

The Timing of Capacity Investment by Start-ups and Established Firms in New Markets*

Robert Swinney

Graduate School of Business, Stanford University, Stanford, CA, 94305, swinney@stanford.edu

G erard P. Cachon, Serguei Netessine

The Wharton School, University of Pennsylvania, Philadelphia, PA, 19104

cachon@wharton.upenn.edu, netessin@wharton.upenn.edu

September, 2009

Abstract

We analyze the competitive capacity investment timing decisions of both established firms and start-ups entering new markets which are characterized by a high degree of demand uncertainty. Firms may invest in capacity early (when the market is highly uncertain) or late (when market uncertainty has been resolved), possibly at different costs. In our model, established firms choose investment timing and capacity level to maximize expected profits. Start-ups are prone to bankruptcy if profit turns out to be too low, and hence choose investment timing and capacity level to maximize the probability of survival. Surprisingly, we find that in monopoly situations, a start-up is more likely to prefer early investment than an established firm, despite the presence of demand uncertainty. In duopoly situations with one start-up and one established firm competing in the same market, we characterize the equilibria of a strategic capacity investment timing game in which firms choose when to build capacity. We find that when demand uncertainty is high, the unique equilibrium of this game is for the start-up to take a leadership role and invest first in capacity while the established firm follows; by contrast, when two established firms compete in an otherwise identical game, high demand uncertainty leads to both firms investing late. We conclude that the threat of firm failure significantly impacts the dynamics of competition involving start-ups, which helps explain multiple phenomena described in the disruptive innovation literature.

1 Introduction

Firms entering new markets face numerous operational challenges. Among the most crucial are issues related to capacity investment. Particularly when the size of a market is uncertain, two common yet difficult decisions are how much capacity to invest in, and when to do it. When choosing how much capacity to build or reserve with a supplier, the trade-off is clear: too much capacity results in underutilized facilities (if output is reduced to match market demand) or depressed prices

*A previous version of this paper was titled "Capacity Investment by Competitive Start-ups." The authors thank the Mack Center for Technological Innovation at the Wharton School for support of this project.

(if output remains high despite low demand), while too little capacity results in reduced sales and suboptimal profit and growth.

Timing the capacity investment decision presents even subtler considerations. Uncertainty surrounding market size typically reduces over time, meaning a firm that invests in capacity early is subject to a higher degree of demand uncertainty than a firm that postpones the investment decision. On the other hand, in competitive situations, a firm investing earlier than its rivals becomes the first-mover in the market, which may yield a strategic advantage. Indeed, the cost of capacity itself may change over time, either increasing (e.g., if contract capacity becomes scarce as the market matures) or decreasing (e.g., if learning enables lower-cost processes). These factors combine to make the decision of *when* to invest in capacity just as difficult and perilous, if not more so, as the decision of *how much* capacity to build or purchase.

The timing of capacity investment when entering new markets is precisely the issue that we consider. We first examine stylized monopoly models in which the sole entrant to a new market must build or source capacity in anticipation of future demand. Eventual market size is uncertain, and the firm is allowed to invest in one of two periods: if the firm invests *early*, then it makes the capacity decision before knowing market size, whereas if it invests *late*, all demand uncertainty is eliminated and capacity is built or sourced after learning market size. The cost of capacity is allowed to vary between periods. Thus, a monopolist firm must trade off the value of information (which is gained if the investment decision is delayed) with potential cost advantages from early investment.

Because new markets are often pursued by nascent firms, we focus on how the timing of capacity investment differs between *start-ups* and *established firms*. We consider the primary difference between these two types of firms to be the threat of bankruptcy or firm failure. Large established firms diversifying into new markets are unlikely to face imminent peril should demand in that market turn out to be low; start-ups, on the other hand, are typically smaller firms wholly invested in a single market, and thus, to a far greater extent than their established counterparts, face potentially disastrous consequences should the market fail to materialize as expected. The presence of this risk, combined with the high degree of demand uncertainty that typically accompanies the development of a new market, implies that start-ups should have a utility function which takes into account the risk and consequences of failure. Hence, in our model, the objective of a start-up is to time

the capacity investment decision to maximize the *probability of survival*. Established firms, by contrast, do not face an imminent risk of failure, and hence make capacity decisions to maximize *expected profit*.

In the monopoly setting, we examine how start-ups differ in their capacity timing decisions from established firms, characterizing how market uncertainty, capacity costs, and the threat of failure influence both capacity levels and investment timing. We find that established firms are likely to prefer late investment even if early investment is cheaper, because the flexibility to respond to market conditions engendered by late investment allows the firm to capture higher profits, particularly in high demand states. By contrast, start-ups prefer to invest in capacity whenever capacity is least expensive—that is, if capacity costs increase over time, start-ups prefer early investment—because lower capacity costs minimize the threshold market size that results in firm survival and hence maximize the probability of survival.

We then proceed to analyze duopoly models in which two firms simultaneously consider entry into a new market. In addition to all of the trade-offs inherent in the monopoly model, the competitive interaction introduces a strategic aspect to the capacity investment timing decision: a firm investing earlier than a competitor may gain a leadership position in a sequential game. We find that when a start-up competes with an established firm, if market uncertainty is high (as in a new market), then the unique equilibrium is for the start-up to invest early, while the established firm invests late. By contrast, when two established firms compete, the only equilibrium when demand uncertainty is high is simultaneous: both firms invest late. We thus conclude that the threat of failure experienced by a start-up tends to push capacity investment earlier—in both monopolistic and competitive situations—and leads to asymmetric investment timing equilibria in which start-up firms, remarkably, act as first-movers in new markets, despite the apparent advantages of established firms in terms of resources and technology (see, e.g., Bower and Christensen 1995).

In this regard, our findings relate to several streams of research, most notably the literature on disruptive innovation. The seminal works on this topic are Bower and Christensen (1995), Christensen and Bower (1996), and Christensen (1997); Schmidt and Druehl (2008) provide a recent review. A *disruptive innovation* is an improvement in a product or service that fundamentally changes its cost, performance, or target market in new or unexpected ways. Such innovations are typically enabled by scientific, technological, or process advancements; for example, the rise of inex-

pensive, physically compact desktop computers enabled the emergence of the personal computing market over the minicomputer and mainframe markets, and the development of cheap, tiny digital flash storage technologies helped contribute to the dominance of digital photography over film photography. A recurring question in this literature is: why do large, established firms typically fail to embrace disruptive innovations early, while smaller start-up firms often take a leadership role in bringing the innovations to market? Our model supports one possible answer to this question, namely, that it is the natural equilibrium of an endogenous timing game between a start-up and an established firm.

The remainder of this paper is organized as follows. §2 provides a brief review of the literature. §3 analyzes the monopoly model, while §§4–5 analyze the duopoly model. §6 presents several extensions to the basic model, and §7 concludes the paper.

2 Related Literature

There are three primary streams of research related to our work: the operations literature on capacity investment under uncertainty; the economics literature on competitive capacity investment; and the strategic management literature on new market entry and disruptive innovation. The latter topic was discussed in §1; here, we briefly review the remaining two broad areas, with further references to relevant works included throughout the remainder of the paper.

Our model is one of capacity investment with stochastic demand. As such, it is related to the extensive operations literature on this topic—see the comprehensive review by Van Mieghem (2003). Some works of particular relevance in this stream include Archibald et al. (2002), Babich et al. (2007), Babich (2008), Swinney and Netessine (2009), and Boyabatli and Toktay (2007), all of which consider the impact of bankruptcy risk on capacity or inventory decisions. Tanrisever et al. (2008) consider the related issue of simultaneous investment into capacity and process improvement in the presence of bankruptcy. While these papers address various consequences of bankruptcy on operational decisions (including process development, capacity levels, financial subsidies to suppliers, and contracting and sourcing strategies), no paper in the literature, to our knowledge, considers the impact of bankruptcy or firm failure on capacity investment *timing*. Indeed, there is a relative lack of research in the operations literature on the topic of capacity investment timing for entry

into new markets.

We analyze duopoly models consisting of two firms strategically investing in capacity before either begins to sell in the market. Similar models, frequently referred to as “endogenous leadership games” in the economics literature, have been studied by Gal-Or (1985), Saloner (1987), Hamilton and Slutsky (1990), Maggi (1996), and Bhaskaran and Ramachandran (2007). Maggi (1996) considers an endogenous leadership game with demand uncertainty, much like ours, although two key differences are that the differing objectives of start-ups (and hence the impact of bankruptcy) are not considered, and further capacity investment may occur in multiple periods (whereas in our model, capacity investment occurs in at most one period, due to, e.g., high fixed costs). Also related along these lines is the long stream of research on capacity investment for entry deterrence, pioneered by Spence (1977).

Lastly, there is an extensive literature on entry timing for reasons not related to strategic capacity investment. Some examples include social influence (Joshi et al. 2009), quality or cost improvements (Lilien and Yoon 1990), product technology (Bayus and Agarwal 2007), and product design (Klastorin and Tsai 2004). Our paper differs from these by focusing solely on the impact of bankruptcy risk on capacity investment timing under demand uncertainty, and exploring how such risk impacts timing in duopolistic settings.

3 Monopolistic Firms

In this section, we introduce and analyze two different monopoly models of capacity investment timing in a new market with uncertain demand: §3.1 discusses an established, profit-maximizing firm, while §3.2 considers a start-up prone to bankruptcy. The established firm model is a relatively standard formulation, and serves as a vehicle to introduce the dynamics of our setting and also as a baseline for comparison with the bankruptcy-prone start-up. We defer all discussion of competition until §4.

3.1 A Monopolistic Established Firm

An established firm (denoted by the subscript e) sells a single product.¹ The quantity of the product released to the market is Q_e . The market price is given by the linear demand curve $p(Q_e) = A - Q_e$. Prior to determining the production quantity, the firm must invest in production capacity K_e which determines its maximum output. This capacity may be internal to the firm (e.g., if the firm in question is a manufacturer) or external (e.g., if the firm outsources production to a contract manufacturer). There is no constraint on the total amount of capacity that can be built or reserved in either case.

Capacity investment may occur at one of two times: either *early* or *late*. Early investment is sufficiently far in advance of the selling season that the total market size is uncertain. The uncertainty in market size is reflected in the demand intercept, A , which is modeled as a continuous random variable with distribution function F , mean μ , and variance σ^2 . Late investment, on the other hand, is sufficiently close to the start of the selling season that all uncertainty in A is eliminated—hence, capacity investment is made after observing the realized value of A . In either case, the production quantity (Q_e) is determined after A has been observed and K_e has been fixed (i.e., just before the selling season), and hence output is subject to the constraint $Q_e \leq K_e$. We will frequently refer to the early period as “period 1” and the late period as “period 2.”

We assume that capacity investment, whenever it is made, is irreversible. Furthermore, capacity investment can occur in at most one period.² The total capacity cost is linear in the amount of capacity reserved, and the marginal cost of capacity may vary over time. The unit cost in period 1 is denoted c_1 and the unit cost in period 2 is denoted c_2 . We make no *ex-ante* assumption on the ordering of c_1 and c_2 . Costs that decrease over time (i.e., $c_1 > c_2$) may be reflective of exogenous technological or process cost improvements, innovation, or raw materials cost decreases; similarly, costs that increase over time ($c_1 < c_2$) could occur if contract manufacturers offer a discount for

¹We implicitly assume that the established firm—diversifying into the new market—has already evaluated the impact (if any) that market entry will have on sales of its existing products, and determined that entry is profitable; Druehl and Schmidt (2008) analyze this related problem of how new market entry can encroach on sales of existing (substitutable) products.

²In reality, firms may be able to invest in capacity in multiple periods. Allowing such an option clearly does not impact the evaluation of deferred (late) investment, though it may increase the value of early investment. If fixed costs of capacity installation or expansion are high, then the value of an option to invest in both periods is relatively low—in the extreme case, if fixed costs are high enough, then firms will only invest in capacity in one period. This is the case that we consider.

early investment, if capacity in the later period is scarce, or if second period capacity must be installed more quickly, incurring expedited construction or configuration costs. The reasons behind inter-temporal cost variation are outside the scope of this paper; rather, we will present results that hold conditional on a particular cost trend.

The marginal production cost is zero, and for analytical tractability, we assume that the firm adheres to a production clearance strategy: that is, the firm always produces up to its capacity and releases the maximum quantity to the market, $Q_e = K_e$.³ The established firm, being a large, diversified company, faces minimal risk of bankruptcy as a result of entry into this new market—hence, facing uncertainty in market size (A), the established firm seeks to maximize *expected profit*, which is denoted $\mathbb{E}(\pi_e(K_e))$, where the absence of the expectation operator, $\pi_e(K_e)$, is used to denote profit for a particular realization of A . Throughout the analysis, optimal values (capacities, profits, etc.) are denoted by the superscript $*$.

Given this formulation, the firm’s optimal expected profit from early capacity investment is

$$\mathbb{E}(\pi_e^*) = \max_{K_e \geq 0} \mathbb{E}((A - K_e - c_1) K_e), \quad (1)$$

while the firm’s optimal expected profit from late capacity investment is

$$\mathbb{E}(\pi_e^*) = \mathbb{E}\left(\max_{K_e \geq 0} ((A - K_e - c_2) K_e)\right). \quad (2)$$

Thus, when the firm is deciding whether or not to invest in capacity in period 1, it must compare (1) with (2). The following theorem provides the details of the optimal capacity timing and investment level.

Theorem 1 *A monopolist established firm prefers early investment if and only if*

$$\sigma^2 < (\mu - c_1)^2 - (\mu - c_2)^2, \quad (3)$$

³The issue of holdback, i.e., producing a quantity less than the total capacity, is discussed in §6. To summarize that discussion: the holdback and clearance models only differ for the established firm when it invests early (because late investment is made after uncertainty is resolved, hence the firm always invests in precisely enough capacity to satisfy its optimal production quantity). For the start-up firm, the holdback and clearance models do not differ at all. We thus find that allowing the firms to engage in holdback does not dramatically impact our results, a fact that is demonstrated in §6 with a numerical study.

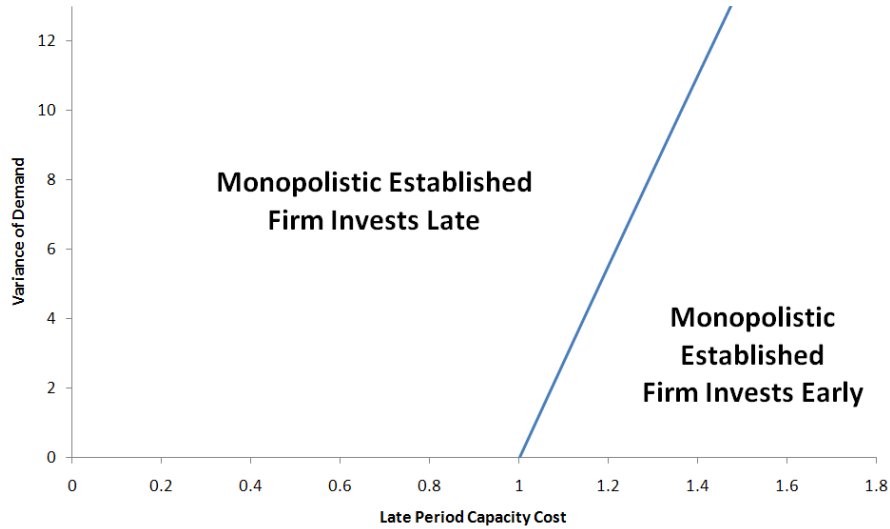


Figure 1. Optimal investment timing for a monopolistic established firm. In this example, $c_1 = 1$ and $\mu = 15$.

yielding optimal capacity $K_e^* = (\mu - c_1)/2$ and expected profit $\mathbb{E}(\pi_e^*) = (\mu - c_1)^2/4$. Otherwise, the firm prefers late investment, yielding optimal capacity $K_e^* = (A - c_2)/2$ and expected profit $\mathbb{E}(\pi_e^*) = (\mu - c_2)^2/4 + \sigma^2/4$.

Proof. *Early Investment:* the profit function implied by (1) is concave and yields a unique maximum at the Cournot monopoly point, $K_e^* = (\mu - c_1)/2$. Expected profit is thus $\mathbb{E}(\pi_e^*) = (\mu - c_1)^2/4$.

Late Investment: the profit function implied by (2) is concave and yields a unique maximum at the Cournot monopoly point, $K_e^* = (A - c_2)/2$. Expected profit is thus $\mathbb{E}(\pi_e^*) = \mathbb{E}\left(\frac{(A - c_2)^2}{4}\right) = (\mu - c_2)^2/4 + \sigma^2/4$. ■

As Theorem 1 demonstrates, an established monopolist prefers early investment if and only if equation (3) holds, i.e., if demand uncertainty is low and early investment is cheaper than late investment ($c_1 < c_2$). Note that if capacity costs decrease over time ($c_1 > c_2$), the firm prefers late investment for any feasible variance (i.e., for any $\sigma^2 \geq 0$). If capacity costs increase over time ($c_1 < c_2$), the firm may prefer early or late investment, depending on the level of demand uncertainty. Figure 1 depicts the optimal investment timing as a function of demand variance and late period capacity costs.

3.2 A Monopolistic Start-up

A common feature of new markets, particularly those enabled by ground-breaking or unforeseen technological innovation, is that they are characterized, *ex ante*, by a large amount of demand uncertainty. Thus, far more so than their established counterparts, smaller start-up firms are exposed to a serious risk: the risk of bankruptcy or firm failure, should demand turn out to be low. Consequently, while it is quite natural to assume that established firms make decisions to maximize expected profits, it is less clear that start-ups should or do behave in the same way: as Radner and Shepp (1996) and Dutta and Radner (1999) demonstrate, a firm prone to bankruptcy that purely maximizes expected profit over an infinite horizon will fail with probability one. The objective of a start-up should, then, take into account the acute risk of failure associated with entry into a new market. This implies that firms particularly prone to bankruptcy—for our purposes, start-ups entering new markets—in fact have a utility function that depends both on *operating profit* and the *risk of failure*, e.g.,

$$\text{Total Utility} = \text{Operating Profit} - \text{Cost of Bankruptcy} \times \text{Probability of Bankruptcy}, \quad (4)$$

where the “cost of bankruptcy” represents either real costs (e.g., default penalties on loans), or a virtual penalty term embodying the expected consequences of bankruptcy.⁴ This type of utility function can be found, for example, in the seminal paper by Greenwald and Stiglitz (1990) and in Walls and Dyer (1996).⁵ If the cost of bankruptcy is small compared to the assets of the firm or if the probability of default is very low, then the firm may safely ignore the last term and simply maximize expected operating profits; this would be the case with large, established firms considering diversifying entry into a new market that represents a small potential fraction of their total business. Our model in the preceding section addressed precisely this scenario.

Alternatively, if the cost of bankruptcy is large compared to the assets of the firm and would result in financial ruin, or if the probability of bankruptcy is high (either of which is likely to

⁴From the accounting and financial points of view, the meaning of the word “bankruptcy” is often complex and does not necessarily imply that the company fails; the actual event of bankruptcy can have varying degrees of consequence to a firm, ranging from reorganization (Chapter 11 bankruptcy) to total liquidation (Chapter 7 bankruptcy). When using this term we simply imply that the company becomes insolvent and ceases to exist due to the negative cash-flow.

⁵As the latter paper points out, “in a homogeneous industry where investment projects have equivalent properties ... competition for scarce capital requires the firm to obtain an appropriate trade-off between maximization of expected profits and the probability of bankruptcy (or financial distress)...” (Walls and Dyer 1996).

be the case for a start-up), the second term dominates the expression; the maximization problem may then be thought of as approximately equal to minimizing the probability of bankruptcy or, equivalently, maximizing the probability of survival. As a result, in what follows, we assume that the presence of failure risk implies that start-ups have a different objective than established firms: instead of maximizing expected profits, they maximize their chance of survival.⁶ Essentially, while any firm has a true profit function that accounts for both operating profits and the chance of bankruptcy as depicted in (4), we examine extreme cases: established firms are entirely concerned with operating profits while start-ups are entirely concerned with the probability of bankruptcy. This dichotomization of the objective function, while stylized, allows us to obtain sharp results.⁷

Consequently, the details of the model are identical to those introduced in §3.1, except for the objective function of the firm. We use the subscript s to denote a start-up firm. The start-up seeks to time its capacity investment and set the precise capacity level in order to maximize the *probability of survival*, denoted $\psi_s(K_s)$. We assume that survival occurs for the start-up if its total profit is greater than α , an exogenous parameter which may represent, for example, overhead costs, existing debt payments, or R&D costs. Consequently, the optimal survival probability from early investment is

$$\psi_s^* = \max_{K_s \geq 0} \Pr((A - K_s - c_1) K_s \geq \alpha), \quad (5)$$

while the optimal survival probability from late investment is

$$\psi_s^* = \Pr\left(\max_{K_s \geq 0} ((A - K_s - c_2) K_s) \geq \alpha\right). \quad (6)$$

Note that, in equation (6), we have assumed that a start-up investing late, no longer subject to any uncertainty in demand, chooses a capacity level to maximize its expected profit. This will, in

⁶As Chod and Lyandres (2008) discuss, the owners of private firms (e.g., start-ups in our model) are typically less diversified than the owners of public firms (established firms in our model), and hence are more sensitive to the risk inherent in a single venture and the corresponding chance of failure. Thus, it's reasonable that start-ups and established firms have different objectives—see Chod and Lyandres (2008) and references therein for a detailed discussion of this matter.

⁷Analysis of our model with the integrated objective function depicted in Equation (4) has proven intractable. Furthermore, an alternative interpretation is the start-up maximizes expected profit subject to a constraint that the probability of survival is at least some minimum value ρ , similar to a VAR constraint in portfolio optimization. If $\rho = 0$, then the firm maximizes expected profit, as in the established firm case. If ρ is set to the maximum feasible threshold, then the start-up maximizes the probability of survival. Hence, our two objective functions (expected profit for the established firm, and survival probability for the start-up) can also be considered to be two extreme versions of the VAR family of objective functions.

turn, result in the largest possible *ex-ante* survival probability. The following theorem describes the optimal investment timing and capacity decisions, given equations (5) and (6).

Theorem 2 *A monopolist start-up prefers early investment if and only if $c_1 < c_2$, yielding optimal capacity $K_s^* = \sqrt{\alpha}$ and survival probability $\psi_s^* = 1 - F(2\sqrt{\alpha} + c_1)$. Otherwise, the firm prefers late investment, yielding optimal capacity $K_s^* = (A - c_2)/2$ and survival probability $\psi_s^* = 1 - F(2\sqrt{\alpha} + c_2)$.*

Proof. *Early Investment:* maximizing the survival probability function in (5) is equivalent to

$$\psi_s^* = \max_{K_s \geq 0} \Pr \left(A \geq \frac{\alpha}{K_s} + K_s + c_1 \right) = \max_{K_s \geq 0} \left(1 - F \left(\frac{\alpha}{K_s} + K_s + c_1 \right) \right),$$

and, consequently, this is equivalent to minimizing $\frac{\alpha}{K_s} + K_s + c_1$. This expression is convex and yields a unique minimizing capacity of $K_s^* = \sqrt{\alpha}$. The corresponding optimal survival probability is thus $\psi_s^* = 1 - F(2\sqrt{\alpha} + c_1)$.

Late Investment: under late investment, the start-up maximizes profit after observing A . This implies the late investment capacity level is identical to the established firm's capacity level until late investment, i.e., $K_s^* = (A - c_2)/2$. The survival probability is thus

$$\psi_s^* = \Pr \left(\frac{(A - c_2)^2}{4} \geq \alpha \right) = 1 - F(2\sqrt{\alpha} + c_2),$$

yielding the result. ■

Theorem 2 demonstrates that a start-up prefers early investment only if costs increase over time ($c_1 < c_2$). If costs decrease over time, the start-up prefers late investment. While the latter result is identical to the established firm case, the former is not; Theorem 1 shows that the established firm can prefer late investment even if costs increase over time, so long as demand uncertainty is large enough. Thus, we conclude from Theorems 1 and 2 that, given any particular set of problem parameters, a monopolistic start-up is more likely to prefer early investment than an established firm.

It is somewhat counter-intuitive that a start-up, prone to such serious consequences should failure occur, is more willing to invest in capacity early than an established firm (given that the two firms have equal capacity costs); moreover, the start-up's decision is curiously unaffected by

the degree of demand uncertainty. The reason for the latter result is that the start-up maximizes the probability of survival by maximizing the range of demand outcomes in which it survives. To accomplish this, it chooses the capacity that leads to survival at the lowest possible demand threshold—with this capacity level, the firm will survive for all higher demand realizations. This threshold demand level is independent of the demand variance, hence variance does not impact the start-up’s survival-maximizing capacity decision.

In addition, because the start-up chooses the capacity level which ensures survival over the largest range of demand outcomes, the ability to respond to demand via late investment is not valuable to the start-up; late investment does not change the minimum demand level which ensures survival, and hence does not increase the start-up’s survival probability. What *does* impact survival probability is capacity cost: lower capacity costs lead to a lower survival threshold and hence a greater survival probability. Consequently, as Theorem 2 shows, when capacity costs change over time, the survival probability will be greater in the lower cost period, which leads to the result that the start-up prefers to invest in the period with the lowest cost.

4 Duopoly Model

We now move to the duopoly model. The details of the model are identical to the monopoly model addressed in the previous section, except there are now *two firms* competing with perfectly substitutable⁸ products in the new market. One firm is a start-up (denoted s) and maximizes the probability of survival, while the other is an established firm (denoted e) that maximizes expected profit. The quantity of the product released to the market by firm i is Q_i , $i \in \{s, e\}$. The market price of the product is given by the linear demand curve $p(Q_i, Q_j) = A - Q_i - Q_j$. As before, A is a random variable with distribution function F , mean μ , and variance σ^2 . Firms have identical capacity costs, which, as in the monopoly case, may vary over time (heterogeneous costs are discussed in §6).

Before the early period (e.g., during an even earlier “decision period”), the firms simultaneously make their capacity timing decision. Each firm has two possible actions: either commit to invest

⁸The model can easily be altered to account for partial substitution, with results that are qualitatively similar to those that follow. The resulting expressions (for capacity, profit, and survival probability) are much more complicated, however. If the substitution between products is sufficiently low, the results resemble those from the monopoly model, whereas if substitution is sufficiently high, the results are similar to the model with perfect substitution.

	Established Firm Early	Established Firm Late
Start-up Early	(E, E)	(E, L)
Start-up Late	(L, E)	(L, L)

Table 1. The four possible sequences of moves and their abbreviations.

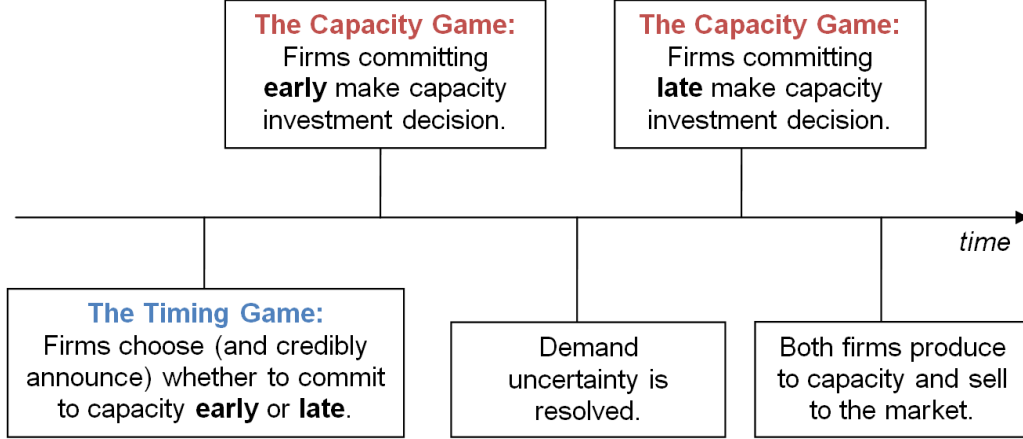


Figure 2. The sequence of events in the duopoly model.

in the early period, or commit to delay until the late period. We assume that these actions are credible and irreversible. This initial game is referred to as the investment timing game, or merely the *timing game*. There are four possible pure-strategy outcomes to the timing game: both firms invest early, both firms defer until the late period, and the two asymmetric outcomes in which one firm invests early and one firm invests late. The timing game, and the abbreviations used to refer to its outcomes, are depicted in Table 1.

The *capacity subgame* then unfolds according to the sequence of moves determined by the timing game. In the late period, we assume all actions from the early period are publicly observable (e.g., if the established firm invests in the early period and the start-up defers, the start-up observes the precise capacity level of the established firm at the beginning of period 2 before choosing its own capacity level). Thus, in addition to the informational and cost considerations from the monopoly model, there are strategic factors in play with the timing of capacity investment: if one firm moves early and the other moves late, the early-moving firm enjoys a leadership position in a sequential game while the late-moving firm is a sequential follower. As before, we assume that capacity investment is irreversible, and firms may invest in capacity in at most one period. The sequence of events is depicted in Figure 2.

In the following four lemmas, we analyze the equilibria to each of the four capacity subgames depicted in Table 1. Once we have derived these equilibria, we may in turn analyze the equilibrium to the investment timing game.

We first consider the case in which both firms invest in capacity late, i.e., after observing A . Because there is no randomness, as in the monopoly model, the start-up will choose capacity to maximize profit. The following lemma describes the equilibrium capacity investments for each firm in this game, in addition to providing the *ex-ante* survival probability of the start-up (ψ_s^*) and the *ex-ante* expected profit of the established firm ($\mathbb{E}(\pi_e^*)$).

Lemma 1 *If both firms invest in capacity late, then equilibrium capacities are $K_e^* = K_s^* = \frac{A-c_2}{3}$. The ex-ante equilibrium expected profit of the established firm is*

$$\mathbb{E}(\pi_e^*) = \frac{\sigma^2 + (\mu - c_2)^2}{9}, \quad (7)$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F(3\sqrt{\alpha} + c_2). \quad (8)$$

Proof. The proofs of all lemmas appear in the appendix. ■

We now move to the game in which the established firm invests late while the start-up invests in capacity early, i.e., prior to observing A . The following lemma describes the equilibrium capacity levels, expected profit, and survival probability.

Lemma 2 *If the start-up invests early while the established firm invests late, equilibrium capacities are $K_e^* = \frac{A-\sqrt{2\alpha}-c_2}{2}$ and $K_s^* = \sqrt{2\alpha}$. The ex-ante equilibrium expected profit of the established firm is*

$$\mathbb{E}(\pi_e^*) = \frac{\sigma^2 + (\mu - c_2 - \sqrt{2\alpha})^2}{4}, \quad (9)$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F(2\sqrt{2\alpha} + 2c_1 - c_2). \quad (10)$$

We next consider the case in which both firms invest in capacity early, i.e., before observing the value of A . Lemma 3 describes the equilibrium.

Lemma 3 *If both firms invest early, equilibrium capacities are $K_e^* = \frac{\mu - \sqrt{\alpha} - c_1}{2}$ and $K_s^* = \sqrt{\alpha}$. The ex-ante equilibrium expected profit of the established firm is*

$$\mathbb{E}(\pi_e^*) = \frac{1}{4} (\mu - c_1 - \sqrt{\alpha})^2, \quad (11)$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F\left(2\sqrt{\alpha} + \frac{\mu - \sqrt{\alpha} + c_1}{2}\right). \quad (12)$$

Lastly, we address the case in which the start-up invests in capacity late while the established firm invests in capacity early. The following lemma describes the resulting equilibrium.

Lemma 4 *If the established firm invests early while the start-up invests late, equilibrium capacities are $K_e^* = \frac{\mu - 2c_1 + c_2}{2}$ and $K_s^* = \frac{2A - \mu + 2c_1 - 3c_2}{4}$. The ex-ante equilibrium expected profit of the established firm is*

$$\mathbb{E}(\pi_e^*) = \frac{(\mu + c_2 - 2c_1)^2}{8}, \quad (13)$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F\left(\sqrt{4\alpha + (c_1 - c_2)^2} + \frac{\mu + c_2}{2}\right). \quad (14)$$

5 Equilibrium to the Timing Game

Having derived equilibria to each of the capacity investment subgames, we may now derive the equilibrium to the investment timing game. The following theorem describes all of the possible equilibria to this game:

Theorem 3 *Let $\Delta c \equiv c_1 - c_2$, let*

$$\Sigma_1 \equiv (\mu - c_1 - \sqrt{\alpha})^2 - (\mu - c_2 - \sqrt{2\alpha})^2 \quad (15)$$

and let

$$\Sigma_2 \equiv \frac{9}{8}(\mu + c_2 - 2c_1)^2 - (\mu - c_2)^2. \quad (16)$$

Then the following pure strategy equilibria to the investment timing game exist:

1. If $\sigma^2 < \Sigma_1$, then both firms invest early.
2. If $\sigma^2 > \Sigma_1$ and $\Delta c < \frac{3-2\sqrt{2}}{2}\sqrt{\alpha}$, then the start-up invests early while the established firm invests late.
3. If $\Delta c > \frac{3-2\sqrt{2}}{2}\sqrt{\alpha}$ and $\sigma^2 > \Sigma_2$, then both firms invest late.

Proof. We will examine the viability of each subgame in Table 1 individually.

(i) (E, L) . First, let us consider the equilibrium in which the start-up invests early and the established firm follows: (E, L) . This is an equilibrium if no firm has incentive to unilaterally deviate: in other words, if the established firm enjoys greater expected profit than in (E, E) , and if the start-up enjoys a greater survival probability than in (L, L) . From Lemmas 1 and 2, comparing the arguments of the distribution function F in each of the equilibrium survival probabilities, we see that if the established firm invests late, the start-up enjoys a (strictly) greater survival probability by investing early if:

$$2\sqrt{2\alpha} + 2c_1 - c_2 < 3\sqrt{\alpha} + c_2.$$

Rearranging this expression, we see it reduces to

$$2\sqrt{2\alpha} < 3\sqrt{\alpha} + 2(c_2 - c_1).$$

If $c_1 < c_2$, the condition holds if $\alpha > 0$. If, on the other hand, $c_1 > c_2$, the start-up *may* unilaterally deviate from (E, L) for some $\alpha > 0$. Examining this expression, we see that the inequality is most likely to hold if α is large—hence, the start-up will deviate from (E, L) if costs decrease over time, and α is sufficiently small.

Next, consider the established firm, which, from Lemmas 2 and 3, will not deviate from (E, L) if

$$\frac{\sigma^2 + (\mu - c_2 - \sqrt{2\alpha})^2}{4} > \frac{(\mu - c_1 - \sqrt{\alpha})^2}{4}.$$

This expression reduces to

$$\sigma^2 > (\mu - c_1 - \sqrt{\alpha})^2 - (\mu - c_2 - \sqrt{2\alpha})^2. \quad (17)$$

In other words, the established firm will not unilaterally deviate from (E, L) if demand is variable enough, where the threshold variability is a function of the problem parameters. This demonstrates case (1) in the theorem.

(ii) (E, E) . We next consider the equilibrium in which both firms build capacity early. From Lemmas 2 and 3, the established firm will not deviate from this equilibrium precisely if (17) is violated. From Lemmas 3 and 4, the start-up will not deviate if

$$\sqrt{4\alpha + (c_1 - c_2)^2} + \frac{\mu + c_2}{2} > 2\sqrt{\alpha} + \frac{\mu - \sqrt{\alpha} + c_1}{2}.$$

This inequality reduces to

$$7\alpha + 3(c_1 - c_2)^2 > 6\sqrt{\alpha}(c_1 - c_2).$$

This inequality always holds for any parameter values, hence the start-up never deviates from (E, E) , and this sequence of moves is an equilibrium if and only if (17) is violated. This demonstrates case (2) in the theorem.

(iii) (L, E) . Because the start-up never has incentive to deviate from (E, E) , it must be the case that (L, E) is not supportable as an equilibrium.

(iv) (L, L) . We lastly consider the equilibrium with both firms building capacity late. In this case, part (i) of the proof demonstrated that the start-up prefers (L, L) to (E, L) if

$$2\sqrt{2\alpha} > 3\sqrt{\alpha} + 2(c_2 - c_1).$$

Similarly, from Lemmas 1 and 4, the established firm will not deviate from this equilibrium if

$$\frac{(\mu + c_2 - 2c_1)^2}{8} < \frac{\sigma^2 + (\mu - c_2)^2}{9}$$

This inequality reduces to

$$\frac{9}{8}(\mu + c_2 - 2c_1)^2 - (\mu - c_2)^2 < \sigma^2.$$

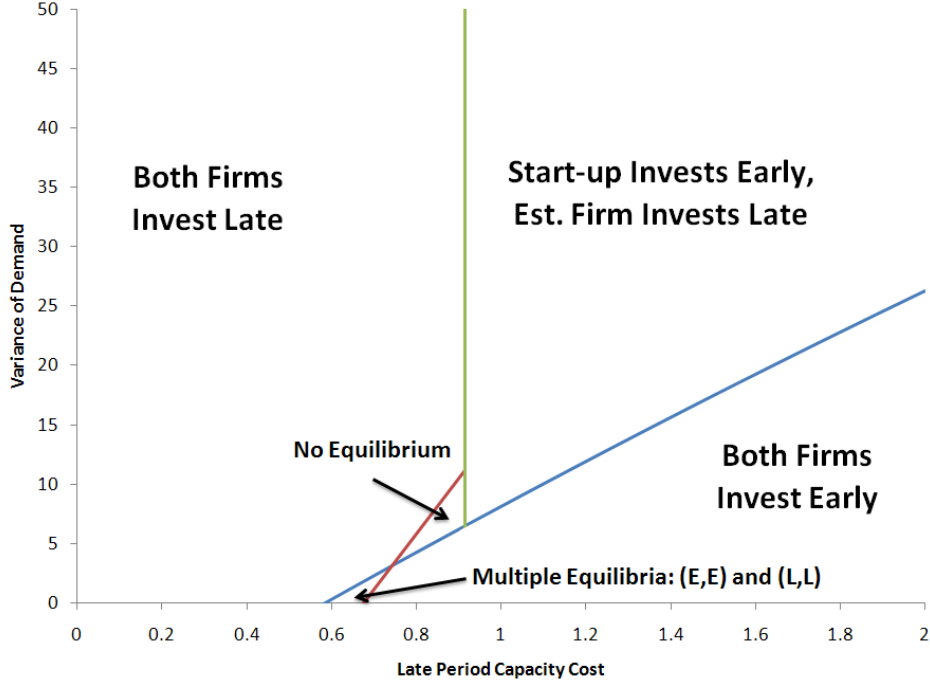


Figure 3. Equilibria to the investment timing game when all five regions are non-empty, as a function of demand variance and second period capacity costs. In this example, $c_1 = 1$, $\alpha = 1$, and $\mu = 13$.

This proves case (3) of the theorem. ■

There are several interesting consequences of these results. First, we note that the equilibrium regions are not exhaustive in covering the parameter space, nor are they mutually exclusive. As a result, regions of no (pure strategy) equilibria can occur, as can regions of multiple equilibria (in particular, regions in which late investment by both firms and early investment by both firms are both possible equilibria). In all, there are five potential equilibrium regions to the investment timing game: one region each for (L, L) , (E, L) , and (E, E) ; one region in which (E, E) and (L, L) are both possible; and one region in which no equilibrium exists. Figure 3 depicts the possible equilibria graphically, as a function of second period capacity costs and demand variance, in an example in which all five equilibrium regions are possible. It may also be the case that the regions of non-existence and multiple equilibria are empty, depending on the other parameter values (e.g., on α , the bankruptcy threshold, and μ , the expected market size).

The key result in Theorem 3 is that, if demand uncertainty is sufficiently high and costs do not decrease too much over time, the unique equilibrium to the investment timing game is for the start-up to invest early and the established firm to invest late. This equilibrium precisely

describes the situation discussed by Bower and Christensen (1995): a new market enabled by disruptive technology with highly uncertain demand, in which a start-up plays the role of leader and the established firm the role of follower. This occurs because of three competing forces in the model. The first is that early investment is valuable due to first-mover advantage in a sequential capacity game (if the competitor invests late). The second is that late investment is valuable due to the ability to exploit demand variance. The third is that the cheaper investment period is valuable due to cost savings, which can impact the value of either period. As we have already seen in the monopoly model, the second reason does not impact a start-up; hence, if costs do not decline severely over time, the start-up prefers early investment due to the leadership position in the capacity game. (Note that, unlike the monopoly model, a start-up facing competition from an established firm may invest early in capacity even if late investment is cheaper.)

By contrast, the established firm *does* value late investment due to the ability to exploit demand variance; hence, if variance is sufficiently high, the established firm prefers late investment even though it cedes a leadership position to the start-up. In particular, the start-up continues to choose the minimum capacity level that ensures survival over the widest range of demand outcomes, and hence does not exploit its leadership position to greatly increase capacity as a profit-maximizing firm might; consequently, it would appear that the established firm does not surrender as much by following a start-up as it might by following another established firm, a hypothesis that we verify in §6 by analyzing a model of two competing established firms.

As the theorem and figures also show: for *no* parameter combination is it a pure strategy equilibrium for the established firm to lead and the start-up to follow, even if period 2 costs are lower than period 1 costs. Indeed, the established firm only invests in capacity early in case (2) of Theorem 3, which, as the figure shows, corresponds to low demand uncertainty and costly second period production. This result is also in agreement with empirical and anecdotal evidence described by Bower and Christensen (1995) and others, namely that established firms rarely invest early in new markets when competing with start-up firms.

We also observe that when costs decrease significantly over time, the picture can become complicated. In particular, a unique equilibrium may exist (either both early or both late), multiple equilibria may exist, or a pure strategy equilibrium may fail to exist. In the region of non-existence, the start-up prefers to invest at the same time as the established firm (i.e., the start-up would like

to exploit cost reduction and information but only if it does not mean giving up a leadership position), while the established firm prefers to invest at the opposite time of the start-up. As a result, the outcome of the game is unclear in this region (although, it should be noted, the region of non-existence typically covers a very small portion of the parameter space).

6 Extensions

6.1 Competition with Two Established Firms

In this section, we analyze an investment timing game identical to the one discussed in §5, with one key difference: rather than competition between a start-up and an established firm, *both* firms are established, profit maximizing firms. We assume, as before, that the firms are *ex-ante* identical in all other respects. This allows us to compare the outcomes of the timing game with heterogeneous firms to an otherwise identical game with two mature firms, thus isolating the impact of bankruptcy-averse behavior on capacity investment timing. The following theorem presents the equilibrium to the timing game in this case.

Theorem 4 *If two established firms compete in an investment timing game, then there exists some threshold σ^* such that, for all $\sigma > \sigma^*$, the unique equilibrium of the investment timing game is for both firms to invest late.*

Proof. We must first analyze several additional aspects of the capacity subgames in order to analyze the investment timing game. First, consider the game in which both firms invest early. This is a Cournot duopoly, hence the equilibrium profits of the (symmetric) established firms are both

$$\mathbb{E}(\pi_e^*) = \frac{(\mu - c_1)^2}{9}.$$

Next, consider the game in which the firms invest sequentially. This is identical to the previously analyzed game in which the established firm invests early and the start-up invests late (because, in that case, the start-up maximized profit due to the elimination of uncertainty). Hence, the profit of the leader is

$$\mathbb{E}(\pi_e^*) = \frac{(\mu + c_2 - 2c_1)^2}{8}$$

while the profit of the follower is

$$\mathbb{E}(\pi_e^*) = \frac{4\sigma^2 + (\mu + 2c_1 - 3c_2)^2}{16}.$$

Finally, the game in which both firms invest late yields identical profits to both firms equal to

$$\mathbb{E}(\pi_e^*) = \frac{\sigma^2 + (\mu - c_2)^2}{9}.$$

Thus, the investment timing game in normal form has payoffs

	Firm 2 Early	Firm 2 Late
Firm 1 Early	$\left(\frac{(\mu - c_1)^2}{9}, \frac{(\mu - c_1)^2}{9}\right)$	$\left(\frac{(\mu + c_2 - 2c_1)^2}{8}, \frac{4\sigma^2 + (\mu + 2c_1 - 3c_2)^2}{16}\right)$
Firm 1 Late	$\left(\frac{4\sigma^2 + (\mu + 2c_1 - 3c_2)^2}{16}, \frac{(\mu + c_2 - 2c_1)^2}{8}\right)$	$\left(\frac{\sigma^2 + (\mu - c_2)^2}{9}, \frac{\sigma^2 + (\mu - c_2)^2}{9}\right)$

First, assume that firm 2 invests early. Firm 1 prefers *late* investment if

$$\frac{4\sigma^2 + (\mu + 2c_1 - 3c_2)^2}{16} > \frac{(\mu - c_1)^2}{9}. \quad (18)$$

Clearly, as σ^2 increases, this inequality is more likely to hold. Similarly, if firm 2 invests late, firm 1 prefers late investment if

$$\frac{\sigma^2 + (\mu - c_2)^2}{9} > \frac{(\mu + c_2 - 2c_1)^2}{8}. \quad (19)$$

Again, as σ^2 increases, this inequality is more likely to hold, thus for large enough σ^2 (i.e., σ^2 above some threshold), late investment is the dominant strategy of both firms and (L, L) is the only possible equilibrium, proving the theorem. ■

As the preceding theorem demonstrates, a high degree of demand uncertainty leads to a unique equilibrium outcome when established, profit-maximizing firms compete: both firms invest in capacity late. This is in stark contrast to the investment timing equilibrium when a start-up competes with an established firm: in that case, we observed that high demand uncertainty can lead to equilibrium outcomes in which the start-up acts as a sequential leader in the investment game. We note that, in the game with two established firms, asymmetric outcomes *can* occur for lower demand variability; however, they can never occur if demand variability is sufficiently large, unlike the

model with one start-up and one established firm. Hence, we conclude that a start-up's propensity to avoid bankruptcy can have a significant effect on the dynamics of competition, particularly when demand uncertainty is high in the context of new markets.

6.2 Firms with Heterogeneous Capacity Costs

In this extension, we return to the base model (one start-up and one established firm) and consider the impact of heterogeneous capacity costs. For the sake of simplicity, we will assume that costs are constant over time for both firms, since we have already explored the impact of time-varying costs. Let the cost of the established firm be c_e and the cost of the start-up be c_s . Our analysis of the asymmetric capacity games in fact already accommodates heterogeneous costs (since costs in the base model varied over time, when firms invest at different times, costs are by definition heterogeneous). Thus, we need only modify our analysis to account for heterogeneous costs in the symmetric investment games. The following lemma summarizes the equilibria to the capacity investment games:

Lemma 5 *If firms have heterogeneous capacity costs that are constant over time, then:*

1. (L,L) *If both firms invest in capacity late, then equilibrium capacities are $K_e^* = \frac{A+c_s-2c_e}{3}$ and $K_s^* = \frac{A+c_e-2c_s}{3}$. The ex-ante equilibrium expected profit of the established firm is*

$$\mathbb{E}(\pi_e^*) = \frac{\sigma^2}{9} + \left(\frac{\mu + c_s - 2c_e}{3} \right)^2,$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F(3\sqrt{\alpha} + 2c_s - c_e).$$

2. (E,L) *If the start-up invests early while the established firm invests late, equilibrium capacities, profits, and survival probabilities are identical to those derived in Lemma 2, with $c_e = c_2$ and $c_s = c_1$.*
3. (E,E) *If both firms invest early, equilibrium capacities are $K_e^* = \frac{\mu - \sqrt{\alpha} - c_e}{2}$ and $K_s^* = \sqrt{\alpha}$.*

The ex-ante equilibrium expected profit of the established firm is

$$\mathbb{E}(\pi_e^*) = \frac{1}{4} (\mu - c_e - \sqrt{\alpha})^2,$$

while the ex-ante equilibrium survival probability of the start-up is

$$\psi_s^* = 1 - F\left(\frac{\mu + 3\sqrt{\alpha} + 2c_s - c_e}{2}\right).$$

4. (L,E) If the established firm invests early while the start-up invests late, equilibrium capacities, profits, and survival probabilities are identical to those derived in Lemma 4, with $c_e = c_1$ and $c_s = c_2$.

Armed with the equilibrium survival probabilities and expected profits, we may derive the equilibrium to the capacity investment timing game:

Theorem 5 *If firms have heterogeneous capacity costs that are constant over time, a unique equilibrium to the timing game exists. Let*

$$\Sigma_1 \equiv (\mu - c_e - \sqrt{\alpha})^2 - (\mu - c_e - \sqrt{2\alpha})^2.$$

Then the following pure strategy equilibria to the investment timing game exist:

1. If $\sigma^2 > \Sigma_1$, the start-up invests early while the established firm invests late.
2. If $\Sigma_1 > \sigma^2$, both firms invest early.

Proof. Similar to the proof in the base model, we will examine each possible equilibrium individually.

(i) (E, L). This is an equilibrium if no firm has incentive to unilaterally deviate: from Lemma 5, the start-up will not deviate if

$$1 - F\left(2\sqrt{2\alpha} + 2c_s - c_e\right) > 1 - F\left(3\sqrt{\alpha} + 2c_s - c_e\right),$$

which always holds. The equilibrium is supportable if the established firm has no incentive to deviate, i.e., if

$$\frac{\sigma^2 + (\mu - c_e - \sqrt{2\alpha})^2}{4} > \frac{1}{4} (\mu - c_e - \sqrt{\alpha})^2,$$

which holds if $\sigma^2 > (\mu - c_e - \sqrt{\alpha})^2 - (\mu - c_e - \sqrt{2\alpha})^2$ —thus, with constant, heterogeneous costs, (E, L) is an equilibrium if σ^2 is sufficiently large.

(ii) (E, E) . This sequence is only an equilibrium if the established firm has no incentive to deviate, which the analysis of (E, L) showed occurs for low σ^2 . It must also be the case that the start-up has no incentive to deviate, which holds if

$$6\sqrt{\alpha}(c_e - c_s) < 7\alpha + 3(c_e - c_s)^2,$$

which, as we saw in the proof of Theorem 3, always holds. Hence, (E, E) is an equilibrium precisely when (E, L) is not.

(iii) (L, E) . We now consider the equilibrium in which the established firm leads and the start-up follows. The start-up has incentive to remain in this equilibrium if

$$1 - F\left(\sqrt{4\alpha + (c_e - c_s)^2} + \frac{\mu + c_s}{2}\right) > 1 - F\left(\frac{\mu + 3\sqrt{\alpha} + 2c_s - c_e}{2}\right)$$

This never holds for any real $(c_e - c_s)$, hence the start-up will always deviate from (L, E) and this cannot be an equilibrium.

(iv) (L, L) . This sequence is only an equilibrium if the start-up has no incentive to deviate, which the analysis of (E, L) showed is never true. Hence, (L, L) cannot be an equilibrium. ■

Intriguingly, when costs are constant over time but differ between the two firms, only two equilibria can exist in the timing game: either the start-up is the leader, or both firms invest early, thus preserving our primary result (that the start-up is the leader when market variability is high). In addition, we note that so long as capacity costs are constant, no regions of non-existence of multiple equilibria may exist, unlike the case with time-varying capacity costs. Thus, we conclude that the main driver of the problematic equilibrium regions is the (potentially) time-varying nature of capacity costs.

6.3 Holdback

Throughout our analysis, we have assumed that firms always produce to their maximum capacity—that is, both firms follow a *production clearance* strategy. From a modeling perspective this allows for a simple and relatively clean analysis of the capacity investment decision—in the absence of this assumption, closed form solutions for equilibrium capacities, profits, and survival probabilities cannot be obtained—and moreover the clearance assumption may be thought of as the outcome of selling the product at a series of different prices until capacity is exhausted or fully utilized, in which case the “market price” is actually an average price. Additionally, firms frequently produce at maximum capacity because of high fixed costs of starting and stopping the production process (e.g., in the chemical or semiconductor industries): see Goyal and Netessine (2007).

From a practical standpoint, though, it may be unwise for a firm to always produce at maximum capacity. Other papers (e.g., Chod and Rudi 2005) have demonstrated that a clearance assumption typically has a negligible impact on analytical outcomes, however, it is useful to verify this result in our setting. Hence, in this section, we discuss the impact of the alternative assumption: a *holdback* strategy, in which the firms may produce any *ex-post* profit maximizing quantity subject to their individual capacity constraints.

It is first useful to consider the qualitative impact of holdback. In fact, our previous analysis accommodates a holdback strategy whenever a firm invests *late*—this is because we assume that late investment occurs after the resolution of demand uncertainty, hence the firm would never invest in more capacity than necessary for maximizing profit (i.e., a firm investing in capacity late always produces to full capacity). Thus, the analysis for any firm investing late is unchanged if holdback is allowed. Furthermore, the analysis of a start-up investing early is also unchanged by the option of holdback. Recall that a start-up investing early chooses the minimum capacity level that supports survival—in other words, if a start-up invests in K_s units of capacity, it must sell all K_s units to survive. As a result, a start-up investing early will always produce up to its maximum capacity level if it survives; holdback could only occur in demand states in which the start-up does not survive, which does not impact the start-up’s subsequent probability of survival.

It follows, then, that holdback only affects an established firm investing early. Intuitively, granting such a firm the option of producing less can only increase the value of early investment

relative to late investment. Some incentive for late investment remains, though, particularly if capacity costs are significant relative to marginal production costs; in that case, there is still value to waiting for the resolution of demand uncertainty to avoid sinking excess money into costly capacity. Hence, we postulate that allowing holdback increases established firm incentives for early investment without completely eliminating incentives for late investment.

While this thought experiment helps to understand the impact of holdback on firm profit, without further analysis, it's unclear how holdback affects the competing start-up's survival probability and the equilibrium to the timing game. To that end, we conducted a numerical study to explore precisely this issue. The model employed in the study is identical to the one analyzed in the rest of the paper, save for the fact that firms are allowed to engage in holdback. The result is that a *quantity game* occurs at the start of the selling season: after observing the realized value of market size (A), firms choose production quantities to maximize expected profit, subject to their individual capacity constraints. (As in the base model, production is assumed to be costless, though positive production costs do not qualitatively change any results).

To analyze this more complicated model, we must make an additional assumption concerning the order of moves in the quantity game. A variety of plausible options exist (e.g., the leader in the capacity game is the leader in the quantity game; the established firm is the leader in the quantity game due to greater market power; firms strategically time their quantity decisions just as they do their capacity decisions). We choose the simplest sequence: simultaneous quantity competition. Thus, the firms engage in capacitated Cournot competition in the quantity game—see Gabszewicz and Poddar (1997) for a proof of existence of an equilibrium in this subgame (as well as an analysis of a similar *ex-ante* capacity investment game, but with two profit-maximizing firms moving simultaneously). We examined 135 parameter instances consisting of every combination in Table 2, which were selected to provide a wide range of possible scenarios (e.g., low/medium/high demand variability, various product margins, etc.). In each case, we calculated the equilibrium to the investment timing game with holdback and with clearance. Comparing the incidence of specific equilibria between the two possible assumptions allows us to determine the impact of holdback on our theoretical results.

Our results are summarized in Tables 3 and 4. We first note that the equilibrium sequence of investment was identical with holdback as with clearance in 62% of cases. In other words, allowing

Parameter	Values
Demand Distribution (A)	Gamma
μ	{10, 20, 30}
σ	{10, 20, 30}
c_1	{1}
c_2	{0.5, 0.75, 1, 1.25, 1.5}
α	{10, 20, 30}

Table 2. Parameter values used in numerical experiments.

	Established Firm Early	Established Firm Late
Start-up Early	3%	77%
Start-up Late	0%	20%

Table 3. Incidence of unique equilibria to the investment timing game with production clearing.

holdback does not change the order of investment in the majority of cases considered. In the remaining 38% of cases, holdback *does* change the equilibrium investment sequence. As the tables demonstrate, when holdback is allowed, the (E, L) equilibrium—with the start-up as leader and the established firm as the follower—is less common than when production clearing is assumed (occurring in 39% of cases with holdback, versus roughly 77% of cases with clearance). The equilibria in which both firms invest late or both invest early are both more common with holdback, as are instances of multiple equilibria. Generally, it appears that holdback shifts equilibria from instances with the start-up acting as a leader to instances in which the firms invest simultaneously. Most crucially, holdback never results in the reverse sequential outcome (i.e., the established firm leading and the start-up following).

We also observed that as the coefficient of variation of demand increases, the (E, L) equilibrium becomes more common, even with holdback: when $\sigma/\mu = 0.33$, none of the cases analyzed with holdback resulted in the (E, L) equilibrium. The frequency of (E, L) equilibria increases as σ/μ increases, and when $\sigma/\mu = 3$, the maximum coefficient of variation in our sample, 53% of the cases analyzed with holdback resulted in the (E, L) equilibrium (with clearance, 80% of the cases analyzed resulted in (E, L) for this value of σ/μ). Thus, the result that the investment timing equilibrium with the start-up acting as a leader is more common with higher demand variability seems to be preserved with holdback, and we conclude that while holdback *does* have an impact on precisely which equilibrium results from the timing game, it does not change the overall structure of our results.

	Established Firm Early	Established Firm Late
Start-up Early	15%	39%
Start-up Late	0%	40%

Table 4. Incidence of unique equilibria to the investment timing game with holdback. In addition to the four pure strategy single equilibria, we observed multiple equilibria in 6% of cases.

7 Conclusion

In this work, our chief goal was to analyze how the threat of bankruptcy impacts the capacity investment and timing decisions of firms entering new markets. We find that in monopoly markets, start-ups are more likely to prefer early capacity investment than profit-maximizing established firms. In competitive markets, when demand uncertainty is large, the outcome of a strategic investment timing game leads to an equilibrium in which the start-up invests early while the established firm invests late—starkly contrasting to a model with two established firms, which leads to simultaneous late investment under high demand uncertainty.

We arrived at these results despite invoking several assumptions intended to minimize the incidence of sequential equilibria. For example, in previous literature, one explanation offered for established firms failing to seize opportunities in disruptive markets is that their demand forecasts are too pessimistic or simply inaccurate. We have found, on the contrary, that even if both firms have *identical* demand forecasts (i.e., identical beliefs about the distribution of market size), sequential equilibria arise if a start-up is present. If we incorporated pessimistic forecasts by established firms into our model, this would have the effect of decreasing the expected market size in the established firm’s profit function, qualitatively preserving our results (and giving the established firm even more incentive to invest late). Similarly, we assumed that both firms have access to the technology that enables the new market at the start of the strategic investment game—in other words, no firm is playing catch-up from a technological standpoint, and both are capable of capacity investment at any time.

We also did not model a variety of other factors that may influence capacity investment timing. For example, greater sales may be enabled by earlier entry (e.g., if late entry results in slower time-to-market). Directionally, the impact of this effect is clear: it increases firm incentives to invest early. While this would likely change the equilibrium thresholds given in Theorem 3, the qualitative impact of the start-up’s bankruptcy-averse objective function remains (as do the consequences of

acting as a first- or second-mover in the capacity game), implying that the strategic investment game will have a similar structure and will yield similar results.

We conclude that capacity competition involving start-ups is fundamentally different in nature from the competition between established firms, and our model offers a plausible explanation of some observed phenomena. Managerially, these results are important because they imply that the optimal strategic investment position differs depending on the nature of the competitor. Thus, blindly following a mantra of seizing the “first-mover advantage” can be a perilous strategy, as any such advantage (or disadvantage) depends critically on the characteristics of the firms in the market.

Our results also relate to the literature on disruptive innovation, which has frequently observed that start-ups tend to pioneer new markets while established firms postpone investment. A variety of reasons for this phenomenon are offered: the established firms are said to be too close to and too trusting of their existing customers, who themselves are ill-equipped to articulate their own changing needs, therefore causing a failure to anticipate opportunities within the existing customer base; the established firms fail to recognize and cultivate entirely new markets; internal incentives at the established firms favor the development and implementation of incremental improvement over radical change. All of these explanations imply that established firms *fail* in some crucial way that newer firms do not. Our results imply that, while it is certainly possible that managerial failures can lead to established firms detrimentally ceding a leadership role to start-ups in new markets, this need not be the case; the operational reality of capacity investment under demand uncertainty, coupled with facing competition from start-ups prone to failure, offers a purely rational explanation for these outcomes.

A Appendix: Proofs

Proof of Lemma 1. Because there is no randomness if both firms invest late, the capacity investment game is a Cournot duopoly with heterogeneous costs. Thus, the profit of each firm is given by

$$\begin{aligned}\pi_e(K_e) &= (A - K_e - K_s - c_2) K_e, \\ \pi_s(K_s) &= (A - K_e - K_s - c_2) K_s.\end{aligned}$$

Both profit functions are concave, yielding unique best replies

$$K_e^*(K_s) = \frac{A - K_s - c_2}{2} \text{ and } K_s^*(K_e) = \frac{A - K_e - c_2}{2}.$$

The equilibrium capacities are found by solving for the intersection of the best replies, which yields the unique equilibrium $K_e^* = K_s^* = \frac{A - c_2}{3}$. Equilibrium profit of each firm is

$$\mathbb{E}(\pi_e^*) = \mathbb{E}(\pi_s^*) = \mathbb{E}\left(\frac{A - c_2}{3}\right)^2 = \frac{\sigma^2 + (\mu - c_2)^2}{9}.$$

Recall that the start-up survives if the total profit level is above α : in other words, if $((A - c_2)/3)^2 \geq \alpha$. Thus, the *ex-ante* survival probability (i.e., the probability of survival *before* learning market size, taking into account the competitive outcome of the capacity game that occurs *after* learning market size) is given by (8), while the (*ex-ante*) equilibrium expected profit of the established firm is given by (7).

Proof of Lemma 2. Recall that the best reply of the established firm is $K_e^*(K_s) = \frac{A - K_s - c_2}{2}$ when both firms invest late: this continues to hold when the start-up invests early and the established firm invests late. The start-up's profit is thus

$$\begin{aligned} \pi_s(K_s) &= (A - K_e^*(K_s) - K_s - c_1) K_s \\ &= \frac{1}{2}(A - K_s - 2c_1 + c_2) K_s. \end{aligned}$$

The survival probability is the probability that $\pi_s(K_s) \geq \alpha$, i.e.,

$$\begin{aligned} \psi_s(K_s) &= \Pr\left(\frac{1}{2}(A - K_s - 2c_1 + c_2) K_s \geq \alpha\right) \\ &= 1 - F\left(\frac{2\alpha}{K_s} + K_s + 2c_1 - c_2\right). \end{aligned}$$

The maximizer of the survival probability is the minimizer of the argument of F in the above equation, i.e., $K_s^* = \sqrt{2\alpha}$, yielding (10) when substituted into the expression for the start-up's survival probability. The established firm's profit is

$$\pi_e(K_e) = \frac{1}{4}(A - c_2 - \sqrt{\alpha})^2,$$

and *ex-ante* expected profit is thus given by the expected value of this expression, yielding (9).

Proof of Lemma 3. Survival for the start-up occurs if $A \geq \frac{\alpha}{K_s} + K_e + K_s + c_1$, so the survival probability is thus

$$\psi_s(K_s, K_e) = 1 - F\left(\frac{\alpha}{K_s} + K_e + K_s + c_1\right).$$

Minimizing the the argument of F in the above expression is equivalent to maximizing the probability of survival. Thus, the start-up's optimal capacity investment is $K_s^* = \sqrt{\alpha}$, a dominant action that is independent of the established firm's capacity level. The established firm's expected profit is

$$\mathbb{E}(\pi_e(K_s, K_e)) = (\mu - K_e - K_s - c_1) K_e.$$

Substituting the equilibrium K_s^* and maximizing this concave function of K_e yields the established

firm's optimal capacity, $K_e^* = \frac{\mu - \sqrt{\alpha} - c_1}{2}$. The associated expected profit is (11), and the equilibrium survival probability of the start-up is hence (12).

Proof of Lemma 4. The best reply of the start-up investing late is the same as in Lemma 1, i.e., $K_s^*(K_e) = \frac{A - K_e - c_2}{2}$. Hence, the established firm's expected profit from early investment is

$$\mathbb{E}(\pi_e(K_e)) = \left(\mu - K_e - \frac{\mu - K_e - c_2}{2} - c_1 \right) K_e$$

Maximizing this expression yields an optimal capacity level of $K_e^* = \frac{\mu - 2c_1 + c_2}{2}$ for the established firm and hence

$$K_s^* = \frac{2A - \mu + 2c_1 - 3c_2}{4}$$

for the start-up. The equilibrium expected profit of the established firm is thus (13), and the start-up's equilibrium survival probability is (14).

Proof of Lemma 5. Omitted—similar to Lemmas 1–4.

References

- Archibald, T. W., L. C. Thomas, J. M. Betts, R. B. Johnston. 2002. Should start-up companies be cautious? Inventory policies which maximise survival probabilities. *Management Sci.* **48**(9) 1161–1174.
- Babich, V. 2008. Independence of capacity ordering and financial subsidies to risky suppliers. Working paper, University of Michigan.
- Babich, V., A. N. Burnetas, P. H. Ritchken. 2007. Competition and diversification effects in supply chains with supplier default risk. *Manufacturing Service Oper. Management* **9**(1) 123–146.
- Bayus, B. L., R. Agarwal. 2007. The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. *Management Sci.* **53**(12) 1887–1902.
- Bhaskaran, S., K. Ramachandran. 2007. Competitive product introductions in technologically dynamic environments. Working paper, Southern Methodist University.
- Bower, J., C. Christensen. 1995. Disruptive technologies: Catching the wave. *Harvard Business Review* **73**(1) 43–53.
- Boyabatli, O., B. Toktay. 2007. Stochastic capacity investment and technology choice in imperfect capital markets. Working paper, Singapore Management University.
- Chod, J., E. Lyandres. 2008. Strategic IPOs and product market competition. Working paper, Boston College.
- Chod, J., N. Rudi. 2005. Resource flexibility with responsive pricing. *Oper. Res.* **53**(3) 532–548.
- Christensen, C. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Press.
- Christensen, C., J. Bower. 1996. Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal* **17** 197–218.

- Druehl, C. T., G. M. Schmidt. 2008. A strategy for opening a new market and encroaching on the lower end of the existing market. *Production & Operations Management* **17**(1) 44–60.
- Dutta, P. K., R. Radner. 1999. Profit maximization and the market selection hypothesis. *The Review of Economic Studies* **66**(4) 769–798.
- Gabszewicz, J., S. Poddar. 1997. Demand fluctuations and capacity utilization under duopoly. *Economic Theory* **10**(1) 131–146.
- Gal-Or, E. 1985. First mover and second mover advantages. *International Economic Review* **26**(3) 649–653.
- Goyal, M., S. Netessine. 2007. Strategic technology choice and capacity investment under demand uncertainty. *Management Sci.* **53**(2) 192–207.
- Greenwald, B., J. Stiglitz. 1990. Asymmetric information and the new theory of the firm: Financial constraints and risk behavior. *Amer. Econom. Rev.* **80**(2) 160–165.
- Hamilton, J. H., S. M. Slutsky. 1990. Endogenous timing in duopoly games: Stackelberg or Cournot equilibria. *Games and Economic Behavior* **2**(1) 29–46.
- Joshi, Y. V., D. J. Reibstein, Z. J. Zhang. 2009. Optimal entry timing in markets with social influence. *Management Sci.* **55**(6) 926–939.
- Klastorin, T., W. Tsai. 2004. New product introduction: Timing, design, and pricing. *Manufacturing Service Oper. Management* **6**(4) 302–320.
- Lilien, G. L., E. Yoon. 1990. The timing of competitive market entry: An exploratory study of new industrial products. *Management Sci.* **36**(5) 568–585.
- Maggi, G. 1996. Endogenous leadership in a new market. *RAND J. Econom.* **27**(4) 641–659.
- Radner, R., L. Shepp. 1996. Risk vs. profit potential: A model for corporate strategy. *Journal of Economic Dynamics and Control* **20**(8) 1373–1393.
- Saloner, G. 1987. Cournot duopoly with two production periods. *Journal of Economic Theory* **42**(1) 183–187.
- Schmidt, G. M., C. T. Druehl. 2008. When is a disruptive innovation disruptive? *Journal of Product Innovation Management* **25**(4) 347–369.
- Spence, A. M. 1977. Entry, capacity, investment and oligopolistic pricing. *The Bell Journal of Economics* **8**(2) 534–544.
- Swinney, R., S. Netessine. 2009. Long-term contracts under the threat of supplier default. *Manufacturing Service Oper. Management* **11**(1) 109–127.
- Tanrisever, F., S. Erzurumlu, N. Joglekar. 2008. Process development and survival of startup firms. Working paper, University of Texas, Austin.
- Van Mieghem, J. 2003. Capacity management, investment, and hedging: Review and recent developments. *Manufacturing Service Oper. Management* **5**(3) 269–302.
- Walls, M., J. Dyer. 1996. Risk propensity and firm performance: A study of the petroleum exploration industry. *Management Sci.* **42**(7) 1004–1021.