

2015/34



Competition in the Market for Flexible Resources:
an application to cloud computing

Wing Man Wynne Lam



CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium

<http://www.uclouvain.be/core>

CORE

Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: immaq-library@uclouvain.be
<http://www.uclouvain.be/en-44508.html>

CORE DISCUSSION PAPER
2015/34

Competition in the Market for Flexible Resources: an application to cloud computing

Wing Man Wynne Lam [†]

July, 2015

Abstract

This paper considers firms' incentives to invest in local and flexible resources when demand is uncertain and correlated. Before demand is realized, two firms invest in their local capacity, and provider(s) of flexible resources invest in their capacity. After demand is realized, firms make their investment decision in flexible resource. I find that market power of the monopolist providing flexible resources distorts investment incentives, while competition mitigates them. The extent of improvement depends critically on demand correlation and the cost of capacity: under social optimum and monopoly, if the flexible resource is cheap, the relationship between investment and correlation is positive, and if it is costly, the relationship becomes negative; under duopoly, the relationship is positive.

Keywords: capacity investment, cloud computing, competition, demand correlation

JEL Classification: D4, L8

[†]University of Liege (ULg), HEC Management School, Liege Competition and Innovation Institute (LCII).

E-mail: wingmanwynne.lam@ulg.ac.be

Competition in the Market for Flexible Resources: an application to cloud computing*

Wing Man Wynne Lam[†]

July, 2015

Abstract

This paper considers firms' incentives to invest in local and flexible resources when demand is uncertain and correlated. Before demand is realized, two firms invest in their local capacity, and provider(s) of flexible resources invest in their capacity. After demand is realized, firms make their investment decision in flexible resource. I find that market power of the monopolist providing flexible resources distorts investment incentives, while competition mitigates them. The extent of improvement depends critically on demand correlation and the cost of capacity: under social optimum and monopoly, if the flexible resource is cheap, the relationship between investment and correlation is positive, and if it is costly, the relationship becomes negative; under duopoly, the relationship is positive.

Keywords: capacity investment, cloud computing, competition, demand correlation

JEL Classification: D4, L8

1 Introduction

For firms in various industries, capacity investment decision involves investing early in their own capacity before demand for their products is realized, and such investment is difficult to reverse. After the demand is realized, firms have the option to undertake a second investment in a flexible resource to accommodate the excess demand. In IT, cloud computing represents a flexible resource: cloud computing is fundamentally the leasing of computer services, including computing power and storage, but on an unprecedented scale. While local computing capacity

*I thank Paul Belleflamme, Giacomo Calzolari, Jacques Crémer, Vincenzo Denicolò, Florian Englmaier, Federico Etro, Neil Gandal, Axel Gautier, Michael Katz, Thomas-Olivier Léautier, Fabio Manenti, Domenico Menicucci, Paul Seabright, Tommaso Valletti, and participants at the CRESSE 2013 conference, the Toulouse 2013 Conference on Intellectual Property, Software and Internet, the Munich 2013 ICT Conference, and the European University Institute 2013 Seminar on Communications and Media Markets, as well as at the seminar at the University of Liege for helpful comments. I also acknowledge the support of Toulouse School of Economics and University of Bologna in earlier versions of this work. Any opinions expressed are those of the author only.

[†]University of Liege (ULg), HEC Management School, Liege Competition and Innovation Institute (LCII).
E-mail: wingmanwynne.lam@ulg.ac.be

can support the average demand of the firm, cloud computing is able to scale services on demand and accommodate the workload that exceeds what the local capacity can handle.¹ Accordingly, firms can use cloud computing as a flexible resource for business continuity and disaster recovery plans.² Electricity offers another example. In developing countries, since the provision of electricity is not very reliable, firms have to decide whether to buy electricity from public utilities, which deliver instant services (hence, the more flexible resource), or to invest in their own private generators. In addition to IT and electricity, in property, firms have to decide whether to rent or buy office space; in transport, people have to decide whether to take flexible public transport or buy their own private cars. Thus, flexibility here can be defined as anything that transforms fixed costs to variable costs (or simply “anything as a service” in the case of the Internet).

Another important characteristic of these markets is that demand is uncertain and correlated, and the providers of flexible resources need to make their capacity choices under these circumstances. For instance, in the cloud computing market, computing demand is uncertain as demand varies daily; and correlated globally and seasonally. For example, a U.S. cloud provider such as Amazon, Google and Microsoft may have customers from Europe as well as Australia. Correlation is therefore driven to some extent by geography: computing demands from countries that are close to each other are positively correlated; demands from countries that are located in different time zones are negatively correlated. Moreover, retailers increase computing demand during the holiday season;³ and businesses need more computing power during the tax season. As argued by Harms and Yamartino (2010), even the largest cloud provider will not be able to fully resolve issues related to uncertainty and correlation by aggregating demand. Similarly, time and weather are well-known common drivers of correlated demand in electricity and transport: household consumption of electricity tends to peak in the evening, while workplace consumption tends to peak in the daytime; and people most want a taxi during rush hour, holidays, and under bad weather condition.

It is likely that under demand uncertainty, the investments in different resources may be inefficient, depending on the strength of demand correlation and the costs of the resources. However, the existing literature has failed to explain why such inefficiencies may happen and to discuss the remedies, which is the subject of this paper. More specifically, this paper focuses on the problem of capacity investment in two resources when demand is uncertain and correlated. In the cloud computing example, capacity is a key part of competition in this industry. In the introductory phase, it is common that cloud providers build far more capacity than needed, and one does not expect capacity to be an issue in this growing phase. However, as cloud computing enters a more mature phase, capacity may become constrained as demand

¹The U.S. National Institute of Standards and Technology provides five defining characteristics of cloud computing: on-demand service, broad network access, resource pooling, rapid elasticity and measured service. This paper focuses on the definition of on-demand service and rapid elasticity.

²Business continuity and disaster recovery plans minimize any disruption of business operation due to insufficient local capacity or failure of critical systems.

³See “Black Friday: Websites swamped by record demand,” *BBC News*, 28 November 2014, available at <http://www.bbc.com/news/business-30242124>.

grows quickly.⁴⁵ For example, on 25 August 2013, Amazon seems to struggle to keep up with the growing computing demand, and an IT problem at one of its datacenters has caused many users of major web services such as Instagram, Netflix, Vine and Airbnb to experience lengthy delays and reduced data transfer speeds for several hours.⁶ Amazon’s web stores, Microsoft Azure and its outlook.com, Google’s Gmail service and its YouTube video site have had faced similar glitches. This raises the interesting questions of what the profit-maximizing investment strategy in flexible resource is, such that the problem of quality degradation can be avoided, and of how we should promote efficient investment from a public policy perspective.

In this model, two firms, whose demand is uncertain and correlated, make their investment decision in local resource. Observing firms’ local investment, providers of flexible resource (e.g. Amazon, Google and Microsoft) decide how much to invest in capacity, and set the price for their flexible resource (e.g. Amazon Web Services (AWS), Google Compute Engine, Microsoft Azure). After demand is realized, firms can buy flexible resources if demand exceeds their local capacity. It is worth noting that I use the “firm” to refer to the seller of final services to the consumers, and the “provider” to refer to the supplier of flexible resources. Thus, there are two parties investing: the “firm” chooses between fixed and flexible resources, and the “provider” invests in its flexible capacity.

I consider both cases of monopoly and duopoly in providing the flexible resource. As should be expected, investment is suboptimal in the monopoly market. Particularly, the provider of the flexible resource tends to underinvest in its capacity with respect to the socially optimal level, whereas firms tend to overinvest in their local capacity. Such inefficiency comes from market power of the monopolist. Firms invest in local capacity to avoid being exploited by the monopolist, which in turn reduces investment incentive of the monopolist.

Competition mitigates the underinvestment problem, and the extent of improvement depends crucially on demand correlation and the cost of capacity. Both socially optimal and monopoly investment in flexible resource increases with correlation if the investment cost of flexible resource is small enough, and decreases with correlation if the flexible resource is costly. The reason is that as correlation increases, the social planner and the provider either “win big” when demand realization of the two firms is high or “lose big” when demand realizations of the two firms is low. If the flexible resource is cheap, the social planner and the provider need not worry about “losing”. Rather, they will focus on reaping benefits from the “winning” outcome, and therefore they invest more as correlation increases. On the contrary, if the flexible resource is expensive, then “losing” is costly, and thus they invest less as correlation increases.

Under duopoly, using numerical examples, I show that investment in flexible resource is increasing in correlation. The reason for not observing the negative relationship between in-

⁴International Data Corporation (IDC) estimates that worldwide spending on public cloud services is expected to reach \$47.4 billion in 2013 and \$107 billion in 2017, which represents a growth rate five times that of the IT industry as a whole.

⁵Capacity can be interpreted in two ways: number of physical servers or service quality. In the former case, there is a maximum traffic that each server can handle. In the latter case, even if the capacity does not hit the limit, high demands can put a costly strain on servers, which results in poor quality of service.

⁶See “Instagram, Vine and Netflix hit by Amazon glitch,” *BBC News*, 26 August 2013, available at <http://www.bbc.co.uk/news/technology-23839901>.

vestment and correlation in this case, as opposed to the social optimum and the monopoly case, is that firms rely more on the flexible resource as competition between providers lowers the price of flexible resources. As correlation increases, firms' incentive to capture the windfall from the "winning" of high demand realizations increasingly outweighs their incentive to avoid the risk of "losing". Knowing this, each provider is willing to build a bigger capacity of flexible resources. These results suggest that information on the cost condition and the degree of demand correlation have important consequences for investment. They also explain the need for far more data on costs and demand in order to underpin the appropriate degree of competition. More generally, these results can be applied to similar issues that arise in other markets (e.g. the markets for cloud computing and the "sharing economy") even in the presence of spot pricing, which I will discuss in more detail in the penultimate section.

This paper extends the literature in three ways. First, it focuses on investment in two types of resources. Second, it considers uncertain and correlated demand. Third, it provides the conditions (in terms of investment costs and demand correlation) under which it is optimal for the social planner and it is profitable for the firms to invest in two different resources. There are three strands of related work, but they focus on only some of these aspects but not all.

First, this paper is closely related to the literature on capacity-pricing games, which dates back to Kreps and Sheinkman (1983), and then extended by Reynolds and Wilson (2000) and de Frutos and Fabra (2011) to incorporate demand uncertainty. Building on their results, we study firms' choices between two types of resources, and introduce both demand uncertainty and demand correlation. This also distinguishes our work from Goyal and Netessine (2007) and Anupindi and Jiang (2008), where firms cannot invest in both fixed and flexible resources at the same time. However, as discussed earlier, it is common for firms (e.g. in IT, electricity, property and transport) to invest in local capacity to accommodate the average demand and use flexible resources to better manage unexpected rises and falls in demand.⁷

Another related strand is the literature on operational management. For example, Lee (2009) studies the optimal capacity investment of a computing service provider in a single resource, and Niyato, Chaisiri and Lee (2009) study the optimal choice between private and public computing services in the monopoly and oligopoly markets, but in a setting without demand correlation. By incorporating demand correlation, an interesting finding of this paper is that investment can increase with correlation, which is in contrast to the common belief that only negative correlations are valuable because the provider can aggregate demand and reduces the risk.⁸ The reason is that when capacity is cheap, the "win big" effect prevails; thus, providers invest more as correlation increases.

Other papers in operational management have failed to address the divergence in the investment behavior between the social planner and the firms, and more importantly, how to correct the problem of suboptimal investment. For example, Van Miegham (1998), and Bish and Wang (2004) study the optimal investment strategy in flexible resources when a monop-

⁷The choice between fixed and flexible resources also relates to the literature on make-or-buy and outsourcing decisions. See Tadelis and Williamson (2013) and Bresnahan and Levin (2013) for excellent surveys of this literature, but the issue of demand correlation has not yet been studied there.

⁸See, for instance, p. 218 of Bayrak et al. (2011).

olist faces uncertain demands for its two products, which corresponds to the socially optimal regime in this paper.

Finally, this paper is loosely related to the literature on real options (RO) in finance, which focuses on the role of RO in providing flexibility to management decisions. Unlike financial assets, capacity investments in IT are not tradable, and are therefore not valued based on levels of risk, but instead they are priced by a third party, which is the resource provider in this model. Thus, the study of the investment incentives of resource providers is important here, but not in RO literature. See, for instance, Angelou and Economides (2005), Benaroch and Kauffman (1999) and Kauffman et al. (2002) for details on the limitation of RO's applicability in IT investments.

2 The Model

Consider two firms, 1 and 2, that need to build capacity in order to serve their customers. To do this, they can either invest in their own local resource L or they can buy flexible resources K from the market. The difference lies in that investments in local resources are irreversible and these resources are for the exclusive use of the investing firm, while flexible resources can be bought from the market instantly when needed and released when not needed. An example of flexible resources is cloud computing as cloud computing power is provisioned as an on-demand service.

Investment technology. The unit cost of local resource and flexible resource are denoted by c_L and c_K respectively. I assume that local resource is supplied competitively, so that firms can buy L at a price c_L . The flexible resource market can be either a monopoly or a duopoly. Hereafter, the seller of final services to the consumers is referred to as the “firm”, and the supplier of flexible resources that may be used in lieu of local resources in the production of such services is referred to as the “provider”.

Demand. Each consumer of a firm has inelastic demand, with willingness to pay of r .⁹ The demand (e.g. the total number of consumers) of each firm is uncertain, and the demands of the two firms are correlated. More specifically, demands for firms 1 and 2, denoted by x and y respectively, are drawn from a joint distribution $h(x, y)$, with support $[0, \infty) \times [0, \infty)$. The demand of firm 1, x , is given by the marginal distribution $f(x) = \int_0^\infty h(x, y) dy$. Similarly, the demand of firm 2, y , is given by $g(y) = \int_0^\infty h(x, y) dx$. In the following analysis, I focus on the case where demands (x, y) follow an exponential distribution with $\lambda = 1$,¹⁰ but in Appendix E, I show that the main results carry through in the linear case. More particularly, the exponential distribution can be described as follows. The marginal distributions $F(x)$ and

⁹Qualitative results for the monopoly case would be similar if we consider elastic demand. As for the duopoly case, however, if we consider elastic demand, we can no longer follow the approach of de Frutos and Fabra (2011), who study a sequential capacity-price game under demand uncertainty with price-inelastic demands. Interested reader can see Reynolds and Wilson (2000) for a discussion of the two-stage game under the assumption of downward-sloping and uncertain demand.

¹⁰A distribution is exponential when $F(\lambda, x) = 1 - \lambda e^{-\lambda x}$ is satisfied.

$G(y)$ and marginal densities $f(x)$ and $g(y)$ are respectively

$$\begin{aligned} F(x) &= 1 - e^{-x}, \\ G(y) &= 1 - e^{-y}, \\ f(x) &= e^{-x}, \\ g(y) &= e^{-y}. \end{aligned}$$

The joint distribution function $H(x, y)$ and joint density function $h(x, y)$ follow Gumbel (1960):

$$\begin{aligned} H(x, y) &= (1 - e^{-x})(1 - e^{-y})(1 + \alpha e^{-x-y}), \\ h(x, y) &= e^{-x-y}[1 + \alpha(2e^{-x} - 1)(2e^{-y} - 1)], \end{aligned}$$

where $-1 < \alpha < 1$ is a measure of correlation.¹¹

We consider the following game:

- Stage 1: firms 1 and 2 invest in their own local capacity L_1 and L_2 simultaneously;
- Stage 2: the provider(s) invest(s) in capacity of flexible resources K ;
- Stage 3: the provider(s) set(s) a per unit price of flexible resource p ;
- Stage 4: demands (x, y) are realized and firms decide whether and how much to buy the flexible resource.¹²

For simplicity, I make the following assumptions. First, $r > c_L$, so there is incentive to purchase local resources. Second, I focus on the more interesting case where $c_K < c_L$ because the value of flexible resources is the highest. If, instead, $c_L < c_K$, then the social planner will invest in a positive amount of local resources; value of flexible resources would be lower. Moreover, it is common in practice that cloud computing exhibits significant economies of scale (the market for cloud computing will be discussed in detail in Section 6). To facilitate our analysis, I focus on the specification with $r = 1$, $c_L = 0.5$ and $c_K \in [0, 0.5]$.¹³ Third, when firms are indifferent between buying and not buying the flexible resource, they will always buy. The solution concept adopted here is subgame perfect equilibrium (SPE).

¹¹Strictly speaking, $\rho = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}}$ is the coefficient of correlation, but since α and ρ move in the same direction (more precisely, $\rho = \frac{\alpha}{4}$, see Equation (3.10) on p. 706 of Gumbel (1960)), there is no loss of generality in saying that α is a measure of correlation.

¹²The timing of this model, where firms invest first, fits the scenario where some flexible resources such as cloud computing offers more flexibility in managing demand uncertainty than local resources. However, one could alternatively consider the case where firms observe the provider's investment in flexible resources before deciding their own local investment. In this setting, firms still overinvest in L , and providers still underinvest in K , provided price is chosen after the capacity decision because the monopoly price will emerge as long as demand is inelastic. Another alternative is to consider the case where p is chosen prior to L , but the underinvestment problem will still occur because the provider will never price at marginal cost (at least under linear pricing) as its profit will become zero and it will not have any incentive to invest. Moreover, it is difficult to think of a situation in practice that fits the scenario of choosing price prior to capacity.

¹³These assumptions are innocuous for two reasons. First, setting $r = 1$ is only a normalization, and it will not affect the qualitative conclusion. Second, the main results hold more generally as long as the flexible resource is more efficient, i.e. $c_K < c_L$.

3 Social Optimum

The social planner chooses L_1, L_2 and K so as to maximize social welfare, which is defined as expected demand served minus investment costs (i.e. $E(x + y) - c_K K - c_L(L_1 + L_2)$). Figure 1 illustrates the basic structure of the demand for flexible resources.

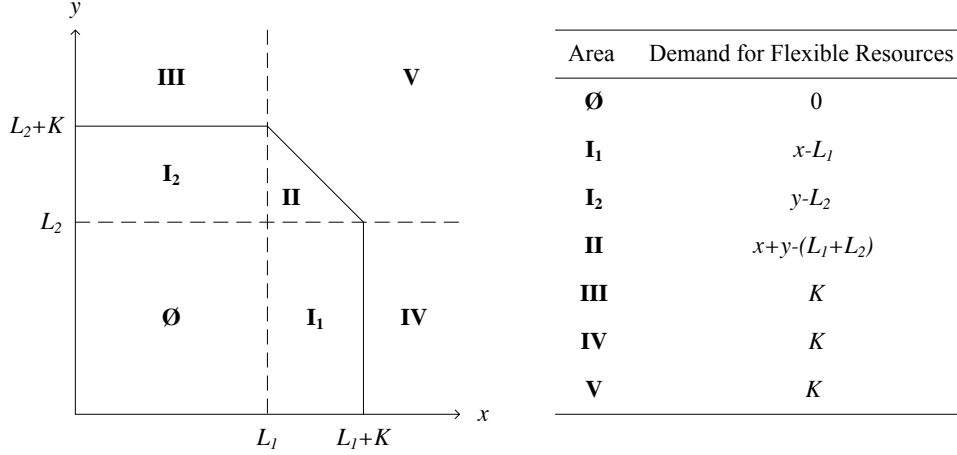


Figure 1: Demand for Flexible Resources.

In *Area $\emptyset$* , both firms have sufficient local capacity to serve their customers, and therefore there is no demand for flexible resources. *Area I_1* captures the situation where firm 2's local capacity is enough to cover its demand, but firm 1's demand exceeds its local capacity and will therefore purchase flexible resources. *Area I_2* illustrates the reverse situation when only firm 2 buys flexible resources. In *Area II* , both firms buy flexible resources. In all the cases above, all demands are served. *Area III* represents the situation where firm 1 has enough local capacity, while firm 2 has too much demand such that the provider of flexible resources is capacity constrained. *Area IV* shows the reverse situation: firm 1 has too much demand, while firm 2's local capacity is sufficient. *Area V* captures the situation where the demands of both firms are extremely high such that it exhausts the capacity of the flexible resource provider. Thus the social welfare is given by

$$\begin{aligned}
 \max_{L_1, L_2, K} S = & \int_{\emptyset + I_1 + I_2 + II} (x + y) h(x, y) dy dx + \int_{III} (x + L_2 + K) h(x, y) dy dx \\
 & + \int_{IV} (L_1 + K + y) h(x, y) dy dx + (L_1 + L_2 + K) \int_V h(x, y) dy dx \\
 & - c_K K - c_L (L_1 + L_2).
 \end{aligned} \tag{1}$$

Let $\Omega(L_1, L_2, K)$ denote the probability of (x, y) falling in areas $\{III\} + \{IV\} + \{V\}$. The social planner only invests in flexible resources, and the socially optimal investment is given by

$$\Omega(0, 0, K) = 1 - \int_0^K \int_0^{K-x} h(x, y) dy dx = c_K.$$

The optimal capacity is such that the social marginal benefit equals the marginal cost. Let L^* and K^* denote the solutions of Equation (1).

Proposition 1. *The social planner only invests in the flexible resource (i.e. $L^* = 0$ and $K^* > 0$). Moreover, there exists $\bar{c}$ such that the socially optimal investment in flexible resource K^* increases with demand correlation α if $c_K \leq \bar{c}$, but decreases with demand correlation if $c_K > \bar{c}$.*

Proof. See Appendix A. □

The social planner only invests in flexible resources because $c_K < c_L$. As the demand correlation increases, the probability of getting either high demand realizations or low demand realizations from both firms increases, meaning firms either “win big” or “lose big”. The impact of an increase in demand correlation therefore depends on the cost of the flexible resource. If the investment cost is sufficiently low, then “losing” is cheap and the social planner would focus on reaping the benefits of high demand realizations. Therefore, investment increases with correlation for low cost. On the contrary, if investment cost is large enough, the social planner aims at minimizing the risk of “losing”, so investment decreases with correlation.

4 Monopoly

Suppose now that there is a monopoly provider for the flexible resource that chooses p and K to maximize its expected profit. Since each consumer has inelastic demand, each firm will set the price equal to r . This means that a firm will buy flexible resources to serve its customers whenever the price for flexible resource is less than r —in other words, firms’ demand for flexible resources is also inelastic. Given L_1 , L_2 , K and monopoly price p^m , the demand for flexible resource is the same as in Figure 1 as long as $p^m \leq r$. As the monopolist can extract all the value of the flexible resource, it is obvious that

$$p^m = r \tag{2}$$

in Stage 3.

The investment of the provider is determined by

$$\Omega(L_1, L_2, K) = 1 - \int_0^L \int_0^{L+K} h(x, y) dy dx + \int_L^{L+K} \int_0^{2L+K-x} h(x, y) dy dx = c_K, \tag{3}$$

In Stage 1, firm 1, expecting that $p^m = r$, chooses its local capacity L_1 so as to maximize its profit:

$$\max_{L_1} \int_0^{L_1} x f(x) dx + \int_{L_1}^{\infty} L_1 f(x) dx - c_L L_1.^{14}$$

The first two terms show that the whole demand is served when demand is below local capacity, whereas capacity is saturated when demand exceeds local capacity. The last term represents the total spending in local capacity.

Then, the first-order condition determines the equilibrium investment of L_1 :

$$1 - F(L_1) \leq c_L. \tag{4}$$

¹⁴The firm only gets positive profit from its local capacity because the surplus of the consumers, who are served by utilizing the flexible resource, are extracted entirely.

The second-order condition is also satisfied.

Analogously, for firm 2, the equilibrium investment of L_2 is determined by

$$1 - G(L_2) \leq c_L. \quad (5)$$

The market equilibrium is characterized by Equations (2), (3), (4) and (5). Let L^m and K^m denote the solutions of this set of equations. It is clear that the monopoly case is different from the social optimum because firms invest in a positive amount of local capacities, and it is also different from the duopoly case because firms' investments are independent of the provider's investment strategy.

Proposition 2. *The monopolistic flexible resource provider underinvests in the flexible resource relative to the social optimum ($K^m < K^*$), while the firms overinvest in their local capacity ($L^m > L^*$).*

Proof. See Appendix B. □

The intuition behind Proposition 2 is as follows. The monopolist sells the flexible resource at the monopoly price, which extracts all consumer surplus. Anticipating this, the firm will invest in L , even if L is the less efficient technology compared with K , in order to gain part of the consumer surplus. As a consequence, the benefit of investing in the flexible resource is lower for the monopolist than for the social planner, and hence the monopolist underinvests. Notice that Proposition 2 holds more generally for any rationing rule. The reason is that firms pay the monopoly price and, thus, all their profits in serving a customer with the flexible resource are extracted, so rationing rules do not affect firms' profits and hence their investment incentives.

To solve the problem, the regulator can ban local investments of the firms. However, this is a rather heavy-handed approach. Firms may prefer local resources for a variety of legitimate reasons. For instance, flexible resources are valuable for the firm as they offer the flexibility to modify a prior investment strategy as more information becomes available over time. More particularly, in case of “good news” the firm can scale up their services, and in case of “bad news” it can scale down. Therefore, firms are willing to pay extra to buy the flexible resource even though it is more expensive ($p^m > c_L$). Indeed, statistics shows that cloud computing is appealing to industries that have high variability in data traffic such as medical research and drug discovery in the healthcare sector.¹⁵

Therefore, I consider an alternative policy of increasing competition. Since surplus appropriation originates from market power, it seems reasonable to investigate whether introducing more competition in the market—thereby forcing down the price—would incentivize the provider and the firms to behave optimally. This issue is particularly relevant to the market for cloud computing, as providers, such as Amazon, Google, Microsoft, IBM, and Dropbox, have been competing aggressively to lower the prices of their cloud storage over the past year. The important question is then how competition would affect investment decisions, and the presence of demand correlation can make it even harder to answer. Indeed, as we will see later,

¹⁵World Economic Forum (2010) identifies the healthcare industry as one of the major sectors which can benefit from cloud computing.

the extent to which competition improves investment incentives is subtler than it appears as it varies with demand correlation and investment cost.

Let us now turn to the impact of correlation.

Proposition 3. *Under monopoly, there is a positive amount of local investment by the firms (i.e. $L^m > 0$). Moreover, there exists $\underline{c}$ such that the monopolist's investment in flexible resource K^m increases with demand correlation α if $c_K \leq \underline{c}$, but decreases with demand correlation if $c_K > \underline{c}$.*

Proof. See Appendix C. □

The impact of an increase in demand correlation on both socially optimal and equilibrium investment depends on whether the flexible resource is significantly more efficient than the local resource. The intuition of Proposition 3 is in the same spirit as Proposition 1. However, the monopolist's investment is more likely to be decreasing in demand correlation as shown in the following corollary:

Corollary 1. *The smallest c_K under which investment in flexible resource decreases with demand correlation is larger at the social optimum than under monopoly, i.e. $\bar{c} > \underline{c}$.*

Proof. See Appendix D. □

The intuition behind Corollary 1 is that local investment is zero at the social optimum but positive in the monopoly case. Thus, while the monopolist has to bear the risk associated with low demand, under which the firms will rely on their local capacity only, there is no such risk under social optimum, making the social planner more willing to invest in flexible resources as demand correlation increases compared to the monopolist.

It is also easy to show that the maximum number of consumers that can be served by each firm is higher under social optimum than monopoly, i.e. $K^* > L^m + K^m$.¹⁶ This is because the social planner invests more in flexible resources, which is cheaper relative to fixed resources and can be used by either firm.

5 Duopoly

Now consider competing providers, who compete in their prices, as in the case of Amazon AWS versus Microsoft Azure. The timing of the game is the same as before.¹⁷ Local investments in the first stage, L_1 and L_2 , generate a stochastic demand function for the flexible resource $G(L_1, L_2)$ that is price-inelastic. The continuation game is a capacity-price game with demand uncertainty, which is considered by de Frutos and Fabra (2011), henceforth FF. However, there are two important differences between this model and FF. First, FF only consider one type of

¹⁶Since one firm's local capacity cannot be used by another firm, it seems more reasonable to use the maximum capacity that one firm can deliver (i.e. $L^m + K^m$) instead of the total capacity that both firms can deliver (i.e. $L^m + L^m + K^m$) as the basis of comparison.

¹⁷Since there is demand uncertainty, the classical result of Kreps and Scheinkman (1983), who prove outcome equivalence between the capacity-price game and the Cournot game, is not applicable.

resources as opposed to two types in this model. Second, FF focus on the characterization of the equilibrium capacity and price decisions under demand uncertainty and imperfect competition (i.e. under competition, how much should a firm invest and what price should it set?), while here I focus on analyzing the same choice variables, but with welfare assessment under monopoly and competition (i.e. will competition promote higher levels of welfare?).

To solve the continuation game, we can apply the following result of FF:

Lemma 1. *(de Frutos and Fabra, 2011, Proposition 7, 8 and 9) Given L_1 and L_2 , the only equilibrium in the pricing stage is a mixed-strategy equilibrium; and capacity choices are asymmetric. In particular, there is a large firm with capacity $k^+(L_1, L_2)$ and a small firm with capacity $k^-(L_1, L_2) < k^+(L_1, L_2)$. Furthermore, if the density function of demand $g(L_1, L_2)$ is non-decreasing, then the equilibrium is unique.*

The introduction of demand uncertainty rules out the existence of symmetric equilibria because given the rival's investment strategy, the marginal return of investment is different between the large firm and the small firm.

Moving to the local investment stage, however, we can no longer follow FF's analysis because in FF, there is only one type of resources, which corresponds to the flexible resource in this model. Moreover, they assume that demand is exogenously given, while the demand for flexible resources in this paper is endogenously determined by investments in local capacity and the strength of demand correlation. Therefore, unlike the monopoly case, firms' investments are no longer independent of the provider's strategy, which poses two significant challenges for my analysis. First, the endogenously determined demand function for the flexible resource is not necessarily non-decreasing, which means that the equilibrium in the continuation game may not be unique. If the equilibrium is not unique, I focus on the most "symmetric" case in the sense that the difference in capacity between the large firm and the smaller firm is minimized, which in turn means that the degree of competitiveness is maximized.

Second, this introduces strategic interaction between the two firms: each firm's investment changes the total demand for flexible resources, which affects providers' investments and in turn affects the rival firm's investment. The complexity of providers' strategies in the pricing stage makes it extremely difficult to characterize the objective function of each firm. Furthermore, the analysis is complicated by the rationing rule, as different rationing rules generate different expected profit for each firm. To simplify the analysis, I assume that L_1 and L_2 are chosen cooperatively such that $L_1 = L_2 = L$, so that the rationing rule does not affect investments in local and flexible resources. The justification for this assumption is that different government departments and different university departments may buy local resources (e.g. on-premise computer servers) cooperatively, but each department cannot use other departments' local capacity.¹⁸ Under this assumption, the objective function of the two firms could be simply represented by the maximization of their joint profit:

$$\max_L [S(L) - \pi^+(L) - \pi^-(L)] - 2c_L L,$$

¹⁸Even though I assume cooperative investment, the two firms act differently from the case with a single firm because the two firms cannot share their local capacity, as illustrated by this example.

where $S(L)$ is the social surplus given by Equation (1), and $\pi^+(L)$ and $\pi^-(L)$ are the equilibrium profits of the two providers of flexible resource. The surplus is shared between the firms and the providers (but not the consumers). This is because demand is inelastic, so firms can extract all consumer surplus.

Solving the above problem yields the equilibrium investment in local capacity L_1^d and L_2^d , where d denotes duopoly. Then, we can also determine the equilibrium aggregate investment in flexible resource $K^d(L_1^d, L_2^d)$.

Proposition 4. *Under duopoly, investment in flexible resources is larger ($K^d \geq K^m$), investment in local resources is smaller ($L^d \leq L^m$), and the price of flexible resources is lower ($p^d \leq p^m$), compared to the monopoly case. Thus, competition always increases social welfare.*

Proof. See Appendix F. □

Competition increases welfare because it mitigates the underinvestment problem in flexible resources and the overinvestment problem in local resources. In markets with price-inelastic demands but without flexible investments, the introduction of competition has no effect on welfare because inelastic demand allows firms to extract the entire consumer surplus. However, with investments, better technology choice improves welfare—a result that is absent in FF. FF focus more on the optimal capacity choice under competition, but not on how competition solves the divergence in investment incentives between the social planner and the monopolist that is the focus of this model. More interestingly, I illustrate through the following numerical example that the extent of improvement brought about by competition depends on the cost of capacity and the degree of correlation.¹⁹

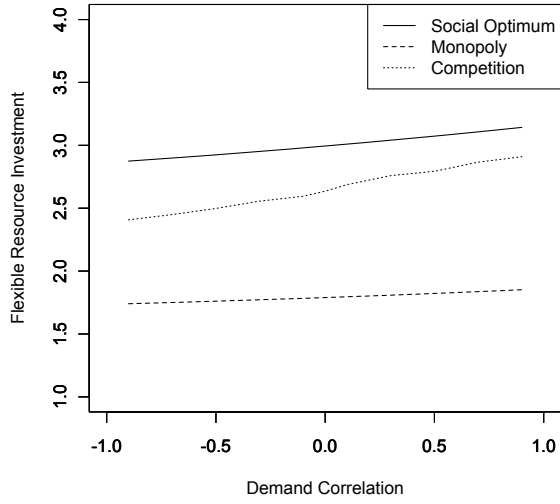
Figure 2 plots, for a given c_K , flexible resource investment against demand correlation. Social optimum is shown with a solid line, the duopoly case is drawn as a dotted line, and the monopoly case is illustrated by a long-dashed line.

The main observations in Figure 2 are summarized in the following remark.

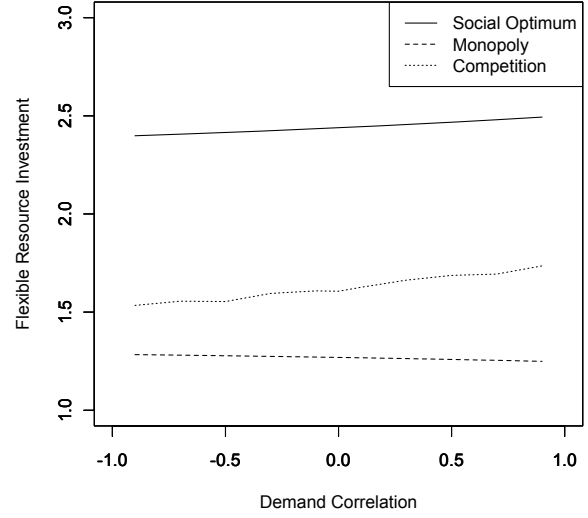
Remark 1. *Comparing the socially optimal, monopoly and duopoly solutions,*

- (i) *When c_K is sufficiently small, both the social planner's investment in flexible resources K^* and the monopolist's investment in flexible resources K^m increase with correlation α . As c_K approaches c_L , both of these investments decrease with correlation. The threshold level such that the impact of correlation changes is larger at the social optimum than it is under monopoly.*
- (ii) *Under duopoly, it can be shown that when demands are very positively or very negatively correlated, the investment in flexible resource K^d is increasing in correlation α .*

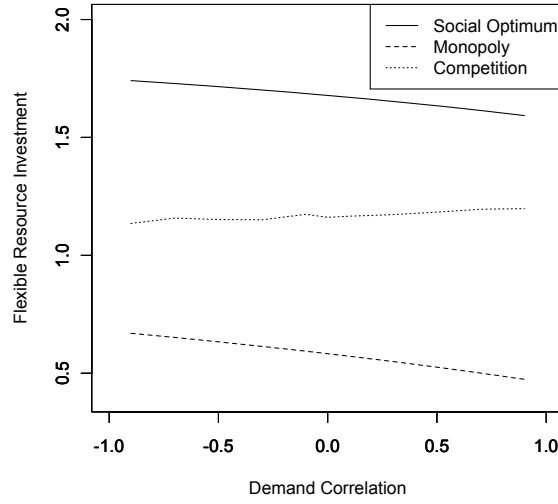
¹⁹The main difficulty in solving for an explicit solution in the duopoly case stems from the fact that the demand for flexible resources is endogenously determined by L and α , and this, in turn, affects the mixed strategy in prices of the provider. Consequently, it is difficult to characterize the profit function of the firm without using a numerical method.



(a) $c_K = 0.2$



(b) $c_K = 0.3$



(c) $c_K = 0.5$

Figure 2: Flexible Resource Investment and Demand Correlation for different values of c_K .

Part (i) is already shown in Propositions 1 and 3, and Corollary 1. As for part (ii), the intuitive reason for not observing a negative relationship between investment and correlation under duopoly, unlike the socially optimal and monopoly regimes, is as follows. Under the socially optimal and monopoly regimes, local investment does not vary with correlation: at the social optimum local investment is zero; in the monopoly case firms pay the monopoly price, and thus their local investment is not affected by correlation. Unlike these regimes, in the duopoly case firms pay less than the monopoly price and are therefore more willing to switch to buying the flexible resource in order to capture the possible windfall of high demand realizations. As a consequence, firms invest less in local capacities, and hence providers invest more in flexible resources as correlation increases.

6 Discussion

In this section I discuss some policy issues in the markets for cloud computing and the “sharing economy”.

6.1 The Market for Cloud Computing

Cloud computing has emerged as a new business model for computing and storage resource management for firms, and a new source of entertainment and communication services for consumers. As the cloud market is still in its infancy, many classic economic issues such as pricing, investment strategies, the appropriate market structure, competition policy, privacy and security concerns are still unclear.²⁰ I take the first step to understand the impact of competition on investment in this industry.

Although there are a number of competitors in the cloud computing market such as AWS, Azure, Google and IBM/SoftLayer, market power exists. For instance, large cloud providers build hyperscale datacenters that exhibit significant increasing returns to scale, which could come from the centralization of computing resources or from volume discount on the components that these providers use to build their datacenter.²¹ As a result, smaller firms may not be able to compete with these incumbents. Moreover, many consumers prefer to buy services from well-known brands because they expect higher quality. This raises concerns about the degree of competition in this market.

This model predicts that the impact of competition on investment depends crucially on the investment cost. It is often argued that cloud computing reduces the cost of investing in computing power significantly. While the marginal cost of producing an extra unit of storage or computing power is close to zero, the costs of electricity for powering up the machines, cooling the systems, as well as management, maintenance and implementation of the software and hardware in a large server farm is far from negligible.²² Also, although storing a large amount

²⁰Recently, there has been a flurry of research on the opportunities and obstacles for the adoption of cloud services; see, for example, Armbrust et al. (2009), Harms and Yamartino (2010), and Marston et al. (2011). They mainly focus on three layers of the cloud architecture: infrastructure, platform, and application. However, as argued in Bayrak et al. (2011), such categorization are useful only in defining technological differences, but not so much in analyzing their economic impact. Indeed the existing literature on cloud computing are mostly descriptive, and only rarely is the problem approached from the theoretical perspective. Fershtman and Gandal (2012) raise important economic issues of cloud computing such as changes in the strength of network effects, compatibility among software applications, the development of standards, and the market structure that should emerge. However, most of these topics have already been well-documented in a separate literature; in order to work on theoretical advancement, one needs to clearly delineate the unique features of the cloud computing market.

Recent efforts to expand the theoretical study of cloud computing include Wang (2014), who studies the adoption of cloud services within a moral hazard framework, and this paper. However, they differ in two respects. First, this paper is about capacity investment, while Wang focuses on the problem of migration, which means that there is no investment on the provider’s side. Second, this paper studies the effect of competition, but such effect is absent in Wang.

²¹See Harms and Yamartino (2010) for more examples of how firms benefit from economies of scale.

²²In September 2012, the New York Times reported that “the digital warehouses use about 30 billion watts

of data is not necessarily expensive, managing large data sets and retrieving the relevant data at the right time can be costly. Therefore, information on the cost structure in the cloud computing market should be gathered and analyzed as it has important consequences for investment.

6.2 Spot Pricing

Although it is always more efficient for firms to use the flexible resource, there are two reasons that prevent everyone from using the flexible resource only: first, the stochastic nature of demand prevents the provider from contracting over the amount of investment *ex ante*; second, the provider of the flexible resource cannot commit to marginal-cost pricing. As a consequence, firms rely more on local capacity and the provider underinvests.

Throughout the paper, I focus on non-contingent pricing.^{23,24} This is because contingent pricing is not very popular in the market for cloud computing: with the exception of AWS, which uses both contingent and non-contingent pricing, other large cloud providers such as Azure, Google and IBM rarely use spot pricing. On the contrary, in the electricity wholesale market, electricity is bought and sold at spot prices.²⁵ Yet, there is only one kind of capacity: firms typically buy energy from electricity companies, but do not generate their own electricity (although some firms may have their own emergency electricity generator, they are not for regular use). As argued by Carr (2005) and Jeff Bezos in Stone (2013), they both envisioned today's IT supply would transform from companies' private capacity into a centralized utility service, just like how electricity became a utility a century ago. It is therefore interesting to think about how spot pricing can change investment incentives in an environment with both flexible and local resources, as in the case of cloud computing, where firms buy flexible resources for its instant scalability and own local resources for reasons of regulatory compliance, data security and privacy. A formal model of contingent pricing would entail a trade-off as follows: the provider tends to price high during peak periods, which induces firms to invest more in local capacity; but it tends to price low during off-peak periods, which induces firms to rely more on flexible resources. Consequently, the extent to which investment is distorted depends on the relative strength of these two effects. If, for instance, the second effect dominates, then contingent pricing can potentially remedy the problem of underinvestment in flexible resources. Despite this additional trade-off created by contingent pricing, investment decision still depends

of electricity, roughly equivalent to the output of 30 nuclear power plants."

²³Non-contingent pricing means that prices are determined before demand is realized, whereas contingent pricing are state-dependent.

²⁴This paper focuses on linear pricing, but one can also think of other pricing structures such as non-linear tariffs. For example, cloud providers such as Amazon, Dropbox and Google use non-linear pricing for their storage service: they provide basic service for free, and then offer additional storage capacity for a fee. It is interesting to analyze the optimal nonlinear tariff under demand uncertainty, but such topic would require substantial extension which is beyond the scope of this paper.

²⁵The electricity literature (see, for instance, Borenstein and Holland (2005), Murphy and Smeers (2005), Joskow and Tirole (2007), and Léautier (2011)) mostly considers a two-stage game, in which firms choose their capacity first, and then they bid prices for each state of the world in a spot market. See also Crew, Fernando and Kleindorfer (1995) for a survey of the literature on peak-load pricing.

fundamentally on the degree of correlation and the cost of capacity, and therefore all the main qualitative results of this paper should remain valid.

Spot pricing is not only common for utilities, but it has also become more widely used in other “sharing economy” marketplaces, where platforms such as Uber and Airbnb vary their prices depending on demand. These platforms can similarly invest in multiple resources (flexible vs. fixed) and demand is correlated across firms. For instance, Uber and Airbnb, instead of owning the hotels/taxis and hiring their own staffs/drivers, act as intermediaries between buyers and a group of independent contractors. Moreover, hotels have huge demand spikes during holiday seasons, and people most want to share a car during busy times and under bad weather. Thus, at a more general level, my results could be extrapolated to gain a better understanding on investment decisions in these markets.

7 Conclusion

This paper has analyzed firms’ incentives to invest in local and flexible resources. I find that market power of the monopolist providing flexible resources distorts investment, while competition restores efficiency. The extent of improvement depends on demand correlation and investment costs. If the investment cost is small, investment under social optimum, monopoly and competition is increasing in correlation; if the cost is large, investment under competition is still increasing in correlation, whereas that under social optimum and monopoly goes in opposite direction.

These results have implications for investment decision on whether to outsource, particularly in the market for cloud computing. Admittedly, the cloud computing market is growing unpredictably, and there is no clear indication or consensus on how it will develop. For now, this paper shows that even if the cloud computing market follows the footsteps of the electricity market and providers eventually adopt spot pricing, a similar trade-off that we derived here will arise. Therefore, analyzing data on cost and demand represents a useful first step towards a fuller understanding of the nascent industry.

I list some important topics that lie beyond the scope of this paper, but would be appropriate for further work. The first is to consider product differentiation. For example, assuming that cloud computing services (such as Dropbox storage services) and local storage services are differentiated—how, then, would the investment strategy change? Second, it would be interesting to study the consequences of vertical integration. For instance, what will happen if upstream cloud computing firms such as Microsoft and Google also enter the downstream market of software applications? Finally, one could examine the potential merits of policies such as a subsidy for investment in flexible resource to remedy the underinvestment problem. While a full analysis of the effect of a subsidy (e.g. how to finance it) is beyond the scope of this paper, preliminary analysis suggests that if the regulator were to introduce a subsidy s for investment in flexible resources, the cost of flexible resources will become $c_K - s$. The provider will then be more willing to offer a lower price, and at the same time, it also has more incentives to undertake investment in flexible resources, which could potentially mitigate the underinvestment problem.

Appendices

A Proof of Proposition 1

The social optimum is obtained by differentiating Equation (1) with respect to L_1 , L_2 and K .

The F.O.C. with respect to L_1 is given by

$$\{IV\} + \{V\} \leq c_L.$$

Similarly, the F.O.C. with respect to L_2 is

$$\{III\} + \{V\} \leq c_L.$$

Finally, the F.O.C. with respect to K is:

$$\{III\} + \{IV\} + \{V\} \leq c_K.$$

As $\{III\} + \{IV\} + \{V\} > \{IV\} + \{V\}$ or $\{III\} + \{V\}$, the marginal benefit of investing in the flexible resource is always higher than that of local capacity. Furthermore, the marginal cost of investing in the flexible resource is lower ($c_K < c_L$). Then we must have $L_1^* = L_2^* = 0$. Since $c_K < c_L < r$, all F.O.C. are satisfied with equality.

The socially optimal investment in the flexible resource is determined by the F.O.C. with respect to K , which can be rewritten as

$$F(K, \alpha, c_K) = \int_0^K \int_0^{K-x} h(x, y) dy dx - 1 + c_K = 0.$$

By implicit function theorem,

$$\frac{\partial K}{\partial \alpha} = -\frac{\frac{\partial F}{\partial \alpha}}{\frac{\partial F}{\partial K}}.$$

We can show that

$$\frac{\partial F}{\partial K} = \int_0^K e^{-K} [1 + \alpha(2e^{-x} - 1)(2e^{x-K} - 1)] dx$$

is positive. Moreover, we have

$$\begin{aligned} \frac{\partial F}{\partial \alpha} &= \int_0^K \int_0^{K-x} e^{-x-y} (2e^{-x} - 1)(2e^{-y} - 1) dy dx \\ &= -e^{-K} [K + 3e^{-K} + 2Ke^{-K} - 3]. \end{aligned}$$

It can be shown that there exists a $\bar{K}^*$ such that $\frac{\partial F}{\partial \alpha} < 0$ when $K > \bar{K}^*$, and $\frac{\partial F}{\partial \alpha} > 0$ when $K < \bar{K}^*$. In addition, it is obvious that K decreases with c_K . Therefore, if c_K is small such that $K > \bar{K}^*$, then $\frac{\partial K}{\partial \alpha} > 0$. On the contrary, if c_K is large, K is small such that $K < \bar{K}^*$, then $\frac{\partial K}{\partial \alpha} < 0$.

B Proof of Proposition 2

For firm 1, its equilibrium investment is determined by

$$1 - F(L_1) = c_L,$$

As $1 - F(L_1) > \{IV\} + \{V\}$, we must have $L_1^m > L_1^* = 0$, and hence there is overinvestment. The same happens for firm 2.

For the flexible resource provider, its equilibrium investment K^m is determined by

$$\begin{aligned} \max_K \Pi = & \int_0^{L_1} \int_{L_2}^{L_2+K} (y - L_2)h(x, y)dydx + \int_{L_1}^{L_1+K} \int_0^{L_2} (x - L_1)h(x, y)dydx \\ & + \int_{L_1}^{L_1+K} \int_{L_2}^{L_1+L_2+K-x} (x + y - L_1 - L_2)h(x, y)dydx \\ & + K \left[\int_0^{L_1} \int_{L_2+K}^{\infty} h(x, y)dydx + \int_{L_1+K}^{\infty} \int_0^{L_2} h(x, y)dydx \right. \\ & \left. + \int_{L_1}^{\infty} \int_{L_2}^{\infty} h(x, y)dydx - \int_{L_1}^{L_1+K} \int_{L_2}^{L_1+L_2+K-x} h(x, y)dydx \right] - c_K K. \end{aligned}$$

which gives us

$$\Omega(L_1^m, L_2^m, K^m) = c_K = \Omega(0, 0, K^*).$$

Suppose that the flexible resource provider invests K such that $L^m + K = K^*$. Since $L^m > 0$, it must be $\Omega(L_1^m, L_2^m, K) < \Omega(0, 0, K^*)$, which means such K cannot be the equilibrium. Therefore, the flexible resource provider must invest K^m such that $L^m + K^m < K^*$, which means that $K^m < K^*$ (underinvestment).

C Proof of Proposition 3

The monopolist's investment is determined by

$$F(K, \alpha, c_K) = \int_0^L \int_0^{L+K} h(x, y)dydx + \int_L^{L+K} \int_0^{2L+K-x} h(x, y)dydx - 1 + c_K = 0.$$

By implicit function theorem,

$$\frac{\partial K}{\partial \alpha} = -\frac{\frac{\partial F}{\partial \alpha}}{\frac{\partial F}{\partial K}}.$$

It is straightforward to show that

$$\begin{aligned} \frac{\partial F}{\partial K} = & \int_0^L e^{-x-L-K} [1 + \alpha(2e^{-x} - 1)(2e^{-L-K} - 1)]dx \\ & + \int_0^L e^{-y-L-K} [1 + \alpha(2e^{-y} - 1)(2e^{-L-K} - 1)]dy \\ & + \int_L^{L+K} e^{-2L-K} [1 + \alpha(2e^{-x} - 1)(2e^{-x-2L-K} - 1)]dy > 0, \\ \frac{\partial F}{\partial \alpha} = & \int_0^L \int_0^{L+K} e^{-x-y} (2e^{-x} - 1)(2e^{-y} - 1)dydx \\ & + \int_L^{L+K} \int_0^{2L+K-x} e^{-x-y} (2e^{-x} - 1)(2e^{-y} - 1)dydx. \end{aligned}$$

Similar to the proof in Appendix A, there exists $\bar{K}^m$ such that $\frac{\partial F}{\partial \alpha} < 0$ when $K > \bar{K}^m$; and $\frac{\partial F}{\partial \alpha} > 0$ when $K < \bar{K}^m$. Moreover, as K^m is decreasing in c_K , then if c_K is small such that $K > \bar{K}^m$, then $\frac{\partial K}{\partial \alpha} > 0$. On the contrary, if c_K is such that $K < \bar{K}^m$, then $\frac{\partial K}{\partial \alpha} < 0$.

D Proof of Corollary 1

From the proof in Appendices A and C, it suffices to show $\frac{\partial F^*}{\partial \alpha}(K^*) < \frac{\partial F^m}{\partial \alpha}(L^m, K^m)$, where both terms integrate the same function over the respective area as shown in Figure 3. The difference between $\frac{\partial F^*}{\partial \alpha}(K^*)$ and $\frac{\partial F^m}{\partial \alpha}(L^m, K^m)$ lies in the shaded area. Comparing integrations over the triangles and the trapezium, we can conclude that the above condition is satisfied because the triangles have higher values of x or y .

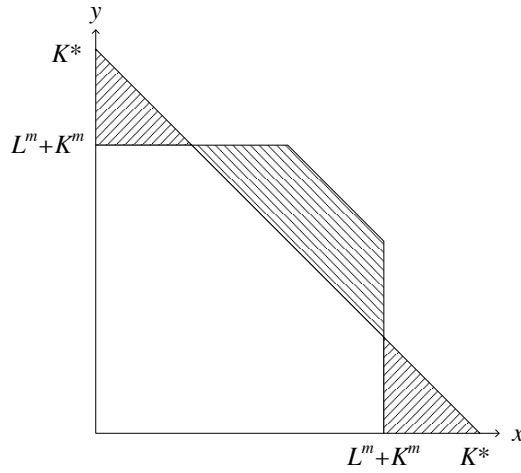


Figure 3: Investment under Social Optimum and Monopoly.

We therefore have

- If $\frac{\partial F^m}{\partial \alpha} < 0$, then $\frac{\partial F^*}{\partial \alpha} < 0$. Both $\frac{\partial K^*}{\partial \alpha}, \frac{\partial K^m}{\partial \alpha} > 0$, which is true for small c_K .
- If $\frac{\partial F^*}{\partial \alpha} > 0$, then $\frac{\partial F^m}{\partial \alpha} > 0$. Both $\frac{\partial K^*}{\partial \alpha}, \frac{\partial K^m}{\partial \alpha} < 0$, which is true for large c_K .
- For medium c_K , $\frac{\partial F^*}{\partial \alpha} < 0$ and $\frac{\partial F^m}{\partial \alpha} > 0$. Then, $\frac{\partial K^*}{\partial \alpha} > 0$ and $\frac{\partial K^m}{\partial \alpha} < 0$.

Thus, under social optimum there is a larger range of c_K under which investment increases with correlation as compared to the monopoly case.

E Linear Example

E.1 Social Optimum

The relationship between investment in flexible resource and demand correlation at the social optimum is slightly different when demands are uniformly distributed. To see this, consider a joint distribution $h(x, y)$ as follows:

- *Positive correlation.* With probability ρ , only pairs of demands on the $x = y$ line are possible (perfect positive correlation). With probability $1 - \rho$, demands are uniformly distributed on a unit square $[0, 1] \times [0, 1]$ (independent demands). We can use ρ as a measure of positive correlation.
- *Negative correlation.* With probability ρ , only pairs of demands on the $x + y = 1$ line are possible (perfect negative correlation). With probability $1 - \rho$, demands are uniformly spread over a unit square $[0, 1] \times [0, 1]$ (independent demands). We can use $-\rho$ as a measure of negative correlation.

Since $c_K < c_L < r$, all the F.O.C. are satisfied with equality. In the case of positive correlation, the optimal capacity is chosen such that the marginal benefit is equal to the marginal cost:

$$\rho(1 - \frac{K}{2}) + (1 - \rho)\frac{1}{2}(2 - K)^2 = c_K.$$

Note that $K \geq 1$ because $c_K \leq 0.5$. Differentiating K with respect to ρ , we find that K^* increases with ρ .

In the case of negative correlation, we have

$$K^* = \max \left\{ 1, 2 - \sqrt{\frac{2c_K}{1 - \rho}} \right\}.$$

Note that $K \geq 1$. The reason is that if demands are perfectly negatively correlated and investment is less than 1, then marginal benefit always exceeds cost. When $K > 1$, the optimal investment is determined by

$$(1 - \rho)\frac{1}{2}(2 - K)^2 = c_K.$$

It is easy to see that K^* increases with $-\rho$.

We therefore have

Result 1. *In the case of uniformly distributed demands, the social planner only invests in the flexible resource (i.e. $L^* = 0$ and $K^* > 0$), and the socially optimal investment K^* always increases with demand correlation ρ .*

The reason is that, for uniformly distributed demands, the marginal benefit of expanding capacity always increases as correlation increases.

E.2 Monopoly Case

In the monopoly case, the result in the linear example is the same as Proposition 3 in the main text. To keep things simple, further assume that $c_K \in [0.25, 0.5]$ such that $L_1 + K$ and $L_2 + K$ are smaller than 1. In the case of positive correlation, the monopolist chooses K such that

$$\rho(\frac{1}{2} - \frac{K}{2}) + (1 - \rho)(\frac{3}{4} - K - \frac{1}{2}K^2) = c_K.$$

In the case of negative correlation, the monopolist choice of K solves

$$\rho(1 - 2K) + (1 - \rho)(\frac{3}{4} - K - \frac{1}{2}K^2) = c_K.$$

We therefore have

Result 2. *In the case of uniformly distributed demands, there is a positive amount of local investment by the firms (i.e. $L^m > 0$); and the monopolist's investment in flexible resource K^m increases with demand correlation ρ if c_K is small, but decreases with demand correlation if c_K approaches c_L .*

F Competition Improves Social Welfare

Competition increases social welfare because it mitigates the underinvestment and overinvestment problem.

- $K^d \geq K^m$: The F.O.C. of K in the monopoly case is

$$\{III\} + \{IV\} + \{V\} = c_K.$$

As for the duopoly case, we refer to Equation (12) in FF: the F.O.C. of K is

$$1 - D(K) = c_K,$$

where $D(K)$ is the demand for flexible resources. Since firms only buy the flexible resource when demand is above their local capacity, this condition can be rewritten as

$$\frac{\{III\} + \{IV\} + \{V\}}{1 - \int_0^{L_1} \int_0^{L_2} h(x, y) dy dx} = c_K.$$

Therefore, $K^d \geq K^m$ because $1 - \int_0^{L_1} \int_0^{L_2} h(x, y) dy dx < 1$. Note that $K^d = K^m$ only when $L_1, L_2 = 0$.

- $p^d \leq p^m$: Under duopoly, providers of the flexible resource randomize over price with the upper bound of r (see Proposition 7 of FF).
- $L^d \leq L^m$: Firms invest less in local resource under duopoly because the price of it is lower.

References

- [1] Georgios Angelou and Anastasios Economides. Flexible ICT investment analysis using Real Options. *International Journal of Technology, Policy and Management*, 5(2):146–166, 2005.
- [2] Ravi Anupindi and Li Jiang. Capacity Investment under Postponement Strategies, Market Competition, and Demand Uncertainty. *Management Science*, 54(11):1876–1890, 2008.
- [3] Michael Armbrust, Armando Fox, Rean Griffith, Anthony Joseph, Randy Katz, Andrew Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. Above the clouds: A Berkeley view of cloud computing. Technical Report No. UCB/EECS-2009-28, 2009. available at <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>.

- [4] Ergin Bayrak, John Conley, and Simon Wilkie. The Economics of Cloud Computing. *The Korean Economic Review*, 27(2):203–230, 2011.
- [5] Michel Benaroch and Robert Kauffman. A Case for Using Real Options Pricing Analysis to Evaluate Information Technology Project Investments. *Information Systems Research*, 10(1):70–86, 1999.
- [6] Ebru Bish and Qiong Wang. Optimal Investment Strategies for Flexible Resources, Considering Pricing and Correlated Demands. *Operations Research*, 52(6):954–964, 2004.
- [7] Severin Borenstein and Stephen Holland. On the Efficiency of Competitive Electricity Markets with Time-invariant Retail Prices. *RAND Journal of Economics*, 36(3):469–493, 2005.
- [8] Timothy Bresnahan and Jonathan Levin. Vertical Integration and Market Structure. In Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, chapter 21, pages 853–890. Princeton University Press, 2013.
- [9] Nicholas Carr. The End of Corporate Computing. *MIT Sloan Management Review*, pages 67–73, Spring 2005.
- [10] Michael Crew, Chitru Fernando, and Paul Kleindorfer. The Theory of Peak-load Pricing: A Survey. *Journal of Regulatory Economics*, 8:215–248, 1995.
- [11] María-Ángeles de Frutos and Natalia Fabra. Endogenous Capacities and Price Competition: The role of demand uncertainty. *International Journal of Industrial Organization*, 29:399–411, 2011.
- [12] Chaim Fershtman and Neil Gandal. Migration to the Cloud Ecosystem: Ushering in a New Generation of Platform Competition. *Communications & Strategies Digiworld Economic Journal*, 85(1):109–123, 2012.
- [13] Manu Goyal and Serguei Netessine. Strategic Technology Choice and Capacity Investment under Demand Uncertainty. *Management Science*, 53(2):192–207, 2007.
- [14] Emil Julius Gumbel. Bivariate Exponential Distributions. *Journal of the American Statistical Association*, 55(292):698–707, 1960.
- [15] Rolf Harms and Michael Yamartino. The Economics of the Cloud. Microsoft White Paper, 2010.
- [16] Paul Joskow and Jean Tirole. Reliability and Competitive Electricity Market. *RAND Journal of Economics*, 38(1):60–84, 2007.
- [17] Robert Kauffman, Henry Lucas, Paul Tallon, Andrew Whinston, and Kevin Zhu. Using Real Options Analysis for Evaluating Uncertain Investments in Information Technology: Insights from the ICIS 2001 Debate. *Communications of the Association for Information Systems*, 9:136–167, 2002.

- [18] David Kreps and Jose Scheinkman. Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *The Bell Journal of Economics*, 14(2):326–337, 1983.
- [19] Thomas-Olivier Léautier. The Invisible Hand: ensuring optimal investment in electric power generation. Working Paper, 2011.
- [20] In Lee. A Model for Determining the Optimal Capacity Investment for Utility Computing. In Dirk Neumann, Mark Baker, Jörn Altmann, and Omer Rana, editors, *Economic Models and Algorithms for Distributed Systems*. Birkhäuser Verlag, Basel, Switzerland, 2009.
- [21] Sean Marston, Zhi Li, Subhajyoti Bandyopadhyay, Juheng Zhang, and Anand Ghalsasi. Cloud Computing - The Business Perspective. *Decision Support Systems*, 51:176–189, 2011.
- [22] Frederic Murphy and Yves Smeers. Generation Capacity Expansion in Imperfectly Competitive Restructured Electricity Markets. *Operations Research*, 53(4):646–661, 2005.
- [23] Dusit Niyato, Sivadon Chaisiri, and Bu-Sung Lee. Economic Analysis of Resource Market in Cloud Computing Environment. Paper presented to the 2009 IEEE Asia-Pacific Services Computing Conference, Singapore, 7-11 December 2009.
- [24] Stanley Reynolds and Bart Wilson. Bertrand-Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes. *Journal of Economic Theory*, 92:122–141, 2000.
- [25] Brad Stone. *The Everything Store: Jeff Bezos and the Age of Amazon*. Little, Brown and Company, United States, 2013.
- [26] Steven Tadelis and Oliver Williamson. Transaction Cost Economics. In Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, chapter 4, pages 159–189. Princeton University Press, 2013.
- [27] Tong Wang. Migration to the Cloud. Working Paper, 2014. available at https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=IIOC2014&paper_id=162.
- [28] Jan Van Miegham. Investment Strategies for Flexible Resources. *Management Science*, 44(8):1071–1078, 1998.
- [29] World Economic Forum. Exploring the Future of Cloud Computing: Riding the Next Wave of Technology-Driven Transformation, 2010. available at <http://members.weforum.org/pdf/ip/ittc/Exploring-the-future-of-cloud-computing.pdf>.

Recent titles

CORE Discussion Papers

2015/01	Yurii NESTEROV and Vladimir SHIKHMAN. Algorithm of Price Adjustment for Market Equilibrium.
2015/02	Claude d'ASPREMONT and Rodolphe DOS SANTOS FERREIRA. Oligopolistic vs. monopolistic competition: Do intersectoral effects matter?
2015/03	Yuuri NESTEROV. Complexity bounds for primal-dual methods minimizing the model of objective function.
2015/04	Hassène AISSI, A. Ridha MAHJOUB, S. Thomas MCCORMICK and Maurice QUEYRANNE. Strongly polynomial bounds for multiobjective and parametric global minimum cuts in graphs and hypergraphs.
2015/05	Marc FLEURBAEY and François MANIQUET. Optimal taxation theory and principles of fairness.
2015/06	Arnaud VANDAELE, Nicolas GILLIS, François GLINEUR and Daniel TUYTTENS. Heuristics for exact nonnegative matrix factorization.
2015/07	Luc BAUWENS, Jean-François CARPANTIER and Arnaud DUFAYS. Autoregressive moving average infinite hidden Markov-switching models.
2015/08	Koen DECANQ, Marc FLEURBAEY and François MANIQUET. Multidimensional poverty measurement with individual preferences
2015/09	Eric BALANDRAUD, Maurice QUEYRANNE, and Fabio TARDELLA. Largest minimally inversion-complete and pair-complete sets of permutations.
2015/10	Maurice QUEYRANNE and Fabio TARDELLA. Carathéodory, helly and radon numbers for sublattice convexities.
2015/11	Takatoshi TABUSHI, Jacques-François THISSE and Xiwei ZHU. Does technological progress affect the location of economic activity.
2015/12	Mathieu PARENTI, Philip USHCHEV, Jacques-François THISSE. Toward a theory of monopolistic competition.
2015/13	A.B. TAYLOR, J.M. HENDRICKX and F. GLINEUR. Smooth strongly convex interpolation and exact worst-case performance of first-order methods.
2015/14	Christian HAFNER, Hans MANNER and Léopold SIMAR. The “wrong skewness” problem in stochastic frontier models: A new approach.
2015/15	Paul BELLEFLAMME, Nessrine OMRANI Martin PEITZ. The Economics of Crowdfunding Platforms.
2015/16	Samuel FEREY and Pierre DEHEZ. Multiple Causation, Apportionment and the Shapley Value.
2015/17	Mehdi MADANI and Mathieu VAN VYVE. A MIP Framework for Non-convex Uniform Price Day-ahead Electricity Auctions.
2015/18	Yukihiro NISHIMURA and Pierre PESTIEAU. Efficient Taxation with Differential Risks of Dependence and Mortality.
2015/19	Louis N. CHRISTOFIDES, Michael HOY, Joniada MILLA and Thanasis STENGOS. Nature or Nurture in Higher Education? Inter-generational Implications of the Vietnam-Era Lottery.
2015/20	Philippe J. DESCHAMPS, Alternative Formulations of the Leverage Effect in a Stochastic Volatility Model with Asymmetric Heavy-Tailed Errors.
2015/21	Jean HINDRIKS and Yukihiro NISHIMURA. Equilibrium Leadership in Tax Competition Models with Capital Ownership: A Rejoinder.
2015/22	Frédéric VRINS and Monique JEANBLANC. The Φ -Martingale.
2015/23	Wing Man Wynne LAM. Attack-Detering and Damage Control Investments in Cybersecurity.

Recent titles

CORE Discussion Papers – continued

- 2015/24 Wing Man Wynne LAM. Switching Costs in Two-sided Markets.
- 2015/25 Philippe DE DONDER, Marie-Louise LEROUX. The political choice of social long term care transfers when family gives time and money.
- 2015/26 Pierre PESTIEAU and Gregory PONTIERE. Long-Term Care and Births Timing.
- 2015/27 Pierre PESTIEAU and Gregory PONTIERE. Longevity Variations and the Welfare State.
- 2015/28 Mattéo GODIN and Jean HINDRIKS. A Review of Critical Issues on Tax Design and Tax Administration in a Global Economy and Developing Countries
- 2015/29 Michel MOUCHART, Guillaume WUNSCH and Federica RUSSO. The issue of control in multivariate systems, A contribution of structural modelling.
- 2015/30 Jean J. GABSZEWICZ, Marco A. MARINI and Ornella TAROLA. Alliance Formation in a Vertically Differentiated Market.
- 2015/31 Jens Leth HOUGAARD, Juan D. MORENO-TERNERO, Mich TVEDE and Lars Peter ØSTERDAL. Sharing the Proceeds from a Hierarchical Venture.
- 2015/32 Arnaud DUFAYS and Jeroen V.K. Rombouts. Sparse Change-Point Time Series Models.
- 2015/33 Wing Man Wynne Lam. Status in Organizations.
- 2015/34 Wing Man Wynne Lam. Competition in the Market for Flexible Resources : an application to cloud computing.

Books

- W. GAERTNER and E. SCHOKKAERT (2012), *Empirical Social Choice*. Cambridge University Press.
- L. BAUWENS, Ch. HAFNER and S. LAURENT (2012), *Handbook of Volatility Models and their Applications*. Wiley.
- J-C. PRAGER and J. THISSE (2012), *Economic Geography and the Unequal Development of Regions*. Routledge.
- M. FLEURBAEY and F. MANIQUET (2012), *Equality of Opportunity: The Economics of Responsibility*. World Scientific.
- J. HINDRIKS (2012), *Gestion publique*. De Boeck.
- M. FUJITA and J.F. THISSE (2013), *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. (2nd edition). Cambridge University Press.
- J. HINDRIKS and G.D. MYLES (2013). *Intermediate Public Economics*. (2nd edition). MIT Press.
- J. HINDRIKS, G.D. MYLES and N. HASHIMZADE (2013). *Solutions Manual to Accompany Intermediate Public Economics*. (2nd edition). MIT Press.
- J. HINDRIKS (2015). *Quel avenir pour nos pensions ? Les grands défis de la réforme des pensions*. De Boeck.

CORE Lecture Series

- R. AMIR (2002), Supermodularity and Complementarity in Economics.
- R. WEISMANTEL (2006), Lectures on Mixed Nonlinear Programming.
- A. SHAPIRO (2010), Stochastic Programming: Modeling and Theory.