pricing therefore reflects an orderly process in which the best firms lead, and weaker ones enter the market if and only if the pioneers do well. This pricing scheme in turn creates a cost to delay; it is better to be treated as a high-quality issuer. Now the comparative static discussed previously becomes clear: As the amount of internal finance increases, this cost of delay becomes irrelevant, so that all firms wait.

In a recent paper, Grenadier, Malenko, and Strebulaev (2014) also consider how timing reveals information. Theirs is a model of disinvestment. Firms strategically delay abandonment in an attempt to signal quality. This delay leads to a large cohort of “living dead” firms that are unwilling to publicly admit defeat. These admissions finally come en masse after a negative public shock, so that firms can (in effect) pin their failures on a common source. Comparing my results to those of Grenadier et al. illustrates an asymmetry related to the passage of time. Grenadier et al.’s bad firms blend in with the crowd by disinvesting after a negative shock. In my context, bad firms would prefer to blend in with the crowd by investing before a positive public shock. Of course, such a strategy is not feasible because positive shocks cannot be predicted.

To summarize the patterns of leadership and imitation in the model:

i) External financing accelerates innovation. Thus, cross-sectionally, within a given industry, the firms that are most reliant on external financing will be most innovative.

ii) There is an adverse selection discount to being a follower. This discount will be reflected in late entrants’ lower valuation ratios (for public firms) and difficulty in raising funds (for private firms) and is justified ex post by lower productivity.

¹Small, young firms are more R&D intensive in financially developed countries (Brown, Martinsson, and Petersen (2013)) than their counterparts in developing countries, and the development in equity markets leads to more technological innovation (Hsu, Tian, and Xu (2014)).
iii) This discount to waiting will be larger when the innovation is less path-breaking. If risk is low, many firms enter the market before uncertainty is resolved. Consequently, the quality of inactive firms is very low. Note that this adverse selection hangover remains (it is very bad news to enter late) even after all of the common signal’s risk is resolved.

iv) Temporally, there is less of an adverse selection discount immediately after new investment opportunities open up because there has been less time for the highest quality firms to be picked off.

v) Overinvestment is guaranteed to occur in the last stage of the game, corresponding to the tail of the wave. Once the wave crashes, financiers are unwilling to supply funding to any agent.

Related Literature

As previously mentioned, the first bullet point is consistent with recent empirical work connecting innovation and external financing. The second point is consistent with Banerjee, GÜÇBILMEZ, and Pawlina’s (2016) recent study of initial public offering (IPO) waves. They argue that followers appear to be of lower quality than leaders on several dimensions, including valuation and under-pricing.

A growing literature introduces asymmetric information into real options signaling models. Grenadier and Wang (2005) allow entrepreneurs to divert cash from shareholders; high-quality agents delay investment. Bergemann and Hege (2005) and Bouvard (2014) consider private benefits from running projects. This friction accelerates investment because managers rush to get their projects under way in order to accelerate their stream of private benefits. Consequently, Bouvard argues that financial contracts should reward investment delay. In the context of my model, such arrangements would be counterproductive because delay is preferred by low-quality issuers. In the works of both Chamley and Gale (1994) and Grenadier (1999), agents try to learn from one another’s private information as revealed via their investment decisions. Such an environment can lead to information collapse in which agents ignore their own information and fail to invest even when doing so is efficient. This collapse is reminiscent of the herd behavior described in earlier literature, but occurs even with endogenous ordering of investment decisions. There is no role played by external financing.

Decamps and Mariotti’s (2004) analysis is perhaps closer still to the present model because it assumes that information is not merely revealed by others’ actions, but rather is actually produced by such actions. Each player would prefer to free-ride on the others’ experimentation costs by letting the other agents go first. Agents strategically delay in order to convince rivals that their opportunity costs of experimentation are high. A discount rate provides a countervailing force to invest early. Agents all have the same gross payoffs from taking the project, so there is no notion of agent quality. No role is considered for external financing.

The connection between innovation incentives and the form of finance is considered by Ferreira, Manso, and Silva (2014). An entrepreneur chooses
between public and private financing. The entrepreneur may be forced to sell his or her shares early in response to a liquidity shock. It is shown that private financing is better suited to motivating innovation. Private financing generates less interim information. These uninformative prices serve, in effect, as a tolerance for failure. In contrast to the current paper, there is a single innovating agent. There is no distinction between internal and external financing.

Perhaps the most general model in this literature is that by Grenadier and Malenko (2011), who show that managerial myopia accelerates innovation, which is the same effect that security issuance has in my paper. Grenadier and Malenko develop a general and useful set of rules for when a particular friction accelerates or slows down innovation, and when the ordering of firms is best-to-worst (as in the current paper) or worst-to-best. Two features distinguish my model from this literature. First, in my model firms learn from one another: If there are no leaders, then no new information is produced. Second, I focus on the role of internal and external financing.

The role of security issuance waves and their connection to issue prices has also been considered in a growing finance literature on IPOs (reviewed in Table 1). A focus in many of these models has been on IPO “underpricing,” defined as the discount that issuers give investors to compensate them for either i) de novo information production, or ii) the winners’ curse due to asymmetric information across investors. These models typically lack either endogenous timing or private information of the selling shareholders, and therefore test a different set of insights than those developed here. For example, Pastor and Veronesi (2005) use an exogenous pace of innovation and no private information.

The literature on the multiarmed bandit problem also has a stream of related papers that consider strategic experimentation, that is, when an agent experiments on which lever to pull while also observing others pull similar levers. This environment leads to underinvestment (Bolton and Harris (1999)). Manso (2011) and Ederer (2012) study optimal compensation schemes in such an environment. My setting involves different economic intuition: With security issuance and asymmetric information, agents care about inferences made by outsiders. As previously discussed, this can provide incentives to enter the market earlier than one would otherwise.

Welch (1996) and Colak and Gunay (2011) offer models in which the ordering of entry is reversed from that shown here. High-quality firms strategically wait. In Welch’s model, the driving force is that firm quality is stochastically revealed. Waiting is therefore an effective signal because high-quality issuers forecast better pricing later. Low-quality firms cannot imitate this strategy because their quality would be revealed. Colak and Gunay (2011) note a constant cost of waiting (i.e., pure time value of money), which higher quality firms are better able to bear. Therefore, they again strategically delay in order to signal their quality.

---

2Bustamante (2012) and Morellec and Schürhoff (2011) also mention this acceleration in earlier work. Morellec and Schürhoff generalize the model in yet another direction, showing that endogenous timing may upset the pecking order.

3In effect, security issuance leads to managerial myopia because of dilution.
Table 1 reviews theoretical models of capital raising in which timing plays a key role.

<table>
<thead>
<tr>
<th>Panel A. Financing Models with Exogenous Ordering of Issuance</th>
<th>Panel B. Financing Models with Endogenous Ordering of Issuance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Issuer Heterogeneity</strong></td>
<td><strong>Allocation of Information</strong></td>
</tr>
<tr>
<td>Hofmann-Burchardi (2001)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Investors produce costly information.</td>
</tr>
<tr>
<td>Boveniste, Busaba, and Wilhelm (2002)</td>
<td>All information is public.</td>
</tr>
<tr>
<td></td>
<td>Underwriters endogenously create one by taxing followers, preventing market failure.</td>
</tr>
<tr>
<td>Person and Warther (1997)</td>
<td>All information is public.</td>
</tr>
<tr>
<td>Alti (2005)</td>
<td>Investors are endowed with heterogeneous information.</td>
</tr>
<tr>
<td>Pastor and Veronesi (2005)</td>
<td>All information is public.</td>
</tr>
<tr>
<td>Banerjee, Gucbilmez, and Pawlina (2016)</td>
<td>All information is public.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
II. The Model

Projects have gross returns $\tilde{X} = iZ$. The random variable $i$ takes values $i \in \{0, 1\}$, where the probability $\pi_i = \Pr(i = 1)$ is private information. These success probabilities are uniformly distributed along the interval $[i_{\text{MIN}}, 1]$. The multiplicative factor $Z \in \mathbb{R}$ is common to all firms. Agents have internal capital $V$ to contribute to the project. However, the project also requires another capital contribution of $K$, which must be raised externally.

Consider what happens if one subset of agents invests before the others. Suppose further that these investments generate more information about $Z$. This process is modeled here as a coin flip where “heads” represents more favorable news about $Z$ and “tails” represents bad news. Upon observing this news, another batch of agents may choose to seek funds if doing so is optimal. Their news may in turn trigger a third wave, and so on. I assume that if no agent enters the market, there is no coin flip at all, whereas if any agent enters the market, then a coin flip occurs, with errors given by the fixed number $e_j$ described as follows:

$$\Pr(\text{Heads on flip } j | Z = 0) = e_j,$$
$$\Pr(\text{Heads on flip } j | Z = 1) = 1.$$  

Consequently, waves can end for two reasons. The first scenario is that upon observing a tail, it follows that $Z = 0$, and therefore no investment can occur. In addition, in some settings, no one is willing to enter the market given current information about $Z$, even when its expectation is positive. Yet, if no one is willing to enter the market, no further coin flips will occur, and information about $Z$ can never change, causing a permanent stalemate. Thus:

$$Z_i = \Pr(Z = 1 | i - 1 \text{ straight heads}) = \frac{p}{p + (1 - p)\prod(e_1 \times \cdots \times e_{i-1})}.$$  

Clearly, $Z_i$ is an increasing function of $i$. In fact, denoting

$$p_i = \Pr(\text{heads} | i - 1 \text{ straight heads}),$$

then by the martingale property, $Z_i = p_iZ_{i+1}$.

Whereas $Z_i$ is increasing, $p_i$ is neither increasing nor decreasing in general. The number $p_i$ is high during times in which very little risk is resolved by leaders, so that $Z_{i+1}$ is close to the prior $Z_i$. By contrast, when $p_i$ is low, innovation is very risky.

The model has equilibria without any investment. These equilibria are supported by the belief that anyone attempting to raise capital is of insufficient quality to merit funding. Because these outcomes are trivial, all subsequent discussion restricts attention to equilibria in which a positive amount of investment occurs.

Every Bayesian Nash equilibrium must have the following form. There is some subset of agents that are willing to go first, before any coin is flipped.

---

4More specifically, I assume that it takes a positive mass of agents to generate a signal. To be more realistic, the errors $e_i(\cdot)$ would be a smoothly decreasing function of the measure of the wave. I have solved some special cases with this property, but have not been able to solve that alternative model in general.
Denote the equity stake demanded by capital suppliers $\alpha_1$ for this wave. Next, another subset of agents is willing to enter the market if and only if there is a heads revealed in the first wave.\footnote{Entry cannot occur if a tail is observed. The payoffs would be 0 for sure, whereas agents would forgo $V$ by entering.} Denote the associated equity stake by $\alpha_2$. The process may continue indefinitely until there is a point at which no agents are willing to enter.

As the following lemma shows, in every Bayesian Nash equilibrium, agent entry must occur in an orderly fashion from highest quality to lowest quality.

**Lemma 1.** In any Bayesian Nash equilibrium,

A) Let $\{\alpha_i\}$ be the equity stakes demanded by investors in wave $i$ in exchange for supplying capital $K$. This sequence satisfies $\alpha_j > \cdots > \alpha_1$.

B) There exists a set of cutoff qualities $\{\pi_i\}$. Agents along the interval $[\pi_1, 1]$ enter the market first, agents along the interval $[\pi_2, \pi_1]$ go next, and so on.

**Proof.** See the Appendix.

The chain of inequalities in Lemma 1A underlies the central tradeoff in this paper. There must be a “punishment” for waiting in equilibrium. This punishment is rational for financiers because waiting is associated with lower quality agents in equilibrium. The tradeoff faced by any agent is as follows: The cost of waiting is that the terms of finance get worse. The benefit of waiting is to improve the information available; learning about $Z$ because it saves a loss of $V$ in some bad states. For high-quality firms, this benefit is smaller because their investments are profitable over a larger range of values for $E(Z)$. Because the benefit of waiting is reduced, high-quality agents are willing to issue earlier.

Were there no punishment to waiting, all agents would prefer to wait. Doing so would yield both better information about $Z$ and preferable terms of finance. Yet, if all agents prefer to wait, there would be no coin flip and therefore no information produced.

**Equilibrium**

The condition for the marginal (lowest quality) agent in the $i$th wave is as follows:

\[
(1 - \alpha_i)Z_i\pi_i = p_i(1 - \alpha_{i+1})Z_{i+1}\pi_i + (1 - p_i)V.
\]

The left-hand side is the expected profit from entering the $i$th wave (i.e., after having observed $(i - 1)$ heads). On the right-hand side, the first term reflects the probability that an $(i + 1)$st wave will occur multiplied by the expected payoff in that event. The last term on the right-hand side reflects the outcome if the agent intends to wait for the $(i + 1)$st wave, only to find that this wave never occurs because the coin flip at time $i$ shows tails.

The participation constraint of investors in the $i$th wave is as follows:

\[
\alpha_iZ_i \left( \frac{\pi_i + \pi_{i+1}}{2} \right) = K.
\]
The parenthetical term is the average firm quality in the waves; therefore, the left-hand side represents the expected payoff to financiers. The equilibrium variables \( \{ \pi_i, \alpha_i \} \) can then be obtained by solving the system of 2\( J \) linear equations given by \( \{(1),(2)\} \) for \( i \in \{1, \ldots, J\} \).

In most of what follows, assume that \( e_1 = 0 \), so that \( Z \) is revealed with certainty after the first wave (see Section III.D for exploration of the more general case). Agents then fall into one of three categories:

i) First wave: Enter the market before \( Z \) is revealed (i.e., agents along the interval \([\pi_1, 1]\)).

ii) Second wave: Enter only if \( Z = 1 \) (i.e., agents along the interval \([\pi_2, \pi_1]\)).

iii) Abstain: Do not enter the market regardless of \( Z \)'s value (i.e., agents along the interval \([0, \pi_2]\)).

In this case, by equation (1), the cutoffs \( \pi_1 \) and \( \pi_2 \) are determined by the following equations:

(3) \[ p(1 - \alpha_1)\pi_1 = p(1 - \alpha_2)\pi_1 + (1 - p)V, \]

(4) \[ (1 - \alpha_2)\pi_2 = V, \]

where \( p \) is the unconditional probability that \( Z = 1 \).

The left-hand side of equation (3) is the payoff from entering in the first wave, before the coin flip. With probability \((1 - p)\), the subsequent coin flip reveals \( Z = 0 \), and so the payoff is 0. With probability \( p \), instead the coin flip reveals \( Z = 1 \). In this state the payoff is the agent’s residual claim \( 1 - \alpha_1 \) times the agent’s quality \( \pi_i \). The right-hand side of equation (3) represents the payoff from waiting and entering the second wave, which occurs only if the coin flip is heads.

Note that \( p \) is absent from equation (4). This is because if a second wave occurs, then \( Z = 1 \) for sure. The agent’s payoff is simply the residual claim \( 1 - \alpha_2 \) multiplied by the (privately known) quality. An agent with \( \pi_2 \) is indifferent between issuing and exiting the market.

Solving \( \{(3),(4)\} \) for \( \{\pi_1, \pi_2\} \) results in the following:

(5) \[ \pi_1 = \frac{(1 - p)V}{p(\alpha_2 - \alpha_1)}, \]

(6) \[ \pi_2 = \frac{V}{1 - \alpha_2}. \]

To complete the description of the equilibrium, I now endogenize \( \alpha_i \). The participation constraints of investors define the equity stakes demanded:

(7) \[ \alpha_1 \left( \frac{\pi_1 + 1}{2} \right) p = K, \quad \text{First wave}, \]

(8) \[ \alpha_2 \left( \frac{\pi_1 + \pi_2}{2} \right) = K, \quad \text{Second wave}. \]

The equilibrium \( \{\pi_1, \pi_2, \alpha_1, \alpha_2\} \) is then determined by the system of four equations (5)–(8). Because this system is nonlinear, existence is neither
guaranteed nor unique. Furthermore, the system of equations (5)–(8) has solutions that are economically uninteresting. For example, if \( \pi_2 > \pi_1 \), then there are no followers; if \( \pi_1 > 1 \), then there are no leaders. In what follows, I further restrict attention to solutions for which \( 0 < \pi_2 < \pi_1 < 1 \) so that there are both leaders and followers. Doing so avoids the degenerate (and uninteresting) stalemate cases described in the Introduction in which no one wants to invest.

*Theorem 1. Comparative Statics for \( p \).* Any Bayesian Nash equilibrium has the following properties:

A) \( \frac{\partial \pi_1}{\partial p} < 0 \),

B) \( \frac{\partial \pi_2}{\partial p} > 0 \).

The result in Theorem 1A is intuitive. If the leaders resolve less risk, there is less benefit to delay. Many firms enter the first wave to avoid being identified as low-quality firms. Thus, high \( p \) is associated with large first waves.

Theorem 1B is somewhat more subtle. There is no direct dependence of \( \pi_2 \) on \( p \) in equation (10). In particular, by the time wave 2 occurs (if at all), the probability \( p \) is no longer directly relevant to financiers. Once it has been publicly revealed that \( Z = 1 \), the ex ante probabilities of this state are sunk. Yet there is an important sense in which history matters in this scenario because of the following indirect effect. As \( p \) rises, \( \pi_1 \) falls, and more agents choose to enter the first wave. Consequently, the quality of agents remaining in the second pool drops. Thus, when \( p \) is high, there is a strong adverse selection effect from being identified as a second-wave agent. This adverse selection effect worsens pricing in the second wave, causing \( \pi_2 \) to rise.

Theorem 1 implies that high \( p \) is associated with a large first wave and small second wave. Low \( p \) is associated with the opposite. These comparative static results on wave sizes are important for subsequent results.

*Theorem 2. Comparative Statics for \( V \).* Any Bayesian Nash equilibrium has the following properties:

A) \( \frac{\partial \pi_1}{\partial V} > 0 \),

B) \( \frac{\partial \pi_2}{\partial V} \) is ambiguous. When \( p \) is low, \( \frac{\partial \pi_2}{\partial V} < 0 \). When \( p \) is high, \( \frac{\partial \pi_2}{\partial V} > 0 \).

The intuition for Theorem 2A is straightforward. As \( V \) rises, the net value of projects falls. Intuitively, one would expect fewer agents to seek financing. The theorem states that this intuition is correct for the first wave.

Theorem 2B reflects competing effects. The direct effect is that higher \( V \) reduces the incentive to seek funding in general (just as in Theorem 2A). This tends to push \( \pi_2 \) upward. However, because fewer agents enter the first wave, the quality of inactive agents is higher, and thus there is less of an adverse selection punishment to waiting. This indirect effect leads to higher equity prices in the second round, pushing \( \pi_2 \) downward.

Which effect dominates? To see this heuristically, rewrite equation (5) as

\[
p (\alpha_2 - \alpha_1) \pi_1 = (1 - p) V.
\]

This formulation emphasizes that when \( p \) is low, the endpoint \( \pi_1 \) is very sensitive to \( V \). Thus, the indirect effect described previously is strong. Loosely, the
market is more forgiving of cautiousness when it recognizes i) that opportunity
costs are high, and ii) a second wave was unlikely anyway.

Overall, the effect of internal capital is to shrink the size of the first wave
but not (necessarily) the second wave. This recalls the intuition of the stalemate
discussed in the Introduction: *Internal capital discourages leadership in risky
situations.* If one can observe the outcome of others’ experiments, it makes more
sense to follow than to lead. It is now shown that the comparative statics with
respect to $K$ (external capital) tell a different story.

**Theorem 3. Comparative Statics for $K$.** Any Bayesian Nash equilibrium has the
following properties:

A) $\frac{\partial \pi_1}{\partial K}$ is ambiguous. When $p$ is low, $\frac{\partial \pi_1}{\partial K} > 0$. When $p$ is high,
$\frac{\partial \pi_1}{\partial K} < 0$.

B) $\frac{\partial \pi_2}{\partial K} > 0$.

Theorems 2 and 3 are highly symmetric. In both cases, one cutoff $\pi_i$ unam-
biguously moves upward in response to an increase in cost, whereas the other cut-
off’s movement is ambiguous. In both cases, the ambiguity results from a tradeoff
between the direct effect and an indirect effect. To see why the ambiguity arises
in different endpoints, note the following:

\[ \pi_2 = \frac{V}{1 - \alpha_2} \quad \text{and} \quad \pi_1 = \frac{(1 - p)V}{p(\alpha_2 - \alpha_1)}. \]

In these equations, $K$ shows up only indirectly through its effect on $\alpha_i$. As
$K$ increases, investors supply more capital and demand larger stakes. Thus, $\alpha_2$
rises, and hence $\pi_2$ also rises. By contrast, as the equation on the right shows, the
change in $\pi_2$ depends on $\alpha_2 - \alpha_1$. This difference is ambiguous: It depends on the
relative size of the direct and indirect effects described after Theorem 2, which in
turn depends on $p$.

### III. Extensions

This section further examines the properties of the model and alters it in
several ways. Section III.A introduces a first-mover advantage into the model.
Financing in the base model can be viewed as creating an endogenous first-mover
advantage because it provides an incentive to go early; Section III.A considers
how these two first-mover advantages interact. Section III.B examines the extent
to which private information matters. Sections III.C and III.D characterize the
incentives for innovation in more detail.

#### A. First-Mover Advantage

I now introduce a first-mover advantage into the base model by assuming
that projects pay off $M > 1$ if funded in the first wave, and 1 otherwise. Consider
the firm’s decision as to which wave to enter (if any) for a given set of $\alpha_i$. Analysis
similar to that of Section II.B yields the following cutoffs:

\[ p(1 - \alpha_1)M \pi_1 = p(1 - \alpha_2)\pi_1 + (1 - p)V, \]
\[ (1 - \alpha_2)\pi_2 = V. \]
Turning to the problem of capital providers, I obtain the following participation constraints, which define the zero-profit conditions for $\alpha_i$ for a given set of entering firms:

$$\alpha_1 M \left(\frac{1 + \pi_1}{2}\right) p = K, \tag{11}$$

$$\alpha_2 \left(\frac{\pi_1 + \pi_2}{2}\right) = K. \tag{12}$$

Taken together, the previous equations constitute a four-equation, four-unknown system to be solved for $\{\pi_1, \pi_2, \alpha_1, \alpha_2\}$. From this system, Theorem 4 follows.

**Theorem 4.** Any Bayesian Nash equilibrium with both leaders and followers has the following properties:

A) $\partial \pi_1 / \partial p < 0$.
B) $\partial \pi_2 / \partial p > 0$.
C) $\partial \pi_1 / \partial V > 0$.
D) $\partial \pi_2 / \partial V$ is ambiguous. When $p$ is high, $\partial \pi_2 / \partial V > 0$. When $p$ is low, $\partial \pi_2 / \partial V < 0$.
E) $\partial \pi_1 / \partial K$ is ambiguous. When $p$ is high, $\partial \pi_1 / \partial K < 0$. When $p$ is low, $\partial \pi_1 / \partial K > 0$.
F) $\partial \pi_2 / \partial K > 0$.
G) $\partial \pi_1 / \partial M < 0$.
H) $\partial \pi_2 / \partial M > 0$.

Properties G and H of Theorem 4 are intuitive. Clearly, the first-mover advantage tips the scales toward investing early. Otherwise, the behavior of the model is unchanged when a first-mover advantage is introduced, as evidenced by properties A–F.

**B. Perfect Information**

One might ask about the extent to which the behavior of the model depends on private information. It would seem to be a robust notion that high-quality agents issue first, and that low-quality agents watch them to extract critical information. To examine how things change in such a scenario, I now assume that $i$ is no longer private information. The quality of each agent is observable.\(^6\) How that this environment again leads to breakdown: All agents decide to wait, and therefore no innovation occurs. In some sense, this result is unsurprising because perfect information enables outside capital to be used as if it were internal capital. Consequently, the model in this subsection behaves like the base model with only internal finance. As shown previously, this implies that no agents want to enter the market.

**Theorem 5. Perfect Information.** If quality $\tilde{i}$ is public information, all agents prefer to wait until after $Z$ is revealed. Therefore, no innovation occurs.

\(^6\)I thank Morten Sorensen for suggesting this extension.
Proof. The gross profit from entering in the first wave is \( p\pi_i \). The investor demands an equity stake \( K/p\pi_i \), leaving the entrepreneur with the following residual claim:

\[
(1 - \alpha)p\pi_i = \left(1 - \frac{K}{p\pi_i}\right)p\pi_i = p\pi_i - K.
\]

If instead the entrepreneur waits, the expected payoff is \( p(\pi_i - K) + (1 - p)V \), which exceeds equation (13) whenever \( p < 1 \). \( \square \)

As another way to understand this result, note that in the base model the benefit of waiting is more accurate information about \( Z \), whereas the cost is that one conveys worse information about \( i \). Perfect information eliminates that cost, making it optimal for all agents to wait.

C. Motivating Innovation

This paper studies the optimal timing of implementing an innovation that already exists. However, another important problem concerns the ex ante incentives to come up with an innovation at all. These incentives should be directly related to the agent’s profit in the game once it is played out. Some computation (omitted) shows that agents’ utility in the game is equal to the following for those who anticipate entering the first round:

\[
U = p\pi_i - K \left(\frac{\pi_i}{\pi_1(V, K) + 1}\right) - V.
\]

The parenthetical term reflects a mispricing drag (i.e., parenthetical term is >1) for agents who are above average. They forfeit some of their net present value by being pooled with lower quality agents. As reflected by the leading term \( K \), this mispricing drag is relevant only to the extent that external finance is employed. Conversely, the expression characterizes a mispricing benefit (i.e., parenthetical term is <1) for agents who are below average within their wave.

Taking the derivative of utility results in the following:

\[
\frac{\partial U_i}{\partial V} = \frac{2\pi_iK}{(\pi_1 + 1)^2} \frac{\partial \pi_1}{\partial V} > 0.
\]

\[
\frac{\partial U_i}{\partial K} = \frac{2\pi_iK}{(\pi_1 + 1)^2} \frac{\partial \pi_1}{\partial K} - 1 \left(\frac{\pi_i}{\pi_1 + 1}\right).
\]

The term \(-1\) in equation (15) is a direct effect. In the absence of mispricing, firms would simply capture the entire net present value of their projects. As \( V \) increases, all net present values decline because more internal capital is required.
This direct effect causes utility to decline on a one-for-one basis with $V$. The analogous $-1$ term in equation (16) is multiplied by a fraction that reflects the fact that firms do not bear $K$ evenly. Because of mispricing, above-average firms pay more than $K$ in expectation to financiers than do their peers. Therefore, marginal increases in $K$ cause high-quality firms to lose utility on a more than one-for-one basis.

The underbraced terms in equations (15) and (16) illustrate indirect effects on mispricing arising through the changes to the endpoint $\pi_1$. For example, consider equation (15). As $V$ rises, the endpoint $\pi_1$ rises, and firms enjoy the benefit of being separated from the lowest quality firms. Note that the magnitude of this effect is proportional to $\partial \pi_1 / \partial V$ (i.e., this indirect effect matters only to the extent that the endpoint actually does move). According to Theorem 2A, $\partial \pi_1 / \partial V > 0$. However, no bound is given for its size, and therefore the expression in equation (15) cannot be signed. Ambiguity also holds for equation (16).

To see how incentives change, it is instructive to consider how $V$ and $K$ affect differences in utility across firm types. In particular, a change may reduce firm utilities across the board, but reduce them faster for low-quality firms. For example, consider two agents with $\pi_i^*$ and $\pi_i^{**} > \pi_i^*$. Equations (15) and (16) imply the following:

\begin{align}
\frac{\partial U_i^{**}}{dV} - \frac{\partial U_i^*}{dV} &= \frac{2(\pi_i^{**} - \pi_i^*)K \partial \pi_1}{(\pi_1 + 1)^2} > 0, \tag{17} \\
\frac{\partial U_i^{**}}{dK} - \frac{\partial U_i^*}{dK} &= \frac{2(\pi_i^{**} - \pi_i^*)(K \partial \pi_1 - \pi_1 - 1)}{(\pi_1 + 1)^2}. \tag{18}
\end{align}

Unlike equation (17), equation (18) cannot be signed, which is reminiscent of the imperfect parallel between Theorem 2A and Theorem 3A. Yet equation (16) does imply the following useful comparison:

\begin{equation}
\text{sign} \left( \frac{\partial U_i^{**}}{dK} \right) = \text{sign} \left( \frac{\partial U_i^*}{dK} \right) \quad \text{and} \quad \left| \frac{\partial U_i^{**}}{dK} \right| > \left| \frac{\partial U_i^*}{dK} \right|. \tag{19}
\end{equation}

For expositional purposes, it is simplest to consider the case in which the derivatives in equations (15) and (16) are negative for both firms. That is, assume that any indirect effects coming through the change of endpoint are modest compared to the $-1$ direct deduction in utility coming from increased capital requirements. In this case, equations (17) and (19) can be restated as follows:

i) Increases in required internal capital lead to a greater decline in utility of low types.

ii) Increases in required external capital lead to a greater decline in utility of high types.

---

7Similar intuition holds even when one or more of these derivatives are positive, but the exposition is complicated by having to break the analysis into subcases.
Thus, increases in internal capital tend to concentrate incentives for innovation in the far right tail of quality. By contrast, increases in external capital tend to distribute innovation incentives more broadly across the entire distribution of firms.

D. Stimulating or Discouraging Innovation

The goal of the paper is to examine: i) how willing agents are to pursue their own innovations, as opposed to waiting for others innovate first, and ii) how the market punishes those who delay innovation. In the process, I identify situations in which the market for innovation breaks down completely because everyone wishes to follow. For example, this occurs when \( p \) is low, or when internal financing is heavily used. I now discuss such breakdowns with more generality. Suppose parameters are such that the following holds:

\[
\pi_3 < \pi_1 < \pi_2 < 1.
\]

The emptiness of \([\pi_2, \pi_1]\) implies that no one would be willing to enter a second wave. Any agent in \([\pi_3, \pi_1]\) must have \(\pi_i < \pi_2\) and so requires two heads before entering the market. Thus, a second wave cannot materialize.

In a sense, the breakdown just described is triggered by the fact that the interval \([\pi_1, 1]\) is too large in this case. In other words, there is too much early innovation, which shuts down the market for later innovation. To rectify this situation, \(\pi_1\) must be raised above \(\pi_2\). One way of doing so is to artificially raise \(\alpha_1\) and decrease \(\alpha_2\), in effect applying a tax on leaders and subsidizing the second wave. Alternatively, following the logic of Theorem 1, \(p\) could be decreased indirectly via policies that affect pioneers’ risk profile.

In fact, transfers of this sort may occur naturally. It has been argued that the supply of informed finance providers (e.g., venture capitalists) is sticky over time, owing to the specialized nature of the human capital of venture capitalists (Gompers and Lerner (2004), Ljungqvist and Richardson (2003)). That is, even if there is a shock to investment demand, the supply side cannot immediately adjust. Depending on which side is in temporary excess, this stickiness can lead to an environment of either “money chasing deals” or “deals chasing money.” When a large first-wave cohort wants funding, the financiers are in short supply and earn rents. This inflates \(\alpha_1\), which is precisely the tax required. A small second wave implies that the financiers are oversupplied, which deflates \(\alpha_2\), which is again precisely the right subsidy. Thus, sticky supply may serve as a governor ensuring that there is neither too much nor too little innovation in each period.

IV. Conclusions

This paper posits a connection between innovation and external finance when firms i) possess private information about their own abilities and ii) learn from the innovation of other firms about the prospects of the innovation.
Internal financing leads to a stalemate in which all agents enter a waiting game, hoping for others to innovate first. This incentive arises because information is a public good, and thus waiting allows others to (partially) resolve risk for their own projects. When using external financing, the terms of financing depend on the investor’s assessment of the issuer’s quality. As shown here, this financing game leads to an orderly and predictable process in which firms enter the market in order of their privately known quality (highest to lowest). This feature creates a rush to innovate early as high-quality agents signal their quality by issuing early and, in so doing, obtain better financing terms.

The fact that outside financing endogenously accelerates innovation is consistent with the work of Kortum and Lerner (2000), who point out that the ratio of venture capital to total industrial R&D is less than 3%, yet venture capital accounts for 15% of industrial innovations. Results such as these beg the question of causality: Are these firms innovative because they are small and nimble, with reliance on external financing being merely a side effect of their size? Instead, this paper suggests a causal channel.

As described in the Introduction, recent empirical research is consistent with such a causal connection (Brown et al. (2009), Brown and Peterson (2009)). Lerner, Sorensen, and Strömberg (2011) find no decrease in R&D expenditures following private equity leveraged buyouts, a time of great pressure for cost-cutting measures. Indeed, they find that the number of influential innovations (as measured by patent citations) significantly increases following the private equity investment. Also consistent with the theme of this paper, Atanassov, Nanda, and Seru (2007) find significant increases in innovative activity following arm’s-length financing, and also find that the innovations are more influential.

The model offers a perspective that may be useful in empirical IPO studies. Within each IPO wave, the model predicts that leaders are of a higher quality than followers. This ordering is consistent with Banerjee et al. (2016), who find that leaders in IPO waves obtain higher market-to-book ratios, despite being similar on other dimensions (e.g., sales, capital expenditures, profitability, and leverage). However, not all waves should be equal in this regard. For example, when the risk of the innovation is low, the number of leaders is (relatively) large compared to the number of followers, and so the model would predict that the market-to-book gap documented by Banerjee et al. should be larger, because in such cases followers are punished more.

In this model, waves are triggered by the actions of leaders rather than due to exogenous shocks. This feature suggests that it could be useful to examine the stock prices of publicly traded competitors. If the wave is simply a reaction to (already known) information, this should be already reflected in the competitors’ stock prices. In turn, the marginal impact of new issues should be to depress incumbents’ stock prices because of increased competition. If instead the new entrants generate positive information about the value of investment opportunities, proxied by the wave generation, the incumbents’ stock prices should rise. Thus, the model suggests a more refined look into the timing of information flows from new issuers to followers and public competitors.

The binary payoffs in the paper imply that, without loss of generality, all securities may be described as equity. The intuition would seem to extend in a
straightforward manner to address the question of how capital structure affects innovation. In particular, riskless debt functions effectively like internal financing. According to Theorem 3, then, debt would tend to decrease the amount of innovation, whereas equity-like securities (more generally, any security that is information-sensitive) would tend to increase innovation. The key property is information-sensitivity because this determines the extent to which the investor cares about the quality of the firm being financed (i.e., the sensitivity of security pricing to firm quality).

Appendix. Proofs

Proof of Lemma 1. I first prove that $\alpha_j > \cdots > \alpha_1$ for all waves that occur with positive probability. To do so, pick an arbitrary wave before the last one, say $i < j$. There is some agent $\pi^* \in (0, 1)$ who strictly prefers this wave. Consequently, the agent prefers wave $i$ to wave $i + 1$, which is the following condition:

$$ (1 - \alpha_i)Z_i\pi^* > \pi_i(1 - \alpha_{i+1})Z_{i+1}\pi^* + (1 - p_i)V. \tag{A-1} $$

Using the martingale condition $Z_i = p_iZ_{i+1}$, this is equivalent to

$$ (\alpha_{i+1} - \alpha_i)p_iZ_i\pi^* > (1 - p_i)V. \tag{A-2} $$

So $(\alpha_{i+1} - \alpha_i) > 0$ as claimed. Similar comparisons hold for all other pairs.

I next prove that agents cannot enter before someone of higher quality in equilibrium. This structure implies the interval form described in the lemma.

Again consider an agent $\pi^* \in (0, 1)$ who prefers some wave $i$. Argue by contradiction and suppose there is an agent $\pi_{alt}$ of higher quality who prefers a later wave, say, the $(i + 1)$st wave. By optimality, agent $\pi^*$ has

$$ (\alpha_{i+1} - \alpha_i)p_iZ_i\pi_{alt} > (1 - p_i)V, \tag{A-3} $$

whereas the assumed preference of agent $\pi_{alt}$ implies

$$ (\alpha_{i+1} - \alpha_i)p_iZ_i\pi^* < (1 - p_i)V. \tag{A-4} $$

Putting these together implies that

$$ (\alpha_{i+1} - \alpha_i)p_iZ_i\pi_{alt} < (1 - p_i)V \quad \text{and} \quad (\alpha_{i+1} - \alpha_i)p_iZ_i\pi^* < (\alpha_{i+1} - \alpha_i)p_iZ_i\pi_{alt}, \tag{A-5} $$

which contradicts $\pi_{alt} > \pi^*$. $\square$

Proof of Theorem 1. I derive the comparative static result $\partial \pi_1 / \partial p < 0$ here. $\partial \pi_2 / \partial p < 0$ is similar.

With some substitution, the system of equations (5)–(8) can be simplified to the following two-equation, two-unknown system.

$$ \pi_1 = \frac{V(1 - p)(\pi_1 + \pi_2)(\pi_1 + 1)}{2K[p - \pi_1(1 - p) - \pi_2]}, \tag{A-6} $$

$$ \pi_2 = \frac{V(\pi_1 + \pi_2)}{\pi_1 + \pi_2 - 2K}. \tag{A-7} $$

Rewrite these equations as

$$ F := \pi_12K[p - \pi_1(1 - p) - \pi_2] - (1 - p)(\pi_1 + \pi_2)(\pi_1 + 1)V = 0, $$

$$ G := \pi_2(\pi_1 + \pi_2 - 2K) - V(\pi_1 + \pi_2) = 0. $$
Totally differentiating this system with respect to $p$, one obtains

\[ \frac{\partial F}{\partial p} + \frac{\partial F}{\partial \pi_1} \frac{\partial \pi_1}{\partial p} + \frac{\partial F}{\partial \pi_2} \frac{\partial \pi_2}{\partial p} = 0, \]  
\[ \frac{\partial G}{\partial p} + \frac{\partial G}{\partial \pi_1} \frac{\partial \pi_1}{\partial p} + \frac{\partial G}{\partial \pi_2} \frac{\partial \pi_2}{\partial p} = 0. \]  

\[ (A-8) \]
\[ (A-9) \]

Now this is a system of two equations to be solved for two unknowns, $\partial \pi_1/\partial p$ and $\partial \pi_2/\partial p$. Solving the system, one obtains the following:

\[ \frac{\partial \pi_1}{\partial p} = \frac{\frac{\partial G}{\partial \pi_2} \frac{\partial F}{\partial \pi_1} - \frac{\partial G}{\partial \pi_1} \frac{\partial F}{\partial \pi_2}}{\frac{\partial \pi_1}{\partial \pi_2} \frac{\partial \pi_2}{\partial \pi_1}}. \]  
\[ (A-10) \]

Reporting only the signs of each of these partial derivatives,

\[ \text{sign} \left( \frac{\partial \pi_1}{\partial p} \right) = 0 \oplus - \oplus \\ \oplus - \oplus < 0. \]  
\[ (A-11) \]

**Proof of Theorem 2.** Part A is similar to Theorem 1. To derive Part B, note that

\[ \frac{\partial \pi_2}{\partial V} = \frac{\frac{\partial G}{\partial \pi_2} \frac{\partial F}{\partial \pi_1} - \frac{\partial G}{\partial \pi_1} \frac{\partial F}{\partial \pi_2}}{\frac{\partial \pi_1}{\partial \pi_2} \frac{\partial \pi_2}{\partial \pi_1}}. \]  
\[ (A-12) \]

Reporting the signs of each of these partial derivatives,

\[ \text{sign} \left( \frac{\partial \pi_2}{\partial V} \right) = \oplus \oplus - \oplus \\ \oplus - \oplus = 0, \]  
\[ (A-13) \]

which is ambiguous. However, one can show that the numerator in expression (A-13) is decreasing in $p$.

When $p = 0$, the numerator in expression (A-13) simplifies to

\[ (\pi_1 + \pi_2)(X + 2\pi_2K) > 0, \]  
\[ (A-14) \]

where $X = \pi_1(\pi_2 + 4K + V) + \pi_2(V + 1 + 2K) > 0$. Therefore, the fraction in expression (A-13) is negative.

When $p = 1$, the numerator in expression (A-13) simplifies to

\[ -(\pi_1 + \pi_2)(1 - \pi_2) < 0. \]  
\[ (A-15) \]

Therefore, the fraction is positive. Once can also solve for the cutoff value $p^*$, which equates the numerator of expression (A-13) to 0:

\[ p^* = \frac{X + 2\pi_2K}{X + 2K}. \]  
\[ (A-16) \]

The proof of Theorem 3 is similar.

**Proof of Theorem 4.** Solving equations (13) and (14) for $\alpha_1$ and $\alpha_2$, respectively, and plugging into equations (11) and (12) collapses the four-equation, four-unknown system of Section III.A to:

\[ F := p \left( 1 - \frac{2K}{M(\pi_1 + 1)p} \right) M\pi_1 - p \left( 1 - \frac{2K}{\pi_1 + \pi_2} \right) \pi_1 - (1 - p)V, \]  
\[ (A-17) \]
\[ G := \left( 1 - \frac{2K}{\pi_1 + \pi_2} \right) \pi_2 - V. \]  
\[ (A-18) \]
These equations imply the following partial derivatives:

\[
\begin{align*}
\frac{\partial F}{\partial p} &> 0, & \frac{\partial G}{\partial p} &= 0, & \frac{\partial F}{\partial M} > 0, & \frac{\partial G}{\partial M} = 0, & \frac{\partial F}{\partial \pi_1} > 0, & \frac{\partial G}{\partial \pi_1} > 0, \\
\frac{\partial F}{\partial \pi_2} > 0, & \frac{\partial G}{\partial \pi_2} = 0, & \frac{\partial F}{\partial V} > 0, & \frac{\partial G}{\partial V} = 0, & \frac{\partial F}{\partial K} > 0, & \frac{\partial G}{\partial K} > 0.
\end{align*}
\]

All of these signs agree with those of the base model, without a first-mover advantage. Therefore, all of the same comparative statics hold. Finally, consider that

\[
\frac{\partial \pi_1}{\partial M} = \frac{\partial G}{\partial M} \frac{\partial F}{\partial \pi_2} \frac{\partial F}{\partial \pi_1} - \frac{\partial G}{\partial \pi_2} \frac{\partial F}{\partial M} \frac{\partial G}{\partial \pi_1}.
\]

Reporting only the signs of each of these partial derivatives,

\[
\text{sign} \left( \frac{\partial \pi_1}{\partial M} \right) = \frac{0}{\oplus} - \frac{\oplus}{\oplus} < 0.
\]

The other cases are similar. □

References


