Decision Support Systems 58 (2014) 31–42

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Capacity planning and performance contracting for service facilities

Yabing Jiang *, Abraham Seidmann

[a] Lutgert College of Business, Florida Gulf Coast University, Fort Myers, FL 33928 USA
[b] Simon Graduate School of Business, University of Rochester, Rochester, NY 14627 USA

ARTICLE INFO

Available online 16 January 2013

Keywords:
Cost allocation
Capacity planning
Franchise contract
Information asymmetry
Mechanism design
Service center

ABSTRACT

Market demand uncertainty and time-based competition make capacity investment and managerial incentive decisions for service facilities such as high-end diagnostic medical imaging centers, modern IT services, or contract manufacturing shops particularly challenging. These facilities compete on service quality, short queuing times and speed. Therefore, having insufficient capacity can be economically devastating for them. Given the high up-front costs involved, firms want to make sure that they neither over- nor under-invest in service capacity. These problems get exasperated by the fact that typically firms are unfamiliar with the local market conditions and do not closely observe the demand-generating efforts of the hired managers. Most prior studies of cost allocation methodologies, contract design, and service resource management tend to address these aspects of the problem separately. They ignore the interaction effects between the capacity decisions and the managerial adverse selection and moral hazard issues, which are crucial elements for successfully running services with fixed capacity, random arrivals, and stochastic service times. Our paper instead focuses on the development of an integrated-approach to the simultaneous design of efficient managerial contracts and of capacity planning for capital intensive service facilities. We derive optimal linear contracting structures under information asymmetry between the firms and management, and analyze their impact on capacity decisions, service levels, service volumes, and the allocations of costs. Surprisingly, we prove that even though a franchise (charge-back) contract induces the first-best effort from the manager, it is not always the best choice for the firms as it may lead to inferior profits for them. In fact, our results explain why a firm's eventual contract choice should be a function of its prior on the probability distribution of the local market demand. We also explain when it may be optimal (for both the firm and the manager) to charge the manager up front a fixed franchising fee that is even greater than the total costs of capacity. Our study applies to many capital-intensive and congestion-prone service systems, where the success of significant up-front capacity investments also hinges on the daily operations of those facilities run by hired managers—who typically possess specific knowledge—that gives them a significant information advantage.

1. Introduction

Effectively managing service centers such as call centers, computerized diagnostic imaging facilities, data centers, SaaS businesses, and telecommunication networks has always been a challenging task. Owners of the centers (firms) are responsible for investing in capacity, which is often capital intensive and involves significant up-front fixed costs for equipment, software, and installation. While high utilization is a critical profitability driver, facing uncertain market demand, firms also have to maintain an acceptable service level and relatively short waiting times in order to compete in the market successfully. Managers of these centers are often contracted to run them as profit centers. As such, they are usually responsible for directing daily activities and generating demand through additional marketing or service quality efforts. For example, we have seen that at most free-standing radiology facilities that are run by the national networks in the US (such as Insight Imaging and Radnet Inc.), the local managers are responsible for staffing, scheduling, marketing, and other demand-generating activities in their territory.

What makes the issue complicated is that managers, as the agents running the service centers, often possess private information about local market conditions, referral patterns, demographic preferences, and their own demand-generating efforts, and they may not willingly or truthfully share that information with their firms. Working with a national firm that operates high-end diagnostic medical imaging centers in various states, we had to make the correct capacity investments in each market and to obtain the true market demand information from the local managers who are typically being incentivized by the patient volume at their center. If the local managers overestimate the demand, to allow ample capacity and fast turnaround times for all exams, the firm will find itself over-investing in capacity, such as MRI, CT, or PET-CT equipment. On the other hand, if the local managers

* Corresponding author.
E-mail addresses: yjiang@fgcu.edu (Y. Jiang), seidmannav@simon.rochester.edu (A. Seidmann).

0167-9236/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2013.01.010

underestimate the demand, in order to have a lower 'service quota' in their contract, the firm will under-invest in capacity, resulting in long lines and lost profit opportunities.

In this paper, we outline the solution to this problem that has been motivated by our practical experience and by the latest research on optimal mechanism design. In our model, the manager (agent) possesses private information regarding the local market and the level of his own demand-generating effort; facing uncertain market demand and the combined agency issues, the firm has to make up-front capacity investment decisions and design incentive contracts to motive the manager to truthfully share his specific market knowledge and to exert the desired level of effort to generate demand. Several studies [4,10,22,27] have pointed out that optimal mechanisms predicated by economic theory are often sophisticated and rarely used in the real world, whereas linear contracts are commonly used in practice. Our focus is on evaluating three effective linear performance-based incentive contracts under information asymmetry: a *uniform contract*, a *variable-rate contract*, and a *charge-back contract*. We show that the *uniform contract*, which offers the same contract terms regardless of market conditions, fails to elicit market information from the manager and leads to a distorted capacity investment decision. The *variable-rate* and *charge-back* (franchise) contracts, on the other hand, can successfully elicit true market information from the manager. We show that the latter even induces the first-best (full-information) effort level from the manager and hence leads to the optimal capacity investment, the same as the full-information level.

We embed a queuing model within the principal-agent framework because most service firms are constrained by a finite installed capacity, yet they have to deliver reasonably fast turnaround times when both demand and service rates could be random and outside their control in the short run. Integrating the capacity decision with the incentive contract design and the service level constraints, therefore, enables us to address a complicated and yet practical problem that has not been fully studied before. We not only derive analytical solutions for this intricate problem but also find some interesting results. For example, we show that if the firm selects the proper fixed and variable charge terms, its franchise contract can induce the first-best demand-generating effort and it will successfully elicit truthful market information from the manager; however franchise contracting fails to produce the first-best profit and it may not always be the best choice for the firm. The fixed charge term in the franchise contract is naturally associated with the classical fixed cost allocation. It has been shown in managerial accounting that firms can benefit from fixed cost allocation because the allocation may influence agents' consumption of perquisites or serve as a proxy for the difficult-to-calculate opportunity costs [28]. Overall, the accounting literature tends to emphasize the role of fixed cost allocation as an effective cost control to hold agents accountable so that they will not over-consume resources at the expense of the firm [2,15,27]. We extend that literature and show here in a service queuing setting firms should also use cost allocation as a way to prevent the managers from under-consuming resources. A high capacity request is associated with a high fixed charge, and the managers have a natural incentive to request low capacity when they privately observe a high-demand market. By doing so, they can shirk in their demand-generating effort and gain greater utility. But, under-consuming resources can be harmful to the firms as well, because low capacity restricts the imaging center from serving more patients, for instance, and also causes long patients' waiting times, which translate to high delay costs and potential losses in future revenue due to a degraded service-level reputation. We show that with proper fixed cost allocations firms can avoid both over-investing and under-investing in capacity.

Our work contributes to the literature in multiple ways. First, we develop and analyze a new contracting problem in a service environment in which firms compete on meeting certain response times and local managers possess valuable private information on market demand. These managers can influence market demand by exerting demand-generating effort, such as marketing promotions or by providing an outstanding customer service experience. Second, we optimally integrate firms' operations capacity decision with their incentive contract decision for a business environment with stochastic arrivals of customers and randomly distributed service times per customer. We investigate three practical mechanisms and show that two of them can successfully solicit the manager's market information. Third, we demonstrate the value of the local manager's private information and prove that the firm's optimal contract choice is a function of the market demand distribution. As expected, we show that the cost of the manager's private information is associated with market demand variance, with a higher demand variance corresponding to a higher information cost. However, the manager's market knowledge is more valuable when there is a low demand variance. Overall, this research provides guidelines for firms that deal with congestion-prone systems and sheds light on how to effectively manage service facilities with combined moral hazard and adverse selection issues.

The rest of the paper is organized as follows. We first review the relevant literature in Section 2. Then we present the model setting in Section 3 and analyze the benchmark case with *full information* in Section 4. In Section 5, we derive three contracts for the information asymmetry case in which the manager's effort is not observable and the manager has private information regarding local market condition. We explain why two of these contracts can effectively solicit true market information from the manager. We then analyze the firm's optimal contract choice and evaluate the factors affecting the installed capacity, the average throughput rates, and the expected waiting times in Section 6 and conclude the paper with practice guidelines for cost allocation and incentive design for service systems in Section 7.

2. Literature review

This paper is related to research that applies agency theory as well as the service resource management and accounting cost allocation literature. Prior research on IT resource management has modeled IT service centers as queuing systems and mostly focused on capacity allocation within a firm through an internal pricing scheme [6,7,18]. In these works, demand is exogenous and there is no agency issue. Clearly, this model setting does not apply to many modern IT service centers that provide services to external clients because external random demand cannot by controlled by the proposed internal transfer pricing scheme, and it fails for any service centers that do experience agency issues.

Marketing literature that applies the principal-agent framework, for example, Basu et al. [3] and Lal and Srinivasan [14], often focuses on the compensation contract design of a sales force selling a commodity good with an unlimited supply and does not work for a service environment. Harris et al. [8] study effective resource allocation in a manufacturing setting under incomplete information. This is one of the early works on mechanism design. In their model, agents have private information on their productivity and effort levels. The objective of the firm is to minimize the costs of producing a given level of output by designing an efficient mechanism. However, since the output level is given there is no demand uncertainty in their model.

Among recent research that combines queuing models with agency theory, Jiang and Seidmann [12] model the capacity management and contract design issue of a service organization with a finite capacity. In their model, the service demand can be affected by the agent's marketing effort, and the agent is both risk and effort averse. They extend prior agency theory literature by introducing delay costs and the capacity decision into the model and taking an integrated approach to derive the optimal incentive contract and capacity investment for the