

Match Supply and Demand in Service Industries

by W. Earl Sasser

From the Magazine (November 1976)

What makes service industries so distinct from manufacturing ones is their immediacy: the hamburgers have to be hot, the motel rooms exactly where the sleepy travelers want them, and the airline seats empty when the customers want to fly. Balancing the supply and demand sides of a service industry is not easy, and whether a manager does it well or not will, this author writes, make all the difference. In this rundown of the juggling feat service managers perform, the author discusses the two basic strategies—"chase demand" and "level capacity"—available to most service companies. He goes on to discuss several ways service managers can alter demand and influence capacity.

The literature on capacity management focuses on goods and manufacturing, and many writers assume that services are merely goods with a few odd characteristics. Unfortunately, these researchers never fully explore the implications of these strange traits:

1. Services are direct; they cannot be inventoried. The perishability of services leaves the manager without an important buffer that is available to manufacturing managers.

2. There is a high degree of producer-consumer interaction in the production of service, which is a mixed blessing; on the one hand, consumers are a source of productive capacity, but on the other, the consumer's role creates uncertainty for managers about the process's time, the product's quality, and the facility's accommodation of the consumer's needs.

3. Because a service cannot be transported, the consumer must be brought to the service delivery system or the system to the consumer.

4. Because of the intangible nature of a service's output, establishing and measuring capacity levels for a service operation are often highly subjective and qualitative tasks.

Whereas the consumption of goods can be delayed, as a general rule services are produced and consumed almost simultaneously. Given this distinction, it seems clear that there are characteristics of a service delivery system that do not apply to a manufacturing one and that the service manager has to consider a different set of factors from those that would be considered by his or her counterpart in manufacturing. And if one looks at service industries, it is quite apparent that successful service executives are *managing* the capacity of their operations and that the unsuccessful are not. So, the "odd characteristics" often make all the difference between prosperity and failure.

Consider the following service managers' actions, which resulted in fiasco:

Increasing the wrong kind of capacity—In studying the battle statistics in the war for market share among airlines, competitors observed that an air carrier in a minority position on a particular route would often get a smaller proportion of the total passengers flown on the route than the share of seats flown.¹ Conversely, the

dominant airline would carry a disproportionately larger share of the total passengers flown. The conclusion was obvious: Fly the seats, and you get the passengers.

In an effort to fly more seats, the airlines lined up to purchase jumbo jets. However, when competitors began flying smaller planes more frequently on the same routes and reaping a good number of passengers, it became painfully apparent to many airlines that frequency (and, to some extent, timing) of departures is the key to market share. Consequently, the airlines “mothballed” many of the jumbos or sold them if they could.

Not increasing all-around capacity—A resort operator decided to increase the number of rooms in a lodging facility and not to expand the central services required to support the additional guests. The fact that room rentals contribute up to 90% of total revenue and that tennis courts, swimming pools, meeting rooms, parking areas, and so on contribute next to nothing, or nothing, convinced the operator to create an imbalance in favor of revenue-producing activities. However, the number of guests adjusted itself to the level of occupancy that the central services could support, not to the level of room capacity. The room capacity beyond the level supported by the central services was wasted.

Not considering the competitive reaction—The Orlando, Florida lodging industry’s response to the announcement of Disney World’s opening is a classic example of this type of service management fiasco. Disney executives had learned well the lessons of Orange County, California, and Disneyland, where revenue is limited to on-site entertainment, food, and souvenir dollars. However, businesses besides Disney have made large profits in lodging, restaurant, and recreational facilities. Correctly perceiving that the same thing would happen in Florida, Disney purchased 200,000 acres south of Orlando, eight times the number owned in Anaheim.

When news broke that Disney would build in central Florida, however, everybody with a hotel or motel in his or her portfolio began plans for Orlando units, even though Disney had preempted all the land within two miles of the Magic Kingdom. The subsequent overbuilding has been well documented. More than 30,000 rooms were built to service a market estimated to need only 19,000. As an Orlando lender moaned, “We had a great little 200-room property there, the only one at the intersection. In less than a year, there were 5,000 rooms either built, under construction, or planned within a quarter mile of that intersection. We had to foreclose, and our occupancy has been running at only 35%.”

Undercutting one's own service—A new entrant in the overnight air freight transportation industry discovered that attempts to capture market share by adding to the existing number of planes and branch offices increased costs faster than revenues. Still looking for market share, the company then offered lower rates for second- and third-day deliveries. Because it had excess capacity, however, the company always delivered packages on the next day. As consumers discovered this fact, the mix of business shifted dramatically to the lower-priced services. So although there was an increase in volume, the resulting lower margins pushed the break-even volume even higher.

These pitfalls are not inevitable. Successful service executives do avoid them, and there are enough examples of well-managed service businesses from which to glean some wisdom on how to match demand for services with capacity to supply them. There are two basic capacity-management strategies available to most companies and a number of ways open to executives to manage both the demand and the supply sides of their businesses. I will discuss the strategies and choices in turn.

Two Basic Strategies

Consider the national operations group of the XYZ brokerage firm. The group, housed in an office building located in the Wall Street area, handles the transactions generated by registered representatives in more than 100 branch offices throughout the United States. As with all firms in the brokerage industry, XYZ's transactions must be settled within five trading days. This five-day period allows operations managers to smooth out the daily volume fluctuations.

But fundamental shifts in the stock market's volume and mix can occur overnight, and the operations manager must be prepared to handle extremely wide swings in volume. For example, on the strength of an "international peace" rumor, the number of transactions for XYZ rose from 5,600 one day to 12,200 the next.

However, managers of XYZ, not unlike their counterparts in other firms, have trouble predicting volume. In fact, a random number generator can predict volume a month or even a week into the future almost as well as the managers can.

How do the operations managers in XYZ manage capacity when there are such wide swings? The answer differs according to the tasks and constraints facing each manager. Here's what two managers in the same firm might say:

- Manager A—"The capacity in our operation is currently 12,000 transactions per day. Of course, what we should gear up for is always a problem. For example, our volume this year ranged from 4,000 to 15,000 transactions per day. It's a good thing we have a turnover rate, because in periods of low volume it helps us reduce our personnel without the morale problems caused by layoffs." (The labor turnover rate in this department is over 100% per year.)

- Manager B—“For any valid budgeting procedure, one needs to estimate volume within 15%. Correlations between actual and expected volume in the brokerage industry have been so poor that I question the value of budgeting at all. I maintain our capacity at a level of 17,000 transactions per day.”

Why the big difference in capacity management in the same firm? Manager A is in charge of the cashiering operation—the handling of certificates, checks, and cash. The personnel in cashiering are messengers, clerks, and supervisors. The equipment—file cabinets, vaults, calculators—is uncomplicated.

Manager B, however, is in charge of handling orders, an information-processing function. The personnel are key-punch operators, EDP specialists, and systems analysts. The equipment is complex—cathode ray tubes, key-punch machines, computers, and communication devices that link national operations with the branches. The employees under B’s control had performed their tasks manually until increased volume and a standardization of the information needs made it worthwhile to install computers.

Because the lead times required to increase the capacity of the information-processing operation are long, however, and the incremental cost of the capacity to handle the last 5,000 transactions is low (only some extra peripheral equipment is needed), Manager B maintains the capacity to handle 17,000 transactions per day. He holds to this level even though the average number of daily transactions for any month has never been higher than 11,000 and the number of transactions for any one day has never been higher than 16,000.

Because a great deal of uncertainty about the future status of the stock certificate exists, the situation is completely different in cashiering. Attempts to automate the cashiering function to the

degree reached by the order-processing group have been thwarted because the risk of selecting a system not compatible with the future format of the stock certificate is so high.

In other words, Manager A is tied to the “chase demand” strategy, and his counterpart, Manager B in the adjacent office, is locked into the “level capacity” strategy. However, each desires to incorporate more of the other’s strategy into his own. A is developing a computerized system to handle the information-processing requirements of cashiering; B is searching for some variable costs in the order-processing operation that can be deleted in periods of low volume.

The characteristics of these two vastly different strategies are outlined in the Exhibit, “Comparison of chase-demand and level-capacity strategies for the XYZ brokerage firm.”

Exhibit Comparison of chase-demand and level-capacity strategies for the XYZ brokerage firm		
	Chase demand	Level capacity
Labor-skill level required	Low	High
Job discretion	Low	High
Compensation rate	Low	High
Working conditions	Sweatshop	Pleasant
Training required per employee	Low	High
Labor turnover	High	Low
Hire-fire costs	High	Low
Error rate	High	Low
Amount of supervision required	High	Low
Type of budgeting and forecasting required	Short-run	Long-run

Exhibit Comparison of chase-demand and level-capacity strategies for the XYZ brokerage firm

Service managers using the chase strategy are usually responsible for unskilled employees performing jobs with little or no discretion for low pay in a relatively unattractive environment. Managers use the level strategy most often where more highly skilled people perform jobs for high pay, with some or a lot of discretion in a relatively pleasant environment.

Because the skill-level requirement for “chase” is lower than that for “level,” the training cost per employee will also be lower for “chase.” However, the annual training costs in a department using the chase strategy could be much higher than for one using the level strategy. The chase strategy requires more employees, and those employees exhibit a higher rate of turnover because of the job characteristics just described.

The chase strategy is usually more costly than the level strategy for other reasons as well. The high turnover rate and the use of unskilled employees both contribute to a high error rate, which means that more supervisors are needed to ensure that jobs are performed according to specifications.

For the chase strategy, the lead times required to attract and train new employees in periods of increased volume and to reduce the work force in periods of contraction are so short that forecasting and budgeting is needed only for the short run. However, because managers using a level strategy need a longer lead time to acquire or dispose of equipment and trained personnel, for them, forecasting and budgeting is a long-run process.

Although the chase demand strategy has many negative connotations for enlightened managers, there are some service delivery systems, such as amusement parks and resort hotels with highly seasonal or random fluctuations in demand, that survive only as a result of its successful application.

Managing Demand & Supply

Besides electing to adopt one of the strategies just described, the service executive may select one or another additional way to cope with a fluctuating demand schedule. To understand how one business did it, see the sidebar, “Finding a Creative Solution.”

Finding a Creative Solution

A southeastern U.S. resort had a problem: in all probability, the demand for its tennis facilities would be very high in July and ...



Altering Demand

The manager can attempt to affect demand by developing off-peak pricing schemes, nonpeak promotions, complementary services, and reservation systems. Let's look at each of these demand-leveling options in turn:

Pricing

One method managers use to shift demand from peak periods to nonpeak ones is to employ a differential pricing scheme, which might also increase primary demand for the nonpeak periods. Examples of such schemes are numerous. They include matinee prices for movies, happy hours at bars, family nights at the ball park on week nights, weekend and night rates for long-distance calls, peak-load pricing by utility companies, and two-for-one coupons at restaurants on Tuesday nights.

Developing nonpeak demand

Most service managers wrestle constantly with ideas to increase volume during periods of low demand, especially in those facilities with a high-fixed, low-variable cost structure. The impact of those incremental revenue dollars on the profitability of the business is tremendous. Examples of attempts to develop non-peak demand are not hard to find. Hamburger chains add breakfast items to their menus, and coffee shops add dinners to theirs. Urban hotels, which cater to the business traveler during the week, develop weekend “minivacation” packages for the suburban population in their geographic areas, while resort hotels, jammed with pleasure travelers during school vacations, develop special packages for business groups during off-seasons.

However, caution must be used in developing plans to increase demand for the underused periods of the service facility. Many companies have made costly mistakes by introducing such schemes and not seeing the impact they would have on existing operations. As Wickham Skinner has noted, for manufacturing companies, there are some real costs associated with “unfocusing” the service delivery system, which is exactly what market-expanding activities have a tendency to do.² New concepts often require equipment and skills not currently found in a service delivery system. The addition of these skills and equipment may require a new type of labor force, a new layout, or more supervision.

Even if the new concept succeeds in creating demand in nonpeak periods, the effects are not always positive. Managers often use slack time productively as a time to train new employees, do maintenance on the equipment, clean the premises, prepare for the next peak, and give the workers some relief from the frantic pace of the peak periods. A new concept, therefore, may have a tendency to reduce the efficiency of the present system at best, or, at worst, to destroy the delicate balance found in most service delivery systems.

Developing complementary services

Another method managers use to shift demand away from peak periods is to develop complementary services, which either attracts consumers away from bottleneck operations at peak times or provides them with an alternative service while they are in the queue for the capacity-restricted operations. For example, restaurant owners have discovered that on busy nights most patrons complain less when sitting in a lounge with cocktail then when standing in line as they wait for tables in the dining area. Also, the profitability of restaurants with bars can more than double.

A diversion can also relieve waiting time. A hotel manager installed mirrors on each floor's central lobby so that customers could check their appearance while they waited for the elevator. Banking by mail or by automated tellers are other ways to cut down customer waiting time.

Creating reservation systems

Service executives can effectively manage demand by employing a reservation system, which in essence presells the productive capacity of the service delivery system. When certain time periods are booked at a particular service facility, managers can often deflect excess demand to other time slots at the same facility or to other facilities at the same company and thereby reduce waiting time substantially and, in some cases, guarantee the customer service.

For instance, if a motel chain has a national reservation system, the clerk can usually find a customer a room in another motel of the chain in a fairly close proximity to his or her desired location if the first-choice motel is full.

In a similar manner, airlines are often able to deflect demand from booked flights to those with excess capacity or from coach demand to first class, especially if their competitors do not have seats available at the consumers' desired flight time.

However, reservation systems are not without their problems, the major one being “no-shows.” Consumers often make reservations they do not use, and, in many cases, the consumer is not financially responsible for the failure to honor the reservation. To account for no-shows, some service companies oversell their capacity and run the risk of incurring the wrath of customers like Ralph Nader, who do show. Many service companies have made it a policy to bill for capacity reserved but not used if the reservation is not cancelled prior to a designated time.

Controlling Supply

The service manager has more direct influence on the supply aspects of capacity planning than he or she does on the demand side. There are several things a service manager can do to adjust capacity to fluctuating demand.

Using part-time employees

Many service companies have found that it is more efficient to handle demand whenever it occurs than it is to attempt to smooth out the peaks. The peaks vary by type of business—during certain hours of the day (restaurant), during certain days of the week (hair styling), during certain weeks of the month (banking), and during certain months of the year (income tax services). These service businesses usually maintain a base of full-time employees who operate the facility during nonrush periods but who need help during peak periods. One of the best-known resources is part-time labor pools, especially high school and college students, parents who desire work during hours when their children are in school, and moonlighters who desire to supplement their primary source of income.

Maximizing efficiency

Many service managers analyze their processes to discover ways to get the most out of their service delivery systems during peak demand periods. In effect, such analyses enable the service

company to increase its peak capacity for little additional cost.

For example, during rush periods employees perform only the tasks that are essential to delivering the service. If possible, managers use slack periods for doing supporting tasks, which in essence they are inventorying for peak periods.

To maximize efficiency, managers examine even peak-time tasks to discover if certain skills are lacking or are inefficiently used. If these skills can be made more productive, the effective capacity of the system can be increased. For example, paramedics and paralegals have significantly increased the productive time of doctors and lawyers. Even rearranging the layout of the service delivery system can have a major impact on the productivity of the providers of the service.

Another way to attack the peak capacity constraint is by cross-training. The service delivery system is composed of various components. When the system is delivering one service at full capacity, some sections of the system are likely to be underused. If the employees in these sections are able to deliver the peak service, they add capacity at the bottleneck. When the demand shifts and creates a bottleneck in other components of the system, the employees can shift back again.

Increasing consumer participation

The more the consumer does, the lower the labor requirements of the producer. Bag-'em-yourself groceries, salad bars at restaurants, self-service gas pumps, customer-filled-out insurance information forms, and cook-it-yourself restaurants are all examples of increased consumer participation in the production of services.

There are, of course, some risks to increasing consumer input: consumers might reject the idea of doing the work and paying for it too; the manager's control over delivery of the service is

reduced; and such a move can create competition for the service itself. A cook-it-yourself restaurant customer might just stay at home.

Sharing capacity

The delivery of a service often requires the service business to invest in expensive equipment and labor skills that are necessary to perform the service but that are not used at full capacity. In such cases, the service manager might consider sharing capacity with another business to use required, expensive, but underused resources jointly.

For example, a group of hospitals in a large urban area might agree that it is unnecessary for each to purchase expensive medical equipment for every ailment and that they ought to share capacity. One would buy cardiac equipment, another gynecological and obstetrical equipment, another kidney machines. Participating doctors would have admitting privileges at all hospitals. By sharing equipment, hospitals would not only better use expensive resources, but as groups of trained and experienced specialists developed at each facility, hospitals would also deliver better medical care.

The shared-capacity concept is possible in the airlines industry in several forms. Several airlines with infrequent flights in and out of a particular airport share gates, ramps, baggage-handling equipment, and ground personnel. In fact, some domestic airlines flying different routes with different seasonal demands exchange aircraft when one's dip in demand coincides with another's peak.

Investing in the expansion ante

Wise service managers often invest in an "expansion ante." When growth occurs, it sometimes becomes clear that some of the new development could have been done when the facility was originally constructed for much less cost and disruption. A careful analysis before the facility is built will show what these items are.

For instance, for a small investment, a restaurateur can build his kitchen with extra space in order to service more diners later on. Contractors can run wiring, plumbing, and air conditioning ducts to the edge of the building where the expansion will take place. The manager can inventory enough land for the expansion and additional parking requirements. These actions will allow the restaurant manager to increase capacity without having to renovate the kitchen, redo the wiring, plumbing, and air conditioning systems, or purchase adjacent land at much higher prices.

Seeking the Best Fit

Managing demand and supply is a key task of the service manager. Although there are two basic strategies for capacity management, the enlightened service manager will, in almost all cases, deviate from these two extremes.

The challenge to the service manager is to find the best fit between demand and capacity. In order to manage the shifting balance that characterizes service industries, managers need to plan rather than react. For example, managers should try to make forecasts of demand for the time periods under question. Then he or she should break the service delivery system down into its component parts, calculate the present capacity of each component, and arrive at a reasonable estimate of what the use of each component will be, given the demand forecast.

Because each system cannot handle infinite demands, the manager needs to question how much of the peak demand the system must handle. Just what is the appropriate level of service for the delivery system to provide? Once the manager can approximate the answer to these questions and has decided which of the basic strategies to employ, he or she is ready to experiment with the different options to alter demand and capacity. Each plan and option a manager arrives at can be costed, and the best fit for the particular service selected.

Ultimately, of course, on the demand side, a manager's true aim is to increase revenues through an existing service delivery system of given capacity. Once the true variable costs are subtracted out, all revenues flow to the bottom line. On the supply side, the manager aims to minimize costs needed to increase or decrease capacity.

When facing increased demand, the business raises its revenues with minimal investment. In times of capital rationing, small investments are often the only ones available to the company. When facing contracting demand, the manager needs to select the best way to adjust the system's capacity to a lower volume.

In following the ideas outlined in this article, service managers need to think creatively about new ways to manage demand and supply. The most important thing to recognize is that they both can be managed efficiently and that the key to doing so lies in planning.

1. See William E. Fruhan, Jr., *The Fight for Competitive Advantage: A Study of the United States Domestic Trunk Air Carrier* (Boston: Division of Research, Harvard Business School, 1972).

2. See Wickham Skinner, "The Focused Factory," HBR May–June 1974, p. 113.

A version of this article appeared in the November 1976 issue of *Harvard Business Review*.

WS

W. Earl Sasser is a Baker Foundation Professor and the chair of the Program for Leadership Development at Harvard Business School.

Recommended For You

How to Choose the Right Forecasting Technique



Putting the Service-Profit Chain to Work



Predicting Consumer Demand in an Unpredictable World



PODCAST The Pros and Cons of Our "Middleman Economy"

