VOLUME AND CAPACITY INTERACTION IN FACILITY DESIGN

ARNOLD H. BUSS

Operations Research Department, Naval Postgraduate School, Monterey, CA 93943-5000

STEPHEN R. LAWRENCE

College of Business Administration, University of Colorado, Boulder, CO 80309-0419

DEAN H. KROPP

John M. Olin School of Business, Washington University, St. Louis, MO 63130

This paper addresses the joint facilities design problem of determining both demand and capacity with stochastic demand arrivals and stochastic processing throughput. Using a simple *M/M/*1 queueing model of a profit maximizing firm, we link marketing and production decision variables by recognizing appropriate congestion costs, and show that coordinated decision-making provides results superior to making demand and capacity decisions sequentially. Sensitivity analysis indicates that the model is robust with respect to its assumptions and parameters. An example illustrates the approach and demonstrates the application of the model.

■ In this paper we address the joint problem of determining demand volume and production capacity under conditions of uncertainty. By introducing congestion costs, we link the demand volume decision (typically determined by the Marketing function) with the capacity design decision (typically made by the Manufacturing/Operations function). Analysis of the resulting model shows that the usual pattern of subordinating the production capacity decision to the demand volume decision is suboptimal when demand arrivals and processing times are uncertain, and further demonstrates that ignoring congestion costs can lead to potentially expensive errors when making volume and capacity decisions.

There is increasing interest in both the research literature and in practice in coordinating manufacturing and marketing functions. Much has been written about the problems of "functional silos" in which functional specialties make important decisions in isolation from other interested parties (Hayes, Wheelwright, and Clark [12]), and about the need for greater functional coordination (Shapiro [18]). Presumably, coordinating activities across functional boundaries will result in lower costs, faster response to customers, greater flexibility, and better utilization of resources (Crittenden [4]). However, much of this literature is anecdotal without theoretical or analytic underpinning. A principal objective of this paper is to provide such support for the facilities planning problem.

A second objective of the paper is to add support to the growing literature which argues that congestion effects must be recognized when addressing production problems. Much of the traditional capacity design literature assumes, either implicitly or explicitly, that capacity utilization rates of 100 percent are feasible and, indeed, desirable. However, recognition that workflow congestion creates negative impacts is increasing (Fry and Blackstone [8]). Manufacturing plants have typically used direct labor and machine utilization as the principle measure of production performance (Eloranta [5]). While direct labor charges are a diminishing fraction of total production costs in most manufacturing firms, the cost of materials is increasing relative to total production costs, thus shifting managerial attention from the control of direct labor expenditures to the control of inventories. Large work-in-process (WIP) inventories, once considered desirable to buffer production and provide high utilization, are currently regarded as costly impediments to effective manufacturing by obscuring quality problems, increasing lead times, reducing flexibility, and tying-up expensive capital. Since workflow congestion leads to increased WIP inventories, attention is shifting to relieving congestion and hence reducing WIP.

To capture the interaction between demand volume and production capacity in facilities design and to include the effects of congestion, we develop a simple profit-maximizing model of an *M/M/*1 production system under assumptions of declining returns to scale for both demand and capacity. With this model the mean arrival rate represents expected demand volume, the mean service rate represents production capacity, and the mean time-in-system captures workflow congestion. We deliberately choose a simple and parsimonious model for this analysis in order to capture the aggregate effects of demand volume and production capacity decisions in the spirit of Manne [15]. Consequently, we do not address other facilities design considerations such as plant layout, product mix, technology selection, and so forth. Other objectives such as return on investment (ROI) could also be considered. Although we briefly consider this, we will assume that we are dealing with a profit-maximizing firm.

The balance of the paper is organized as follows: In the next section we discuss the relevant capacity design literature, for which the decision variables have been either the demand rate or the service rate. The following section contains our basic model, which has both the demand and service rates as decision variables. We then analyze the model and give results about profitability and optimal utilization. Sensitivity analysis is undertaken, where the effects of errors in key model parameters are considered, and we illustrate an appli-

 0740-817X/94/$3.00x.00