EMPIRICAL INVESTIGATIONS INTO THE CAUSAL IMPACT OF HEALTHCARE PROVIDER BEHAVIOR ON PATIENT CARE

Lesley Meng

A DISSERTATION

in

Operations, Information and Decisions

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

Christian Terwiesch
Andrew M. Heller Professor; Professor of Operations, Information and Decisions

Graduate Group Chairperson

Catherine Schrand, Celia Z. Moh Professor; Professor of Accounting

Dissertation Committee

Hummy Song, Assistant Professor of Operations, Information and Decisions
Serguei Netessine, Professor of Operations, Information and Decisions

EMPIRICAL INVESTIGATIONS INTO THE CAUSAL IMPACT OF HEALTHCARE PROVIDER BEHAVIOR ON PATIENT CARE

Dedicated to Lloyd, Mom, and Dad

ACKNOWLEDGEMENT

First and foremost, I'd like to thank my PhD advisor, Professor Christian Terwiesch. Through the countless research meetings we've had over the past five years, I have learned a tremendous amount and grown substantially as a researcher. Thank you, Christian, for encouraging me to take the challenging courses, for inspiring me to tackle the important questions, and for all of the time and effort you invested in guiding and mentoring my work. Words cannot quite capture how grateful and fortunate I am for having been your student; thank you so much for your unwavering support and encouragement throughout my PhD.

I'd like to thank the rest of my dissertation committee, without which this would not have been possible: Professor Hummy Song and Professor Serguei Netessine. Professor Hummy Song was an early role model and mentor, and our conversations throughout the past three years have kept me motivated, focused, and inspired. Beyond being a supportive and wonderful PhD Coordinator, Professor Serguei Netessine's often-open office door meant that during the job market I received constant reassuring guidance to what I thought were 'pressing' job market questions. Though he was not on my committee, I'd also like to thank Professor Bob Batt, a collaborator on the first chapter of my dissertation since the first year of my PhD. Professor Bob Batt was not only an engaged and insightful co-author that taught me a tremendous amount about the publication process, but he was also a supportive and patient mentor who took the time to provide guidance and morale during the academic job market. Thank you, Bob, Hummy, and Serguei!

My warm thanks go to the other OID faculty that helped shape my knowledge and chatted about my research with me throughout the years: Professor Gad Allon, Professor Hamsa Bastani, Professor Gérard Cachon, Professor Morris Cohen, Professor Santiago Gallino, Professor Noah Gans, Professor Etan Green, Professor Simone Marinesi, Professor Katy Milkman, Professor Ken Moon, Professor Maria Rieders, Professor Sergei Savin, and Professor Senthil Veeraraghavan. My thanks also go out to two other Wharton faculty who inspired my learning in econometrics and took the time to discuss my research with me: Professor Paul Shaman in the Statistics Department and Professor Christophe Van den Bulte in the Marketing Department. On the clinical front, I'd like to thank Doctor Kit Delgado, Doctor Krzysztof Laudanski, and Ms. Ann Huffenberger, who helped open doors at the hospital units we were studying so I could spend time watching how nurses and physicians worked. This list would not be complete without mentioning the OID Administrative and Computing Staff. It is thanks to you all that our research and teaching activities are seamless, so many many thanks for all that you do!

I am lucky to have also forged several lifelong friendships during my time at Penn, and without these wonderful people the PhD experience would not have been the same. Baye, Edward, Gemma, Lucy, and Özge, thank you so much for your unwavering friendship and support throughout the years - I could not imagine going through this without you! Aman, Celia, Chloe, Dawson, Jiding, Park, Rowena, Tan, and Vanitha - thank you for forming the rest of my OID PhD student family, it is all of you that made coming to work each day a joy, and I look forward to staying in touch as we continue to grow as academics!

Of course, none of this would have been possible without my family. Mom and Dad, thank you for teaching me to chase my own goals, and for always being my loudest cheerleaders. Your love and support throughout the years have made all the difference. Thank you to my grandparents, for not only showing me what a life of learning looks like, but also the importance of being good and doing good within the world we live in. To my little brother Lloyd, thank you for your undying support of my continuous pursuit of education - it would not have been the same without your humor and encouragement. I promise that this will be the last graduation ceremony of mine you'll have to sit through.

Lastly, a heartfelt thanks to Min, for challenging me when I needed a push, for supporting me when things got tough, and for reminding me to believe in myself when it slipped my mind. You inspire me each day with your thirst for knowledge, and I look forward to always pushing each other to be the best researchers we can be.

ABSTRACT

EMPIRICAL INVESTIGATIONS INTO THE CAUSAL IMPACT OF HEALTHCARE PROVIDER BEHAVIOR ON PATIENT CARE

Lesley Meng

Christian Terwiesch

This dissertation in operations management focuses on the study of healthcare operations management using large-scale empirical datasets and econometric methods. In chapter one, we utilize infrared location tracking data to study the impact of physical facility layout on how service workers organize their tasks. We focus on the hospital emergency department as a service setting where nurses (servers) have discretion over how they interact with their patients (customers) in a facility that introduces significant heterogeneity in necessary walking distance. Our findings show that even in services, the spatial organization of a facility can lead to servers with discretion over task timing using that discretion in ways that help the server but that lead to reduced customer quality. In chapter two, we examine the hospital intensive care unit (ICU) to investigate the impact of exogenous medication delays, introduced by shift changes, on granular patient health outcomes. The ICU is an ideal setting for this research because patients are often in critical condition and require medications to remain in healthy states (as measured by vital signs). Using patient vital sign data electronically archived every few minutes, merged with the electronic medical record and the medication order/delivery database, we are able to estimate the marginal impact of a minute of medication delay on patient vital status following the late medication. Beyond providing actionable, data-driven insight to managers and healthcare practitioners surrounding how we can better enable workers to maximize effectiveness and efficiency, the research in this dissertation utilizes novel large-scale datasets, unique econometric techniques, and innovative measurement of health outcomes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Introduction

In my dissertation, I use novel large-scale datasets to understand how healthcare providers can make operational improvements to achieve better outcomes. In chapter one, I study the impact of spatial organization-level factors such as facility layout on individual healthcare provider behavior in the emergency department (ED). In chapter two, I focus on the patient health implications of delayed medications due to temporal organization-level factors such as shift changes. More generally, I am interested in uncovering the (often hidden) causal impact of management decisions at the organization level on healthcare worker behavior, and subsequently, its impact on the effectiveness and efficiency of patient care. My research interest in empirical healthcare operations management combines my past education in medical science and public health with my doctoral training in operations management, econometrics, and causal inference.

The data that has enabled my research comes from the electronic collection and storage of increasingly larger and more granular types of information. Examples of the terabyte-scale datasets used in my work include infrared nurse location tracking data in the emergency department, and patient vital sign data electronically archived every few minutes for the duration of a patient's stay in the intensive care unit (ICU). Combining detailed datasets like these with in-person interviews, nurse shadowing, and rigorous econometric methods, I am able to investigate and identify more details surrounding provider behavior and patient outcomes than previously possible using only the electronic medical record (EMR).

In chapter one of my dissertation, **"An Econometric Analysis of How Facility Layout Impacts Care Provision in the Emergency Department"**, we utilize infrared location tracking data to study the impact of physical facility layout on how service workers organize their tasks. We focus on the hospital emergency department as a service setting where nurses (servers) have discretion over how they interact with their patients (customers) in a facility that introduces significant heterogeneity in necessary walking distance. We combine the tracking data with patient EMR data, bedside call data, and the architectural floor plan, to show that nurses reduce their total walking distance by decreasing the frequency of visits to patient rooms far away. We find that this behavior is consistent with batching their tasks for these patients in rooms farther away rather than reducing their tasks. While this behavior decreases necessary nurse walking, it comes at the expense of diminished care quality. We find that patients in rooms farther away press the call button more frequently, an action that is linked with poor patient satisfaction. Our findings show that even in services, the spatial organization of a facility can lead to servers with discretion over task timing using that discretion in ways that help the server but that lead to reduced customer quality. Examples of other service settings where our results would generalize include restaurants, airlines, maintenance services, and surveillance security services, among others.

In chapter two, **"The Impact of Medication Delays on Patient Health in the ICU: Estimating Marginal Effects Under Endogenous Delays"**, we examine the ICU to investigate the impact of exogenous medication delays, introduced by shift changes, on granular patient health outcomes. The ICU is an ideal setting for this research because patients are often in critical condition and require medications to remain in healthy states (as measured by vital signs). Using patient vital sign data electronically archived every few minutes, merged with the electronic medical record and the medication order/delivery database, we are able to estimate the marginal impact of a minute of medication delay on patient vital status following the late medication. We use temporal organization-level factors such as shift changes, physician rounding, and care coordination activities to identify the exogenous portion of delays to find that delaying certain groups of medications result in an increased probability that the patient will enter an unhealthy state following the delay. The interesting, and practically relevant, finding is that the magnitudes of these effects vary significantly by medication type, which allows us to generate a priority list of medications that could help providers focus their resources during busy times (such as shift changes and physician rounding). This work generalizes to services where delay can lead to an undesirable outcome and prioritization of attention is important due to limited resources;

examples of this include emergency services such as firetrucks, ambulances, and project management, among others.

My goal in conducting this research is to provide actionable, data-driven insight to managers and healthcare organizations surrounding how we can better enable workers to maximize effectiveness and efficiency in light of organizational constraints. I am excited at the possibility of expanding my future contributions to other clinical settings, different econometric methods, and possibly even other areas of personal interest, such as environmental sustainability.

CHAPTER 2 : The Impact of Facility Layout on Service Worker Behavior: An Empirical Study of Nurses in the Emergency Department

Joint work with Robert J. Batt and Christian Terwiesch

> "In the ordinary hospital the nurses make many useless steps. More of their time is spent in walking than in caring for the patient. This hospital is designed to save steps. Each floor is complete in itself, and just as in the factories we have tried to eliminate the necessity for waste motion, so we have also tried to eliminate waste motion in the hospital." -*(Ford et al., 1922), Henry Ford describing his vision for the hospital of the future*

2.1. Introduction

The physical layout of an operation can have a substantial impact on process flow and quality. The Lean Operations movement identified *motion* as one of the seven sources of waste, and thus improving facility layout to reduce the cost of walking and wasted motion is a key element of the Lean improvements (Ohno, 1988). While such effects have been well studied in manufacturing environments (e.g., Vollmann and Buffa, 1966; Rosenblatt, 1986; Benjaafar, 2002) and recently have also been studied in the field of warehouse facility design (Roodbergen, 2001; Heragu et al., 2005; Tompkins et al., 2010) and pick-worker performance (Batt and Gallino, 2017), the effects of physical layout on the productivity and quality of services has received much less attention. This is surprising because in many service settings, especially those in knowledge intensive domains such as healthcare, engineering, banking, or education, the work to be executed by an employee is subject to a substantial amount of discretion, which could be used by the employee to reduce motion waste. For example, it seems unlikely that a warehouse employee picking items for an online order will decide to leave an order unfulfilled when fulfilling it might be associated with a substantial amount of walking. In the domain of research and engineering, however, it has long been documented through self-reported surveys on communication patterns that the architectural design of the office building can prevent cross-functional communication from happening (Allen, 1970), thus preventing two employees from communicating with each other because their offices are located far away from each other. The aim of this article is thus to analyze the effects of the physical facility on the productivity and quality of work provided by skilled service workers.

To study this aim, we choose the hospital emergency department (ED) as the empirical domain for our research. In this environment, customers (patients) remain largely in one location, and servers (nurses, doctors, etc.) travel to the customer repeatedly to provide a face-to-face service. Though one might believe that the care provided by the nurses and doctors is entirely determined by the medical needs of the patients, hospital facilities have a physical layout in which the care for some patients requires more walking and movement by care providers compared to other patients. Many EDs are constructed with a central work area for care providers (commonly known as the nurses' station), which serves as a hub for most work, much of which is digital and can thus be done away from the bedside. The distance from this hub to the point of care can vary substantially across patients and hence might lead to some variation in care. In particular, since walking is "costly" in terms of time and energy for the provider, it is plausible that care providers attempt to reduce the total distance walked by factoring such movements into their decision making process of which patient to see at what time. Such discretion about which patient to see at what time is an important difference between skilled service workers and plant and warehouse operations.

This article is based on a dataset that captures how nurses move through the emergency department of a large hospital. For a period of five months, our dataset combines patient and encounter level data (extracted from the electronic medical record system), detailed data of nurse movements throughout the ED (each nurse was equipped with a wireless device that submitted location data to a central database every 6 seconds), patient call-button activation and response timestamps, and measurements from the architectural floor plan of the ED. A large-scale dataset as granular and objective as this one uniquely enables us to compute the nurse-station-to-patient-room distance (or simply "distance") for each patient and to analyze how this distance impacts clinical variables such as the length of stay (LOS), and how often patients press the nurse call button, both of which have been shown to be important determinants of the quality of the patient experience (Pines et al., 2008; Tzeng et al., 2012). This allows us to make the following contributions.

1. We find that nurses in our study setting walk 4.9 miles per shift, on average, which occupies 85 minutes per day, corresponding to 12% of their total work time. This is consistent with prior literature documenting how far nurses walk. We extend this work to show how nurses make adaptations to their work behavior to mitigate some of the fatigue associated with all this walking distance.

2. We find that nurses make fewer visits to patient rooms that are farther away from the nurses' station, but the average visit duration is longer, suggesting a batching of tasks. However, the total time a nurse spends with the patient each hour remains unchanged with room distance. The net effect of this behavior on the nurse is a reduction in the total distance walked per hour.

3. We test for indirect effects of distance on the patient experience and find that patients in more distant rooms observe longer waits between nurse visits and press the nurse call button more often, an action that differs in frequency by patient race and has previously been shown to be associated with decreased patient satisfaction. While some call button activations lead to a physical nurse visit, many do not. Further, we show that nurses are less likely to physically visit a patient in response to a call if the patient is in a distant room.

4. We do not find an association between room distance and patient LOS.

Together, these findings show that skilled service workers, just like their counterparts in plant or warehouse operations, spend a sizable portion of their work time in transit. However, there exists a second, more subtle effect of distance in the case of skilled service workers. In the absence of clear work instructions, skilled service workers have substantial discretion over how they organize their work, which they use to reduce walking distance by batching tasks. Interestingly, we find that nurses working in a busy ED are able to provide an equal amount of care to patients in rooms of varying distances away while batching their tasks. However, this results in longer wait times between nurse visits for patients in rooms farther away, which results in increased nurse call button activations coming from these rooms. Patients may be using the call button as a means to minimize the time they have to spend waiting, however this strategy is not entirely effective because nurses are less likely to react to calls coming from distant rooms with a physical visit. Thus, patients in rooms farther away are not only made to wait longer between visits, but their calls for attention are more likely to be ignored. Short of redesigning the facility layout to mitigate these effects, management can assist workers in minimizing the negative effects associated with task batching by creating an infrastructure for additional operational transparency to the customer, a strategy that has been shown to improve customer satisfaction with services where there exist delays (Buell et al., 2016).

2.2. Literature Review

Our work builds upon three main streams of literature: empirical studies of worker behavior in service operations, the medical literature studying healthcare provider behavior, and facility layout studies in the architecture domain.

Many of the prior empirical studies of worker behavior in the Operations Management literature have relied on customer-centered data. This is at least partly because modern record-keeping systems, such as electronic medical record (EMR) systems in hospitals and point-of-sale systems in restaurants, capture granular time-stamped data on customer encounters. Using such data, prior work has analyzed such topics as the impact of workload on productivity (KC and Terwiesch, 2009; Tan and Netessine, 2014) and worker discretion (Jaeker and Tucker, 2017; Batt and Terwiesch, 2016; Freeman et al., 2017), the impact of interruptions (Cai et al., 2017), delays (Chan et al., 2016), and performance feedback (Song et al., 2017) on productivity, and the impact of multitasking on performance (KC, 2013).

In contrast, much of the literature studying healthcare provider behavior directly using provider-centered data has, until recently, relied on observational or survey data. For example, a study using direct observation of ED nurses found that on average nurses spend 32% of their time on direct patient care, 47% on indirect patient care, and 21% on non-patient care activities (Hollingsworth et al., 1998). Similarly, Tucker (2004) observed nurses across multiple hospitals to describe the impact of operational failure on their work. Studies on hospital facility layout in the healthcare architecture literature have utilized provider survey data to compare the impact of nurse-patient visibility (Bosch et al., 2016) and ward layout (Hua et al., 2012) on patient satisfaction and length of stay (Soriano-Meier et al., 2011).

The recent introduction of pedometer, radiofrequency identification (RFID) and infrared (IR) tracking of individuals has allowed for the large-scale collection of a much more granular and objective source of data to describe provider behavior. For example, Welton et al. (2006) use pedometer data to show that nurses walk an average of 4.1 miles during a typical 12-hour inpatient unit shift, which translates to approximately 70 minutes of time that is not available for care or rest. Hendrich et al. (2009); Choudhary et al. (2010) and Fahey et al. (2013) show using RFID tracking data that the layout of a hospital inpatient medical unit significantly influences nurse walking patterns, and Staats et al. (2017) uses this type of data to examine nurse adherence to process compliance such as hand-washing. Digital location tracking data has also been used in the marketing literature, where by attaching location tracking devices to shopping carts, researchers have been able to study the shopping path behavior of grocery store customers (Hui et al., 2009, 2013). Collecting datasets as granular and objective as these using observation or survey methods would have required an insurmountable number of labor hours.

We contribute to this work along a methodological as well as a theoretical dimension. On the methodological dimension, we compile a unique and novel dataset that combines provider tracking data, patient EMR data, and patient bedside call data. Compared to the observational or survey based work done before, such data is objective and available at a very large scale. Compared to the cited studies of hospital provider behavior using only location tracking data (such as in Hendrich et al. (2009); Choudhary et al. (2010); Fahey et al. (2013)), we expand the research scope beyond descriptive analyses of provider behavior by linking this type of location tracking data to the patient electronic medical record and patient bedside call data to cleanly study the impact of facility layout on provider behavior, and subsequently, patient care quality. For example, we track 217 nurses, 29,430 patients, and a total of 4,150,000 recordings of a provider location at a particular time; this enables us to perform a very different type of econometric analysis.

On the theoretical dimension, we propose a framework that disentangles two effects of distance on work behavior. Prior work has entirely focused on the direct effect of distance, demonstrating that distance leads to workers spending a large portion of their time in transit leading to reduced productivity. This effect, as we discussed above, is common across manufacturing plants and hospitals alike. However, we also identify an indirect effect of distance in settings where workers have discretion surrounding how they organize their work. We propose that service workers use this discretion to reduce the time they spend walking. In particular, we show that they batch tasks, which trades off how long they have to spend walking with the service requirement of the customers they are serving. Patients in rooms of varying distances from the nurses' station all receive the same amount of nurse time per hour, but patients in rooms far away end up experiencing a longer wait time between visits. This comes at the expense of customer impatience, and a lower perceived service quality.

2.3. Empirical Setting & Data

Our data come from a large, urban, academic medical center with an ED patient volume of over 72,000 patients per year. This ED contains 33 patient rooms, 7 hallway beds, a trauma bay for high acuity patients, and a fast-track area for low acuity patients. Figure 1 shows the floor plan of the 33 main patient rooms. The rooms are labeled and the nurses' stations

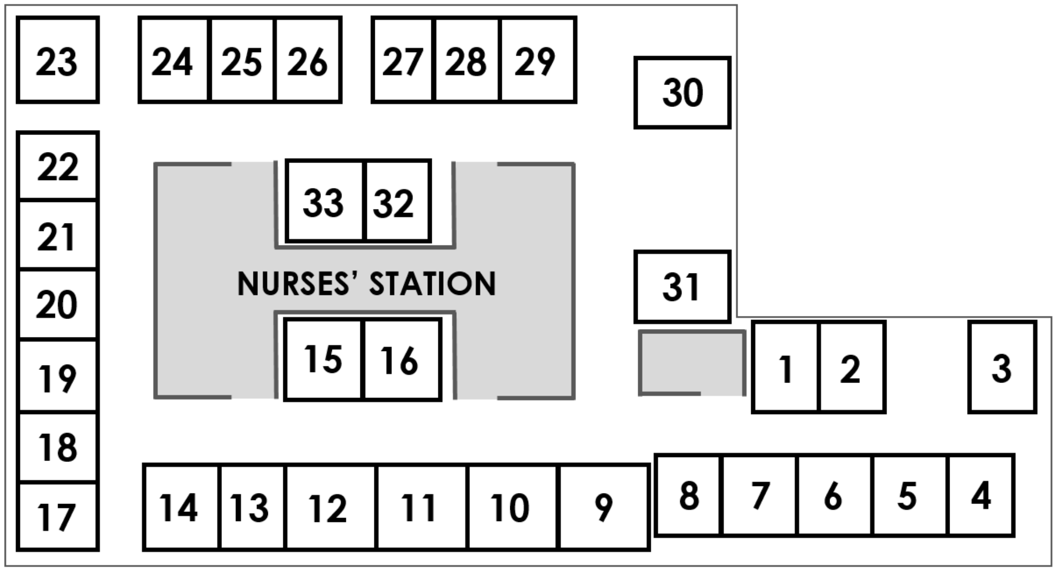


Figure 1: Layout of the 33 Core ED Patient Rooms; shaded region indicates nurses' station.

are shaded in grey. To reduce patient heterogeneity and to ensure accurate location data, we focus only on the patients placed in the 33 main patient rooms, however we include all patients in the relevant census variables.

The patient experience follows a process typical to many EDs in the United States (e.g., Song et al., 2017; Batt and Terwiesch, 2016). Upon arrival to the ED, patients undergo a triage process where they are assigned an Emergency Severity Index (ESI) value ranging from 1 (most severe) to 5 (least severe) (Gilboy et al., 2005). After triage, patients wait in the waiting room to be assigned a treatment bed. Bed assignment is generally made on a first-come-first-served basis by decreasing severity level. The assignment of nurse to patient is based on the treatment room to which the patient is assigned. Each treatment room belongs to a fixed cluster of three or four adjacent rooms, and the ED nurses sign up first-come-first-served for a cluster at the beginning of each shift. Nurses have primary responsibility for any patient placed in the rooms for which they signed up. There are three nursing shifts at this ED, which run from 7am through 7pm, 3pm through 3am, and 11pm through 11am.

Once in a treatment bed, treatment occurs over multiple interactions with care providers who come to the patient's room. Once ED treatment is complete, patients are either discharged home or moved to a bed in an inpatient unit in the hospital. Following (Song et al., 2015b), we refer to the time from bed assignment to being ready to leave the ED as the "length of stay" (LOS).

While most care provider interactions are initiated by the provider, the patient has the ability to summon a nurse by pressing a call button located on a bedside controller. Pressing the call button both illuminates a call light outside the room and initiates a phone call to a clerk at the nurses' station. The clerk speaks with the patient and then either assigns the call to a nurse (generally, the nurse with primary responsibility for that room and patient) to visit the patient, or handles the call themselves from the nurses' station. Sometimes a nurse sees the call light outside the patient room and visits the patient without prompting from the clerk. The call is recorded as "resolved" when either a nurse physically visits the patient in the room and turns off the call alert, or the patient's concern or request is addressed by the clerk over the phone without a physical visit and the clerk turns off the call alert. Approximately 62% of calls result in an immediate physical visit by the nurse.

We bring together four sources of data from our study hospital to create our analysis dataset: patient-level EMR data, nurse real-time location tracking data, call button data, and measurements from the architectural floorplan of the ED. The data spans five months in 2013, during which time the ED saw 29,430 patients. After removing the critically ill patients seen in the trauma bay and the more stable patients seen in the fast-track area, we have sample of 15,595 patients seen in the main treatment rooms throughout our study period.

In this ED, patient-to-bed assignment occurs based on (1) the severity of the patient, and (2) load balancing patients across working nurses by the "charge nurse". To avoid potential endogeneity concerns resulting from room assignment based on variables observed by the charge nurse but not by us researchers, we reduce the dataset to patient room assignments that occurred when there was only one free room in the ED. In these scenarios, the charge

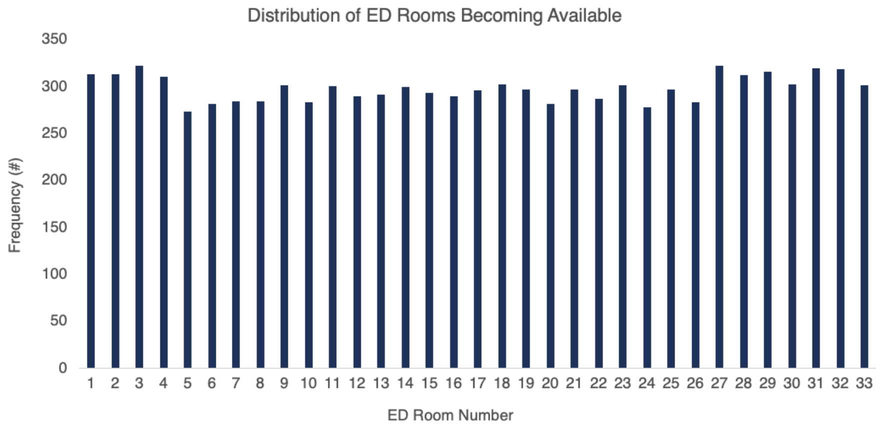


Figure 2: Distribution of ED Rooms Becoming Available

nurse has only one option for where to place the new patient. The ED was in this state for 63% of patient encounters during our study period (9,880 patient encounters of 15,595 total patient encounters); we conduct all of our analyses on this subset of the data. To ensure that the one free room is consistently randomly sampled across the 33 available rooms, we plot the frequency of observations across the rooms to find that they are roughly equal in frequency (Figure 2). We repeat the main analyses on the full dataset in the Appendix and find qualitatively similar results.

Table 1: Summary Statistics of Key ED Patient-Care Variables

Statistic (n = 9,880)	Mean	Median	5^{th} percentile	95^{th} percentile
Patient ESI	2.7	3.0	2.0	4.0
Age	46.9	47.0	20.0	79.0
Admitted (binary)	29.1%			
Female (binary)	59.0%			
Lab Orders (#)	5.2	5.0	0.0	13.0
Wait Room Census (#)	12.2	11.0	2.0	26.0
ED Bed Census (#)*	47.3	47.0	40.0	56.0

**Census includes Fast Track & Trauma*

The nurse location tracking data comes from unique infrared identification tags that are worn by each nurse. This location tracking system was installed as part of a phone and communication system which allows calls for nurses to be routed directly to their location. Every six seconds the tag emits an infrared signal which is picked up by one of the 147 receivers located throughout the ED, including in each patient room. This generates a location record in the database. Our data contain over four million such location records created by 217 unique nurses over the five months. Combining this data with the architectural floor plan of the ED, we can identify the distance traveled by nurses each shift as well as the frequency and duration of each visit to a patient room. Summary statistics of our data show that for each hour of work in the ED, nurses walk an average of 669 meters, with a median of 520 meters, and a standard deviation of 554 meters. This translates to nurses walking 4.9 miles per shift, on average, which occupies 85 minutes per day, corresponding to 12% of their total work time. These statistics are consistent with published metrics of nurse walking distance; Welton et al. (2006) use pedometers to find that nurses walk around 4.1 miles during an equally long shift in an inpatient unit.

Besides distance, the 33 patient rooms in the study ED vary over two other characteristics: physical room size and physician staffing. 27 of the patient rooms are 'normal' sized, while four rooms are larger and two rooms are smaller. Despite the difference in square-footage, all the rooms are identically equipped and are capable of serving any ED patient. Regarding staffing, the ED rooms are split almost evenly into three different levels of physician staffing. 'Normal' physician staffing involves one attending and one resident. 'Low' physician staffing involves one resident and one physician assistant, and 'high' physician staffing involves one attending and two residents. The reason for these differences in staffing is to allow for matching of the severity of patients to the physician staffing level by way of room assignment. However, due to high levels of patient load in the ED, matching is often not possible and physicians help each other out when there is a complex patient in the unit. We control for these room size and physician staffing fixed differences in our model estimation; the distribution of these attributes across room distances are outlined in the Appendix.

To identify the effect of patient room location on nurse work behavior, we code the distance of each patient room from the nurses' station based on actual travel distance; we measure the distance from the closest seated position in the nurses' station to the patient's bed in each of the patient rooms in meters. The farthest patient room is approximately 45 meters away, whereas the closest room is 10 meters away.

2.4. Research Framework & Hypotheses Development

To explore the impact of patient room distance on nurse work behavior and patient care, we separate our hypotheses into two sections. Section 2.4.1 outlines our hypotheses surrounding how room distance impacts nurse work behavior. Section 2.4.2 outlines our hypotheses surrounding how room distance indirectly affects the patient care experience.

2.4.1. Impact of Room Distance on Nurse Behavior

To explore the relationship between patient room distance and nurse visit behavior, we break down nurse-patient visit patterns into its various components and examine the impact of distance on these elements separately.

Nurses perform many types of tasks as part of treating each patient. Some of these tasks, such as note-making, preparing medications, and labeling labs are performed at the nurses' station. Indeed, the nurses' station serves as a "home base", with nurses spending 49% of their time at the station. However, many nursing tasks are performed at the bedside in the treatment room. Much of this work is routine, for example patient interviews, physical examinations, vital sign readings, administering medications, and managing pain and comfort, among others. While these tasks are of varying levels of importance and urgency, many of them are not time-critical and can be delayed somewhat without major adverse effects to the patient (Tucker, 2004). Therefore, nurses have some discretion over when they visit the patient to perform these tasks, and perhaps even over which tasks are performed. Because nurses have discretion over the "what" and "when" of most routine tasks, they might use this discretion to reduce distance traveled.

Prior literature using only observational data to describe nurse work behavior has shown that "the number of entries into patient rooms was negatively correlated with average time per visit [...] The data describe two overall strategies of nurse mobility patterns: fewer, longer visits versus more frequent, shorter visits" (Hendrich et al., 2009). However, the authors do not explore when or why nurses might vary between these two styles of visits. This leads to our first hypothesis.

Hypothesis 1 ***Hypothesis 1****: The number of nurse visits to a patient room per patient hour decreases with room distance.*

The null hypothesis is that distance has no impact on the number of times a nurse visits a patient and that the delivery of care is entirely driven by medical variables. If Hypothesis 1 is supported and nurses are visiting distant rooms less frequently, it is possible that this is because nurses are skipping some tasks for those patients. Oliva and Sterman (2001) refer to this as "cutting corners" and Batt and Terwiesch (2016) refer to it as "task reduction." Another possibility is that nurses are batching tasks together. For example, rather than walking to a distant patient room every time a routine task arises, the nurse may wait until a few tasks need to be done and then make a single trip to the patient to perform multiple tasks. Prior work shows that healthcare providers sometimes batch tasks when they believe it saves them time (Ibanez et al., 2017). Nurses could also batch tasks across nearby rooms rather than within a room, however this is rarely observed in our dataset. From conversations with the nurse team at the study hospital we learn that nurses typically like to return to the nurses' station between visits to patient rooms in order to write patient notes from the previous visit and prepare new medications, labs, etc. for the next patient visit. While we cannot directly observe the tasks performed by the nurse in the patient room, we do observe the amount of time the nurse spends in the room, which we take as a proxy for the number of tasks being performed. If nurses are engaging in task reduction then we would expect the total time nurses spend with patients per hour to be lower for distant rooms. Otherwise, distance should have no effect on the time spent with patients each hour. We test for this in our next hypothesis.

Hypothesis 2 *The total nurse-patient interaction time per patient hour decreases with* *room distance.*

If Hypothesis 2 is supported and the total amount of nurse-patient interaction per hour decreases with an increase in the distance of patient rooms from the nurses' station, this would suggest that nurses are engaging in task reduction for the patients in rooms far away. If Hypothesis 2 is not supported, then all patients receive the same amount of aggregate time with the provider each hour. However, if patients in rooms far away receive this nurse time in fewer visits, the visits made to patient rooms far away must be of longer duration. The pattern of fewer visits but each of longer duration to distant rooms resembles a batching strategy to reduce the amount of necessary walking. We test for this batching behavior by also looking at the average duration of visits to the patient room. If nurses are batching tasks for patients in distant rooms (and not cutting corners) then the average duration of a visit should be longer as the nurse performs more tasks in a single visit.

Hypothesis 3 *Mean nurse visit duration increases with room distance.*

Finding support for both Hypothesis 1 and Hypothesis 3, that is distant rooms receive fewer but longer nurse visits, is evidence of nurses batching tasks and reducing walking.

2.4.2. Impact of Room Distance on the Patient Experience

We next consider the impact on patients of the above hypothesized changes in nursing behavior due to room distance. If nurses batch tasks for patients in distant rooms to save on walking, then the time interval between visits likely grows.

Hypothesis 4 *The mean time between nurse visits to the patient room increases with room distance.*

If Hypothesis 4 is supported, this can have at least two effects on the patient's care experience. First, as the patient is left alone for longer periods of time, the patient is more likely to press the call button to signal for help.

Hypothesis 5 *The number of call button activations per patient hour increases with room* *distance.*

The second possible impact of distance-based task batching is increased LOS. As nurses batch tasks, some tasks happen later than they otherwise would have. If some of these delayed tasks are on the critical path of treatment, then the LOS of patients in distant rooms will be longer than patients in nearby rooms. In addition, increased distance might directly impact LOS simply due to the additional time it takes care providers to walk to a distant room.

Hypothesis 6 *The patient length of stay increases with room distance.*

Prior work has shown that both calls and the response time to calls are drivers of patient dissatisfaction with care (Tzeng et al., 2012). In addition, patient calls are perceived by nurses as being disruptive of work flow (Kalisch and Aebersold, 2010; Gurvich et al., 2017) and cause the delay of other tasks (Cai et al., 2017). Increased LOS has also been shown to be a driver of patient dissatisfaction (Pines et al., 2008; Herring et al., 2009) and also reduces the productivity of the ED. Further, ED LOS is a key hospital performance metric that is reported to the Centers for Medicare & Medicaid Services for public review (Carrier et al., 2014).

2.5. Econometric Specification

To test these hypotheses, we estimate six regression models with the following dependent variables: Visits per Hour, Duration per Hour, Duration per Visit, Time Between Visits, Calls per Hour, and the ED LOS. These dependent variables are constructed at the patient encounter level and we utilize ordinary least squares (OLS) regression to estimate these models. The summary statistics for these variables are shown in Table 2. The independent variables are identical across the six models. We choose to use OLS to allow us to keep a linear relationship between distance and our response variables. To ensure the robustness of our estimates, we repeat the analysis using robust regression (Huber's M-estimator, (Huber et al., 1964; Huber, 2011)) in Section 2.8 and find qualitatively identical results.

Table 2: Summary Statistics of Dependent Variables

Statistic (n = 9,880)	Mean	Median	5^{th} percentile	95^{th} percentile
Visits per Hour	2.54	2.13	0.44	5.95
Duration per Hour (nurse-mins)	5.60	4.13	0.29	11.90
Duration per Visit (mins)	2.18	1.86	0.33	5.07
Time Between Visits (mins)	25.87	19.26	5.30	66.91
Patient Calls Placed (calls/hour)	1.31	1.02	0.00	4.09
LOS (hours)	5.66	4.70	1.67	12.56

Our main estimation equation is:

$$Y_{ij} = \beta_o + \beta_p DIST_{ij} + \mathbf{W}_j\beta_q + \mathbf{X}_{ij}\beta_r + \mathbf{Z}_i\beta_s + \epsilon_{ij}$$

We use index i for each unique patient visit to the ED and j for rooms. Y stands for the relevant dependent variable we are testing. The room-specific variables ($\mathbf{W}$) control for the room size and staffing characteristics described in Section 2.3. The encounter-specific controls ($\mathbf{X}$) include the ESI level of the patient, the patient's age, the patient's gender (male or female), the patient's race (Asian, Black, Hispanic, White, or Unknown), the physician assigned to the encounter, the nurse assigned to the encounter, the number of diagnostic orders for the encounter, the chief complaint of the patient, whether the patient was admitted, the census in the ED wait room, and the bed census in the ED. The values captured by the categorical variables of patient gender and race were based on the dataset; reference categories used in estimation for the patient's gender and race are male and White, respectively. The time-related controls ($\mathbf{Z}$) include the month, the hour of day, and the shift during which the patient arrived. These controls are listed in Table 3. We define patient room distance ($DIST_{ij}$) as the actual walking distance from the closest seated position in the nurses' station to the patient's bedside, measured in units of 10 meters in length.

Table 3: Control Variables for Main Estimation Models

Control Group	Controls
Room-Specific Controls **W**	Physician Staffing Level (High, Regular, Low)
	Room Size (Large, Regular, Small)
Encounter-Specific Controls **X**	Patient ESI
	Patient Age
	Patient Gender
	Patient Race
	Assigned Physician
	Assigned Nurse
	Number of Diagnostic Orders
	Patient's Chief Complaint
	Whether Patient was Admitted
	ED Wait Room Census
	ED Bed Census
Time-Related Controls **Z**	Month
	Hour of Day
	Nurse Shift

We now turn to testing the hypotheses laid out in Section 2.4 by estimating Equation 2.5 with each of the dependent variables as specified.

2.6. Results: The Impact of Room Distance on Nurse Behavior

The results for this estimation are presented in Table 4; all controls are listed in Table 16 and standard errors are clustered at the room level. Model 1 is estimated with the number of nurse visits per hour as the dependent variable, and our results show that rooms more distant from the nurses' station receive fewer nurse visits per patient hour ($\beta_p = -0.390$, $p < 0.001$). This supports Hypothesis 1. We find that every additional 10 meters of distance from the nurses' station leads to 0.39 fewer visits per hour (a 15% reduction from the mean). Recall that the patient rooms farthest away from the nurses' station are approximately 35 meters farther away than the nearest rooms (45m vs. 10m), thus the farthest rooms receive an average of 1.4 fewer visits per hour (35m $\times$ 0.039). This suggests that nurses are likely reducing their walking by decreasing the number of visits they make to distant patient rooms.

Table 4: Results: OLS Regression Estimation on Nurse Visit Models

	(1) Vis/Hr	(2) Dur/Hr	(3) Dur/Vis
Distance (10m)	-0.390^{***} (0.024)	0.034 (0.047)	0.343^{***} (0.020)
Patient Race: Asian	0.151 (0.128)	0.117 (0.248)	-0.007 (0.104)
Patient Race: Black	0.079 (0.043)	-0.043 (0.084)	-0.114^{**} (0.035)
Patient Race: Hispanic	0.120 (0.127)	0.028 (0.247)	-0.098 (0.104)
Female Patient	-0.065 (0.040)	0.109 (0.077)	0.086^{**} (0.032)
Room Size: Large	0.198^{**} (0.073)	0.341^{*} (0.142)	0.072 (0.060)
Room Size: Small	-0.176^{*} (0.082)	0.242 (0.160)	0.139^{*} (0.067)
MD Staffing: High	0.225^{***} (0.059)	0.226^{*} (0.114)	-0.039 (0.048)
MD Staffing: Low	-0.293^{***} (0.051)	0.451^{***} (0.100)	0.463^{***} (0.042)
Diagnostic Lab Orders	0.060^{***} (0.006)	0.162^{***} (0.011)	0.008 (0.005)
Admitted Patient	0.313^{***} (0.051)	0.673^{***} (0.100)	0.009 (0.042)
Shift 2 (3p-3a)	0.500^{***} (0.090)	0.954^{***} (0.175)	-0.027 (0.074)
Shift 3 (11p-11a)	0.873^{***} (0.127)	1.944^{***} (0.246)	-0.020 (0.103)
Constant	1.747^{***} (0.420)	1.754^{*} (0.816)	1.271^{***} (0.343)
Observations	9,880	9,880	9,880
Adjusted R^2	0.241	0.170	0.139

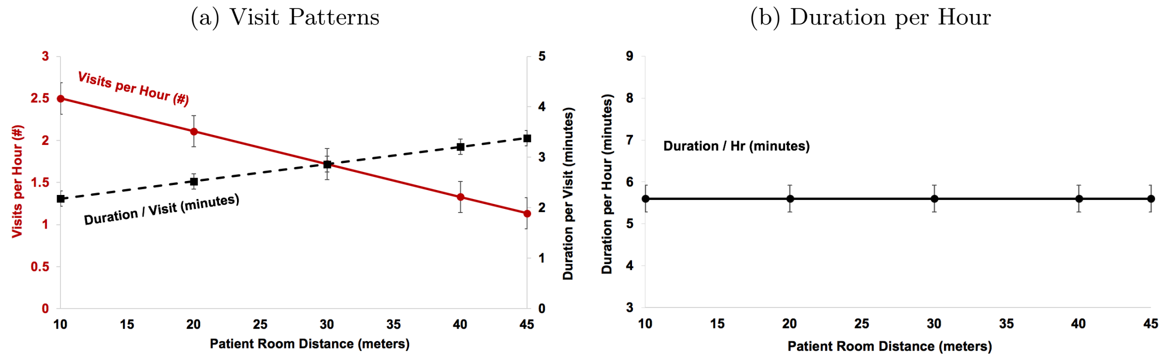
*p<0.05; **p<0.01; ***p<0.001

Given that nurses are reducing the number of visits made to patient rooms farther away, we use Model 2 to determine whether or not task reduction is the reason we observe this visit reduction. We use the total minutes per patient hour spent in the patient room as the dependent variable to estimate the aggregate amount of time the patients had with providers. If task reduction is the reason behind the observed reduction in visit frequency, then the total number of minutes per hour spent with the provider should be reduced for distant patients. Our results show that distance does not have a significant effect on the total time a nurse spends with a patient per patient hour ($\beta_p = 0.034$, $p > 0.05$), suggesting that distance does not affect the aggregate amount of time with nurses. Thus we do not find support for Hypothesis 2.

If all patients receive the same amount of aggregate nurse time and patients in rooms farther away receive fewer visits, then it follows that these patients must also receive longer visits. Model 3 studies the average duration of individual nurse visits. We find that the average duration of nurse visits is longer for more distant rooms ($\beta_p = 0.343$, $p < 0.001$). This supports Hypothesis 3 and to the extent that time in the room correlates with the number of tasks performed, this suggests that nurses are batching tasks for patients in distant rooms. We find that nurses spend a mean of 0.343 minutes (21 seconds) longer per visit for each 10 meters of distance from the nurses' station. This translates to an additional 1.2 minutes spent per visit for the patients in rooms farthest away. As shown in Table 2, the average visit duration over all visits is 2.2 minutes, so this marginal effect due to distance is large, relative to the mean.

To illustrate our results, we calculate the expected nurse visit frequencies across different patient rooms for a typical patient. In our study sample, a typical patient is a 45-year-old female patient of White race visiting the ED during a weekday between 7am and 7pm, presenting with abdominal pain, the most common chief complaint in the ED. We use our estimated coefficients to calculate and plot the predicted values of our main dependent variables given the covariates of our typical patient (Figure 3). As patient room distance

Figure 3: Effects of Distance on Nurse Visit Patterns



Note: Error bars indicate 95% confidence interval

increases, the number of visits the nurses make to the room per hour decreases (Figure 3a, solid line). However, nurses compensate these patients with more time spent each time a visit is made (Figure 3a, dashed line). As a result, the total minutes per hour spent with the patient does not change across room distances (Figure 3b).

2.6.1. Post hoc Analysis: Nurse Walking Distance

As shown in prior work (e.g., Welton et al., 2006; Hendrich et al., 2008), nurses spend a significant portion of their shift duration walking. Presumably, the total distance a nurse walks in a shift is a function of the room distance he or she is working that shift. It is reasonable to assume that nurses working more distant rooms walk greater distances each hour. However, because walking is both time consuming and tiring, nurses have an incentive to reduce the amount of walking they do. If nurses are making decisions to minimize the distance traveled by batching tasks for patients in far away rooms, it is possible that such efforts result in less walking when responsible for patients in rooms farther away. We construct a model to test this.

We specify the following OLS regression model to test this at the nurse-hour level. We use index n for each unique nurse and t for each working hour; all variables are constructed at the nurse-hour level based on the rooms the nurse is responsible for that hour.

$$DistWalked_{nt} = \beta_0 + \beta_1 AvgPtDist_{nt} + \mathbf{RN}_n\beta_k + \mathbf{RM}_{nt}\beta_l + \mathbf{PT}_{nt}\beta_m + \mathbf{TC}_t\beta_n + \epsilon_{nt} \quad (2.1)$$

The main independent variable of interest, $AvgPtDist_{nt}$, is the average of the patient room distances from the nurses' station that the nurse is responsible for each hour. The room-specific controls ($\mathbf{RM}_{nt}$) contain a fixed effect for the physician assigned to these rooms, an indicator for whether or not there was a small or large room in the set, and an indicator for the staffing level of the rooms (these room characteristics were explained in detail in Section 2.3). The patient-specific controls ($\mathbf{PT}_{nt}$) include the average ESI level of the patients the nurse is responsible for, the average age of these patients, the number of diagnostic orders required by these patients, and the number of these patients who are admitted. The census and time-related controls ($\mathbf{TC}_t$) include the month, the hour of day, the shift, the census in the ED wait room, and the bed census in the ED. Lastly, we apply a fixed-effect on the nurse ID to control for heterogeneity across nurses ($\mathbf{RN}_n$).

We find that as the average distance between the nurses' station and a nurses' patients increase by 10 meters, nurses tend to walk on average 31 meters *less* during each hour of work (Table 5). Thus, this finding provides additional support that nurses might be engaging in distance-based task batching to save on walking.

2.7. Results: The Impact of Room Distance on the Patient Experience

Given the evidence in support of distance-based task batching by nurses, we are interested in understanding the impact this has on the patient. Hypothesis 4 focuses on the waiting time experienced by patients between nurse visits, Hypothesis 5 considers call button activations, and Hypothesis 6 considers the patient LOS. Again, we estimate Equation 2.5 using the relevant dependent variables, and the estimation results for these three models are shown in Table 6; all controls are listed in Table 16, and standard errors are clustered at the room level.

Table 5: Nurse Walking Distance per Nurse-Hour (OLS)

	(4) Distance per Nurse-Hour (m)	
Mean Patient Distance	-31.070***	(4.030)
Room Size: Large	212.615***	(9.425)
Room Size: Small	12.935	(11.765)
MD Staffing: High	-249.535***	(8.320)
MD Staffing: Low	-4.420	(7.345)
Total Labs Required	2.600***	(0.195)
Total Admitted Patients	132.535***	(1.625)
Shift 2 (3p-3a)	-7.735	(6.630)
Shift 3 (11p-11a)	-29.510***	(8.515)
Constant	284.310***	(47.190)
Observations	29,526	
Adjusted R^2	0.372	

*p<0.05; **p<0.01; ***p<0.001

Table 6: Results: OLS Regression Estimation on Call Frequency and LOS Models

	(5) Time btw. Visits (Mins)	(6) Calls/Hr	(7) LOS (Hrs)
Distance (10m)	3.058*** (0.331)	0.112*** (0.008)	0.016 (0.044)
Female Patient	0.008 (0.540)	−0.008 (0.013)	−0.019 (0.071)
Patient Race: Asian	−1.713 (1.739)	−0.130** (0.043)	−0.007 (0.231)
Patient Race: Black	−0.030 (0.586)	−0.066*** (0.014)	−0.055 (0.078)
Patient Race: Hispanic	3.804* (1.727)	−0.125** (0.047)	−0.193 (0.230)
Room Size: Large	−1.049 (0.995)	0.067** (0.025)	−0.068 (0.132)
Room Size: Small	−0.540 (1.120)	−0.021 (0.026)	0.077 (0.147)
MD Staffing: High	−2.180** (0.802)	−0.102*** (0.019)	−0.278** (0.105)
MD Staffing: Low	1.918** (0.700)	−0.079*** (0.016)	−0.034 (0.092)
Diagnostic Lab Orders	−0.472*** (0.080)	−0.022*** (0.002)	0.324*** (0.011)
Admitted Patient	0.559 (0.698)	−0.009 (0.015)	1.389*** (0.094)
Shift 2 (3p-3a)	−5.133*** (1.227)	−0.100*** (0.017)	0.726*** (0.114)
Shift 3 (11p-11a)	−7.052*** (1.719)	−0.057* (0.024)	0.662*** (0.152)
Constant	29.807*** (5.712)	0.362*** (0.106)	2.060*** (0.606)
Observations	9,880	9,880	9,880
Adjusted R^2	0.062	0.143	0.336

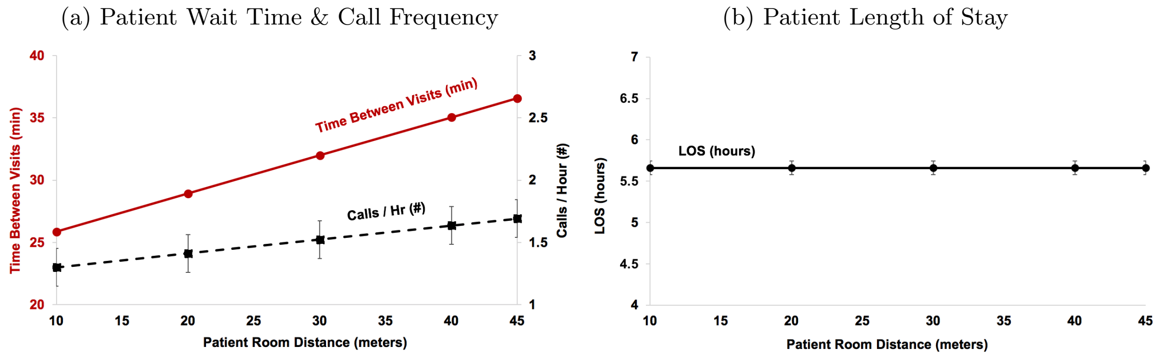
*p<0.05; **p<0.01; ***p<0.001

Model 5 shows that patients in more distant rooms experience longer wait times between visits by the nurse. Specifically, for every additional 10 meters farther away a patient is, the wait time between nurse visits increases by approximately 3 minutes ($\beta_p = 3.058$, $p < 0.001$). For the patients in the rooms farthest away from the nurses' station, this represents an additional wait time of 10.7 minutes between visits. Compared to the empirical average of 26 minutes, the magnitude of this effect is substantial.

Model 6 shows that the mean number of call button activations per hour increases with room distance ($\beta_p = 0.112$, $p < 0.001$). This supports Hypothesis 5 and suggests that patients might be attempting to affect their own care by pressing the call button when they are made to wait longer between visits. We find that the mean number of calls per hour increases by 0.112 for each 10 meters of distance, an 8.5% increase from the mean value of 1.3 calls per hour. With 35 meters between the nearest and farthest rooms, this results in distant rooms generating 0.39 more calls per hour, on average. Interestingly, the rate of call button activations differs by patient race. Under-represented minorities press the call button less frequently compared to White patients ($\beta_{r,Asian} = -0.130$, $p < 0.01$, $\beta_{r,Black} = -0.066$, $p < 0.001$, $\beta_{r,Hispanic} = -0.125$, $p < 0.001$). To the extent that patients are using the call button to influence the care they receive, patients that are less prone to use the call button may receive reduced nurse visits. We examine these effects further in the Appendix, where we repeat the estimation on the interaction of race and gender to obtain more granular group estimates.

The finding that patients in distant rooms generate additional nurse calls is particularly relevant in the ED, where nurses are frequently interrupted by patients (28% of interruptions are due to patient calls) (Kalisch and Aebersold, 2010). Due to the high cost of errors in the ED, nurses are advised to try and 'decrease external interruptions' (Skaugset et al., 2016). Not only are these calls disruptive to nurse work flow (Cai et al., 2017; Gurvich et al., 2017), but they are also associated with decreased patient satisfaction with care (Tzeng et al., 2012). Yet, evidence has previously not linked these interruptions with the physical

Figure 4: Effects of Distance on the Patient



Note: Error bars indicate 95% confidence interval

facility layout of the unit.

Model 7 shows that distance is not significantly related to LOS, and Hypothesis 6 is not supported ($\beta_p = 0.016$, $p > 0.05$). Thus, despite finding that patients in distant rooms receive batched tasks, and longer inter-visit wait times, this behavior does not translate to an increased LOS. Given the importance of the patient length of stay as a hospital performance metric, one possible explanation is that nurses are acutely aware of the tasks that must not be delayed and prioritize those on the critical path.

To illustrate the impact of this behavior on the patient, we calculate the expected frequency of patient call button activations per hour and show that patients make more calls as the distance between their room and the nurses' station increases (Figure 4a, dashed line). This is likely driven by the extended wait time between visits that patients in the distant rooms experience (Figure 4a, solid line). If nurses are making fewer visits, then the patients wait longer between visits, on average. We do not see an impact on the patient length of stay across room distances (Figure 4b), which suggests that tasks on the critical path of ED treatment are usually delivered in a timely manner, regardless of room distance.

2.7.1. Post hoc Analysis: Patient Calls

These findings highlight a hidden side-effect to nurses using their discretion to minimize the cost of motion in their workplace. Despite the equal allocation of total care time across patients of varying room distances, the patients in rooms far away observe longer times between visits by the nurse and press the call button more, perhaps in an attempt to mitigate this additional waiting. We are interested in better understanding two metrics related to the patient call button activations: (1) the extent to which the equal distribution of care time across patient rooms is a result of patients being proactive in engaging with the nurse, and (2) the extent to which patient calls are successful in triggering a physical visit by the nurse.

We first look at the proportion of nurse visits that were a result of a patient call across room distances. To study this question we specify an OLS model using Equation 2.5, and the same controls as listed in Table 16. The dependent variable of interest is $\%VisitsDueToCall_{ij}$, and represents the proportion of nurse visits during a patient's stay that occurred during an 'active call'. We define an active call as the time duration between when a call is placed, and when the call has been resolved. Visits made to the patient during these timeframes are likely to have been the result of the patient call. The second metric is the proportion of patient calls that resulted in a physical visit by the nurse. Similar to the first question, we specify an OLS model using Equation 2.5, and the same controls as listed in Table 16. The dependent variable of interest is $\%CallTriggeredAVisit_{ij}$.

Results are presented in Table 7. Model 8 shows that the marginal effect of 10 meters of distance is a 1.5 percentage point increase in the proportion of visits that come from a call button activation ($\beta_p = 0.015$, $p < 0.001$). The empirical average proportion of patient call-driven visits across all nurse visits is approximately 15%. The patients in the most distant rooms would observe a proportion of call-driven visits that is 5.25 percentage points higher compared to patients in the closest rooms. Thus, patients in distant rooms play a larger role in affecting their own care. However, when we look at the results in Model 9, we

Table 7: Post-Hoc Analysis on Patient Calls

	(8) % Call-Driven Visits	(9) % Visit-Triggering Calls
Distance (10m)	0.015*** (0.003)	−0.041*** (0.003)
Female Patient	0.016*** (0.004)	−0.003 (0.005)
Patient Race: Asian	−0.004 (0.014)	0.052*** (0.015)
Patient Race: Black	0.001 (0.005)	−0.007 (0.005)
Patient Race: Hispanic	−0.006 (0.014)	0.044** (0.017)
Room Size: Large	0.018* (0.008)	0.033*** (0.009)
Room Size: Small	−0.0002 (0.009)	−0.014 (0.010)
MD Staffing: High	−0.023***(0.007)	0.033*** (0.007)
MD Staffing: Low	0.006 (0.006)	0.0004 (0.006)
Diagnostic Lab Orders	0.002* (0.001)	0.004*** (0.001)
Admitted Patient	0.016** (0.006)	0.038*** (0.005)
Constant	0.071 (0.038)	0.537*** (0.039)
Observations	9,880	9,880
Adjusted R^2	0.032	0.148

*p<0.05; **p<0.01; ***p<0.001

see that patient calls are also less likely to result in a physical visit by the nurse when these calls are coming from far away ($\beta_p = -0.041$, $p < 0.001$). Specifically, the proportion of patient calls that result in a physical visit by the nurse decreases by 4.1 percentage points for every 10 meters farther away the room is located. The rooms farthest away observe a reduction in the proportion of calls that trigger physical visits of 14.35% compared to patients in the rooms closest to the nurses' station. On average, 62% of patient calls trigger a physical visit by the nurse.

We know from our findings in Section 2.6 that the total minutes of time nurses spend with patients each hour remains constant with room distance. For the patients in rooms farther away, this care is being delivered in fewer visits, but each of longer duration. Patients are likely getting frustrated with the additional wait time and placing more calls. We find that nurse visits to patient rooms due to patient calls increases with room distance; this suggests that patients are utilizing the nurse call button in an attempt to mediate their own care. However, these calls are not always successful in triggering a visit, and the probability that a call generates a visit is lowest for the most distant rooms. Therefore, despite attempts at rushing the nurse using the call button, the patient ends up receiving the same amount of time with the provider, but with longer inter-visit wait times. It is possible that not responding to these patient calls with a physical visit results in the patient feeling an even greater sense of being neglected, which likely results in poor patient satisfaction with the stay.

2.8. Robustness Analysis

To ensure the robustness of our estimates, we repeat our analysis using robust regression (Ripley, 2002). This method uses Huber's M-estimator and is appropriate for situations where there exist outliers in the response variables of interest (Huber et al., 1964; Huber, 2011). The distribution of our response variables are right-skewed with several outliers. While many of these outliers seem to be representative of normal ED operations, we conduct robust regression on the data to ensure our findings are consistent even when using a much more conservative estimator that is robust to outliers. We choose to use OLS and robust regression to avoid needing to log-transform the data just to satisfy OLS assumptions, therefore preserving the linear relationship between distance and our response variables.

Using this more conservative method, our results are qualitatively identical (see Appendix for results). The number of visits made to patient rooms per hour decreases by 0.315 visits for every additional 10 meters of distance. Similar to our OLS findings, there appears to be no significant difference in the total nurse-patient interaction minutes per hour across room distances. As a patient's room distance increases by 10 meters, the duration spent per visit by a nurse increases by about 0.303 minutes. The rooms farthest away from the nurses' station receive visits that are, on average, 1.06 minutes longer compared to the rooms closest to the nurses' station. Similar to our OLS estimates, we find evidence suggestive of nurses using a batching heuristic when managing patients staying in rooms farthest away from the nurses' station. When using robust regression to understand the impact of this behavior on perceived care quality, we find that patients in rooms every 10 meters farther away experience 2.65 additional minutes of wait time between visits, and make on average 0.040 more calls per hour for the nurse. In the rooms farthest away this translates to 9.28 minutes of additional inter-visit wait time and 0.126 more calls per hour. The length of stay of patients across rooms of varying distances remains unchanged.

2.9. Managerial Implications

In our study, we find that nurses spend 12% of each shift walking, covering a total distance of almost 5 miles. Given the physical demands of this work, it is only human that nurses make an effort to reduce the distance they walk by reducing the number of visits they make to patient rooms far away from the nurses' station, which we show. Interestingly, the data shows that nurses compensate for the reduced number of visits by spending more time with the patient, on average, during the visits they do make, which suggests a batching of tasks. In fact, we find no change in the mean total nurse-patient interaction time.

Despite the nurses successfully managing their work so that patients in rooms of varying distances receive the same total amount of care time, the reduced visit rate leads to increased time between nurse visits, which in turn triggers an increase in call button activations. Call button activations are correlated with decreased patient satisfaction (Tzeng et al., 2012) and are disruptive to work happening at the nurses' station, potentially incurring significant changover costs (Gurvich et al., 2017), and increasing the rate of errors (Cai et al., 2017). We also examine patient LOS and do not find evidence of the reduced visit rate leading to a change in patient LOS.

According to the handbook on patient safety and quality for nurses by the US Agency for Healthcare Research and Quality, "cognitive psychologists have identified the physical environment as having a significant impact on safety and human performance", more specifically, "some of the effects of long work hours and increasing workload can be mitigated by minimizing the distances staff must travel between patient rooms" (Reiling et al., 2008).

Our findings are especially relevant when planning the layout of new facilities. Consider the situation currently facing the study ED. To increase ED capacity, the study ED will soon be moving to a new building that introduces significantly more heterogeneity in patient room distance compared to the status quo. The new ED will be elongated and narrowed to double the length and half the width. Some members of hospital management have expressed concerns surrounding this new layout, even without the support of data. If the new ED is designed with a central nurses' station, similar to the current ED, this new facility layout would double the distance required to walk to the patient rooms farthest from the nurses' station from 45 meters to 90 meters. While we hesitate to extrapolate our estimated results to 90 meters (well beyond the range of our data), it seems likely that such extreme distances would lead to further visit rate reduction, increased inter-visit wait times and increased call button activations.

Of course, investments into new hospital buildings are multi-million dollar projects and changing the layout of the ED is not an option for most. So, another option to counteract the increased waiting times experienced by patients in distant rooms is to improve operational transparency. Showing patients what is taking so long between moments of interaction is known to decrease frustration and increasing ratings of the service, sometimes so much so that the patient actually values the wait time more than receiving an instantaneous service (Buell and Norton, 2011). The experiments conducted by Buell and Norton (2011) translate well into this setting because patients do not observe what nurses are doing when they are outside the patient's room. If nurses provided patients with information between each visit suggesting that the providers are working hard towards completing their care (i.e. "I am going to send these labs in now and will come back once I get your results from the technician", or, "I am going to consult the doctor about what medications to send you home with, I will be back once I get that information"), then patients might place fewer calls, or at the very least feel less frustrated (Maister et al., 1984; Larson, 1987).

2.10. Conclusion

In this study, we provide detailed empirical evidence from a hospital ED on how the physical layout of an operation can influence the work behavior of service workers and subsequently, the resulting service quality perceived by patients (customers). While our results are specific to the setting we study, we believe that these findings have implications for other service settings in which (1) the service is carried out in a physical environment (as opposed to being all digital), (2) there exists heterogeneity in the location of customers (not all customers are equidistant to where servers spend the majority of their time), and (3) workers have discretion over when to complete their work (as opposed to having to adhere to a strict operating protocol) and how much effort to exert.

Several other service settings meet these three conditions. Certainly many other areas of healthcare delivery, such as inpatient wards and post-anesthesia care units ("recovery rooms"), meet the conditions and face similar challenges as the ED. We note that part of what makes the distance-based variation in service quality troubling in the healthcare setting is that the patient generally has no choice over the location of their room and thus some patients experience longer wait times due to "bad luck" in room assignment.

Beyond healthcare, services such as full-service restaurants, many retail stores, and even airplanes, likewise have heterogeneity in distance between customers and server and the servers have discretion over the timing of service provided. Consider full-service restaurants, a setting that has been studied with a productivity focus by Tan and Netessine (2014). Servers spend a lot of time walking through the restaurant and face heterogeneity in walking distances to their customers. Thus, they might engage in distance-based visit reduction similar to the ED nurses in this study. Diners located farther away would thus be made to wait longer by the server (and would need to "press the call button" by flagging down the server), likely decreasing perceived service quality, which could translate into lost revenue or tips. Similar to the healthcare settings, customers often have little choice over the location of their table relative to the server station and thus can experience longer wait times due to bad luck.

In retail settings such as department stores, large furniture stores, and car dealerships, salespeople often have a "home base" (e.g., cash register, front desk, showroom, etc.) where they spend much of their time, periodically venturing out to help customers (Fisher and Raman, 2010; Musalem et al., 2016). If the store layout is such that there is significant heterogeneity in distance of browsing areas, then customers who are browsing in distant locations may experience reduced and delayed sales help, leading to reduced sales and customer satisfaction.

Lastly, one can imagine the same dynamics at work on a large commercial airliner, especially on a long-haul flight. Passengers in seats close to the galleys often are served food and drinks first and can benefit from the frequent passing-by of the flight attendants as the attendants come and go from their seats near the galleys and doors. Passengers in more distant seats may be stuck waiting longer for a drink or to get rid of garbage. And much like the ED setting, passengers with longer inter-visit times may have to resort to pressing the call button to summon assistance leading to interruptions of other tasks and reduced customer satisfaction.

While in the ED, the data show that visit rate reduction is offset by task batching leading to no change in total nurse-patient interaction time, it is not known if such offsetting actions occur in these other settings. This is an area for possible future study. In all these settings, as in the ED of our study, the negative effects of distance-based visit rate reduction can likely be addressed through careful facility design (i.e., reducing distance heterogeneity), and in the cases where service workers have successfully traded off the fatigue associated with facility layout and the service needs of the customer, improved operational transparency.

2.11. Appendix

2.11.1. OLS on All Data: Nurse Visit Models

To compare our findings estimated from the subset of data against the full dataset that includes instances where there were multiple patient beds available when a patient was assigned to a bed, we repeat the OLS regression analysis using all of the data and the same models. The coefficient estimates are qualitatively identical (Table 8). The number of visits made to patient rooms per hour decreases by 0.385 visits for every additional 10 meters of distance. This translates to 1.34 fewer visits for the patients in the rooms farthest away

Table 8: OLS Results Using All Patient Encounters - Nurse Visit Models

	(1) Vis/Hr	(2) Dur/Hr	(3) Dur/Vis
Distance (10m)	-0.385^{***} (0.021)	0.002 (0.040)	0.309^{***} (0.017)
Female Patient	-0.022 (0.034)	0.146^{*} (0.065)	0.076^{**} (0.027)
Patient Race: Asian	0.100 (0.105)	-0.155 (0.202)	-0.084 (0.085)
Patient Race: Black	0.046 (0.037)	-0.069 (0.071)	-0.102^{***} (0.030)
Patient Race: Hispanic	0.093 (0.111)	-0.018 (0.213)	-0.104 (0.089)
Room Size: Large	0.172^{**} (0.061)	0.196 (0.118)	0.003 (0.049)
Room Size: Small	-0.259^{***} (0.071)	0.132 (0.137)	0.179^{**} (0.057)
MD Staffing: High	0.223^{***} (0.050)	0.204^{*} (0.096)	-0.057 (0.040)
MD Staffing: Low	-0.253^{***} (0.044)	0.361^{***} (0.084)	0.389^{***} (0.035)
Diagnostic Lab Orders	0.055^{***} (0.005)	0.154^{***} (0.010)	0.011^{**} (0.004)
Admitted Patient	0.356^{***} (0.045)	0.635^{***} (0.086)	-0.029 (0.036)
Shift 2 (3p-3a)	0.390^{***} (0.081)	0.648^{***} (0.156)	-0.043 (0.065)
Shift 3 (11p-11a)	0.667^{***} (0.104)	1.373^{***} (0.201)	-0.047 (0.084)
Constant	2.011^{***} (0.329)	2.857^{***} (0.632)	1.477^{***} (0.264)
Observations	15,595	15,595	15,595
Adjusted R^2	0.219	0.150	0.126

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

from the nurses' station. Similar to our findings on the truncated dataset, there appears to be no significant difference in the total nurse-patient interaction minutes per hour across room distances. This suggests that nurses are reducing visits, but not reducing the total time spent; we find support for the hypothesis that they are batching tasks. As a patient's room distance increases by 10 meters, the duration spent per visit by a nurse increases by about 0.309 minutes. The rooms farthest away from the nurses' station receive visits that are, on average, 1.08 minutes longer compared to the rooms closest to the nurses' station.

Similarly, when using the full dataset to understand the impact of this behavior on care quality (Table 9), we find that patients in rooms every 10 meters farther away experience a 2.9 minute increase in their inter-visit wait time, and make on average 0.120 more calls per hour for the nurse. In the rooms farthest away this translates to 10.15 minutes of additional waiting time between visits and 0.42 more calls per hour. The length of stay of patients across rooms of varying distances remains unchanged.

Table 9: OLS Results Using All Patient Encounters - Call Frequency and LOS Models

	(4) Time btw. Visits (Mins)	(5) Calls/Hr	(6) LOS (Hrs)
Distance (10m)	2.942*** (0.266)	0.120*** (0.007)	0.029 (0.033)
Female Patient	0.022 (0.434)	0.012 (0.012)	−0.088 (0.053)
Patient Race: Asian	−2.107 (1.359)	−0.145*** (0.039)	−0.032 (0.168)
Patient Race: Black	−0.059 (0.476)	−0.039** (0.013)	−0.040 (0.059)
Patient Race: Hispanic	2.603 (1.431)	−0.164*** (0.044)	−0.203 (0.177)
Room Size: Large	−0.364 (0.793)	−0.016 (0.023)	−0.033 (0.098)
Room Size: Small	1.453 (0.918)	−0.011 (0.024)	0.152 (0.112)
MD Staffing: High	−0.256 (0.164)	−0.020 (0.017)	−0.237** (0.078)
MD Staffing: Low	1.590** (0.566)	−0.055*** (0.015)	−0.036 (0.069)
Diagnostic Lab Orders	−0.383*** (0.066)	−0.021*** (0.002)	0.314*** (0.008)
Admitted Patient	0.977 (0.581)	−0.041** (0.014)	1.378*** (0.073)
Shift 2 (3p-3a)	−3.205** (1.047)	−0.078*** (0.016)	0.746*** (0.089)
Shift 3 (11p-11a)	−3.966** (1.350)	−0.007 (0.021)	0.858*** (0.104)
Constant	25.483*** (4.249)	0.455*** (0.086)	1.551*** (0.389)
Observations	15,595	15,595	15,595
Adjusted R^2	0.057	0.109	0.340

*p<0.05; **p<0.01; ***p<0.001

2.11.2. Call Frequency by Gender and Race Patient Groups

We test for differences in call button use across gender and race using an interaction term, and find differences between White, male patients and the following groups: White female patients seem to call more frequently compared to White male patients, whereas Asian male patients, Black female patients, and Hispanic female patients tended to call less frequently compared to White male patients. Results are presented in Table 10.

Table 10: OLS Results on Call Frequency by Patient Groups

	(7) Calls/Hr
Distance (10m)	0.110^{***} (0.008)
White Female Patient	0.079^{***} (0.021)
Asian Male Patient	-0.153^{*} (0.076)
Asian Female Patient	0.020 (0.092)
Black Male Patient	0.020 (0.022)
Black Female Patient	-0.141^{***} (0.027)
Hispanic Male Patient	-0.010 (0.073)
Hispanic Female Patient	-0.202^{*} (0.095)
Room Size: Large	0.066^{**} (0.025)
Room Size: Small	-0.025 (0.026)
MD Staffing: High	-0.103^{***} (0.019)
MD Staffing: Low	-0.080^{***} (0.016)
Triage Level 1	-0.161 (0.103)
Triage Level 2	-0.032^{*} (0.015)
Triage Level 4	-0.042 (0.038)
Triage Level 5	-0.018 (0.119)
Patient Age	-0.002^{***} (0.0004)
Admitted Patient	-0.010 (0.015)
Shift 2 (3p-3a)	-0.096^{***} (0.017)
Shift 3 (11p-11a)	-0.056^{*} (0.024)
Diagnostic Lab Orders	-0.021^{***} (0.002)
Constant	0.309^{**} (0.106)
Observations	9,880
Adjusted R^2	0.145
	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

2.11.3. Core ED Rooms: Room Distances and Fixed Effects

Table 11 outlines the core ED rooms we consider in our analysis, along with the room's distance from the nurses' station and its size and physician staffing level. These variables were used to code the room type fixed effects and our main independent variable, $DIST_{ij}$.

Table 11: ED Room Distances & Fixed Effect Categories

Room Number	Distance	Room Size	Room Staffing Level
1	20m	Regular	High
2	25m	Regular	High
3	45m	Regular	High
4	45m	Regular	High
5	35m	Regular	High
6	25m	Regular	High
7	15m	Regular	High
8	10m	Regular	High
9	15m	Large	High
10	10m	Large	High
11	15m	Large	High
12	20m	Large	High
13	10m	Regular	Low
14	15m	Regular	Low
15	25m	Regular	Low
16	25m	Regular	Low
17	30m	Regular	Low
18	25m	Regular	Low
19	20m	Regular	Low
20	20m	Regular	Low
21	15m	Regular	Regular
22	20m	Regular	Regular
23	30m	Regular	Regular
24	20m	Regular	Regular
25	15m	Small	Regular
26	20m	Regular	Regular
27	20m	Regular	Regular
28	10m	Regular	Regular
29	10m	Regular	Regular
30	20m	Regular	Regular
31	25m	Regular	Regular
32	25m	Small	Regular
33	25m	Regular	Regular

2.11.4. Robust Regression Estimation Tables

The robust regression estimates are outlined below in Tables 12 and 13.

Table 12: Robust Linear Regression Results on Nurse Visit Variables

	(8) Vis/Hr	(9) Dur/Hr	(10) Dur/Vis
Distance (10m)	-0.315^{***} (0.018)	0.069 (0.040)	0.303^{***} (0.016)
Female Patient	-0.005 (0.030)	0.149^{*} (0.065)	0.077^{**} (0.025)
Patient Race: Asian	0.072 (0.095)	0.137 (0.208)	-0.055 (0.082)
Patient Race: Black	0.066^{*} (0.032)	-0.057 (0.070)	-0.075^{**} (0.028)
Patient Race: Hispanic	0.112 (0.094)	-0.049 (0.207)	-0.093 (0.081)
Room Size: Large	0.221^{***} (0.054)	0.348^{**} (0.119)	0.067 (0.047)
Room Size: Small	-0.157^{*} (0.061)	0.122 (0.134)	0.152^{**} (0.053)
MD Staffing: High	0.224^{***} (0.044)	0.187 (0.096)	-0.067 (0.038)
MD Staffing: Low	-0.240^{***} (0.038)	0.346^{***} (0.084)	0.415^{***} (0.033)
Diagnostic Lab Orders	0.055^{***} (0.004)	0.157^{***} (0.010)	0.020^{***} (0.004)
Admitted Patient	0.233^{***} (0.038)	0.572^{***} (0.083)	0.057 (0.033)
Shift 2 (3p-3a)	0.543^{***} (0.067)	0.972^{***} (0.147)	-0.028 (0.058)
Shift 3 (11p-11a)	0.931^{***} (0.094)	1.874^{***} (0.206)	0.061 (0.081)
Constant	1.418^{***} (0.312)	0.877 (0.683)	0.891^{***} (0.268)
Observations	9,880	9,880	9,880

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 13: Robust Linear Regression Results on Patient Length of Stay and Calls

	(11) Time btw. Visits (Mins)	(12) Calls/Hr	(13) LOS (Hrs)
Distance (10m)	2.650*** (0.177)	0.040*** (0.005)	0.025 (0.030)
Female Patient	−0.066 (0.289)	0.011 (0.009)	−0.075 (0.049)
Patient Race: Asian	−0.336 (0.932)	−0.129*** (0.028)	0.082 (0.159)
Patient Race: Black	−0.036 (0.314)	−0.034*** (0.009)	0.069 (0.053)
Patient Race: Hispanic	−0.295 (0.926)	−0.112*** (0.031)	0.001 (0.158)
Room Size: Large	−0.544 (0.533)	0.067*** (0.017)	0.011 (0.091)
Room Size: Small	0.468 (0.600)	0.00002 (0.017)	−0.039 (0.101)
MD Staffing: High	−1.649*** (0.430)	−0.059*** (0.013)	−0.232** (0.072)
MD Staffing: Low	1.541*** (0.375)	−0.023* (0.011)	0.029 (0.063)
Diagnostic Lab Orders	−0.204*** (0.043)	−0.012*** (0.001)	0.265*** (0.007)
Admitted Patient	−0.548 (0.374)	−0.010 (0.010)	1.034*** (0.065)
Shift 2 (3p-3a)	−3.583*** (0.657)	−0.036** (0.011)	0.097 (0.078)
Shift 3 (11p-11a)	−4.290*** (0.921)	−0.067*** (0.016)	−0.068 (0.104)
Constant	21.561*** (3.061)	0.486*** (0.070)	3.112*** (0.416)
Observations	9,880	9,880	9,880

*p<0.05; **p<0.01; ***p<0.001

CHAPTER 3 : The Impact of Medication Delays on Patient Health in the ICU: Estimating Marginal Effects Under Endogenous Delays

Joint work with Ann Huffenberger, Krzysztof Laudanski, and Christian Terwiesch

3.1. Introduction

In operations, we know that customers value the timely delivery of goods and services. Timeliness is one of the key measures of service quality and has been explored in all kinds of service industries. In fact, there is an entire subfield in operations known as queueing theory that studies how the dynamics of a system affect customer wait times. Many of these theories have been applied to call centers, airports, restaurants, hospitals, and even traffic with the goal of improving capacity or prioritization decisions.

From the perspective of service providers, key to any capacity or prioritization decision is to ensure that the customer does not spend too much time waiting. This is primarily driven by the fact that customers incur a cost of waiting, and if the customer is made to wait too long, they might leave. In operations management, there exist many great papers that empirically quantify the customer's cost of waiting. Akşin and colleagues structurally estimate the rewards and waiting cost values for customers calling a bank call center, and find, among other things, that high priority customers observe a cost of waiting equivalent to a dollar per minute (Akşin et al., 2013). Allon and colleagues study the cost of waiting in line at a fast-food drive thru, and find that customers attribute a very high cost to the time they spend waiting. Specifically, to overcome an additional second of waiting time, the restaurant would need to compensate an average customer by as much as five cents (Allon et al., 2011). Batt and Terwiesch study the phenomenon of patients leaving the emergency department without being seen, and how observing the stocks and flows of patients in the waiting room could contribute to this behavior. They find that the predicted probability of abandonment for a medium severity patient increases by 2% with a one-hour increase in wait time (Batt and Terwiesch, 2015).

Remarkably, while the cost incurred by customers waiting in service settings such as call centers and restaurants have been well studied, the health costs incurred by patients waiting for medical interventions has received less attention. Researchers have examined the impact of additional wait time in the emergency department waiting room on the likelihood that patients will leave the queue (Batt and Terwiesch, 2015), and the impact of congestion in the emergency department on delays in medical interventions (Pines et al., 2007), however no work has quantified the causal impact of delays in medical interventions on patient health outcomes.

The causal impact of delays in medical interventions on patient health is both interesting and important, however it is also challenging to study, which may explain why few researchers have examined this in the past. There are two reasons for this: endogeneity in observed delays and health measurement. These challenges are prevalent in the medical literature where empirical questions of this nature are studied cross-sectionally with each datapoint representing a patient's visit to the hospital. Delays are then captured by some one-dimensional best practice target, for example the time to first antibiotic (Gaieski et al., 2010). Then, a regression is run that measures the association between antibiotic delay and patient in-hospital mortality, controlling for patient-specific variables. The concern with this approach is that the observed antibiotic delays in the data are not exogenous, since providers could be making prioritization decisions that impact the recorded delays in a way that is impossible for the researcher to observe. As a result, the measured effects are associations rather than causal relationships. The second challenge is the measurement of health outcomes, which has largely been the result of a lack of availability of more granular health data. Mortality is one of the most commonly used outcomes in healthcare-related research. While mortality as an outcome is clean, it is also crude. Often, the medication delays observed by patients visiting a well-run hospital are too short to create real variation in mortality.

We overcome these two limitations by using instrumental variables to exploit exogenous variation in medication delays introduced by shift changes and unit care coordination activities that allow us to cleanly identify the causal impact of a medication delay on granular patient health outcomes. To avoid using patient mortality as an outcome, we utilize a novel, large-scale dataset that contains patient vital signs electronically archived every 15 minutes during the entire duration of a patient's stay in the intensive care unit (ICU). This allows us to make the following contributions:

1. We find that urgent and unplanned medications scheduled for immediate delivery in the ICU are delayed by 88 minutes on average.

2. We use an identification strategy that uses instrumental variables to capture exogenous variation in medication delays introduced through nurse shift changes and unit care coordination activities. Our instrumental variables explain a significant portion of observed variation in medication delays.

3. We measure the impact of these exogenous delays in medication on novel and granular measures of patient health constructed using patient vital signs during their stay in the ICU. We do this for multiple medication groups and multiple vital sign thresholds so that we can better understand the delay costs of several commonly administered medications across a broad spectrum of health conditions.

This study is one of the first to causally quantify the impact of medication delays on patient health. Our construction of novel patient health metrics based on real-time tracking of patient vital signs during their stay allows us to establish the immediate impact of a medication delay on the patient's health state over the next couple of hours. While this measure of health outcome is correlated to more commonly used outcomes such as the patient length of stay and mortality, we are uniquely able to also capture patient suffering.

This is particularly relevant in the intensive care unit, where patients are highly unstable and rely on healthcare providers to ensure they remain in healthy states. Our findings can assist healthcare providers to (1) better understand the impact of these delays in medications on patient health, and (2) improve medication prioritization such that any necessary delays experienced by the patient incurs minimal harm to the patient's health.

3.2. Related Work

Our work builds upon two main streams of literature: empirical studies of service delays in operations management, and empirical studies of medical intervention delays on patient health outcomes in health services research.

The operations management literature separates neatly into articles studying non-healthcare service settings and those focusing on healthcare settings. Articles from the former group have quantified the customer's empirical cost of waiting across a wide variety of service settings. Akşin and colleagues structurally estimate the rewards and waiting cost values for customers calling a bank call center and find, among other things, that high priority customers observe a cost of waiting equivalent to a dollar per minute (Akşin et al., 2013). Yu and colleagues show that delay announcements can actually impact customers' waiting costs in a call center setting. They find that the per-unit waiting cost of customers actually decreases with the offered waiting time associated with the announcements (Yu et al., 2016). Allon and colleagues study the cost of waiting in line at a fast-food drive-thru, and find that customers attribute a very high cost to the time they spend waiting. Specifically, to overcome an additional second of waiting time, the restaurant would need to compensate an average customer by as much as five cents (Allon et al., 2011). The impact of customer waiting on purchase behavior has also been examined in the retail setting by studying customer purchase behavior at a deli counter in a supermarket (Lu et al., 2013). Craig and colleagues have studied the impact of the wait time experienced during a blood donation encounter on the likelihood of a subsequent donation. They find that a 38% increase in the average wait time to donate blood results in a 10% decrease in the number of donations per year (Craig et al., 2016).

Operations management studies focusing on service delays in the hospital setting have examined the impact of queue configuration on patient wait times and length of stay in the emergency department. Song and colleagues compare the wait times experienced by patients visiting the emergency department under a pooled queueing system versus a dedicated queueing system and find a 9% decrease in the average wait time experienced by patients under the dedicated queueing system. One of the main drivers of this reduction in wait time with the dedicated queue configuration is the physician's ownership over their own set of patients and resources (Song et al., 2015a). Much of the literature in this space has examined how healthcare provider behavior can be a lever through which to control the delays that exist in the healthcare delivery process. Articles studying this have shown that focus (Kc and Terwiesch, 2011), multitasking (Kc, 2013), early-task-initiation (Batt and Terwiesch, 2016), and speeding up (Jaeker and Tucker, 2017) are mechanisms through which the healthcare provider adapts their work behavior to mitigate the impacts of long queues and high workload. Further, studies in healthcare operations management have also examined the impact of intervention delays on the patient. Batt and Terwiesch study the phenomenon of patients leaving the emergency department without being seen, and how observing the stocks and flows of patients in the waiting room could contribute to this behavior. They find that the predicted probability of abandonment for a medium severity patient increases by 2% with a one-hour increase in wait time (Batt and Terwiesch, 2015). Song and colleagues study the impact of delays in care introduced by off-service placement and find that the additional distance that physicians must travel to get to patients placed in off-service units result in a longer length of stay, and is associated with a higher likelihood of readmission, in-hospital mortality, and adverse health events (Song et al., 2018).

Studies of delays in the health services research literature have focused their efforts on understanding the impact of service delays in healthcare on patient health outcomes. However, due to limitations in the types of data available for researchers to do this type of work, the focus has been on utilizing the available data to conduct inference on the delays of interventions such as medications and ambulances on well-documented patient outcomes such as mortality. This includes work showing an association between delays in transfer from the emergency department to the ICU and patient mortality (Chalfin et al., 2007), and an association between the time to antibiotics in the emergency department on mortality from sepsis in the emergency department (Gaieski et al., 2010). Other associations surrounding delayed medical interventions that have been studied include delays in ambulance transit times on mortality (Jena et al., 2017), delays in emergency surgery on mortality (McIsaac et al., 2017), and time spent waiting for health services on mortality (Prentice and Pizer, 2007). Upstream from this literature are medical journal articles that have established associations between hospital crowding and subsequent delays in medical interventions such as antibiotics (Pines et al., 2007) and pain treatment (Pines and Hollander, 2008).

The impact of such work has been profound in understanding the cost of delays on customer behavior across a wide variety of service settings; we build upon the work done by our operations management and clinical colleagues by assembling a dataset that allows us to causally study the impact of a marginal minute of medication delay on granular patient health metrics constructed using vital signs. This is doubly important for patients spending time in the ICU or in any inpatient hospital bed, since these patients are often in critical condition and are unable to leave if they receive an unsatisfactory service. We contribute to the existing literature by generating insights surrounding how service delays in healthcare could directly impact the patient's health.

3.3. Empirical Setting & Data

Our data come from an academic medical center that serves a large metropolitan city. We focus on 4 of the main intensive care units in this medical center, which are located across 2 different hospitals. Together, these ICUs maintain 100 critical care beds, and observe a patient volume of 41,370 patients over our 4 year study period spanning 2012 through 2015. In our study sample, 43% of patients were female, and the average age was 61, with

Table 14: ICU Patient Vital Sign Thresholds

Vital Sign	Lower Threshold	Upper Threshold
Mean Arterial Pressure (MAP) (mmHg)	60	90
Systolic Blood Pressure (SBP) (mmHg)	90	120
Diastolic Blood Pressure (DBP) (mmHg)	none	80
Heart Rate (bpm)	60	100
Temperature (F)	96	100
Oxygen Saturation	98	none
Respiratory Rate (Bpm)	6	20

a standard deviation of 15 years. The ICU is typically the unit in the hospital that sees the sickest patients, and patients are usually admitted through the emergency department (Chan et al., 2016), after surgery, or because the patient's condition worsened in another unit or facility.

Patients in our study sample have an average ICU length of stay of 2.7 days, with a median length of stay of 1.8 days. The distribution of the patient length of stay is shown in Figure 5. 58.4% of patient discharges from the ICU are to a long term care facility for patients to continue receiving care after leaving the ICU. 8.0% of patients die during their time in the ICU, and 32.2% of patients are discharged to a regular (non-'intensive-care') unit within the hospital. Thus, patients spending time in these units are typically very unstable, and require intensive monitoring by healthcare providers. The nurse to patient staffing ratios in these units are 1 patient to 1 nurse, and at most 2 patients to 1 nurse. The ICU care environment is ideal to study the clinical impact of medication delays, since it is often a goal of ICU care to keep patients in healthy vital states, and this is typically done through the administration of medications. Determining whether or not a patient is in a healthy vital state is based on established clinical guidelines. The actual clinical vital sign thresholds used in our study hospitals are outlined in Table 14; note that the mean arterial pressure (MAP) is equivalent to [(2*the value of the diastolic blood pressure (DBP)) + the value of the systolic blood pressure (SBP)] $\div$ 3. We also include all three measures of blood pressure here to test the impact of medications on all of them separately, in case effects vary.

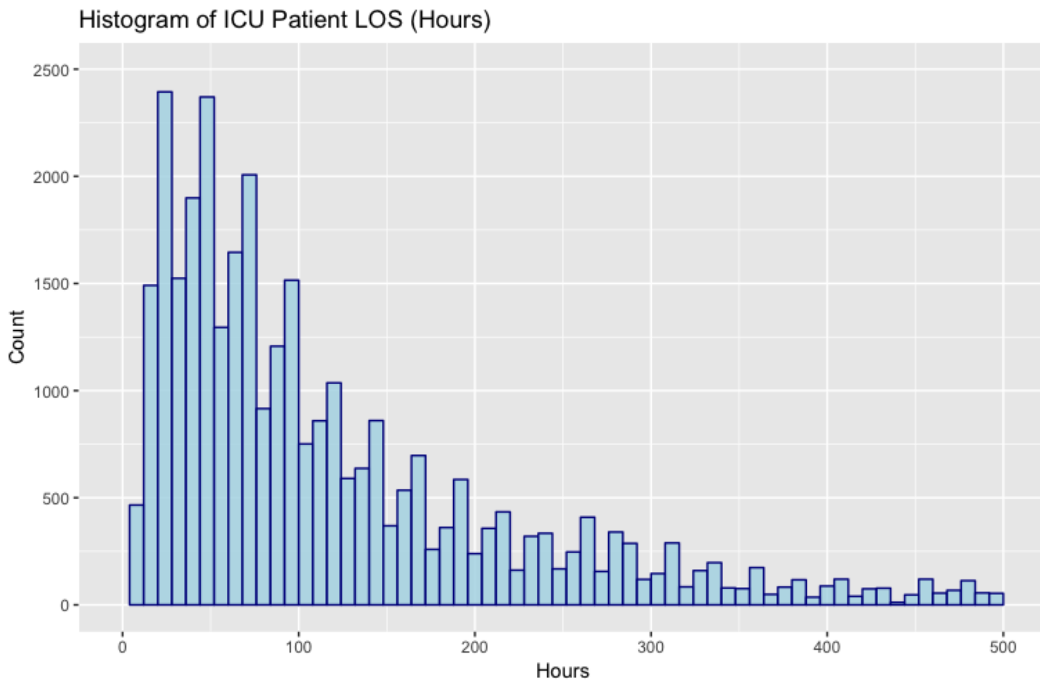


Figure 5: Empirical distribution of the ICU patient length of stay (hours)

We compile three sources of data to create our analysis dataset: the patient medical record, patient vital sign tracking data, and the database of all medications ordered and delivered during a patient's stay in the ICU. To illustrate how we measure medication delay, it is important to understand how medications are ordered and delivered in the ICU. The moment a physician places an order for a medication, a timestamp is generated in the medication database for that patient and labeled the "medication order time". When the medication is about to be delivered to the patient, the nurse administering the medication scans the medication and the patient identification wristband, prior to actual medication delivery. This process of bar-code scanning medications and patients prior to medication delivery was implemented as a way to reduce errors in medication delivery (Poon et al., 2010). It also allows us to cleanly identify the moment a medication was delivered to the patient; when the nurse scans the medication in, a second timestamp is created in the medication database labeled "medication delivery time". In practice, physicians could place an order for a medication to be delivered at a subsequent time, this would be identified in our dataset as

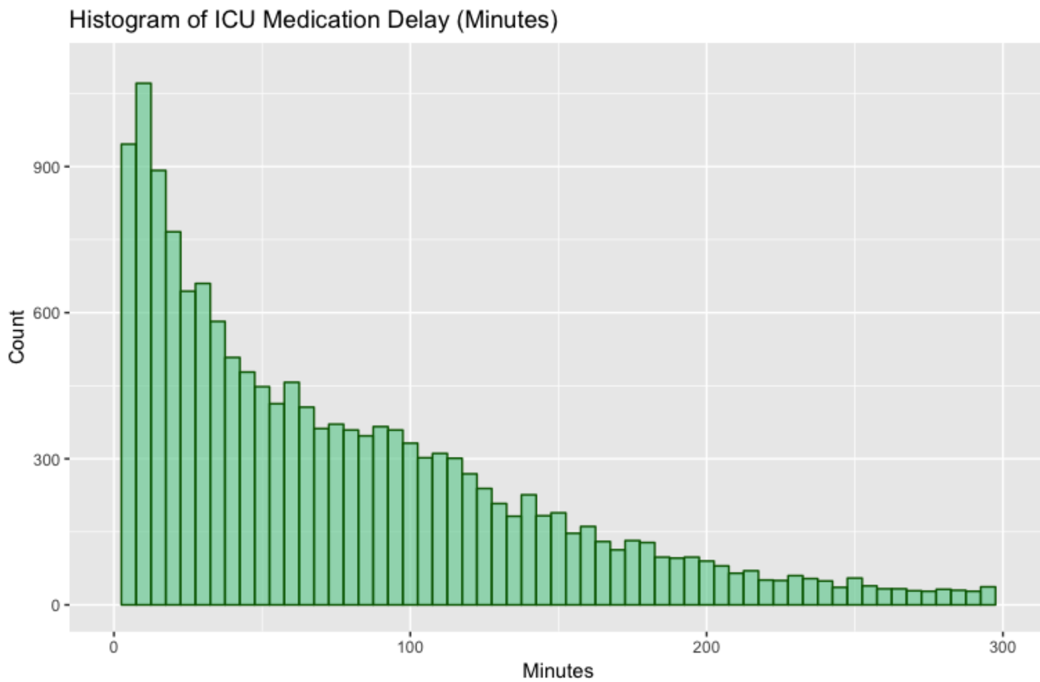


Figure 6: Empirical distribution of ICU medication delays

a medication having a later "medication scheduled time" compared to its "medication order time". For the purposes of our study, in order to cleanly measure the impact of medication delays on patient health, we focus only on unplanned, urgent medications ordered to be delivered to the patient immediately (i.e. the "medication scheduled time" is identical to the "medication order time"). This represents about 30% of medication orders in this setting. Figure 6 shows the distribution of medication delays for unplanned, urgent medications in the ICU. The average delay is 88 minutes, with a median delay of 62 minutes.

For identification purposes, we focus only on the medications that are typically unplanned, and ordered urgently for immediate delivery. We also require that the medication be stocked and available in the ICU such that the delay observed is not confounded by a delay in medication retrieval from the pharmacy. The dataset originally was comprised of 870 unique medications, which we group into 50 medication groups based on clinical function. Of these, we focus on 9 medication groups that satisfy the criteria of (1) being universally prescribed to a wide variety of patients to assist in keeping them in healthy states in an unplanned

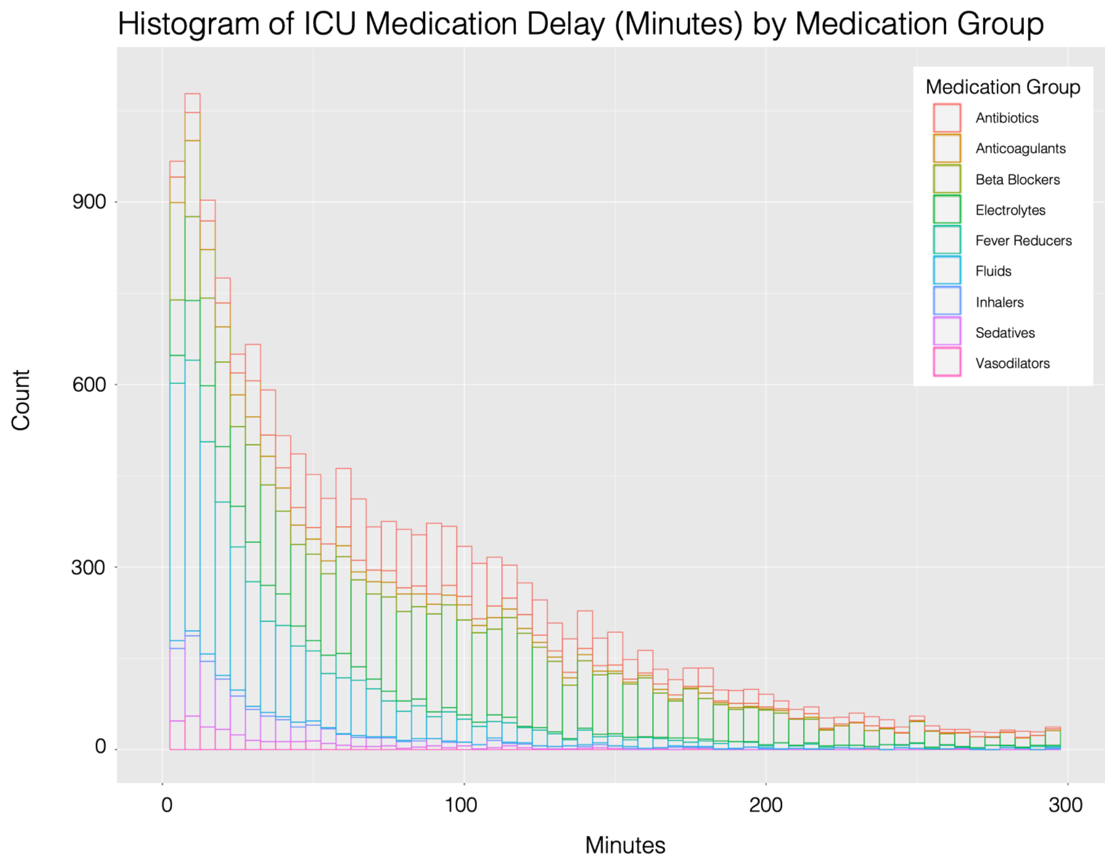


Figure 7: Empirical distribution of ICU medication delays by medication group

and urgent fashion, and (2) typically being stocked and available for quick retrieval on the unit where the patient resides. The list of medication names within each group is shown in the Appendix. The histogram of observed delays separated by these 9 medication groups is shown in Figure 7 and the empirical average delays observed by the same groups are outlined in the Table 15.

3.4. Identification Strategy

Ideally, to study this question we would set up the following experiment: we would randomize ICU patients into two groups with the control group receiving low to zero medication delays and the treatment group receiving large exogenous medication delays. We would ask nurses to monitor patients to study the impact of these delays on their health. Un-

Table 15: Empirical Medication Delay by Medication Group

Medication Group	Mean Empirical Delay
Electrolytes	121 minutes
Antibiotics	111 minutes
Anticoagulants	75 minutes
Inhalers	75 minutes
Fever Reducers	61 minutes
Fluids	57 minutes
Sedatives	52 minutes
Beta Blockers	51 minutes
Vasodilators	45 minutes

fortunately, this is both ethically and practically infeasible, so we utilize an identification strategy that captures only the exogenous portion of delays so that we can measure the causal impact of such delays on patient health. This approach is necessary because delays are endogenous in nature; nurses who deliver the medications make prioritization decisions that impact the delays we observe, but in a way that is not identifiable in the data. As a result, any estimation directly using medication delays as an explanatory variable without addressing endogeneity will result in biased estimates.

To measure the causal impact of delays in medication on patient health, we use as our unit of analysis each medication order for each patient (using only the orders of the unplanned & urgently scheduled medications we include in our analysis). Using the illustration in Figure 8, t_1 represents the moment a medication was ordered to be delivered immediately. Since each individual vital reading is noisy, we average the patient's vital sign readings over two hours ($h = 2$) immediately before and after the medication was ordered to create the patient's 'pre-health state' and 'post-health state', respectively (for each vital threshold, as outlined in Table 14). We code each of these states into binary indicators where the healthy state is zero (if the patient's two-hour average vitals were inside the relevant threshold), and the unhealthy state is one (if the patient's two-hour average vitals were outside the relevant threshold) for each vital threshold. For our main analysis we choose the vital average time horizon (h) to be two hours; since medication delays average 88 minutes, a two-hour window would capture the vital readings that could have been the result of most

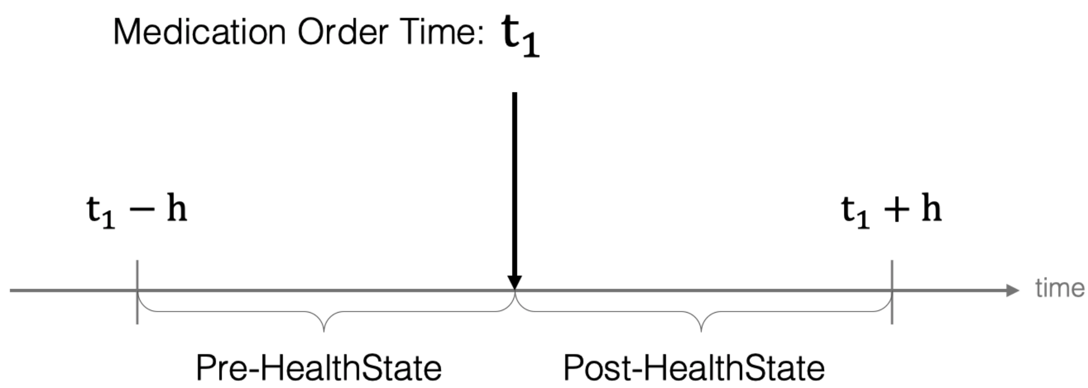


Figure 8: Granular health measurement using vital signs

observed medication delays. In Section 3.7, we vary h and show results for when the time period is 1 hour and 3 hours.

Using these average measures of granular patient health, we use the patient's pre-health state across all 12 vital thresholds (as outlined in Table 14) as control variables and the patient's post-health state as the dependent variable of interest. Controlling for the pre-vital states in this way controls for the impact of any correlation between vital states on the estimated coefficients on medication delay, and ensures that the value we estimate for medication delay is as accurate as possible. Using this measure of granular patient health, we can identify the causal impact of a marginal minute of medication delay on the odds that the patient enters an unhealthy vital state. In addition to controlling for the pre-health state of the patient, we control for other exogenous patient covariates as outlined in Table 16.

Using this identification strategy for each instance of an ordered and subsequently delivered medication, we can capture four scenarios. If the patient started in a healthy state, they can either remain there after medications were ordered, or transition into an unhealthy state. If the patient started in an unhealthy state, they can either remain there, or transition into a healthy state. By using the binary post-health state as the dependent variable, we can measure how a delay in medications could impact the odds that the patient ends up

Table 16: Control Variables for Main Estimation Models

Control Group	Controls
Unit/Hospital Level **W**	Hospital
	Unit
Patient-Medication Level **X**	Patient Gender
	Patient Age
	Patient Complexity
	Primary Diagnosis Code
	Percent of Total Stay Elapsed
	Origin Health State (12)
Time-Related Controls **Z**	Year Medication Ordered
	Month Medication Ordered
	Hour Medication Ordered

in an unhealthy state. In addition, using the patient's origin health states as exogenous explanatory variables, we can measure the difference in the odds of transitioning from either origin state.

We use this identification strategy to separately estimate the impact of medication delays across all 9 medication groups for all 12 vital thresholds. This will allow us to determine the impact of delays in all medication groups across all vital thresholds. If there exist differences in the health cost of delaying certain medications, we can assist providers in better prioritizing the medications to be delivered during busy times.

3.5. Instrumental Variable Estimation

Despite controlling for the patient's pre-health state across multiple vital thresholds as well as clinical variables from the patient's medical record, we cannot rule out the possibility that there remains unobserved variables that impact the medication delays we observe, as well as the patient's vital health status. This is largely because decisions made in the ICU are complex, and what we observe in the medical record is only a fraction of the information that nurses have when making these decisions with the patient physically in front of them. As a result, any unobserved variables that impact both medication delay and patient health will enter the error term and bias our estimated impact of delay.

To overcome this challenge, we use an intrumental variable approach to identify the exogenous portion of medication delay, instruments that we found initially through observing nurse workflow in these units. Our first instrument is the nurse shift change, and our second instrument is the average per-medication delay observed by other patients in the same unit within the same hour that medication was ordered. The nurse shift change occurs daily at 7am and 7pm. During these time windows (6:30am - 7:30am, 6:30pm - 7:30pm), nurses can be found conducting a thorough hand-off procedure with the oncoming staff, which takes their attention off the patient for at least half an hour. We identify this instrument during each instance of a medication order by using a binary indicator for whether or not the medication was ordered within the 6:30am - 7:30am or 6:30pm - 7:30pm time windows. Our second instrument, the average medication delay experienced by other patients on the same unit within the same hour, captures other exogenous shocks on the unit that affect medication delay. These shocks include, but are not limited to, census and high patient load effects, unit level care coordination activities, and any external distractions between providers. Both instruments were identified by the authors after personally observing activities on the unit during shift change hours and outside shift change hours. Care-coordination activities on the unit include meetings held within the central provider staffing area (away from patient rooms) to discuss the status of patients (often with regards to which patient can be discharged to make room for a more severe incoming patient).

Since our measure of patient health is binary, we use the control function approach to obtain consistent coefficient estimates in a two-stage fashion when our second stage is a non-linear logistic regression (Wooldridge and Imbens, 2007). This method uses the error term to adjust for the potential bias introduced when the error term from the first stage model is withheld from the second stage logistic regression. Our first stage equation is the following:

$$MedDelay_{ijt} = \beta_0 + \beta_1 ShiftChange_{ijt} + \beta_2 OtherPtDelay_{ijt} + \mathbf{W}_j\beta_r + \mathbf{X}_{ijt}\beta_e + \mathbf{Z}_t\beta_h + \epsilon_{ijt} \tag{3.1}$$

Since our unit of analysis is at the medication order level, we index i as the patient, j as the medication, and t as the time when the medication was ordered. In Equation 3.1, $ShiftChange_{ijt}$ represents a binary indicator for whether or not the medication order occurred during the shift change time windows. $OtherPtDelay_{ijt}$ represents the average medication delay observed by other patients on the same unit within the same hour that a medication was ordered for the index patient of interest. We take the fitted values of medication delay from this first stage regression ($\widehat{MedDelay}_{ijt}$), and the fitted residuals from the first stage regression ($\widehat{\epsilon}_{ijt}$), to estimate the second stage regression, which in this case is a logistic regression on the binary outcome of post-health state:

$$\begin{aligned} &\log\left(\frac{p(PostUnhealthyState_{ijt})}{1 - p(PostUnhealthyState_{ijt})}\right) \\ &\quad = \beta_\theta + \beta_\gamma \widehat{MedDelay}_{ijt} + \widehat{\epsilon}_{ijt} + \beta_2 OtherPtDelay_{ijt} + \mathbf{W}_j\beta_\alpha + \mathbf{X}_{ijt}\beta_\delta + \mathbf{Z}_t\beta_p \end{aligned} \tag{3.2}$$

Using this estimation strategy, β_γ represents the causal impact of a marginal minute of medication delay on the log odds that a patient enters an unhealthy state as a result.

3.5.1. Instrumental Variable Validity

For these instruments to be valid they must satisfy the relevance condition and the exclusion restriction condition (Wooldridge, 2010), we discuss these in turn.

The relevance condition requires that the chosen instruments explain sufficient variation in our endogenous variable (medication delay). To test this statistically, we perform ANOVA F-tests comparing the first stage model without our instruments and the first stage model with our instruments. The results are shown in Section 3.6.1, and show that our instruments explain a significant portion of variation in observed medication delay. Both of our instruments were chosen because they were observed in practice to exogenously drive increases in medication delay. We find evidence of this in the data; the time window around the nurse shift change is associated with increased medication delays, and increases in the delay observed by other patients on the same unit drive increases in the delay observed by the index patient.

The exclusion restriction condition requires that our instruments impact the patient's health condition only through its effect on medication delays. Support for this assumption comes from how care is delivered in the ICU. Patient vitals are a real-time measure of the health stability of the patient. Providers in the ICU monitor these conditions, and the mechanism through which they can affect this is through medications (which includes the administration of fluids and electrolytes, which are not 'medications' *per se*, but are included due to their therapeutic effect in the ICU). Therefore, our instruments affect medication delays, and this is the primary mechanism through which they affect patient health. The nursing shift change does not alter any other aspect of patient care, except that it is a period of time where the nurses' attention is taken off the patient. Similarly, delays experienced by other patients on the same unit around the same timeframe represent moments where the unit requires the attention of the nurses and so similar to our first instrument, the nurses' attention is taken off the patient and his or her needs. By itself, nurse attention is not able to ensure the patient remains in healthy states. However, combined with timely medication delivery, nurses can ensure that patients remain in healthy vital states because they received their medication on time. Therefore, we have reason to believe that our instruments satisfy the exclusion restriction assumption, and affect patient health only through its effects on medication delay.

3.6. Instrumental Variable Estimation Results

Results for our first stage estimation are shown in Tables 17, 18, and 19. We present coefficients from the first stage estimation across all 9 medication groups.

Table 17: Results: OLS First Stage Results

	(1) Antibiotics	(2) Inhalers	(3) Beta Blockers
Shift Change IV	55.35*** (3.44)	39.51*** (3.91)	39.97*** (1.79)
Other Patient Delay IV	0.79*** (0.01)	0.76*** (0.02)	0.36*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.35	0.08	0.24
F-Statistic	3027.3***	1117.6***	1971.9***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 18: Results: OLS First Stage Results

	(4) Anticoagulants	(5) Fluids	(6) Electrolytes
Shift Change IV	89.68*** (2.88)	35.56*** (3.42)	47.69*** (2.89)
Other Patient Delay IV	0.49*** (0.01)	0.34*** (0.01)	0.65*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.27	0.08	0.28
F-Statistic	2141.3***	522.7***	1859.9***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

We observe strong first stage regressions, with F-Statistic estimates ranging from 522.7 (Model 5, F-Statistic = 522.7***) to 3027.3 (Model 1, F-Statistic = 3027.3***), suggesting that the addition of our two instrumental variables explain significantly more variation in the observed delays compared to models without them. From our first stage results, we find that the time window around a nurse shift change in the ICU is associated with

Table 19: Results: OLS First Stage Results

	(7) Fever Reducers	(8) Sedatives	(9) Vasodilators
Shift Change IV	60.81*** (3.14)	48.81*** (3.49)	57.11*** (2.06)
Other Patient Delay IV	0.56*** (0.01)	0.37*** (0.01)	0.21*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.25	0.07	0.12
F-Statistic	1752.7***	748.52***	1015.0***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 20: First Stage Estimated Exogenous Delays by Medication Group

Medication Group	Mean Empirical Delay	Shift Change Delay IV	Other Patient Delay IV
Electrolytes	121	47.69	0.65
Antibiotics	111	55.35	0.79
Anticoagulants	75	89.68	0.49
Inhalers	75	39.51	0.76
Fever Reducers	61	60.81	0.56
Fluids	57	35.56	0.34
Sedatives	52	48.81	0.37
Beta Blockers	51	39.97	0.36
Vasodilators	45	57.11	0.21

Units are in minutes.

an increase in the medication delay of 35 minutes for fluids (Model 5, $\beta_1 = 35.56^{***}$) to 89 minutes for anticoagulants (Model 4, $\beta_1 = 89.68^{***}$). For each additional minute of the average medication delay experienced by other patients on the unit, an index patient observes a delay between 0.21 minutes for vasodilators (12.6 seconds; Model 9, $\beta_2 = 0.21^{***}$) to 0.79 minutes for antibiotics (47.4 seconds; Model 1, $\beta_2 = 0.79^{***}$). We summarize our instrumented exogenous delays alongside the observed empirical mean endogenous delays in Table 20.

Our second stage estimation results are presented in Tables 21 and 22; only the coefficients on β_γ are presented here. The coefficients here represent the percentage change in the odds of entering the relevant adverse vital threshold shown in the corresponding column due to a 1 minute delay in a medication from the group in the corresponding row. Since we run many models to determine these results (each cell represents a model), we apply the false discovery rate correction using the Benjamini-Hotchberg procedure to address multiple testing concerns. This is discussed in the Appendix.

Table 21: Second Stage Logistic Regression: β_γ Coefficient Estimates (Blood Pressure Vitals)

	MAP Low	MAP High	SBP Low	SBP High	DBP High
Antibiotics	$-0.57\%^{*}$	$0.19\%^{*}$	0.07%	0.10%	0.19%
Inhalers	-1.05%	-0.22%	-1.73%	-0.31%	0.68%
Beta Blockers	-1.10%	$0.67\%^{**}$	-0.10%	$0.63\%^{*}$	$0.82\%^{**}$
Anticoagulants	-0.30%	-0.10%	-0.19%	0.08%	-0.17%
Fluids	0.50%	$-0.50\%^{*}$	$0.50\%^{*}$	$-0.40\%^{*}$	-0.37%
Electrolytes	-0.25%	$0.15\%^{*}$	$-0.30\%^{**}$	$0.15\%^{*}$	0.01%
Fever Reducers	-0.06%	-0.36%	0.06%	-0.27%	-0.39%
Sedatives	-0.81%	-0.04%	0.32%	-0.15%	0.32%
Vasodilators	-97.10%	$2.99\%^{***}$	-0.43%	$3.19\%^{***}$	$1.54\%^{**}$

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 22: Second Stage Logistic Regression: β_γ Coefficient Estimates (All Other Vitals)

	HR Low	HR High	TEMP Low	TEMP High	O2 Low	RR Low	RR High
Antibiotics	0.26%	-0.07%	0.27%	-0.09%	0.06%	-0.25%	0.20%*
Inhalers	1.71%	1.18%*	-98%	0.76%	0.28%	-78.09%	1.06%*
Beta Blockers	-0.12%	0.65%*	-2.76%	-0.18%	0.25%	-100%	0.23%
Anticoagulants	0.56%	-0.28%	-99.95%	-1.00%	-0.05%	-0.34%	0.02%
Fluids	0.61%	-0.13%	-0.34%	-0.53%	-0.09%	-1.78%	−0.47%**
Electrolytes	-0.40%	−0.17%*	0.60%*	-0.18%	-0.07%	-0.57%	0.04%
Fever Reducers	-1.11%	0.14%	-2.41%	0.54%*	0.00%	0.91%	-0.06%
Sedatives	-0.14%	0.06%	-0.51%	-0.28%	-0.17%	-100%	0.03%
Vasodilators	-0.18%	0.13%	3.97%	-0.78%	0.66%	-97.84%	1.35%**

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

This estimation method has allowed us to quantify the causal impact of a marginal minute of delay across multiple commonly prescribed medication groups on the odds that a patient enters an adverse vital state across multiple vital thresholds that are used in clinical practice. Many of the coefficients estimated make intuitive sense: a 1 minute delay of vasodilators increases the odds of the patient entering a adverse high blood pressure state by 2.99% (MAP High, $\beta_\gamma = 2.99$, $p < 0.001$) and 3.19% (SBP High, $\beta_\gamma = 3.19$, $p < 0.001$). Vasodilators are prescribed to lower blood pressure so it should follow that delaying this medication would result in the patient entering a high blood pressure state. Similarly, a minute of delay in fever reducing medication results in an increased odds of entering a fever state by 0.54% (Temperature High, $\beta_\gamma = 0.54$, $p < 0.05$). Beta blockers are often prescribed to lower blood pressure and heart rate, so it makes meaningful sense that our results show that delaying beta blockers results in an increase in the odds of increased blood pressure (MAP High, $\beta_\gamma = 0.67$, $p < 0.01$; SBP High, $\beta_\gamma = 0.63$, $p < 0.05$; DBP High, $\beta_\gamma = 0.82$, $p < 0.01$) and heart rate (HR High, $\beta_\gamma = 0.65$, $p < 0.05$). Inhalers are typically given for asthma and breathing conditions, so delaying the delivery of inhalers increases the odds of entering a high respiratory rate state (RR High, $\beta_\gamma = 1.06$, $p < 0.01$).

To show the direction of the bias introduced by endogeneity, we present estimates from models without our instrumental variables in Appendix 3.10.3. To ensure the robustness of our estimates, we also show the second stage results using a linear probability model with errors clustered at the patient level in Appendix 3.10.4.

3.6.1. Instrumental Variable Validity Test

We test the relevance of our instruments by performing an ANOVA F-test that compares models with only our exogenous variables to models with our exogenous variables as well as our two instrumental variables (Hall et al., 1996). From this analysis we observe F-Statistic estimates that range from 522.7 (Model 5, F-Statistic = 522.7***) to 3027.3 (Model 1, F-Statistic = 3027.3***); all estimates are shown in Tables 17, 18, and 19. This suggests that our instrumental variables explain a sufficient amount of variation in our endogenous regressor, and the coefficients are highly unlikely to equal zero. In fact, our two instrumental variables are among the two variables in our first-stage regressions with the strongest signal in explaning delay. This makes sense, since our instruments were selected because of their exogenous impact on medication delay.

3.7. Robustness Analysis

3.7.1. Vital Average Time Window

We vary the time period over which we average the patient vitals immediately before and after a medication was ordered, and find the following results from the second stage estimation using a time period of 1 hour (Tables 23 and 24), and 3 hours (Tables 25 and 26). Results from the first stage estimation are in the Appendix (Tables 31 through 36).

The results we obtain using these two thresholds are slightly different. We average vital readings over an hourly period instead of using individual vital readings as a way to minimize the variance in our estimate of the patient's health state. Patient vital signs can fluctuate quite a lot (oftentimes because of errors in recording or patient movement that results in abnormal readings), so using an hourly average reduces the impact of these noisy values on our estimate of patient health. We initially chose two hours as the hourly cutoff because (1) the average medication delay in our dataset is 88 minutes (or 1 hour and 28 minutes), and (2) there is usually a slight lag after medication administration before medications begin to work. If we assume that most medications begin to work within 30 minutes, then an hourly average cutoff of 2 hours would capture the impact of most medication delays on patient vital signs after the medication was ordered, while also capturing the clinical efficacy of the medication given by including several instances of vital readings after the medication begins to take effect. When we change this to average vital signs over a 1-hour window, we only capture patient vitals the hour immediately before and after medications were ordered, which is noisy. Since the patient's vital signs have not had sufficient time to either reach an unhealthy state and stabilize there, or remain in a healthy state, our identification strategy does not have sufficient power in identifying an effect. Many of our effects disappear. When we use a 3-hour window, we identify more effects, since the measures of health (when averaged over a longer time period) tend to be more stable, and the efficacy of delayed medications are more likely to be captured in later vital readings. The effects found using this time-window are very similar to the effects found using the 2-hour time-window. However, with respect to identification, the 2-hour window is a much cleaner choice since patient vitals over a 3-hour window could be affected by additional interventions due to the prolonged time-window.

Table 23: Second Stage Logistic Regression Using 1-Hour Window: β_γ Coefficient Estimates (Blood Pressure Vitals)

	MAP Low	MAP High	SBP Low	SBP High	DBP High
Antibiotics	-0.05%	0.26%	0.27%	-0.04%	0.10%
Inhalers	-1.27%	-1.88%	-1.29%	-0.35%	-3.75%
Beta Blockers	-0.99%	0.77%	0.22%	0.76%	0.35%
Anticoagulants	-1.99%	-1.21%	-1.93%	-1.65%	-1.48%
Fluids	1.06%	-0.31%	0.57%	-0.20%	-0.15%
Electrolytes	-0.41%	-0.04%	-0.75%	-0.05%	-0.04%
Fever Reducers	0.83%	0.00%	-0.81%	-0.45%	0.28%
Sedatives	-5.06%	0.72%	-2.04%	1.58%	1.35%
Vasodilators	-0.66%	2.44%*	-0.75%	1.69%	2.02%

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 24: Second Stage Logistic Regression Using 1-Hour Window: β_γ Coefficient Estimates (All Other Vitals)

	HR Low	HR High	TEMP Low	TEMP High	O2 Low	RR Low	RR High
Antibiotics	1.33%*	-0.07%	0.52%	0.44%	0.21%	-1.34%	0.16%
Inhalers	-0.34%	0.55%	-0.22%	-1.27%	0.07%	-0.97%	-0.49%
Beta Blockers	-3.21%	-0.58%	0.58%	-1.30%	0.54%	0.49%	-0.32%
Anticoagulants	-0.42%	-0.27%	-96.66%	-0.65%	-0.01%	-0.10%	-0.24%
Fluids	2.81%	0.35%	-0.99%	-0.29%	-0.13%	-2.62%	-0.52%
Electrolytes	-0.43%	-0.58%	1.37%*	-0.23%	-0.05%	-0.47%	-0.05%
Fever Reducers	-0.98%	0.06%	-0.79%	0.05%	-0.44%	1.28%	0.24%
Sedatives	1.82%	0.80%	4.31%	-1.82%	0.14%	-3.45%	0.20%
Vasodilators	-0.96%	-0.34%	-4.29%	-0.48%	-0.91%	-0.38%	1.76%

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 25: Second Stage Logistic Regression Using 3-Hour Window: β_γ Coefficient Estimates (Blood Pressure Vitals)

	MAP Low	MAP High	SBP Low	SBP High	DBP High
Antibiotics	-0.10%	0.04%	0.16%	0.07%	-0.13%
Inhalers	-0.96%	0.27%	-1.64%	-0.22%	0.35%
Beta Blockers	-0.15%	0.71%***	-0.38%	0.56%***	0.77%***
Anticoagulants	-0.60%	0.01%	-0.24%	0.07%	-0.09%
Fluids	0.68%***	-0.74%	0.64%***	-0.44%	-0.36%
Electrolytes	-0.49%	0.12%	-0.30%	0.16%***	0.10%
Fever Reducers	-0.44%	-0.39%	-0.29%	-0.20%	-0.51%
Sedatives	-0.37%	0.07%	0.00%	-0.07%	0.24%
Vasodilators	-27.13%	2.56%***	-0.85%	4.17%***	0.87%*

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 26: Second Stage Logistic Regression Using 3-Hour Window: β_γ Coefficient Estimates (All Other Vitals)

	HR Low	HR High	TEMP Low	TEMP High	O2 Low	RR Low	RR High
Antibiotics	0.09%	-0.06%	0.78%**	-0.03%	0.02%	0.17%	0.13%*
Inhalers	1.11%	0.40%	-0.04%	0.31%	0.07%	-0.81%	0.32%
Beta Blockers	-0.53%	1.02%***	-2.00%	0.13%	0.05%	-21.85%	0.31%
Anticoagulants	0.34%	0.04%	-2.04%	-0.17%	-0.11%	-0.05%	-0.15%
Fluids	0.51%	-0.03%	-0.02%	0.06%	-0.22%	0.95%*	-0.36%
Electrolytes	0.07%	-0.13%	0.19%	-0.10%	-0.04%	-0.60%	0.02%
Fever Reducers	-0.51%	0.04%	0.28%	0.60%***	0.08%	0.35%	-0.03%
Sedatives	-0.07%	-0.03%	-0.49%	-0.18%	-0.01%	-13.08%	0.08%
Vasodilators	0.20%	0.49%	3.64%**	-0.37%	0.16%	-0.09%	0.68%

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

3.8. Managerial Implications

In this study, we find that patients in the ICU experience an average medication delay of 88 minutes for unplanned, urgently scheduled medications meant to be delivered immediately. We quantify the causal impact of some of this delay on patient health through an examination of 9 separate medication groups and 12 vital sign thresholds that are regularly monitored in the intensive care unit. We do this using an instrumental variable approach that captures exogenous variation in medication delay during shift changes and care coordination activities on the unit.

Despite the existence of literature showing an association between medication delays and patient mortality in the emergency department, studies have not previously been able to causally quantify the impact of a delay in medication on patient health in a way that prescribes an actionable insight for providers to do better. For example, knowing that delays in medications are associated with a slightly higher risk of in-hospital mortality might convince providers to improve medication delivery times, but without information about exactly what the causal effect of such delays are on patient health across different medication groups, the finding is less actionable. In this work, we quantify the exact increase in the odds that any patient will enter an unhealthy vital state across 12 closely monitored vital thresholds in the ICU due to an additional minute of medication delay across 9 medication groups. Because we do this for 9 groups of commonly used and stocked medications in the ICU, we can (using our estimated coefficients), create a prioritization list of medications based on their adverse effects on the patient. Let us assume, for illustration purposes, that the medications that impact more vital thresholds and cause a larger effect when delayed are worse to delay. Then, we can order the medications based on their "estimated clinical urgency". We can then compare our estimated clinical urgency against the empirical mean delays that we observe in the data. This would allow us to classify medications as being "statistically under-prioritized" or "statistically over-prioritized", this is shown in Figure 9.

This comparison shows us that providers are in fact prioritizing heart and vascular medica-

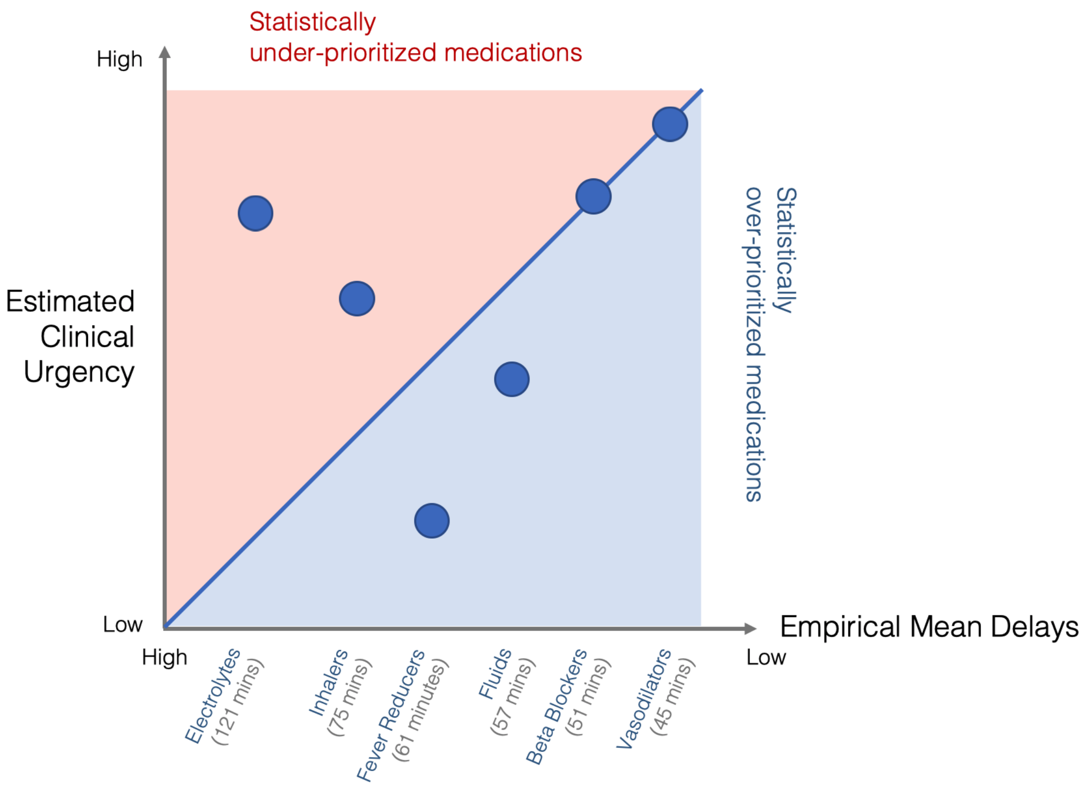


Figure 9: A re-prioritization of medications based on the clinical impact of delays

tions such as vasodilators and beta blockers. In our results these two medications have the most impact on patient vitals (specifically blood pressure and heart rate vital thresholds), and it appears that providers prioritize these medications above others since they have the lowest average empirical delay. Other medications, such as fever reducers and fluids appear to be 'over-prioritized' in that they have lower delays than medications that we estimate would impact the patient more if delayed, such as electrolytes and inhalers (which we label as 'statistically under-prioritized'). Of course, this exercise is simply an illustrative example that shows the potential for these findings to assist providers in better prioritizing medication delivery in the ICU. This would be particularly useful if providers know something about the patient's condition as it relates to vital sign thresholds. For example, a patient arriving in the ICU after a heart surgery would require their blood pressure and heart rate to be kept in healthy vital sign ranges for optimal recovery. Thus, providers knowing this can use our estimated clinical cost of delays to prioritize medications such that delays do not adversely affect the patient's blood pressure and heart rate vital signs.

Given that delays are inevitable in healthcare delivery, particularly in a setting as complex and fast-paced as the intensive care unit, our findings can enable providers to take a step towards mitigating the negative impact of some of this medication delay on the patient's health in a personalized way. This could be one step towards a more personalized approach to healthcare delivery, one where observed delays in medication might only minimally impact patient health.

Further, this work generalizes to services where delays can lead to undesirable outcomes and the prioritization of attention is important due to limited resources; examples of this include emergency services such as firetrucks, ambulances, and project management, among others. Central to the application of this work to these other settings is causally quantifying the cost of delays. In the case of firetrucks or ambulances, one could imagine a scenario where there are multiple locations needing assistance. Prioritization decisions would need to be made given limited resources and the first destination of these rescue vehicles in this scenario would be the one with the highest cost of delay. This study provides a causal framework to study these problems using observational data and instrumental variables; findings using this approach have the potential to assist service providers to make better prioritization decisions and mitigate the negative impact of delays on their customers.

3.9. Conclusion

We empirically quantify the cost of medication delays in the ICU on patient health. This is especially relevant in the ICU setting since patients are typically in critical condition and require medications to remain in healthy states for optimal recovery. Through quantifying these effects, we find that vasodilators and beta blockers are two of the worst medication groups to delay. While it is good to know that vasodilators and beta blockers are among the least delayed medications in the ICU, they still have an average delay of 45 minutes and 51 minutes, respectively. This is only slightly shorter than the average delay across all medication groups in our study, which is 88 minutes. Our contributions can assist providers in making better prioritization decisions when it comes to medication delivery during busy times such as shift changes. By understanding the exact health impact of delaying certain medications on patient vitals across multiple vital thresholds, providers can attempt to better prioritize medication delivery so that patient harm is minimized.

3.10. Appendix

3.10.1. Medication Groupings

Antibiotics: sulfamethoxazole-trimethoprim, tigecycline IVPB, tobramycin IVPB, Vancomycin Injection, amikacin IVPB, amoxicillin oral liquid, ampicillin , Ampicillin/Sulbactam, azithromycin, aztreonam IVPB, ceFAZolin IVPB, cefepime IVPB, cefTAZidime IVPB, ceftazidime-avibactam IVPB, cefTRIAXone IVPB, chloramphenicol IVPB, ciprofloxacin, ceFAZolin IVPB, cefepime IVPB, cefTAZidime IVPB, ceftazidime-avibactam IVPB, cefTRIAXone IVPB, chloramphenicol IVPB, ciprofloxacin, clindamycin, colistimethate IVPB, DAPTOmycin IVPB, doxycycline, erythromycin IVPB, gentamicin IVPB, imipenem and cilastatin IVPB, levoFLOXacin, linezolid, meropenem IVPB, metroNIDAZOLE, nafcillin injection, penicillin G K IVPB, Piperacillin/Tazobactam Inj

Inhalers: acetylcysteine nebulizer solution, albuterol inhaler, albuterol-ipratropium inhaler, ipratropium-albuterol nebulization, beclomethasone 80 MICROgrams inhaler, ipratropium inhaler, levalbuterol 0.63 mg/3mL solution, theophylline, tiotropium bromide inhaler.

Beta Blockers: Atenolol Tablet, carvedilol, esmolol infusion, labetalol, Metoprolol Injection, metoprolol tartrate, nadolol, pindolol, propranolol.

Anticoagulants: argatroban infusion, bivalirudin bolus, dabigatran, enoxaparin 120 mg injection, heparin inf 25000 Units/250 mL, rivaroxaban, warfarin.

Fluids: albumin 5% IVPB, dextrose 10% in water, fluid challenge lactated ringers, fluid challenge, Free Water Bolus, lactated ringers, sodium chloride.

Electrolytes: calcium acetate, calcium carbonate-vitamin D, calcium chloride IVPB, potassium chloride IVPB, potassium phosphate IVPB.

Fever Reducers: acetaminophen, acetaminophen oral liquid, acetaminophen suppository, acetaminophen oral liquid, acetaminophen oral liquid, Acetaminophen Tablet.

Sedatives: ALPRAZolam, chloral hydrate oral liquid, clonazePAM, LORazepam, midazolam infusion, temazepam, traZODone.

Vasodilators: hydrALAZINE, hydrALAZINE injection.

3.10.2. Multiple Testing Correction

We do the false discovery rate correction using the Benjamini-Hotchberg procedure (Benjamini and Hochberg, 1995) on our main results from Section 3.6 to address concerns surrounding multiple testing, and find that all of our coefficients are significant at an α level of 0.26. This suggests that if multiple hypothesis testing is an issue in this study, 26% of our significant findings were a result of such effects.

3.10.3. Estimated Coefficients without Instrumental Variables

To show the direction of the bias, we estimate our models without instrumental variables, instead using the endogenous delays as the predictor. The results of this estimation show 'medication delay' coefficients that are lower, or less significant. This suggests that the endogeneity in patient prioritization is such that nurses prioritize patients who are more sick, and therefore the endogenous medication delays show longer delays for patients who are less sick. As a result, our estimated cost of medication delay is less than it is in reality, since sicker patients are typically getting their medication quicker.

Table 27: Second Stage Logistic Regression Without IVs: β_γ Coefficient Estimates (Blood Pressure Vitals)

	MAP Low	MAP High	SBP Low	SBP High	DBP High
Antibiotics	-0.09%	0.11%*	0.07%	-0.02%	0.07%
Inhalers	-0.10%	0.08%	-1.68%	0.19%	0.02%
Beta Blockers	-0.79%	0.39%***	-0.49%	0.39%***	0.36%**
Anticoagulants	-0.38%	0.06%	-0.23%	0.15%	-0.08%
Fluids	0.06%	$-0.19\%^{***}$	-0.04%	-0.07%	$-0.17\%^{*}$
Electrolytes	$-0.15\%^{*}$	0.04%	$-0.19\%^{**}$	0.07%**	-0.05%
Fever Reducers	-0.13%	0.03%	0.10%	0.01%	-0.03%
Sedatives	0.10%	0.09%	-0.13%	-0.05%	-0.07%
Vasodilators	-97.79%	1.05%***	-0.06%	1.01%***	0.56%**

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 28: Second Stage Logistic Regression Without IVs: β_γ Coefficient Estimates (All Other Vitals)

	HR Low	HR High	TEMP Low	TEMP High	O2 Low	RR Low	RR High
Antibiotics	-0.08%	0.01%	0.37%*	0.04%	0.01%	0.20%	0.15%***
Inhalers	0.18%	0.10%	-97.82%	0.03%	0.06%	-2.10%	-0.05%
Beta Blockers	-0.08%	0.22%	-0.95%	-0.17%	0.08%	-4.18%	0.05%
Anticoagulants	0.21%	0.02%	-0.53%	0.12%	0.15%	-0.72%	-0.16%
Fluids	0.16%	-0.02%	-0.20%	0.00%	-0.07%	0.01%	-0.06%
Electrolytes	-0.02%	$-0.11\%^{**}$	0.24%*	-0.03%	-0.03%	-0.17%	-0.01%
Fever Reducers	-0.97%	0.23%*	-0.08%	0.10%	0.08%	0.10%	0.17%*
Sedatives	-0.10%	0.11%	-1.08%	-0.01%	-0.04%	-4.08%	-0.04%
Vasodilators	-0.49%	-0.35%	1.43%**	-0.11%	0.03%	-98.06%	-0.16%

$^*p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Note: Each cell represents the percentage change in the odds of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

3.10.4. Linear Probability Model Coefficients with Clustered Errors

To ensure the robustness of our results under error clustering at the patient level, we utilize a linear probability model that clusters errors at the patient level. Our results show similar significance levels and the direction of the coefficients are consistent.

Table 29: Second Stage Linear Probability Model: β_γ Coefficient Estimates (Blood Pressure Vitals)

	MAP Low	MAP High	SBP Low	SBP High	DBP High
Antibiotics	$-0.014\%^{*}$	$0.022\%^{*}$	0.005%	0.014%	0.012%
Inhalers	-0.025%	-0.012%	-0.052%	-0.033%	0.051%
Beta Blockers	-0.009%	$0.096\%^{**}$	-0.003%	$0.091\%^{**}$	$0.092\%^{***}$
Anticoagulants	-0.010%	-0.009%	-0.018%	0.007%	-0.012%
Fluids	$0.024\%^{*}$	$-0.046\%^{*}$	$0.034\%^{*}$	$-0.054\%^{*}$	-0.020%
Electrolytes	$-0.008\%^{*}$	0.015%	$-0.014\%^{**}$	$0.020\%^{*}$	-0.002%
Fever Reducers	-0.002%	-0.043%	0.002%	-0.038%	-0.035%
Sedatives	-0.016%	-0.002%	0.011%	-0.025%	0.033%
Vasodilators	-0.013%	$0.328\%^{***}$	-0.011%	$0.221\%^{**}$	$0.262\%^{***}$

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Note: Each cell represents the percentage change in the probability of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

Table 30: Second Stage Linear Probability Model: β_γ Coefficient Estimates (All Other Vitals)

	HR Low	HR High	TEMP Low	TEMP High	O2 Low	RR Low	RR High
Antibiotics	0.003%	-0.007%	0.000%	-0.010%	0.007%	-0.002%	0.029%*
Inhalers	0.025%	0.125%*	-0.030%	0.051%	0.047%	-0.004%	0.142%*
Beta Blockers	0.000%	0.064%*	-0.003%	-0.009%	0.049%	-0.005%	0.027%
Anticoagulants	0.013%	-0.024%	0.002%	0.022%	-0.009%	0.005%	0.008%
Fluids	0.007%	-0.008%	0.010%	-0.025%	-0.017%	-0.007%	−0.061%**
Electrolytes	-0.005%	−0.015%*	0.001%	-0.007%	-0.011%	-0.002%	0.007%
Fever Reducers	-0.006%	0.014%	-0.003%	0.046%**	0.001%	0.016%***	-0.008%
Sedatives	-0.010%	0.015%	-0.007%	-0.019%	-0.028%	-0.010%	0.018%
Vasodilators	-0.007%	0.026%	0.045%*	-0.039%	0.103%	-0.004%	0.200%**

*p<0.05; **p<0.01; ***p<0.001

Note: Each cell represents the percentage change in the probability of entering an adverse health state in the corresponding column due to one additional minute of delay in a medication from the corresponding row.

3.10.5. Additional Tables from Robustness Checks Section

First-Stage Results Using Vital Sign Window of 1 Hour

Table 31: Results: OLS First Stage Results - 1 Hour Window

	(10) Antibiotics	(11) Inhalers	(12) Beta Blockers
Shift Change IV	69.10*** (7.31)	55.23*** (2.35)	28.89*** (3.23)
Other Patient Delay IV	0.91*** (0.02)	0.19*** (0.01)	0.40*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.34	0.16	0.25
F-Statistic	732.7***	158.1***	527.7***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 32: Results: OLS First Stage Results - 1 Hour Window

	(13) Anticoagulants	(14) Fluids	(15) Electrolytes
Shift Change IV	117.90*** (7.99)	27.42*** (3.42)	65.51*** (6.51)
Other Patient Delay IV	0.36*** (0.02)	0.37*** (0.01)	0.57*** (0.02)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.17	0.25	0.21
F-Statistic	317.38***	146.37***	416.80***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 33: Results: OLS First Stage Results - 1 Hour Window

	(16) Fever Reducers	(17) Sedatives	(18) Vasodilators
Shift Change IV	60.13*** (6.74)	47.05*** (8.43)	82.59*** (2.06)
Other Patient Delay IV	0.376*** (0.02)	0.28*** (0.03)	0.09*** (0.009)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.28	0.14	0.24
F-Statistic	231.43***	71.54***	290.10***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

First-Stage Results Using Vital Sign Window of 3 Hours

Table 34: Results: OLS First Stage Results - 3 Hour Window

	(19) Antibiotics	(20) Inhalers	(21) Beta Blockers
Shift Change IV	60.89*** (2.52)	18.58*** (2.37)	46.41*** (1.50)
Other Patient Delay IV	0.81*** (0.01)	0.76*** (0.01)	0.36*** (0.005)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.35	0.18	0.22
F-Statistic	6105.7***	3296.3***	3505.1***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 35: Results: OLS First Stage Results - 3 Hour Window

	(22) Anticoagulants	(23) Fluids	(24) Electrolytes
Shift Change IV	78.39*** (1.92)	45.60*** (2.46)	46.22*** (1.86)
Other Patient Delay IV	0.56*** (0.01)	0.32*** (0.01)	0.57*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.31	0.11	0.26
F-Statistic	4478.7***	1106.9***	3453.2***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

Table 36: Results: OLS First Stage Results - 3 Hour Window

	(25) Fever Reducers	(26) Sedatives	(27) Vasodilators
Shift Change IV	61.93*** (2.36)	47.27*** (2.73)	56.84*** (2.06)
Other Patient Delay IV	0.57*** (0.01)	0.43*** (0.01)	0.19*** (0.01)
Origin State Controls	✓	✓	✓
Unit Controls	✓	✓	✓
Patient Controls	✓	✓	✓
Time Controls	✓	✓	✓
Observations	14,474	14,474	14,474
Adjusted R^2	0.25	0.11	0.13
F-Statistic	3619.1***	1445.9***	1804.8***

*p<0.05; **p<0.01; ***p<0.001

Note: F-Statistic and significance from ANOVA F-tests comparing models with our instruments included with models without instruments.

BIBLIOGRAPHY

Z. Akşin, B. Ata, S. M. Emadi, and C.-L. Su. Structural estimation of callers' delay sensitivity in call centers. *Management Science*, 59(12):2727–2746, 2013.

T. J. Allen. Communication networks in R & D laboratories. *R&D Management*, 1(1): 14–21, 1970.

G. Allon, A. Federgruen, and M. Pierson. How much is a reduction of your customers' wait worth? an empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management*, 13(4):489–507, 2011.

R. Batt and S. Gallino. Finding a needle in a haystack: The effects of searching and learning on pick-worker performance. 2017.

R. J. Batt and C. Terwiesch. Waiting patiently: An empirical study of queue abandonment in an emergency department. *Management Science*, 61(1):39–59, 2015.

R. J. Batt and C. Terwiesch. Early task initiation and other load-adaptive mechanisms in the emergency department. *Management Science*, 63(11):3531–3551, 2016.

S. Benjaafar. Modeling and analysis of congestion in the design of facility layouts. *Management Science*, 48(5):679–704, 2002.

Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300, 1995.

S. J. Bosch, M. Apple, B. Hiltonen, E. Worden, Y. Lu, U. Nanda, and D. Kim. To see or not to see: Investigating the links between patient visibility and potential moderators affecting the patient experience. *Journal of Environmental Psychology*, 47:33–43, 2016.

R. W. Buell and M. I. Norton. The labor illusion: How operational transparency increases perceived value. *Management Science*, 57(9):1564–1579, 2011.

R. W. Buell, T. Kim, and C.-J. Tsay. Creating reciprocal value through operational transparency. *Management Science*, 63(6):1673–1695, 2016.

X. Cai, J. Gong, Y. Lu, and S. Zhong. Recover overnight? work interruption and worker productivity. *Management Science*, 2017.

E. Carrier, J. Khaldun, and R. Y. Hsia. Association between emergency department length of stay and rates of admission to inpatient and observation services. *JAMA internal medicine*, 174(11):1843–1846, 2014.

D. B. Chalfin, S. Trzeciak, A. Likourezos, B. M. Baumann, and R. P. Dellinger. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. *Critical Care Medicine*, 35(6):1477–1483, 2007.

C. W. Chan, V. F. Farias, and G. J. Escobar. The impact of delays on service times in the intensive care unit. *Management Science*, 63(7):2049–2072, 2016.

R. Choudhary, S. Bafna, Y. Heo, A. Hendrich, and M. Chow. A predictive model for computing the influence of space layouts on nurses' movement in hospital units. *Journal of Building Performance Simulation*, 3(3):171–184, 2010.

A. C. Craig, E. Garbarino, S. A. Heger, and R. Slonim. Waiting to give: Stated and revealed preferences. *Management Science*, 63(11):3672–3690, 2016.

L. Fahey, K. D. Lopez, J. Storfjell, and G. Keenan. Expanding potential of radiofrequency nurse call systems to measure nursing time in patient rooms. *Journal of Nursing Administration*, 43(5):302–307, 2013.

M. Fisher and A. Raman. *The new science of retailing: how analytics are transforming the supply chain and improving performance*. Harvard Business Review Press, 2010.

H. Ford, S. Crowther, and H. Ford. My life and work, 1922.

M. Freeman, N. Savva, and S. Scholtes. Gatekeepers at work: An empirical analysis of a maternity unit. *Management Science*, 63(10):3147–3167, 2017.

D. F. Gaieski, M. E. Mikkelsen, R. A. Band, J. M. Pines, R. Massone, F. F. Furia, F. S. Shofer, and M. Goyal. Impact of time to antibiotics on survival in patients with severe sepsis or septic shock in whom early goal-directed therapy was initiated in the emergency department. *Critical Care Medicine*, 38(4):1045–1053, 2010.

N. Gilboy, P. Tanabe, D. Travers, A. Rosenau, D. Eitel, et al. Emergency severity index, version 4: implementation handbook. *Rockville, MD: Agency for Healthcare Research and Quality*, pages 1–72, 2005.

I. Gurvich, K. O'Leary, L. Wang, and J. A. Van Mieghem. Collaboration, interruptions and changeover times: Model and empirical study of hospitalist processing times. 2017.

A. R. Hall, G. D. Rudebusch, and D. W. Wilcox. Judging instrument relevance in instrumental variables estimation. *International Economic Review*, pages 283–298, 1996.

A. Hendrich, M. Chow, B. A. Skierczynski, and Z. Lu. A 36-hospital time and motion study: how do medical-surgical nurses spend their time? *RCHE Publications*, page 50, 2008.

A. Hendrich, M. P. Chow, S. Bafna, R. Choudhary, Y. Heo, and B. A. Skierczynski. Unit-related factors that affect nursing time with patients: Spatial analysis of the time and motion study. *HERD: Health Environments Research & Design Journal*, 2(2):5–20, 2009.

S. S. Heragu, L. Du, R. J. Mantel, and P. C. Schuur. Mathematical model for warehouse design and product allocation. *International Journal of Production Research*, 43(2):327–338, 2005.

A. Herring, A. Wilper, D. U. Himmelstein, S. Woolhandler, J. A. Espinola, D. F. Brown, and C. A. Camargo Jr. Increasing length of stay among adult visits to us emergency departments, 2001–2005. *Academic Emergency Medicine*, 16(7):609–616, 2009.

J. C. Hollingsworth, C. D. Chisholm, B. K. Giles, W. H. Cordell, and D. R. Nelson. How do physicians and nurses spend their time in the emergency department? *Annals of Emergency Medicine*, 31(1):87–91, 1998.

Y. Hua, F. Becker, T. Wurmser, J. Bliss-Holtz, and C. Hedges. Effects of nursing unit spatial layout on nursing team communication patterns, quality of care, and patient safety. *HERD: Health Environments Research & Design Journal*, 6(1):8–38, 2012.

P. J. Huber. Robust statistics. In *International Encyclopedia of Statistical Science*, pages 1248–1251. Springer, 2011.

P. J. Huber et al. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.

S. K. Hui, P. S. Fader, and E. T. Bradlow. Path data in marketing: An integrative framework and prospectus for model building. *Marketing Science*, 28(2):320–335, 2009.

S. K. Hui, J. J. Inman, Y. Huang, and J. Suher. The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing*, 77(2):1–16, 2013.

M. R. Ibanez, J. R. Clark, R. S. Huckman, and B. R. Staats. Discretionary task ordering: Queue management in radiological services. *Management Science*, 2017.

J. A. B. Jaeker and A. L. Tucker. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Science*, 63(4):1042–1062, 2017.

A. B. Jena, N. C. Mann, L. N. Wedlund, and A. Olenski. Delays in emergency care and mortality during major us marathons. *New England Journal of Medicine*, 376(15):1441–1450, 2017.

B. J. Kalisch and M. Aebersold. Interruptions and multitasking in nursing care. *The joint commission journal on quality and patient safety*, 36(3):126–132, 2010.

D. S. KC. Does multitasking improve performance? evidence from the emergency department. *Manufacturing & Service Operations Management*, 16(2):168–183, 2013.

D. S. Kc. Does multitasking improve performance? evidence from the emergency department. *Manufacturing & Service Operations Management*, 16(2):168–183, 2013.

D. S. KC and C. Terwiesch. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science*, 55(9):1486–1498, 2009.

D. S. Kc and C. Terwiesch. The effects of focus on performance: Evidence from california hospitals. *Management Science*, 57(11):1897–1912, 2011.

R. C. Larson. Or forum—perspectives on queues: Social justice and the psychology of queueing. *Operations research*, 35(6):895–905, 1987.

Y. Lu, A. Musalem, M. Olivares, and A. Schilkrut. Measuring the effect of queues on customer purchases. *Management Science*, 59(8):1743–1763, 2013.

D. H. Maister et al. *The psychology of waiting lines*. Citeseer, 1984.

D. I. McIsaac, K. Abdulla, H. Yang, S. Sundaresan, P. Doering, S. G. Vaswani, K. Thavorn, and A. J. Forster. Association of delay of urgent or emergency surgery with mortality and use of health care resources: a propensity score–matched observational cohort study. *Canadian Medical Association Journal*, 189(27):E905–E912, 2017.

A. Musalem, M. Olivares, and A. Schilkrut. Retail in high definition: Monitoring customer assistance through video analytics. 2016.

T. Ohno. *Toyota production system: beyond large-scale production*. CRC Press, 1988.

R. Oliva and J. D. Sterman. Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, 47(7):894–914, 2001.

J. M. Pines and J. E. Hollander. Emergency department crowding is associated with poor care for patients with severe pain. *Annals of emergency medicine*, 51(1):1–5, 2008.

J. M. Pines, A. R. Localio, J. E. Hollander, W. G. Baxt, H. Lee, C. Phillips, and J. P. Metlay. The impact of emergency department crowding measures on time to antibiotics for patients with community-acquired pneumonia. *Annals of emergency medicine*, 50(5): 510–516, 2007.

J. M. Pines, S. Iyer, M. Disbot, J. E. Hollander, F. S. Shofer, and E. M. Datner. The effect of emergency department crowding on patient satisfaction for admitted patients. *Academic Emergency Medicine*, 15(9):825–831, 2008.

E. G. Poon, C. A. Keohane, C. S. Yoon, M. Ditmore, A. Bane, O. Levtzion-Korach, T. Moniz, J. M. Rothschild, A. B. Kachalia, J. Hayes, et al. Effect of bar-code technology on the safety of medication administration. *New England Journal of Medicine*, 362(18): 1698–1707, 2010.

J. C. Prentice and S. D. Pizer. Delayed access to health care and mortality. *Health services research*, 42(2):644–662, 2007.

J. Reiling, R. G. Hughes, and M. R. Murphy. The impact of facility design on patient safety. 2008.

B. D. Ripley. *Modern Applied Statistics with S*. Springer, 2002.

K.-J. Roodbergen. *Layout and routing methods for warehouses*. Number EPS-2001-004-LIS. 2001.

M. J. Rosenblatt. The dynamics of plant layout. *Management Science*, 32(1):76–86, 1986.

L. M. Skaugset, S. Farrell, M. Carney, M. Wolff, S. A. Santen, M. Perry, and S. J. Cico. Can you multitask? evidence and limitations of task switching and multitasking in emergency medicine. *Annals of emergency medicine*, 68(2):189–195, 2016.

H. Song, A. L. Tucker, and K. L. Murrell. The diseconomies of queue pooling: An empirical investigation of emergency department length of stay. *Management Science*, 61(12):3032–3053, 2015a.

H. Song, A. L. Tucker, and K. L. Murrell. The diseconomies of queue pooling: An empirical investigation of emergency department length of stay. *Management Science*, 61(12):3032–3053, 2015b.

H. Song, A. L. Tucker, K. L. Murrell, and D. R. Vinson. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science*, 2017.

H. Song, A. Tucker, R. Graue, S. Moravick, and J. Yang. Capacity pooling in hospitals: The hidden consequences of off-service placement. *Working Paper*, 2018.

H. Soriano-Meier, P. L. Forrester, S. Markose, and J. Arturo Garza-Reyes. The role of the physical layout in the implementation of lean management initiatives. *International Journal of Lean Six Sigma*, 2(3):254–269, 2011.

B. R. Staats, H. Dai, D. Hofmann, and K. L. Milkman. Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science*, 63(5):1563–1585, 2017.

T. F. Tan and S. Netessine. When does the devil make work? an empirical study of the impact of workload on worker productivity. *Management Science*, 60(6):1574–1593, 2014.

J. A. Tompkins, J. A. White, Y. A. Bozer, and J. M. A. Tanchoco. *Facilities planning*. John Wiley & Sons, 2010.

A. L. Tucker. The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management*, 22(2):151–169, 2004.

H.-M. Tzeng, D. L. Ronis, and C.-Y. Yin. Relationship of actual response time to call lights and patient satisfaction at 4 us hospitals. *Journal of nursing care quality*, 27(2):E1–E8, 2012.

T. E. Vollmann and E. S. Buffa. The facilities layout problem in perspective. *Management Science*, 12(10):B–450, 1966.

J. M. Welton, M. Decker, J. Adam, and L. Zone-Smith. How far do nurses walk? *Medsurg Nursing*, 15(4):213, 2006.

J. Wooldridge and G. Imbens. Lecture notes 6, summer 2007, 2007.

J. M. Wooldridge. *Econometric analysis of cross section and panel data.* MIT press, 2010.

Q. Yu, G. Allon, and A. Bassamboo. How do delay announcements shape customer behavior? an empirical study. *Management Science*, 63(1):1–20, 2016.