

Abstract

Capacity decisions in service operations often involve a trade-off between the operating cost and the service level offered to customers. Although the cost of attaining a pre-specified service level has been well-studied, there isn't much research studying how customer service levels affect revenue and profit. This paper conducts an empirical study to analyze how waiting in queue in the context of a retail store affects customer purchasing behavior. Our methodology uses a novel technology based on digital imaging to record periodic information about the queuing system. Our econometric method uses queuing theory combining these data with point-of-sales information to estimate the effect of queues on purchases. We find that waiting in queue has a non-linear impact on purchase incidence and that customers appear to focus mostly on the length of the queue, without adjusting enough for the speed at which the line moves. We also find that customers sensitivity to waiting is heterogeneous and negatively correlated with price sensitivity.

These empirical findings have important implications for service design, staffing and pricing. First, pooling multiple queues into a single queue may increase the length of the queue observed by customers and thereby lead to an increase in lost sales. Second, the model can be used to make staffing decisions, balancing labor costs with revenues.

Third, we show that the relationship between price and waiting sensitivity is an important factor for pricing decisions in a multi-product category when congestion effects are present.