Getting Big Too Fast: Strategic Dynamics with Increasing Returns and Bounded Rationality

John D. Sterman, Rebecca Henderson
Sloan School of Management, Massachusetts Institute of Technology, 30 Wadsworth Street, Cambridge, Massachusetts 02142 {jsterman@mit.edu, rhenders@mit.edu}

Eric D. Beinhocker
McKinsey Global Institute, 30 Kensington Church Street, London W8 4HA, United Kingdom, eric.beinhocker@mckinsey.com

Lee I. Newman
Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, Michigan 48109, leenewm@umich.edu

Neoclassical models of strategic behavior have yielded many insights into competitive behavior, despite the fact that they often rely on a number of assumptions—including instantaneous market clearing and perfect foresight—that have been called into question by a broad range of research. Researchers generally argue that these assumptions are “good enough” to predict an industry’s probable equilibria, and that disequilibrium adjustments and bounded rationality have limited competitive implications. Here we focus on the case of strategy in the presence of increasing returns to highlight how relaxing these two assumptions can lead to outcomes quite different from those predicted by standard neoclassical models. Prior research suggests that in the presence of increasing returns, tight appropriability, and accommodating rivals, in some circumstances early entrants can achieve sustained competitive advantage by pursuing “get big fast” (GBF) strategies: Rapidly expanding capacity and cutting prices to gain market share advantage and exploit positive feedbacks faster than their rivals. Using a simulation of the duopoly case we show that when the industry moves slowly compared to capacity adjustment delays, boundedly rational firms find their way to the equilibria predicted by conventional models. However, when market dynamics are rapid relative to capacity adjustment, forecasting errors lead to excess capacity—overwhelming the advantage conferred by increasing returns. Our results highlight the risks of ignoring the role of disequilibrium dynamics and bounded rationality in shaping competitive outcomes, and demonstrate how both can be incorporated into strategic analysis to form a dynamic, behavioral game theory amenable to rigorous analysis.

Key words: marketing; applications; simulation; strategy; bounded rationality

History: Accepted by Bruno Cassiman and Pankaj Ghemawat, special issue editors; received February 15, 2006. This paper was with the authors 3 months for 1 revision.

1. Introduction

The tools of classical game theory have been a source of great insight into a wide variety of strategic problems. In studies of phenomena as dispersed as entry, pricing, diffusion, and investment, models employing the standard assumptions of perfect rationality and full information have demonstrated their usefulness in predicting behavior (for a small sampling of a huge and diverse literature see, for example, Besanko et al. 2003, Bresnahan and Reiss 1991, Hendricks and Porter 2003, MacDonald and Ryall 2004, and Tirole 1990). However, a significant body of research suggests that under many circumstances the core behavioral assumptions of neoclassical economics are inconsistent with empirical observation (see, for example, Beinhocker 2006, Camerer et al. 2004, Gilovich et al. 2002, Kahneman and Tversky 2000, Colander et al. 2004, Camerer and Fehr 2006). Despite this, the mainstream literature has generally assumed that in practice these assumptions are “good enough,” and has focused on the characteristics of equilibrium—assuming that while adjustment may be costly and decision making imperfect, in the end industries will converge to the equilibrium predicted by conventional theory.

This stance has, of course, been hotly contested by evolutionary theorists for many years, who argue that holding fast to the traditional simplifying assumptions of neoclassical theory may be dangerous in formulating normative policies, particularly in settings with high dynamic complexity (Simon 1982, Nelson and Winter 1982, Dosi 1997, Gavetti and Levinthal 2000). Nevertheless, scholars have only begun to sketch out how alternative assumptions and tools might yield superior and robust implications for managerial action.
Here we explore the conditions under which disequilibrium dynamics and limited information-processing capability may lead to normative conclusions quite different than those of models assuming full rationality. We focus on the particular case of increasing returns, and the commonly associated recommendation to “get big fast” (GBF) as a particularly compelling example of the risks of assuming that the assumptions of neoclassical theory are “good enough” to provide a basis for action.

Research in strategy and economics has long identified increasing returns, or positive feedback effects, as a potentially potent source of competitive advantage. These positive feedbacks include learning by doing, scale economies, network effects, information contagion, and the accumulation of complementary assets. A large and fruitful literature suggests that in the presence of such positive feedbacks it may be possible for aggressive firms to build persistent competitive advantage.

Spence (1979, 1981), for example, showed that learning effects create asymmetric advantage and thus an incentive to preempt rivals, particularly if firms can appropriate the benefits of learning. His work has been the basis for a lively literature exploring a range of extensions (e.g., Kalish 1983, Tirole 1990, Majd and Pindyck 1989, Ghemawat and Spence 1985), which clarified the conditions under which aggressive strategies are likely to succeed.

Moving beyond the learning curve, research exploring industries with strong network effects has also identified increasing returns as a central source of competitive advantage (Katz and Shapiro 1994, Shapiro and Varian 1999, Fudenberg and Tirole 2000, Parker and van Alstyne 2005). Arthur (1989, 1994) shows how positive feedbacks can lead to lock-in and path dependence. Sutton (1991) shows that increasing returns flowing from economies of scope in advertising can lead a few firms to dominate an industry. Sutton (1998) also suggests that under some circumstances learning via R&D can have similar effects. Jovanovic (1982) and Klepper (1996) both develop models in which dominant firms emerge as initially heterogeneous costs are amplified by positive feedbacks.

Within this tradition, many scholars have been careful to highlight the limitations of these insights as a basis for action. In general, the literature suggests that if the effects of increasing returns are privately appropriable and rivals are likely to accommodate aggressive behavior, then it may be rational for a firm to pursue an aggressive strategy and seek to grow faster than their rivals (e.g., Shapiro and Varian 1999, Fudenberg and Tirole 2000). Typical tactics include pricing below the short-run profit-maximizing level, rapidly expanding capacity, advertising heavily, and forming alliances to build positional advantage and deter entry (Spence 1981; Fudenberg and Tirole 1983, 2000; Tirole 1990).

Intuitively, such aggressive strategies are superior because they increase both industry demand and the aggressive firm’s share of that demand, boosting cumulative volume, reducing future costs, and building the firm’s positional advantage until it dominates the market. Aggressive strategies appear to have led to durable advantage in industries with strong learning curves such as synthetic fibers, chemicals, and disposable diapers (Shaw and Shaw 1984, Lieberman 1984, Ghemawat 1984, Porter 1984), and in markets with network externalities and complementary assets, such as VCRs and personal computers.

Of course, as a lively tradition in the strategic management literature has suggested, translating these results into a blanket prescription to get big fast may be extremely dangerous. Porter (1980) presents an extensive list of circumstances under which a strategy of aggressive preemption is likely to fail—suggesting, for example, that if capacity must be added in large lumps, or is only available with long lead times, or if information flow is asymmetric or distorted, then attempting to exploit a first mover advantage may be disastrous. Similarly, Ghemawat (1987) explores the effects of divergent beliefs about the future on dynamic games in capacity. He shows that although many formal models assume perfect information, firms may hesitate to invest if there is significant uncertainty, (rationally) fearing a significant adverse selection problem. He gives examples of industries in which dominant firms either hesitated to invest aggressively or regretted having done so (a dynamic directly analogous to the winner’s curse). Goldfarb et al. (2006) show how information imperfections and delays led to the emergence of what they term a “Get Big Fast Cascade” during the Internet boom, and subsequent low or negative returns to equity investments in that period. Shapiro and Varian (1999) and Lieberman and Montgomery (1998) echo these concerns, pointing out that a blanket prescription to move first can be dangerous given the subtleties of many industries.

Here we build on these insights to build a formal model of the circumstances under which the combination of delays in capacity adjustment and locally rational decision making can undo the results of the conventional neoclassical models. Following Porter (1980), we focus on the twin assumptions of instantaneous capacity adjustment and perfect foresight as particularly problematic.

If firms were well informed and could forecast accurately, capacity would match orders well (at least on average). Alternatively, even if forecasting ability were poor, capacity could match demand well if it
could be adjusted rapidly and at low cost. We show that the ability of firms to exploit increasing returns can be compromised by realistic adjustment rigidities and commonly used forecasting heuristics even when conditions otherwise favor the aggressive strategy.

When the dynamics of the market are sufficiently slow, delays in information acquisition, decision making, and system response are sufficiently short, and the cognitive demands on the firm’s managers are sufficiently low, our model yields predictions observationally indistinguishable from those of neoclassical models. In these conditions, the traditional assumptions of the mainstream game-theoretic literature are, indeed, “good enough.”

However, in more complex and dynamic environments—particularly those in which demand evolves quickly relative to capacity adjustment—aggressive strategies may lead to disaster, even when the conditions for success specified in the neoclassical literature have been met. In these circumstances, managers are not able to anticipate the saturation of the market in time to reduce capacity. As long as the industry is growing, all is well, but when sales peak and fall, firms find themselves with excess capacity. The more aggressive the firm’s strategy, the more pronounced the overcapacity and resulting losses. We show that the failure of the aggressive strategy when the market dynamics are rapid is not due to the failure of increasing returns to confer advantage on the aggressive firm. Rather, the failure of the aggressive strategy is due to the interaction of capacity adjustment lags with the firm’s boundedly rational forecasting heuristic.

As a typical example, consider fiber optic equipment maker JDS Uniphase (JDSU; Figure 1). The firm faced few serious rivals and could plausibly hope to appropriate the benefits of increasing returns through proprietary technology and capabilities—both conditions identified as critical to a GBF strategy by the neoclassical literature. Throughout the boom of the late 1990s JDSU aggressively expanded capacity and employment, both internally and through acquisition. Lags in capacity expansion, however, constrained production growth, and the backlog of unfilled orders ballooned. The collapse of demand caught the firm by surprise. Lags in reducing capacity meant costs could not drop as fast as sales: Though it eventually cut employment by more than 23,000 (81%), JDSU operated with a negative gross margin for most of the next year, posted losses of roughly $60 billion between 2001 and 2003, and saw its stock price fall 99%.

Our model reproduces these results without the need to assume that people are naive automata, making myopic decisions without regard to strategic considerations. The model’s agents monitor market conditions, including the plans and actions of their competitors, and adjust their behavior accordingly. However, their rationality is bounded: In the tradition of Simon (1982), Cyert and March (1963/1992), Forrester (1961), and Nelson and Winter (1982), the agents make decisions using routines and heuristics because the complexity of the environment exceeds their ability to optimize even with respect to the limited information available to them.

The paper begins with a brief presentation of the model (the online supplement provides full documentation and the model itself is available online with...
the software needed to run it; the online supplement is provided in the e-companion1). For simplicity, we do not attempt to capture all sources of increasing returns, but focus on the learning curve, which is a source of positive feedback prevalent in many industries and well explored in the literature. Section 3 presents our results and explores their sensitivity to key assumptions. We conclude with a discussion of implications and avenues for further research. Several researchers have suggested that “lumpy” capacity and imperfect information may make GBF strategies problematic. We show formally how, even with continuous capacity, relatively short delays in capacity acquisition coupled with locally rational behavioral decision rules that are strongly supported by the empirical literature can lead to results significantly different from those predicted by models based on the traditional assumptions of full information and perfect rationality. We believe that delineating the circumstances under which the standard neoclassical assumptions yield misleading results is one of the major contributions of our paper, and more generally we suggest that in cases of high dynamic complexity, a reliance on the standard assumptions underlying much modern game-theoretic research is not inconsequential. We suspect that the use of analytical techniques that can formalize many of the intuitions now current in the literature and in management practice may permit a wider (and more realistic) set of assumptions that could be of significant utility to the field.

2. A Boundedly Rational, Disequilibrium Model

The model consists of firms that compete against one another as they interact with a customer sector. We begin by describing firm behavior and then turn to a discussion of the evolution of demand. The firm model captures order fulfillment, capacity acquisition, costs, and pricing. The customer sector generates industry demand as a function of product adoption, price, and initial and replacement purchases.

As founding assumptions, we assume capacity adjusts with a lag, and that firms do not have the ability to forecast sales perfectly. These assumptions are consistent with a long tradition of experimental and empirical evidence (Armstrong 2001; Brehmer 1992; Collopy and Armstrong 1992; Diehl and Sterman 1995; Kampmann 1992; Paich and Sterman 1993; Parker 1994; Rao 1985; Sterman 1989a, b; 1994). In traditional models the market-clearing price can be derived as a necessary property of equilibrium, given the capacity decision. In disequilibrium settings, however, both price and capacity targets must be determined. Here we draw on the literature cited above and the well-established tradition of bounded rationality (Cyert and March 1963/1992, Forrester 1961, Simon 1982, Morecroft 1985), and assume that firms set prices with intentionally rational decision heuristics. We demonstrate the local rationality of firms in the model by showing that the model generates the neoclassical results when capacity can be adjusted in a sufficiently quick manner relative to the dynamics of demand such that the firms’ demand forecasts and estimates of their competitors’ capacity plans are reasonably accurate.

The model is formulated in continuous time as a set of nonlinear differential equations. Because no analytic solution is known, we use simulation to explore its dynamics. Although the model portrays an industry with an arbitrary number of firms, \( i \in \{1, \ldots, n\} \), we restrict ourselves to \( n = 2 \) in the simulation experiments below.

2.1. The Firm

Firm profits are revenue, \( R \), less fixed and variable costs, \( C^f \) and \( C^v \), respectively (the firm index \( i \) is deleted for clarity):

\[
\pi = R - (C^f + C^v). \tag{1}
\]

Fixed costs depend on unit fixed costs, \( U^f \), and current capacity, \( K \); variable costs depend on unit variable costs, \( U^v \), and production, \( Q \).

\[
C^f = U^f K; \quad C^v = U^v Q. \tag{2}
\]

Both fixed and variable costs per unit fall as cumulative production experience, \( E \), grows, according to a standard learning curve:

\[
U^f = U^f_0 (E/E_0)^\gamma; \quad U^v = U^v_0 (E/E_0)^\gamma; \tag{3}
\]

\[
dE/dt = Q; \tag{4}
\]

where \( U^f_0 \) and \( U^v_0 \) are the initial values of unit fixed and variable costs, respectively. \( E_0 \) is the initial level of production experience and \( \gamma \) is the strength of the learning curve.

Production, \( Q \), is the lesser of desired production, \( Q^* \), and capacity, \( K \).

\footnote{An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.} For simplicity we assume that the firm maintains no inventories and makes all product to order. Shipments thus equal production. Including inventories and distribution channels would intensify disequilibrium dynamics through the well-known “bullwhip effect” (Sterman 1989b, Lee et al. 1997).
the backlog of unfilled orders, $B$, and target delivery delay $\tau^*$. Backlog accumulates orders, $O$, less production:

$$Q = \text{MIN}(Q^*, K), \quad (5)$$

$$Q^* = B/\tau^*, \quad (6)$$

$$dB/dt = O - Q. \quad (7)$$

Capacity cannot be changed instantly, but adjusts to the target level $K^*$ with an average lag $\lambda$.

2.2. Firm Strategy

Under the traditional assumption of full rationality, each firm’s target capacity and pricing behavior would be given by the solution to the differential game defined by the structure of the firm and of customer demand. In reality, however, managers do not make decisions by solving dynamic programming problems of such complexity (e.g., Camerer et al. 2004, Camerer and Fehr 2006). Rather, they use intendedly rational heuristics to set prices and acquire capacity, and the game theoretic models reach managers in the form of case studies and rules of thumb. In the presence of increasing returns, books and consultants prescribe rules such as, “By slashing prices below costs, winning the biggest share of industry volume, and accelerating its cost erosion, a company [can] get permanently ahead of the pack… [and build] an unchallengeable long-term cost advantage” (Rothschild 1990, p. 181). Similarly, in 1996 the Wall Street Journal noted the popularity of “the notion of increasing returns, which says that early dominance leads to near monopolies as customers become locked in and reluctant to switch to competitors. Now, dozens of companies are chasing market share” (Hill et al. 1996). In this spirit, we model target capacity and price with realistic boundedly rational heuristics; heuristics that allow us to capture different strategies for managing the product lifecycle and learning curve, including the “market share advantage leads to lower costs which leads to greater market share advantage” logic derived from the increasing returns literature.

2.3. Target Capacity and Demand Forecasting

Due to the capacity-acquisition delay each firm must forecast future industry demand and then determine what share of that demand it seeks to capture. Firms pursuing GBF strategies will seek the dominant share of the market. Such a firm must acquire capacity sufficient to supply its target share, $S^*$, of the industry demand it forecasts, $D^*$ (adjusted by the normal capacity utilization rate, $u^*$):

$$K^* = \text{MAX}(K_{\text{min}}, S^*D^*/u^*), \quad (8)$$

where $K_{\text{min}}$ is the minimum efficient scale of production.

The capacity-acquisition delay requires the firm to forecast demand $\lambda$ years ahead. Many studies show that forecasts are dominated by smoothing and extrapolation of recent trends (e.g., Cololly and Armstrong 1992, Sterman 2000). We capture such heuristics by assuming firms extrapolate demand $\lambda$ years ahead on the assumption that recent growth will continue. The expected growth rate in demand, $g^*$, is estimated from reported industry demand, $D^*$, over a historical horizon, $h$.

$$D^* = D^* \exp(\lambda g^*), \quad (9)$$

$$g^* = \ln(D_t^*/D_{t-h})/h. \quad (10)$$

It is easy to show that this forecasting heuristic provides a correct, unbiased forecast when demand grows at a constant rate (see the online supplement). However, the reported demand data available as an input for forecasting is noisy. Underlying trends are obscured by transient variations in demand arising from weather, changes in consumer sentiment and liquidity, seasonal factors, and other sources of high-frequency noise. Forecasters therefore face a strong trade-off between responsiveness and overreaction. The longer the historical horizon $h$ used to assess growth, the less vulnerable the firm will be to forecast errors arising from high-frequency noise in demand, but the greater the lag in responding to new trends. Sterman (1987, 2000) provides empirical evidence consistent with such forecasting procedures and shows how changes in growth trends lead to significant overreaction in various industries. Note also that the instantaneous, current industry order rate is not available. Rather, firms rely on consultants and industry associations to estimate current demand. It takes time to collect, analyze, and report such data, so the reported order rate lags current orders (see the online supplement).

The firm’s target market share, $S^*$, depends on the firm’s strategy. We consider two strategies, denoted aggressive and conservative. In the aggressive strategy, the firm follows the recommendation of the popular interpretation of the increasing returns literature by seeking greater market share than its rivals, lowering prices, and expanding capacity. In contrast, the conservative firm seeks accommodation with its rivals and sets a modest market-share goal.

Firms also monitor the plans of their competitors. Because demand forecasts are based on a heuristic, there is in general no guarantee that the capacity acquired by each firm will sum to total industry demand. If a firm’s competitors underforecast demand, the firm may estimate that there is uncontested demand; that is, a gap between its forecast of industry demand $\lambda$ years ahead, when the capacity it orders today will be available, and its forecast of
the capacity its competitors will have at that time. In such a situation the aggressive player seeks to exploit increasing returns not only by setting an aggressive market share goal but also by taking advantage of tidiness, delay, or forecasting error on the part of its rivals by opportunistically increasing its target capacity when it believes competitor capacity will fall short of industry demand. In contrast, the conservative firm seeks accommodation with its rivals, but fears over-capacity and will cede additional share to avoid it.

In the base case we assume firms can accurately assess each competitor’s target capacity, including capacity under construction and capacity plans not yet publicly announced, with only a short delay required for the firm to carry out the required competitive intelligence. Assuming that capacity plans are known favors the GBF strategy by limiting overbuilding due to failure to account for the competitors’ supply line of capacity on order or under construction (Sterman 1989a, b; 2000).

In sum, the capacity-acquisition lag requires firms to forecast future demand. Firms forecast by extrapolating recent trends in demand. Aggressive firms pursuing a GBF strategy seek capacity to command a dominant share of the industry demand they forecast; conservative players constrain their capacity plans to avoid overcapacity. All firms monitor the capacity plans of their rivals, with aggressive firms building more when they detect that their rivals are building too little, and conservative players cutting back when they detect that their rivals are building too much. The online supplement provides full documentation.

2.4. Pricing
Firms do not have the ability to determine the optimal price and instead must search for an appropriate price level. Due to decision making and administrative lags, price, $P$, adjusts to a target level, $P^*$, with an adjustment time $\tau$: 

$$\frac{dP}{dt} = (P^* - P)/\tau^*.$$  

(11)

We assume firms use the anchoring and adjustment heuristic to estimate target prices. The current price forms the anchor, which is then adjusted in response to unit costs, the demand-supply balance, and market share.

$$P^* = \text{MAX}[U^*, P \cdot f(\text{unit costs, demand-supply balance, market share})]$$

$$= \text{MAX}[U^*, P \cdot f(U^* + U^1, Q^*/(u^*K), S^* - S)],$$  

(12)

where the MAX function prevents the firm from pricing below unit variable cost $U^*$. The price discovery process constitutes a hill-climbing heuristic in which the firm searches for better prices in the neighborhood of the current price, using price relative to unit costs, demand-supply balance, and market share relative to its target to assess the gradient (Sterman 2000). The first term ensures that, ceteris paribus, prices fall as unit costs $(U^* + U^1)$ decline through the learning curve. The firm also responds to the adequacy of its current capacity, measured by the ratio of desired production $Q^*$ to the rate of output defined by current capacity and normal capacity utilization, $u^*K$. When capacity is insufficient the firm raises its price; excess capacity causes prices to fall. Finally, the firm prices strategically in support of its capacity goals by adjusting prices when there is a gap between its target and current market share, $S^* - S$. When the firm desires a greater share than it currently commands, it will lower price; conversely, if the market share exceeds the target the firm increases price—trading share for higher profits and signaling rivals its desire to achieve a cooperative equilibrium. The price formulation is consistent with the behavioral model of price in Cyert and March (1963/1992), and experimental evidence (Paich and Sterman 1993, Kampmann 1992).

2.5. Industry Demand
Total orders for the product evolve according to the standard Bass diffusion model, modified to include both initial and replacement purchases (Bass 1969, Mahajan et al. 1990). The population, $POP$, is divided into adopters of the product, $M$, and potential adopters, $N$. Adoption arises from an autonomous component, representing the impact of advertising and other external influences, and from social exposure and word of mouth (WOM) encounters with those who already own the good,

$$\frac{dM}{dt} = N(\alpha + \beta M/POP),$$  

(13)

where $\alpha$ captures the strength of external influences such as advertising and $\beta$ is the strength of social exposure and WOM generated by adopters. The population that will ultimately adopt the product is a function of product price. We assume a linear demand curve.

Industry orders consist of initial and replacement purchases. Each household orders $\mu$ units when they adopt, so initial purchases are $\mu(dM/dt)$. Households also order replacements as their units reach the end of their useful life.

2.6. Market Share
Each firm receives orders $O_i$ equal to a share of the industry order rate, $S^*_O$, determined by a standard logit choice model. Share depends on both price and availability. Availability does not vary in models where markets clear at all times. In reality product availability varies substantially. For example, rapid growth often causes unintended backlog accumulation, product allocations, and long delivery delays,
shows the evolution of industry orders for each case, 
\( \text{as illustrated by the case of JDSU (Figure 1). Availability is measured by the firm’s average delivery delay, which, by Little’s Law, is the ratio of backlog, } \), B, to shipments, Q. The logit choice model is then given by:

\[
S_i^0 = A_i / \sum A_j, \tag{14}
\]

\[
A_i = \exp(\epsilon_p P_i / P^*) \exp(\epsilon_s (B_i / Q_i) / \tau^*), \tag{15}
\]

where A is product attractiveness, and \( \epsilon_p \) and \( \epsilon_s \) are the sensitivities of attractiveness to price and availability, respectively. Both price and delivery delay are normalized by reference values, \( P^* \) and \( \tau^* \), respectively, so that the sensitivities \( \epsilon \) are comparable dimensionless quantities. Note that because orders and shipments need not be equal, market share, defined as each firm’s share of industry shipments, \( S_i = Q_i / \sum Q_j \), will in general equal the firm’s order share only in equilibrium.

3. Results
We begin by confirming that under conditions of perfect foresight and instantaneous capacity adjustment the model reproduces the conclusions of the neoclassical literature. We then explore the ways in which these conclusions change as these assumptions are relaxed. The dynamics of industry demand depend on the strength of the advertising and word-of-mouth effects, the slope of the demand curve, the fractional product replacement rate, and the strength of the learning curve (Table EC.1 reports the base-case parameters).

For the base case the model is calibrated to capture the dynamics of typical consumer electronics items such as camcorders (Table EC.1).3 We assume a 70% learning curve (costs fall 30% for each doubling of cumulative production), which is a typical value. We also assume that the sensitivity of order share to price is high, implying products are only moderately differentiated by nonprice factors, and that the delays in reporting industry orders and estimating competitor target capacity are only three months. These parameters all favor the success of an aggressive strategy (we present sensitivity analyses below). For illustration, we define three industry demand scenarios: fast, medium, and slow, defined by different strengths of the word-of-mouth effect, \( \beta = 2.0, 1.0, \) and 0.5, respectively. These values generate product lifecycles that span much of the variation in observed diffusion rates (Parker 1994, Klepper and Graddy 1990). Figure 2 shows the evolution of industry orders for each case, assuming that there are no capacity constraints and that prices follow unit costs down the learning curve. In all cases a period of rapid growth is followed by a peak and decline to the replacement rate of demand. The stronger the word-of-mouth feedback, the faster the growth, the earlier and higher the peak in orders, and the larger the decline from peak to equilibrium demand. Demand in the slow scenario peaks after about 20 years, while in the fast scenario, the peak comes at about year 6. Even faster dynamics have been documented (Parker 1994), often with only a few years from boom to bust.

For ease of comparison, both firms have identical parameters and initial conditions. Note in particular that the forecasting procedure used by each firm is identical, so the two firms have consistent beliefs about industry demand and competitor capacity. Only the strategy each pursues may differ. We contrast aggressive and conservative strategies. In the aggressive strategy, the firm seeks at least 80% of the market, and will seek more if it believes its rival is underbuilding. The conservative player is willing to split the market with its rival, but will cede if it perceives that a 50% share would result in excess capacity.

We first replicate the standard results by assuming that capacity can instantly adjust to the level required to provide the target rate of capacity utilization at all times, \( K = Q^\tau / \mu^* \). This perfect-capacity case corresponds to the assumption that the market always clears, either because capacity can be adjusted instantly, or because agents have perfect foresight so

\[ \text{Figure 2} \quad \text{Diffusion Dynamics for Three Values of the Word-of-Mouth Parameter (Slow, Medium, Fast: } \beta = 0.5, 1.2 \text{ Respectively), for the Perfect Capacity Case with Target Market Share for Both Firms = 50\%} \]

\[ \text{Note. Top, industry orders. Bottom, price.} \]

\[ \text{Table EC.1 reports the base-case parameters.} \]

3 As (arbitrary) scaling parameters we set the initial price at $1,000/unit, and the potential size of the market at that price to 60 million households, each seeking \( \mu = 1 \) unit. The replacement rate is 10% per year, and the ratio of fixed to variable costs is 3:1.
that they can anticipate any capacity acquisition lag. Because the market always clears, capacity utilization always equals the target rate, and delivery delays are always normal. The share of orders going to each firm therefore responds only to price, and prices respond only to unit costs and to the gap between the firm’s target and actual market share. Table 1 shows the net present value (NPV) of cumulative profits for the three scenarios.4

In all market scenarios the results are identical to those predicted by the simple neoclassical models. Although the NPV of industry profit is maximized when both firms play the conservative strategy, each firm has a strategic incentive to defect and play the aggressive strategy to exploit increasing returns. However, a firm facing the prospect of playing the conservative strategy against an aggressive competitor can improve its position by switching to the aggressive strategy. Doing so, however, lowers industry profits. The payoffs form a prisoner’s dilemma. As the literature suggests, industry profitability is destroyed when both firms indulge in a high stakes game of chicken by playing aggressively throughout. Of course in an asymmetric situation where one player has an initial lead, the optimal strategy is to preempt rivals.

The faster the dynamics of the market unfold, the greater industry profits are for any strategy combination. Figure 3 shows payoffs to each strategy combination in the market clearing case as the word of mouth parameter $\beta$ varies. Consistent with the literature (e.g., Kalish 1983), stronger word of mouth brings people into the market sooner, boosting profits and the advantage of the GBF strategy. Also consistent with the literature, the faster the product lifecycle unfolds, the greater is the strategic incentive to defect and play the aggressive strategy (Table 1). In the slow demand-growth scenario a player can improve the NPV of its profits by $1.3 billion by defecting and playing the aggressive strategy while its rival plays conservatively. In the fast demand-growth case defecting improves the NPV of profits by $1.8 billion.

Figure 4 shows the dynamics in the perfect-capacity case where one firm plays the aggressive strategy and the other plays conservatively. As one would expect, all the key variables trend smoothly toward their equilibrium values. Accurate forecasting and knowledge of competitor capacity plans mean capacity is always exactly at the desired level, so capacity utilization never varies. The aggressive firm rapidly achieves lower costs, the dominant market share, and higher net income.

We next consider the case in which firms face capacity adjustment lags and must therefore forecast industry demand and competitor responses. Table 2 shows the payoff matrices for each demand scenario; Figure 5 shows how the payoffs depend on the speed of product diffusion. Both show that the predictions of the standard neoclassical model break down when demand evolves rapidly in the presence of capacity adjustment lags and imperfect forecasting.

When the market dynamics are sufficiently slow, the firm’s demand forecasts and knowledge of its competitor’s capacity plans are reasonably accurate, and capacity closely tracks the required level. As in the perfect-capacity case, players have a temptation to play aggressive when their rival plays the conservative strategy (see the slow scenario in Table 2). However, for market dynamics faster than those given by a

---

Footnote 4: We use a real discount rate of 4% per year and simulate the model for 40 years. The results are robust to discount rates from 0 to at least 20% per year.

Table 1: Payoffs to Aggressive and Conservative Strategies in Each Industry Demand Scenario (See Figure 2): The Case of Instantaneous Capacity Adjustment (NPV of Cumulative Profits, Billion $)

<table>
<thead>
<tr>
<th>Industry demand scenario</th>
<th>Aggressive (A)</th>
<th>Conservative (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow ($\beta = 0.5$)</td>
<td>3.2, 3.2</td>
<td>5.1, 2.1</td>
</tr>
<tr>
<td></td>
<td>2.1, 5.1</td>
<td>3.8, 3.8</td>
</tr>
<tr>
<td>Medium ($\beta = 1$)</td>
<td>4.8, 4.8</td>
<td>7.3, 3.2</td>
</tr>
<tr>
<td></td>
<td>3.2, 7.3</td>
<td>5.7, 5.7</td>
</tr>
<tr>
<td>Fast ($\beta = 2$)</td>
<td>6.5, 6.5</td>
<td>9.4, 4.8</td>
</tr>
<tr>
<td></td>
<td>4.8, 9.4</td>
<td>7.6, 7.6</td>
</tr>
</tbody>
</table>

Figure 3: Firm Payoffs as They Depend on the Speed of the Product Lifecycle: No Capacity Acquisition Lag

Note. A | [A, C] indicates the profits of the aggressive (A) firm assuming the other firm plays the conservative strategy, C | [A, C] give the profits of the conservative player given that the other firm plays the aggressive strategy, etc.
critical value of the word-of-mouth parameter, $\beta^{\text{CRIT}} \approx 1.3$, the conservative strategy dominates. For example, in the fast scenario, industry and firm profits are maximized when both firms play conservatively: Neither firm has any incentive to defect (Table 2). Indeed, in contrast to the predictions of models based upon standard neoclassical assumptions, the faster the market grows, the worse the performance of the aggressive strategy.

At first sight this result might seem surprising, because faster diffusion brings faster sales, increasing the returns to playing aggressively in the instantaneous adjustment/perfect forecast case. Closer examination, however, illustrates the ways in which rapid

Table 2 Payoffs to Aggressive and Conservative Strategies in Each Industry Demand Scenario (Figure 2)

<table>
<thead>
<tr>
<th>Industry demand scenario</th>
<th>Capacity adjusts with a lag</th>
<th>Aggressive (A)</th>
<th>Conservative (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow ($\beta = 0.5$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>$-7.0$</td>
<td>$4.8$</td>
<td>$0.9$</td>
</tr>
<tr>
<td>C</td>
<td>$0.9$</td>
<td>$4.8$</td>
<td>$3.5$</td>
</tr>
<tr>
<td>Medium ($\beta = 1$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>$-11.1$</td>
<td>$5.2$</td>
<td>$1.0$</td>
</tr>
<tr>
<td>C</td>
<td>$1.0$</td>
<td>$5.2$</td>
<td>$4.4$</td>
</tr>
<tr>
<td>Fast ($\beta = 2$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>$-19.7$</td>
<td>$-1.7$</td>
<td>$1.9$</td>
</tr>
<tr>
<td>C</td>
<td>$0.2$</td>
<td>$-1.7$</td>
<td>$1.9$</td>
</tr>
</tbody>
</table>

Note. Results when capacity adjusts with a lag (NPV of cumulative profits, billion $).
diffusion compounds the errors caused by slow capacity adjustment and imperfect forecasting to create serious problems for the aggressive firms.

Figure 6 shows the dynamics for the case in which capacity adjusts with a lag and the market evolves rapidly. As in the perfect-capacity case (Figure 4), one firm plays aggressively while the other cedes share. Also as in the perfect-capacity case, the aggressive firm immediately cuts price to gain market share. In the case with the capacity lag, the aggressive firm also sets target capacity to 80% of its forecast of industry demand. Due to the delays in perceiving industry orders and capacity acquisition, actual capacity lags behind orders, and both firms quickly reach full utilization. Capacity remains inadequate until about year 1.5. During this time, excess backlogs accumulate and customers are forced to wait longer than normal for delivery. The capacity crunch causes both firms to boost prices above normal levels, although the aggressive firm continues to price below the conservative firm. Such transient shortages and price bubbles are often observed during the growth phases of successful products, for example radio, black and white television, color television (Dino 1985), and more recently, DRAM, iPods, Harley-Davidson motorcycles, and Playstation 3 video games.

Demand continues to grow rapidly, albeit at a declining fractional rate. As industry order data are reported, both firms gradually adjust their demand forecasts. However, due to the adjustment lags, capacity begins to overshoot the required level, and utilization falls below normal. When industry orders peak

Note. Compare to Figure 4.
and decline, shortly before year 6, both firms find themselves with significant excess capacity. Excess capacity causes large losses both directly, as fixed costs remain high while sales fall, and indirectly, as excess capacity forces prices down. The aggressive firm suffers the most, because it has expanded capacity faster. As boom becomes bust, the aggressive firm finds utilization drops below 50%. The conservative firm also experiences excess capacity, but the magnitude and duration of the problem is smaller because it has been steadily giving up market share during the growth phase. Both firms experience excess capacity as the market saturates, but the aggressive firm loses far more than its conservative rival. Such capacity overshoot is widespread in maturing industries, and was frequently observed in Paich and Serman’s (1993) experimental product lifecycle task.

Note that the poor performance of the aggressive strategy when the market dynamics are rapid is not due to the failure of increasing returns to confer cost advantage on the aggressive firm. As in the perfect-capacity case, the aggressive strategy achieves its intended goal: Low prices and rapid expansion quickly give the aggressor a cost advantage, which steadily widens as the industry moves through its lifecycle. Indeed, at the end of the simulation, the aggressive firm has unit costs 57% less than its rival, a larger advantage than it enjoyed in the perfect-capacity case. Instead, the failure of the aggressive strategy arises from the interaction of the disequilibrium dynamics of the market and the boundedly rational heuristics each firm’s managers use to forecast demand, plan capacity, and set prices.

When capacity adjusts perfectly, firms have an incentive to exploit increasing returns by playing the aggressive strategy, and faster market evolution increases the incentive to do so (Figure 3). However, when firms face a capacity adjustment lag, the excess capacity induced by forecast error and the underestimation of the competitor’s capacity plans increases with the speed of the product lifecycle. Eventually, the costs of excess capacity overwhelm the advantage conferred by increasing returns, and the aggressive strategy becomes inferior (Figure 5).

### 3.1 Sensitivity Analysis.

Before turning to conclusions we explore the sensitivity of the results to key assumptions (Table 3 and Table EC.2). Reducing the capacity-adjustment and information-reporting delays raise the value of $\beta^{\text{CRIT}}$, favoring the aggressive strategy, as one would expect. But reducing the strength of the learning curve raises the value of $\beta^{\text{CRIT}}$, so that the recommendation to “get big fast” is, paradoxically, more robust when increasing returns are relatively less powerful. This counter-intuitive result arises from the interaction of increasing returns, demand, and boundedly rational demand forecasting: The stronger the learning curve, the faster price falls as the market evolves. Rapidly falling prices accelerate the growth of demand, worsening the eventual capacity overshoot and increasing the losses arising from the aggressive strategy.

Sensitivity analysis of other model parameters (reported in the online supplement) shows that the critical value of the word-of-mouth parameter above which the aggressive strategy becomes inferior remains in the range from 2.0 to less than 0.5, corresponding to sales peaks from five to twenty years after product launch, well within the range documented for many real products (Paker 1994).

We have made a number of assumptions that reduce the attractiveness of aggressive GBF strategies. First, to the extent capacity can be used to make follow-on products, the costs of capacity overshoot will be mitigated. Second, we assume there are no economies of scope allowing related products to share in the benefits of learning. Third, we assume there is no growth in the underlying pool of potential customers; such growth would reduce the severity of the saturation peak.

On the other hand a number of our assumptions favor the aggressive strategy. We assume learning is perfectly appropriable, increasing the ability of firms to gain sustained cost advantage. Spillovers allow conservative firms to benefit from the production experience of larger rivals, dissipating the advantages aggressors pay so dearly to acquire (Ghamawat and Spence 1985). We assume that production adjusts instantaneously at constant marginal cost (until capacity utilization reaches 100%), and that capacity can be adjusted smoothly and continuously. Lumpy capacity that can only be added in large chunks relative to demand would worsen the excess capacity incurred after market saturation. Further, capacity adjusts with an average lag of just one year, less than the typical lags estimated in the literature. There are no capacity adjustment costs or exit costs. The capacity acquisition lag is symmetric; faster decommissioning would favor the aggressive strategy, but if excess capacity is sold to other players at low prices because its costs are
sunk, industry capacity remains high even as individual firms shed their own surplus, depressing prices and profits. We omit balance-sheet considerations and thus the risk of bankruptcy: Aggressive firms that ultimately do well in the simulation may not survive the losses of the transition from boom to bust (Oliva et al. 2003), a common phenomenon in the collapse of the dot-com bubble. We assume that the competitor’s planned capacity target is fully known with only a short delay, whereas in reality the determination of competitor plans is difficult and time consuming. The online supplement considers the more realistic case in which competitor capacity plans are imperfectly known. The results show the GBF strategy becomes inferior with much milder forecasting errors. The information on which the firm bases its decisions is free of noise, measurement error, bias, or other distortion. We assume firms can base their forecasts on industry orders, reported with only a one-quarter year lag, when in most industries order data are unavailable and firms must rely on estimates of industry revenues or shipments, introducing an additional delay and also confounding demand (orders) with capacity (which may constrain shipments below the rate of incoming orders during periods of rapid demand growth).

Relaxing any of these assumptions causes the aggressive strategy to be dominated by the conservative strategy at lower rates of market growth and for less-durable products, strengthening our results (see the online supplement for sensitivity analysis).

4. Discussion and Conclusions

Conventional models have been a potent source of insight into many of the dynamics surrounding increasing returns. Our results, however, suggest limitations of these models with implications both for practicing managers and for the choice of tools appropriate for the study of strategic behavior. On the practice front, existing theory recommends aggressive preemption in the presence of strong learning curves, network externalities, and other positive feedbacks that confer cumulative positional advantage—provided the firm can move first, the gains from learning and other sources of increasing returns are privately appropriable, and uncertainty is modest. Several researchers have also identified additional circumstances in which aggressive preemption may not be an optimal strategy. Porter (1980), for example, outlined a long list of factors that seemed likely to make aggressive preemption dangerous, and his work has been echoed by scholars who have highlighted the ways in which, for example, lumpy capacity and private information can lead to a winner’s curse type of dynamic and can cause aggressive strategies to be suboptimal even in the presence of increasing returns (Ghemawat 1987, Lieberman and Montgomery 1998).

Our results reinforce the risks of aggressive preemption in situations of increasing returns, and offer some new insights into the dangers of taking the results of models assuming perfect information and full rationality as blueprints for practical action. Our results suggest that aggressive preemption can be suboptimal even in situations where capacity can be adjusted continuously and heterogeneity in private information is absent. We find that realistic procedures for demand forecasting and the assessment of competitor actions can interact with delays in the reactions of firms to changes in demand to cause substantial capacity overshoot when demand growth slows. The relevant delays not only include lags in adjusting firm resources, but also delays in gathering market data, in carrying out competitive intelligence, and in adjusting forecasts.

Clearly, better forecasting would favor the aggressive strategy, but here the evidence is not encouraging. In Paich and Sterman’s (1993) product-lifecycle experiment, subjects consistently failed to forecast the sales peak, leading to excess capacity and large losses similar to those simulated here—even after extensive experience with the task. Outside the laboratory, a wide range of new product-diffusion models have been developed that, in principle, allow forecasting of the sales peak (Parker 1994, Mahajan et al. 1990, Armstrong 2001). In practice, diffusion models often miss the turning point, because, as Mahajan et al. (1990, p. 9) comment, “by the time sufficient observations have developed for reliable estimation, it is too late to use the estimates for forecasting purposes.” Rao (1985) examined the ability of ten popular models to predict sales of typical durable goods. Mean absolute forecast errors averaged more than 40%, and the extrapolative models generally outperformed diffusion models.

These results suggest a firm electing to pursue a GBF strategy must devote significant effort to understanding the dynamics of market demand so that it is not caught unprepared by market saturation. Experience and experimental studies suggest that this is both hard medicine to take and difficult to carry out successfully. Alternatively, when the risk of capacity overshoot is high, firms should consider conservative strategies even in the presence of increasing returns, allowing less-sensible rivals to play the aggressive strategy, then buying these rivals at distress prices when they fail during the transition from boom to bust.

On the methodological front, our results suggest that the presumption that the rationality assumptions of game theory are “good enough” to provide robust frameworks for action is not always appropriate.
Equilibrium models of rational agents provide an approximation to the behavior of real people in real markets. Our results show that when the system dynamics are slow enough, delays in information acquisition, decision making, and implementation are short enough, and the environment is simple enough relative to managers’ cognitive capabilities, the traditional assumptions of the mainstream game-theoretic literature are, indeed, “good enough.” The results also show, however, that the validity of these assumptions should not be assumed. Rather, determining what constitutes “slow,” “short,” and “simple” in models of competitive strategy requires the development of models that capture the disequilibrium dynamics resulting from realistic adjustment rigidities and behavioral decision-making processes.

In situations with high dynamic complexity, boundedly rational people can and do behave differently from their neoclassical counterparts. The case of increasing returns in a dynamic market shows that these differences matter and that their impact can be examined rigorously. The results highlight the ways in which small departures from full rationality can change optimal strategy, and lead to results that are—by the standards of neoclassical models—quite counterintuitive. Our finding that, in the presence of high dynamic complexity, increasing the strength of increasing returns significantly reduces the odds that aggressive preemption will be a dominant strategy is an example of one such result. Although further work is required to explore the relationship between behavior and dynamic complexity beyond the two-firm case under increasing returns, we speculate that relaxing the assumptions of full information and perfect rationality may lead to similar differences in a variety of other contexts. The increasing complexity and turbulence of the world suggests that more and more industries are likely to face extended periods of disequilibrium. We believe dynamic behavioral models are thus likely to have an increasingly important role to play in our understanding of strategic dynamics.

5. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

Financial support was provided by the Project on Innovation in Markets and Organizations at the MIT Sloan School of Management. The authors thank Bruno Cassiman, Pankaj Ghemawat, Bob Gibbons, Nelson Repenning, and the referees for helpful suggestions.

References


