Risk Aversion and Information Asymmetry in the Pricing of Capacity-on-Demand and Pay-per-Use Computing Products

Phillip M. Yelland
Sun Microsystems Laboratories
Menlo Park, California

1 Introduction

In recent years, an increasing number of vendors of information technology have begun to offer their customers the option of purchasing equipment with computer processing, storage or other capacity that exceeds the customers basic requirement, allowing the customer to refrain from paying for the surplus capacity until it's actually required (as the customer's business grows, experiences seasonal fluctuations or the like). Thus a vendor might supply a 12-processor computer server with only 8 operational processors (at a lower initial price), with the customer paying to use the idle processors "on demand". Examples of particular COD products are described in (Sun 2005), (IBM 2005) and (HP 2005); a survey is presented in (Partridge 2004).

Almost invariably, these so-called *capacity-on-demand* offerings also incorporate some form of *pay-per-use* arrangement, so that the equipment monitors the extra capacity activated during some period (a month, for example), and the customer is charged accordingly at the end of the period. These charges are based on a rate per unit of capacity used (normally expressed as the use of a resource — a CPU or bank of processor memory, say, or area of disk storage — for a given time), such as "dollars per CPU-hour" or "dollars per gigabyte-month" (Shankland 2005). In almost all cases, the buyer is required to commit to using a certain minimum number of units at the given rate within the period (IBM 2005, HP 2005). For the remainder of this paper, we use the term "capacity-on-demand" (abbreviated *COD*) to denote the rather than the more cumbersome — if more precise — "capacity-on-demand with pay-per-use".[1]

Much of the appeal of capacity-on-demand — at least according to its vendors — is that matching a buyer's expenditures on IT equipment to the use of that equipment helps manage the risks that

[1] In fact the terminology surrounding capacity-on-demand products and their brethren is somewhat vexed: The terms "capacity-on-demand", "on-demand", "pay-per-use", "utility computing" and "agile computing" (amongst others) have all been used by vendors to describe the sort of products examined in this paper. To save confusion, we use the term "capacity-on-demand" to describe these products, which we take to comprise computing resources dedicated exclusively to one customer (and usually situated on the customer's premises) that offer additional capacity for an additional fee based upon recorded usage.

the buyer faces; this is a point made explicitly in (IBM 2005) and (HP 2005). COD products allow the buyer to purchase additional capacity should the need arise (reducing the risks due to business disruption), while not requiring payment for capacity that's never used (and which is presumably not contributing to the buyer's business).

This paper examines some of the subtleties that surround the pricing of capacity-on-demand products. In particular, it examines conditions under which provision of such products can be expected to improve a vendor's expected profits. Certain features of COD products are of particular interest in the analysis:

- Regardless of the amount of extra capacity actually used by the buyer, the seller's cost of providing that capacity is essentially fixed (it's more-or-less the cost of producing the entire piece of equipment). This significantly simplifies the analysis, in that it allows us to focus entirely on the revenues generated for the seller when calculating the effects on profits.
- As indicated above, COD products are usually offered as a form of "insurance" — in buying such products, the buyer expects to reduce the deleterious effects of uncertain events. Thus the attractiveness of COD products to a buyer generally depends upon the buyer's appetite for risk.
- As we will demonstrate, COD products can expose their vendor to the possibility of *adverse selection*; since buyers may be better informed about their future usage than is the seller, they may be able to make use of this private information at the time of the purchase to the detriment of the seller.

The remainder of the paper is as follows: In the next section, we look at some of the existing research related to the work detailed here. Then a simple model is described that attempts to capture the salient aspects of capacity-on-demand products and the issues involved in their pricing. This model is examined in detail with different assumptions concerning the information available to the seller and the buyer. In each case, the COD arrangement is compared with a more conventional purchase in which the vendor essentially sells the entire piece of equipment to the buyer upfront. We finish with general conclusions, some recommendations for practical approaches to COD pricing and for future research.

2 Related Work

The capacity-on-demand products described in this paper share many of the characteristics of so-called *utility computing* offerings.[2] The latter (as we discern the term here) provide access to *shared* centralized computing facilities on a pay-per-use basis, the idea being that a buyer of such services pays to have individual transactions executed by the computing center, in a manner reminiscent of the time-shared computing arrangements of an earlier era. Like capacity-on-demand, utility computing normally involves some sort of usage-based pricing, and also shifts the seller from variable to fixed costs of provision. For an example of utility computing products, see (Shankland 2005) for a brief overview of one vendor's line.

Investigations of the economics of utility computing have a long lineage — perhaps unsurprisingly, given that it at least partially recapitulates time-sharing. (Mendelson 1985) is a seminal reference that examines different approaches to pricing such shared computing services. More recently, research including (Paleologo 2004), (Liu et al. 2001) and (Hellerstein et al. 2004) has revisited some of the same issues, focusing particularly on the more modern context of Web service hosting.[3]

There are a number of key differences between the economics of utility computing and those of capacity-on-demand. Most notable of these is that utility computing centers are usually subject to multiple service requests from many buyers at the same time, so that many of the models used to describe them involve extensive use of queuing theory — see (Mendelson 1985), for example, or (Liu et al. 2001). By contrast, capacity-on-demand (at least the term is used here) involves a single buyer's use of a non-shared resource, eliminating much of the technical machinery associated with modeling utility computing. With utility computing, sharing of the provider's computing center alters the cost-structure of the services provided depending on the number and nature of the requests made by the purchasers — see (Paleologo 2004) for a discussion of "multiplexing gains" and Hellerstein et al.'s (2004) use of inventory-control concepts to deal with multiple requests. Again, such considerations are irrelevant in the capacity-on-demand setting, where the vendor's costs are essentially fixed. Finally, we note that in contrast to this paper, work to date on the pricing of utility computing products has paid little attention to the role of risk aversion — indeed (Mendelson 1985) mentions it only insofar as he deliberately elects to ignore it.

---

[2] Indeed, as we observed earlier, some COD products have actually been sold under the "utility computing" moniker (Sun 2005).

[3] It is perhaps interesting to contrast the literature's almost exclusive attention to the economics of utility computing systems with customers' apparent preference for COD products, as indicated in (Lyons 2005).

COD products also share some of the characteristics of *information goods* — goods such as books, music, films and computer software that can be digitized and copied at little or no cost. A prime distinguishing feature of information goods is that while there are usually fixed costs associated with their production (up-front investments in writing, filming, etc.), the corresponding variable costs (of copying) are very low. Similarly, with COD products, the costs of providing capacity (up to the maximum capacity of the equipment provided) are fixed (being the cost of producing the equipment itself), while the variable cost of enabling the idle capacity is almost negligible. Works such as (Varian 2000) and (Sundararajan 2004) examine some of the issues involved in pricing information goods, focusing particularly on the role played by transaction costs (costs associated with monitoring, billing, etc.). It is notable that Sundararajan (2004) specifically mentions IBM's "pay-per-use" zSeries software (though not hardware) as an example of an information good.

While related to that here, work to date concerning information goods differs in some important respects: First, research on information goods concentrates on primary demand in consumer markets rather than derived demand in an industrial setting. Thus in a consumer market, future demand is determined entirely by the buyer's preferences (which are usually assumed to be fully known), whereas here, the buyer's future demand may be stochastic, since it is driven in turn by the demand faced by the buyer's business. And the uncertain nature of the buyer's future demand and the role of COD products in mitigating such uncertainty also raises issues concerning the buyer's risk aversion, which is normally omitted from the treatment of information goods.

Finally, as we noted above, some of the investigations described in this paper concern the effects of information asymmetry in dealing with uncertain demand. Similar issues have recently been of concern to researchers in supply chain management, too. As with the work here, in the supply chain literature it is normally assumed that the buyer possesses private information (or is at least in an advantageous position with regard to obtaining such information) about the nature of future demand, and may be tempted to use it to the disadvantage of the supplier and the supply chain as a whole. Papers such as (Miyaoka 2005), (Lariviere 2002) and (Cachon and Lariviere, 2001) discuss the structure of contracts intended to ensure that the buyer of a product shares forecasts of stochastic demand for the product with its suppliers. An overview of the literature in this area is provided in (Cachon 2003).

Of course, though the distribution of information in these supply chain scenarios (where the buyer may be in possession of a private forecast) parallels that in this study, many of the physical details involved differ considerably. For example, production of physical inventory (in contrast to the activation of spare capacity in COD) is a costly and time-consuming activity, giving very different incentives to the participants.

3 A Simple Model of Capacity-on-Demand

The model used in this paper attempts to capture the distinguishing aspects of capacity-on-demand products in as simple a setting as possible. It features two actors: A *seller* (designated female for purposes of identification) and a (male) *buyer*. It is assumed that the seller is a monopoly (this is not too far from reality in much of the commercial information technology market, where issues such as software compatibility and training requirements normally give vendors a fair amount of "lock-in", at least in the short term). Interaction takes place over a single period. At the beginning of the period, the seller offers a number of units of "capacity" at a specified price, on a take-it-or-leave-it basis. As discussed in Section 1, this "capacity" may take the form of CPU's, processor memory, data storage, or the like that are placed at the buyer's disposal throughout the period. Also in keeping with the observation we made earlier, we assume that the seller's costs are the same regardless of the amount of capacity used by the buyer.

We consider two separate pricing schemes:

- With the *bulk* pricing scheme, the buyer pays a single price for the right to use *all* the capacity throughout the period.
- According to the *capacity-on-demand* or *COD* scheme, the capacity provided is divided into two amounts: *Base* capacity, which the buyer purchases for the duration of the period (this is the minimum usage commitment discussed in the introduction), and *contingent* capacity, which the buyer purchases (presumably after the fact) only if he uses it during the period.

For the purposes of comparing the two pricing schemes, we assume that it is possible to delimit two units of capacity — one *base* unit and one *contingent* unit — amongst those on offer. Again, such units may represent an hour of CPU time, a day's use of a gigabyte of storage, and so on; it is not required that they be the only units of capacity provided to the buyer.

With bulk pricing, the buyer is charged p_1 dollars for the right to use both base and contingent units. Under COD pricing, the buyer pays p_2 dollars for the base unit only, and pays an additional p_2 dollars in the event that he also uses the contingent unit — thus his total payment is either $\$p_2$ or $\$2p_2$ depending on whether or not he uses the contingent unit.

Since capacity-on-demand products are intended to be sold to firms, it is logical to assume that they are employed in a production process (where "production" may include the provision of a service, of course). The buyer, therefore, realizes a monetary *payoff* from his use of the capacity furnished by the seller. For convenience (and since both monetary denominations and units of capacity are arbitrary), we will assume that the payoff from the first unit is $1. Since the buyer's use of the contingent unit is not prescribed (by definition), however, we will make its payoff dependent on the “state of the world”, as described by the binary indicator σ. Thus if $\sigma = 0$, the payoff of the contingent unit is \$0 (in which case the buyer has no incentive to use it), whereas if $\sigma = 1$, the contingent unit has a strictly positive payoff $\$v > 0$ (so that the buyer may choose to use it, depending on the price demanded by the seller).

We assume that the buyer’s production process exhibits strictly decreasing returns to scale, so that the maximum payoff from the contingent unit is strictly less than that of the base unit. (This postulate is not insupportable, in that capacity-on-demand is supposed to provide “extra” capacity above the base, and also in that we can choose — say — the first unit of the base capacity as our base unit, and the last unit of contingent capacity actually used as our contingent unit.) Therefore, since the payoff from the base unit is \$1, we require that for contingent unit’s payoff, $\$v < 1$.

Finally, so as to make the potential of COD products to mitigate risks attractive to him, we assume that the buyer is *risk-averse*. Thus we accord him a utility function, $U(\bullet)$, that expresses his preference for monetary outcomes, and which is continuous, strictly increasing and strictly concave.[4] We let y stand for the buyer’s initial wealth.

An aside: One might object that the model’s assumption of a linear pricing scheme for the COD offering is simplistic. In fact, the assumption that base and contingent units are charged at the same rate is consonant with the “minimum unit commitments” described in (IBM 2005), for example, or (HP 2005). Furthermore, the “units” of base and contingent capacity need not be physically identical — the analysis here is unaltered so long as both units cost the same under COD, and the maximum payoff from the contingent unit is strictly less than that from the base unit.

[4] To keep the analysis tractable, the seller is assumed risk-neutral.

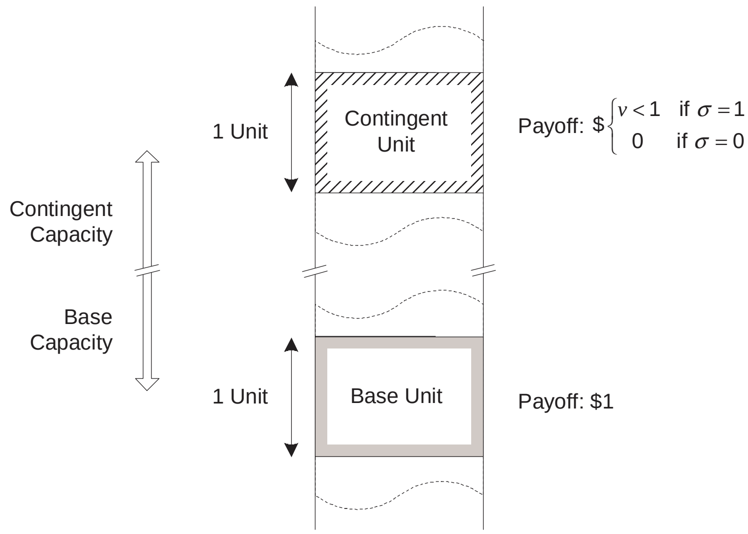

Figure 1: Model Illustration

The model is illustrated in Figure 1, and possible outcomes are summarized in Table 1.

		Offer Accepted		Declined
		$\sigma=0$	$\sigma=1$	$\sigma=0$ or 1
Payment	Bulk	p_1	p_1	0
	COD	p_2	$2p_2$	0
Buyer's payoff	Base unit	1	1	0
	Contingent unit	0	v	0
Buyer's net $ benefit	Bulk	$1-p_1$	$1+v-p_1$	0
	COD	$1-p_2$	$1+v-2p_2$	0
Buyer's utility	Bulk	$U(y+1-p_1)$	$U(y+1+v-p_1)$	$U(y)$
	COD	$U(y+1-p_2)$	$U(y+1+v-2p_2)$	$U(y)$

Table 1: Model Summary

3.1 Variations on the Simple Model

As suggested in the introduction, the effect of pricing decisions for COD products depends crucially on the information that the seller and buyer have about the buyer's future usage. In the simple model outlined above, this corresponds to information about the actual value of the indicator variable σ — the "state of nature" that determines the buyer's payoff from the contingent unit, upon which his usage depends. Thus we analyze three variations of the simple model that encapsulate three different assumptions about the information available to buyer and seller:

- In the *No Buyer Forecasts* case, both the buyer and seller share the same information about the value of σ, which takes the form of a probability distribution.
- With *Infallible Buyer Forecasts*, the seller still knows σ only up to a probability distribution, but the buyer knows the actual value of σ, which allows him to determine his future usage perfectly at the time the COD contract is offered him.
- Finally, *Fallible Buyer Forecasts* arrive in the shape of an additional state variable, θ, whose value is known only to the buyer, and which is (positively, but imperfectly) correlated with σ. The seller, it is assumed, knows nothing about θ. Both seller and buyer still share a probability distribution of the value of σ, but the buyer is able to use his additional information to refine this distribution.

3.2 Game-Theoretic Framework

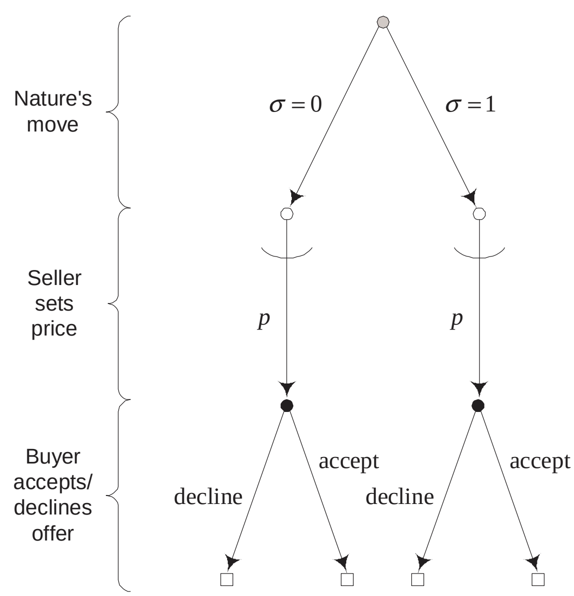

Figure 2: Basic Game Tree

Analysis of the model variations in the previous section is conducted within the framework of *Bayesian game theory*. Full introductions to the topic of Bayesian games (also known as *games with incomplete information*) are given in (Fudenberg and Tirole, 1991) and (Myerson, 1997); only the most fundamental elements of the framework are required in the analysis conducted here.

In its basic form, the simple model gives rise to a game tree resembling that depicted in Figure 2. The latter sets out a sequential game with three players: the buyer and seller discussed above, and *nature*, which chooses the value of the state variable σ . The game proceeds in three steps:

1. Nature chooses the value of σ , thus determining the buyer's payoff from the contingent unit.
2. The seller offers either bulk or COD with a given price. (Since the offer price is continuous, the seller chooses from an infinite number of alternatives at this stage, indicated by circular arcs in the diagram.)
3. The buyer accepts or declines the offer.

The moves selected at each stage of the game determine the final payoffs of the buyer and seller, the risk-averse buyer's payoff taking the form of the utility of his wealth, and the seller's comprising her revenue.

In analyzing the game, we seek a *Bayesian Nash equilibrium*, wherein both buyer and seller look to maximize their expected payoff, assuming that the other player does likewise. Discussions of the descriptive, predictive and normative implications of Nash equilibria (and thus of Bayesian Nash equilibria) may be found in (Mas-Colell et al., 1995, pp. 248 – 249) or (Fudenberg and Tirole, 1991, pp. 18 – 29).

For reference during the discussion in the next sections, Table 2 summarizes the parameters of the basic model and its variations, along with their assumed ranges.

Quantity	Interpretation	Range/Assumptions
σ	"State of the world" indicator	$\sigma \in \{0,1\}$
$1	Payoff from base unit in either state	
v	Payoff from contingent unit when $\sigma = 1$	$0 < v < 1$
y	Buyer's initial wealth	$y \gg 0$
$U(\bullet)$	Buyer's utility function	Continuous, increasing, concave
R_A	Buyer's risk aversion index	$R_A > 0$
ϕ	Probability that $\sigma = 1$	$0 < \phi < 1$
p_1	Bulk offer price	$p_1 > 0$
p_2	COD offer price	$p_2 > 0$
p	Offer price (bulk or COD unspecified)	$p > 0$
$e_1(p_1)$	Seller's revenue as a function of bulk price	
$e_2(p_2)$	Seller's revenue as a function of COD price	
p_1^*	Seller's optimum bulk offer price	$p_1^* > 0$
p_2^*	Seller's optimum COD offer price	$p_2^* > 0$
e_1^*	Seller's maximum revenue with bulk pricing	$e_1^* \geq 0$
e_2^*	Seller's maximum revenue with COD pricing	$e_2^* \geq 0$
θ	Buyer's private signal variable	$\theta \in \{0,1\}$
ρ	Accuracy of buyer's private signal	$\frac{1}{2} < \rho < 1$
$C_{10}(\bullet)$, etc.	Predicates involving offer prices (Section 6)	
$I(\bullet)$	Characteristic/indicator function	$I(\mathbf{true}) \triangleq 1,\ I(\mathbf{false}) \triangleq 0$

Table 2: Summary of Notation

(Parentheses, brackets, braces and the like are interchanged freely in the interests of clarity.)

4 Variation I: No Buyer Forecasts

4.1 Overview

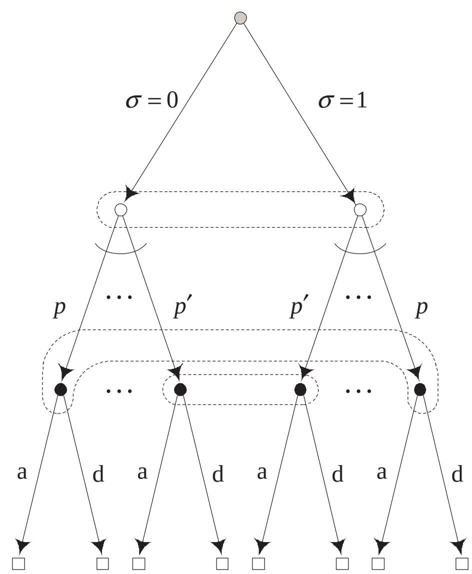

Figure 3: Game Tree — Fallible Buyer Forecasts

In this variation, neither buyer nor seller has definitive information about the value of σ, as reflected by the regions drawn around sets of nodes in the figure.[5] When the buyer moves, therefore, she is unable to distinguish between the two nodes resulting from nature's moves. Correspondingly, the buyer is able to distinguish between nodes only on the basis of the offered price (only two sample prices, p and p' are shown in the figure), not according to the state of nature they reflect.

Both buyer and seller, however, do share a probability distribution (termed the *prior distribution*) over σ. Since σ is a binary variable, this distribution is completely described by a single model parameter:

$$\phi \triangleq \Pr(\sigma = 1), \quad \text{so that } \Pr(\sigma = 0) = 1 - \phi \qquad (1)$$

We now examine how each player might maximize his/her payoff under bulk and COD pricing.

[5] In game-theoretic parlance, they delimit the *information sets* of the game.

4.2 Bulk Case

Recalling Table 1, when offered a bulk price p_1, the buyer with initial wealth y faces the following outcomes, contingent on his choice to accept or decline the offer and the state of nature as indicated by σ:

Accept/decline offer	State	Buyer's Utility	Probability
accept	$\sigma = 0$	$U(y+1-p_1)$	$1-\phi$
	$\sigma = 1$	$U(y+1+v-p_1)$	ϕ
decline	$\sigma = 0,1$	$U(y)$	1

Thus if the buyer is to maximize his expected utility (a condition for the Bayesian Nash equilibrium we seek), his best response to the offer is as follows:[6]

$$\begin{cases} \textbf{accept} & \text{if } (1-\phi)U(y+1-p_1)+\phi\,U(y+1+v-p_1) \geq U(y) \\ \textbf{decline} & \text{otherwise} \end{cases} \tag{2}$$

Since the seller always earns $\$p_1$ in the event the buyer accepts in (2), her revenues (as a function of the price offered) are:

$$e_1(p_1) = \begin{cases} p_1 \text{ if } (1-\phi)U(y+1-p_1)+\phi\,U(y+1+v-p_1) \geq U(y) \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

This characterization may be further refined using a standard result from utility theory. Specifically, it can be shown (see Nicholson 2001, pp. 232-243, for example) that for random variable *w* with mean μ and variance τ^2, expected utility may be approximated:

$$\mathrm{E}\left[U(w)\right] \simeq U(\mu - \tfrac{1}{2}R_A\,\tau^2), \quad \text{where } R_A = -\frac{U''(\mu)}{U'(\mu)} \tag{4}$$

Here, the *Arrow-Pratt coefficient of absolute risk aversion*, R_A (a measure of the curvature of the buyer's utility function at μ), is an index of the buyer's antipathy toward uncertain outcomes; it is strictly positive for the risk-averse buyer is assumed in the model.[7] The expression $\frac{1}{2}R_A\sigma^2$ constitutes a *risk premium*, and is determined by the discount that the (risk-averse) buyer associates with the uncertainty in the outcome of his selection in (2).

Since buyer's payoffs in (2) may be considered as a random variable with mean $y+1+\phi v-p_1$ and variance $v^2\phi(1-\phi)$, (4) implies that the following equivalences hold up to approximation:

---

[6] For the sake of convenience, we resolve indifference in favor of purchase.

[7] In the notation used here, the general dependence of R_A on μ is elided.

$$
\begin{aligned}
& (1-\phi)U(y+1-p_1)+\phi U(y+1+v-p_1) \ \geq \ U(y) \\
\Leftrightarrow \ & U(y+1+\phi v-p_1-\tfrac{1}{2}R_A v^2\phi(1-\phi)) \ \geq \ U(y) \\
\Leftrightarrow \ & y+1+\phi v-p_1-\tfrac{1}{2}R_A v^2\phi(1-\phi) \ \geq \ y \\
\Leftrightarrow \ & p_1 \ \leq \ 1+\phi v-\tfrac{1}{2}R_A v^2\phi(1-\phi)
\end{aligned}
\tag{5}
$$

Given (3), the seller attains maximum earnings when the inequalities in (5) are binding, so her optimum price offer is:

$$p_1^* = 1+\phi v-\tfrac{1}{2}R_A v^2\phi(1-\phi) \tag{6}$$

Thus the seller's maximum earnings are:

$$e_1^* = e_1(p_1^*) = p_1^* = 1+\phi v-\tfrac{1}{2}R_A v^2\phi(1-\phi) \tag{7}$$

4.3 COD Case

Calculations proceed analogously to those in the bulk case, so that the buyer's strategy is described:

$$
\begin{cases}
\textbf{accept} & \text{if } (1-\phi)U(y+1-p_2)+\phi U(y+1+v-2p_2) \geq U(y) \\
\textbf{decline} & \text{otherwise}
\end{cases}
\tag{8}
$$

In this case, when the buyer accepts, the seller is paid p_2 (for the base unit alone) with probability $1-\phi$ and $2p_2$ (for the base and contingent units) with probability ϕ:

$$
\begin{aligned}
e_2(p_2) &= \begin{cases}
(1-\phi)p_2+\phi\times 2p_2 & \text{if } (1-\phi)U(y+1-p_2)+\phi U(y+1+v-2p_2) \geq U(y) \\
0 & \text{otherwise}
\end{cases} \\
&= \begin{cases}
(1+\phi)p_2 & \text{if } (1-\phi)U(y+1-p_2)+\phi U(y+1+v-2p_2) \geq U(y) \\
0 & \text{otherwise}
\end{cases}
\end{aligned}
\tag{9}
$$

The buyer's payoffs have mean $y+1+\phi v-(1+\phi)p_2$ and variance $(v-p_2)^2\phi(1-\phi)$, so using the risk aversion coefficient as in the bulk case, approximately:

$$
e_2(p_2) = \begin{cases}
(1+\phi)p_2 & \text{if } (1+\phi)p_2 \leq 1+\phi v-\tfrac{1}{2}R_A(v-p_2)^2\phi(1-\phi) \\
0 & \text{otherwise}
\end{cases}
\tag{10}
$$

Now formally, by deriving the appropriate Kuhn-Tucker conditions, or informally by graphical inspection, we can show that the seller's optimum price must satisfy:

$$(1+\phi)p_2^* = 1+\phi v-\tfrac{1}{2}R_A(v-p_2^*)^2\phi(1-\phi) \tag{11}$$

And using (9), the seller's maximum expected earnings are derived:

$$\begin{aligned} e_2^* &= e_2(p_2^*) \\ &= (1+\phi)\,p_2^* \\ &= 1+\phi v - \tfrac{1}{2} R_A \left(v - p_2^*\right)^2 \phi(1-\phi) \\ &= 1+\phi v - \tfrac{1}{2} R_A \left(v - \frac{e_2^*}{1+\phi}\right)^2 \phi(1-\phi) \end{aligned} \tag{12}$$

This quadratic equation has two roots, the larger of which is given by the rather formidable expression:

$$e_2^* = \frac{1+\phi}{R_A\,\phi(1-\phi)}\left(\phi\left(R_A v(1-\phi)-1\right)-1+\sqrt{\phi\left(\phi+2+2R_A(1-v)(1-\phi)\right)+1}\right) \tag{13}$$

For the parameter ranges assumed for the model, it is evident that the term in the radical is positive, so we are assured that e_2^* is real. Furthermore, it is positive, i.e.:

$$\phi\left(R_A v(1-\phi)-1\right)-1+\sqrt{\phi\left(\phi+2+2R_A(1-v)(1-\phi)\right)+1} > 0 \tag{14}$$

For if we were to assume the contrary, then:

$$\sqrt{\phi\left(\phi+2+2R_A(1-v)(1-\phi)\right)+1} \le 1-\phi\left(R_A v(1-\phi)-1\right) \tag{15}$$

Squaring both sides and simplifying, this implies that:

$$1+\phi v - \tfrac{1}{2} R_A v^2 \phi(1-\phi) \le 0 \tag{16}$$

In addition, since the left hand side of (15) is positive:

$$\begin{aligned} & 1-\phi\left(R_A v(1-\phi)-1\right) \ge 0 \\ \Rightarrow\quad & 1+\phi - R_A v\phi(1-\phi) \ge 0 \\ \Rightarrow\quad & \tfrac{1}{2} R_A v^2 \phi(1-\phi) \le \tfrac{1}{2} v(1+\phi) \end{aligned} \tag{17}$$

Substituting (17) into (16):

$$\begin{aligned} & 1+\phi v - \tfrac{1}{2} v(1+\phi) \le 1+\phi v - \tfrac{1}{2} R_A v^2 \phi(1-\phi) \le 0 \\ \Rightarrow\quad & 1-\tfrac{1}{2} v(1-\phi) \le 0 \end{aligned} \tag{18}$$

But the latter is a contradiction, given the ranges of the model parameters.

4.4 Commentary

Comparing (7) and (12), given the range restrictions on the model parameters:

$$
\begin{aligned}
e_2^* > e_1^* &\Rightarrow \left(v - \frac{e_2^*}{1+\phi}\right)^2 < v^2 \\
&\Rightarrow \left| v - \frac{e_2^*}{1+\phi} \right| < v \\
&\Rightarrow 0 < \frac{e_2^*}{1+\phi} < 2v \\
&\Rightarrow e_2^* < 2v\,(1+\phi) && \text{since } e_2^* \text{ and } \phi \text{ are positive} \\
&\Rightarrow e_1^* < e_2^* < 2v\,(1+\phi) && \text{by assumption} \\
&\Rightarrow 1 + \phi v - \tfrac{1}{2} R_A v^2 \phi (1-\phi) < 2v\,(1+\phi) && \text{substituting for } e_1^* \\
&\Rightarrow 1 + \phi v - 2v\,(1+\phi) < \tfrac{1}{2} R_A v^2 \phi (1-\phi) \\
&\Rightarrow \tfrac{1}{2} R_A > \frac{1-(2+\phi)v}{v^2 \phi (1-\phi)}
\end{aligned}
\tag{19}
$$

A similar argument beginning with the assumption $e_2^* \le e_1^*$ establishes the equivalence:

$$
e_2^* > e_1^* \;\Leftrightarrow\; \tfrac{1}{2} R_A > \frac{1-(2+\phi)v}{v^2 \phi (1-\phi)} \tag{20}
$$

Note that the final inequality is satisfied if — as assumed — R_A is strictly positive and $v \ge 1/(2+\phi)$, so that with model parameters within their assumed ranges:

$$
e_2^* > e_1^* \Leftrightarrow\; v \ge \frac{1}{2+\phi} \;\text{ or }\; \tfrac{1}{2} R_A > \frac{1-(2+\phi)v}{v^2 \phi (1-\phi)} \tag{21}
$$

Informally, this may be interpreted as increasingly favoring capacity-on-demand over bulk as a selling strategy to the extent that the following hold:

1. The return to the buyer from use of the contingent capacity is high.
2. The probability that the contingent capacity will be used is high.
3. The buyer is relatively more risk averse.

Note, however, that neither bulk nor capacity-on-demand is a superior approach in all circumstances. This might seem at odds with intuition; after all, a buyer purchasing capacity-on-demand pays additional fees only if he uses (and presumably profits from) the additional capacity, reducing the variability of his net dollar benefit — an agreeable prospect for the risk-averse buyer. However, the derivation in (19) demonstrates that in order to effect such a reduction in variability, the price charged in the capacity-on-demand scheme must be such that overall earnings cannot exceed $2v(1+\phi)$, which may be less than the seller can earn offering a bulk purchase under the same circumstances.

5 Variation II: Infallible Buyer Forecasts

5.1 Overview

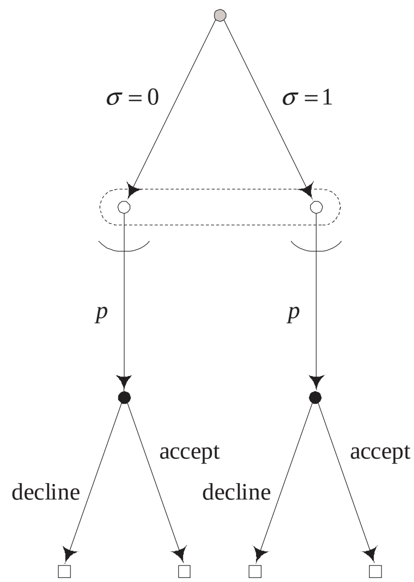

Figure 4: Game Tree — Infallible Buyer Forecasts

This variation is depicted in Figure 4. The seller is still unable to distinguish the states in which her moves are made, but the buyer now has precise information about the value of σ.

5.2 Bulk Case

Since the buyer is now able to observe σ, he can predict his payoffs with certainty. Accordingly, the buyer's best response is given by the following composite piecewise function:

$$\begin{cases} \left.\begin{cases} \textbf{accept} & \text{if } U(y+1-p_1) \geq U(y) \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 0 \\ \left.\begin{cases} \textbf{accept} & \text{if } U(y+1+v-p_1) \geq U(y) \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 1 \end{cases} \tag{22}$$

Since the buyer's utility function, $U(\bullet)$, is assumed to be continuous and strictly increasing, this expression may be simplified:

$$\begin{cases} \left.\begin{cases} \textbf{accept} & \text{if } p_1 \le 1 \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 0 \\ \left.\begin{cases} \textbf{accept} & \text{if } p_1 \le 1 + v \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 1 \end{cases} \tag{23}$$

Given such a strategy on the part of the buyer (with $I(\bullet)$ denoting the indicator function), the seller can expect revenue:

$$e_1(p_1) = (1-\phi)\, p_1\, I(p_1 \le 1) + \phi\, p_1\, I(p_1 \le 1+v) \tag{24}$$

We have assumed that $0 < v < 1$, so this expression is has local maxima:

$$\begin{aligned} p_1 &= 1 && \Rightarrow \quad e_1(p_1) = 1 \\ p_1 &= 1 + v && \Rightarrow \quad e_1(p_1) = (1+v)\phi \end{aligned} \tag{25}$$

Finding the pricing scheme's global maximum entails comparing these two local maxima; the (trivially) derived conditions on the model parameters are as follows:

$$\begin{aligned} p_1^* &= \begin{cases} 1+v & \text{if } v > \dfrac{1}{\phi} - 1 \\ 1 & \text{otherwise} \end{cases} \\ e_1^* &= \begin{cases} (1+v)\phi & \text{if } v > \dfrac{1}{\phi} - 1 \\ 1 & \text{otherwise} \end{cases} \end{aligned} \tag{26}$$

5.3 COD Case

As before, the buyer's best response is a piecewise function:

$$\begin{cases} \left.\begin{cases} \textbf{accept} & \text{if } U(y+1-p_2) \ge U(y) \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 0 \\ \left.\begin{cases} \textbf{accept} & \text{if } U(y+1+v-2p_2) \ge U(y) \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 1 \end{cases} \tag{27}$$

Which may be simplified:

$$\begin{cases} \left.\begin{cases} \textbf{accept} & \text{if } p_2 \le 1 \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 0 \\ \left.\begin{cases} \textbf{accept} & \text{if } 2p_2 \le 1 + v \\ \textbf{decline} & \text{otherwise} \end{cases}\right\} & \text{if } \sigma = 1 \end{cases} \tag{28}$$

Thus the seller can expect revenue:

$$e_2(p_2) = (1-\phi)\, p_2\, I(p_2 \le 1) + \phi\, 2p_2\, I(2p_2 \le 1+v) \tag{29}$$

This has the following two local maxima:

$$\begin{aligned} p_2 &= 1 && \Rightarrow \quad e_2(p_2) = 1-\phi \\ p_2 &= \frac{1+v}{2} && \Rightarrow \quad e_2(p_2) = (1+\phi)\frac{1+v}{2} \end{aligned} \tag{30}$$

Globally:

$$\begin{aligned} p_2^* &= \begin{cases} \dfrac{1+v}{2} & \text{if } \phi > \dfrac{1}{3} \text{ or } v > \dfrac{1-3\phi}{1+\phi} \\ 1 & \text{otherwise} \end{cases} \\ e_2^* &= \begin{cases} (1+\phi)\dfrac{1+v}{2} & \text{if } \phi > \dfrac{1}{3} \text{ or } v > \dfrac{1-3\phi}{1+\phi} \\ 1-\phi & \text{otherwise} \end{cases} \end{aligned} \tag{31}$$

5.4 Commentary

As in the previous section, it is possible to determine which pricing scheme (bulk vs. COD) produces greater expected earnings for the seller. Like before, neither bulk nor COD is superior under all circumstances. To demonstrate this, begin by expanding the definition of e_1^* to yield the following:

$$e_2^* > e_1^* \iff e_2^* > 1 > (1+v)\,\phi \;\vee\; e_2^* > (1+v)\,\phi > 1 \tag{32}$$

Since the alternate value of e_2^* is $1-\phi$, and since we have assumed that $\phi > 1$, it must be that $1-\phi < 1$, so:

$$e_2^* > e_1^* \iff e_2^* = (1+\phi)\frac{1+v}{2} \;\wedge\; \left[e_2^* > 1 > (1+v)\,\phi \;\vee\; e_2^* > (1+v)\,\phi > 1\right] \tag{33}$$

With model parameters within their assumed ranges, it is straightforward to show that, $(1+\phi)\frac{1+v}{2} > (1+v)\,\phi$, so that the above reduces to:

$$e_2^* > e_1^* \iff e_2^* = (1+\phi)\frac{1+v}{2} \;\wedge\; e_2^* > 1 \tag{34}$$

Next, we have:

$$\begin{aligned} (1+\phi)\frac{1+v}{2} > 1 &\iff (1+\phi)(1+v) > 2 \\ &\iff (1+\phi)\,v > 2-(1+\phi) \\ &\iff (1+\phi)\,v > 1-\phi \\ &\iff v > \frac{1-\phi}{1+\phi} \end{aligned} \tag{35}$$

Combining (34) and (35) establishes the conditions under which the COD scheme is superior:

$$e_2^* > e_1^* \iff v > \frac{1-\phi}{1+\phi} \iff \phi > \frac{1-v}{1+v} \tag{36}$$

Stepping back and comparing the pricing schemes offered in this section with those of the previous one, it's obvious that the players' potential winnings are altered significantly by the buyer's possession of a private forecast of his usage. First, note that since the buyer knows the outcome of his purchase decision ahead of time, the utility of any outcome is not impaired by his risk aversion.

The seller, by contrast, suffers from her ignorance of the buyer's forecast. For consider how the seller might price both bulk and capacity-on-demand offerings were she to share the buyer's forecast of his actual use (as embodied in the value of the state indicator σ) — a situation termed the *first best solution* in the literature. It is a relatively minor exercise to prove that the seller's optimal first best strategy is to offer a bulk price of $1+v$ dollars or a COD price of $(1+v)/2$ dollars when $\sigma=1$ and a price of $1 in either scheme when $\sigma=0$. In these circumstances, expected revenues are $1+\phi v$ dollars with either pricing scheme, which is demonstrably greater than e_1^* and e_2^*.

So with knowledge of the usage forecast, the seller can earn revenues strictly exceeding those of a seller who is ignorant of the buyer's projected usage. This difference in revenues — the surplus that the buyer is able to obtain from his private forecast — is termed the buyer's (expected) *information rent*; it is strictly positive, regardless whether the bulk or COD pricing scheme is adopted.

However, under certain circumstances, a seller offering COD may suffer particularly from this information asymmetry. Observe that from (35), (and from the fact that $1-\phi<1$, as mentioned above) it is perfectly possible to have $e_2^*<1$ — in fact the minimum value of e_2^* approaches $\$\,3/4$ when $\phi=1/2$ as $v\to 0$. However, recall that we know (as does the seller) that the buyer will certainly have use for at least one unit of capacity in either state of the world, and that therefore he's bound to achieve a payoff of at least $1. In other words, *for a range of model parameters, the COD seller's expected earnings are strictly less than $1 — despite the fact that she knows that the buyer is <u>surely</u> able to extract a payoff of at least $1.* (By contrast, for the bulk scheme, it is always the case that $e_1^* \ge 1$.)

To see informally how this apparent paradox arises, consider the seller's price-setting decision in the COD case, as set out in (31):

– She could select a price of $1, in which case the buyer will only purchase if he knows his usage will be less than two units (for in the case that he does need 2 units, he'd be compelled to make a payment of 2 dollars, which is more than his payoff of $1+v$ dollars). The probability that the buyer needs only one unit is $1-\phi$. Therefore in this instance, the seller can expect earnings of $1-\phi<1$ dollars.

– Or she could set the price at $\frac{1}{2}(1+v)<1$ dollars. At this price, the buyer will purchase regardless of whether he needs 1 or 2 units. However, if he needs only one unit, he pays an amount strictly less than his payoff (of $1). For the parameter values specified in (36), this loss of surplus to the customer requiring only 1 unit is sufficient to depress expected earnings below 1 dollar.

In Section 7, we discuss approaches to pricing COD products that avoid this anomaly. Meanwhile, the next section looks at a model that combines elements of both of those we've examined heretofore.

6 Variation III: Fallible Buyer Forecasts

6.1 Overview

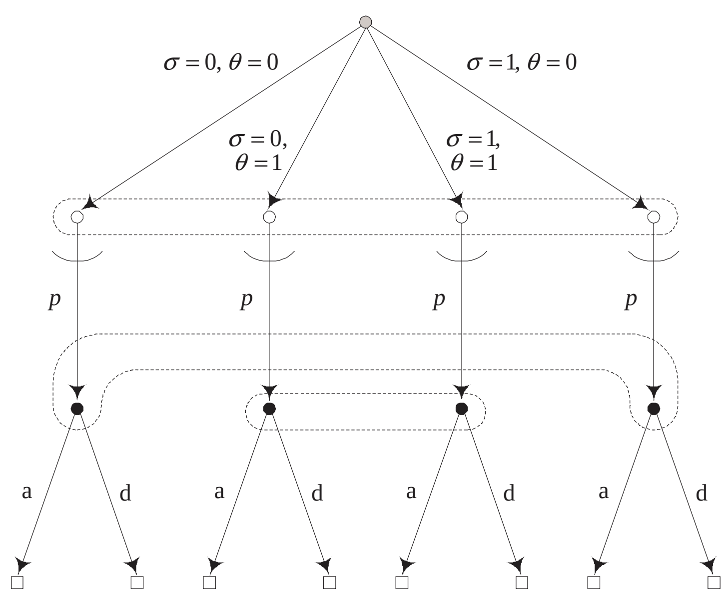

Figure 5: Game Tree — Fallible Buyer Forecasts

In this game, the seller is still unsure about the value of σ, but the buyer can observe a private "signal" variable, θ, that is positively correlated with σ. In the diagram of Figure 5, such an arrangement is represented by enumerating the combined values of σ and θ; while the seller cannot distinguish between any of these combinations, the buyer can separate them, but only based on the value of θ.

As before, both seller and buyer possess common knowledge about σ and θ in the form of a (joint) prior probability distribution. Since the seller knows nothing about the buyer's signal, it is logical to assume that (marginally) in the common prior $\Pr(\theta = 1) = \Pr(\theta = 0) = 1/2$. To complete the prior's joint distribution, we need to specify the conditional distribution of σ given θ, $\Pr(\sigma \mid \theta)$, which determines $\Pr(\theta, \sigma) = \Pr(\sigma \mid \theta)\Pr(\theta)$. For simplicity, this conditional distribution is specified by a single model parameter, ρ:

$$\begin{aligned} &\Pr(\sigma = 1 \mid \theta = 1) = \rho, \\ &\Pr(\sigma = 1 \mid \theta = 0) = 1 - \rho \end{aligned} \tag{37}$$

Amongst other things, this leaves the seller wholly uninformed about the value of σ (a rather more restrictive assumption than that made in Section 4), since its marginal distribution in the prior is:

$$\begin{aligned} \Pr(\sigma = 1) &= \Pr(\sigma = 1 \mid \theta = 1)\Pr(\theta = 1) + \Pr(\sigma = 1 \mid \theta = 0)\Pr(\theta = 0) \\ &= \tfrac{1}{2}\rho + \tfrac{1}{2}(1 - \rho) \\ &= \tfrac{1}{2} \\ \Pr(\sigma = 0) &= \Pr(\sigma = 0 \mid \theta = 1)\Pr(\theta = 1) + \Pr(\sigma = 0 \mid \theta = 0)\Pr(\theta = 0) \\ &= \tfrac{1}{2}(1 - \rho) + \tfrac{1}{2}\rho \\ &= \tfrac{1}{2} \end{aligned} \tag{38}$$

To make the buyer's forecast informative but less than perfect, we require that $1/2 < \rho < 1$.[8]

6.1.1 Bulk Case

For the buyer, the expected payoffs of his choices in this game can be summarized as follows:

Value of θ	Buyer's move	Buyer's expected utility
0	**decline**	$U(y)$
	accept	$\rho U(y + 1 - p_1) + (1 - \rho)U(y + 1 + v - p_1)$
1	**decline**	$U(y)$
	accept	$(1 - \rho)U(y + 1 - p_1) + \rho U(y + 1 + v - p_1)$

[8] Providing that $\rho \neq 1/2$, this can always be assured by inverting the value of θ as necessary.

Abbreviate:

$$\begin{aligned} C_{10}(p) &\triangleq \rho U(y+1-p)+(1-\rho)U(y+1+v-p) \geq U(y) \\ C_{11}(p) &\triangleq (1-\rho)U(y+1-p_1)+\rho U(y+1+v-p_1) \geq U(y) \end{aligned} \tag{39}$$

These predicates can also be approximated using the buyer's risk aversion index:

$$\begin{aligned} C_{10}(p) &\Leftrightarrow p \leq 1+(1-\rho)v-\tfrac{1}{2}R_A v^2 \rho(1-\rho) \\ C_{11}(p) &\Leftrightarrow p \leq 1+\rho v-\tfrac{1}{2}R_A v^2 \rho(1-\rho) \end{aligned} \tag{40}$$

Thus the buyer's best response to an offer price p_1 is:

$$\begin{cases} \textbf{accept} & \text{if } [\theta=0 \text{ and } C_{10}(p_1)] \text{ or } [\theta=1 \text{ and } C_{11}(p_1)] \\ \textbf{decline} & \text{otherwise} \end{cases} \tag{41}$$

Offering price p_1, the seller receives:

$$\begin{cases} p_1 & \text{if buyer } \textbf{accept}\text{s} \\ 0 & \text{otherwise} \end{cases} \tag{42}$$

Substituting from (41), (42) becomes:

$$\begin{cases} p_1 & \text{if } [\theta=0 \text{ and } C_{10}(p_1)] \text{ or } [\theta=1 \text{ and } C_{11}(p_1)] \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

Thus the expected revenue resulting from an offer price p_1 is:

$$\begin{aligned} e_1(p_1) &= p_1\left[\Pr(\theta=0)\, I\{C_{10}(p_1)\}+\Pr(\theta=1)\, I\{C_{11}(p_1)\}\right] \\ &= \tfrac{1}{2} p_1\left[I\{C_{10}(p_1)\}+I\{C_{11}(p_1)\}\right] \end{aligned} \tag{44}$$

From the approximations in (40), we derive two local maxima for this expression:

$$\begin{aligned} p_{10} &= 1+(1-\rho)v-\tfrac{1}{2}R_A v^2\rho(1-\rho), & e_1(p_{10}) &= \tfrac{1}{2}p_{10}\left[I\{C_{10}(p_{10})\}+I\{C_{11}(p_{10})\}\right] \\ p_{11} &= 1+\rho v-\tfrac{1}{2}R_A v^2\rho(1-\rho), & e_1(p_{11}) &= \tfrac{1}{2}p_{11}\left[I\{C_{10}(p_{11})\}+I\{C_{11}(p_{11})\}\right] \end{aligned} \tag{45}$$

To expand out the indicator functions in the expressions above, note first that by definition, $C_{10}(p_{10})$ and $C_{11}(p_{11})$. Furthermore, since $p_{11} = p_{10} + (2\rho-1)\,v$, with the assumptions we have made for the ranges of the model parameters, it is also the case that $C_{11}(p_{10})$ and $\neg C_{10}(p_{11})$.

Thus:

$$\begin{aligned} e_1(p_{10}) &= \tfrac{1}{2}p_{10}[1+1] = p_{10} = 1+(1-\rho)v-\tfrac{1}{2}R_A v^2\rho(1-\rho) \\ e_1(p_{11}) &= \tfrac{1}{2}p_{11}[0+1] = \tfrac{1}{2}p_{11} = \tfrac{1}{2}\left[1+\rho v-\tfrac{1}{2}R_A v^2\rho(1-\rho)\right] \end{aligned} \tag{46}$$

Comparing these two expressions leads straightforwardly to the global maximum:

$$p_1^* = \begin{cases} 1+(1-\rho)v - \frac{1}{2}R_A v^2 \rho(1-\rho) & \text{if } \frac{1}{2}R_A \le \dfrac{1+(2-3\rho)v}{v^2\rho(1-\rho)} \\ 1+\rho v - \frac{1}{2}R_A v^2 \rho(1-\rho) & \text{otherwise} \end{cases}$$

$$e_1^* = \begin{cases} 1+(1-\rho)v - \frac{1}{2}R_A v^2 \rho(1-\rho) & \text{if } \frac{1}{2}R_A \le \dfrac{1+(2-3\rho)v}{v^2\rho(1-\rho)} \\ \frac{1}{2}\left[1+\rho v - \frac{1}{2}R_A v^2 \rho(1-\rho)\right] & \text{otherwise} \end{cases} \tag{47}$$

6.1.2 COD Case

Value of θ	Buyer's move	Buyer's expected utility
0	**decline**	$U(y)$
	accept	$\rho U(y+1-p_2)+(1-\rho)U(y+1+v-2p_2)$
1	**decline**	$U(y)$
	accept	$(1-\rho)U(y+1-p_2)+\rho U(y+1+v-2p_2)$

In this case, we abbreviate:

$$\begin{aligned} C_{20}(p) &\triangleq \rho\, U(y+1-p)+(1-\rho)U(y+1+v-2p) \ge U(y), \\ C_{21}(p) &\triangleq (1-\rho)U(y+1-p)+\rho\, U(y+1+v-2p) \ge U(y) \end{aligned} \tag{48}$$

Approximately:

$$\begin{aligned} C_{20}(p) &\Leftrightarrow [1+(1-\rho)]\,p \le 1+(1-\rho)v - \tfrac{1}{2}R_A(v-p)^2\rho(1-\rho) \\ C_{21}(p) &\Leftrightarrow (1+\rho)\,p \le 1+\rho v - \tfrac{1}{2}R_A(v-p)^2\rho(1-\rho) \end{aligned} \tag{49}$$

The buyer's best response to an offer price p_2 is:

$$\begin{cases} \textbf{accept} & \text{if } [\theta=0 \text{ and } C_{20}(p_2)] \text{ or } [\theta=1 \text{ and } C_{21}(p_2)] \\ \textbf{decline} & \text{otherwise} \end{cases} \tag{50}$$

When the buyer accepts a COD product, the seller's revenues are dependent on the buyer's actual usage, as determined by the value of σ. Thus, she receives:

$$\begin{cases} p_2 & \text{if buyer } \textbf{accept}\text{s and } \sigma=0 \\ 2p_2 & \text{if buyer } \textbf{accept}\text{s and } \sigma=1 \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

Substituting from (50) and rearranging, (51) becomes:

$$\begin{cases} p_2 & \text{if } [\theta=0 \text{ and } \sigma=0 \text{ and } C_{20}(p_2)] \text{ or } [\theta=1 \text{ and } \sigma=0 \text{ and } C_{21}(p_2)] \\ 2p_2 & \text{if } [\theta=0 \text{ and } \sigma=1 \text{ and } C_{20}(p_2)] \text{ or } [\theta=1 \text{ and } \sigma=1 \text{ and } C_{21}(p_2)] \\ 0 & \text{otherwise} \end{cases} \tag{52}$$

Again, the expected revenue resulting from an offer price p_2 is:

$$\begin{aligned} e_2(p_2) &= [p_2 \times \Pr(\theta=0,\sigma=0) + 2p_2 \times \Pr(\theta=0,\sigma=1)]\, I\{C_{20}(p_2)\} + \\ &\qquad [p_2 \times \Pr(\theta=1,\sigma=0) + 2p_2 \times \Pr(\theta=1,\sigma=1)]\, I\{C_{21}(p_2)\} \\ &= \tfrac{1}{2}\big([\rho\, p_2 + 2(1-\rho)p_2]\, I\{C_{20}(p_2)\} + [(1-\rho)p_2 + 2\rho\, p_2]\, I\{C_{21}(p_2)\}\big) \\ &= \tfrac{1}{2} p_2 \big[(2-\rho)\, I\{C_{20}(p_2)\} + (1+\rho)\, I\{C_{21}(p_2)\}\big] \end{aligned} \tag{53}$$

This time there are two local maxima, each satisfying the following:

$$\begin{aligned} &[1+(1-\rho)]\, p_{20} = 1 + (1-\rho)v - \tfrac{1}{2} R_A (v - p_{20})^2 \rho(1-\rho), \\ &e_2(p_{10}) = \tfrac{1}{2} p_{20}\big[(2-\rho)\, I\{C_{20}(p_{20})\} + (1+\rho)\, I\{C_{21}(p_{20})\}\big] \\ \\ &(1+\rho)p_{21} = 1 + \rho v - \tfrac{1}{2} R_A (v - p_{21})^2 \rho(1-\rho), \\ &e_2(p_{21}) = \tfrac{1}{2} p_2\big[(2-\rho)\, I\{C_{20}(p_{21})\} + (1+\rho)\, I\{C_{21}(p_{21})\}\big] \end{aligned} \tag{54}$$

Taking the larger of the two roots in each case yields:

$$\begin{aligned} p_{20} &= \frac{1}{R_A\, \rho(1-\rho)}\Big[\rho\big(R_A v(1-\rho)+1\big) - 2 + \sqrt{\rho\big(2R_A(1-v)(1-\rho)+\rho-4\big)+4}\,\Big] \\ p_{21} &= \frac{1}{R_A\, \rho(1-\rho)}\Big[\rho\big(R_A v(1-\rho)-1\big) - 1 + \sqrt{\rho\big(2R_A(1-v)(1-\rho)+\rho+2\big)+1}\,\Big] \end{aligned} \tag{55}$$

Arguments similar to that in Section 4.3 demonstrate that both prices are real and positive. As before, $C_{20}(p_{20})$ and $C_{21}(p_{21})$ by definition. Next, substituting (55) into (49) and simplifying produces:

$$\begin{aligned} C_{20}(p_{21}) &\Leftrightarrow \frac{2\rho-1}{R_A\, \rho(1-\rho)}\Big[1+\rho-\sqrt{\rho\big(2R_A(1-v)(1-\rho)+\rho+2\big)+1}\,\Big] \le 0 \\ &\Leftrightarrow (1+\rho)^2 \le \rho\big(2R_A(1-v)(1-\rho)+\rho+2\big)+1 \\ &\Leftrightarrow 2R_A\,\rho\,[\rho(1-v)-(1-v)] \le 0 \\ &\Leftrightarrow \rho(1-v) \le 1-v \\ &\Leftrightarrow \rho \le 1 \\ &\Leftrightarrow \textbf{true} \qquad \text{Since it is assumed that } \tfrac{1}{2} < \rho < 1 \end{aligned} \tag{56}$$

Similarly, we can show that $C_{21}(p_{20})$ is false.

Thus the values of maxima in (54) may be simplified:

$$e_2(p_{20}) = \tfrac{1}{2}(2-\rho)\ p_{20}$$
$$e_2(p_{21}) = \tfrac{3}{2}\ p_{21} \tag{57}$$

Furthermore a tedious but straightforward expansion of both of the above expressions along the same lines as (56) establishes that for the model parameter ranges assumed, it is always the case that $e_2(p_{21}) > e_2(p_{20})$.

Therefore:

$$e_2^* = e_2(p_{21}) = \tfrac{3}{2}\ p_{21} \tag{58}$$

6.2 Commentary

Comparing (47) and the expansion of (58), more algebra proves that:

$$e_2^* > e_1^* \Leftrightarrow\ v \geq \frac{1}{2+\rho}\ \text{ or }\ \tfrac{1}{2}R_A > \frac{1-(2+\rho)v}{v^2\rho(1-\rho)} \tag{59}$$

Comparing this with (21), we see that these are precisely the conditions associated with Variation I (the "no forecasting" case), but with the buyer's forecasting accuracy, ρ, substituted for the common prior probability of contingent usage, ϕ, from Variation I.[9] In other words, the seller's selection of pricing scheme depends on the risk that the buyer associates with his forecast. In this case too, we can deduce conditions favoring the seller's offering COD rather than bulk:

1. The return to the buyer from use of the contingent capacity is high.
2. The accuracy of the buyer's forecast (as distinct from the prior probability that he will actually use the contingent capacity) is high.
3. The buyer is relatively more risk averse.

Note that as the accuracy of the buyer's forecast approaches 1, the conditions in (59) converge to the following:

$$e_2^* > e_1^* \Leftrightarrow\ v \geq \frac{1}{3} \tag{60}$$

As would be expected, this is precisely the condition that attaches to the "infallible buyer forecast", (36), when the prior probability of contingent use, ϕ, is $\frac{1}{2}$, which corresponds to the prior probability that $\sigma = 1$ in this case, as established in (38).

[9] Naturally, the actual values of the seller's maximal revenues differ in each case.

The effects of adverse selection can be seen at work in COD pricing for this case, too. For consider the limiting case as R_A tends to zero; such a risk neutral buyer, faced with a payoff of at least $1 in either state of the world, would be willing to pay at least $1 for the prospect. In such an instance, it follows directly from (47) that the bulk seller's revenues are $1+(1-\rho)v$ dollars, which is always greater than $1. For the COD seller, however, maximum expected revenues are:

$$e_2^* = \frac{3}{2}\frac{1+\rho v}{1+\rho} \tag{61}$$

Note then that:

$$v < \frac{2\rho - 1}{3\rho} \quad \Rightarrow \quad \frac{3}{2}\frac{1+\rho v}{1+\rho} < 1 \tag{62}$$

Again, we see that under certain circumstances, the COD seller is unable to extract consumer surplus that she knows must be present. (In fact other things being equal, such circumstances are more likely to arise as the buyer's forecast quality, ρ, increases.)

7 Conclusions

This paper has investigated some very simple models intended to capture some of the salient aspects of capacity-on-demand computing products. Such products have the peculiar property of exacting a fixed cost from the seller while returning variable revenues depending on the buyer's use of the product from period to period. Using the framework of Bayesian games, we compared the maximum revenues that a seller could expect from such products with those from a more conventional offering whose price was not contingent on usage. We explored three possible scenarios: One in which both seller and buyer were privy to the same information concerning the buyer's future usage; one in which the buyer (but not the seller) was certain of his future usage; and one in which the buyer had more information than the seller (but not certainty) concerning his usage. In each case, we were able to characterize the conditions under which the seller could expect greater revenues with the capacity-on-demand product than with the conventional one (neither was uniformly superior). Broadly speaking, the seller could expect to fare better selling capacity-on-demand to a buyer who was relatively more risk averse, who placed a relatively greater value on the contingent capacity in the product and who was either more likely to use the contingent capacity (in the event the buyer had no private information) or was more certain as to whether or not he would need the capacity (when the buyer had a private forecast). However, we were also able to show that in certain circumstances, a COD seller would be unable reliably to capture surplus that she is sure the buyer must possess — a problem that the seller of the conventional product would not face.

The latter effect is, as we indicated earlier, a manifestation of *adverse selection* — a phenomenon studied at length in the literature, beginning with the seminal (Akerloff 1970); see (Riley 2001) for a comprehensive survey. In capacity-on-demand pricing, adverse selection issues from the fact that the possession of private information on the part of the buyer of his future usage means that he is only likely to pay higher per-unit charges for his contingent capacity if he knows he's unlikely to use many such units. Such behavior, as we have seen, limits the expected revenues that the seller realizes.

Since adverse selection has been the object of sedulous investigation for over thirty years now, it's probably not surprising that an array of mechanisms have been proposed to address it. For problems like those discussed in this paper, these mechanisms amount to pricing schemes involving a menu of different product-price combinations, where products in the combinations differ in terms of quality or quantity — see (Wilson 1997) or (Laffont and Martimont 2001) for details. Properly designed, these pricing schemes implement *second degree price discrimination*, whereby buyers "self-select" based on their private information, allowing the seller to extract the maximum benefit possible, given that she is ignorant of the buyer's information at the outset.

For the very simple model presented in this paper, a correspondingly simple COD pricing scheme entirely obviates the effects of adverse selection. This scheme prices both units of the product at their maximum marginal benefit — i.e., a buyer using a single unit pays \$1, and a user of both units pays $\$1+v$. It is straightforward to demonstrate that with such a pricing scheme, whether the buyer possesses no forecast, a fallible forecast or an infallible one, he is never left with consumer surplus, and in particular, that the seller is never left unable to extract utility that she knows the buyer must possess.

Naturally, the near-triviality of this optimal pricing scheme results from the very minimal nature of the model itself — the model postulates only one unit of base capacity and one unit of contingent capacity, whose maximum benefits are known to both parties. Of course in reality, the situation is likely to be much more complex, with multiple units of both types of capacity and benefits that are uncertain to the seller and/or to the buyer. Therefore a logical extension of the work presented here would be the exploration of COD pricing strategies in more complicated settings; a comprehensive discussion of many of the issues to be tackled can be found in (Wilson 1997).

Another possible direction for future work would be to apply the insights gained here to the utility computing products mentioned earlier. In many ways, these can be viewed as a development of capacity-on-demand/pay-per-use products, which apply similar pricing concepts (i.e., payment per unit of capacity used) to *shared* computing resources (normally located at the vendor's site). A promising start might be to extend the minimal models in this paper, so as to incorporate multiple buyers whose demands for contingent capacity are more or less correlated. Techniques like those used in risk modeling for insurance, investment portfolio and credit provision (Bluhm et al. 2002), (Cherubini et al. 2004) could be used to explore the ramifications of the correlation structure of the buyers' demands.

8 References

Akerloff, George, 1970. The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, August, pp. 488-500.

Bluhm, Christian, Overback, Ludger and Wagner, Christoph, 2002. *An Introduction to Credit Risk Modeling*. Chapman & Hall/CRC.

Cachon, Gerard, 2003. Supply chain coordination with contracts. *Handbooks in Operations Research and Management Science: Supply Chain Management*. Graves, Steve and de Kok, Ton (eds.), North Holland.

Cachon, Gerard and Lariviere, Martin, 2001. Contracting to assure supply: how to share demand forecasts in a supply chain. *Management Science*, 47(5), pp. 629-46.

Cherubini, Umberto, Luciano, Elisa and Vecchiato, Walter, 2004. *Copula Methods in Finance*. John Wiley & Sons.

Fudenberg, Drew and Tirole, Jean, 1991. *Game Theory*, MIT Press.

Hellerstein, Joseph L., Katircioglu, Kaan, Surendra, Maheswaran, 2004. *A Framework for Applying Inventory Control to Capacity Management for Utility Computing*. IBM Research Report, RC23373, October.

HP, 2005. *On-demand solutions from HP: pay-as-you-go IT for an unpredictable world*. Retrieved February, 2005 from http://h10016.www1.hp.com/hpfinancialservices/pdf/OD_Final_5981-5499EN.pdf

IBM, 2005. *Capacity on Demand for pSeries servers: Types of COD*. Retrieved February, 2005 from http://www-1.ibm.com/servers/eserver/pseries/ondemand/cod/types.html

Laffont, Jean-Jacques, Martimont, David, 2001. *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, 2001.

Lariviere, Martin A. 2002. *Inducing Forecast Revelation through Restricted Returns*. Working Paper, June, retrieved February 2005 from: http://www.kellogg.northwestern.edu/faculty/lariviere/research/ForecastRevelation.htm.

Liu, Zhen, Squillante, Mark S., Wolf, Joel L., 2001. On maximizing service-level-agreement profits. *Proceedings of the 3rd ACM Conference on Electronic Commerce*.

Lyons, Daniel. 'On Demand' letdown. *Forbes.com*, February 24, 2005. Retrieved February 2005 from http://www.forbes.com/enterprisetech/2005/02/24/cz_dl_0224ibmsidebar.html

Mas-Colell, Andreu, Whinston, Michael and Green, Jerry, 1995. *Microeconomic Theory*. Oxford University Press.

Mendelson, Haim, 1985. Pricing computer services: Queueing effects. *Communications of the ACM*, Vol 28, No. 3 (March).

Myerson, Roger, 1997. *Game Theory: Analysis of Conflict*. Harvard University Press.

Miyaoka, Julia, 2005. *Implementing Collaborative Forecasting Through Buyback Contracts*. Retrieved February 2005 http://bctim.wustl.edu/topics/topics.cfm?categories_id=33&searchid=45.

Nicholson, Walter, 2001. *Microeconomic Theory: Basic Principles and Extensions*. South-Western College Publishers.

Paleologo, Giuseppe A., 2004. Price-at-Risk: A methodology for pricing utility computing services. *IBM Systems Journal*, Vol. 43, No. 1.

Partridge, Richard, 2004. HP, IBM, and Sun COD offerings : An overview. *D.H. Brown Technology Trends Report*, D.H. Brown, Port Chester, NY.

Riley, John G., 2001. Silver Signals: Twenty-Five Years of Screening and Signaling. *Journal of Economic Literature*, 39, 432-478.

Shankland, Stephen, 2005. Sun bangs utility computing drum. *CNET News*, January 31, 2005. Retrieved, February, 2005 from http://news.com.com/Sun+bangs+utility+computing+drum/2100-7339_3-5557902.html

Sun, 2005. *Sun in the Data Center – Capacity on Demand*. Retrieved February, 2005 from http://www.sun.com/datacenter/cod/index.html

Sundararajan, Arun, 2004. Nonlinear pricing of information goods. *Management Science*, Volume 50, Number 12, December.

Varian, Hal R., 2000. Buying, sharing and renting information goods. *Jounal of Industrial Economics*, Vol XLVIII, No. 4, December 2000.

Wilson, Robert B., 1997. *Nonlinear Pricing*. Oxford University Press.

About the Author

Phillip M. Yelland is a senior staff engineer at Sun Microsystems Laboratories. He has worked in a wide variety of settings, ranging from formal methods research to software development management, both during his time at Sun and previously. Throughout, his activities have reflected a preoccupation with the application of theoretical results to the practical conduct of business, and his research currently centers on the use of statistical and microeconomic techniques for operations management. He has an M.A. and Ph.D. in Computer Science from the University of Cambridge in England, and an M.B.A. from the University of California at Berkeley.