

# Competitive Capacity Investment under Uncertainty

Xishu Li, Rob Zuidwijk, M.B.M de Koster, Rommert Dekker  
Erasmus University Rotterdam, the Netherlands x.li@rsm.nl

March 3, 2016

We consider a long-term capacity investment problem in a competitive market under demand uncertainty. Two firms move sequentially in the competition and a firm's capacity decision interacts with the other firm's current and future capacity. Throughout the investment race, a firm can either choose to plan its investments proactively, taking into account possible responses from the other firm, or decide to respond reactively to the competition. In both cases, the optimal decision at each period is determined according to an ISD (*Invest, Stayput, Disinvest*) policy. We develop two algorithms to efficiently derive proactive ISD policies for the leader and follower firms. Using data from the container shipping market (2000-2015), we show that the optimal capacity determined by our competitive strategy is consistent with the realized investments in practice. By revealing strategical flexibility of proactive strategies, our results demonstrate that firms in the competition can gain more capacity and profit through such a strategy. Using Monte Carlo simulations, we explore the impact of different market conditions and investment irreversibility levels on capacity strategies. In particular, by comparing the results of competitive strategies and strategies that separate firms into different markets, we show that both firms can benefit from the competition and that market downturns likely lead to investment cascades.

*Key words:* capacity, dynamic investment, proactive strategy, Stackelberg competition, feedback policies

---

## 1. Introduction

Capacity investment refers to the change in a firm's stocks of various processing resources over time (Van Mieghem, 2003). Firms face a number of challenges in such decisions, since capital assets are costly and an investment is usually irreversible while there is uncertainty in future rewards. As a discrepancy between a firm's capacity and demand results in inefficiency and losses, either through under-utilized resources or unfulfilled demand, the goal of capacity planning is often to minimize this discrepancy in a profitable way. However, doing this is not always possible when firms compete in quantity and their decisions interact with each other (Del Sol and Ghemawat, 1999). In a competitive market where dominant firms exist and product price fluctuates with these firms' capacity, the decisions of one firm directly impact those of the other firms. Investment strategies that ignore competition can have fundamental problems, as they either tend to recommend waiting too long before making an investment, or underestimate the likely countermoves of the other dominant firms towards the decision firm's investment. Without a proper theory, investment decisions in a competitive market can lack guidance.

The container shipping market, where the shipping service price is controlled by a small number of liner operators through their capacity<sup>1</sup>, is a good example of a competitive market. Over the past few years, we have observed a striking investment race among shipping firms for fleet capacity: the world fleet capacity in fully cellular containerships has increased by 60.3% since 2009 (Alphaliner, 2015). However, in stark contrast to the enormous increase in fleet capacity, the shipping industry has had a difficult ride since the 2008 global recession (Barnato, 2015). The battle of survival for shipping firms can only be partially attributed to the crisis or to “buying too many ships” before the crisis started, in anticipation of continued growth. The situation was aggravated by post-crisis investment cascades. Except Maersk’s E class, all Ultra Large Container Vessels were ordered after 2008 (Wikipedia, 2016). For example, in 2011 CMA-CGM increased the capacity option of its three on-order vessels by 15.7%<sup>2</sup> (CMA-CGM, 2011) and this capacity record only stood for a short while. In the same year, Maersk spent \$3.8 billion to build 20 Triple-E-class vessels, causing the size of the largest containership to instantly rise by another 14.2% (Macguire, 2013).

On the one hand, these large investments during market downturn cannot be explained by generic investment frameworks. *First*, the investments are not supported by demand. Market economy theories recommend firms to order more new ships when they expect demand to outpace supply growth (Olhager et al., 2001; Van Mieghem, 2003). However, the continuing recession in Europe and the slowdown in China have led firms to downgrade their demand growth estimation from 9%-13% before the crisis to 3%-5% now (Drewry, 2005, 2014). *Second*, the competition for vessel size cannot be justified by economies of scale. Larger vessels are more cost efficient as they result in lower unit costs in many categories, e.g., operating cost and building cost (Cullinane and Khanna, 2000). However, as the gap between global fleet capacity and trade volume has increased to over 144% since 2005 (Søndergaard and Eismark, 2012), the advantage of economy-of-scale cannot always be realized. In fact, carriers face more losses if they sail large vessels with insufficient cargo. *Third*, the outcome of fleet investments in the container shipping market does not meet the general expectation of investment, which is to boost profit. Instead, these investments cause high volatility in freight rate and losses in profits (UNCTAD, 2012, 2013, 2014). For instance, after CMA-CGM’s Marco Polo vessels and Maersk’s Triple-E-class were ordered, the spot rates in the Asia-Europe market hit rock bottom, dropping from an average value of \$1789 per TEU in 2010 to \$450 per TEU in December 2011 (Odell, 2012; UNCTAD, 2013). Consequently, in 2011 many carriers suffered huge losses and depleted their cash reserves; some were close to bankruptcy (Sanders, 2012).

<sup>1</sup> By December 2015, the top 10 shipping lines controlled over 62% of the world container fleet (www.Alphaliner.com).

<sup>2</sup> CMA-CGM’s Marco Polo vessels were originally ordered in 2007 with a capacity of 13,830 TEU. In 2011, CMA-CGM increased the capacity to 16,000 TEU.

On the other hand, these post-crisis investments clearly have a competitive feature since firms invested to get ahead of their competitors during a slow-growth economy and to get ready to win when good times return. Facing market downturn since 2008, leading carriers didn't lay off capacity to mitigate declining freight rates. Instead, they further deflated rates by ordering more vessels, resulting in lower profits and suppressed capacity for competitors (Søndergaard and Eismark, 2012). Competitors with a strong financial position managed to survive by increasing their capacity likewise, whereas weaker firms lacked business to balance the cash flow and thus went bankrupt. Besides the shipping industry, other oligopolistic markets such as the airline industry and the semiconductor industry have also shown similar investment races. For instance, airlines compete in buying larger aircraft while facing meager profits (Chen and MacMillan, 1992) and semiconductor manufacturers invest aggressively in capacity during market downturns (Ghemawat, 2009).

Obviously, competitive capacity investment is risky for any firm as the future is uncertain. One way to reduce the risk is to divide the investment project into several sub-projects and execute them in phases, which gives firms possibilities to respond to future market changes. Because capital assets have long lifetimes, investment decisions that are made in earlier phases influence decision making in the future. A long-term investment strategy should address optimal timing and size of a capacity adjustment. Capturing the optimal investment timing in a competitive market requires firms to balance the financial risk of investing and the competitive risk of not investing. Once-in-a-cycle delays can create a lasting competitive disadvantage in a multi-round investment race. Moreover, practice examples, such as the lack of success of the Airbus 380, have shown that competition does not always require firms to have a high-capacity fleet.

An investment strategy, which help firms survive a recession and thrive in the competition, should achieve an "elusive" balance between being too defensive and too aggressive (Gulati et al., 2010). However, little research has been done on revealing the structure of such an optimal competitive investment strategy, especially in which firms sequentially respond to each other's investments by adjusting its own capacity. Our study fills this gap by investigating optimal long-term investment strategies of two firms moving sequentially in a competitive market where 1) uncertainty exists in exogenous demand growth; 2) competitive interaction between a firm's decision and the other firm's current and future decisions is considered; and 3) a firm's objective at each period is to maximize the expected value of its long-term plan by adapting it to the evolving market.

We contribute to research and practice as follows. *First*, we contribute to the literature by providing a theory that can explain the competitive investment phenomena observed in practice, which cannot be fully understood from current models. *Second*, we derive full optimal long-term

investment policies in a competitive market, in terms of investment timing and size. Existing research either focuses on timing only or studies investment as a single-shot game. *Third*, we explicitly take the competition effect into account by developing methods which derive all stayput regions for both the leader and the follower in their optimal policies. Each stayput region represents an optimal solution set of a firm, taking into account the impact of the competitor's responses on the firm's current and future rewards. *Fourth*, we allow both firms to choose either a proactive or a reactive strategy and develop different methods to efficiently derive the optimal policy in each case. By revealing strategical flexibility of a proactive strategy, our methods show the potential advantage of firms being proactive in the competition. *Fifth*, using data from the container shipping market over a timespan of 16 years (2000-2015), we show that the investment decisions computed by our model are consistent with what happened in practice. *Last*, using Monte Carlo simulations, we show that market and investment cost parameters influence capacity strategies differently.

## 2. Literature review

Literature on strategic capacity investment is concerned with determining the timing and size of buying or selling additional capacity under uncertainty (see Van Mieghem, 2003 and Chevalier-Roignant et al., 2011 for a detailed literature review). Models that study the optimal capacity type often consider a one-period problem where a firm sells two products and has the option to invest in two types of resources: flexible vs dedicated (Van Mieghem, 1998; Goyal and Netessine, 2007). A flexible resource is able to produce either product, at the expense of a higher investment cost than the dedicated resources. These models study the impact of capacity characteristics and demand correction between the two products on the optimal capacity strategy. Since our focus is on the impact of competition on the long-term investment strategy, we limit the review to models that consider only a single type of capacity resources.

Brennan et al. (2000) address the three stages in the development of capacity models: (1) *static* models, (2) *dynamic* models, and (3) *combined real options and game-theoretic* models. As *combined real options and game-theoretic* models also study investment dynamics, they can be considered as a stream within *dynamic* models. *Static* models investigate the optimal locations and sizes of capacity in a processing network for a single or multiple decision makers in a stationary environment where there is no managerial flexibility to cope with market changes (Bish and Wang, 2004; Van Mieghem, 2007). It collapses the problem to a single initial capacity investment where the optimal capacity remains constant over time. This category of capacity models adopts queuing (Lederer and Li, 1997; Cachon and Harker, 2002) and newsvendor network formulations (Van Mieghem and Rudi, 2002; Netessine et al., 2002; Kulkarni et al., 2004). While losing dynamics in capacity decisions,

some models extend the single-period solution to a situation with dynamic i.i.d. demand and hence investigate further dynamics in inventory (Van Mieghem and Rudi, 2002). *Static* models that involve multiple players often study coordination between “vertical” players such as manufacturers and retailers (Cachon and Lariviere, 1999; Armony and Plambeck, 2005; Plambeck and Taylor, 2005; Caldentey and Haugh, 2006), or competition between “horizontal” players (Lederer and Li, 1997; Van Mieghem and Dada, 1999) who supply a common market. Two main aspects considered in these multi-player *static* models are supply network partitioning and information asymmetry. Although these models consider competition effect between firms, they are still restricted to a stationary setting, emphasizing optimal capacity size rather than timing.

*Dynamic* models allow time-dependent investments to respond to the resolution of uncertainty. They emphasize the timing of capacity adjustment in a single-shot<sup>3</sup> or a multi-round game and derive a structured policy for investments at different time points (Burnetas and Gilbert, 2001; Angelus and Porteus, 2002; Narongwanich et al., 2002; Ryan, 2004; Huh and Roundy, 2005; Huh et al., 2006). Some noted approaches in this category are decision-tree analysis, dynamic programming, control theory, and real options approach. Often, optimal investment dynamics follow a so-called ISD (invest, stayput and disinvest) policy, which is characterized by a continuation region: when current capacity falls in this region, it is optimal to stay put; otherwise, it should be adjusted to an appropriate point on the region’s boundary (Eberly and Van Mieghem, 1997). Although traditional *dynamic* models have been refined over time to incorporate many real-world features, such as hedging, the lack of consideration given to competitive interactions between firms limits their applications in a competitive setting. Our work is built on Eberly and Van Mieghem (1997) and we extend their method to incorporate the competition impact on investments.

Our model belongs to the most recent development of *dynamic* models, i.e., *combined real options and game-theoretic* models (Chevalier-Roignant et al., 2011). In these models, firms not only condition their decisions on the resolution of exogenous uncertainties but also on the (re)actions of competitors. The focus is on determining the investment timing of different players and explaining competitive behavior. The most widely used method is the “option games” approach, which involves several decision makers and an uncertain market (Ferreira et al., 2009). Most models consider two types of players only (i.e., leader vs. follower or  $n$  players moving simultaneously<sup>4</sup>), while only a few incorporate a third player (Bouis et al., 2009). There are two major types of *combined real options and game-theoretic* models. The first studies a single-shot investment with lumpy capacity

<sup>3</sup> A single-shot game can contain multiple time periods, where “single” refers to the number of capacity investments a firm can make during the game.

<sup>4</sup> In the case where  $n$  players moving simultaneously, the two types of players refer to the decision firm and its rivals.

(Dixit, 1994; Hoppe, 2000; Murto, 2004; Pawlina and Kort, 2006; Thijssen et al., 2006; Swinney et al., 2011). Investment is viewed as an optimal stopping problem, focusing on finding the demand values at which capacity should be adjusted in order to maximize the expected reward.

The second type of *combined real options and game-theoretic* models allows multiple rounds of investments and explores an optimal capacity strategy that contains a sequence of decisions. Most models in this category focus on incremental capacity expansion (Grenadier, 2002; Aguerrevere, 2003, 2009), while a few investigate repeated lumpy investment decisions (Novy-Marx, 2007). Numerical results of a multi-round investment problem can be derived using stochastic dynamic programming and Monte Carlo simulation (Murto et al., 2004). For analytical results, control theory is used to derive equilibrium investment strategies in a Nash framework. The key feature is that each firm determines its optimal capacity strategy while taking its competitors' strategies as given. Up to now, research has mostly focused on strategies where both firms invest simultaneously and in the same proportion, like in a Cournot-Nash equilibrium strategy (Grenadier, 2002; Smit and Trigeorgis, 2012). The concept of the Nash equilibrium, especially a Cournot-Nash equilibrium, is challenged by Back and Paulsen (2009). They prove that the Cournot-Nash equilibrium presented by Grenadier (2002) fails subgame perfection: if one firm invests more than the equilibrium strategy specifies, driving the price down, other firms ignore this and continue to invest as they would have done otherwise. Another drawback of Cournot capacity models is that the optimal policy does not explicitly reflect the competition impact on firms' investments. Strategies that form a subgame perfect equilibrium and recognize competition should allow firms to react to each other's investments. These strategies have a Stackelberg nature, because all firms respond to the investments of any other firm like a Stackelberg follower (Back and Paulsen, 2009).

Research that examines dynamic competitive investments is an emerging trend in the literature of capacity models. So far they have been applied mostly in financial studies and have some limitations. First, most studies specify fixed capacity sizes as actions available to a firm and use the real options approach to determine only the timing of taking a particular action. Second, current studies have been limited to simultaneous investment strategies and are considered as "open-loop" strategies in the sense that there is no feedback from the investment of any firm to the investment of any other firm, neither in the same period nor in the next ones (Back and Paulsen, 2009). Although "open-loop" strategies are mathematically tractable, they are dynamically inconsistent as decisions are derived at the initial time, without accounting for the state evolution beyond that time. We contribute to the extant literature by introducing sequential feedback strategies, modeled by a Stackelberg game. When using feedback strategies, firms have information on their competitors' current capacity and react to capacity perturbations through their own investments. Moreover,

we allow the size of an investment to be determined by the optimal policy and hence study more complete features of a capacity strategy, i.e., timing and size. Table 1 gives an overview of some existing capacity models and our model.

Capacity models	Examples	Multi-round	Investment decisions		Firm interaction		
			<i>timing</i>	<i>size</i>	<i>vertical</i>	<i>horizontal</i>	
						<i>simultaneous</i>	<i>sequential</i>
Static	Netessine et al., 2002			✓			
	Kulkarni et al., 2004			✓			
	Van Mieghem and Rudi, 2002	✓		✓			
	Plambeck and Taylor, 2005			✓	✓		
	Caldentey and Haugh, 2006			✓	✓		
	Lederer and Li, 1997			✓		✓	
	Van Mieghem and Dada, 1999			✓		✓	
Dynamic	Eberly and Van Mieghem, 1997	✓	✓	✓			
	Ryan, 2004	✓	✓	✓			
	Huh et al., 2006	✓	✓	✓			
	Dixit, 1994		✓				✓
	Hoppe, 2000		✓				✓
	Pawlina and Kort, 2006		✓				✓
	Thijssen et al., 2006		✓				✓
	Swinney et al., 2011		✓			✓	
	Grenadier, 2002	✓	✓			✓	
	Aguerrevere, 2003, 2009	✓	✓			✓	
	<b>Our research</b>	✓	✓	✓			✓

**Table 1** Overview of some existing capacity models and our model

### 3. The model

The notations used in our model are listed in Table 2. To illustrate the model, consider the following example: two firms ( $l$  and  $f$ ) sell a homogeneous product (e.g., shipping service) in an oligopolistic market within a finite time horizon  $\Gamma = \{1, \dots, T\}$ , assuming capacity is instantaneously adjustable but investment is partially irreversible. At the beginning of  $\Gamma$ , firm  $j$  is referred to as the leader if  $j = l$  and as the follower if  $j = f$ . Let  $k_{tj} \in \mathcal{K}_{tj}$  represent firm  $j$ 's capacity level at period  $t$ .  $k_{0l}$  and  $k_{0f}$  are the initial capacity levels. The finite set  $\mathcal{K}_{tj} \subseteq \mathbb{R}_{\geq 0}$  denotes the set of capacity choices available to firm  $j$  at period  $t$ . At the beginning of each period  $t \in \Gamma$ , both firms have the option to change their capacity from  $k_{t-1j}$  to  $k_{tj}$  or continue their current capacity (i.e.,  $k_{tj} = k_{t-1j}$ ). The decisions will take the form of a Stackelberg competition: one firm moves first and the other firm observes the opponent's decision and then makes its own decision. We assume that the sequence is fixed among all periods: the leader always moves first.

At the beginning of each period  $t \in \Gamma$ , the optimal capacity is determined such that the value of a firm's long-term strategy, which is a sequence of actions from the current one to the one at the end of  $\Gamma$ , is maximized. Therefore, the capacity decision is based on the demand, supply, and investment cost information then available to the firm and on its assessment of the uncertain future. Let  $\omega_\tau \in \Theta$  represent the demand at the beginning of period  $\tau$  or an estimation of the demand outlook of period  $\tau$ , where  $\Theta \subseteq \mathbb{R}$  is the set of demand realizations. We allow exogenous uncertainty to exist in  $\omega_\tau$  for  $\tau > t$  and it possesses a Markov property. Thus, there is a transition probability function  $Pr : \Theta \times \Theta \times \Gamma \rightarrow [0, 1]$  such that the conditional transition probability is  $\Pr\{\omega_{t+1} = x_{t+1} \mid \omega_t = x_t\} = Pr(x_t, x_{t+1}, t)$ , independent of  $x_{t_0}$  for  $t_0 < t$ . Thus, the demand information relevant to the capacity decision at period  $t$  contains only the current demand, i.e.,  $\omega_t$ .

The supply information comprises both firms' current capacity levels, available capacity choices, capacity utilization parameters, and utilization cost functions at period  $t$ . Let time period  $t$ , the value of the demand  $\omega_t$ , and the firms' capacity  $k_{tj}$  (or  $k_{t-1j}$ ) define the state of the system. Denote the state space by  $\Omega = \mathcal{K}_{tl} \times \mathcal{K}_{tf} \times \Theta \times \Gamma$ . At the beginning of period  $t$ , the state vector that the leader observes is  $\mathbf{Y}_{tl} = (k_{t-1l}, k_{t-1f}, \omega_t, t)$ , and after the leader's first move  $k_{tl}$ , the state vector observed by the follower is  $\mathbf{Y}_{tf} = (k_{tl}, k_{t-1f}, \omega_t, t)$ . Hereinafter, we omit the time variable  $t$  in the state vector. After capacity decisions  $k_{tl}$  and  $k_{tf}$ , the two firms then engage in a short-term production competition which takes the form of a Cournot competition. The optimal production quantities can be found by allowing firms to set their production simultaneously to maximize their own operating profit of the current period. It is worth mentioning that the production decisions are different in nature from the capacity decisions as the former only affect the current period, whereas the latter influence firms permanently (see also *Murto et al., 2004*).

Firm  $j$ 's production quantity at period  $t$  is  $q_{tj}$ , which is determined by a function of the state  $(k_{tl}, k_{tf}, \omega_t)$ , i.e.,  $q_{tj} = Q_{tj}(k_{tl}, k_{tf}, \omega_t)$ . The total production quantity at period  $t$  is  $q_t = \sum_{j \in \{l, f\}} Q_{tj}(k_{tl}, k_{tf}, \omega_t)$ . Given the total production quantity and the demand of period  $t$ , the price of the homogeneous product is given by an inverse demand function:  $p_t = P_t(q_t, \omega_t)$ . As  $q_t$  can be represented as a function of the state  $(k_{tl}, k_{tf}, \omega_t)$ , we can write the price function in the same manner, i.e.,  $P_t(k_{tl}, k_{tf}, \omega_t)$ . Given the production quantity and capacity, the production cost of firm  $j$  at period  $t$ ,  $h_{tj}$ , is given by a function:  $h_{tj} = H_{tj}(q_{tj}, k_{tj})$ . Let  $a_{tj}$  denote the capacity utilization parameter, which represents the capacity usage per produced unit of firm  $j$  at period  $t$ . Firm  $j$ 's operating profit at period  $t$  is given as follows:

$$\begin{aligned} \pi_{tj}(k_{tl}, k_{tf}, \omega_t) &= p_t q_{tj} - h_{tj}, & \forall t \leq T \\ \text{s.t.} \quad a_{tj} q_{tj} &\leq k_{tj} \end{aligned} \tag{1}$$



As the marginal profit of an investment is usually nonincreasing in practice, we make the following assumption on the operating profit function:

ASSUMPTION 1. *For any given and fixed capacity of the other firm  $k_{ti} \in \mathcal{K}_{ti}$  ( $i \neq j$ ) and for each  $\omega_t \in \Theta$ , firm  $j$ 's operating profit function  $\pi_{tj}(k_{ti}, \cdot, \omega_t)$  is concave in its own decision  $k_{tj}$ .*

Examples of operating profit functions that satisfy Assumption 1 include profit functions associated with market-clearing prices and those with isoelastic prices. Details of the profit function we use will be given in the case study section (Section 7).

In addition to the demand and supply information, which are used for determining the profit of each period, the investment cost information is also used in the capacity decision. This includes the discount rate and both firms' marginal investment costs and marginal disinvestment revenues at period  $t$ ,  $c_{tj}$  and  $r_{tj}$ . It may seem quite strict to assume firms know each other's cost parameters, available capacity choices, capacity utilization parameters, and utilization cost functions mentioned above. However, in industries like shipping, such parameters are published by market observers such as *Drewry* in their annual reports on container census and on carrier financials (Drewry Maritime Research, 2015a,b). These values are not only useful for the investment decisions at the current period, but also for planning future investments since shipping firms as well as firms in other industries often neglect potential changes in future parameters. This is to say, they presume  $\mathcal{K}_{\tau j} = \mathcal{K}_{tj}$ ,  $a_{\tau j} = a_{tj}$ ,  $H_{\tau j} = H_{tj}$ ,  $c_{\tau j} = c_{tj}$ , and  $r_{\tau j} = r_{tj}$  for  $\tau > t$ . This assumption is also in line with existing dynamic capacity models, which only allow univariate uncertainty (e.g., demand uncertainty). We define the investment cost function of firm  $j$  at period  $t$  as a kinked piece-wise linear function:  $C_{tj}(k_{tj}) = c_{tj} \times (k_{tj} - k_{t-1j})^+ - r_{tj} \times (k_{t-1j} - k_{tj})^+$  where  $(x)^+$  denotes  $\max\{0, x\}$ . As purchasing capital assets or technology is partially irreversible, we make the following assumption on the investment cost parameters:

ASSUMPTION 2. *Capacity investment is costly to reverse as  $c_{tj} > r_{tj}$ . In addition, the present value of a unit of used capacity cannot be higher than a new unit, i.e.,  $c_{tj} \geq \delta^{\tau-t} r_{\tau j}$  for  $\tau \in \{t, \dots, T\}$  where  $\delta > 0$  is the single-period discount factor.*

At the end of  $\Gamma$ , the salvage value function of firm  $j$  is  $F_j(k_{Tl}, k_{Tf}, \omega_{T+1})$ . Analogous to the operating profit function, we make the following assumption on the salvage value function:

ASSUMPTION 3. *For any given and fixed capacity of the other firm  $k_{Ti} \in \mathcal{K}_{Ti}$  ( $i \neq j$ ) and for each  $\omega_{T+1} \in \Theta$ , firm  $j$ 's salvage value function  $F_j(k_{Ti}, \cdot, \omega_{T+1})$  is concave in its own capacity  $k_{Tj}$ .*

Parameter	Description
$\Gamma$	set of time periods
$k_{tj}$	firm $j$ 's capacity at period $t$
$\mathcal{K}_{tj}$	set of capacity choices available to firm $j$ at period $t$
$\mathbf{K}_{tj}$	firm $j$ 's investment strategy vector from period $t$ to the end of $\Gamma$
$\mathcal{K}_{tj}$	set of all investment strategy vectors $\mathbf{K}_{tj}$
$\omega_t$	demand indicator of period $t$
$\Theta$	set of demand realizations
$Pr$	transition probability function of the demand
$\mathbf{Y}_{tj}$	state vector observed by firm $j$ at the beginning of period $t$
$q_{tj}, Q_{tj}$	firm $j$ 's production quantity and production quantity function at period $t$
$q_t$	total production quantity at period $t$
$p_t, P_t$	market price and price function of the product at period $t$
$h_{tj}, H_{tj}$	production cost and production cost function of firm $j$ at period $t$
$a_{tj}$	firm $j$ 's capacity utilization parameter at period $t$
$\pi_{tj}$	firm $j$ 's operating profit function at period $t$
$c_{tj}, r_{tj}$	firm $j$ 's marginal investment cost and marginal disinvestment revenue at period $t$
$C_{tj}$	investment cost function of firm $j$ at period $t$
$\delta$	single-period discount factor
$F_j$	salvage value function of firm $j$
$V_{tj}, V_{tj}^*$	firm $j$ 's value function and optimal value function at the beginning of period $t$
$S_{tj}$	firm $j$ 's stayput region at period $t$
$k_{tj}^L, K_{tj}^L$	lowerbound and lowerbound function of firm $j$ 's stayput region at period $t$
$k_{tj}^H, K_{tj}^H$	upperbound and upperbound function of firm $j$ 's stayput region at period $t$

Table 2 Model parameters

## 4. The optimal value function

### 4.1. The follower's value function

Let  $\mathbf{K}_{tj} = (k_{tj}, k_{t+1j}, \dots, k_{Tj})$  denote firm  $j$ 's investment strategy vector from period  $t$  to the end of  $\Gamma$  and  $\mathcal{K}_{tj}$  denote the set of all investment strategy vectors  $\mathbf{K}_{tj}$ . At the beginning of period  $t$ , given the state  $\mathbf{Y}_{tf} = (k_{tl}, k_{t-1f}, \omega_t)$ , and strategies  $\mathbf{K}_{t+1l}$  and  $\mathbf{K}_{tf}$  will be implemented, the follower's expected net present value (NPV) conditioned on the current demand is:

$$V_{tf}(k_{tl}, k_{t-1f}, \omega_t, \mathbf{K}_{t+1l}, \mathbf{K}_{tf}) = E \left[ \sum_{\tau=t}^T \delta^{\tau-t} (\pi_{\tau f}(k_{\tau l}, k_{\tau f}, \omega_{\tau}) - C_{\tau f}(k_{\tau f})) + \delta^{T+1-t} F_f(k_{Tl}, k_{Tf}, \omega_{T+1}) \mid \omega_t \right] \quad (2)$$

Here  $\mathbf{K}_{t+1l}$  is the follower's opinion of the leader's future strategy. In order to derive a structured optimal strategy, we assume that there is a stationary relationship between the follower's strategy  $\mathbf{K}_{tf}$  and his opinion of the leader's response strategy  $\mathbf{K}_{t+1l}$ . This is to say, for every  $k_{\tau f}$  in  $\mathbf{K}_{tf}$ , a value for  $k_{\tau+1l}$  is assumed according to a rule and this rule is consistent for  $\tau \in \{t, \dots, T-1\}$ . All existing oligopoly capacity models (e.g., Grenadier (2002); Novy-Marx (2007)) implicitly assume that both players know ex ante the opponent's optimal response to the player's own strategy.

This assumption may hold true for the leader, as he can exert some control over the market and thus know the follower's possible responses. However, the same assumption is not always applicable to the follower in a multi-round game. We relax the optimality assumption and allow the follower to either respond reactively to the competition or plan its investment proactively, taking into consideration the leader's possible future responses. In cases where the follower proactively considers the leader's strategy, we further elaborate on two situations: the follower has partial or full information on the leader's future strategy.

After specifying the case, we can omit  $\mathbf{K}_{t+1l}$  in the follower's value function, i.e.,  $V_{tf}(k_{tl}, k_{t-1f}, \omega_t, \mathbf{K}_{tf})$ . The follower's optimal value function at the beginning of period  $t$  is:

$$V_{tf}^*(k_{tl}, k_{t-1f}, \omega_t) = \sup_{\mathbf{K}_{tf} \in \mathcal{K}_{tf}} V_{tf}(k_{tl}, k_{t-1f}, \omega_t, \mathbf{K}_{tf}) \quad (3)$$

#### 4.2. The leader's value function

Similar to the follower, the leader can either adopt a reactive strategy or proactively plan his investments, considering the follower's responses. As the leader moves first, we assume that in the proactive case, he has full information on the follower's responses. Full information includes the follower's value function  $V_{tf}$ , as well as the follower's opinion of the leader's future strategy in  $V_{tl}$ . At the beginning of period  $t$ , given the state  $\mathbf{Y}_{tl} = (k_{t-1l}, k_{t-1f}, \omega_t)$  and a strategy  $\mathbf{K}_{tl}$  will be implemented, the leader's expected NPV conditioned on the current demand is:

$$V_{tl}(k_{t-1l}, k_{t-1f}, \omega_t, \mathbf{K}_{tl}) = E \left[ \sum_{\tau=t}^T \delta^{\tau-t} (\pi_{\tau l}(k_{\tau l}, k_{\tau f}, \omega_{\tau}) - C_{\tau l}(k_{\tau l})) + \delta^{T+1-t} F_l(k_{Tl}, k_{Tf}, \omega_{T+1}) \mid \omega_t \right] \quad (4)$$

In the reactive strategy,  $k_{\tau f} = k_{t-1f} \in \mathcal{K}_{t-1f}$ ,  $\forall \tau \in \{t, \dots, T\}$ <sup>5</sup>, whereas in the proactive strategy,  $k_{\tau f}$  is the first item in the vector  $\mathbf{K}_{\tau f}^*$  that is determined by  $V_{\tau f}^*(k_{\tau l}, k_{\tau-1f}, \omega_{\tau})$ .

The leader's optimal value function at the beginning of period  $t$  is:

$$V_{tl}^*(k_{t-1l}, k_{t-1f}, \omega_t) = \sup_{\mathbf{K}_{tl} \in \mathcal{K}_{tl}} V_{tl}(k_{t-1l}, k_{t-1f}, \omega_t, \mathbf{K}_{tl}) \quad (5)$$

Mixing the two players' strategies, four cases are considered: (a) the leader is proactive, while the follower reacts to the competition by assuming that the leader will *stay put* in the next period, i.e.,  $k_{\tau l} = k_{tl}$ ,  $\forall \tau \in \{t+1, \dots, T\}$  in equation (2); (b) both players are proactive, however, the follower only has partial information on the leader's strategy and assumes that the leader is *adversarial*, i.e.,  $k_{\tau l} = \arg \min_{k \in \mathcal{K}_{\tau l}} V_{\tau f}^*(k, k_{\tau-1f}, \omega_{\tau})$ ,  $\forall \tau \in \{t+1, \dots, T\}$  in equation (2); (c) both

<sup>5</sup> Instead of observing  $\mathcal{K}_{tf}$  at the beginning of period  $t$ , the leader in the reactive case assumes  $\mathcal{K}_{tf} = \mathcal{K}_{t-1f}$ .

players are proactive and the follower has full information on the leader's *optimal* strategy, i.e.,  $k_{\tau l} = \arg \max_{k \in \mathcal{K}_{\tau l}} V_{\tau l}^*(k_{\tau-1l}, k_{\tau-1f}, \omega_{\tau})$ ,  $\forall \tau \in \{t+1, \dots, T\}$  in equation (2); (d) both players are *reactive* by assuming that the other player will stay put in the next period. Case (b) corresponds to one of the situations where the follower is proactive, but guesses incorrectly about the leader's strategy. A case in which the leader is reactive and the follower is proactive is not mentioned here, as it is identical to case (a) with a delayed starting point (i.e., the follower moves first). We denote cases (c) and (d) as symmetric cases since both players adopt the same type of strategies and have the same amount of information on each other's strategy, whereas cases (a) and (b) are asymmetric as the situations are different for the two players. In any proactive case, equations (3) and (5) suffer the curse of dimensionality. Below, we use recursive optimality equations to get the optimal value and derive an ISD policy which determines the optimal action of the current period.

### 4.3. The optimality equations

The value of a long-term strategy can be broken into the value of the current action and the value of the future ones. According to Bellman's principle of optimality,  $V_{tj}$  (equations (3) and (5)) satisfy the following recursive optimality equations for  $t \in \Gamma$ :

At the beginning of period  $T+1$ , firm  $j$ 's optimal value associated with the state  $\mathbf{Y}_{T+1j} = (k_{Tl}, k_{Tf}, \omega_{T+1})$  is equal to its salvage value:

$$V_{T+1j}^*(k_{Tl}, k_{Tf}, \omega_{T+1}) = F_j(k_{Tl}, k_{Tf}, \omega_{T+1}) \quad (6)$$

At the beginning of period  $t$ , the follower's optimal value function associated with the state  $\mathbf{Y}_{tf} = (k_{tl}, k_{t-1f}, \omega_t)$  is:

$$V_{tf}^*(k_{tl}, k_{t-1f}, \omega_t) = \sup_{k_{tf} \in \mathcal{K}_{tf}} \left\{ \pi_{tf}(k_{tl}, k_{tf}, \omega_t) - C_{tf}(k_{tf}) + \delta E[V_{t+1f}^*(k_{t+1l}, k_{tf}, \omega_{t+1}) \mid \omega_t] \right\} \quad (7)$$

where  $k_{t+1l}$  depends on whether it is case (a) (or (d)), (b) or (c).<sup>6</sup>

At the beginning of period  $t$ , the leader's optimal value function associated with the state  $\mathbf{Y}_{tl} = (k_{t-1l}, k_{t-1f}, \omega_t)$  is:

$$V_{tl}^*(k_{t-1l}, k_{t-1f}, \omega_t) = \sup_{k_{tl} \in \mathcal{K}_{tl}} \left\{ \pi_{tl}(k_{tl}, k_{t-1f}, \omega_t) - C_{tl}(k_{tl}) + \delta E[V_{t+1l}^*(k_{tl}, k_{t-1f}, \omega_{t+1}) \mid \omega_t] \right\} \quad (8)$$

where  $k_{t-1f}$  depends on whether it is case (d) or the other cases.

<sup>6</sup> In the e-companion, we list the follower's optimal value functions in the four above-mentioned cases, respectively.

Without specifying the case here, we define a function  $G_{tj}$  as firm  $j$ 's expected NPV evaluated at period  $t$ , given that its capacity has been adjusted to  $k_{tj}$  and an optimal follow-up investment strategy will be implemented (i.e.,  $V_{t+1j}^*(\cdot)$ ). The follower's expected NPV at period  $t$  is:

$$G_{tf}(k_{tl}, k_{tf}, \omega_t) = \pi_{tf}(k_{tl}, k_{tf}, \omega_t) + \delta E[V_{t+1f}^*(k_{t+1l}, k_{t+1f}, \omega_{t+1}) | \omega_t] \quad (9)$$

The follower's optimization problem at period  $t$  then becomes the following:

$$V_{tf}^*(k_{tl}, k_{t-1f}, \omega_t) = \sup_{k_{tf} \in \mathcal{K}_{tf}} \left\{ G_{tf}(k_{tl}, k_{tf}, \omega_t) + r_{tf} \times (k_{t-1f} - k_{tf})^+ - c_{tf} \times (k_{tf} - k_{t-1f})^+ \right\} \quad (10)$$

The leader's expected NPV at period  $t$  is:

$$G_{tl}(k_{tl}, k_{t-1f}, \omega_t) = \pi_{tl}(k_{tl}, k_{t-1f}, \omega_t) + \delta E[V_{t+1l}^*(k_{t+1l}, k_{t+1f}, \omega_{t+1}) | \omega_t] \quad (11)$$

The leader's optimization problem at period  $t$  equals the following:

$$V_{tl}^*(k_{t-1l}, k_{t-1f}, \omega_t) = \sup_{k_{tl} \in \mathcal{K}_{tl}} \left\{ G_{tl}(k_{tl}, k_{t-1f}, \omega_t) + r_{tl} \times (k_{t-1l} - k_{tl})^+ - c_{tl} \times (k_{tl} - k_{t-1l})^+ \right\} \quad (12)$$

Eberly and Van Mieghem (1997) solve the optimization problem for a single-firm case. They show that if the optimal value function  $V^*$  is strictly concave, the optimal policy (also called as ISD policy) can be represented in the form of a unique stayput region, which is a continuum of optimal solutions to the investment problem. The boundaries of the stayput region define the decision rule for investments at each period: if capacity falls within the boundaries (i.e., inside the stayput region), it is optimal not to adjust capacity; otherwise, capacity should be adjusted to an appropriate point on the region's boundary. Extending their method to a two-firm setting, we first derive the ISD policy for a firm that behaves reactively in the competition. In cases where a proactive strategy is adopted, we compute the follower's ISD policy which can be seen as a function of the leader's first move and the current demand. We then obtain the leader's ISD policy which takes into account the follower's response policy.

## 5. The reactive ISD policy

In Proposition 1, we show that a player's optimal value function ( $V_{tf}^*$  or  $V_{tl}^*$ ) is concave if he adopts the reactive strategy (e.g., the follower in case (a) and both players in case (d)). This allows us to efficiently find an optimal solution to the investment problem. The proactive cases will be dealt with separately.

PROPOSITION 1. *If firm  $j$  adopts the reactive strategy, the optimal value function  $V_{tj}^*$  is jointly concave in  $(k_{t-1j}, k_{tj})$  for any given current capacity<sup>7</sup> of the other firm  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and for each  $\omega_t \in \Theta$ .*

*Proof.* The proof will be given in the e-companion to this paper.

In Theorem 1, we present firm  $j$ 's ISD policy at period  $t$  in cases where the optimal value function  $V_{tj}^*$  is jointly concave in  $(k_{t-1j}, k_{tj})$  for any given  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and for every  $\omega_t \in \Theta$ . According to Proposition 1, the follower in the reactive case (a) and both players in case (d) satisfy the condition, thus the corresponding ISD policy will take the form as shown in Theorem 1. Let  $\frac{\nabla_- G(x)}{\nabla x}$  and  $\frac{\nabla_+ G(x)}{\nabla x}$  denote the infimum of all left-sided difference quotients and the supremum of all right-sided difference quotients of the function  $G$  at the point  $x$ , i.e.,  $\frac{G(y)-G(x)}{y-x} \geq \frac{\nabla_- G(x)}{\nabla x}$ ,  $\forall y < x$  and  $\frac{G(y)-G(x)}{y-x} \leq \frac{\nabla_+ G(x)}{\nabla x}$ ,  $\forall y > x$ , where  $x$  and  $y$  are in the domain of  $G$ .  $\frac{\nabla_- G_{tj}(k_{tj})}{\nabla k_{tj}}$  and  $\frac{\nabla_+ G_{tj}(k_{tj})}{\nabla k_{tj}}$ <sup>8</sup> can be seen as firm  $j$ 's (minimal) marginal value of investment and (maximal) marginal value of disinvestment at capacity  $k_{tj}$ . Define the origin and end of the firm  $j$ 's capacity space as  $k_{jo} = \inf \mathcal{K}_{tj}$  and  $k_{je} = \sup \mathcal{K}_{tj}$ .

THEOREM 1. *Given the other firm's current capacity  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and  $\omega_t \in \Theta$ , if firm  $j$ 's optimal value function  $V_{tj}^*$  is jointly concave in  $(k_{t-1j}, k_{tj})$  and there exists a unique solution to the optimization problem in equation (10) if  $j = f$  or (12) if  $j = l$ , then the solution is an ISD policy which can be characterized by the following lowerbound and upperbound functions:*

$$K_{tj}^L(k_{ti}, \omega_t) = \sup \left\{ \{k_{jo}\} \cup \{k_{tj} : \frac{\nabla_- G_{tj}(k_{ti}, k_{tj}, \omega_t)}{\nabla k_{tj}} \geq c_{tj}, \quad k_{tj} \in \mathcal{K}_{tj} \} \right\} \quad (13)$$

$$K_{tj}^H(k_{ti}, \omega_t) = \inf \left\{ \{k_{je}\} \cup \{k_{tj} : \frac{\nabla_+ G_{tj}(k_{ti}, k_{tj}, \omega_t)}{\nabla k_{tj}} \leq r_{tj}, \quad k_{tj} \in \mathcal{K}_{tj} \} \right\} \quad (14)$$

*Proof.* If firm  $j$ 's optimization problem  $V_{tj}^*$  is jointly concave in  $(k_{t-1j}, k_{tj})$  for any given  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and for every  $\omega_t \in \Theta$ , the function  $G_{tj}(k_{ti}, k_{tj}, \omega_t)$  is concave in  $k_{tj}$  under Assumption 1 as a sum of concave functions is concave. The rest of the theorem then follows the reasoning of *Theorem 2* in Eberly and Van Mieghem (1997).

COROLLARY 1. *Let  $k_{tj}^L$  and  $k_{tj}^H$  denote the lowerbound and upperbound computed by the two boundary functions in Theorem 1, i.e.,  $k_{tj}^L = K_{tj}^L(k_{ti}, \omega_t)$  and  $k_{tj}^H = K_{tj}^H(k_{ti}, \omega_t)$ , then  $k_{tj}^L \leq k_{tj}^H$ .*

*Proof.* Under Assumption 2, which specifies  $r_{tj} < c_{tj}$ , and concavity of the optimization problem in equation (10), it guarantees that  $k_{tj}^L \leq k_{tj}^H$ .

<sup>7</sup> If  $j = l$ , the leader in the reactive case observes the follower's current capacity  $k_{t-1f} \in \mathcal{K}_{t-1f}$  and assumes  $k_{tf} = k_{t-1f}$ .

<sup>8</sup> We write  $G_{tj}(k)$  to represent  $G_{tf}(k_{ti \neq j}, k_{tj} = k, \omega_t)$ .

According to *Property 1* in Eberly and Van Mieghem (1997), the ISD policy defined in Theorem 1 can be expressed in terms of a set<sup>9</sup>  $S_{tj}(k_{ti}, \omega_t) \subset \mathbb{R}_{\geq 0}$  for each  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and each  $\omega_t \in \Theta$ :

$$S_{tj}(k_{ti}, \omega_t) = [k_{tj}^L, k_{tj}^H] = \left\{ k_{tj} : r_{tj} \leq \frac{\nabla_+ G_{tj}(k_{ti}, k_{tj}, \omega_t)}{\nabla k_{tj}} \wedge \frac{\nabla_- G_{tj}(k_{ti}, k_{tj}, \omega_t)}{\nabla k_{tj}} \leq c_{tj}, \quad k_{tj} \in \mathcal{K}_{tj} \right\} \quad (15)$$

At the beginning of period  $t$ , firm  $j$ 's investment decision rule indicated by  $S_{tj}(k_{ti}, \omega_t)$  is described as follows:

- if  $k_{t-1j} \in S_{tj}(k_{ti}, \omega_t)$ , no adjustment should be made, i.e.,  $k_{tj} = k_{t-1j}$ ;
- if  $k_{t-1j} \notin S_{tj}(k_{ti}, \omega_t)$  and  $k_{t-1j} < K_{tj}^L(k_{ti}, \omega_t)$ , an investment should be made such that the new capacity hits the boundary of  $S_{tj}(k_{ti}, \omega_t)$  at the lower side, i.e.,  $k_{tj} = K_{tj}^L(k_{ti}, \omega_t)$ ;
- if  $k_{t-1j} \notin S_{tj}(k_{ti}, \omega_t)$  and  $k_{t-1j} > K_{tj}^H(k_{ti}, \omega_t)$ , a disinvestment should be made such that the new capacity hits the boundary of  $S_{tj}(k_{ti}, \omega_t)$  at the higher side, i.e.,  $k_{tj} = K_{tj}^H(k_{ti}, \omega_t)$ .

The economic meaning behind the stayput region  $S_{tj}(k_{ti}, \omega_t)$  is: capacity should be adjusted when the marginal value of investment is larger than the marginal investment cost  $c_{tj}$  or the marginal value of disinvestment is smaller than the marginal disinvestment revenue  $r_{tj}$ . The marginal value of investment equals  $c_{tj}$  at the investment boundary, and the marginal value of disinvestment equals  $r_{tj}$  at the disinvestment boundary.

## 6. The proactive ISD policy

If firm  $j$  adopts a proactive strategy, the optimal value function  $V_{tj}^*$ , as well as the function  $G_{tj}$ , may not be concave in  $k_{tj}$ . Non-concavity of the optimization problem may arise from the fact that the firm in a proactive case strategically plans its investments, considering their impact on the other firm's present and future value. As a result, a proactive ISD policy may contain multiple stayput intervals at a period. Each interval indicates an optimal solution set of the firm, taking into account the impact of the other firm's responses on the firm's strategy. The more stayput intervals a proactive ISD policy contains, the more strategical flexibility the player has at that period. Depending on the player's current capacity and the proactive ISD policy, the optimal solution to the investment problem can then be determined.

### 6.1. The follower's proactive ISD policy

In a proactive case, the follower perceives that his current decision may influence the leader's future investments. Regardless of the accuracy of the follower's information on the leader's future strategy, if the follower expects that  $\frac{\nabla V_{t+1l}(k_{tl}, k_{tf}, \omega_{t+1})}{\nabla k_{tf}} \neq 0$ , he can strategically choose  $k_{tf}$  to influence the

<sup>9</sup> In the case where the capacity space is discrete, the stayput region is a subset containing a finite number of values in the interval  $[k_{tj}^L, k_{tj}^H]$ .

leader's future capacity  $k_{t+1}, \dots, k_{Tl}$ , with the purpose of maximizing his own total value of the present and future. We develop an *elimination algorithm* (Algorithm 1) to compute the follower's proactive ISD policy at period  $t$ . There are two steps in our algorithm. First, a unique range is identified, in which the stayput region is contained, and capacity values that fall outside of this range are excluded from the final solution set. Second, using two sequential rolling capacity procedures over the range computed at the first step, we further eliminate capacity values, of which the associated objective function in equation (10) can be improved. Given a capacity space  $\mathcal{X}_{tj} = [k_{jo}, k_{je}]$ , the range of the stayput region computed at the first step is denoted as  $S_{[k_{jo}, k_{je}]}^r$  and the final solution set at the second step, i.e., the stayput region, is denoted as  $S_{[k_{jo}, k_{je}]}$ .

To solve the follower's investment problem in equation (10), we first use the two boundary functions (equations (13) and (14)) in Theorem 1 to identify the range  $S_{[k_{fo}, k_{fe}]}^r$ , i.e.,  $S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ . Note that as  $G_{tf}$  is not necessarily concave,  $k_{tf}^L$  is not guaranteed to be smaller than or equal to  $k_{tf}^H$ . However, in Proposition 2 we show that the stayput region  $S_{[k_{fo}, k_{fe}]}$  is exclusively contained in the range  $S_{[k_{fo}, k_{fe}]}^r$ . Therefore,  $k_{tf}^L \leq k_{tf}^H$  ( $S_{[k_{fo}, k_{fe}]}^r \neq \emptyset$ ) if solutions exist to the optimization problem, i.e.,  $S_{[k_{fo}, k_{fe}]} \neq \emptyset$ . Such a condition is guaranteed if the capacity space  $\mathcal{X}_{tf}$  is nonempty and closed. Hereinafter, we assume the existence of an optimal solution  $k_{tf}^*$  (not necessarily unique). In Proposition 3, we show that the boundaries of the range  $S_{[k_{fo}, k_{fe}]}^r$  is contained in the final solution set, i.e.,  $k_{tf}^L, k_{tf}^H \in S_{[k_{fo}, k_{fe}]}$ . In the proofs of the propositions, we use the argument that a capacity value  $k$  is a non-stayput point (i.e.,  $k \notin S_{[k_{fo}, k_{fe}]}$ ) if there exists another value  $k' > k$  such that  $\frac{G_{tf}(k) - G_{tf}(k')}{k - k'} \geq c_{tf}$  or  $k' < k$  such that  $\frac{G_{tf}(k) - G_{tf}(k')}{k - k'} \leq r_{tf}$ . This means the value of the objective function in equation (10) can be increased by adjusting  $k$  to  $k'$ .

**PROPOSITION 2.**  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ .

*Proof.* The proof will be given in the e-companion to this paper.

**PROPOSITION 3.**  $k_{tf}^L$  and  $k_{tf}^H$  are the lowerbound and the upperbound of the stayput region  $S_{[k_{fo}, k_{fe}]}$ <sup>10</sup>.

*Proof.* The proof will be given in the e-companion to this paper.

After identifying the range  $S_{[k_{fo}, k_{fe}]}^r$ , two sequential rolling capacity procedures are implemented at the second step of the *elimination algorithm* to eliminate all remaining non-stayput points from  $S_{[k_{fo}, k_{fe}]}^r$ . Below, we define the two types of sequential rolling capacity procedures using an example of capacity interval  $[k_o, k_e]$ . Given  $[k_o, k_e]$  and a capacity value  $k_i \in [k_o, k_e]$ , denote  $\Delta k_i$  as the smallest possible value after  $k_i$ :  $\Delta k_i = \inf(k_i, k_e]$  if  $k_i < k_e$ ; and  $\Delta k_i = k_e$  if  $k_i = k_e$ . Denote  $\Delta^- k_i$  as the largest possible value before  $k_i$ :  $\Delta^- k_i = \sup[k_o, k_i)$  if  $k_i > k_o$ ; and  $\Delta^- k_i = k_o$  if  $k_i = k_o$ .

<sup>10</sup>  $S_{[k_{fo}, k_{fe}]}$  may contain several disconnected stayput intervals. Therefore,  $k_{tf}^L, k_{tf}^H$  can also be interpreted as the lowerbound of the first stayput interval and the upperbound of the last stayput interval.



**Rolling up procedure** Starting with the capacity interval origin  $k_o$ , the first iteration interval is set as  $\mathcal{K}_1 := [k_o, k_o]$  and the  $i$ th iteration interval is denoted as  $\mathcal{K}_i = [k_o, k_i]$  for  $i > 1$ . At each iteration  $i$ , we extend the previous iteration interval  $\mathcal{K}_{i-1}$  by adding the smallest possible value after  $k_{i-1}$  to the right end of the iteration interval, i.e.,  $k_i = \Delta k_{i-1}$ , until  $k_i = k_e$ <sup>11</sup>. Given each iteration interval  $\mathcal{K}_i = [k_o, k_i]$ , we identify the upperbound of the range  $S_{[k_o, k_i]}^r$  using the upperbound function (equation (14)) in Theorem 1: if  $k_i$  is not the upperbound of  $S_{[k_o, k_i]}^r$ , eliminate  $k_i$  from  $[k_o, k_e]$ ; otherwise, keep  $k_i$ .

**Rolling down procedure** Starting with the capacity interval end  $k_e$ , the first iteration interval is set as  $\mathcal{K}_1 := [k_e, k_e]$  and the  $i$ th iteration interval is denoted as  $\mathcal{K}_i = [k_i, k_e]$  for  $i > 1$ . At each iteration  $i$ , we extend the previous iteration interval  $\mathcal{K}_{i-1}$  by adding the largest possible value before  $k_{i-1}$  to the left end of the iteration interval, i.e.,  $k_i = \Delta^- k_{i-1}$ , until  $k_i = k_o$ . Given each iteration interval  $\mathcal{K}_i = [k_i, k_e]$ , we identify the lowerbound of the range  $S_{[k_i, k_e]}^r$  using the lowerbound function (equation (13)) in Theorem 1: if  $k_i$  is not the lowerbound of  $S_{[k_i, k_e]}^r$ , eliminate  $k_i$  from  $[k_o, k_e]$ ; otherwise, keep  $k_i$ .

Starting with the origin of the range  $S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ , we first implement the *rolling up procedure*. This procedure is finished once the upperbound of the range  $S_{[k_{tf}^L, k_{tf}^H]}^r$  is checked. The set of remaining capacity values after the *rolling up procedure* is denoted as  $S_{[k_{fo}, k_{fe}]}^{r1}$ , which may consist of several disconnected intervals. According to Proposition 3,  $k_{tf}^L, k_{tf}^H \in S_{[k_{fo}, k_{fe}]}^r$ . Hence,  $k_{tf}^L, k_{tf}^H \in S_{[a, b]}$ ,  $\forall [a, b] \subset [k_{fo}, k_{fe}]$ . This means that if a capacity value  $k$  is a stayput point in a space  $\mathcal{K}$ ,  $k$  is also a stayput point in any space that contains  $k$  and is a subset of the original space  $\mathcal{K}$ . As a result,  $k_{tf}^L$  and  $k_{tf}^H$  will not be eliminated after the *rolling up procedure*. We then implement the *rolling down procedure*, starting with the end point in the set  $S_{[k_{fo}, k_{fe}]}^{r1}$ , i.e.,  $k_{tf}^H$ . This procedure is finished once the lowerbound of the range  $S_{[k_{tf}^L, k_{tf}^H]}^r$  is checked. In Proposition 4 we show that the remaining capacity values after the *rolling up procedure* and the *rolling down procedure* compose the stayput region  $S_{[k_{fo}, k_{fe}]}^r$ . In Theorem 2, we outline the elimination method for computing the follower's proactive ISD policy.

**PROPOSITION 4.** *The rolling up and rolling down procedures eliminate all the non-stayput capacity values from the range  $S_{[k_{fo}, k_{fe}]}^r$ .*

*Proof.* The proof will be given in the e-companion to this paper.

**THEOREM 2.** *Given  $k_{tl} \in \mathcal{K}_{tl}$  and  $\omega_t \in \Theta$ , the solution to the follower's optimization problem at period  $t$  (see equation (10)) can be represented in the form of an ISD policy that may contain multiple disconnected stayput intervals. The set of the stayput intervals, i.e., the stayput region  $S_{tf}(k_{tl}, \omega_t)$ , can be derived by Algorithm 1.*

<sup>11</sup> A point of convergence can occur because the capacity space is discrete.

*Proof.* The first part of theorem follows directly as the optimal solution to the follower's investment problem at period  $t$  can be expressed by a set of capacity values, of which the associated value of equation (10) cannot be improved. According to Propositions 2, 3 and 4, the set of capacity values computed by Algorithm 1 is the follower's stayput region at period  $t$ . For consistency purposes, we denote the stayput region as  $S_{tf}(k_{tl}, \omega_t)$ , i.e.,  $S_{tf}(k_{tl}, \omega_t) = S_{[k_{fo}, k_{fe}]}$ .

---

**Algorithm 1** Elimination algorithm for computing the follower's stayput region

---

1. Compute  $k_{tf}^L$  and  $k_{tf}^H$  using equations (13) and (14). If solutions exist to the follower's optimization problem in equation (12), then  $k_{tf}^L \leq k_{tf}^H$ . The range  $S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$  contains the stayput region, i.e.,  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r$ .
2. Use two sequential rolling procedures over the range  $S_{[k_{fo}, k_{fe}]}^r$ :
  - (a) First implement the *rolling up procedure* in  $S_{[k_{fo}, k_{fe}]}^r$ . This eliminates capacity values of which the associated objective function in equation (12) can be improved. Denote the set of remaining capacity as  $S_{[k_{fo}, k_{fe}]}^{r1}$ ;
  - (b) Second implement the *rolling down procedure* in  $S_{[k_{fo}, k_{fe}]}^{r1}$ . This eliminates capacity values of which the associated objective function in equation (12) can be improved.

The remaining capacity values compose the stayput region  $S_{[k_{fo}, k_{fe}]}$ , i.e.,  $S_{tf}(k_{tl}, \omega_t)$ .

---

In Proposition 5 we prove that the optimal investment policy indicated by multiple disconnected stayput intervals assigns a non-stayput capacity value to the closest boundary of a close-by stayput interval. In Corollary 2, we show a decision rule, which uses a unique capacity value  $k$ , on how to adjust a non-stayput capacity that falls between two consecutive stayput intervals. A binary search can be used to identify this value  $k$ , which we refer to as the investment threshold between the two consecutive stayput intervals. In the case where  $Threshold(k, k^{H1}, k^{L2}) = 0$ , there is no difference between adjusting  $k$  to  $k^{H1}$  and adjusting  $k$  to  $k^{L2}$ ; for consistency, we select  $k$  as the threshold. If  $Threshold(k, k^{H1}, k^{L2}) < 0$ ,  $\forall k \in (k^{H1}, k^{L2})$ , we specify  $k^{L2}$  to be the threshold, indicating all capacity in  $(k^{H1}, k^{L2})$  should be adjusted to  $k^{H1}$ . If  $Threshold(k, k^{H1}, k^{L2}) > 0$ ,  $\forall k \in (k^{H1}, k^{L2})$ , we specify  $k^{H1}$  to be the threshold, indicating all capacity in  $(k^{H1}, k^{L2})$  should be adjusted to  $k^{L2}$ .

**PROPOSITION 5.** *Given  $k_{t-1f} = k$  and  $k \notin S_{tf}(k_{tl}, \omega_t)$ , if the optimal investment policy indicated by  $S_{tf}(k_{tl}, \omega_t)$  assigns  $k_{tf} = b$ , then no stayput values exist in the interval  $[k, b)$  for the case  $b > k$ , or in the interval  $(b, k]$  for the case  $b < k$ . In other words, the interval  $[k, b)$  or  $(b, k] \not\subseteq S_{tf}(k_{tl}, \omega_t)$ .*

*Proof.* The proof will be given in the e-companion to this paper.

**COROLLARY 2.** *Between any two consecutive stayput intervals (e.g.,  $[k^{L1}, k^{H1}]$  and  $[k^{L2}, k^{H2}]$ ) in the stayput region  $S_{tf}(k_{tl}, \omega_t)$ , a unique capacity value  $k$  exists (i.e.,  $k^{H1} < k < k^{L2}$ ) such that the*

optimal investment policy assigns all capacity in  $(k^{H1}, k]$  to be adjusted to the upperbound of the lower stayput interval (i.e.,  $k^{H1}$ ) and assigns all capacity in  $(k, k^{L2})$  to be adjusted to the lowerbound of the higher stayput interval (i.e.,  $k^{L2}$ ).

*Proof.* For any non-stayput capacity value  $k$  that is between two stayput intervals,  $[k^{L1}, k^{H1}]$  and  $[k^{L2}, k^{H2}]$ , i.e.,  $k^{H1} < k < k^{L2}$ , an adjustment should be made either to  $k^{L2}$  or  $k^{H1}$  according to Proposition 5. By comparing the expected NPV of adjusting  $k$  to  $k^{H1}$  with the one of adjusting  $k$  to  $k^{L2}$ , a decision whether to invest or disinvest can be made. Under Assumption 2, which specifies  $r_{tj} < c_{tj}$ , the function  $Threshold(k, k^{H1}, k^{L2}) = G_{tf}(k^{L2}) - c_{tf} \times (k^{L2} - k) - G_{tf}(k^{H1}) - r_{tf} \times (k - k^{H1})$  is monotonously increasing in  $k$ . The rest of corollary then follows.

## 6.2. The leader's proactive ISD policy

Similar to the follower's value functions in cases (b) and (c), the leader's optimal value function  $V_{tl}^*$  and the function  $G_{tl}$  in a proactive case, e.g., cases (a), (b) and (c), may not be concave in  $k_{tl}$ . This is because the leader's first move influences the follower's responses at the current period and in the future (see equation (7)). Hence, the leader can strategically choose  $k_{tl}$  in order to maximize his own value, which depends on the follower's current and future capacity (see equation (8)). We develop a *decomposition algorithm* (Algorithm 2) to compute the leader's proactive ISD policy. The mechanism of this type of algorithm is to "divide and conquer" (see also Groenevelt (1991)). Since the follower's optimal response at the current period,  $k_{tf}^*$ , is determined by a function of the leader's first move (i.e.,  $S_{tf}(k_{tl}, \omega_t)$ ), we divide the leader's capacity space into separate intervals, each of which is associated with a value of  $k_{tf}^*$ . Within the same interval where the value of  $k_{tf}^*$  is fixed, the leader can still choose different actions to influence the follower's future capacity  $k_{t+1f}, \dots, k_{Tf}$ , which influence the leader's own future value. The *elimination algorithm* designed in Section 6.1 determines firm  $j$ 's stayput region at period  $t$  in a situation where given a fixed current action from the opponent, firm  $j$  can influence the opponent's future capacity by choosing  $k_{tj}$ . Thus, after dividing the capacity space into intervals, the *elimination algorithm* is used to identify the stayput region in each interval.

Our *decomposition algorithm* contains three steps. The first step is to compute the range  $S_{[k_{lo}, k_{le}]}^r$  which contains the stayput region by using the same method as the first step of Algorithm 1. As the leader's capacity space is nonempty and closed, the existence of an optimal solution to the leader's optimization problem in (12) is guaranteed. Hence, the range  $S_{[k_{lo}, k_{le}]}^r$  exists and the boundaries of the range are denoted as  $k_{tl}^L$  and  $k_{tl}^H$ , i.e.,  $S_{[k_{lo}, k_{le}]}^r = [k_{tl}^L, k_{tl}^H]$ . We then divide the range  $S_{[k_{lo}, k_{le}]}^r$  into separate *terminal intervals*, each of which ends at a *terminal*. Denote the start of the  $n$ th *terminal interval* as  $I_o^n$  and the  $n$ th *terminal* as  $I_e^n$ . The  $n$ th *terminal interval* is then  $[I_o^n, I_e^n]$ . Below, we describe how *terminals* are identified and how *terminal intervals* are determined.

Similar to the above-mentioned *rolling up procedure*, we roll up the capacity domain in order to sequentially identify all the *terminals*. Let a function  $K_{tf}(k_{tl}, k_{t-1f}, \omega_t)$  represent the optimal investment policy indicated by  $k_{t-1f}$  and  $S_{tf}(k_{tl}, \omega_t)$ , i.e.,  $k_{tf}^* = K_{tf}(k_{tl}, k_{t-1f}, \omega_t)$ . The first *terminal interval* starts with the origin of  $[k_{tl}^L, k_{tl}^H]$ , i.e.,  $I_o^1 = k_{tl}^L$ . If there exists  $k$  such that  $k = \inf\{k_{tl} : K_{tf}(k_{tl}, k_{t-1f}, \omega_t) \neq K_{tf}(\Delta k_{tl}, k_{t-1f}, \omega_t), I_o^n \leq k_{tl} \leq k_{tl}^H\}$ , then the  $n$ th *terminal* is  $I_e^n = k$ ; otherwise,  $I_e^n = k_{tl}^H$ . If  $I_e^1 \neq k_{tl}^H$ , the  $n$ th *terminal interval* starts with  $I_o^n = \inf(I_e^{n-1}, k_{tl}^H)$  for  $n > 1$ . The searching process for *terminals* is continued until  $I_e^n = k_{tl}^H$ . In order to explain the next steps in our *decomposition algorithm*, we assume that at the first step  $N$  *terminals* are identified in the range  $[k_{tl}^L, k_{tl}^H]$  and the range is thus divided into  $N$  *terminal intervals*.

The second step of our *decomposition algorithm* is to determine the stayput region in each *terminal interval* by using Algorithm 1. Denote the stayput region in the  $n$ th *terminal interval*,  $[I_o^n, I_e^n]$  as  $S_n$  and the origin and end of  $S_n$  as  $k_o^{s_n}$  and  $k_e^{s_n}$ , i.e.,  $S_n = [k_o^{s_n}, k_e^{s_n}]$ .  $S_n$  represents the leader's optimal solution set in the capacity interval  $[I_o^n, I_e^n]$  where the follower's action at the current period is fixed, i.e.,  $K_{tf}(k, k_{t-1f}, \omega_t)$  gives the same value  $\forall k \in [I_o^n, I_e^n]$ . If the value of the objective function in equation (12) can be improved by adjusting a capacity value  $k$  in  $S_n$  to a value in  $S_{i \neq n}$ , then it is said that  $k$  is a non-stayput value in the capacity space, i.e.,  $k \notin S_{[k_{io}, k_{ie}]}$ ; if  $k$  does not need to be adjusted to any value in  $S_i$ ,  $\forall i \neq n$ , then  $k \in S_{[k_{io}, k_{ie}]}$ .

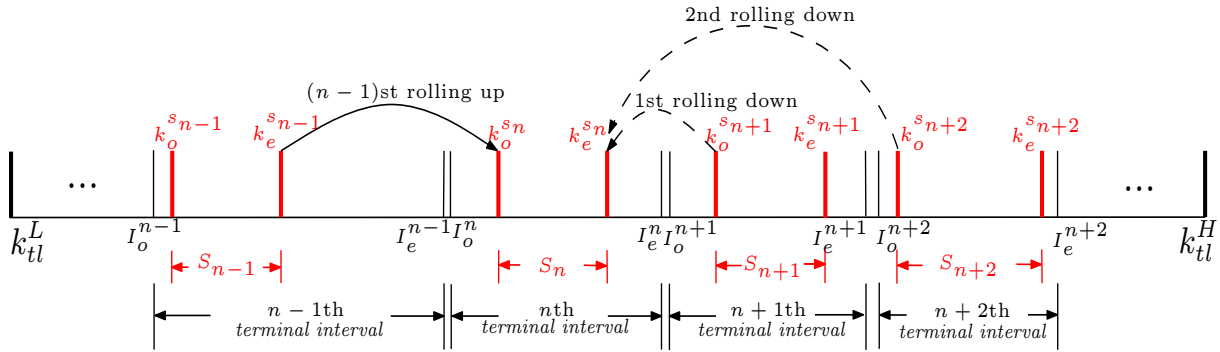


Figure 1 Illustration of the cross-interval comparison

Using cross-interval comparisons, the third step of our *decomposition algorithm* is to eliminate all the non-stayput values from  $S_n$ ,  $\forall n = 1, \dots, N$ . For a capacity  $k \in S_n$ , Proposition 6 shows that whether  $k \in S_{[k_{io}, k_{ie}]}$  can be efficiently determined by comparing  $k$  with the upperbound of  $S_i$ ,  $\forall i = 1, \dots, n-1$  and with the lowerbound of  $S_i$ ,  $\forall i = n+1, \dots, N$ . Thus, we apply the *rolling up procedure*  $n-1$  times for the upperbounds (see black solid lines in Figure 1) and apply the *rolling down procedure*  $N-n$  times for the lowerbounds (see black dash lines in Figure 1). For  $i = 1, \dots, n-1$ , the  $i$ th *rolling up procedure* starts with the end of  $S_i$  (i.e.,  $k_e^{s_i}$ ) and stops at the  $j$ th iteration if  $k_j \in S_{[k_e^{s_i}, k_j]}$ , where  $[k_e^{s_i}, k_j]$  is the  $j$ th iteration interval. For  $i = n+1, \dots, N$ , the

$(i-n)$ th procedure starts with the start of  $S_i$  (i.e.,  $k_o^{s_i}$ ) and stops at the  $j$ th iteration if  $k_j \in S_{[k_j, k_o^{s_i}]}^r$ , where  $[k_j, k_o^{s_i}]$  is the  $j$ th iteration interval. Proposition 7 shows that these two stopping rules are sufficient for eliminating non-stayput points in  $S_n$ . After cross-interval comparisons, no capacity values need to be adjusted. The remaining capacity values in each *terminal interval* compose the leader's stayput region, and we write it as a function of the follower's capacity and his ISD policy at the beginning of period  $t$ :  $S_{tl}(S_{tf}, k_{t-1f}, \omega_t)$ . In Theorem 3, we outline the decomposition method for computing the leader's proactive ISD policy.

**PROPOSITION 6.** *Given the stayput region in the  $i$ th terminal interval, i.e.,  $S_i = [k_o^{s_i}, k_e^{s_i}]$ , if there exists a capacity value  $k_1 > k_e^{s_i}$  satisfying  $k_1 \in S_{[k_e^{s_i}, k_1]}^r$  (or a capacity value  $k_1 < k_o^{s_i}$  satisfying  $k_1 \in S_{[k_1, k_o^{s_i}]}^r$ ), then  $k_1 \in S_{[k_2, k_1]}$  (or  $k_1 \in S_{[k_1, k_2]}$ ),  $\forall k_2 \in S_i$ .*

*Proof.* The proof will be given in the e-companion to this paper.

**PROPOSITION 7.** *Given the stayput region in the  $i$ th terminal interval, i.e.,  $S_i$ , if a capacity value  $k_1 \in S_i$  satisfying the following:  $k_1 \in S_{[k, k_1]}^r$ ,  $\exists k < k_1$  (or  $k_1 \in S_{[k_1, k]}^r$ ,  $\exists k > k_1$ ), then we have:  $k_2 \in S_{[k_1, k_2]}$ ,  $\forall \{k_2 : k_1 < k_2 \in S_n\}$  (or  $k_2 \in S_{[k_2, k_1]}$ ,  $\forall \{k_2 : k_2 < k_1 \wedge k_2 \in S_n\}$ ).*

*Proof.* The proof will be given in the e-companion to this paper.

**THEOREM 3.** *Given  $S_{tf}$ ,  $k_{t-1f} \in \mathcal{K}_{t-1f}$ , and  $\omega_t \in \Theta$ , the solution to the leader's optimization problem at period  $t$  (see equation (12)) can be represented in the form of an ISD policy that may contain multiple disconnected stayput intervals. The set of the stayput intervals, i.e., the stayput region  $S_{tl}(S_{tf}, k_{t-1f}, \omega_t)$ , can be derived by Algorithm 2.*

*Proof.* The first part of theorem follows directly as the optimal solution to the leader's investment problem at period  $t$  can be expressed by a set of capacity values, of which the associated value of equation (12) cannot be improved. According to Theorem 2, Propositions 6 and 7, the set of capacity values computed by Algorithm 2 is the leader's stayput region at period  $t$ .

---

**Algorithm 2** Decomposition algorithm for computing the leader's stayput region

---

1. Compute the range  $S_{[k_{lo}, k_{le}]}^r$  of the stayput region using the same method as the first step of Algorithm 1:  $S_{[k_{lo}, k_{le}]} \subseteq S_{[k_{lo}, k_{le}]}^r = [k_{tl}^L, k_{tl}^H]$ . Next, identify *terminals* in the range  $[k_{tl}^L, k_{tl}^H]$  and divide the range into separate *terminal intervals*. Assume that there are  $N$  *terminals* in  $[k_{tl}^L, k_{tl}^H]$  and the range is divided into  $N$  *terminal intervals*.
  2. Use Algorithm 1 to search for the stayput region in each *terminal interval*. Denote the stayput region in the  $n$ th *terminal interval* as  $S_n$ .
  3. Eliminate all the non-stayput values from  $S_n$ ,  $\forall n = 1, \dots, N$  through cross-interval comparisons. The remaining capacity values constitute the leader's stayput region  $S_{[k_{lo}, k_{le}]}^r$ , i.e.,  $S_{tl}(S_{tf}, k_{t-1f}, \omega_t)$ .
-

## 7. Case study on the container shipping market

We first use data from the container shipping market to determine the optimal ISD investment strategies for a leader and a follower in all four cases (i.e., (a) *stay put*, (b) *adversarial*, (c) *optimal*, and (d) *reactive*). Abbreviations are used to refer to a player in a specific case, e.g., the leader in case *stay put* is abbreviated as *SL*. The results are then compared with the realized capacity investments of the two players in practice. Second, we study the underlying structures of the optimal ISD strategies between different players and cases. Third, using Monte Carlo simulations, we investigate the impact of different market and investment cost parameters on capacity strategies. Lastly, based on our results, we draw several implications for investors in a competitive market.

### 7.1. Optimal ISD investment strategies in the container shipping market

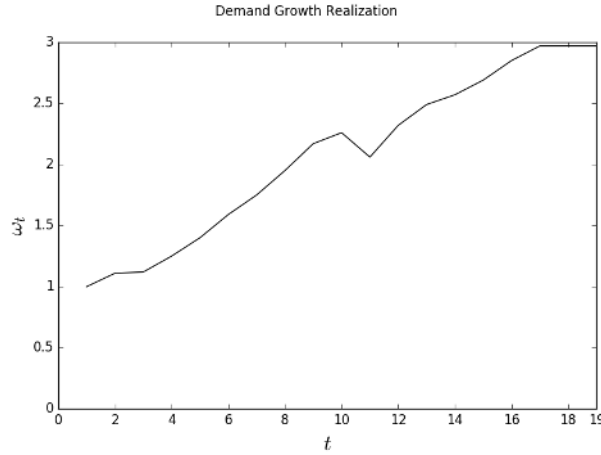
We use the demand and supply data of the container shipping market over a timespan of 16 years (2000-2015). To avoid the border effect in a long-term strategy, we extend the timespan to 2017, i.e.,  $\Gamma = \{1, \dots, 18\}$ . Maersk and MSC, which are currently ranked as the first and second liner operators based on their fleet capacity (Alphaliner, 2015), are chosen as the leader and the follower. At the beginning of each year, firms observe the current demand and predict the demand growth of the coming year. In order to mitigate the risk, a demand forecast often consists of several future scenarios with different probabilities, each scenario taking into account forecast errors to a varying degree. We use a categorical distribution as an approximation of a normal distribution to represent the demand forecast. At the beginning of 2000, the first observed demand is set as 1.  $\omega_{t+1}, \forall t \in \Gamma$ , evolves according to the following transition rule:

$$\omega_{t+1} = \begin{cases} \omega_t(1 + \mu_t + 2\sigma_t) & \text{with } pr = 2.28\% \\ \omega_t(1 + \mu_t + \sigma_t) & \text{with } pr = 13.59\% \\ \omega_t(1 + \mu_t) & \text{with } pr = 68.26\% \\ \omega_t(1 + \mu_t - \sigma_t) & \text{with } pr = 13.59\% \\ \omega_t(1 + \mu_t - 2\sigma_t) & \text{with } pr = 2.28\% \end{cases}$$

where  $\mu_t$  is the industry-wide accepted forecast of demand growth at period  $t$  and  $\sigma_t$  is the forecast error at period  $t$ , which is represented by the average discrepancy between all previous forecasts  $\mu_\tau$  and the realized demand growths  $\mu'_\tau, \forall \tau = 1, \dots, t$ . At each period  $t \in \Gamma$ , the values of  $\mu_t$  and  $\mu'_t$  are the average forecast and the average realization, which are based on the two half-yearly *Clarkson Shipping Review and outlook*<sup>12</sup> (Clarkson Research Services, 2000-2015a, 2000-2015b). Note that the forecast error is measured on realized demand growths in practice, rather than on simulated growths  $\omega_\tau$ . A maximal forecast deviation in the container shipping market, 5%, is chosen as the first value in  $\sigma = (\sigma_1, \dots, \sigma_{18})$ . The transition probabilities, i.e.,  $pr$ , are taken from the Z-score

<sup>12</sup> We use the same forecast of 2015 for 2016 and 2017, i.e.,  $\mu_{18} = \mu_{17} = \mu_{16}$ , and assume an zero demand growth in 2016 and 2017, i.e.,  $\mu'_{18} = \mu'_{17} = \mu'_{16}$ .

table. A large sample of random demand paths is generated and Figure 2 shows an example that closely matches realized demand growths in practice, i.e.,  $(\mu'_1, \dots, \mu'_{18})$ .



**Figure 2** A simulated demand growth path

Supply and investment cost parameters are kept constant throughout the whole timespan. Discrete market share values are used to represent the capacity choices available to firm  $j$  at period  $t$  and thus firms' capacity decisions can be interpreted as changes in their respective market shares. An upper limit of 19 is set to both firms' capacity space, due to the fact that to date, a liner operator's market share (based on its fleet capacity) has never exceeded 19% (Alphaliner, 2015). At each period  $t \in \Gamma$ , both firms have the same set of available capacity choices:  $\mathcal{K}_{tl} = \mathcal{K}_{tf} = \{0, 1, \dots, 19\}$ . At period  $t \in \{1, \dots, 16\}$  (i.e., 2000-2015), the realized capacity  $k'_{tj}$  of Maersk and MSC are the end-of-year capacity, which are extracted from *UNCTAD Review of Maritime Transport* (UNCTAD, 2001-2015) and from Alphaliner (2015). At the beginning of 2000, Maersk and MSC owned around 12% and 5% market share, i.e.,  $k_{0l} = k'_{0l} = 12$  and  $k_{0f} = k'_{0f} = 5$  (UNCTAD, 2000).

Assume for simplicity that at each period both firms' capacity utilization parameters equal 1, i.e.,  $a_{tj} = 1$ , and utilization cost functions  $H_{tj}(\cdot)$  take the following form: if  $a_{tj}q_{tj} \leq k_{tj}$ ,  $H_{tj}(q_{tj}, k_{tj}) = 0$ ; otherwise,  $H_{tj}(q_{tj}, k_{tj}) = \infty$ ,  $\forall j \in \{l, f\}, \forall t \in \Gamma$ . In cases where  $h_{tj} = H_{tj}(\cdot) = 0$ , the product price  $p_t$  can be seen as the marginal profit of a unit capacity (see equation (1)). As the price of shipping services fluctuates heavily with the firms' capacity investment, a linear market-clearing price is used, i.e.,  $q_{tj} = k_{tj}$  and  $P_t(k_{tl}, k_{tf}, \omega_t) = \alpha \cdot \omega_t - k_{tl} - k_{tf}$  where  $\alpha > 0$  is a given market parameter. As  $a_{tj} = 1$  and  $q_{tj} = k_{tj}$ ,  $H_{tj}(q_{tj}, k_{tj}) = 0$ . Therefore, firm  $j$ 's operating profit function is:

$$\pi_{tj}(k_{tl}, k_{tf}, \omega_t) = (\alpha \cdot \omega_t - k_{tl} - k_{tf}) \times k_{tj} \quad (16)$$

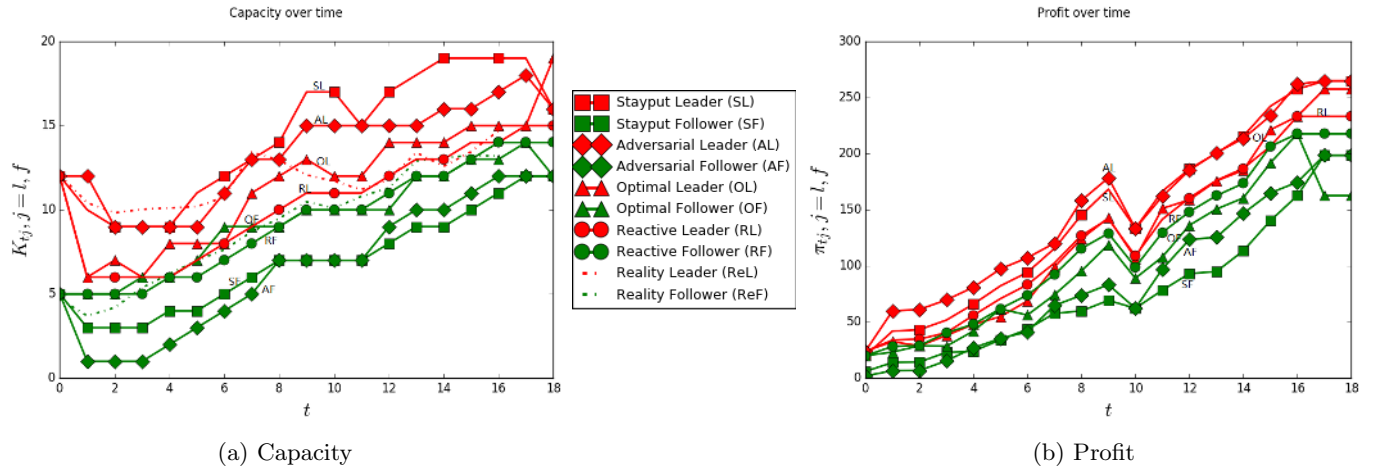
We use the same form for the salvage value function, i.e.,  $F_j(k_{Tl}, k_{Tf}, \omega_{T+1}) = (\alpha \cdot \omega_{T+1} - k_{Tl} - k_{Tf}) \times k_{Tj}$ . Assumptions 1 and 3 are satisfied with this example of profit functions.. In addition, as

$P_t(\cdot) < 0$  is possible, a firm can have negative profits for a period. This is consistent with common practice in the shipping industry, i.e., in order to remain active in the market, shipping lines continue to operate even if they are facing a loss. With this profit function, a firm benefits from all its capacity when the marginal profit of a unit capacity is positive and suffers losses from all its capacity when supply in the market exceeds demand (i.e., the marginal profit is negative). The  $\alpha$  value is determined using the historical freight rates (dollars per TEU), demand and supply data (both in thousand TEUs) on the three major liner trade routes, i.e., transpacific, Europe-Asia, and transatlantic. Assuming the linear market-clearing price, at the beginning of 2000 the marginal demand impact on price ranged from 1.2 to 2.1 on the three routes (UNCTAD, 2001). Since we use single-digit demand values (e.g.,  $\omega_1 = 1$ ) and double-digit capacity choices (e.g.,  $k_{0l} = 12$ ), we scale the marginal demand effect by multiplying 10, i.e.,  $\alpha = 15$ .

Based on the latest released vessel prices in *Maersk's container market weekly report* (Maersk, 2015), which states that the average second-hand vessel price is \$4,837 per TEU and the average newbuilding vessel price is \$10,741 per TEU, both firms' investment cost parameters are set as  $c_{tl} = c_{tf} = 10.7$  and  $r_{tl} = r_{tf} = 4.8$ . Notice that these are not the prices of changing 1% market share, however, the difference between  $c$  and  $r$  represents the investment irreversibility level in the container shipping market. Using Monte Carlo simulations, we will further investigate the impact of different investment irreversibility levels on the optimal strategies in Section 7.3. We adopt a discount rate of 0.89 as this value is frequently used by shipping firms to calculate the present value of future earnings (Gullaksen, 2012; Greenwood and Hanson, 2013).

The leader's and follower's optimal value functions (see equations (7) and (8)) can be computed using a backward induction method. The optimal capacity at each period are then determined based on the respective ISD policy as discussed in Sections 5 and 6. As  $\pi_{tf}$  in equation (16) is nonincreasing in the leader's capacity  $k_{tl}$  for any given  $k_{tf} \in \mathcal{K}_{tf}$  and for each  $\omega_t \in \Theta$ . It is easy to show by backwards induction that the follower's value function in case (b) adversarial is jointly concave in  $(k_{t-1f}, k_{tf})$ . Thus, the follower's ISD policy in case (b) can be computed in the same way as the reactive ISD policy using Theorem 1. Using the same demand path as shown in Figure 2, we compute the optimal capacity and profits of the four cases in Figure 3. The realized capacity investments of the two liner operators are represented by the red and green dash lines in Figure 3a. Considering value and adjustment pattern, the leader's realized capacity come close to  $AL$  from period 1 to 8 and close to  $OL$  and  $RL$  from period 9 to the end, while the follower's realized capacity are best matched by  $OF$  and  $RF$ . At the end of 2015, the two liner operators had 14.7% and 13.2% market share, which are exactly matched by  $OL$  and  $OF$ . While the dominant players followed their respective competitive investment strategies, other players were reducing their capacity. This is indeed what happened in practice: many small shipping firms have left the market (add ref).





**Figure 3** Optimal capacity and profit of the leader and the follower

## 7.2. Underlying structures of the ISD strategies

Using Algorithms 1 and 2 for *OF* and *OL* or Theorem 1 for *SF*, *AF*, *RF* and *RL*, we reveal the underlying ISD policies of the leader and follower’s optimal capacity and explain the differences between the four cases. With the black, red and green whisker lines indicating the first, second, and third stayput intervals (if they exist), respectively, and the purple dots indicating the investment threshold between two consecutive intervals, Figure 4 shows the stayput regions of the leader and the follower in all four cases. In each subfigure, the red solid line depicts the leader or follower’s optimal capacity over time. Taking Figure 4a as an example, at the beginning of period 4 with  $k_{3l} = 9$ , the stayput region at this period consists of two intervals: [9] and [11]. According to the ISD policy, it is optimal for the leader to stay put at his current capacity, i.e.,  $k_{4l}^* = 9$ .

First, as a result of the first mover advantage, the leader always acquires a same or higher capacity than the follower in all four cases, except *OL* at period 6 (see Figure 3a). This is justified by the observation in Figure 4: the leader’s stayput regions are usually higher than the follower’s ones. In this example of demand path, higher capacity leads to more profits: the leader gains more profits than the follower, except *OL* at period 5 (see Figure 3b).

Second, comparing a player’s capacity and profits in the symmetric cases (i.e., *optimal* and *reactive*) with those in the asymmetric cases (i.e., *stay put* and *adversarial*), *OL* and *RL* have lower capacity and profits than *SL* and *AL*, while *OF* and *RF* have higher capacity and profits than *SF* and *AF* in Figure 3. As shown in Figure 4, a player’s ISD policy at a period can contain multiple intervals in cases where he proactively considers the other player’s strategy (e.g., *SL*, *AL*, *OL* and *OF*). Multiple stayput intervals suggest more flexibility in a player’s strategy and thus a higher probability that his capacity needs to be adjusted. In the asymmetric cases, the follower’s ISD policies only contain a single interval at each period. It further gives the leader advantages

and allows the ISD policies of  $SL$  and  $AL$  to have multiple stayput intervals more frequently than the ISD policies of  $OL$ . This explains the difference between the leader's capacity and profits in symmetric and those in asymmetric cases. The follower favors the symmetric cases as both players in the same symmetric case have similar ISD policies in terms of the number of intervals.

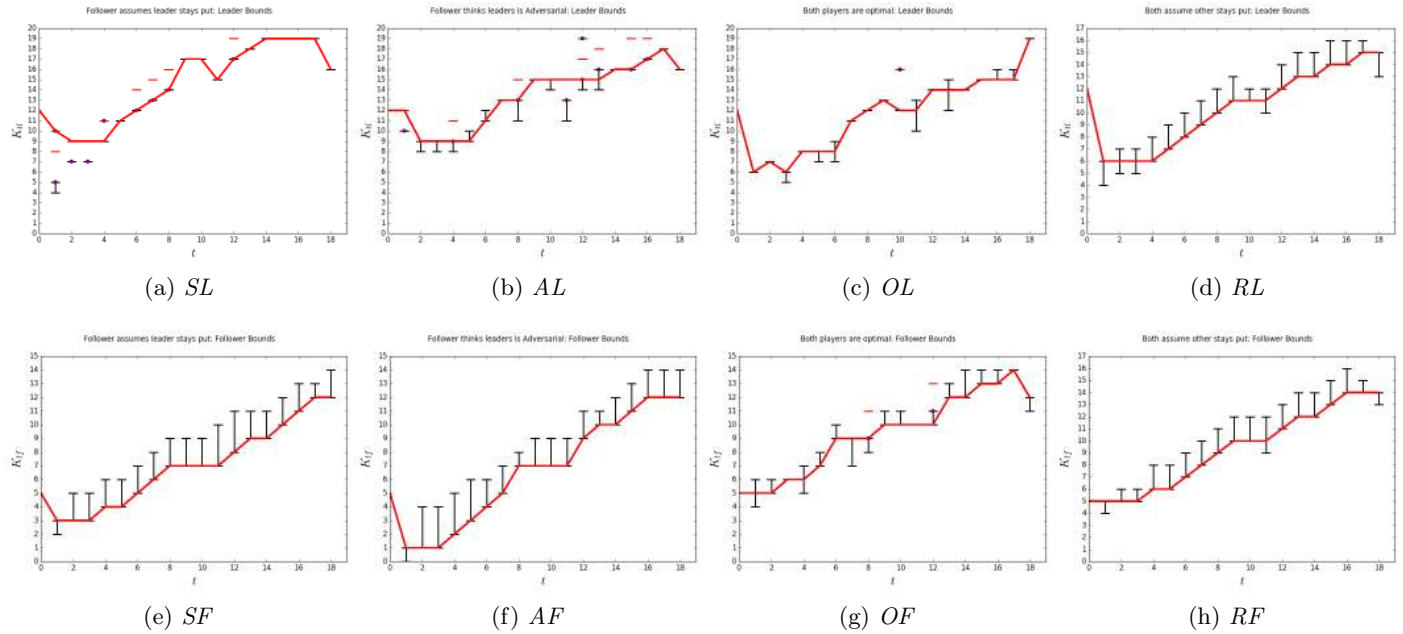
Third, comparing the symmetric cases in Figure 3,  $OL$  has similar profits but higher capacity than  $RL$ , while  $OF$  has similar capacity but lower profits than  $RF$ . Compared to the ISD policies of  $RL$  and  $RF$  in Figure 4, the ISD policies of  $OL$  and  $OF$  have smaller<sup>13</sup> stayput regions and multiple intervals at some periods. This indicates that proactive thinking in *optimal* gives both players strategical flexibility and makes them more likely to have higher capacity than they do in *reactive*. As the difference between the ISD policies of  $OL$  and those of  $RL$  is larger than the difference between the ISD policies of  $OF$  and those of  $RF$ , it is more likely that  $OL$  will have higher capacity than  $RL$  compared to the situation where  $OF$  surpasses  $RF$ .

Fourth, comparing the asymmetric cases in Figure 3,  $SL$  has higher capacity than  $AL$ . As time proceeds, the profits of  $SL$  gradually catch up with those of  $AL$ , while both the capacity and profits of  $SF$  are gradually surpassed by  $AF$ . In *adversarial* where the follower's information on the leader's strategy appears to be inaccurate, the follower can still benefit from proactive thinking. As indicated in Figure 4, the ISD policies of  $AL$  at some periods have larger stayput regions than those of  $SL$  and thus strategical flexibility of  $AL$  is restrained. This explains why  $AL$  has lower capacity than  $SL$  in Figure 3a. With our example of profit functions (equation (16)), the follower's assumption on the leader being adversarial directly suppresses the follower's own capacity: the stayput regions of  $AF$  are lower than those of  $SF$ . This describes why in the capacity and profits of  $AF$  gradually surpass those of  $SF$  only during later periods (when the demand is picking up).

### 7.3. Impact of market condition and investment irreversibility

Using Monte Carlo simulations, we investigate the impact of different market conditions and investment irreversibility levels on the optimal strategies. Using the same remaining parameters as in Section 7.1, we change  $\alpha$  values to represent different market conditions in the first experiment and change the value of disinvestment unit price, i.e.,  $r_{tj}$ , to represent different levels of investment irreversibility in the second experiment. Besides the four cases, a single-firm case is added in order to distinguish the competitive effect. This is to say, with different parameter values, we identify the occasions when a player's strategy under duopoly changes significantly differently from his strategy under monopoly. To allow direct comparison, we let both the leader and the follower operate as a monopolist in their separate market (i.e.,  $\frac{k_{ol}}{k_{ol}+k_{of}}\omega_t$  or  $\frac{k_{of}}{k_{ol}+k_{of}}\omega_t$ ) and refer to them

<sup>13</sup> The size of a stayput region depends on the number of contained capacity values, without considering whether the values are continuous. Thus, a large stayput region does not necessarily contain few intervals.



**Figure 4 Stayput region of the leader and the follower**

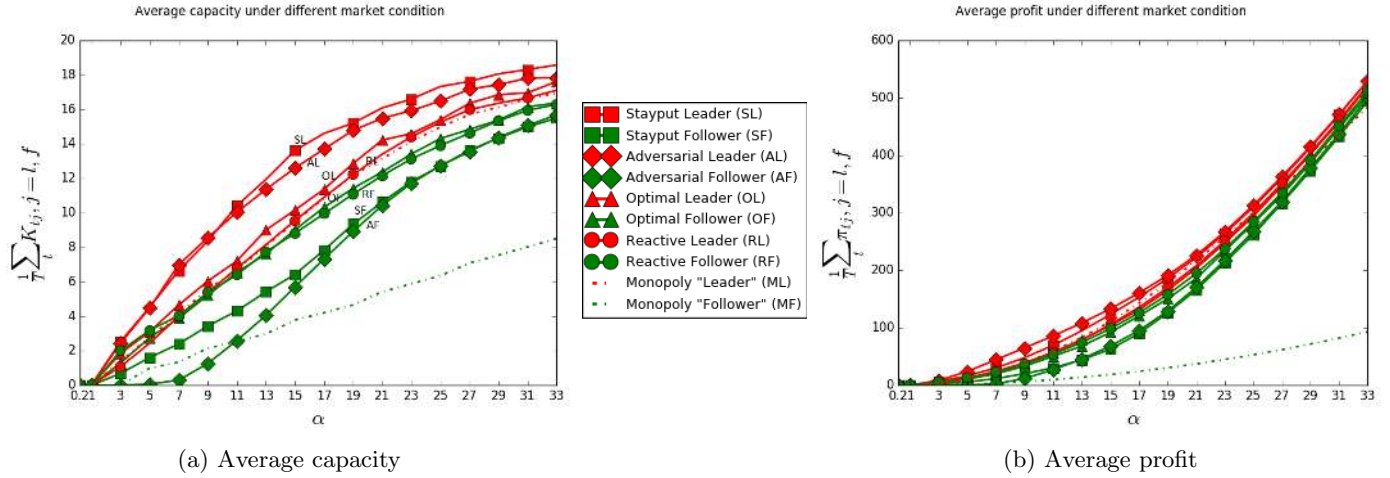
as monopoly “leader” ( $ML$ ) and monopoly “follower” ( $MF$ ). The method developed by Eberly and Van Mieghem (1997) is used to determine the monopolist’s optimal ISD investment strategy. In each experiment, 10,000 random demand paths are simulated and the same demand path is applied to all five cases with different parameter values. At each period, the optimal capacity and profit are computed and then averaged over the 10,000 simulations. For each case, we then average the mean optimal capacity and profits of each period over the entire timespan and use them to indicate the overall performance of this case. Figures 5 and 6 show the average capacity and profits of all five cases with different  $\alpha$  and  $r_{tj}$  values.

First, both players, especially the follower, perform better in cases where they adopt a competitive strategy compared to the case where they act as a monopolist in a separate market. Comparing the four cases to the single-firm case with different  $\alpha$  and  $r_{tj}$  values,  $SL$ ,  $AL$ ,  $OL$  have higher average capacity than  $ML$ , whereas  $RL$  has similar average capacity as  $ML$ . In all four cases, the follower has much higher average capacity than  $MF$  (see Figures 5a and 6a).  $SL$  and  $AL$  have higher average profits than  $ML$ , while  $OL$  has more average profits than  $ML$  during either market downturn (i.e., small  $\alpha$  values) or revival (i.e., large  $\alpha$  values). In all four cases, the follower has much higher average profits than  $MF$  (see Figures 5b and 6b).

Second, the previous results using a single demand path are robust with different  $\alpha$  and  $r_{tj}$  values. For example, the leader generally performs better than the follower in terms of average capacity and profit and the leader performs better in asymmetric cases than he does in symmetric cases, while the follower has the opposite situation. Comparing the symmetric cases, both players

in *optimal* have slightly higher average capacity than they do in *reactive* with different  $\alpha$  values (see Figure 5a). Comparing the asymmetric cases, *SF* has higher average capacity during market downturns, while *AF* gradually catches up as the market revives (see Figure 5a).

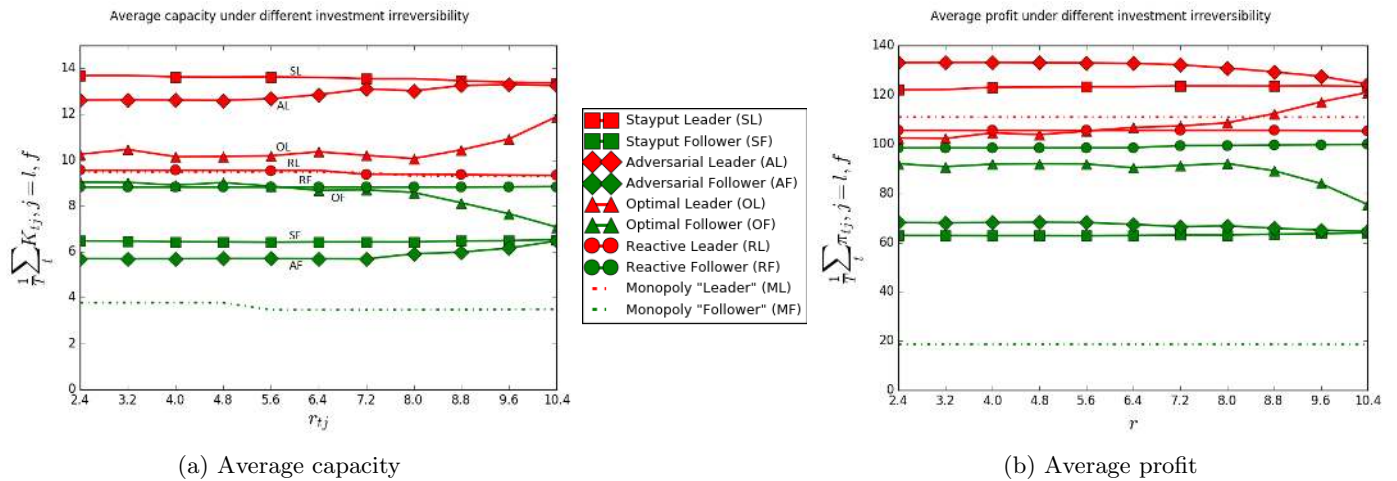
Third, when reducing the investment irreversibility (i.e., increasing the  $r$  value), only *OL* benefits from the cost advantage and *OF*'s average capacity and profits decrease.



(a) Average capacity

(b) Average profit

Figure 5 Impact of market parameters on the optimal strategies



(a) Average capacity

(b) Average profit

Figure 6 Impact of investment cost parameters on the optimal strategies

### 7.4. Managerial insights

In a market where dominant firms can influence the product price through their capacity investments, a competitive strategy allows these firms to benefit more in terms of capacity and profit,

compared to the strategy which separates firms into different markets. While dominant firms exercise their competitive strategies and thus gradually increase capacity, other small players' capacity are suppressed and the market continues to consolidate. During an investment race, it is crucial to be the first one to act as it increases the competitive advantage over the other players. We consider two types of competitive strategy (i.e., proactive and reactive) for the leader and follower firms in the race. Figure 7 shows the ranking of the four cases based on a player's average capacity, e.g.,  $1l$  and  $4l$  indicating the leader holds the highest and lowest average capacity among the four cases. The shadow area where the follower is proactive while the leader is reactive can be seen as identical to *stay put*, but with reverse players. Our results show that the best competitive strategy for any player features proactive thinking. It allows a firm to strategically plan its investments considering the possible responses from the other firm. As shown in Figure 7, the leader's average capacity is higher in the proactive cases than in the reactive case. As the leader moves first and exercises his best strategy, the follower's best response is to collect information on the leader's future decisions and to adopt a proactive strategy as well.

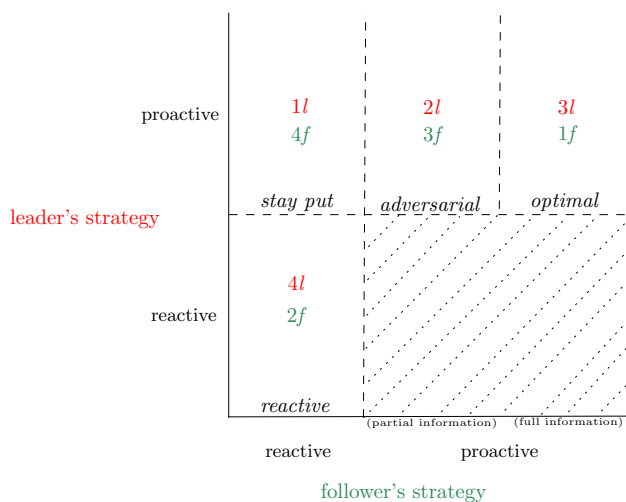


Figure 7 Ranking based on a player's average capacity

## 8. Conclusion

We study a competitive capacity investment problem under uncertainty and derive the optimal strategy in the form of an ISD policy. Different from the majority of oligopoly capacity models in the literature which focus on a single-shot investment, we investigate a long-term problem where firms sequentially adjust their capacity at each period in order to maximize the sum of current reward and expected future rewards. In a competitive market where firms' decisions interact with each other, our optimal long-term investment policies give different insights into the competitive strategy by explicitly featuring the interaction between a firm's capacity and the other firms'

current and future capacity. Moreover, by allowing for flexible capacity choices, our optimal policies do not only determine the timing of capacity adjustments, but also characterize the size of an investment. In the two-player case, the follower's ISD policy at period  $t$  is determined by the leader's first move  $k_{tl}$  and the uncertainty level, while the leader's ISD policy is a function of the uncertainty level, the follower's current capacity  $k_{t-1f}$ , and the follower's optimal investment policy at period  $t$ . We consider two types of competitive investment strategies for both players: proactive and reactive. The optimal ISD policies in each case are determined by different methods. In cases where both players are proactive and strategically influence each other, we developed two algorithms (i.e., *decomposition algorithm* and *elimination algorithm*) to compute the leader's and follower's ISD policies. A proactive strategy gives a player strategic flexibility, as indicated by multiple stayput intervals in the ISD policies.

We demonstrate the optimal ISD investment strategies using data from the container shipping market. Our results show that the realized investments in practice, which are questioned as "irrational" decisions, follow an optimal structure. In addition, firms' "overinvestments" during a market downturn are also shown to be optimal decision according to their long-term competitive strategies. Compared to the strategies that separate firms into different market, a competitive strategy allows both firms to benefit from the competition. Our results show that the leader can hold more capacity and make more profits when adopting a proactively rather than a reactively competitive strategy. The follower's best response to the leader's proactive strategy is also to be proactive. Regardless of the accuracy of the follower's information on the leader's strategy, the follower can benefit from proactive thinking if he strategically plans his investments while taking into account the their impact on the leader's future. Besides a firm's position in the investment race and whether the firm adopts a proactive strategy, market condition and investment costs can also influence the result of its ISD investment strategy.

Our model investigates investment strategies in a competitive market and provides insights into the structure of an optimal investment policy. However, investment practice is more complex. The competition often involves more than two firms and the middle player can adopt a leader, follower, or mix strategy. A direction for future research is to investigate the optimal investment policy of the middle player.

## References

- Aguerrevere, Felipe L. 2003. Equilibrium investment strategies and output price behavior: A real-options approach. *Review of Financial Studies* **16**(4) 1239–1272.
- Aguerrevere, Felipe L. 2009. Real options, product market competition, and asset returns. *The Journal of Finance* **64**(2) 957–983.

- Alphaliner. 2015. The top 100 liner operators (december, 2015) Retrieved from <http://www.alphaliner.com/top100>.
- Angelus, Alexandar, Evan L Porteus. 2002. Simultaneous capacity and production management of short-life-cycle, produce-to-stock goods under stochastic demand. *Management Sci.* **48**(3) 399–413.
- Armony, Mor, Erica L Plambeck. 2005. The impact of duplicate orders on demand estimation and capacity investment. *Management Science* **51**(10) 1505–1518.
- Back, Kerry, Dirk Paulsen. 2009. Open-loop equilibria and perfect competition in option exercise games. *Review of Financial Studies* 46.
- Barnato, Katy. 2015. Rough seas ahead for container shipping industry. Retrieved from [www.cnbc.com](http://www.cnbc.com).
- Bish, Ebru K, Qiong Wang. 2004. Optimal investment strategies for flexible resources, considering pricing and correlated demands. *Operations Research* **52**(6) 954–964.
- Bouis, Romain, Kuno JM Huisman, Peter M Kort. 2009. Investment in oligopoly under uncertainty: The accordion effect. *International Journal of Industrial Organization* **27**(2) 320–331.
- Brennan, Michael J, Lenos Trigeorgis, Petter Bjerksund. 2000. *Project flexibility, agency, and competition: New developments in the theory and application of real options*. Oxford University Press New York.
- Burnetas, Apostolos, Stephen Gilbert. 2001. Future capacity procurements under unknown demand and increasing costs. *Management Sci.* **47**(7) 979–992.
- Cachon, Gérard P, Patrick T Harker. 2002. Competition and outsourcing with scale economies. *Management Sci.* **48**(10) 1314–1333.
- Cachon, Gérard P, Martin A Lariviere. 1999. Capacity choice and allocation: Strategic behavior and supply chain performance. *Management Sci.* **45**(8) 1091–1108.
- Caldentey, René, Martin Haugh. 2006. Optimal control and hedging of operations in the presence of financial markets. *Math. Oper. Res.* **31**(2) 285–304.
- Chen, Ming-Jer, Ian C MacMillan. 1992. Nonresponse and delayed response to competitive moves: The roles of competitor dependence and action irreversibility. *Academy of Management Journal* **35**(3) 539–570.
- Chevalier-Roignant, Benoît, Christoph M Flath, Arnd Huchzermeier, Lenos Trigeorgis. 2011. Strategic investment under uncertainty: A synthesis. *Eur. J. Oper. Res.* **215**(3) 639–650.
- Clarkson Research Services. 2000-2015a. Shipping review and outlook (autumn version) .
- Clarkson Research Services. 2000-2015b. Shipping review and outlook (spring version) .
- CMA-CGM. 2011. Marco polo Retrieved from [www.ship-technology.com/projects/cma-cgm-marco-polo](http://www.ship-technology.com/projects/cma-cgm-marco-polo).
- Cullinane, Kevin, Mahim Khanna. 2000. Economies of scale in large containerships: Optimal size and geographical implications. *Journal of transport geography* **8**(3) 181–195.
- Del Sol, Patricio, Pankaj Ghemawat. 1999. Strategic valuation of investment under competition. *Interfaces* **29**(6) 42–56.

- Dixit, Avinash K. 1994. *Investment under uncertainty*. Princeton university press.
- Drewry. 2005. Container market quarterly forecast, 4th quarter 2005 .
- Drewry. 2014. Container market quarterly forecast, 4th quarter 2014 .
- Drewry Maritime Research. 2015a. Container census 2015: Survey and forecast of global container units .
- Drewry Maritime Research. 2015b. Ship operating costs annual review and forecast 2015/16 .
- Eberly, Janice C, Jan A Van Mieghem. 1997. Multi-factor dynamic investment under uncertainty. *Journal of economic theory* **75**(2) 345–387.
- Ferreira, Nelson, Jayanti Kar, Lends TRIGEORGIS. 2009. Option games: The key to competing in capital-intensive industries. *Harvard Business Review* **87**(3).
- Ghemawat, Pankaj. 2009. The risk of not investing in a recession. *MIT Sloan Management Review* **50**(3) 31.
- Goyal, Manu, Serguei Netessine. 2007. Strategic technology choice and capacity investment under demand uncertainty. *Management Sci.* **53**(2) 192–207.
- Greenwood, Robin, Samuel Hanson. 2013. Waves in ship prices and investment. Tech. rep., National Bureau of Economic Research.
- Grenadier, Steven R. 2002. Option exercise games: An application to the equilibrium investment strategies of firms. *Review of financial studies* **15**(3) 691–721.
- Groenevelt, H. 1991. Two algorithms for maximizing a separable concave function over a polymatroid feasible region. *Eur. J. Oper. Res.* **54**(2) 227–236.
- Gulati, Ranjay, Nitin Nohria, Franz Wohlgelegen. 2010. Roaring out of recession. *Harvard Business Review* **88**(3) 62–69.
- Gullaksen, JG Consultant Engineers. 2012. Ship acquisition and investment analysis models using spreadsheets: A computer modeling approach .
- Hoppe, Heidrun C. 2000. Second-mover advantages in the strategic adoption of new technology under uncertainty. *International journal of industrial organization* **18**(2) 315–338.
- Huh, Woonghee Tim, Robin O Roundy. 2005. A continuous-time strategic capacity planning model. *Naval Research Logistics (NRL)* **52**(4) 329–343.
- Huh, Woonghee Tim, Robin O Roundy, Metin Çakanyildirim. 2006. A general strategic capacity planning model under demand uncertainty. *Naval Research Logistics (NRL)* **53**(2) 137–150.
- Kulkarni, Shailesh S, Michael J Magazine, Amitabh S Raturi. 2004. Risk pooling advantages of manufacturing network configuration. *Production and Operations Management* **13**(2) 186–199.
- Lederer, Phillip J, Lode Li. 1997. Pricing, production, scheduling, and delivery-time competition. *Oper. Res.* **45**(3) 407–420.



- Macguire, Eoghan. 2013. Maersk 'triple e': Introducing the world's biggest ship Retrieved from [www.cnn.com/2013/06/26/business/maersk-triple-e-biggest-ship](http://www.cnn.com/2013/06/26/business/maersk-triple-e-biggest-ship).
- Maersk. 2015. Container market weekly report (august 14-21, 2015) .
- Murto, Pauli. 2004. Exit in duopoly under uncertainty. *RAND Journal of Economics* 111–127.
- Murto, Pauli, Erkkä Näsäkkälä, Jussi Keppo. 2004. Timing of investments in oligopoly under uncertainty: A framework for numerical analysis. *Eur. J. Oper. Res.* **157**(2) 486–500.
- Narongwanich, Wichai, Izak Duenyas, John R Birge. 2002. Optimal portfolio of reconfigurable and dedicated capacity under uncertainty. *Preprint. University of Michigan* .
- Netessine, Serguei, Gregory Dobson, Robert A Shumsky. 2002. Flexible service capacity: Optimal investment and the impact of demand correlation. *Oper. Res.* **50**(2) 375–388.
- Novy-Marx, Robert. 2007. An equilibrium model of investment under uncertainty. *Review of Financial Studies* **20**(5) 1461–1502.
- Odell, Mark. 2012. Shipping sector caught in choppy waters Retrieved from [www.ft.com](http://www.ft.com).
- Olhager, Jan, Martin Rudberg, Joakim Wikner. 2001. Long-term capacity management: Linking the perspectives from manufacturing strategy and sales and operations planning. *International Journal of Production Economics* **69**(2) 215–225.
- Pawlina, Grzegorz, Peter M Kort. 2006. Real options in an asymmetric duopoly: Who benefits from your competitive disadvantage? *Journal of Economics & Management Strategy* **15**(1) 1–35.
- Plambeck, Erica L, Terry A Taylor. 2005. Sell the plant? the impact of contract manufacturing on innovation, capacity, and profitability. *Management Science* **51**(1) 133–150.
- Ryan, Sarah M. 2004. Capacity expansion for random exponential demand growth with lead times. *Management Science* **50**(6) 740–748.
- Sanders, Ulrik. 2012. Restoring profitability to container shipping Retrieved from [www.bcgperspectives.com](http://www.bcgperspectives.com).
- Smit, Han TJ, Lenos Trigeorgis. 2012. *Strategic investment: Real options and games*. Princeton University Press.
- Søndergaard, Jens, Lars R. Eismark. 2012. Balancing the imbalances in container shipping Retrieved from [www.atkearney.com](http://www.atkearney.com).
- Swinney, Robert, Gérard P Cachon, Serguei Netessine. 2011. Capacity investment timing by start-ups and established firms in new markets. *Management Science* **57**(4) 763–777.
- Thijssen, Jacco JJ, Kuno JM Huisman, Peter M Kort. 2006. The effects of information on strategic investment and welfare. *Economic Theory* **28**(2) 399–424.
- UNCTAD. 2000. Review of maritime transport. *United Nations Conference on Trade & Development* .
- UNCTAD. 2001. Review of maritime transport. *United Nations Conference on Trade & Development* .

- UNCTAD. 2001-2015. Review of maritime transport. *United Nations Conference on Trade & Development* .
- UNCTAD. 2012. Review of maritime transport. *United Nations Conference on Trade & Development* .
- UNCTAD. 2013. Review of maritime transport. *United Nations Conference on Trade & Development* .
- UNCTAD. 2014. Review of maritime transport. *United Nations Conference on Trade & Development* .
- Van Mieghem, Jan, Nils Rudi. 2002. Newsvendor networks: Inventory management and capacity investment with discretionary activities. *Manufacturing Service Oper. Management* **4**(4) 313–335.
- Van Mieghem, Jan A. 1998. Investment strategies for flexible resources. *Management Science* **44**(8) 1071–1078.
- Van Mieghem, Jan A. 2003. Capacity management, investment, and hedging: Review and recent developments. *Manufacturing Service Oper. Management* **5**(4) 269–302.
- Van Mieghem, Jan A. 2007. Risk mitigation in newsvendor networks: Resource diversification, flexibility, sharing, and hedging. *Management Sci.* **53**(8) 1269–1288.
- Van Mieghem, Jan A, Maqbool Dada. 1999. Price versus production postponement: Capacity and competition. *Management Science* **45**(12) 1639–1649.
- Wikipedia. 2016. List of largest container ships Retrieved from [https://en.wikipedia.org/wiki/List\\_of\\_largest\\_container\\_ships](https://en.wikipedia.org/wiki/List_of_largest_container_ships).

## Appendix. Proofs of Propositions

PROPOSITION 1. *In cases where firm  $j$  responds reactively to the competition, his optimal value function  $V_{tj}^*$  is jointly concave in  $(k_{t-1j}, k_{tj})$  for any given the other firm's current capacity  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and for each  $\omega_t \in \Theta$ .*

*Proof.* Follow the reasoning of *Theorem 1 from Eberly and Van Mieghem (1997): satisfying Assumptions 1, 2, and 3, they show by induction and a concavity preservation lemma that a single firm's optimal value function  $V_t(k_{t-1}, \omega_t)$  is jointly concave in  $(k_{t-1}, k_t)$ . In the reactive case with  $k_{\tau i \neq j}$  given and invariant  $\forall \tau > t$ , firm  $j$ 's optimal value function  $V_{tj}^*(k_{ti}, k_{t-1j}, \omega_t)$  is jointly concave in  $(k_{t-1j}, k_{tj})$  for any given the other firm's current capacity  $k_{ti \neq j} \in \mathcal{K}_{ti}$  and for each  $\omega_t \in \Theta$ .*

PROPOSITION 2.  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ .

*Proof.* According to equations (13) and (14),  $\frac{\nabla_- G_{tf}(k_{tf}^L)}{\nabla k_{tf}^L} \geq c_{tf}$  and  $\frac{\nabla_+ G_{tf}(k_{tf}^H)}{\nabla k_{tf}^H} \leq r_{tf}$ . Thus,  $\frac{G_{tf}(k) - G_{tf}(k_{tf}^L)}{k - k_{tf}^L} \geq \frac{\nabla_- G_{tf}(k_{tf}^L)}{\nabla k_{tf}^L} \geq c_{tf}$ ,  $\forall k < k_{tf}^L$  and  $\frac{G_{tf}(k) - G_{tf}(k_{tf}^H)}{k - k_{tf}^H} \leq \frac{\nabla_+ G_{tf}(k_{tf}^H)}{\nabla k_{tf}^H} \leq r_{tf}$ ,  $\forall k > k_{tf}^H$ . It means that capacity  $k$  should be adjusted in order to maximize the value of the objective function in equation (10), i.e.,  $k \notin S_{[k_{fo}, k_{fe}]}$ ,  $\forall \{k : k < k_{tf}^L \vee k > k_{tf}^H\}$ . If  $k_{tf}^L > k_{tf}^H$ ,  $S_{[k_{fo}, k_{fe}]} = [k_{tf}^L, k_{tf}^H] = \emptyset$ . If  $k_{tf}^L \leq k_{tf}^H$ ,  $S_{[k_{fo}, k_{fe}]} \subseteq [k_{tf}^L, k_{tf}^H] \neq \emptyset$ . In conclusion,  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ .

PROPOSITION 3.  $k_{tf}^L$  and  $k_{tf}^H$  are the lowerbound and the upperbound of the stayput region  $S_{[k_{fo}, k_{fe}]}$ .

*Proof.* Assume there exists  $k_1, k_2$ , such that  $k_1 = \inf\{k : \frac{G_{tf}(k_{tf}^L) - G_{tf}(k)}{k_{tf}^L - k} \geq c_{tf}, k_{tf}^L < k < k_{tf}^H\}$  and  $k_2 = \sup\{k : \frac{G_{tf}(k_{tf}^H) - G_{tf}(k)}{k_{tf}^H - k} \leq r_{tf}, k_{tf}^L < k < k_{tf}^H\}$ . Thus,  $G_{tf}(k_{tf}^L) - G_{tf}(k_1) \leq c_{tf} \times (k_{tf}^L - k_1)$  and  $G_{tf}(k_{tf}^H) - G_{tf}(k_2) \leq r_{tf} \times (k_{tf}^H - k_2)$ . In addition,  $G_{tf}(k_{tf}^L) - G_{tf}(k) > c_{tf} \times (k_{tf}^L - k), \forall \{k : k_{tf}^L < k < k_1\}$  and  $G_{tf}(k_{tf}^H) - G_{tf}(k) > r_{tf} \times (k_{tf}^H - k), \forall \{k : k_2 < k < k_{tf}^H\}$ . According to equations (13) and (14),  $\frac{G_{tf}(k) - G_{tf}(k_{tf}^L)}{k - k_{tf}^L} \geq \frac{\nabla - G_{tf}(k_{tf}^L)}{\nabla k_{tf}^L} \geq c_{tf}, \forall k < k_{tf}^L$  and  $\frac{G_{tf}(k) - G_{tf}(k_{tf}^H)}{k - k_{tf}^H} \leq \frac{\nabla + G_{tf}(k_{tf}^H)}{\nabla k_{tf}^H} \leq r_{tf}, \forall k > k_{tf}^H$ . Adding the inequalities, we get the followings:  $\frac{G_{tf}(k) - G_{tf}(k_1)}{k - k_1} > c_{tf}, \forall \{k : k_{tf}^L < k < k_1\}$  and  $\frac{G_{tf}(k) - G_{tf}(k_2)}{k - k_2} < r_{tf}, \forall \{k : k_2 < k < k_{tf}^H\}$ ;  $\frac{G_{tf}(k) - G_{tf}(k_1)}{k - k_1} \geq c_{tf}, \forall k < k_{tf}^L$  and  $\frac{G_{tf}(k) - G_{tf}(k_1)}{k - k_1} \leq r_{tf}, \forall k > k_{tf}^H$ . Hence, if  $k_1, k_2$  exist,  $\frac{G_{tf}(k) - G_{tf}(k_1)}{k - k_1} \geq c_{tf}, \forall k < k_1$  and  $\frac{G_{tf}(k) - G_{tf}(k_1)}{k - k_1} \leq r_{tf}, \forall k > k_2$ .

According to the definitions of  $k_{tf}^L$  and  $k_{tf}^H$  (see equations (13) and (14)), the marginal values at  $k_1$  and  $k_2$  are:  $\frac{\nabla - G_{tf}(k_1)}{\nabla k_1} < c_{tf}$  and  $\frac{\nabla + G_{tf}(k_2)}{\nabla k_2} > r_{tf}$ . Thus, there exists some  $k_a < k_1$  such that  $\frac{G_{tf}(k_a) - G_{tf}(k_1)}{k_a - k_1} < c_{tf}$  and exists some  $k_b > k_2$  such that  $\frac{G_{tf}(k_b) - G_{tf}(k_2)}{k_b - k_2} > r_{tf}$ . This contradicts the claim that  $\forall k < k_1$  or  $\forall k > k_2$ ,  $k$  should be adjusted to  $k_1$  or  $k_2$ . Therefore,  $k_1$  and  $k_2$  do not exist and  $k_{tf}^L, k_{tf}^H \in S_{[k_{fo}, k_{fe}]}$ . According to Proposition 2,  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ . Hence,  $k_{tf}^L$  and  $k_{tf}^H$  are the lowerbound and the upperbound of the stayput region  $S_{[k_{fo}, k_{fe}]}$ .

PROPOSITION 4. The rolling up and rolling down procedures eliminate all the non-stayput capacity values from the range  $S_{[k_{fo}, k_{fe}]}^r$ .

*Proof.* At each iteration of the rolling up procedure,  $k$  is eliminated if  $k \notin S_{[k_{tf}^L, k]}^r$ . According to Proposition 2,  $k \notin S_{[k_{tf}^L, k]}$  if  $k \notin S_{[k_{tf}^L, k]}^r$ . If  $k$  is a non-stayput point in a space  $\mathcal{X}$ ,  $k$  is also a non-stayput point in any space that contains the original space  $\mathcal{X}$ . Hence,  $k \notin S_{[a, b]}, \forall [a, b] \supseteq [k_{tf}^L, k]$ , e.g.,  $k \notin S_{[k_{fo}, k_{fe}]}$ . This is to say, capacity values that are eliminated by the rolling up procedure are the non-stayput points in the space  $[k_{fo}, k_{fe}]$ . By the same token, capacity values that are eliminated by the rolling down procedure are also the non-stayput points.

Next, we prove that all of the remaining values are the stayput points in the space  $[k_{fo}, k_{fe}]$ . At each iteration of the rolling up procedure,  $k$  is kept if  $k \in S_{[k_{tf}^L, k]}^r$ . According to Proposition 3,  $k \in S_{[k_{tf}^L, k]}$  if  $S_{[k_{tf}^L, k]}^r = [\cdot, k]$ . If  $k \in S_{[k_{tf}^L, k]}$ , it means that the value of the objective function in equation (10) cannot be improved by adjusting  $k$  to any point in the interval  $[k_{tf}^L, k]$ . The rolling down procedure follows the same argument as the rolling up procedure:  $k \in S_{[k, k_{tf}^H]}$  if  $S_{[k, k_{tf}^H]}^r = [k, \cdot]$ . After the rolling up and rolling down procedures, any remaining capacity value  $k$  is a stayput point in the interval  $[k_{tf}^L, k_{tf}^H]$  as  $k$  should neither be adjusted to any point in the interval  $[k_{tf}^L, k]$  nor to any point in the range  $[k, k_{tf}^H]$ . According to Proposition 2,  $S_{[k_{fo}, k_{fe}]} \subseteq S_{[k_{fo}, k_{fe}]}^r = [k_{tf}^L, k_{tf}^H]$ . Hence, the stayput region in the range  $[k_{tf}^L, k_{tf}^H]$  is also the stayput region in the space  $[k_{fo}, k_{fe}]$ , i.e.,  $S_{[k_{tf}^L, k_{tf}^H]} = S_{[k_{fo}, k_{fe}]}$ . In conclusion, the rolling up and rolling down procedures eliminate the non-stayput capacity values from the range and all of the remaining capacity compose the stayput region  $S_{[k_{fo}, k_{fe}]}$ .

PROPOSITION 5. Given  $k_{t-1f} = k$  and  $k \notin S_{tf}(k_{tl}, \omega_t)$ , if the optimal investment policy indicated by  $S_{tf}(k_{tl}, \omega_t)$  assigns  $k_{tf} = b$ , then no stayput values exist in the interval  $[k, b]$  if  $b > k$ , or in the interval  $(b, k]$  if  $b < k$ . In other words, the interval  $[k, b)$  or  $(b, k] \not\subseteq S_{tf}(k_{tl}, \omega_t)$ .

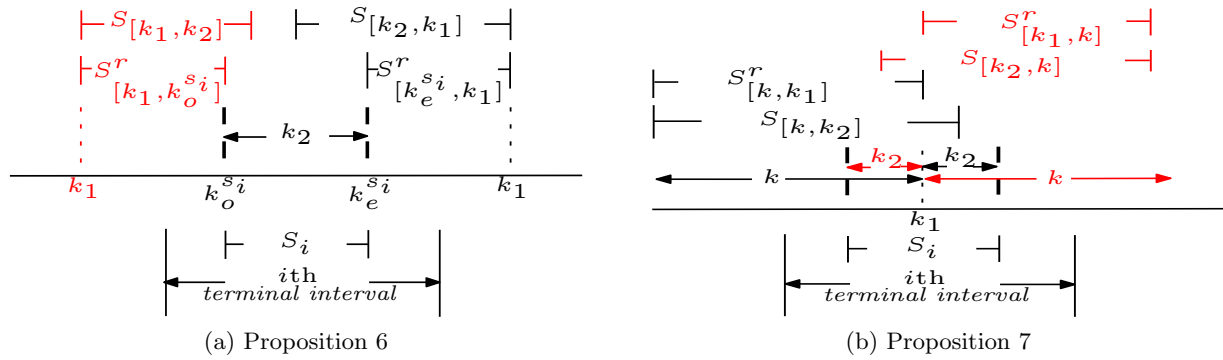
*Proof.* The optimal investment policy assigns  $k_{tf} = b$  when  $k_{t-1f} = k$ , thus  $G_{tf}(b) - c_{tf} \times (b - k) \geq G_{tf}(k') - c_{tf} \times (k' - k)$ ,  $\forall k' \geq k$  in the case  $b > k$ . Therefore, capacity values in the interval  $[k, b)$  need to be adjusted to  $b$  in order to maximize the value of the objective function in equation (10) as  $\frac{G_{tf}(b) - G_{tf}(a)}{b - a} \geq c_{tf}$ ,  $\forall a \in [k, b)$ . In other words,  $a \notin S_{tf}(k_{tl}, \omega_t)$ ,  $\forall a \in [k, b)$ . The proof for the case  $b < k$  follows the same argument. In conclusion, the interval  $[k, b)$  or  $(b, k] \not\subseteq S_{tf}(k_{tl}, \omega_t)$  if the optimal investment policy indicated by  $S_{tf}(k_{tl}, \omega_t)$  assigns  $k_{tf} = b$  when  $k_{t-1f} = k \notin S_{tf}(k_{tl}, \omega_t)$ .

**PROPOSITION 6.**<sup>14</sup> Given the stayput region in the  $i$ th terminal interval, i.e.,  $S_i = [k_o^{s_i}, k_e^{s_i}]$ , if there exists a capacity value  $k_1 > k_e^{s_i}$  satisfying  $k_1 \in S_{[k_e^{s_i}, k_1]}^r$  (or a capacity value  $k_1 < k_o^{s_i}$  satisfying  $k_1 \in S_{[k_1, k_o^{s_i}]}^r$ ), then  $k_1 \in S_{[k_2, k_1]}$  (or  $k_1 \in S_{[k_1, k_2]}$ ),  $\forall k_2 \in S_i$ .

*Proof.* Given  $S_i = [k_o^{s_i}, k_e^{s_i}]$ ,  $\frac{G(k_e^{s_i}) - G(k)}{k_e^{s_i} - k} > r$ ,  $\forall k \in [k_o^{s_i}, k_e^{s_i})$ . If there exists a capacity value  $k_1 > k_e^{s_i}$  satisfying  $k_1 \in S_{[k_e^{s_i}, k_1]}^r$ , then  $k_1 \in S_{[k_e^{s_i}, k_1]}$  according to Proposition 3. Thus,  $\frac{G(k_1) - G(k_e^{s_i})}{k_1 - k_e^{s_i}} > r$ . Adding inequality  $G(k_1) - G(k_e^{s_i}) > r(k_1 - k_e^{s_i})$  to inequality  $G(k_e^{s_i}) - G(k) > r(k_e^{s_i} - k)$ ,  $\forall k \in [k_o^{s_i}, k_e^{s_i})$ , we get:  $\frac{G(k_1) - G(k)}{k_1 - k} > r$ ,  $\forall k \in [k_o^{s_i}, k_e^{s_i})$ . In conclusion,  $k_1 \in S_{[k_2, k_1]}$ ,  $\forall k_2 \in S_i$ . The proof for the case  $k_1 < k_o^{s_i}$  follows the same argument.

**PROPOSITION 7.**<sup>15</sup> Given the stayput region in the  $i$ th terminal interval, i.e.,  $S_i$ , if a capacity value  $k_1 \in S_i$  satisfying the following:  $k_1 \in S_{[k, k_1]}^r$ ,  $\exists k < k_1$  (or  $k_1 \in S_{[k_1, k]}^r$ ,  $\exists k > k_1$ ), then we have:  $k_2 \in S_{[k, k_2]}$ ,  $\forall \{k_2 : k_1 < k_2 \in S_n\}$  (or  $k_2 \in S_{[k_2, k]}$ ,  $\forall \{k_2 : k_2 < k_1 \wedge k_2 \in S_n\}$ ).

*Proof.* If a capacity value  $k_1 \in S_i$  satisfying the following:  $k_1 \in S_{[k, k_1]}^r$ ,  $\exists k < k_1$ , then  $k_1 \in S_{[k, k_1]}$  according to Proposition 3. Thus,  $\frac{G(k_1) - G(k)}{k_1 - k} > r$ . For any  $k_2 > k_1$  and  $k_1, k_2 \in S_i$ ,  $\frac{G(k_2) - G(k_1)}{k_2 - k_1} > r$ . Adding inequality  $G(k_1) - G(k) > r(k_1 - k)$  to inequality  $G(k_2) - G(k_1) > r(k_2 - k_1)$ , we get:  $\frac{G(k_2) - G(k)}{k_2 - k} > r$ , i.e.,  $k_2 \in S_{[k, k_2]}$ ,  $\forall \{k_2 : k_1 < k_2 \in S_n\}$ . The proof for the case  $k > k_1$  follows the same argument.



**Figure 8** Figures for illustrating Propositions 6 and 7

<sup>14</sup> See Figure 8a for an illustration; in Figure 8, the red symbols indicate the second situation which is stated within the brackets in a proposition.

<sup>15</sup> See Figure 8b for an illustration.