Capacity Pooling in Hospitals: The Hidden Consequences of Off-Service Placement*

Allocating inpatient bed capacity across various services is a common challenge for hospital managers. Typically, hospital managers must allocate a fixed number of beds to each service in advance, yet the number of patients arriving to each service is highly variable. In order to better manage the supply-demand mismatches between bed types and patient types, many hospitals engage in capacity pooling by assigning patients from a service with no available beds to an available bed in a unit designated for a different service. Although physically located elsewhere, these patients are still cared for by a physician of their home service. This is known as “off-service placement”. This is a common occurrence in many hospitals (Best et al., 2015, Dai and Shi, 2017), with approximately 20 to 25 percent of patients being placed off service. Yet, its effects are not well understood.

Capacity pooling in hospitals, manifested in the form of off-service placement, offers some clear benefits. First, incoming patients are placed in a bed and begin receiving care earlier than otherwise possible. This may reduce the in-hospital length of stay (LOS), improve the quality of care, and increase the level of patient satisfaction. Second, it also enables the hospital to increase revenue by utilizing bed capacity that might otherwise go unused. However, there are some drawbacks as well. One downside is that the increased travel distance to a different unit increases the time required from physicians to care for the patient. This erodes physician capacity during an already busy period. Furthermore, off-service placement may decrease quality of care and patient satisfaction due to the fact that the physicians do not regularly interact with nurses on the off-service unit, increasing the coordination effort required to provide patient care (Gittell, 2002, Shi et al., 2014). As such, it is not a priori clear whether the potential benefits of off-service placement outweigh the potential downsides.
In this paper, we study the effects of off-service placement on various patient and operational outcomes. In doing so, we add to the rich operations management literature on healthcare capacity decisions. Green (2002) examines the question of how many inpatient hospital beds are needed for a service, given an upper limit on acceptable waiting times for patients of that type. Best and his colleagues (2015) model the tradeoff between creating pooled wings, which can treat a large number of different patient care types and have the benefit of being able to absorb variability in demand, versus focused wings, which specialize in treating one or two patient types and yield shorter LOS and higher quality of care. Dai and Shi (2017) propose a Markov decision process model of patient flow that incorporates off-service placement, in which they use an approximate dynamic programming algorithm to minimize total costs of waiting for an on-service bed to become available versus the cost of off-service placement. These prior studies demonstrate that many hospitals struggle with capacity management.

For our study, we collaborate closely with a large academic medical center that has, on average, 20 percent of patients that are placed off service. We focus on the hospital’s 473 medical/surgical beds (which are distributed across 8 services and 17 units), as those are the beds that are pooled for off-service placement. For our analyses, we combine multiple sources of patient and operational data to create a dataset that treats a patient encounter as a single observation. We limit our sample to patient encounters during which a patient was either always on-service or always off-service, so that we are able to cleanly identify the effect of being placed off-service. Using a final sample that consists of 25,206 patient encounters, we test a set of hypotheses that predict that, relative to patients placed in on-service beds, those placed in off-service beds will have a longer hospital LOS, experience lower quality of care, and report lower levels of patient satisfaction.
A key estimation challenge is that patients who are placed off service may be different from those who are placed on service, in ways that are observable to the admission controller (who determines bed placement) but are unobservable to us in the data. In order to address this endogeneity concern, we use a two-stage least squares approach with instrumental variables to causally identify the effect of off-service placement. Specifically, we employ two instrumental variables: service busyness (i.e., whether the service exceeded 99 percent utilization in the hour preceding admission) and the service-to-hospital busyness ratio (i.e., the ratio of the service’s utilization relative to the hospital’s utilization in the hour preceding admission). Using this approach and controlling for several patient and operational characteristics, we find that off-service placement is associated with a 25.6% increase in LOS (defined as the time spent in the hospital following the first transfer into a medical/surgical bed) ($p < 0.01$). These patients are discharged slightly later in the day and experience slightly more trigger activations ($p < 0.05$ for each). They also experience an increased likelihood of hospital readmission ($p < 0.01$) and in-hospital mortality ($p < 0.05$). However, they do not experience significantly different levels of overall patient satisfaction.

Our findings highlight the importance of better understanding and properly accounting for the effects of off-service placement. This work has important implications for practice, as quantifying the effects of off-service placement on patient and operational outcomes enables hospitals to make better informed decisions around capacity allocation. We also contribute to the theoretical literature by providing a data-driven estimation of the underlying costs of off-service placement, which can be used for more accurately modeling patient flow.

* References available upon request