

The Value Contribution of Service Capacity Management to the Company's Operations and Performance

by George Cosmin Tănase



Abstract

For producing and delivering a specific service unit, certain service resources are necessary. The service capacity describes the amount of service resources which are available for service production and delivery. Due to the direct contact between these service resources and the customer (or their objects), which are integrated into the service process, a service provider can only deliver as many service units as service capacities are available. However, in every service setting there are often situations where there is a *gap between demand and capacity*, meaning the gap between the number of potential users of the service and the amount of service resources available.

Keywords: demand level, service capacity, resources, perishability, revenues, customer perception, highly competitive market, predictability, demand patterns

JEL Classification: D24, L80

There are many examples which describe situations where service demand and capacity do not match. Actually, for most services it is unlikely that demand and capacity will match exactly. The underlying reason for this is that the *characteristics of services* are derived from the customer integration into the service process. This main characteristic results in a so-called perishability of service capacity. **Service capacity management** has various significant effects on value:

- When service capacities are not used, certain parts of fixed costs are not covered. A hotel potentially pays interest on loans financing the hotel facilities and rooms. When rooms are not used, a portion of these costs are not covered by the revenues generated by stays at the hotel.

- Vice versa, when a service provider's capacity is not enough to satisfy the existing demand, the firm 'realises' lost sales and revenues. When a hotel has enquiries for 200 stays, but can only offer beds to 150 guests, the potential revenues of the 50 inquiries are lost.
- Moreover, not being able to provide a room possibly results in dissatisfaction of these potential customers and this negative experience thus prevents them from approaching the hotel in future.
- Customers currently using a service might perceive a bad service quality when there are not enough service resources available.

According to this comprehensive value relevance of service capacity management there are many *links to other elements of the Service Value Chain*. Regarding the *primary value processes*, the interaction process where the customer meets the service resources in order to produce and deliver a service is negatively affected when the service capacities are overcrowded or overused during periods of high demand. In consequence, a bad service quality is perceived by the customer. In an extreme situation, an interaction with a specific customer cannot take place at all. On the level of the customer relationship, this can lead to an impeded customer acquisition when a prospective customer planned to try a service for the first time and was not able to use it because of capacity constraints. Also, customer retention can be affected when a customer is annoyed because of not receiving a service due to capacity constraints.

While the service capacity and consequently capacity management affect the primary value processes of a service provider, conversely, the *secondary value processes* are utilised by service capacity management. Regarding the service product, service providers apply value-added services in order to facilitate reservations or waiting times. Service pricing and communications are used in order to manage service demands according to existing patterns of demand.

Finally, *service resources* compose a service provider's capacity. Depending on the service, this capacity is constrained by the number of employees or the tangible or technological resources of a provider. Consequently, managing the amount of these resources is the object of service capacity management.

Gaps between service demand and capacity

Service capacity problems are determined by the existing service capacity on the one hand and the current service demand on the other hand. By contrasting these two variables, two generic *forms of capacity gaps* can be described which have different causes as well as different consequences for customer and provider. These forms are:

- demand gap;
- capacity gap.

When there are more capacities than are requested by consumers, the service capacity exceeds the demand and a **demand gap** results. The consequences are vacant service facilities (e.g. empty hotel rooms) or non-occupied service employees (e.g. train conductor waiting for more people to enter the train at the next stop). Vacant facilities result in fixed costs which are sunk without a respective revenue. Moreover, variable costs are generated because the service facility possibly still has to be maintained. An aeroplane has to be checked for technical problems regardless of whether the plane is fully booked or not. A hotel room which was not in use for several days probably is cleaned more than it is used. Unoccupied employees tend to feel underutilised, resulting in lower employee motivation, satisfaction and productivity.

Conversely, a **capacity gap** is the result of demand exceeding the service capacity available. In this case, there are more prospects for a specific service delivery at a certain time than can be served by the provider. A capacity gap has consequences on two levels. On the level of core service, a capacity gap leads to the inability of the respective customers to use the service. When a flight is booked on a certain day and at a certain time, then prospects who would like to use this flight cannot be served. Either this results in lost revenues because the prospect then books the flight with another airline or – given a certain time flexibility of the prospect – the customer reorganises their time schedule with dissatisfaction: as a consequence at the need for schedule reorganisation or at least the effort to do it. On the level of a partial process within the whole core service process, a capacity gap results in waiting times for the customer (e.g. waiting to order drinks in a crowded bar) and possibly also customer dissatisfaction.

The tricky thing for service providers with these gaps is that they are dynamic. For any service, there might not always be a capacity gap or a demand gap. Unfortunately, the dynamics are often unpredictable. Consequently, the *optimal service level* is a central objective of service managers. The service level describes the percentage of capacity utilised compared to the capacity available. Because of the profit relevance of the service level – both from a cost and a revenue perspective – service firms aim at realising a high utilisation of service resources. For example, for the call centre channel it was found that an aggregation of demand (e.g. concentrating demand in a few, large call centres) and the absolute staff number, positively affect service levels. More specifically, service providers aim at managing the dynamics of the gaps between service demand and capacity. The starting point for doing this is the knowledge of the factors which influence these gaps.

Determinants of service capacity management

Based on the general reasons for the capacity problem, like the perishability of capacities and the heterogeneity of service demand, the determinants of service capacity management, the factors which influence the opportunities and decisions of capacity management, can be identified.

Regarding the *demand-related determinants*, it is principally the *demand level* that influences capacity management decisions. Knowing the demand level or at least having some knowledge of the demand level is a basic requirement in order to be able to determine service capacities. As service demand is uncertain in services, providers define the optimal service capacity satisfying the service demand to be lower than the maximum capacity. This is because, if a service location is too busy, service quality is perceived as being poor, but if it is too slow, often the atmosphere suffers. For the most part, service demand is not constant over time. In contrast there can be strong *demand dynamics* due to seasonal, daytime-related or weekday related differences between demand levels. There are various reasons for these variations in time, such as habits and culture (e.g. lunch at noon as a determinant for restaurants) or external influencing factors: patients contracting more colds in winter, as a determinant for doctors. When there are specific reasons for the demand dynamics, so-called *demand patterns* are identifiable and are utilised by capacity management. However, there still remains a relatively distinct degree of *accidental demand dynamics*.

Options for managing service capacity

As the capacity problem in service production and delivery is determined by the two factors, demand and capacity, these represent the levers for managing service capacities systematically by managing demand and/or capacity. Managing demand means applying measures in order to influence the customers' perceptions and behaviors, while managing capacity concerns changes in the amount of service resources available. A further fundamental differentiation of capacity management activities is the *point in time* when the measure is applied compared to the time of the service usage. Following the concept of the interaction process, capacity management activities can be applied in two of the three phases: before the service interaction starts (*pre-process*; e.g. reservation system) and during or at the beginning of the service interaction (*in-process*; e.g. providing a waiting room). Crossing the time dimension with the object of the measures (i.e. demand versus capacity) results in four *types of service capacity management activities*:

- determining the capacity level;
- short-term capacity adjustments;
- demand adjustments;
- waiting time management.

Determining the capacity level

A service provider's long-term capacity management defines the quantity and the quality of service capacities to be employed by the service firm. This decision concerns the number of employees employed, the dimension of the service facilities, etc. and concerns especially capacities which can only slowly be modified, extended or downsized, such as hotel building, or the general level of the workforce. The determination of the long-term capacity is affected by the following influence factors:

- expected demand level;
- expected demand dynamics.

The *expected demand level* plays a role in determining the long-term capacity because the capacities are established in order to satisfy the demand. The relevant demand level is composed of the number of customers as well as the expected service usage of these customers (e.g. number of flights per year, number of nights in a hotel, level of utility usage).

Furthermore, the *expected number of customers* is affected by the customer segments targeted by the service provider and their own expected market share. Consequently, the market structure affects the ability to assess the expected number of customers. In a monopoly (e.g. state utility provider), the demand level is much easier to determine than in a highly competitive market with a lot of customer churn (e.g. mobile communications, airlines). Furthermore, the expected service usage is a component of the expected demand level. Independent of the number of customers, the expected demand varies with the amount of services used by a specific customer. For example, telephone companies target two customer segments, private and corporate clients. In general, a telephone provider serves more private clients than corporate ones in number but the single corporate customer generates more revenues than the single private customer. Overall, often, it is easier and more valid to forecast the expected demand by many small customers compared to forecasting the demand by a few big customers.

A further determinant of the long-term capacity is the *expected demand dynamics*. For most services, demand is heterogeneous and varies over time according to certain patterns. Ski lifts are used frequently in the winter time, flower demand increases on days like Valentine's day and Mother's day, many entertainment and leisure services are used more frequently at weekends, and generally, most services are used more frequently in the daytime compared to night-time. Because there are many reasons influencing demand dynamics which are known by service providers from experience or logic, demand dynamics can be forecast to a certain extent. Although there still remains a more or less dynamic demand which can not be foreseen, service providers have some indicators of the probable demand patterns.

Conclusion

Capacity problems and consequently service capacity management are determined by the demand level, demand dynamics and the adaptability of service resources. Even if a service provider estimates the probable demand and demand dynamics, it is unrealistic in most service industries to expect that the exact demand and demand dynamics can be foreseen. A bank cannot forecast exactly how many clients will visit a branch at a specific time on a specific day for every single day. And even if the bank knew the customer behavior exactly, it cannot hire new employees for hours with a lot of client traffic and release them in hours with less traffic.

Therefore, in many service industries, providers realise short-term capacity adjustments. By installing capacities in a flexible way, the generally given capacities can be adjusted flexibly to short-term, predictable or accidental demand dynamics. A flexible capacity utilisation is principally concerned with the assignment of employees, but there are also other aspects of how tangible or technological resources can be utilised flexibly.

For *predictable short-term demand patterns*, service providers employ approaches of short-term *workforce scheduling*. By considering the expected appearance of customers in a service outfit, the tasks of service employees are planned according to the demand pattern.

References

- [1] Adenso-Díaz, B., González-Torre, P. and Garcia, V. (2002) 'A capacity management model in service industries' *International Journal of Service Industry Management* **13** (3/4) pp. 286–302.
- [2] Hur, D., Mabert, V.A. and Bretthauer, K. (2004) 'Real-time schedule adjustment decisions: A case study' *Omega – The International Journal of Management Science* **32** (5) pp. 333–44.
- [3] Klassen, K.J. and Rohleder, T.R. (2001) 'Combining operations and marketing to manage capacity and demand in services' *Service Industries Journal* **21** (2) pp. 1–30.
- [4] Klassen, K.J. and Rohleder, T.R. (2002) 'Demand and capacity management decisions in services: How they impact on one another' *International Journal of Operations and Production Management* **22** (5) pp. 527–48.
- [5] Lovelock, C. and Wirtz, J. (2004) *Services Marketing: People, Technology, Strategy*, Upper Saddle River, NJ: Pearson/Prentice-Hall.
- [6] Liljander, V. (1994) 'Modeling perceived quality using different comparison standards' *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior* **7** pp. 126–42.
- [7] Martin, C.R. and Horne, D.A. (1993) 'Services innovation: Successful versus unsuccessful firms' *International Journal of Service Innovation Management* **4** (1) pp. 49–65.