The Effects of Autoscaling in Cloud Computing on Entrepreneurship *

Amir Fazli Amin Sayedi Jeffrey D. Shulman

September 21, 2016

*University of Washington, Foster School of Business. All authors contributed equally to this manuscript. Amir Fazli (fazli@uw.edu) is a Ph.D. Student in Marketing, Amin Sayedi (aminsa@uw.edu) is an Assistant Professor of Marketing and Jeffrey D. Shulman (jshulman@uw.edu) is an Associate Professor of Marketing and Michael G. Foster Faculty Fellow, all at Foster School of Business, University of Washington.

The Effects of Autoscaling in Cloud Computing on Entrepreneurship

Abstract

Technology startups often rely on computational resources to serve their customers, though rarely is the number of customers they will serve known at the time of market entry. Today, many of these computations are run using cloud computing. A recent innovation in cloud computing known as autoscaling allows companies to automatically scale their computational load up or down as needed. We build a game theory model to examine how autoscaling will affect entrepreneurs' decisions to enter a new market and the resulting equilibrium prices, profitability, and consumer surplus. Prior to autoscaling, startups needed to set their computational capacity before realizing their computational demands. With autoscaling, a company can be assured of meeting demand and pay only for the demand that is realized. Though autoscaling decreases expenditures on unneeded computational resources and therefore should make market entry more attractive, we find this is not always the case. We highlight strategic forces that determine the equilibrium outcome. Our model identifies the likelihood of a startup's success in a market, differentiation among potential entrants in a market, and a product's potential value to consumers as key drivers of whether autoscaling increases or decreases market entry, prices, and consumer surplus.

1 Introduction

Consider entrepreneurs deciding whether to launch a web startup in a new market. Getting started requires an investment of time and money into research, development, legal, and other startup expenses prior to knowing whether the company will ever succeed. Furthermore, web and mobile-based startups rely on computational capacity to serve customers. Every interaction a customer has with an application such as a page load, data transfer, and object viewing requires computational resources. This is particularly relevant to companies launching a Software as a Service (SaaS), an industry estimated at $49 billion and expected to grow to $67 billion by 2018 (Columbus 2015). For any company relying on computational resources, cloud computing allows the company to outsource the server set-up and maintenance to a cloud provider.

More and more companies are adopting the cloud to handle their computational needs. Examples of companies using the cloud include big firms such as Netflix, Airbnb, Pinterest, Samsung, Expedia, and Spotify as well as many small businesses and startups (Gaudin 2015; Bort 2015). In fact, end user spending on cloud services in 2015 was estimated to be above $100 billion (Flood 2013) with an expected annual growth of 44% in workloads (Ray 2013).

While cloud computing allows startups to outsource their computational costs, startups who ran their tasks in the cloud often needed to decide, and pre-commit to, their computational capacity at the time of purchasing the cloud service. As such, due to the unpredictable nature of startups' demand, the pre-purchased capacity may be excessive or insufficient for the traffic. For example, BeFunky, an online photo editing startup, was featured on a popular social media site three weeks after launch and saw 30,000 visitors in three hours, crashing their servers (Nickelsburg 2016). Such events happen regularly enough that there is a term for it: *the Slashdot effect,* which as Klems, Nimis, and Tai (2008) describe occurs when a Web startup company is featured on a popular network, resulting in a significant increase in traffic load and causing the startup's servers to slow down or crash. Such a problem can be quite costly as an Aberdeen study found that "a 1-second delay in page load time can result in a 7% loss in conversion and a 16% decrease in customer satisfaction" (Poepsel 2008). Kissmetrics, an analytics company, reports that 1 in 4 people abandon a page if it takes longer than 4 seconds to load (Work 2011). Though capacity can later be increased, the missed demand can be costly. As Amazon CEO Jeff Bezos describes, startups face a serious challenge when choosing computational capacity:[1]

> "And you do face this issue (demand uncertainty) whenever you have a startup company. You want to be prepared for lightning to strike because if you're not, that generates a big regret. If lightning strikes and you weren't ready for it, that's kind of hard to live with. At the same time, you don't want to prepare your physical infrastructure to hubris levels either in the case that lightning doesn't strike."

To address this challenge, cloud providers such as Amazon, Microsoft, and Google have begun to offer *autoscaling*, a feature that allows startups to scale their computational capacity up or down automatically in real time. Using autoscaling, startups can maintain application availability

[1] https://animoto.com/blog/news/company/amazon-com-ceo-jeff-bezos-on-animoto/, accessed September 2016.

and scale their computational capacity for serving consumers without having to make capacity pre-commitments.

For companies, autoscaling means having just the right number of servers required for meeting the demand at any point in time, which can provide an attractive solution to handling uncertain demand. Autoscaling is offered with no additional fees and has been celebrated as one of the most beneficial features of cloud computing. Users of autoscaling such as AdRoll and Netflix find that when a new customer comes on board, they can handle the additional traffic instantly.[2] As Mikko Peltola, the Operations Lead at Rovio, noted regarding the benefits of autoscaling, "We can scale up as the number of players go up ... so we can automatically increase the processing power for our servers."[3] The Chief Technology Officer for Cloud at General Electric has mentioned this feature as one of the main reasons for the company's move to the cloud, stating "Running inside a public cloud environment, you're able to consume unlimited capacity as needed" (Weaver 2015). Anecdotally, some entrepreneurs such as Animoto CEO, Brad Jefferson, who used Amazon Web Services (AWS), see the scaling it offers as a game changer for their startup: "We simply could not have launched Animoto.com and our professional video rendering platform at our current scale without massive CapEx and a lot of VC funding. The viral spike in Animoto video creations we experienced this week would have been disastrous without AWS."[4]

In this paper, we develop an analytical model to examine how the emergence of autoscaling in cloud computing will affect entrepreneurs' decisions regarding market entry and prices. In particular, despite popular belief, we identify conditions for when autoscaling negatively affects entrepreneurship. In other words, fewer startups will enter in certain markets due to the advent of autoscaling. Autoscaling has several properties that make it unique from some previously explored areas in marketing and operations. In particular, autoscaling:

- removes a capacity decision that otherwise has to be made prior to pricing;
- converts computational capacity costs from fixed costs to variable costs at the time of pricing;
- allows capacity to be set after the uncertainties regarding consumers' level of interest and competitors' strategies are resolved; however, autoscaling still

[2]See https://aws.amazon.com/solutions/case-studies/adroll/
[3]See https://aws.amazon.com/solutions/case-studies/rovio/
[4]https://animoto.com/blog/news/company/amazon-com-ceo-jeff-bezos-on-animoto/, accessed September 2016.

- preserves the uncertainty that exists at the time of making the entry decision.

Given these properties, the effects of autoscaling on company strategies cannot be addressed by prior research on capacity choices and demand uncertainty. In fact, *our model and predictions diverge from prior literature substantively.*

We uniquely incorporate these properties of autoscaling into a game theory model in which two horizontally differentiated startups have the option to enter a market upon incurring an entry cost. We compare a model in which startups choose computational capacity to a model in which startups can choose to adopt autoscaling in cloud computing. The model is constructed to address the following research questions:

1. How does autoscaling affect entrepreneurs' profits?

2. How does autoscaling affect entrepreneurs' pricing strategies?

3. How does autoscaling affect market entry decisions?

4. How does autoscaling affect consumer surplus?

We explore the roles of several important market factors in determining the answer to each of the research questions. First, we model the startups' ex ante likelihood of a successful venture. As Griffith (2014) suggests, the value that a startup brings to the market is unknown before entry. In some markets, consumer needs are well known and established, therefore yielding a higher likelihood of successfully creating a product that matches consumer needs. For other markets, the consumer needs are less understood and there is a lower likelihood of a successful venture. We show that the likelihood of a successful venture plays a critical role in determining how autoscaling affects equilibrium strategies and profits.

Secondly, we model the degree of potential success a venture may enjoy. For some markets, the consumer need addressed by the startup may have greater value to the customer than other markets. A greater value to the customer can translate to a highly successful venture in terms of the ability to charge higher prices. We examine how the magnitude of a successful venture (in terms of how much a consumer would pay to satisfy the need if the startup successfully addressed it) will affect pricing decisions, market entry, and consumer surplus.

In answering the first research question, we find that autoscaling can increase or even decrease entrepreneurs' expected profits. We identify strategic consequences of autoscaling and the conditions that lead to each possibility. In particular, when the probability of success is sufficiently low, autoscaling increases the entrepreneurs' expected profits. However, autoscaling may also create a prisoner's dilemma situation where startups choose autoscaling, but autoscaling lowers their equilibrium profits. In particular, when entry costs are sufficiently small and the probability of success is moderately high, startups adopt autoscaling in equilibrium; however, their equilibrium profits would be higher if autoscaling was not available.

In addressing the second research question, we find that autoscaling may increase or decrease average prices charged by competing startups. Existing economic theory would suggest that removing the capacity decision prior to pricing would result in a shift from a Cournot game to a Bertrand game, thereby decreasing prices (e.g., Kreps and Scheinkman 1983). However, our model shows that when entry costs are low enough such that both startups enter the market, autoscaling increases the average prices set by each startup if the probability of a successful venture is not too high. On the other hand, if the probability of a successful venture is sufficiently high, then autoscaling decreases the average prices set by each startup.

With regard to the third research question, we find that autoscaling can sometimes decrease market entry. Though we confirm common intuition that the likelihood of a market being served by at least one startup is improved with autoscaling, we find that, under certain conditions, entry by multiple startups will not occur because of autoscaling. The counter-intuitive result occurs when entry costs are moderately high and there is a high probability that entrants will have a successful venture.

Finally, in addressing the fourth research question, we show that autoscaling may increase or decrease expected consumer surplus depending on the likelihood of a successful venture and entry costs. Given the fact that autoscaling guarantees companies have the capacity to serve consumers in case of high demand, thereby resolving issues such as the Slashdot effect, one might expect that consumers benefit from autoscaling. However, our model shows that when entry costs are low enough such that both startups enter the market, autoscaling decreases expected consumer surplus if and only if the probability of a successful venture is moderate.

The findings of this study provide implications for various players in new markets, including startups, cloud providers and consumers. Our analysis informs startup managers on how autoscaling affects competitive dynamics in pricing and entry. The results suggest that a startup manager should consider not only the positive direct effect of autoscaling in reducing costs, but also the negative strategic effect in altering the nature of competition. By evaluating the probability of success, the value to the consumers, and entry costs, managers can use the findings from this study to determine whether autoscaling increases or decreases the likelihood of monopoly power over the new market. Our findings also inform cloud providers about how autoscaling affects not only the number of firms using the cloud, but also the number of servers each of those firms purchases. Our model provides insights for consumers on how autoscaling affects the prices charged in the market, showing that for high probabilities of success average prices decrease with autoscaling and for lower probabilities of success they increase with autoscaling. We also find conditions for which autoscaling will decrease or increase consumer surplus, which can be used for consumer surplus maximizing policy design. To the best of our knowledge, this paper is the first to study the marketing aspects of cloud computing, and how it can affect prices and market entry. With the growing trend of adopting the cloud by firms, cloud computing is expected to become a major part of any business and this provides the field of marketing with a variety of related new topics to explore.

In addition to contributions to practice, our work contributes to economic theory regarding capacity commitments. Our benchmark model uniquely solves a capacity choice game with demand uncertainty and horizontal differentiation between sellers. Contrary to the previous literature where capacity commitments lead to higher prices, we show that under demand uncertainty, capacity commitments (relative to autoscaling) can intensify the competition and cause lower equilibrium prices. We also uniquely study the effects of removing capacity commitments made under demand uncertainty (via autoscaling) on firms' market entry decisions.

The rest of this paper is organized in the following order. In Section 2, we review the literature related to our research problem. In Section 3, we introduce the model. In Section 4, we present the analysis of the model and derive the results. Finally, the discussion of our findings is presented in Section 5.

2 Literature Review

Academic research on cloud computing is still relatively new and most of the work done on this topic focuses on technological issues of the cloud (e.g., Yang and Tate 2012). The few existing business and economics studies of cloud computing have mainly offered conceptual theories and evidence from surveys and specific cases (e.g., Leavitt 2009; Walker 2009; Gupta, Seetharaman, and Raj 2013). Sultan (2011) suggests that cloud computing can benefit small companies due to its flexible cost structure and scalability. Regarding the ability to autoscale, Armbrust et al. (2009) suggest that elasticity in the cloud shifts the risk of misestimating the workload from the user to the cloud provider. Regarding market entry, Marston, Li, Bandyopadhyay, Zhang, and Ghalsasi (2011) conceptually argue that cloud computing can reduce costs of entry and decrease time to market by eliminating upfront investments. Our paper is the first to model autoscaling in the cloud and show how its effects on market entry are not always positive, and depend on market characteristics.

In addition to cloud computing, our research is related to a number of topics in the literature. In particular, previous research shows uncertainty in demand plays an important role in capacity and production decisions. Che, Narasimhan, and Padmanabhan (2010) find how demand uncertainty and consumer heterogeneity affect whether a firm optimally adopts a make-to-stock system, a backorder system, or a combination of both. Ferguson and Koenigsberg (2007) examine how a firm should sell its deteriorating perishable inventory and compare this option to discarding the previously unsold stock. Desai, Koenigsberg, and Purohit (2007) find the optimal inventory with demand uncertainty as a function of a product's durability. Desai, Koenigsberg, and Purohit (2010) find a strategic reason for retailers to carry inventory larger than the expected sales in both high and low demand states. Biyalogorsky and Koenigsberg (2014) consider product introductions by a monopolist facing uncertainty about consumer valuations and find whether the firm offers multiple products simultaneously or sequentially. Anupindi and Jiang (2008) consider capacity investments by competing firms with demand uncertainty and look at the effect of flexibility to postpone production until demand is realized. They find the flexibility increases capacity investment and profitability, whereas we find when autoscaling may decrease equilibrium computational expenditures and when it may decrease profitability. Relative to these models, autoscaling affects more than the level of uncertainty in demand at the time capacity is set, it eliminates the capacity decision entirely and converts the capacity costs from being sunk at the pricing decision to being variable costs. Moreover, autoscaling does not eliminate uncertainty about firms' success at the entry stage.

Our examination of how cloud computing with autoscaling affects a startup's entry decision also relates to the literature on market entry. A body of literature looks at the timing of entry and how an incumbent can deter entry (e.g., Spence 1977; Narasimhan and Zhang 2000; Joshi, Reibstein, and Zhang 2009; Milgrom and Roberts 1982; and Ofek and Turut 2013). In contrast, our model examines simultaneous entry decisions by startups. Our model adds to entry literature by jointly considering both entry and capacity decisions, such that each firm's decision to enter depends on the expected future capacity of both firms and whether this capacity will be chosen ex ante or autoscaled to demand.

We compare autoscaling with cases where firms commit to their computational capacity before pricing and realizing demand. This relates to other papers examining capacity commitments. Kreps and Scheinkman (1983) find that a Bertrand pricing game is equivalent to a Cournot game when capacity is chosen prior to pricing. Reynolds and Wilson (2000) extend this model to include uncertainty about market size and find there is no symmetric, pure-strategy equilibrium capacity choice when there is significant demand variation. Nasser and Turcic (2015) find symmetric horizontally differentiated firms use asymmetric strategies on whether to commit to capacity or not. This is consistent with our finding in Lemma 2. However, since they do not allow for demand uncertainty, capacity commitments always alleviate competition in their model, whereas, in ours, capacity commitments sometimes intensify competition. Furthermore, at least one firm commits to capacity in any equilibrium in their model, whereas, in our model, both firms may use autoscaling. In asymmetric games, Daughety (1990) finds the Stackelberg leader will commit to a greater quantity, whereas Shulman (2014) finds an unauthorized seller procuring diverted units from authorized retailers will unilaterally choose a lesser quantity than it could profitably sell. Swinney, Cachon, and Netessine (2011) examine the optimal timing of capacity investment in a model in which market price is given by a demand curve and firms can choose to set capacity early at one marginal cost of capacity or after demand is realized at a different cost of capacity. Van Mieghem and Dada (1999) allow firms to choose the time of their pricing decisions and find that postponing pricing until after demand is realized makes the capacity decision less sensitive to demand uncertainty. Our research expands this literature by considering demand uncertainty and market entry decisions in a horizontally differentiated market with and without capacity commitments. In contrast to the previous literature where capacity commitments always alleviate competition, we show that, depending on the level of demand uncertainty, capacity commitments may indeed intensify the competition. Furthermore, in our model, firms decide whether to adopt autoscaling or to pre-commit to capacity; we find that firms do not always follow symmetric strategies in regards to adoption of autoscaling. Finally, we uniquely explore how entry decisions are affected by existence of capacity commitments under demand uncertainty.

Autoscaling in cloud computing also has the effect of converting up-front capacity costs to variable costs that change with demand. Prior research examining the effect of converting fixed to variable costs via outsourcing (e.g., Shy and Stenbacka 2003; Buehler and Haucap 2006; Chen and Wu 2013) have found that prices and profitability rise with this conversion. Relative to these models, demand is uncertain in our model and our up-front capacity cost is endogenous since firms can choose their capacity. We also find, in contrast, that average prices can fall when cloud computing with autoscaling is used even in conditions for which entry is unaffected.

In summary, our paper uniquely compares market entry, pricing, and profitability between computing resources requiring capacity pre-commitments and cloud computing with autoscaling. The advent of cloud computing with autoscaling has several properties: 1. Autoscaling allows for capacity decisions to be made after the demand uncertainty is resolved. 2. Autoscaling converts the fixed cost of computational resources that is sunk prior to the firms' pricing decision into variable costs that are directly affected by the pricing decision. 3. Autoscaling makes it such that capacity and pricing decisions are made simultaneously rather than sequentially. However, autoscaling does not affect the level of uncertainty at the time a firm makes its entry decision. Though previous research has separately examined demand uncertainty, the timing of capacity and pricing decisions, or the conversion of fixed costs to variable costs, our paper is unique in its comprehensive examination of the effect of autoscaling. In particular, our paper is the first to solve for entry, capacity and pricing decisions in a model of horizontally differentiated firms with demand uncertainty and to compare this equilibrium to the entry and pricing decisions of horizontally differentiated firms who make entry decisions with demand uncertainty but whose capacity can autoscale to the demand realized upon setting prices.

3 Model

We consider two symmetric startups who could potentially enter a particular web or mobile application market. Following the convention and for ease of exposition, we subsequently refer to these startups as firms. To enter the market, the firms would incur a fixed entry cost, F. This cost includes start-up expenses such as legal, research and development, and human capital investments. To model post-entry competition, we adopt a discrete horizontal differentiation model (e.g., Narasimhan 1988; Iyer, Soberman, and Villas-Boas 2005; Zhang and Katona 2012; Zhou, Mela, and Amaldoss 2015) with three consumer segments, each consumer demanding at most one unit of the product. Upon entry, each Firm i will find a segment of consumers, Segment i with $i \in \{1, 2\}$, who will buy from Firm i if and only if the price p_i is below their reservation value v_i and who will derive zero value from the competitor's product. This captures the reality that consumers vary in their taste preferences regardless of firm entry and that firms have idiosyncratic differences that will allow them to serve these tastes differently from each other upon successful entry. The size of each Segment i is given by $\alpha < 1/2$ for $i \in \{1, 2\}$. The remaining $1 - 2\alpha$ consumers, Segment 3, are indifferent between firms and will prefer to buy from the firm with the lowest price. The parameter α can be interpreted as the extent to which consumers vary in their taste preferences.

In the absence of autoscaling, the timing of the game is as follows:

Stage 1: Firms each make a simultaneous observable decision of whether or not to enter the market and thereby incur the cost F. We allow for uncertainty in whether a firm will find the venture successful in terms of whether v_i is high or low. We assume the ex ante probability of a firm finding success in this market is γ, which is common knowledge. In other words, if Firm i enters the market, v_i is an i.i.d. draw from a binary distribution in which $v_i = V$ with probability γ and $v_i = 0$ with probability $1 - \gamma$. This assumption reflects the idea that the value provided to customers is unclear for potential entrants. As Lilien and Yoon (1990) argue, the fit between market requirements and the offering of the new entrant is highly unpredictable and is critical to the success of the entrant. In a survey of 101 startups, it was reported that the number one reason for the failure of a startup is the lack of market need for the offered product (Griffith 2014), suggesting that the value created for customers is unknown to many startups before entry.

To remark on the structure of demand and uncertainty, notice that our model set up has several desirable properties. In particular, it allows a firm to be uncertain about the size of the potential market and the effect of its price on realized demand: the firm may find itself a monopolist, the firm may find itself with very low demand (normalized to zero), or the firm may find itself competing head-to-head. Though one can explore alternative model specifications to capture these same properties, the current specification allows for tractability while uncovering a novel mechanism.

Stage 2: Firms that enter make a simultaneous observable decision of computational capacity. Firms incur a computational cost, c, for each consumer they serve and choose their computational capacity k_i to maximize expected profit.

Stage 3: The reservation value for each Firm i, v_i, becomes common knowledge and each firm in the market simultaneously chooses p_i to maximize profit.

Stage 4: Demand is realized. In the case a firm experiences demand greater than its computational capacity, we assume an efficient rationing rule (see Tirole 1988, p. 213) in which demand from Segments 1 and 2 is satisfied prior to demand from Segment 3. Residual demand from Segment 3 is allocated to the competing firm, provided it has available capacity.

When autoscaling is possible, we allow firms to decide whether to use autoscaling or choose a computational capacity in Stage 2. If a firm chooses autoscaling, it incurs the computational cost c only on each unit of realized demand. In practice, changing capacity decisions in the absence of autoscaling takes at least a few hours, and in some cases days, before coming into effect on the cloud servers. The Befunky example, the Slashdot effect, and the "lightning strike" analogy by Amazon's CEO, discussed in the introduction, highlight the fact that demand often changes faster than what firms can respond to in terms of computational capacity. Our assumption that capacity decision is made before the demand is realized captures this reality. However, our main results are robust to this assumption. In particular, even if firms can choose to adopt autoscaling after the demand is realized, our results in Propositions 2, 3, 4 and 5 still hold.

The timing of the game is depicted in Figures 1 and 2. A summary of notation is in Table 1.

4 Analysis

Our research objective is to identify how the advent of autoscaling affects equilibrium prices, profits, and market entry. To this end, we first examine equilibrium entry and prices in the situation

Symbol	Description
α	Size of each of Segments 1 and 2
k_i	Capacity chosen by Firm i
v_i	Reservation value consumers have for Firm i
V	The consumer reservation value in the high-value condition
γ	Probability that $v_i = V$
c	Cost per unit of computational capacity
p_i	Price chosen by Firm i
F	Cost of entry

Table 1: Summary of notation

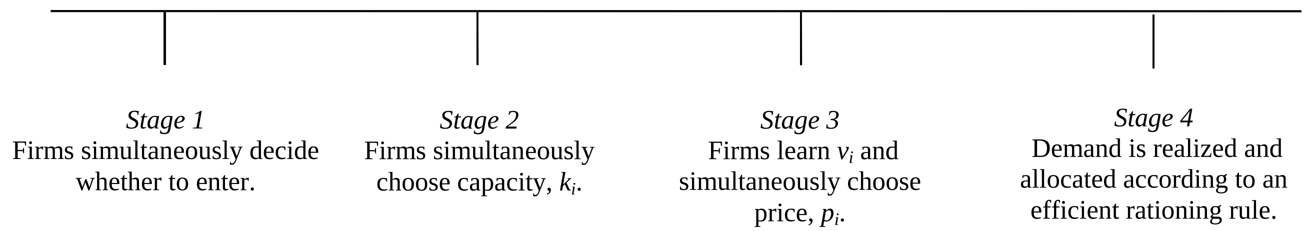

Figure 1: Sequence of events with no autoscaling

in which computational capacity must be determined prior to demand realization. We will subsequently characterize the equilibrium when autoscaling is available. We will conclude with a comparison across these possibilities.

4.1 No Autoscaling

We solve the model in which there is no autoscaling via backward induction beginning with the pricing subgame equilibrium. First suppose that $k_1 + k_2 > 1$. We want to calculate equilibrium prices of this game. Without loss of generality, assume that $k_2 \geq k_1$. Also, it is easy to see that

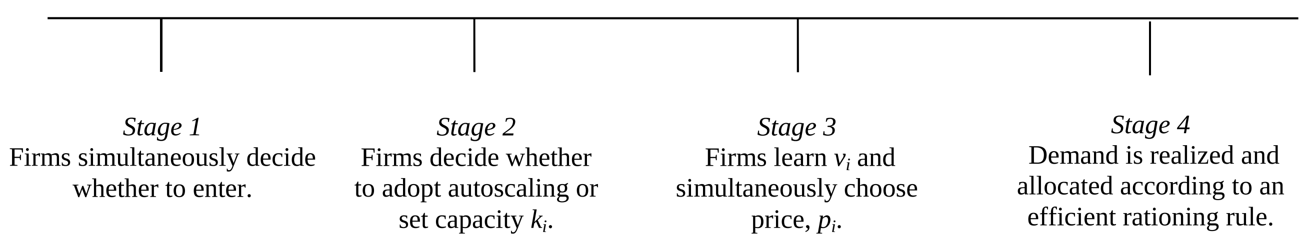

Figure 2: Sequence of events with autoscaling

firms never set their capacity to $k_i > 1-\alpha$ or $k_i < \alpha$; therefore, it is sufficient to consider the case where $k_i \in [\alpha, 1-\alpha]$ for $i \in \{1,2\}$.

We start by showing that this game does not have a pure strategy equilibrium. Assume for sake of contradiction that the firms use prices p_1 and p_2 in a pure strategy equilibrium. If $p_1 \neq p_2$, then the firm with a lower price can benefit from deviating by increasing its price to $\frac{p_1+p_2}{2}$. If $p_1 = p_2$, then Firm 2 can benefit from deviating by decreasing its price to $p_2 - \varepsilon$, for sufficiently small ε, to acquire more consumers from Segment 3. Therefore, a pure strategy equilibrium cannot exist.

Next, we find a mixed strategy equilibrium for this game. Mixed strategies can be interpreted as sales or promotions and are common in the marketing literature (e.g., Chen and Iyer 2002, Iyer, Soberman, and Villas-Boas 2005, Zhang and Katona 2012). Provided $k_1 \leq 1-\alpha$, Firm 2 can choose to *attack* with a price that clears its capacity or *retreat* with a price V that harvests the value from the $1-k_1$ consumers that Firm 1 cannot serve due to its capacity constraint. Let z be the price at which Firm 2 is indifferent between *attacking* to sell to k_2 consumers at price z and *retreating* to sell to $1-k_1$ consumers at price V. We have $z = \frac{V(1-k_1)}{k_2}$. Figures 3 and 4 demonstrate the different appeals of these two pricing strategies. The choice between *retreating* and *attacking* for each firm depends on the choice of the other firm. If Firm 1's price is high, it becomes easier for Firm 2 to attract the consumer segment that is in both firms' reach resulting in Firm 2 choosing to *attack*. On the other hand, if Firm 1's price is low, Firm 2 would prefer to *retreat* than to compete with Firm 1 over the overlapping consumers. In the equilibrium that we find, both firms use a mixed strategy with prices ranging from z to V. Suppose that $F_i(.)$ is the cumulative distribution function of price set by Firm i and $F_j(.)$ is the cumulative distribution function of price set by competing firm j. The profit of Firm i earned by setting price x, excluding the sunk cost of capacity, is

$$\pi_i(x) = F_j(x)(1-k_j)x + (1-F_j(x))k_i x.$$

Using equilibrium conditions, we know that the derivative of this function must be zero for $x \in (z, V)$. Therefore, we have

$$-x(k_i + k_j - 1)F'_j(x) - F_j(x)(k_i + k_j - 1) + k_i = 0.$$

The solution to this differential equation is

$$F_j(x) = \frac{k_i}{k_i + k_j - 1} + \frac{C_j}{x}$$

where constant C_j is determined by the boundary conditions. As for the boundary conditions, we use $F_1(V) = 1$. Therefore, we get

$$F_1(x) = \begin{cases} 0 & \text{if } x < z \\ \frac{(k_1-1)V+k_2x}{x(k_1+k_2-1)} & \text{if } z \le x < V \\ 1 & \text{if } x \ge V \end{cases} \tag{1}$$

This implies that Firm 1 mixes on prices between z and V such that Firm 2 is indifferent between using any two prices in this range. Furthermore, given $F_1(.)$, Firm 2 strictly prefers any price in $[z, V]$ to any price outside this interval. To have an equilibrium, the strategy of Firm 2 should be such that Firm 1's strategy is not suboptimal. In other words, Firm 1 should be indifferent between any two prices in $[z, V]$, and should weakly prefer any price in $[z, V]$ to any price outside this interval. Therefore, we have to use boundary condition $F_2(z) = 0$ to make sure that (1) Firm 1's indifference condition is satisfied in $[z, V]$, and (2) Firm 2 does not set the price to lower than z, as we already know from $F_1(.)$ that such prices are suboptimal for Firm 2. As such, we get

$$F_2(x) = \begin{cases} 0 & \text{if } x < z \\ \frac{k_1((k_1-1)V+k_2x)}{k_2x(k_1+k_2-1)} & \text{if } z \le x < V \\ 1 & \text{if } x \ge V \end{cases} \tag{2}$$

Note that $F_2(x)$ is discontinuous at $x = V$, and jumps from $\frac{k_1}{k_2}$ to 1. This implies that Firm 2 uses price V with probability $1 - \frac{k_1}{k_2}$. In other words, $f_2(V) = (1 - \frac{k_1}{k_2})\delta(0)$, where $f_2(.)$ is the probability density function for price of Firm 2 and $\delta(.)$ is Dirac delta function.[5] [6]

Given $F_i(.)$, we can calculate the expected profit of each firm in this mixed strategy equilibrium.

[5] See Hassani (2009), pp 139-170.

[6] One might wonder if the probability density $1 - \frac{k_1}{k_2}$ allocated to price V by Firm 2 could be instead allocated to price z. The answer is that it cannot. While such strategy would still keep Firm 1 indifferent between any two prices in $[z, V]$, it would make price $z - \varepsilon$ (for sufficiently small ε) a strictly better strategy for Firm 1, which violates equilibrium conditions.

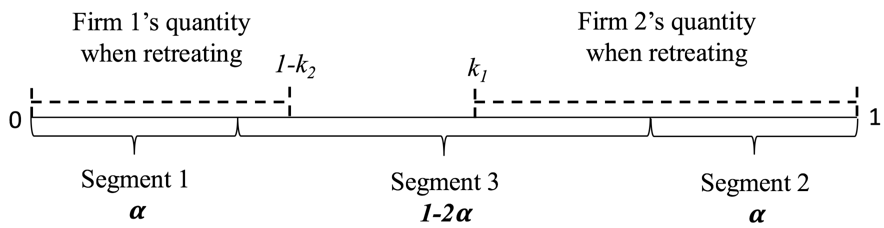

Figure 3: How k_1 and k_2 affect firm incentives to adopt a *retreating* price

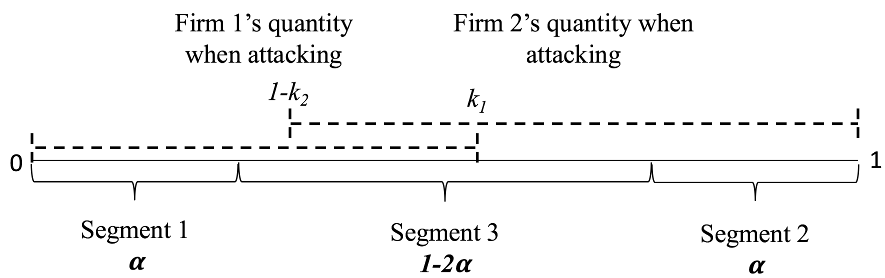

Figure 4: How k_1 and k_2 affect firm incentives to adopt an *attacking* price

Excluding the sunk cost of capacity, we have

$$\pi_1 = \frac{(1-k_1)k_1 V}{k_2} \quad \text{and} \quad \pi_2 = V(1-k_1).$$

Note that the higher capacity firm, Firm 2, earns a profit equal to the profit it would have made if it had chosen a *retreating* strategy while pricing at V; as such, the expected profit of Firm 2 is independent of its capacity k_2. On the other hand, the lower capacity firm, Firm 1, earns more than what it would have earned if *retreating* was chosen, since $k_1 < k_2$ requires $\frac{(1-k_1)k_1 V}{k_2} > V(1-k_2)$. As expected, after excluding sunk costs, the higher capacity firm makes a higher profit than the lower capacity firm.

Note that the mixed strategy pricing equilibrium bears some resemblance to Chen and Iyer (2002) who find in a model of customized pricing the ratio of profits is equal to the ratio of consumer addressability. In our model, the profit ratio is equal to the ratio of capacities. However, the model in Chen and Iyer (2002) is conceptually very different from ours. In particular, the overlap between the customers of the two firms is always non-zero in Chen and Iyer (2002), whereas in our model the overlap is non-zero only if the sum of the capacities is larger than the market size, i.e., $k_1 + k_2 > 1$. Furthermore, even though the ratio of profits is the same in both papers, the actual profit functions are very different. For example, as mentioned above, and in contrast to Chen and Iyer (2002), the profit of the firm with larger capacity does not depend on its own capacity in our model.

Now suppose $k_1 + k_2 \leq 1$. This implies that each firm that enters the market can sell to its capacity without competing with the other firm for consumers in Segment 3. As such, each firm that successfully enters the market can charge $p_i = V$ and sell k_i units for profit $(V - c)k_i$. Increasing the price will result in zero sales and profit, decreasing the price will still sell k_i units but at lower revenue.

Next consider the capacity subgame equilibrium. The capacity decision is made in anticipation of the possible combinations of values for v_1 and v_2. If both firms find success (i.e., $v_1 = v_2 = V$), then the profit depends on how k_i and k_j relate to each other and relate to α. The expected profit for Firm i depends on its capacity relative to the capacity of competing Firm j and can be written as follows:

$$E(\pi_i) = \begin{cases} \gamma V k_i - c k_i & \text{if } k_i \leq 1 - k_j \leq 1 - \alpha \\ \gamma(1-\gamma)V k_i + \gamma^2 \left(\frac{(1-k_i)k_i V}{k_j}\right) - c k_i & \text{if } 1 - k_j < k_i < k_j \leq 1 - \alpha \\ \gamma(1-\gamma)V k_i + \gamma^2 V(1-k_j) - c k_i & \text{if } 1 - k_i < k_j \leq k_i \leq 1 - \alpha \\ \gamma(1-\gamma)V(1-\alpha) + \gamma^2 V(1-k_j) - c k_i & \text{if } \alpha \leq k_j \leq 1 - \alpha < k_i \\ \gamma(1-\gamma)V(1-\alpha) + \gamma^2 V\alpha - c k_i & \text{if } k_i > 1 - \alpha \text{ and } k_j > 1 - \alpha \\ \gamma V(1-\alpha) - c k_i & \text{if } k_i > 1 - \alpha \text{ and } k_j < \alpha \end{cases}$$

where index j indicates the other firm. The equilibrium capacity choices are summarized in the following Proposition.

Proposition 1 *Suppose both firms enter the market initially. The equilibrium capacity choices depend on γ as follows:*

- *If there is a low probability of a successful venture (i.e., $\gamma < c/V$), then both firms choose $k_i = 0$ and earn zero profit.*

- *If there is a moderate probability of a highly successful venture (i.e., $\gamma > c/V$, $\gamma(1-\gamma)V - c > 0$, and $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$), then both firms choose $k_i = 1 - \alpha$ and earn expected profit equal to $(1-\alpha)(V(\gamma(1-\gamma)) - c) + \gamma^2 V\alpha$.*

- *If there is a moderate probability of a moderately successful venture (i.e.,* $\gamma > c/V$*,* $\gamma(1-\gamma)V - c > 0$ *and* $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$*), then one firm sets* $k = 1-\alpha$ *and the other firm sets* $k = \frac{\gamma V(1-\alpha(1-\gamma))-c(1-\alpha)}{2\gamma^2 V}$*. The higher capacity firm earns expected profit equal to* $\frac{c(1-\alpha)+\gamma V(1-\alpha(1-\gamma))}{2}$ *and the lower capacity firm earns expected profit equal to* $\frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)}$*.*

- *If there is a high probability of a successful venture (i.e.,* $\gamma > c/V$ *and* $\gamma(1-\gamma)V - c < 0$*), then the unique symmetric equilibrium is* $k_1 = k_2 = 1/2$*. Firms earn expected profit equal to* $(\gamma V - c)/2$*.*

Proposition 1 highlights a non-monotonic effect of γ on the equilibrium capacity choice. Intuitively, if there is a low probability of success then neither firm wishes to invest in computational capacity because there is a high probability of it going unused. Interestingly, when there is a high probability of a successful venture, firms dampen competition by choosing a capacity that just covers the market. To understand this, consider the extreme case in which $\gamma = 1$. If firms choose computational capacity such that $k_1 + k_2 = 1$, both firms can charge their monopoly price for all of their consumers. The moment capacities are such that firms compete even for a single consumer, the firms are unable to avoid intense price competition for that consumer, thereby affecting revenues from all of their customers. Though an additional unit of capacity can result in an additional sale, the subsequent effect on price competition is severe enough such that firms refrain from competing directly.

Another interesting facet of Proposition 1 is that a moderate probability of a successful venture leads to excessive capacity choices. Therefore, a reduction in the probability of success can actually cause an increase in capacity. To understand this result, consider two competing effects of decreasing γ. On the one hand, lower γ implies greater downside risk that the chosen capacity will go completely unused due to $v_i = 0$ and the resulting failure in the market. This effect would suggest that capacity should decrease as γ decreases. On the other hand, a firm's chance at having monopoly power over all of Segment 3 is maximized at moderate levels of γ. The latter *monopoly harvesting* effect dominates the former *downside risk* effect at moderate levels of γ.

We also see in Proposition 1 that symmetric firms may adopt asymmetric strategies. Since the probability of a successful venture is reasonably high, the potential to serve as a monopolist with probability $\gamma(1-\gamma)$ is lucrative enough such that one firm wishes to expand its capacity beyond

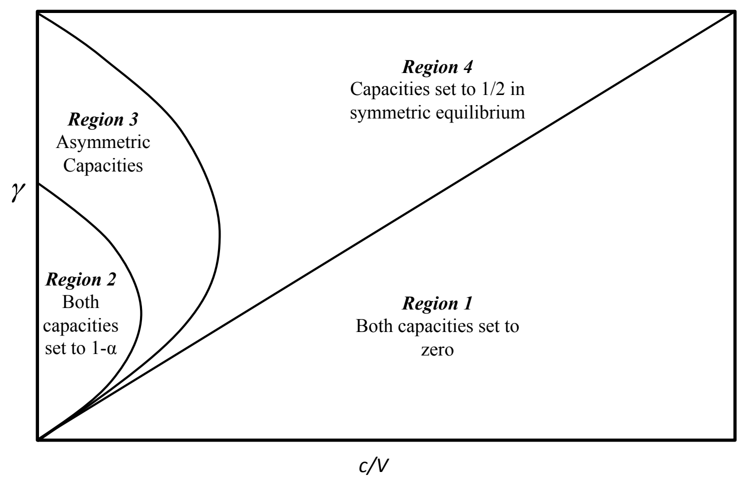

Figure 5: Equilibrium capacities as a function of γ and c/V

the non-competing level of $k_i = 1/2$. However, since the value of being a monopolist is contained (i.e., V is not too high), the second firm loses more in the competitive scenario than it gains in the monopoly scenario by matching the high capacity. As such, one firm will choose capacity to be able to serve all of Segment 3 and the other firm will choose its best response recognizing the drawback of maximum capacity in the case of competition.

The results of Proposition 1 are depicted in Figure 5.[7] Note that Region 2 in which both firms set capacity $k = 1 - \alpha$ disappears when $\alpha = 0$. In other words, as the firms become undifferentiated in the minds of consumers, they will either differentiate in their capacity choice (Region 3) or dampen competition via restricted capacity choices (Region 4). As α increases, Region 2 expands at the expense of Region 3 in which firms choose asymmetric capacities. In fact, the border between Region 3 and Region 4 in which both firms choose $k = 1/2$ becomes the border between Region 2 and Region 4 as α approaches $1/2$, thereby eliminating Region 3. In other words, as the firms become more differentiated in the minds of consumers, they become less likely to adopt asymmetric strategies.

We now turn our attention to the entry decision. If only one firm enters the market, the firm will be a monopolist optimally choosing $k = 1 - \alpha$ and earning expected profit $(\gamma V - c)(1 - \alpha)$ if

[7] In Figures 5–9, we use parameters $\alpha = \frac{1}{4}$, $c = \frac{1}{2}$, $V = \frac{7}{2}$ and $F = 0$, unless that parameter is being used as a variable in the figure.

$\gamma > c/V$ and optimally choosing $k = 0$ to earn zero profit otherwise.

Lemma 1 *If $F > (\gamma V - c)(1-\alpha)$, then there is no entry in the conventional model. Otherwise, the entry decisions depend on F, γ, and V as follows:*

- *If there is a moderate probability of a highly successful venture (i.e., $\gamma > c/V$ and $\gamma(1-\gamma)V - c > 0$ and $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$), both firms enter if $F < (1-\alpha)(V\gamma(1-\gamma)-c) + \gamma^2 V\alpha$. Otherwise only one firm enters.*

- *If there is a moderate probability of a moderately successful venture (i.e., $\gamma > c/V$ and $\gamma(1-\gamma)V - c > 0$ and $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$), both firms enter if $F < \frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)}$. Otherwise only one firm enters.*

- *If there is a high probability of a successful venture (i.e., $\gamma > c/V$ and $\gamma(1-\gamma)V - c < 0$), then both firms enter if $F < (\gamma V - c)/2$. Otherwise only one firm enters.*

Lemma 1 shows how the probability of success can affect the firms' decision to enter. Next, we consider the effect of autoscaling on firms' entry decisions and compare it to our findings from this Lemma.

4.2 The Effects of Autoscaling

We now turn our attention to the equilibrium when autoscaling is available. We allow for firms to decide whether to use autoscaling or prepay for computational capacity after the entry decision. Major cloud providers offer autoscaling with no additional fees.[8] We first solve for the equilibrium pricing supposing both firms enter and choose autoscaling. We will then look at the equilibrium pricing supposing only one firm chooses autoscaling to identify the equilibrium decisions of whether to adopt autoscaling. Upon identifying the equilibrium adoption of autoscaling conditional on entry, we look at how firms' prices, profits, and entry decisions are affected by autoscaling.

First, consider the case where both firms enter and choose autoscaling. With probability γ^2, we have $v_1 = v_2 = V$, and it is straightforward to show there is no pure strategy pricing equilibrium; instead the pricing subgame leads to a mixed strategy equilibrium where the prices of both firms range between z' and V. Similar to our analysis of mixed strategy equilibrium without autoscaling, z'

[8]https://aws.amazon.com/autoscaling/pricing/, accessed September 2016.

is the price for which each firm is indifferent between *attacking*, resulting in a profit of $(z'-c)(1-\alpha)$, and *retreating*, resulting in a profit of $\alpha(V-c)$. This results in $z' = \frac{\alpha(V-c)}{1-\alpha} + c$.

Supposing that $G_i(.)$ is the cumulative distribution function for the price of Firm i, the profit of Firm i when setting price x, is

$$\pi_i(x) = \alpha G_j(x)(x-c) + (1-G_j(x))(1-\alpha)(x-c)$$

Setting the derivative of this function equal to zero for $x \in (z', V)$ and using the boundary conditions $G(z') = 0$ or $G(V) = 1$, we find

$$G_j(x) = \begin{cases} 0 & \text{if } x < z' \\ \frac{(1-\alpha)(x-c)-\alpha(V-c)}{(1-2\alpha)(x-c)} & \text{if } z' \leq x \leq V \\ 1 & \text{if } x > V \end{cases}$$

which results in the expected profit $\pi_{AA} = \alpha(V-c)$ for each firm, when they both use autoscaling.

With probability $\gamma(1-\gamma)$, $v_1 = V$ and $v_2 = 0$, giving Firm 1 monopoly power over all of Segment 3 and profit of $(V-c)(1-\alpha)$. Thus, expected profit with autoscaling when both firms enter is $(V-c)(\gamma^2\alpha + \gamma(1-\gamma)(1-\alpha))$.

Finally, consider the case where both firms enter but only one firm adopts autoscaling. Without loss of generality, we assume that Firm 2 adopts autoscaling and Firm 1 chooses capacity k_1. Using the same techniques as in Section 4.1, we prove in the Technical Appendix that there is no pure strategy equilibrium for prices. The mixed strategy, represented by cumulative distribution function H_i, is given by

$$H_1(x) = \begin{cases} 0 & \text{if } x < z'' \\ \frac{-ck_1+(-1+k_1)V+\alpha(c-x)+x}{(k_1-\alpha)(x-c)} & \text{if } z'' \leq x < V \\ 1 & \text{if } x \geq V \end{cases}$$

$$H_2(x) = \begin{cases} 0 & \text{if } x < z'' \\ \frac{k_1(-ck_1+(-1+k_1)V+\alpha(c-x)+x)}{(-1+\alpha)(\alpha-k_1)x} & \text{if } z'' \leq x < V \\ 1 & \text{if } x \geq V \end{cases}$$

where $z'' = \frac{\alpha c - c k_1 + (k_1-1)V}{\alpha-1}$ is the price at which Firm 1 is indifferent between attacking at price z'' and retreating at price V. Note that $H_2(x)$ is discontinuous at $x = V$. This implies that Firm 2 uses price V with probability $1 - \frac{k_1(V-c)}{(1-\alpha)V}$. In other words, $h_2(V) = (1 - \frac{k_1(V-c)}{(1-\alpha)V})\delta(0)$, where $h_2(.)$ is the probability density function for price of Firm 2 and $\delta(.)$ is Dirac delta function. Prior to the pricing game, the optimal capacity k_1 that maximizes expected profit for Firm 1 is given by:

$$k_1^* = \frac{V\gamma(1-\alpha+\alpha\gamma) - c\left(1+\alpha\left(-1+\gamma^2\right)\right)}{2(V-c)\gamma^2}$$

We may now examine the equilibrium adoption of autoscaling. The payoffs from each possible firm choice of autoscaling or capacity k are summarized in Table 2.

	Firm 2 uses Autoscaling	**Firm 2 uses capacity *k***
Firm 1 uses Autoscaling	$\pi_1 = (V-c)(\gamma^2\alpha + \gamma(1-\gamma)(1-\alpha))$ $\pi_2 = (V-c)(\gamma^2\alpha + \gamma(1-\gamma)(1-\alpha))$	$\pi_1 = \gamma^2(1-k)(v-c) + \gamma(1-\gamma)(1-\alpha)(v-c)$ $\pi_2 = \gamma^2\frac{k(c(k-\alpha)+V(1-k))}{1-\alpha} + \gamma(1-\gamma)kV - ck$
Firm 1 uses capacity *k*	$\pi_1 = \gamma^2\frac{k(c(k-\alpha)+V(1-k))}{1-\alpha} + \gamma(1-\gamma)kV - ck$ $\pi_2 = \gamma^2(1-k)(v-c) + \gamma(1-\gamma)(1-\alpha)(v-c)$	Profits are the same as in Section 4.1

Table 2: Payoffs from autoscaling adoption strategies assuming both firms enter the market

Comparing payoffs across cases in Table 2, the equilibrium adoption of autoscaling is summarized in the following lemma.

Lemma 2 *A firm's decision on using autoscaling depends on the number of market entrants, α, γ, c, and V as follows:*

- *If both firms enter the market, both will choose autoscaling for low enough probability of success, $\gamma < (1-\alpha)/(2-3\alpha)$, or high enough computational costs, such that $c/V > L$;*
- *If both firms enter the market, only one firm will choose autoscaling for high enough probability of success and low enough computational costs, such that $\gamma > (1-\alpha)/(2-3\alpha)$ and $c/V < L$;*
- *If one firm enters the market, it will choose autoscaling;*

where L is defined as:

$$L \equiv \frac{\gamma^2(-3\alpha\gamma+\alpha+2\gamma-1)^2}{2(1-\gamma)\sqrt{(1-\alpha)^3\gamma^3(2\alpha\gamma-\alpha-\gamma+1)} + \gamma\left((2-3\alpha)^2\gamma^3 + ((9-5\alpha)\alpha-4)\gamma^2 - (\alpha-1)\alpha\gamma + (\alpha-1)^2\right)}$$

Lemma 2 establishes the equilibrium adoption of autoscaling, which will be used in subsequent analysis of the effect of autoscaling's availability on prices, entry, profits, and consumer surplus. Autoscaling allows the firms to avoid over- or under-spending on computational capacity. However, there is also a strategic effect of autoscaling that can lead to dampened or intensified competition. Interestingly, the Lemma shows that symmetric firms may make asymmetric decisions regarding the use of autoscaling. When competition is sufficiently intensified by autoscaling, one firm chooses a fixed capacity in order to alleviate competition to some extent.

Effect of Autoscaling on Equilibrium Prices

Given the firms' equilibrium strategies, we can examine how autoscaling will affect equilibrium prices in the event that entry costs are low enough such that multiple firms enter.

Proposition 2 *Suppose entry costs are such that both firms enter the market with or without autoscaling. The effect of autoscaling on average prices depends on* γ *as follows:*

- *If the probability of a successful venture is not too high (i.e.,* $\gamma(1-\gamma)V - c > 0$ *and* $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$*), then autoscaling increases the average price set by each firm.*
- *If the probability of a successful venture is high enough (i.e.,* $\gamma > c/V$ *and* $\gamma(1-\gamma)V - c < 0$*), then autoscaling decreases the average price set by each firm.*

Proposition 2 shows the effect of autoscaling on the average prices charged by each firm in Regions 2 and 4 of Figure 5. In Region 2, both firms would choose enough capacity to serve their own segment (Segment i for Firm i) and all of Segment 3 without autoscaling. The cause of the change in average prices with the addition of autoscaling is that autoscaling turns the cost of each server from a sunk cost to a cost that depends on the number of consumers served by the firm. Without autoscaling, firms do not consider the cost of servers in their pricing decision, as this cost is sunk. Thus, they receive no negative utility from serving a larger portion of the market and are more flexible to do so by decreasing price. However, with autoscaling, each additional customer adds an additional cost, resulting in firms having less incentive to decrease their price to get more customers compared to when costs were sunk. Therefore, autoscaling decreases the firms' incentive to *attack* aggressively with price and increases the average price in Region 2. Thus demand uncertainty creates an important distinction from previous literature on capacity choice (e.g., Kreps and Scheinkman 1983), as it results in a capacity game that *decreases* average prices relative to the pricing game that arises with autoscaling.

On the other hand, the price-competition dampening effect of autoscaling is mitigated in Region 4 where firms would choose not to *attack* without autoscaling since they restrict their capacity to dampen competition. The equilibrium choice of capacity results in both firms charging the maximum price of V. Autoscaling removes this separation of targeted consumers and increases competition between the two firms, resulting in decreased average prices in Region 4.

Proposition 2 thus shows that the probability of a successful venture is critical in determining whether autoscaling increases or decreases prices; a result that is new to the literature.

Note that Region 3 in Figure 5 represents a case where the two firms have different average prices without autoscaling. Since with both firms using autoscaling they both set the same price, evaluating the effect of autoscaling on average charged prices is not as straight forward as in Regions 2 and 4. Later in this section, we use consumer surplus as a proxy to average prices to study the effects of autoscaling in this region. But first, we examine the firms' entry decisions.

Effect of Autoscaling on Entry Decisions

Given the firms' equilibrium strategies and their expected profits, we can derive their entry decisions. We summarize the entry decisions with autoscaling in the following lemma.

Lemma 3 *If* $F > \gamma(V-c)(1-\alpha)$*, then there is no entry in the autoscaling model. Both firms enter the market if* $F < Max[(V-c)(\gamma^2\alpha + \gamma(1-\gamma)(1-\alpha)), \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma v(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)}]$*. Otherwise, only one firm enters.*

By comparing the results presented in Lemma 1 to Lemma 3, we can determine the effect of autoscaling on firm entry. We first consider the effect of autoscaling on participation in the market by any of the firms, i.e., how autoscaling affects whether at least one of the two firms enters.

Proposition 3 *Autoscaling increases the range of entry costs, F, such that at least one firm enters the market.*

Proposition 3 confirms common intuition that autoscaling can make entry more attractive for at least one firm. The reason is that it allows firms with uncertain likelihood of success to incur the cost of computational needs after demand is realized. This highlights the *downside risk reducing* effect of autoscaling. Without autoscaling, firms have to invest in the cost of entry F and the cost of computational capacity ck_i prior to realizing whether the venture will be successful. Autoscaling increases the range of entry costs for which at least one firm enters by $c(1-\gamma)(1-\alpha)$. Thus, the higher the cost of capacity and the lower the probability of success, the more effective autoscaling will be in guaranteeing that the market will be served by at least one firm. We now examine how autoscaling affects whether multiple firms enter the market.

Proposition 4 *There exists a cut-off $\hat{\gamma}$ for which autoscaling decreases the range of entry costs, F, such that both firms enter if and only if $\gamma > \hat{\gamma}$.*

Proposition 4 finds the counter-intuitive result that autoscaling can decrease market entry. Though autoscaling has a *downside risk reducing* effect, it also has a *competition intensifying* effect. In other words, autoscaling makes it less costly for a firm to find out if it has a successful venture on its hands, but also less costly for a firm to fight aggressively for consumers in Segment 3. To further explain these effects and when each is dominant, we consider the three potential outcomes if both firms enter.

When both firms enter, there is a $\gamma(1-\gamma)$ probability that a firm finds itself a monopolist, a γ^2 probability that a firm finds itself competing head-to-head, and a $1-\gamma$ probability that a firm finds $v_i = 0$. In the former case, autoscaling weakly benefits firms because they are assured of having the computational capacity to satisfy the demand of all consumers in Segment 3. Without autoscaling, firms acknowledging the downside risk choose $k_i \leq 1-\alpha$ and thus cannot satisfy all demand when given monopoly power over all consumers in Segment 3. This is the problem that startup Befunky experienced without autoscaling in the earlier example and represents a positive *demand satisfaction* effect of autoscaling. In the latter case, autoscaling weakly benefits firms because it prevents them from over-purchasing capacity. Without autoscaling $k_i \geq 0$ and thus firms have excess computational capacity when $v_i = 0$. This is the positive *downside risk reducing* effect of autoscaling. However, autoscaling weakly disadvantages firms if they compete head-to-head. The same fact that benefits firms when they are monopolists is damaging when they compete: with autoscaling firms are assured of having computational capacity to satisfy demand of all consumers in Segment 3. As such, autoscaling intensifies competition. Without autoscaling, the fact that

$k_i \leq 1 - \alpha$ allows firms to include a more profitable *retreating* price in the equilibrium mixed strategy. In this case, the demand satisfaction effect actually leads to the negative *competition intensifying* effect of autoscaling.

The *downside risk reducing* effect is most dominant when γ is low. The *demand satisfaction* effect is most dominant when $\gamma(1-\gamma)$ is high (i.e., moderate γ) and the *competition intensifying* effect is most dominant when γ is high. A high γ increases the probability of competition and also makes it such that firms without autoscaling choose capacities such that this competition is avoided. Therefore, autoscaling decreases market entry when the probability of a successful venture is sufficiently high. Propositions 3 and 4 suggest that to find the effect of autoscaling on the number of new entrants, we must consider the probability of success as well as the cost of entry, which goes against the intuition that autoscaling always facilitates market entry.

The results of Propositions 3 and 4 are graphically depicted in Figure 6. As shown in this figure, there are four regions of interest. In region A, autoscaling decreases entry due to the *competition intensifying* effect. Autoscaling allows one firm to be a monopolist because the other firm cannot profitably enter given the anticipated level of competitive intensity. In regions B and D, the market will not be served by either firm unless there is autoscaling. In region C, a firm would have monopoly power because the downside risk of capacity pre-purchase makes it unprofitable for a second entrant, but autoscaling alleviates these effects and results in competing firms entering the market.

Corollary 1 *For low costs of entry, autoscaling can create a prisoner's dilemma. Both firms choose autoscaling even though they earn greater expected profit in the absence of autoscaling.*

As noted previously, the *competition intensifying* effect can outweigh the *demand satisfaction* effect and the *downside risk reducing* effect for sufficiently high γ. If F is sufficiently low, both firms will choose to enter with or without autoscaling. Furthermore, as shown in Corollary 1, they both choose autoscaling in equilibrium. Interestingly, this leads to a prisoner's dilemma situation where the firms' adoption of autoscaling results in diminished expected profitability of both firms. This result is depicted in Figure 7. The dashed lines in Figure 7 correspond to regions when autoscaling is not available (from Figure 5), and show how autoscaling affects firms' equilibrium profits in different regions. When the probability of success, γ, is very high, only one firm uses autoscaling

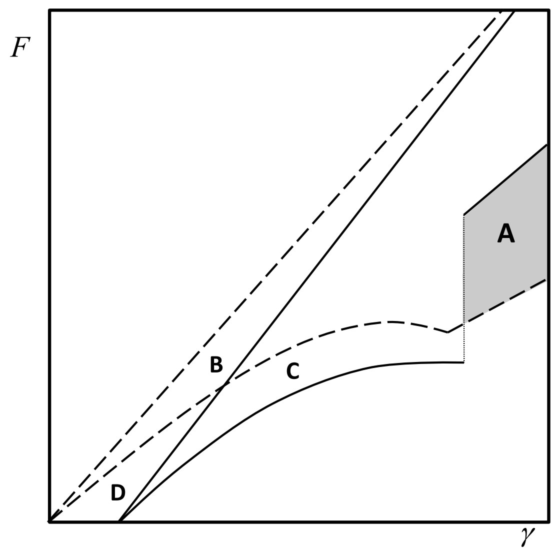

Figure 6: The effect of autoscaling on entry: Region A denotes autoscaling decreases entry from 2 to 1 firms. Region B denotes autoscaling increases entry from 0 to 1 firm. Region C denotes autoscaling increases entry from 1 to 2 firms. Region D denotes autoscaling increases entry from 0 to 2 firms.

while the competing firm can strategically limit its computational capacity to soften competition. Also, when γ is sufficiently low, both firms use autoscaling, but due to the downside risk reducing effect of autoscaling, both firms get higher profits with autoscaling. However, a moderately high γ creates a prisoner's dilemma situation where the competition intensifying effect of autoscaling dominates the downside risk reducing effect, but the firms still use autoscaling. Therefore, both firms would be better off if autoscaling was not available in this region. We now turn our attention to the impact of autoscaling on consumers.

Effect of Autoscaling on Consumer Surplus

So far, we studied the effects of autoscaling on firms' strategies and their profit. Autoscaling can increase price competition between firms. It can also increase market entry. Both of these effects, intuitively, should lead to higher surplus for consumers. However, autoscaling also changes the capacity cost from sunk cost at the time of pricing to variable cost. Therefore, as shown in Proposition 2, autoscaling can lead to higher average prices, and thus lower surplus, for consumers. Furthermore, as shown in Proposition 4, autoscaling can also increase the likelihood of a monopoly market. In this section, we study the effect of these opposing forces on consumer surplus.

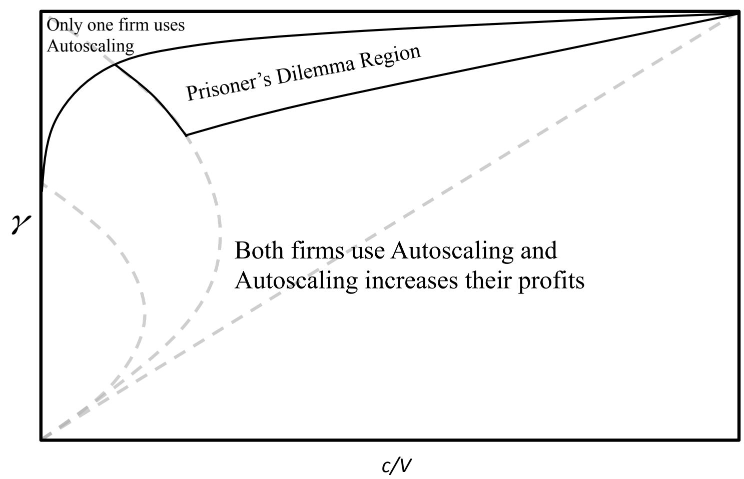

Figure 7: When existence of autoscaling can lead to a prisoner's dilemma effect for the firms

Expected consumer surplus can be derived from calculating the difference between expected social welfare and combined expected firm profit. Social welfare is equal to the combined value consumers get (i.e., V times the number of purchases) minus the cost to deliver that value (i.e., c times the computational capacity). Therefore, the expected consumer surplus can be written as

$$E[CS] = V \times E[\#\text{purchases}] - c \times E[\text{computational capacity}] - E[\pi_1] - E[\pi_2] \tag{3}$$

where #purchases and computational capacity indicate the total number of consumers who purchase the product and the total computational capacity reserved by the firms, respectively. If there is only one firm in the market, in both cases with and without autoscaling, that firm sets the price to V and results in zero consumer surplus. If both firms are in the market, consumer surplus with autoscaling available depends on whether both firms choose to adopt autoscaling or if only firm adopts this feature. We present the values of consumer surplus derived from Equation (3) in the Technical Appendix. Comparing across conditions, we have the following result.

Proposition 5 *For sufficiently small F, sufficiently small c and a moderate value of γ, autoscaling decreases expected consumer surplus. If F is sufficiently large, autoscaling has no effect on the expected consumer surplus. Otherwise, autoscaling increases the expected consumer surplus.*

Proposition 5 shows the effect of autoscaling on consumer surplus. The result is also depicted in Figures 8 and 9. Figure 8 is analogous to Figure 5 with the contours from Figure 5 marked with dashed gray lines. As we can see, for Region 2 of Figure 5, where both firms set their capacities to $1-\alpha$ in the absence of autoscaling, and part of Region 3, where firms set asymmetric capacities, autoscaling decreases consumer surplus. Intuitively, since autoscaling increases average prices in this region, it decreases consumer surplus.

Figure 9 is analogous of Figure 6 with the contours from Figure 6 marked as dashed gray lines. The figure shows that autoscaling can only affect consumer surplus if the cost of entry is sufficiently low. In particular, unless both firms enter the market in at least one of the two cases (i.e., autoscaling or no autoscaling available), consumer surplus will be zero; therefore, if F is not low enough, autoscaling can have no effect on consumer surplus. Figure 9 also shows that when autoscaling increases market entry to two firms, it also increases expected consumer surplus.

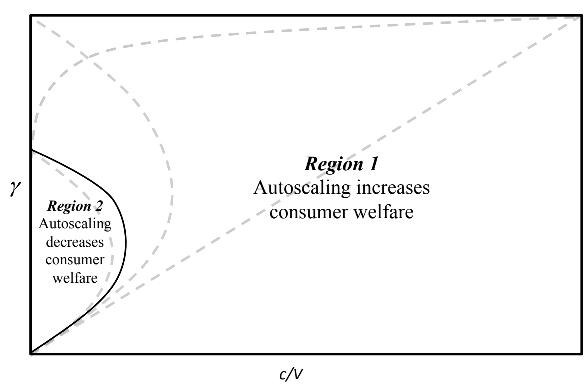

Figure 8: Effect of autoscaling on consumer surplus as a function of V, c and γ

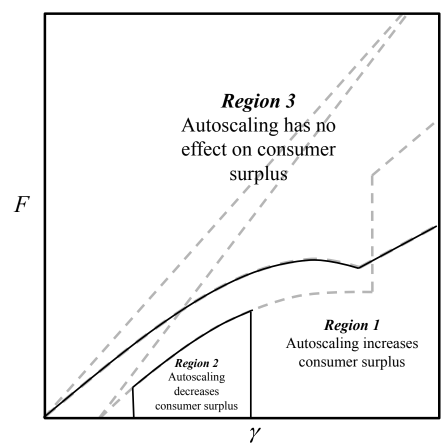

Figure 9: Effect of autoscaling on consumer surplus as a function of F and γ

5 Discussion

The emergence of cloud computing and its feature of autoscaling has the potential to influence the entry of startups into new markets. Before autoscaling was available, startups needed to invest heavily in computational capacity before entering the market so that they could serve highly unpredictable demand levels. The uncertainty in demand meant that the capacity chosen by firms could be more or less than the actual needs, resulting in extra costs or unfulfilled demand. However, autoscaling allows the new entrants a flexible capacity such that the cloud provider will designate the specific computational resources as the actual demand and required resources are realized. This feature can help new entrants to the market pay only for the capacity that is required to meet their demand, reducing the disadvantages of uncertainty in demand. In this paper, we find the effects of autoscaling on entry decisions of startups, their prices, their capacity decisions, and their profits.

Our model shows how the *demand satisfaction* effect of autoscaling, in which startups can be assured of having capacity to satisfy demand, turns out to be a double-edged sword in that it frees competing startups to aggressively pursue customers. Our research identifies this *competition intensifying* effect of autoscaling and establishes the conditions under which this effect will outweigh the positive effects of autoscaling. This effect has several important implications for startups.

The results can help guide pricing decisions by startups who enter the market with autoscaling available. We find that autoscaling will increase the average prices startups charge if the probability of a successful venture is not too high. In this case, autoscaling decreases a startup's costs, but these savings are not passed on to consumers because autoscaling converts a sunk fixed cost into a variable cost that a startup incurs if and only if they attract more customers. As a consequence, price competition is dampened by autoscaling in the case that startups would otherwise choose excessive capacities. However, if the probability of a successful venture is sufficiently high, startups would optimally limit their capacities to dampen competition in the absence of autoscaling. In this case, autoscaling gives the startups freedom to aggressively pursue customers with price and thus can decrease average prices.

The trade-off between the intensified competition and the reduced downside risk of failure affects startups' decisions on whether to enter the market. In particular, when the probability of success is sufficiently high, the increased competition effect of autoscaling dominates the decreased downside risk of failure. As a result, while autoscaling facilitates the entry for the first startup, it deters the second startup from entering the market. Our findings expand the market entry literature by looking at the topic from the new angle of capacity choice before versus after demand is realized.

Our research also guides managers of startups with their response to the introduction of autoscaling and cloud computing. The research shows if the probability of success is moderately high, autoscaling leads to a prisoner's dilemma situation where both startups adopt autoscaling, but they would have been better off if both had avoided it. Our findings suggest that in industries where cost of entry is relatively low, the introduction of the autoscaling feature in cloud computing can reduce the profit of new entrants to the market even though it decreases their downside risk of failure.

Autoscaling could also have positive or negative effects on consumer surplus. On the one hand, it can increase price competition between the firms and facilitate their entry, both of which lead to higher consumer surplus. On the other hand, autoscaling can lower consumer surplus by dampening price competition or by discouraging the second startup from entering the market if one startup has already entered. Our results show that when cost of entry is low and probability of success is moderate, autoscaling decreases expected consumer surplus.

Overall, our results highlight how startups should use information on cost of entry, probability of success, cost of capacity and level of differentiation to evaluate their entry decisions, capacity commitments and pricing strategies in the presence of autoscaling. Our research is one of the first to consider the marketing aspects of cloud computing and autoscaling. With the rapid adoption of cloud computing by firms across different industries, marketing and economics research on the cloud can be a rich and important topic of study for future research.

References

[1] Armbrust, M., A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica. 2009. "Above the clouds: A Berkeley View of Cloud Computing." *Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS* 28, 13.

[2] Anupindi, R. and L. Jiang. 2008. "Capacity Investment Under Postponement Strategies, Market Competition, and Demand Uncertainty." *Management Science* 54(11): 1876-1890.

[3] Biyalogorsky, E., and O. Koenigsberg. 2014. "The Design and Introduction of Product Lines When Consumer Valuations are Uncertain." *Production and Operations Management* 23(9): 1539-1548.

[4] Bort, C. 2015. "Google Has Nabbed Some Huge Customers for its Most Important New Business Cloud Computing." *http://www.businessinsider.com/google-names-huge-companies-using-its-cloud-2015-6*, accessed September 2016.

[5] Buehler, S. and J. Haucap. 2006. "Strategic Outsourcing Revisited." *Journal of Economic Behavior and Organization* 61(3) 325-338.

[6] Che, H., C. Narasimhan, and V. Padmanabhan. 2010. "Leveraging Uncertainty Through Backorder." *Quantitative Marketing and Economics* 8(3): 365-392.

[7] Chen, P. and S. Wu. 2013. "the Impact and Implications of On-Demand Services on Market Structure." *Information Systems Research* 24(3): 750-767.

[8] Chen, Y. and G. Iyer. 2002. "Consumer Addressability and Customized Pricing." *Marketing Science* 21(2): 197-208.

[9] Columbus, L., 2015. "Roundup of Cloud Computing Forecasts and Market Estimates Q3 Update, 2015." *www.forbes.com/sites/louiscolumbus/2015/09/27/roundup-of-cloud-computing-forecasts-and-market-estimates-q3-update-2015/*, accessed September 2016.

[10] Desai, P., O. Koenigsberg, and D. Purohit. 2007. "The Role of Production Lead Time and Demand Uncertainty in Marketing Durable Goods." *Management Science* 53(1): 150-158.

[11] Desai, P., O. Koenigsberg, and D. Purohit. 2010. "Forward Buying by Retailers." *Journal of Marketing Research* 47(1): 90-102.

[12] Daughety, A. 1990. "Beneficial Concentration." *American Economic Review* 80(5): 1231-1237.

[13] Ferguson, M. and O. Koenigsberg. 2007. "How Should a Firm Manage Deteriorating Inventory." *Production and Operations Management* 16(3): 306-321.

[14] Flood, G. 2013. "Gartner Tells Outsourcers: Embrace Cloud Or Die." *http://www.informationweek.com/cloud/infrastructure-as-a-service/gartner-tells-outsourcers-embrace-cloud-or-die/d/d-id/1110991*, accessed September 2016.

[15] Gaudin, S. 2015. "AWS Assures Customers it's Atop the Cloud." *http://www.computerworld.com/article/2990119/cloud-computing/aws-assures-customers-theyre-atop-the-cloud.html*, accessed September 2016.

[16] Griffith, E. 2014. "Why Startups Fail, According to Their Founders." *Forbes* Sep, 25.

[17] Gupta, P., A. Seetharaman, and J. R. Raj. 2013. "The Usage and Adoption of Cloud Computing by Small and Medium Businesses." *International Journal of Information Management* 33(5): 861-874.

[18] Hassani, S. 2009. *Mathematical Methods for Students of Physics and Related Fields* (2nd ed.). New York, NY: Springer. Chapter 5.

[19] Iyer, G., D. Soberman, and J. Villas-Boas. 2005. "The Targeting of Advertising." *Marketing Science* 24(3): 461-476.

[20] Joshi, Y., D. Reibstein, and J. Zhang. 2009. "Optimal Entry Timing in Markets with Social Influence." *Management Science* 55(6): 926-939.

[21] Kreps, D. and J. Scheinkman. 1983. "Quantity Precommitments and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics* 14(2): 326-337.

[22] Leavitt, N. 2009. "Is Cloud Computing Really Ready For Prime Time?" *Computer* (1): 15-20.

[23] Lilien, G. and E. Yoon. 1990. "The Timing of Competitive Market Entry: An Exploratory Study of New Industrial Products." *Management Science* 36(5): 568-585.

[24] Marston, S., Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi. 2011. "Cloud Computing: The Business Perspective." *Decision Support Systems* 51(1): 176-189.

[25] Milgrom, P. and J. Roberts. 1982. "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis." *Econometrica* 50(2): 443-459.

[26] Narasimhan, C. and J. Zhang. 2000. "Market Entry Strategy Under Firm Heterogeneity and Asymmetric Payoffs." *Marketing Science* 19(4): 313-327.

[27] Narasimhan, C. 1988. "Competitive Promotional Strategies." *Journal of Business* 61(4): 427-449.

[28] Nasser, S. and D. Turcic. 2015. "To Commit or Not to Commit: Revisiting Quantity vs. Price Competition in a Differentiated Industry." *Management Science* 62(6): 1719-1733..

[29] Nickelsburg, M. 2016. "Startup Spotlight: BeFunky makes online photo editing and graphic design simple." *http://www.geekwire.com/2016/befunky/*, accessed September 2016.

[30] Ofek, E. and O. Turut. 2013. "Vaporware, Suddenware, and Trueware: New Product Preannouncements Under Market Uncertainty." *Marketing Science* 32(2): 342-355.

[31] Poepsel, M. 2008. "Customers Are Won or Lost in One Second, Finds New Aberdeen Report" *http://www.prnewswire.com/news-releases/customers-are-won-or-lost-in-one-second-finds-new-aberdeen-report-65399152.html*, accessed September 2016.

[32] Ray, T. 2013. "Salesforce, Google, Amazon Cloud Winners, Says Piper; Microsoft Straddles the Line." *blogs.barrons.com* Oct, 17.

[33] Reynolds, S. and B. Wilson. 2000. "Bertrand-Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes." *Journal of Economic Theory* 92: 122-141.

[34] Shulman, J. 2014. "Product Diversion to a Direct Competitor." *Marketing Science* 33(3):422-436.

[35] Shy, O. and R. Stenbacka. 2003. "Strategic outsourcing." *Journal of Economic Behavior and Organization* 50(2):203-224.

[36] Spence, M. 1977. "Entry, Capacity, Investment and Oligopolistic Pricing." *The Bell Journal of Economics* 8(2): 534-544.

[37] Sultan, N. A. 2011. "Reaching for the Cloud: How SMEs Can Manage." *International journal of information management* 31(3):272-278.

[38] Swinney, R., G. Cachon., and S. Netessine. 2011. "Capacity Investment Timing by Start-ups and Established Firms in New Markets." *Management Science* 57(4): 763-777.

[39] Tirole, J. 1988. *The Theory of Industrial Organization.* Cambridge, MA: MIT Press.

[40] Van Mieghem, J. and M. Dada. 1999. "Price Versus Production Postponement: Capacity and Competition." *Management Science* 45(12):1631-1649.

[41] Walker, E. 2009. "The Real Cost of a CPU Hour." *Computer* (4): 35-41.

[42] Weaver, L. 2015. "Why GE is moving nearly 9,000 apps to the public cloud." *https://enterprisersproject.com/article/2015/10/why-ge-s-moving-nearly-9000-apps-public-cloud*, accessed September 2016.

[43] Work, S. 2011. "How Loading Time Affects Your Bottom Line." *https://blog.kissmetrics.com/loading-time/*, accessed September 2016.

[44] Yang, H., and M. Tate. 2012. "A Descriptive Literature Review and Classification of Cloud Computing Research." *Communications of the Association for Information Systems* 31(2): 35-60.

[45] Zhang, K. and Z. Katona. 2012. "Contextual Advertising." *Marketing Science* 31(6): 980-994.

[46] Zhou, B., C. Mela, and W. Amaldoss. 2015. "Do Firms Endowed with Greater Strategic Capability Earn Higher Profits?" *Journal of Marketing Research* 52(3): 325-336.

A Technical Appendix

Proof of Proposition 1

If $\gamma < c/V$, then expected profit is strictly decreasing in capacity for any capacity chosen by the competition. Therefore, each firm optimally chooses zero capacity.

If $\gamma > c/V$, then for any $k_2 < 1-\alpha$, expected profit is increasing in k_1 for any $k_1 < 1-k_2$. Therefore, we can rule out any $k_1+k_2 < 1$. First consider $\gamma(1-\gamma)V - c < 0$. Suppose there were a symmetric equilibrium in which $k_1 + k_2 > 1$. By definition, this implies that $k_1 > 1/2$ which means Firm 1 earns greater expected profit by deviating downward. Therefore, the only potential symmetric equilibrium requires $k_1 + k_2 = 1$. If Firm 1 deviates upward, its expected change in profit is $\gamma(1-\gamma)V - c < 0$.

If $\gamma(1-\gamma)V - c > 0$, then $k_1 + k_2 = 1$ is no longer an equilibrium because either firm can profitably deviate to harvest the potential of monopoly power. We can therefore focus our attention on $k_1 + k_2 > 1$. If $k_2 > k_1$, then Firm 2's expected profit is strictly increasing in k_2 until $k_2 = 1-\alpha$ and is strictly decreasing for any $k_2 > 1-\alpha$. If $k_2 > k_1$, Firm 1's expected profit for any $k_1 < 1-\alpha$ is given by $\gamma(1-\gamma)V + \gamma^2\frac{(1-k_1)k_1V}{k_2} - k_1c$ which is concave in k_1 and maximized at $k_1 = \frac{\gamma V(1-\alpha(1-\gamma))-c(1-\alpha)}{2\gamma^2 V}$. Note this value of k_1 is in fact less than $1-\alpha$ if and only if $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$. We verify that Firm 2 cannot benefit from a global deviation undercutting k_1 at this level. Supposing that Firm 2 deviates to a lower capacity than Firm 1, its profit would become

$k_2\left(-c+\gamma V\left(1-\gamma-\frac{2\gamma^3(-1+k_2)V}{c(-1+\alpha)+\gamma V(1+(-1+\gamma)\alpha)}\right)\right)$, which is maximized at

$$\tilde{k}_2=-\frac{c^2(-1+\alpha)+c\gamma V(2+2\alpha(-1+\gamma)-\gamma)+V^2\gamma^2\left(-1+\alpha+\gamma-2\alpha\gamma+(-2+\alpha)\gamma^2\right)}{4V^2\gamma^4}.$$

It is easily shown that any values of c that allow for $\tilde{k}_2<k_1$ will result in Firm 2's profit at $\tilde{k}_2$ (i.e., $\frac{\left(c^2(-1+\alpha)+c\gamma V(2-\gamma+2(-1+\gamma)\alpha)+\gamma^2V^2\left(-1+\gamma+\gamma^2(-2+\alpha)+\alpha-2\gamma\alpha\right)\right)^2}{8\gamma^4V^2(c(-1+\alpha)+\gamma V(1+(-1+\gamma)\alpha))}$) to be less than Firm 2's profit at our equilibrium.

If $V>\frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$ then neither firm benefits from deviating from $k_j=1-\alpha$.

Proof of Lemma 1

A firm will enter if its expected profit is greater than its entry cost. If and only if $F<(\gamma V-c)(1-\alpha)$, then a firm's best response to its competitor not entering the market is to enter the market. Supposing one firm enters the market, the remaining firm's best response to its competitor's entry is to also enter the market provided the expected profit earned when competing is greater than the cost of entry. The anticipated payoffs associated with being one of two firms entering are reported in Proposition 1 and the conditions on F are presented in the Lemma.

Analysis of Pricing Subgame When Only One Firm Uses Autoscaling

We start by showing that this game does not have a pure strategy equilibrium. Assume for sake of contradiction that the firms use prices p_1 and p_2 in a pure strategy equilibrium. If $p_1\neq p_2$, then the firm with a lower price can benefit from deviating by increasing its price to $\frac{p_1+p_2}{2}$. If $p_1=p_2$, then Firm 2 can benefit from deviating by decreasing its price to $p_2-\varepsilon$, for sufficiently small ε, to acquire all consumers in Segment 3. Therefore, a pure strategy equilibrium cannot exist.

Next, we find a mixed strategy equilibrium for this game. Provided $k_1\leq 1-\alpha$, Firm 2 can choose to *attack* with a price that clears its capacity or *retreat* with a price V that harvests the value from the $1-k_1$ consumers that Firm 1 cannot serve due to its capacity constraint. Let z'' be the price at which Firm 2 is indifferent between *attacking* to sell to $1-\alpha$ consumers at price z'' and *retreating* to sell to $1-k_1$ consumers at price V. We have $(1-k_1)(V-c)=(1-\alpha)(z''-c)$ which gives us $z''=\frac{\alpha c-ck_1+(k_1-1)V}{\alpha-1}$. In equilibrium, both firms use a mixed strategy with prices ranging from z'' to V. Suppose that $H_i(.)$ is the cumulative distribution function used by Firm i.

The profit of Firm 1 earned by setting price x is

$$\pi_1(x) = H_2(x)\alpha x + (1 - H_2(x))k_1 x - k_1 c$$

Using equilibrium conditions, we know that the derivative of this function with respect to x must be zero for $x \in (z'', V)$. By solving the differential equation we get

$$H_2(x) = \begin{cases} 0 & \text{if } x < z'' \\ \frac{k_1(-ck_1+(-1+k_1)V+\alpha(c-x)+x)}{(-1+\alpha)(\alpha-k_1)x} & \text{if } z'' \leq x < V \\ 1 & \text{if } x \geq V \end{cases}$$

Similarly, the profit of Firm 2 earned by setting price x is

$$\pi_2(x) = H_1(x)(1 - k_1)(x - c) + (1 - H_1(x))(1 - \alpha)(x - c)$$

By setting the derivative with respect to x to zero for $x \in (z'', V)$ and solving the differential equation, we get

$$H_1(x) = \begin{cases} 0 & \text{if } x < z'' \\ \frac{-ck_1+(-1+k_1)V+\alpha(c-x)+x}{(k_1-\alpha)(x-c)} & \text{if } z'' \leq x < V \\ 1 & \text{if } x \geq V \end{cases}$$

Proof of Lemma 2

We first show that both firms not using autoscaling cannot be an equilibrium. The expected profit of using autoscaling, when the opponent uses capacity k, is

$$\pi_{AN} = \gamma^2(1 - k)(V - c) + \gamma(1 - \gamma)(1 - \alpha)(V - c).$$

The expected profit of using fixed capacity k', when the opponent uses capacity k, according to the proof of Proposition 1, is

$$\pi_{NN} = \gamma(1-\gamma)k'(V-c) - (1-\gamma)k'c + \gamma^2 \times \begin{cases} \frac{(1-k')k'V}{k} - k'c & \text{if } k' \leq k \\ \min(k', 1-k)V - k'c & \text{if } k' > k \end{cases}$$

First, note that if $k \leq \frac{1}{2}$, then $\begin{cases} \frac{(1-k')k'V}{k} - k'c & \text{if } k' \leq k \\ \min(k', 1-k)V - k'c & \text{if } k' > k \end{cases}$ is optimized at $k' = 1-k$. Therefore, we have $\pi_{AN} > \pi_{NN}$, i.e., deviating to autoscaling is profitable when the opponent uses fixed capacity $k \leq \frac{1}{2}$. Thus, for an equilibrium in which neither firm uses autoscaling to exist, both firms must be using fixed capacity larger than half.

We know from Proposition 1 that the only possible equilibrium in which both firms use fixed capacity larger than $\frac{1}{2}$ (i.e., that satisfies best response equilibrium condition) is the one in which they both set their capacities to $1-\alpha$. But in this case, we have $\pi_{NN} = (1-\alpha)(V(\gamma(1-\gamma)) - c) + \gamma^2 V\alpha$, which is strictly less than π_{AN} for feasible k. Therefore, deviating to autoscaling is strictly profitable for each firm. Therefore, both firms using fixed capacities cannot be an equilibrium.

Given the fact that at least one firm uses autoscaling in equilibrium, we examine whether both firms use autoscaling or only one. If Firm 2 adopts autoscaling, Firm 1's best response is to adopt autoscaling if and only if

$$(V-c)(\gamma^2\alpha + \gamma(1-\gamma)(1-\alpha)) \geq \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right) + \gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(1-\alpha)\gamma^2(V-c)}$$

where the right hand side of the inequality is Firm 1's profit if it chooses the best possible capacity k, given by $k_1^* = \frac{c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha-1-\alpha\gamma)}{2\gamma^2(c-V)}$, to Firm 2's autoscaling decision. By rearranging this inequality for c/V, we get the condition in Lemma 2. Finally, we can confirm that when the above inequality does not hold and $k_1^* = \frac{c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha-1-\alpha\gamma)}{2\gamma^2(c-V)}$, Firm 2 does not benefit from deviating to fixed capacity. Therefore, when the above inequality does not hold, Firm 1 uses fixed capacity k_1^* and Firm 2 uses autoscaling in equilibrium.

If only one firm enters, this firm will enjoy monopoly power over $(1-\alpha)$ consumers with probability γ. Using autoscaling, the firm's profit equals to $\gamma(V-c)(1-\alpha)$, which is strictly larger than $(V\gamma - c)(1-\alpha)$, the profit earned without autoscaling, for any $\gamma > 0$.

Proof of Proposition 2

We start by calculating the average prices without autoscaling when $\gamma > c/V$, $\gamma(1-\gamma)V - c > 0$, and $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$. Based on Proposition 1, both firms set their capacities equal to $1-\alpha$. The cumulative distribution function of prices set by each firm is $F(x) = \frac{(1-\alpha)x+V\alpha}{(1-2\alpha)x}$. Thus, the probability density function equals $f(x) = \frac{V\alpha}{x^2(1-2\alpha)}$ and the average price of each firm is $\bar{p}_{NN} = \int_z^V f(x)x\,dx = \frac{\alpha V \log(\frac{1-\alpha}{\alpha})}{1-2\alpha}$. With both firms using autoscaling, the probability density function equals $g(x) = \frac{\alpha(V-c)}{(1-2\alpha)(x-c)^2}$ and the average price of each firm becomes $\bar{p}_{AA} = \int_{z'}^V g(x)x\,dx = \frac{(1-2\alpha)c+\alpha\log(\frac{1-\alpha}{\alpha})(V-c)}{1-2\alpha}$. Thus we have $\bar{p}_{AA} - \bar{p}_{NN} = c(1 - \frac{\alpha\log(\frac{1-\alpha}{\alpha})}{1-2\alpha})$. For all $s > 1$, we know that $\log(s) < s-1$. Allowing $s = \frac{1-\alpha}{\alpha}$, we show that $\log(\frac{1-\alpha}{\alpha}) < \frac{1-2\alpha}{\alpha}$, thus $\bar{p}_{AA} > \bar{p}_{NN}$. When only one firm (Firm 2) uses autoscaling and the other (Firm 1) chooses capacity k_1, the probability density function of Firm 1's price equals $h_1(x) = -\frac{(k_1-1)(c-V)}{(\alpha-k_1)(c-x)^2}$ and its average price is $\bar{p}_{NA} = \int_{z''}^V h_1(x)x\,dx = c + \frac{(1-k_1)(V-c)\log\left(\frac{1-\alpha}{1-k_1}\right)}{k_1-\alpha}$. We show $\frac{\partial(\bar{p}_{NA})}{\partial k_1} = -\frac{(c-V)\left((\alpha-1)\log\left(\frac{\alpha-1}{k_1-1}\right)-\alpha+k_1\right)}{(\alpha-k_1)^2}$ is negative. For all $s > 0$, we know that $s\log(s) > s-1$. Allowing $s = \frac{1-\alpha}{1-k_1}$, we have $\frac{\alpha-k_1}{k_1-1} < \frac{(1-\alpha)Log((1-\alpha)/(1-k_1))}{(1-k_1)}$ and $\frac{\partial(\bar{p}_{NA})}{\partial k_1} < 0$. Therefore, the minimum of $\bar{p}_{NA}$ occurs at $k_1 = 1-\alpha$ and is equal to $\bar{p}_{AA}$. Thus, $\bar{p}_{NA} \geq \bar{p}_{AA} > \bar{p}_{NN}$.

For Firm 2, which uses autoscaling, the probability density function equals

$$h_2(x) = \frac{k_1(-\alpha c + ck_1 - k_1V + V)}{(\alpha-1)x^2(\alpha-k_1)} + \left(1 - \frac{k_1(c-V)}{(\alpha-1)V}\right)\delta(x-V)$$

and average price equals

$$\bar{p}_{AN} = \int_{z''}^V h_2(x)x\,dx = \frac{(\alpha-k_1)(V(\alpha+k_1-1)-ck_1) + k_1(-\alpha c+ck_1-k_1V+V)\log\left(\frac{(1-\alpha)V}{-\alpha c+ck_1-k_1V+V}\right)}{(\alpha-1)(\alpha-k_1)}$$

We show $\frac{\partial(\bar{p}_{AN})}{\partial k_1} = \frac{\left(V\left(-2\alpha k_1+\alpha+k_1^2\right)-c(\alpha-k_1)^2\right)\log\left(\frac{(1-\alpha)V}{c(k_1-\alpha)+V(1-k_1)}\right)-\alpha(\alpha-k_1)(c-V)}{(\alpha-1)(\alpha-k_1)^2}$ is negative. For all $s > 0$, we know that $s\log(s) > s-1$. Allowing $s = \frac{(1-\alpha)V}{c(k_1-\alpha)+V(1-k_1)}$, we find $\frac{\partial(\bar{p}_{AN})}{\partial k_1} < \frac{(\alpha-k_1)(c-V)^2}{(\alpha-1)^2V} < 0$. Therefore, the minimum of $\bar{p}_{AN}$ occurs at $k_1 = 1-\alpha$ and we have $Min[\bar{p}_{AN}] - \bar{p}_{NN} = c\log\left(\frac{(1-\alpha)V}{-2\alpha c+\alpha V+c}\right) + \frac{\alpha V\log\left(\frac{-2\alpha c+\alpha V+c}{\alpha V}\right)}{2\alpha-1} + c$. Since $\log\left(\frac{-2\alpha c+\alpha V+c}{\alpha V}\right) < \frac{-2\alpha c+\alpha V+c}{\alpha V} - 1$, we know $Min[\bar{p}_{AN}] - \bar{p}_{NN} > \frac{(\alpha V)\left(\frac{-2\alpha c+\alpha V+c}{\alpha V}-1\right)}{2\alpha-1} + c\left(\log\left(\frac{(1-\alpha)V}{-2\alpha c+\alpha V+c}\right)+1\right) = c\log\left(\frac{(1-\alpha)V}{-2\alpha c+\alpha V+c}\right) > 0$. Thus, $\bar{p}_{AN} > \bar{p}_{NN}$.

When $\gamma > c/V$ and $\gamma(1-\gamma)V - c < 0$, we know from Proposition 1 that without autoscaling, $k_1 + k_2 = 1$ and both firms set their price equal to V. As we showed, $\log(\frac{1-\alpha}{\alpha}) < \frac{1-2\alpha}{\alpha}$. Thus, $\bar{p}_{AA} = \frac{(1-2\alpha)c+\alpha\log\left(\frac{1-\alpha}{\alpha}\right)(V-c)}{1-2\alpha} < V$. Since $\log(\frac{1-\alpha}{1-k_1}) < \frac{1-\alpha}{1-k_1} - 1 = \frac{k_1-\alpha}{1-k_1}$, we have $\bar{p}_{NA} < V$. Finally, we have $\log\left(\frac{(1-\alpha)V}{-\alpha c+ck_1-k_1V+V}\right) < \frac{(1-\alpha)V}{-\alpha c+ck_1-k_1V+V} - 1$ and therefore, $\bar{p}_{AN} < V$.

Proof of Lemma 3

Given the best capacity response to autoscaling, $k_1^* = \frac{V\gamma(1-\alpha+\alpha\gamma)-c\left(1+\alpha\left(-1+\gamma^2\right)\right)}{2(V-c)\gamma^2}$, the expected profit for Firm 1 choosing not to autoscale, assuming that Firm 2 uses autoscaling, is

$$\pi_{NA} = \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(1-\alpha)\gamma^2(V-c)}$$

and the expected profit for Firm 2 choosing to autoscale, assuming that Firm 1 uses fixed capacity k_1^*, is

$$\pi_{AN} = \frac{(\gamma V(\alpha(\gamma-1)+1) - c(\alpha(\gamma-1)^2+2\gamma-1))}{2}.$$

It is straightforward to show $\pi_{NA} < \pi_{AN}$. Thus, if the subgame equilibrium involves only one firm choosing autoscaling, both firms will enter if $F < \pi_{NA}$.

Both firms will choose autoscaling if $\pi_{NA} < \pi_{AA}$. As such, both firms will enter if $F < Max[\pi_{NA}, \pi_{AA}]$.

Proof of Proposition 3

With autoscaling, at least one firm enters the market if and only if $F < F_A \equiv \gamma(V-c)(1-\alpha)$ whereas without autoscaling at least one firm enters if and only if $F < F_N \equiv (\gamma V - c)(1-\alpha)$.

Proof of Proposition 4

We compare the cut-offs on F such that both firms enter. Let F_{AA} denote the cut-off for two firms to enter with autoscaling, F_{NA} denote the cut-off for two firms to enter when only one firm uses autoscaling, and F_{NN} denote the cut-off for two firms to enter when autoscaling is not available.

First consider the two cases in which $\gamma > c/V$ and $\gamma(1-\gamma)V - c > 0$. If $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, then $0 < c < V$ requires $\gamma < (1-\alpha)/(2-3\alpha)$. Therefore, based on Lemma 2, both firms will use autoscaling if autoscaling is available. Thus, $F_{AA} - F_{NN} = c(1-\alpha+\gamma(1-\gamma-\alpha+2\alpha\gamma))$ which is decreasing in α and positive at $\alpha = 1/2$ and therefore positive for all $\alpha < 1/2$. If $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, the derivative of $F_{AA} - F_{NN}$ with respect to V is $\frac{c^2(1-\alpha)^2-\gamma^2V^2(1-\alpha+\gamma(3\alpha-2))^2}{4\gamma^2V^2(1-\alpha)}$, which is positive for all $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$. At $V = \frac{c}{\gamma}$, we have $F_{AA} - F_{NN} = \frac{1}{4}c\left(4(1-2\alpha)\gamma^2 + \frac{(\alpha(13\alpha-20)+8)\gamma}{\alpha-1} - 4\alpha + 4\right)$ which is positive for $c/V > L$, where both firms use autoscaling. Therefore, $F_{AA} - F_{NN}$ is positive for $V > \frac{c}{\gamma}$, which includes all possible values of V in this case. Also for $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, $F_{NA} - F_{NN}$ is increasing in V and equal to $\frac{c\gamma^2(\alpha(\alpha(\gamma-1)\gamma+2)-1)}{4(\alpha-1)((\gamma-1)\gamma+1)}$ at $V = \frac{c}{\gamma(1-\gamma)}$. Since $\alpha < \frac{1}{2}$, we have $\frac{c\gamma^2(\alpha(\alpha(\gamma-1)\gamma+2)-1)}{4(\alpha-1)((\gamma-1)\gamma+1)} > 0$, and therefore, $F_{NA} > F_{NN}$ for $V > \frac{c}{\gamma(1-\gamma)}$, which includes all possible values of V in this case.

Now consider when $\gamma > c/V$ and $\gamma(1-\gamma)V - c < 0$. In this case, $F_{AA} - F_{NN} = \gamma(V-c)(1-\alpha-\gamma(1-2\alpha) - (\gamma V - c)/2$ which is convex in γ, equal to $-(V-c)(1-2\alpha)/2 < 0$ when evaluated at $\gamma = 1$, decreasing in γ through $\gamma = 1$, and equal to zero at

$$\hat{\gamma}_{AA} = \frac{V(1-2\alpha) - 2c(1-\alpha) + \sqrt{8c(V-c)(1-2\alpha) + (V - 2c(1-\alpha) - 2V\alpha)^2}}{4(V-c)(1-2\alpha)}.$$

Therefore, $F_{AA} - F_{NN}$ is negative for any $\gamma > \hat{\gamma}_{AA}$. Also in this case,

$$F_{NA} - F_{NN} = \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right) + \gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)} - \frac{1}{2}(\gamma V - c)$$

which is equal to $((-1+2\alpha)(c-V))/(4(-1+\alpha)) < 0$ at $\gamma = 1$ and $\frac{\partial(F_{NA}-F_{NN})}{\partial\gamma}$ equals $\frac{(2\alpha-1)(c-V)}{2(\alpha-1)} < 0$ at $\gamma = 1$. $F_{NA} = F_{NN}$ has one root between $\gamma = 0$ and $\gamma = 1$ which is

$$\hat{\gamma}_{NA} = 1/2\Big(-(((-1+\alpha)(1-\alpha+\sqrt{1-2\alpha})V)/(\alpha^2(c-V))) + \sqrt{\frac{(\alpha-1)\left(\alpha^3(V-2c)^2+\alpha^2\left(4\left(\sqrt{1-2\alpha}-1\right)c^2-4\left(\sqrt{1-2\alpha}-1\right)cV+\left(2\sqrt{1-2\alpha}-5\right)V^2\right)+2\left(3-2\sqrt{1-2\alpha}\right)\alpha V^2+2\left(\sqrt{1-2\alpha}-1\right)V^2\right)}{\alpha^4(c-V)^2}}\Big)$$

Therefore, $F_{NA} - F_{NN}$ is negative for any $\gamma > \hat{\gamma}_{NA}$. Thus, autoscaling decreases the range of F such that both firms enter the market for $\gamma > Max[\hat{\gamma}_{AA}, \hat{\gamma}_{NA}]$.

Proof of Corollary 1

We show that for sufficiently small F (such that both firms enter the market), when $\gamma > c/V$, $\gamma(1-\gamma)V - c < 0$, and $\pi_{NA} < \pi_{AA} < \frac{\gamma V - c}{2}$, we have a prisoner's dilemma situation where both firms use autoscaling, even though their profits would be higher if autoscaling was not available.

When $\gamma > c/V$ and $\gamma(1-\gamma)V - c < 0$, both firms set their capacity to $\frac{1}{2}$, and each earn expected profit $\frac{\gamma V-c}{2}$, if autoscaling is not available. However, when autoscaling is available, since $\pi_{NA} < \pi_{AA}$, they both use autoscaling and each earn $\pi_{AA} < \frac{\gamma V-c}{2}$ in equilibrium, which creates the prisoner's dilemma. Now, we have to prove that all these conditions can be satisfied at the same time to show that the described region, in which prisoner's dilemma happens, actually exists.

Let $\gamma = \frac{1-\alpha}{2-3\alpha}$. After algebraic simplifications, we have both firms using autoscaling in equilibrium (i.e., $\pi_{NA} < \pi_{AA}$) if and only if $\frac{V}{c} > \frac{10\alpha^2-15\alpha+6}{4(\alpha-1)^2}$. Furthermore, using algebraic simplifications, the expected profit of autoscaling equilibrium for each firm is lower than that when firms do not use autoscaling (i.e., $\pi_{AA} < \frac{\gamma V-c}{2}$) if and only if $\frac{V}{c} < \frac{\alpha^2+2\alpha-2}{(\alpha-1)\alpha}$. It is easy to see that $\frac{10\alpha^2-15\alpha+6}{4(\alpha-1)^2} < \frac{\alpha^2+2\alpha-2}{(\alpha-1)\alpha}$ for any $\alpha \leq 1/2$. Therefore, for when $\frac{V}{c} \in (\frac{10\alpha^2-15\alpha+6}{4(\alpha-1)^2}, \frac{\alpha^2+2\alpha-2}{(\alpha-1)\alpha})$, the prisoner's dilemma situation holds.

Proof of Proposition 5

If $c/V > L$ where L is defined in Lemma 2, then both firms would choose autoscaling upon entry and

$$E[CS_{AA}] = \begin{cases} \gamma^2(V-c)(1-2\alpha) & \text{if } F < (V-c)((1-\alpha)(1-\gamma)\gamma + \alpha\gamma^2) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

If $c/V < L$, then only one firm would choose autoscaling upon entry and

$$E[CS_{NA}] = \begin{cases} \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)\left(c\left((\alpha-2)\gamma^2-4(\alpha-1)\gamma+\alpha-1\right)+\gamma V(\alpha\gamma+\alpha-1)\right)}{4(\alpha-1)\gamma^2(c-V)} & \text{if } F < F_{NA} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $F_{NA} \equiv \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)}$.

Similarly, when autoscaling is not available, the expected consumer surplus is

$$E[CS_{NN}] = \begin{cases} \gamma^2 V(1-2\alpha) & \text{if } c < \gamma(1-\gamma)V \text{ and} \\ & V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))} \text{ and} \\ & F < (1-\alpha)(\gamma(1-\gamma)V-c)+\alpha\gamma^2 V \\ \frac{(1-\alpha)c^2}{4\gamma^2 V} + \frac{(\alpha-1)c(2\gamma+1)}{2\gamma} + \frac{V(\alpha(\gamma-1)+1)(\alpha\gamma+\alpha-1)}{4(\alpha-1)} & \text{if } c < \gamma(1-\gamma)V \text{ and} \\ & V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))} \text{ and} \\ & F < \frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Using Equations (4), (5), and (6), we compare consumer surplus with and without autoscaling. We start with the region $c < \gamma(1-\gamma)V$ and $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, which represents Region 2 in Figure 5. Note that $V > \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$ and $0 < c < V$ require $\gamma < (1-\alpha)/(2-3\alpha)$, therefore based on Lemma (2) it is not possible that only one firm uses autoscaling in this region. Thus, we only compare cases when both firms use autoscaling with cases when autoscaling is not available. Using Equations (4) and (6), we can see that if $F < (1-\alpha)(\gamma(1-\gamma)V-c)+\alpha\gamma^2 V$, then both firms enter the market with or without autoscaling and $E[CS_{NN}] > E[CS_{AA}]$. For $(1-\alpha)(\gamma(1-\gamma)V-c)+\alpha\gamma^2 V < F < (V-c)((1-\alpha)(1-\gamma)\gamma+\alpha\gamma^2)$, $E[CS_{AA}] > E[CS_{NN}] = 0$. For $F > (V-c)((1-\alpha)(1-\gamma)\gamma+\alpha\gamma^2)$, $E[CS_{AA}] = E[CS_{NN}] = 0$.

Next, consider when $c < \gamma(1-\gamma)V$ and $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, which represents Region 3 in Figure 5. First, we analyze the cases when both firms use autoscaling in Region 3. Comparing Equations (4) and (6) and assuming $c/V < \gamma$, we find that $E[CS_{NN}] > E[CS_{AA}]$ when $c/V < \frac{\gamma\left(4\alpha\gamma^3-\sqrt{\gamma}\sqrt{4(1-2\alpha)^2\gamma^5-8(\alpha-1)(2\alpha-1)\gamma^3-4(\alpha-1)(2\alpha-1)\gamma^2+(\alpha(13\alpha-20)+8)\gamma+4(\alpha-1)^2}-2\alpha\gamma-\alpha-2\gamma^3+2\gamma+1\right)}{1-\alpha}$ and $F < \frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)}$. Note that these conditions only hold when both firms use autoscaling, since $\frac{\gamma\left(4\alpha\gamma^3-\sqrt{\gamma}\sqrt{4(1-2\alpha)^2\gamma^5-8(\alpha-1)(2\alpha-1)\gamma^3-4(\alpha-1)(2\alpha-1)\gamma^2+(\alpha(13\alpha-20)+8)\gamma+4(\alpha-1)^2}-2\alpha\gamma-\alpha-2\gamma^3+2\gamma+1\right)}{1-\alpha}$ is positive if and only if $\gamma < (1-\alpha)/(2-3\alpha)$. Also in Region 3 of Figure 5, if both firms use autoscaling, $E[CS_{AA}] > E[CS_{NN}] = 0$ if $\frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)} < F < (V-c)((1-\alpha)(1-\gamma)\gamma+\alpha\gamma^2)$, and $E[CS_{AA}] = E[CS_{NN}] = 0$ if $F > (V-c)((1-\alpha)(1-\gamma)\gamma+\alpha\gamma^2)$.

Next, we compare the consumer surplus without autoscaling in Region 3, which occurs for

$c < \gamma(1-\gamma)V$ and $V < \frac{(1-\alpha)(2V\gamma^2+c)}{\gamma(1+\alpha(\gamma-1))}$, with consumer surplus when only one firm uses autoscaling, which occurs for $\gamma > (1-\alpha)/(2-3\alpha)$ and $c/V < L$. Comparing Equations (5) and (6), we find $E[CS_{NA}] > E[CS_{NN}]$ for $F < \frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)}$ and

$$
\begin{aligned}
c/V < \; & \frac{\alpha^2(2-\gamma(\gamma((\gamma-4)\gamma-4)+2)) + 2\alpha(\gamma(\gamma((\gamma-2)\gamma-5)+2)-2)+2(\gamma(3\gamma-1)+1)}{2(\alpha-1)^2} \\
& - ((\alpha^2(2-\gamma(\gamma((\gamma-4)\gamma-4)+2)) + 2\alpha(\gamma(\gamma((\gamma-2)\gamma-5)+2)-2)+6\gamma^2-2\gamma+2)^2 \\
& + 4(\alpha-1)^2\gamma(\alpha^2(\gamma^3-6\gamma^2+\gamma-2)+\alpha(4-2(\gamma-4)\gamma^2)-\gamma(2\gamma+1)-2))^{1/2}/(2(\alpha-1)^2)
\end{aligned}
$$

Since the term on the right hand side of the above inequality is greater than γ, thus for all $c/V < \gamma$, we have $E[CS_{NA}] > E[CS_{NN}]$. Furthermore in this region, $E[CS_{NA}] > E[CS_{NN}] = 0$ if $\frac{(\gamma V(1-\alpha(1-\gamma))-c(1-\alpha))^2}{4\gamma^2 V(1-\alpha)} < F < \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)}$, and $E[CS_{NA}] = E[CS_{NN}] = 0$ if $F > \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)}$.

Finally, if $c > \gamma(1-\gamma)V$, then when autoscaling is not available, both firms either do not enter the market or enter and set their prices equal to V. Both of these cases result in zero consumer surplus. Therefore autoscaling increases consumer surplus when both firms use autoscaling and $F < (V-c)((1-\alpha)(1-\gamma)\gamma + \alpha\gamma^2)$, or when one firm uses autoscaling and $F < \frac{\left(c\left(\alpha\left(\gamma^2-1\right)+1\right)+\gamma V(\alpha(-\gamma)+\alpha-1)\right)^2}{4(\alpha-1)\gamma^2(c-V)}$. Otherwise, autoscaling does not affect consumer surplus.