Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Does Multitasking Improve Performance? Evidence from the Emergency Department

Diwas Singh KC

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 16, No. 2, Spring 2014, pp. 168–183
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2013.0464


# Does Multitasking Improve Performance? Evidence from the Emergency Department

Diwas Singh KC
Goizueta Business School, Emory University, Atlanta, Georgia 30322, diwas.kc@emory.edu

This paper examines the effect of multitasking on overall worker performance, as measured by processing time, throughput rate, and output quality using microlevel operational data from the field. Specifically, we study the multitasking behavior of physicians in a busy hospital emergency department (ED). By drawing on recent findings in the experimental psychology literature and the nascent work in cognitive neuroscience, we develop several hypotheses for the effect of multitasking on worker performance. We first examine how multitasking affects a physician's processing time. We find that the total time taken to discharge a given number of patients has a U-shaped response to the level of physician multitasking; that is, multitasking initially helps to reduce the time taken, but only up to a certain threshold level, after which it increases in the level of multitasking. In addition, multitasking significantly impacts quality of care. Although lower levels of multitasking are associated with improved quality of care, at higher levels, additional multitasking leads to a smaller number of detected diagnoses and an increased likelihood of a 24-hour revisit rate to the ED. These findings have important implications for the design and organization of work in general and for the delivery of critical care in particular.

*Keywords*: multitasking; productivity; quality; emergency department; capacity planning
*History*: Received: July 7, 2011; accepted: August 24, 2013. Published online in *Articles in Advance* November 27, 2013.



## 1. Introduction

> There is time enough for everything, in the course of the day, if you do but one thing at once; but there is not time enough in the year, if you will do two things at a time.
>
> —Philip Dormer Stanhope, 4th Earl of Chesterfield, *Letters to His Son* (1774)

In the workplace, people frequently perform multiple tasks over any given interval of time, often alternating between activities that demand their time and attention. The term multitasking (or time sharing) was first used to describe the sharing of computing processor capacity among a number of distinct jobs (e.g., Denning 1971). In the workplace today, human multitasking is a natural response to the increased demand from a growing number of competing activities. Although the 18th century English nobleman Lord Chesterfield advised that multitasking would negatively impact performance, many today would argue that multitasking has made them more productive.

An apparent source of productivity gains from multitasking is the ability to utilize idle time between tasks. By attending to several different tasks over a given interval of time, a multitasker can reduce the time spent waiting between jobs, particularly when the individual task completion times are uncertain. Switching to a new task rather than idly waiting on a pending task can thus increase worker utilization and improve overall productivity. In fact, most multitaskers state that their ability to switch tasks has made them more productive, with some multitaskers grossly overestimating the purported benefits (Pennebaker 2009, Ophir et al. 2009, Vega et al. 2008).

However, recent research in experimental psychology finds limits to human multitasking, positing that it lowers productivity and quality. For example, Gladstones et al. (1989) and Pashler (1994) find that experimental subjects are unable to focus their attention when asked to multitask. Insufficient attention to the tasks at hand in turn increases the likelihood of errors. Given the human brain's inability to focus attention on two or more tasks simultaneously, many researchers (Hallowell 2005, Mayer and Moreno 2003, Pashler 1994) believe that multitasked activities are eventually reduced to a set of sequential cognitive tasks. Consequently, excessive multitasking can lead to a cognitive bottleneck, and taking on several tasks concurrently may actually decrease productivity. Furthermore, multitasking is associated with a "warm-up" or setup cost that is incurred each time an individual switches tasks. In other words, the brain is compelled to restart and refocus (Rubinstein et al. 2001) between tasks. This recurring setup reduces effective processing capacity, and excessive multitasking can thus be detrimental to individual productivity. Recent advances in functional magnetic resonance imaging (fMRI) and cognitive neuroscience (e.g., Hoshino et al. 2006) corroborate these findings and point to an underlying physiological basis for the throughput losses that arise from the recurring setups.

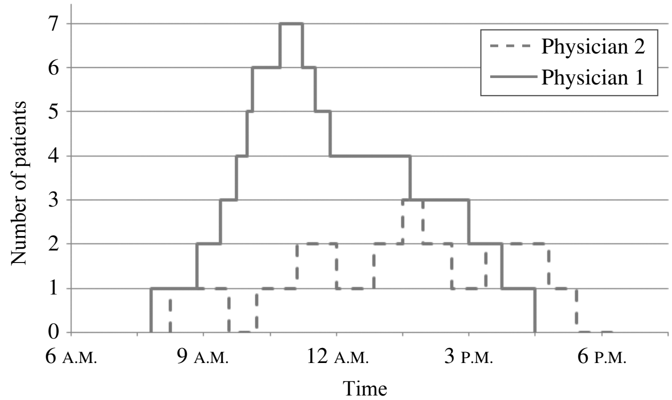
These studies thus far have primarily explored the performance implications of multitasking for relatively straightforward cognitive tasks in experimental settings (Pashler 1994); the implications of multitasking for more complex day-to-day business functions involving decision making under uncertainty are likely to be more significant. Given the natural tendency for workers to multitask when confronted with an increase in workload (Pennebaker 2009), there may be significant productivity and quality implications of multitasking in the workplace and for the operational performance of firms. The implications for a knowledge economy, where attention and focus are significant drivers of productivity and quality, may be particularly severe.

The study of productivity has a long tradition in the operations management literature (Smith 1776, Taylor 1914). Yet the impact of multitasking on productivity and quality remains relatively unexplored. In particular, there are few studies that have examined the operational effects of multitasking using transactional data from a field study, involving activities over longer periods of time. Does multitasking make you more productive? Does it improve the quality of your work?

To address these questions, we collect patient flow and clinical data from an East Coast metropolitan hospital's emergency department (ED) to examine the multitasking behavior of the ED physicians and the resulting effect on their productivity and output quality. The ED is an ideal setting for this study for a number of reasons. It is a high-paced work environment where multitasking is standard business practice, and important decisions are frequently made under pressing time and operating constraints. For this reason, the ED has often been called a natural lab for the study of errors (Croskerry et al. 2004) because of the workload, high levels of uncertainty, and multiple moving parts. From a behavioral standpoint, decision makers are under pressure to provide timely medical evaluations and appropriate delivery of care. ED patients are at the risk of dying, so the need to better understand the performance implications of multitasking is critical. Finally, the ED provides natural variability in multitasking over the course of the year for the same physician, which allows us to identify performance changes.

We follow prior work in the experimental psychology literature that defines multitasking as the carrying

**Figure 1** **Multitasking by Two Physicians**




out of two or more activities or responsibilities concurrently by an individual; in the context of this study, we define multitasking as the number of distinct patients cared for by a physician at a given point in time. As a motivating preview to our findings, consider Figure 1, which shows the activities of two different physicians over the course of the same day. We find varying degrees of multitasking; physician 1 handles a up to seven patients during her shift, whereas physician 2 consistently handles a fewer number of patients simultaneously, peaking at a maximum of three patients. Based on over 146,211 patient discharges under varying levels of multitasking over a period of three years, we examine several key dimensions of physician performance and offer the following contributions.

We first explore the consequences of multitasking for the productivity of the individual physician. Multitasking is fundamentally about *how* a given amount of work is sequenced and processed. As such, differing levels of multitasking can influence the total time taken to complete the work. To examine this aspect of worker performance, we introduce the term *busy period*, which, for a physician working continuously to discharge a given set of patients, simply denotes the total time taken from the moment that the first patient is picked up to the instant that the last patient is discharged. We find that the busy period has a U-shaped response to the level of physician multitasking; that is, multitasking initially helps to reduce the time taken to process the patients, but only up to a certain threshold level. After a critical threshold is exceeded, productivity decreases and the busy period starts to increase. That is, multitasking improves productivity as long as the losses due to nonvalue-added activities (e.g., setups and other costs) are exceeded by gains from reducing idle time. We find that multitasking has significant implications for quality of care; multitasking has an inverted U-shaped effect on the quality outcome. Although some multitasking is beneficial, excessive multitasking results in a smaller number of detected diagnoses for the patient. Similarly, although beneficial at low levels of multitasking, at higher levels, incremental multitasking incurs an increased number of patient revisits to the ED. Collectively, these two results suggest that cognitive limitations and time constraints due to multitasking result in the quality lapses.

These findings offer important new insights into the design and organization of work. Although some overlap of tasks may help worker performance, there are limits to multitasking. In the context of the ED, selecting the appropriate level of multitasking allows the physician to complete work faster, and with improved output quality. Our findings suggest that more effective assignment of patients can thus improve ED physician performance.

The rest of this paper is organized as follows. In §2 we review the related literature. In §3 we present the clinical context and describe our data. In §4 we develop the hypotheses, and in §5 we provide the empirical model specifications. In §6 we present our findings, and in §7 we conclude the paper.

## 2. Literature Review

The study of productivity has been of great interest in operations management, with some of the early scientific and quantitatively rigorous studies dating to Taylor's (1914) analysis of worker output in an assembly line. Recent work has further examined the microfoundations and behavioral underpinnings of productivity. For example, Schultz et al. (1999) measure worker productivity in serial systems and find that the processing rate of workers is dependent on their workload. Furthermore, workers adjust their processing speed based on available capacity (Hasija et al. 2010, KC and Terwiesch 2009) or on the perceived impact of their effort on system-level throughput (Powell and Schultz 2004). Schultz et al. (2003) conduct a set of experiments to study worker flexibility in a low-inventory serial assembly system. They find that worker flexibility can decrease productivity, either because of a break in the worker's rhythm or because the worker forgets task-specific knowledge during the switch.

Various organizational and behavioral factors can have a significant bearing on worker-level productivity (Gino and Pisano 2008, Bendoly et al. 2009). For example, Staats and Gino (2012) examine microlevel productivity data in which workers complete one task at a time. They find that productivity on the current task is significantly improved if a worker has had a related experience earlier in the day. However, this benefit decreases with time. In addition, a sequence of a variety of tasks (as opposed to routine and repetitive tasks) performed earlier in the day has a positive impact on current productivity. Similarly, Narayanan et al. (2009) have found that limited amounts of varied experience (i.e., tasks related, but not identical, to the focal task) may improve productivity.

There is growing interest for exploring the tradeoff between output quality and speed of service. For example, Oliva and Sterman (2001) find that speedup can reduce both quality and revenue. Similarly, KC and Terwiesch (2012) find that a high level of occupancy in the cardiac intensive care unit (ICU) leads to an early discharge of patients, but also increases the likelihood of a revisit to the ICU, incurring further capacity usage. Tucker and Edmondson (2003) also identify numerous instances of rework by nurses. Because of the demands on their time, nurses often try to patch problems quickly in the short term. Rather than identifying the underlying causes to prevent problems from occurring again, overworked nurses instead end up adopting myopic policies. Similarly, Needleman et al. (2006) and Aiken et al. (2002) find that increased workload for care providers is associated with lower quality of care.

Although this prior body of work has examined the effects of workload over prolonged periods of time, few studies have investigated the effect of multitasking. Aral et al. (2012) recently examined information technology (IT) worker productivity at an executive recruiting firm and found that multitasking is associated with increased project output but diminishing marginal returns. However, in contrast to IT projects, the care of ED patients involves tasks completed over a shorter period of time. Therefore, the cognitive limitations, particularly due to working memory (Baddeley and Hitch 1994), are likely to be more important. In our analysis, we control for workload while varying the level of multitasking for individual workers over time to examine the effects on task completion times as well as quality of output. Chisholm et al. (2000) and (2001) document the presence of multitasking in the ED but do not examine the impact on productivity and quality.

Experimental psychologists have identified the limitations of multitasking among human subjects in laboratory settings. For example Gladstones et al. (1989) and Pashler (1989) found that individuals demonstrate severe interference when asked to perform simple tasks at the same time. Hallowell (2005) describes multitasking as a "mythical activity." A key insight from this stream of work is that multitasking incurs a productivity loss due to task switching. When presented with new information, the brain is forced to pause and refocus. Thus, setup (or warm-up) cost is incurred each time a worker switches between tasks. Such recurring setups reduce effective capacity for the worker.

More recent advances in fMRI have allowed scientists to observe the neural activity of subjects and to examine their cognitive limitations in great detail. For example, Hoshino et al. (2006), who recently examined the brains of multitasking individuals under fMRI, concluded that there is a neurological basis for the decline in productivity of multitaskers. Charron and Koechlin (2010) found that people slow down and make many more mistakes when asked to multitask in the lab. This research demonstrates that multitasking is in fact sequential processing. Given the brain's inability to perform high-level cognitive functions simultaneously, a set of multitasked activities is broken down into a sequence of cognitive functions. Excessive toggling leads in fact to a cognitive bottleneck.

The results from these studies are compelling. Multitasking hinders productivity in three specific ways. First, there is the setup cost associated with each switch or toggle between tasks. Second, multitaskers have difficulty sifting through all of the available information to focus on the information that is relevant to the task at hand. Finally, multitaskers often have difficulty keeping the relevant information well organized for short-term memory retrieval. The net impact is an adverse effect not only on the number of tasks completed, but also on the quality of the work.

Despite overwhelming laboratory evidence and pervasive multitasking in the service economy among workers engaged in high levels of cognitive decision making, there are few empirical studies of multitasking from the field. In terms of structure, processor sharing (Parekh and Gallager 1993) comes closest to capturing the division of a limited amount of attention (or service time) across competing activities. However, the processor-sharing literature is normative in its scope, outlining how a processor's capacity should be allocated among a number of competing requests. Similarly, in the economics literature, Holmstrom and Milgrom (1991) examine the design of incentive contracts in a principal–agent framework with the assumption that multitasking leads to performance degradation. However, few studies have empirically examined the performance degradation.

## 3. Empirical Setting

In this section, we first describe the process flow, followed by a description of the data.

### 3.1. Clinical Context and Process Flow

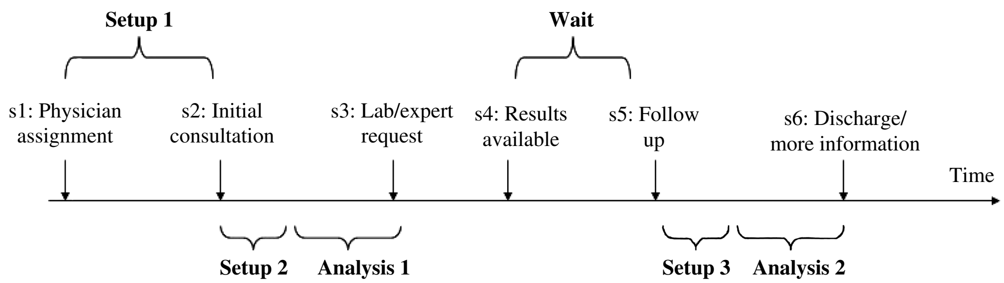
The ED that we study is part of an East Coast metropolitan hospital and treats a large volume of patients each year. Like other EDs across the United States (Institute of Medicine 2006), this ED is grappling with physician shortages and increased patient volume, resulting in a high level of physician multitasking. An ED physician in (Salamon 2008, p. 18) describes the job requirements as follows:

> They need to be able to multitask. In most medicine you proceed diagnostically and then therapeutically at the same time. You have patients coming to you with acute pain; they're sick, and we start treatment not knowing what we're treating. We'll give you an aspirin because we think this may be your heart failure before I figure out whether it really is your heart. You have to be able to do that with multiple patients at the same time. You need to be able to maintain situational awareness.

Consider the flow process for patient $j$ and the role of ED physicians in the care process. Upon arrival, patient $j$ is seen by the triage nurse, who evaluates the patient's condition, determines the triage severity level, and creates an electronic record and a physical folder for the patient. The patient's electronic record, which is color coded for one of three severity levels (some ED's use up to five severity levels), is then placed in a virtual queue to be processed by a physician, who is specialized in emergency medicine. ED physicians use a computer terminal for monitoring the queue, picking up new patients and reviewing their clinical information (including triage notes and historical medical records), ordering and viewing diagnostic tests, and providing updates to the patient's record. Physicians continuously monitor the queue to pick up new patients based on triage severity level and arrival time when not attending to their existing patients. Specifically, triage level 1 (most severe) patients are always picked up before level 2 patients, who receive precedence over level 3 patients. For each of the three severity levels, patients are sorted by time of triage. Physicians generally pick up a new patient from the top of their category queue, which ensures that for a given severity level, the earliest arrival is picked up first.

The service process begins once patient $j$ is picked up by a physician. As illustrated in Figure 2, the patient flow involves several key events. The physician first performs a behind-the-scenes evaluation of the patient's medical record (including reviewing previous history for prior medications, allergies and treatments, notes by the triage nurse, and possible need for expert opinion) prior to actually meeting the patient face-to-face. Consequently, the physician incurs an initial setup (setup 1) while the patient waits for the physician. The initial consultation includes an additional setup phase (setup 2) where the physician gets acquainted with the details of the patient's case. This typically includes introductions, evaluating the patient's overall disposition, gathering relevant information, and verifying the patient's medical history. The physician's encounter involves questioning the patient and family members, examining the patient in detail, diagnosing the patient, and determining the type and scope of care needed. Following the face-to-face evaluation, the physician records her observations and formulates a possible course of

**Figure 2** **Patient Flow Process**




treatment. We call this the analysis stage (analysis 1). If the patient's condition is not severe, he may be discharged. In most cases, the initial consultation is followed by requests for expert opinion or diagnostic tests (e.g., blood tests, CAT (computerized axial tomography) scans, MRI scans, or X-rays). From a process flow perspective, waiting for the expertise of a specialist is identical to waiting for the diagnostic results. If the physician is not immediately available when the test results (or expert opinion) become available, the patient incurs an additional wait (*wait*).

While waiting for a specialist or the results from diagnostic tests, the physician has available time to either pick up a new patient or attend to one of her existing patients. When patient $j$'s lab tests become available and the specialists provide their expert opinion, the physician can then return to patient $j$. The physician often spends time reacquainting herself with the specifics of the case. In other words, there is an additional setup phase (setup 3). After evaluating the new information, the patient is either scheduled for additional tests and consultations or is discharged from the ED after appropriate treatment and instructions. Although a specialist may be called to the ED to evaluate a patient, the primary physician is ultimately responsible for treatment and discharge decisions. Finally, we note that in the case of more complex patients, there may be more cycles of wait and setup than shown in Figure 2.

### 3.2. Description of Data

We assembled our data from the emergency department for 12 consecutive quarters beginning in October 2004 and ending in September 2007. Our data collection was supplemented with field observations and interviews with clinical staff. In addition, we shadowed ED physicians for three different shifts. These observations helped to shape our hypotheses and to map out the patient flow process.

Our data include information on each patient visit, treatments, and outcomes. Most importantly, we observe the unique patient–physician pair for each patient in the ED, meaning we know the specific physician who treated an individual patient. This information thus allows us to examine the service encounter of each individual patient, as well as the productivity of each individual physician.

The emergency department has a total of 14 licensed beds. As with a large number of other EDs, this hospital has responded to bed shortages by adding makeshift hallway beds. During the course of our site visits, we observed five hallway beds. However, the number of hallway beds has varied over the course of the three-year study period. A total of 146,221 patients were seen in this emergency room over this period of study. Of these, 286 ($<0.2\%$) patients did not have accurate physician information and were dropped from the sample; this left a usable sample of 145,935 patients. Our data set includes various measures of microlevel productivity observations. In particular, we observe the time a patient is picked up by a physician and when he is discharged. The time that a patient is under the care of a given physician (physician assignment period) is simply the difference between the instant the patient is assigned and the instant that he is discharged from the care of the ED physician. Table 1 displays the patient-level summary statistics. We find that the average patient's physician assignment period is 2.07 hours. However, the standard deviation is quite large (2.94 hours), reflecting the underlying heterogeneity in the patient population. For each patient discharged from the ED we observe various demographic factors, including age,

**Table 1** **Summary Statistics**

| Variable | Mean | Std. dev. | Bottom quartile | Median | Top quartile |
|---|---|---|---|---|---|
| *Physician assignment period* (hours) | 2.07 | 2.94 | 0.48 | 1.08 | 2.73 |
| *Age* | 34.3 | 22.6 | 20 | 32 | 47 |
| *Gender* (female = 1) | 0.527 | 0.499 | 0 | 1 | 1 |
| *Emergency status* | 0.151 | 0.358 | 0 | 0 | 0 |
| *Relative value unit* (workload) | 6.44 | 9.10 | 0.10 | 2.47 | 11.02 |
| *Revisit probability within 24 hours* | 0.0253 | 0.157 | 0 | 0 | 1 |
| *Number of diagnoses* | 1.28 | 1.31 | 0 | 1 | 2 |

*Note.* $N = 145{,}935$ patients between October 2004 and September 2007.

**Table 2 Hospital Summary Statistics**

| Variable | Mean | Std. dev. | Bottom quartile | Median | Top quartile |
|---|---|---|---|---|---|
| *System load* | 14.3 | 5.4 | 10 | 14 | 18 |
| *Multitasking* | 3.77 | 2.05 | 2 | 3 | 5 |
| *Number of physician staffed in four-hour window* | 4.28 | 1.47 | 3 | 4 | 5 |
| *Busy period duration* (hour) | 4.97 | 5.42 | 0.8 | 2.32 | 7.8 |
| *Busy period number of patients* | 7.46 | 8.89 | 2.0 | 5 | 10 |
| *Monthly volume* | 4,053 | 220 | 3,898 | 4,063 | 4,258 |
| *Number of busy periods* | 19,562 | — | | — | |
| *Number of beds* | 14 | — | | — | |

*Notes.* The number of beds does not include temporary hallway beds. Statistics are based on 145,935 patients discharged by 26 physicians.

gender, and race. The average patient is 34.4 years of age. Slightly more females (ratio = 0.527) visit the ED than males.

**3.2.1. Outcomes.** The raw variables in our data set allow us to construct our key outcome measures. In particular, the data set includes the set of diagnoses made for each patient visit, which allows us to measure the number of diagnoses recorded by the physician. A unique patient identifier allows us to determine any incidence of revisit to the same ED within 24 hours. The arrival and discharge time stamps allow us to construct the busy periods. Recall that for a given set of patients, the busy period is simply the total time from the moment that the first patient is picked up to the time that the last patient is discharged. Table 2 provides summary statistics for these variables. We find that the 24-hour revisit rate for any complaint is 2.53%. We see that during the period of study, 19,562 busy periods were completed; in an average busy period, 7.46 patients were discharged.

**3.2.2. Explanatory Variables.** The microlevel operational data allow us to construct the time series for the number of patients seen by a given physician at a specific point in time. Specifically, once a patient is picked up by physician $i$ at time $t$, the level of multitasking ($M_i(t)$) for that physician is immediately increased by 1. Similarly, at the instant a patient is discharged, the level of multitasking decreases by 1. In other words, the value of $M_i(t)$ remains constant until either a pickup or discharge event occurs as illustrated in Figure 1 for two distinct physicians during a day.

The arrival and discharge times also allow us to construct the load in the ED at specific points in time. The system load (or occupancy) in the ED at time $t+dt$ is given by the relationship $SystemLoad(t+dt) = SystemLoad(t) + Arrivals_{t,\,t+dt} - Discharges_{t,\,t+dt}$, where $Arrivals_{t,\,t+dt}$ and $Discharges_{t,\,t+dt}$ denote the number of arrivals and departures between $t$ and $t+dt$, respectively. This simple formula, in conjunction with the microlevel time-stamp data, allows us to construct minute-by-minute *SystemLoad* for the ED during the entire period of study. The above formula, however, requires us to know the occupancy at the beginning (i.e., at time 0) or at a specific point in time. To estimate the occupancy at a given moment, we consider all the patients in the system at time $t$ and simply count all the arrivals that occurred before time $t$ and all the departures that occurred after time $t$. The difference between these two measures (*Discharges* − *Arrivals*) should yield the exact occupancy at time $t+dt$, provided that $t$ is larger than the longest possible patient stay. We drop the first two days and the last two days of observations to deal with this truncation issue.

To control for the workload associated with the discharge of a given patient, we use the relative value unit (RVU) designation for that patient's visit. Specifically, the RVU is a measure of the physician's work associated with the care of a given patient and takes into account the physician's time, physical and mental effort, as well as clinical and technical judgment and skill. RVUs are tied to the primary diagnosis code for the patient and are currently used by payers, including Medicare, to reimburse physicians for services rendered.[1] We also observe each patient's major diagnostic category (MDC). Patients assigned to a given MDC category generally fall under a specific medical specialty, and the MDC classification allows us to group similar patients into a manageable number of meaningful categories. Therefore, the number of distinct patient MDC categories provides a measure of the task variety concurrently managed by the physician. A list of the MDC categories, as well as additional information and references on MDC categories, is provided in the online supplement (available at http://dx.doi.org/10.1287/msom.2013.0464). To further account for task complexity, we collect the triage severity level determined by the triage nurse at the time of admission. Factoring in these patient characteristics enables us to account for heterogeneity in treatments and outcomes of a range of patients in the ED.

Table 2 displays the summary statistics for the ED system load and physician levels of multitasking. We find that the average level of physician multitasking for patients in the sample is 3.77. There are, on average, 14.3 patients in the ED. We also find that,

[1] Detailed information on the construction of RVUs can be obtained from the Centers for Medicare and Medicaid Services at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/Downloads/RVU12D.zip.

**Table 3** **Summary Statistics: Longitudinal Variation in Multitasking Among Physicians**

| Variable | Mean | Std. dev. | Median | 1% | 99% |
|---|---|---|---|---|---|
| *Top third* | 4.06 | 2.07 | 4 | 1 | 10 |
| *Middle third* | 3.74 | 2.05 | 3 | 1 | 10 |
| *Bottom third* | 3.07 | 1.79 | 3 | 1 | 8 |

*Notes.* The mean level of physician multitasking during the course of the study was used to first assign physicians into thirds. The reported statistics were then produced for the set of physicians in the respective third.

on average, 4.28 physicians are available in the ED during an average four-hour period. Table 3 displays the variation in multitasking among physicians. The physicians with the highest (lowest) average levels of multitasking are represented in the top (bottom) third. We find that the physicians in the top third (highest average multitasking) also tend to have greater levels of intertemporal variation in the levels of multitasking, as evidenced by the high standard deviation (2.07). Physicians in the bottom third tend to maintain more consistent levels of multitasking over time.

## 4. Hypotheses Development

In the discussions below, we follow the prior literature that has defined multitasking as the number of distinct activities or responsibilities an agent concurrently manages. For example, Aral et al. (2012) and Chisholm et al. (2000) define multitasking as the number of projects concurrently undertaken by a given worker and the number of patients concurrently seen by a given physician, respectively. Here, the level of multitasking ($M_i(t)$) is defined similarly as the number of patients that physician $i$ has under her care at time $t$, as described in §3.2.2.

### 4.1. Impact on Physician's Productivity

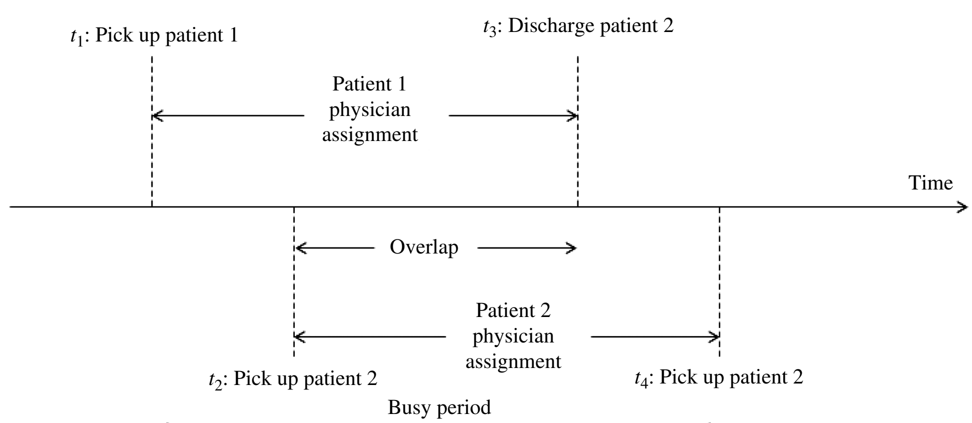
We first consider the effect of multitasking on the overall time taken by physician $i$ to discharge a given number of patients under her care. To motivate our discussion, we present an example of a typical scenario encountered by ED physicians. Suppose that physician $i$ ends up caring for and discharging two patients in the ED; patient 1 presents with a broken collar bone and patient 2 presents with symptoms of abdominal bleeding. Our objective is to understand how temporal overlap in the treatment for these two patients impacts the time taken by the physician to complete the workload for these two patients. Figure 3 illustrates the work flow for physician $i$.

Physician $i$ picks up patients 1 and 2 at times $t_1$ and $t_2$ and discharges them at $t_3$ and $t_4$, respectively. Therefore, $t_4 - t_1$ represents the busy period, or total time taken for her to discharge these two patients, and $t_3 - t_2$ captures the extent of overlap of activities associated with patients 1 and 2. Figure 3 raises an important question about the design and organization of work in general and in the critical care setting in particular. What is the optimal way to organize and sequence a given amount of work in order to complete it faster? Should one pursue multiple activities in parallel or focus on a single activity until it is completed? In short, how does multitasking between activities for patients 1 and 2 affect the busy period? We note that varying the degree of overlap does not directly impact the aggregate work for physician $i$, since all of the activities required to diagnose, treat, and discharge both patients 1 and 2 have to be completed. Therefore, by construction, a busy period holds workload constant while varying the extent of task overlap.

To answer the question above, we consider the trade-off between two competing effects—the creation of nonvalue-added activities and the utilization of idle time between patients—due to multitasking.

In many processes, tasks are often separated by periods of inactivity (or idle time, or waiting for some dependent task to complete). In the ED, as described

**Figure 3** **Physicians Multitasking**



in Figure 2, a patient's flow incurs periods of inactivity or wait. But by managing multiple patients concurrently, a physician can minimize or eliminate the idle time by simply switching to a pending activity for a different patient. For example, from Figure 2, physician $i$ can pick up a new patient (patient 2) at time $s_3$, rather than wait idly for a previous patient's (patient 1) test results. By switching between patients, physician $i$ minimizes her idle time, and allocates her time more effectively towards activities that help to discharge the other patients under her care.

On the other hand, excessive multitasking may generate additional nonvalue-added activities. In particular, multitasking incurs a switchover (setup 3) whenever the physician revisits a patient under her care (e.g., Pashler 1994). In explaining the role of setups on performance, Rubinstein et al. (2001) suggest that the human "executive control" processes involve two distinct stages. Stage one involves "goal shifting" (individual now does activity A instead of B), whereas stage two involves "rule activation" (turn off rules for B and turn on rules for A). Although both of these stages help people switch between tasks (even without their knowing it), the switching costs can impact performance. For example, the physician may have to become reacquainted with the specifics of the case, reevaluate the patient's notes, restart a test order from scratch, and gauge any changes in the patient's disposition. Higher levels of multitasking thus incur excessive toggling and higher setups.

In addition, overseeing multiple patients concurrently makes the physician more liable to be interrupted by nurses, specialists, or even patients and family members. In our study site, during a particular four-hour period of field observation that involved shadowing an ED physician, we noted 23 instances of physician interruption from nurses, clinical specialists, and other medical providers. The greater the number of patients under her care, the greater the frequency of interruption. For example, Chisholm et al. (2000) conducted a time-motion study and found that physicians are constantly interrupted by other care providers in the ED. Such interruptions often require the physician to put patient $j$ "on hold" temporarily for more pressing demands. This ultimately incurs further setups when care for patient $j$ is resumed; in many cases, the task has to be restarted from the beginning. Multitasking can also lead to another source of delays, which we call "coordination waits." As an illustration from Figure 2, consider patient $j$, whose test results have come in at time $s_4$. The physician, however, may be busy attending to other patients and can only get to patient $j$ at time $s_5$. As the number of patients under the care of physician $i$ increases, the likelihood of increased coordination waits (*wait*) for patient $j$ increases. Collectively, the resulting disruption of work flow can increase the busy period.

The net effect of multitasking on the busy period is therefore determined by the relative magnitudes of these two opposing sets of forces. We postulate that at lower levels of multitasking, more idle time is incurred. In the extreme case of managing a single patient, the physician has to wait after test orders have been submitted, before anything else can be done on the patient. Multitasking offers the opportunity to significantly reduce physician idle time and hence reduce the busy period. Once the physician is already operating at high levels of multitasking, idle times have been reduced or eliminated, and sources of waste associated with high levels of multitasking begin to play a more dominant role. In particular, the physician encounters greater setups, interruptions, and coordination losses. With excessive multitasking, any minor gains in idle time minimization are more than offset by the costs of task switching. Based on these discussions, we postulate the following hypothesis:

Hypothesis 1A. *A busy period* (*time to discharge a given number of patients*) *has a U-shaped response to multitasking*.

The preceding discussions also imply that multitasking has an inverted U-shaped effect on the physician's throughput rate. At low levels, multitasking enables the physician to reduce her idle time and to spend time on value-added activities such as diagnosis, testing, and treatment, which collectively contribute to a faster rate of discharge of patients. At high levels of multitasking, nonvalue-added activities are generated. So even though the physician is heavily utilized, an increasing share of her work is on nonvalue-added tasks that do not contribute to a faster rate of patient discharge. This leads to the following hypothesis:

Hypothesis 1B. *Multitasking has an inverted U-shaped effect on throughput rate*.

### 4.2. Impact on Physician's Quality of Care

In addition to overall productivity, quality of output is an important consideration in evaluating physician performance. To illustrate, consider Figure 3. Instead of *Busy Period* as the outcome variable, we are now interested in the total output quality (*Total Quality*) associated with the care for both patients 1 and 2. The multitasking of the activities for patients 1 and 2 has important implications for this outcome variable. How should one organize and sequence work to maximize *Total Quality*?

Multitasking may initially improve quality of care for several reasons. First, multitasking is associated with an increase in the cognitive load for the physician (Edwards and Gronlund 1998, Vega et al. 2008), leading to an increase in mental stress. A significant amount of scientific research has shown that low levels of stress can aid cognitive function. For example, Lupien et al. (2007) found that at low levels, the release of stress hormones can improve cognitive functioning by directly impacting regions of the brain responsible for learning and recall. Much of this research has explored the physiological underpinnings of performance, including the chemical pathways of hormones and their interactions with neurons (e.g., Lupien et al. 2007, McEwen 2002). For example, Kuntz et al. (2013) find that at low levels of stress, additional stress can improve hospital outcomes. Second, performing multiple tasks may offer motivational benefits (Hackman and Oldham 1976). Switching between tasks may force the physician to remain in a state of mindful activity (Levinthal and Rerup 2006), rather than execute the same routine without thinking. This stream of research suggests that some multitasking may increase physician engagement and improve performance.

At increased levels of multitasking, however, these gains are likely to be offset by the inefficiencies brought on by excessive multitasking. First, as discussed previously, excessive multitasking generates nonvalue-added activities. This effectively puts additional pressure on the physician to reduce the overall time spent diagnosing and treating a patient, particularly given the competing demands of other patients under her care. Under these conditions, the physician may discharge a patient early to alleviate her workload. Second, the time spent on a patient may be less focused, particularly if the physician incurs more frequent interruptions related to other patients. For example, Tucker and Spear (2006) find that interruptions to the routines for health workers are associated with an increased likelihood of medical errors. Third, the increased load on working memory from having to manage information on several distinct patients may also impair the physician's ability to discern relevant information, and to effectively diagnose and treat patients (Baddeley and Hitch 1994). In contrast, a physician who sees only one patient at a time is only required to retain information on the single patient; this information can be flushed from working memory after the patient is discharged. Laboratory studies have confirmed that subjects who switch between cognitive-intensive tasks experience information overload, which leads to a degradation in performance (Allport et al. 1994, Speier et al. 1999). All of these factors potentially impair decision making (Lin 2009, Ophir et al. 2009, Bendoly et al. 2009, Pashler 2000, Aiken et al. 2002). We thus hypothesize that multitasking is beneficial at lower levels of care, whereas at increased levels it ultimately reduces the physician's performance. In the ED, patient revisit within 24 hours is a frequently used measure of quality of care, and is therefore an appropriate measure of the physician's performance. We thus postulate the following:

Hypothesis 2A. *Multitasking has a U-shaped effect on the number of patient revisits.*

An additional measure of physician output quality is the number of diagnoses recorded by the physician. In absolute terms, the number of diagnoses that should be made is a function of the patient's presenting condition. Specifically, for a given set of patients with similar medical workload and risk level, the number of required diagnoses should, on average, be the same. However, once the patient's medical conditions are accounted for, any systematic reductions in the number of diagnoses is indicative of potential quality lapses. Our analysis adjusts for the patient's primary presenting condition and severity through the use of risk and workload scores. Moreover, the arrival of patients to the ED is a random process, which further ensures that the presenting condition of the next patient at the top of the queue is relatively independent of the level of physician multitasking.

The competing forces that affect patient revisit also apply to the number of diagnoses recorded. Low levels of stress can aid physician performance. Similarly, performing multiple tasks may offer motivational benefits (Hackman and Oldham 1976), and switching between tasks may force the physician to remain in a state of mindful activity (Levinthal and Rerup 2006). All of this suggests that some multitasking may increase physician engagement and improve performance, which in turn allows for a more complete patient diagnosis. On the other hand, higher multitasking can limit the time spent with a patient. This reduces the likelihood of a complete and thorough examination, leading to fewer recorded diagnoses. Furthermore, a physician who is spread thin may be forced to "cut corners" (Oliva and Sterman 2001). Reporting additional diagnoses often requires further tests and expert consultations, particularly given the stringent restrictions placed by payers on compliance and accuracy with respect to claims data submission (Crocker 2006). This places further administrative burden on the physician. In a recent study, Powell et al. (2012) found that increased workload reduces the number of diagnoses recorded by physicians in a trauma center. We postulate that even after controlling for workload, the sources of waste due to additional setups and interruptions associated with multitasking lead to fewer diagnoses.

An alternative hypothesis is that a physician whose cognitive capabilities are impaired by multitasking may actually record additional diagnoses because of her inability to discriminate between competing diagnoses. However, we argue that the clinical and administrative burden required in reporting additional diagnoses (coupled with the limitations on time and resource availability, as well as claims data submission restrictions) more than offsets this tendency to increase the number of diagnoses. Given these arguments, we postulate the following:

Hypothesis 2B. *Multitasking has an inverted U-shaped effect on the number of diagnoses.*

## 5. Empirical Specifications

In the discussions below, subscript $i$ denotes physician, $j$ denotes patient, and $t$ denotes time.

### 5.1. Effect on Busy Period

To test Hypothesis 1A, we need to control for the total number of patients discharged. Specifically, the physician needs to discharge the same number of patients over different instances of busy periods but with varying levels of multitasking. Our first order of business is to thus generate such comparable instances. We denote the set of distinct busy periods encountered by physicians by $B$. Let $\tau_{\text{admit},\,i,\,b}$ denote the particular point in time at the beginning of a busy period $b \in B$, in which physician $i$ goes from having zero patients under her care to picking up her first patient. After $\tau_{\text{admit},\,i,\,b}$ the physician may go on to admit and discharge one or more patients. Let $\tau_{\text{disch},\,i,\,b}$ denote the next point in time in which physician $i$ has zero patients again. In other words, discharging her patient at time $\tau_{\text{disch},\,i,\,b}$ demarcates the completion of $b$. Note that immediately before the start and immediately after the end of $b$, a physician has zero patients under her care, and that $b$ can begin anywhere during a shift (not necessarily at the start of the shift). Therefore, $T_{i,\,b} = \tau_{\text{disch},\,i,\,b} - \tau_{\text{admit},\,i,\,b}$ denotes the *Busy Period* (Figure 3) taken by physician $i$ to discharge all of the patients associated with $b$, and is our outcome measure of interest.

Given that our unit of analysis is a busy period $b$, our independent variable is defined at the level of $b$. Since the instantaneous levels of physician multitasking $M_i(t)$ may vary during the busy period, we construct a single composite measure that captures the effective level of multitasking experienced by physician $i$ over the course of the entire busy period $b$. Specifically, we define $\bar{M}_{i,\,b} = (1/T_{i,\,b}) \int_{\tau_{\text{admit},\,i,\,b}}^{\tau_{\text{disch},\,i,\,b}} M_i(t)\,dt$ to denote the time-averaged level of multitasking encountered by physician $i$ over the course of the busy period $b$. To examine how multitasking affects the busy period, we use the following empirical specification:

$$\begin{aligned}\log(T_{i,\,b}) = {}& \alpha_i + N_b + \rho_1 \bar{M}_{i,\,b} + \rho_2 (\bar{M}_{i,\,b})^2 \\ &+ \rho_3 \overline{AdjSysLoad_b} + \rho_4\, Month_b \\ &+ \rho_5\, Day_b + \rho_6\, RVU_b + \rho_7\, Sev_b \\ &+ \rho_8\, NumCat_b + \varepsilon_{i,\,b}, \qquad (1)\end{aligned}$$

where $\alpha_i$ is the physician fixed effect. Since our unit of analysis is the busy period, our control variables are also defined at the level of $b$. In particular, $N_b$ is a fixed effect that denotes the number of patients discharged during $b$. In other words $N_b$ is the same for a set of busy periods with the same number of total patients discharged; for example, if busy periods 1, 6, and 8 all involved discharging seven patients, they would have the same fixed effect $N_b$, which allows us to do an apples to apples comparison of equivalent busy periods.

$RVU_b$ controls for the workload associated with $b$. As described earlier, the RVU for a given patient captures the physician's workload for that patient. We therefore estimate $RVU_b$ by summing the RVUs for each of the patients in $b$. Similarly, higher severity might carry a higher cognitive load for the physician. We account for the composite severity (denoted by $Sev_b$) of the patients in the busy period by aggregating the triage severity levels for all the patients in $b$. In addition, $NumCat_b$ accounts for the variety of patients in $b$, and is measured by the number of different types of patients. Specifically, $NumCat_b$ is equal to the number of distinct MDC categories of patients included in $b$. As an illustrative example, physician Smith is assigned to her first patient (patient 1) at 10 a.m. on a given day. She then picks up patients 2 and 3, and continues working on all three patients, until the last of these three patients is discharged at 2 p.m. that day. The busy period associated with these three patients is four hours. Table 4 presents the individual characteristics of the three patients, as well as the associated composite measures of $N_b$, $RVU_b$, $Sev_b$, and $NumCat_b$.

Our model also accounts for system-level effects. Since the instantaneous value of the system load ($SystemLoad(t)$) varies over the course the busy period, we define

$$\overline{AdjSysLoad_b} = \left(\frac{1}{PhyNum_b T_{i,\,b}}\right) \int_{\tau_{\text{admit},\,i,\,b}}^{\tau_{\text{disch},\,i,\,b}} SystemLoad(t)\,dt,$$

**Table 4 Illustrative Example of Patient Characteristics During a Given Busy Period**

| Patient | Pickup (a.m.) | Discharge (p.m.) | RVU | Emergency status | Patient MDC category |
|---|---|---|---|---|---|
| 1 | 10:00 | 1:00 | 4.2 | 1 | Circulatory system |
| 2 | 11:00 | 12:30 | 3.0 | 0 | Kidney, urinary tract |
| 3 | 11:30 | 2:00 | 5.0 | 1 | Circulatory system |

Control variables for busy period analysis

| $T_{i,\,b}$ | $N_b$ | $RVU_b$ | $Sev_b$ | NumCat$_b$ |
|---|---|---|---|---|
| Four hours | 3 | 12.2 | 2 | 2 |

to denote the time averaged and physician adjusted ($PhyNum_b$ is the average number of physicians on staff during $b$) value of system load over the course of busy period $b$. Finally, we account for the daily and monthly seasonality by including dummy variables corresponding to day of week and month of year associated with $b$. Equation (1) is estimated by least squares. A negative value for $\rho_1$ and a positive value for $\rho_2$ provide support for Hypothesis 1A.

### 5.2. Effect on Throughput Rate

We next estimate the effect of multitasking on individual physician throughput rate. In contrast to the previous analysis, we now hold the time period of observation constant, and examine the number of patients discharged, while varying the level of multitasking. Specifically, we examine patient throughput over the course of the physician's shift. Let $\mu_{i,s}$ denote the number of patients discharged by physician $i$ during a given shift $s$. The relationship between multitasking and throughput rate is captured by the following:

$$\log\left(\frac{\mu_{i,s}}{T_s}\right) = \lambda_{0i} + \lambda_1 \overline{ShiftMT}_{i,s} + \lambda_2 \overline{ShiftMT}^2_{i,s} + \lambda_3 \overline{AdjSysLoad}_s + \lambda_4 Month_s + \lambda_5 Day_s + \lambda_6 RVU_s + \lambda_7 Sev_s + \lambda_8 NumCat_s + e_{i,s}, \quad (2)$$

where $\mu_{i,s}/T_s$ is the number of patients during a given shift, adjusted for the duration of the shift ($T_s$). Over the course of the shift, the instantaneous levels of multitasking vary, as older patients are discharged and new patients are picked up. We therefore quantify the extent of multitasking experienced by physician $i$ over the course of the entire shift $s$ by generating a composite and representative measure of multitasking. This is obtained by computing $\overline{ShiftMT}_{i,s} = (1/T_s)\int_{t_{s,\text{start}}}^{t_{s,\text{end}}} M_i(t)\,dt$, where $t_{s,\text{start}}$ and $t_{s,\text{end}}$ denote the start and end times for shift $s$, respectively, and $T_s = t_{s,\text{end}} - t_{s,\text{start}}$ denotes the duration of the shift. Similarly, since the instantaneous level of system load, $SystemLoad(t)$, varies over the course of the shift, we construct a time-averaged physician adjusted measure of ED system load for the shift $s$ using $\overline{AdjSysLoad}_s = (1/NumPhy_s T_s)\int_{t_{s,\text{start}}}^{t_{s,\text{end}}} SystemLoad(t)\,dt$, where $NumPhy_s$ is the average number of physicians in the shift. Finally, we account for shift-level heterogeneity due to the characteristics of the underlying patients. Since the outcome measure (throughput rate) is adjusted for the duration of the shift, our controls $RVU_s$, $Sev_s$, and $NumCat_s$, corresponding to the time-averaged measures of the aggregate relative value units, patient severity, and number of distinct MDC patient types, respectively, are adjusted for the duration of the shift. For example, $RVU_s$ is given by the total number of $RVU$s in shift $s$ divided by the duration of the shift $T_s$. Similarly, $Sev_s$ is obtained by dividing the sum of the severity scores of the patients in shift $s$ by $T_s$.

A shift lasts longer than a busy period. By definition, a shift can contain one or more busy periods. In addition, the end of a shift presents a constraint on the time that the physician would like to spend in the ED. In particular, physicians can organize their care process to complete their tasks by the end of the shift. However, we do not expect these shift-level effects to impact the inverted U-shaped effect of multitasking on throughput rate. In other words, we expect a positive value for $\lambda_1$ and a negative value for $\lambda_2$ in specification (2).

### 5.3. Effect on Physician's Quality of Care

To test Hypotheses 2A and 2B, we employ an estimation strategy identical to the one described in §5.1, using the same set of explanatory variables. Recall that (1) estimated the effect of multitasking on the busy period, conditional on same number of patient discharged, and after accounting for patient severity and workload. Instead of the length of the busy period as the outcome variable as in (1), we now look at the aggregate quality outcome for the given busy period. For comparable sets of busy periods, we consider two measures of service quality outcome: incidences of patient revisit within 24 hours and the total number of recorded diagnoses. In the specification below, $Y_{i,b}$ denotes the service quality outcome associated with physician $i$ for *all* patients who were discharged during busy period $b$. Specifically, $Y_{i,b}$ is obtained by computing the total number of diagnoses (or the total number of patient revisits) for all patients who were seen in $b$:

$$Y_{i,b} = \alpha_i + N_b + \varphi_1 \bar{M}_{i,b} + \varphi_2(\bar{M}_{i,b})^2 + \varphi_3 \overline{AdjSysLoad}_b + \varphi_4 Month_b + \varphi_5 Day_b + \varphi_6 RVU_b + \varphi_7 Sev_b + \varphi_8 NumCat_b + \varepsilon_{i,b}, \quad (3)$$

where $\alpha_i$ is the physician fixed effect, and $\bar{M}_{i,b}$ is the time-averaged level of multitasking over the course of $b$. As defined previously, $\overline{AdjSysLoad}_b$ is the physician adjusted system load associated with $b$. The specification adjusts for workload, severity, and variety by including $RVU_b$, $Sev_b$ and $NumCat_b$, respectively, as defined in §5.1. Based on the specification above, a negative (positive) value for $\varphi_1$ and a positive (negative) value for $\varphi_2$ provide support for Hypothesis 2A (Hypothesis 2B).

## 6. Results

### 6.1. Key Estimates

Table 5 provides the results for the effect of physician multitasking on busy period (model 1). Multitasking is initially associated with a reduction in busy

**Table 5 Effect of Multitasking on Busy Period Duration**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Physician F.E.* | No | Yes | Yes | Yes | Yes |
| *Number of patients F.E.* | Yes | Yes | Yes | Yes | Yes |
| *Triage severity* | No | No | 0.019*** (0.001) | 0.019*** (0.001) | 0.019*** (0.001) |
| *Day of week* | Yes | Yes | Yes | Yes | Yes |
| *Month* | Yes | Yes | Yes | Yes | Yes |
| *RVU* | Yes | Yes | Yes | 0.486*** (0.09) | 0.486*** (0.09) |
| *Types of patients* | — | — | — | — | −0.0186*** (0.005) |
| *Physician adjusted system load* | −0.058*** (0.007) | 0.056*** (0.005) | 0.056*** (0.005) | 0.056*** (0.0533) | 0.056*** (0.0053) |
| *Physician multitasking* | −0.460*** (0.041) | −0.77*** (0.028) | −0.78*** (0.028) | −0.78*** (0.028) | −0.772*** (0.028) |
| *Physician multi-tasking squared* | 0.086*** (0.006) | 0.091*** (0.0039) | 0.089*** (0.0038) | 0.089*** (0.0039) | 0.088*** (0.0039) |
| Model fit $F$-test ($\Pr > F$) | $<0.001$ | $<0.001$ | $<0.001$ | $<0.001$ | $<0.001$ |
| Adjusted $R$-squared | 0.639 | 0.826 | 0.827 | 0.829 | 0.829 |

*Notes.* Asymptotic standard errors are in parentheses. The RVU coefficient is divided by a factor of 1,000. F.E., fixed effects.
***1% statistical significance.

**Table 6 Effect of Multitasking on Shift Throughput Rate**

| | (1) | (2) | (3) |
|---|---|---|---|
| *Physician F.E.* | No | Yes | Yes |
| *Day of week* | Yes | Yes | Yes |
| *Month* | Yes | Yes | Yes |
| *Shift severity* | — | — | −0.0046*** (0.001) |
| *Shift RVU* | — | — | 0.1 (0.1) |
| *Shift patient types* | — | — | −0.0048*** (0.0004) |
| *System load* | −0.0076*** (0.001) | −0.0137*** (0.001) | −0.0128*** (0.0014) |
| *Physician multitasking* | 0.749*** (0.025) | 0.888*** (0.030) | 0.802*** (0.029) |
| *Physician multitasking squared* | −0.0695*** (0.004) | −0.087*** (0.004) | −0.078*** (0.0037) |
| Model fit $F$-test ($\Pr > F$) | $<0.001$ | $<0.001$ | $<0.001$ |
| Adjusted $R$-squared | 0.693 | 0.766 | 0.802 |

*Notes.* $N = 3{,}457$ shifts. Asymptotic standard errors are in parentheses. The RVU coefficient is divided by a factor of 1,000. F.E., fixed effects.
***1% statistical significance.

period, as evidenced by the negative value for the linear coefficient term for multitasking ($-0.77, p < 0.01$) in specification (5). The coefficient estimate for the quadratic term is positive ($0.089, p < 0.01$), indicating decreasing returns to multitasking. Overall physician adjusted ED system load is also associated with an increase in the busy period ($0.056, p < 0.01$); this reflects the fact that the patient's care process involves multiple resources (e.g., lab testing facilities, specialists, nursing staff, etc.); when the ED is congested, the demand for these other resources also increases, leading to an overall increase in the time taken to discharge the patient. Our results are robust to various specifications (1)–(5), which include the physician fixed effect as well as workload, severity, and patient mix controls. Our coefficient estimates and our results for the U-shaped effect of multitasking (Hypothesis 1A) are largely unchanged by including heterogeneity due to workload and patient mix.

Multitasking is initially associated with improvement in physician throughput rate (Table 6) based on our estimation of model 2; at low levels of multitasking a physician can improve throughput rate by taking on additional patients and utilizing idle time. At high levels of multitasking, the sources of throughput loss play a dominant role. Multitasking thus has an inverted U-shaped effect on throughput rate, providing support for Hypothesis 1B. The corresponding linear and quadratic coefficient terms from specification (3) are (0.802, $p < 0.01$) and (0.078, $p < 0.01$), respectively. Based on our model specification, the optimal level of physician multitasking that maximizes throughput rate is around five patients ($0.802/(2 \times 0.078)$). Table 7 illustrates the predicted values for throughput rate for the average physician as she increases the level of multitasking from three to six patients. We see that patient throughput rate (the number of patients discharged per hour) improves up to five patients, but decreases from five patients on. The corresponding throughput rate difference between multitasking levels of 4 and 5 is 0.11 patient per hour, and the throughput rate difference between multitasking levels of 5 and 6 is 0.14 patients per hour. These values correspond to a difference of approximately one patient discharge over the course of 7–10 hours, if the physician deviates from the optimal multitasking level of five patients.

Next we examine the other dimension of physician performance: quality of care (Tables 8 and 9, based on model 3). We first considered estimating Equation (3) by Poisson regression. However, we find the dispersion parameter to be statistically significant in all of our specifications. Therefore, we use the more general and appropriate negative binomial regression

**Table 7 Predicted Average Effect of Multitasking on Shift Throughput Rate**

| Multitasking | Throughput (patients/hour) |
|---|---|
| 3 | 1.81 |
| 4 | 2.26 |
| 5 | 2.37 |
| 6 | 2.23 |

**Table 8 Effect of Multitasking on Busy Period Revisits**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Physician F.E.* | No | Yes | Yes | Yes | Yes |
| *Number of patients F.E.* | Yes | Yes | Yes | Yes | Yes |
| *Triage severity* | — | — | 0.004<br>(0.01) | 0.001<br>(0.01) | −0.004<br>(0.01) |
| *Day* | Yes | Yes | Yes | Yes | Yes |
| *Month* | Yes | Yes | Yes | Yes | Yes |
| *RVU* | — | — | — | 0.1 (0.5) | −0.0 (0.5) |
| *Types of patients* | — | — | — | — | 0.81***<br>(0.028) |
| *Adjusted system load* | −0.0939***<br>(0.028) | 0.08**<br>(0.03) | 0.0784**<br>(0.029) | −0.0787***<br>(0.029) | −0.0780***<br>(0.029) |
| *Physician multitasking* | −0.336***<br>(0.046) | −0.405***<br>(0.146) | −0.406***<br>(0.147) | −0.407***<br>(0.147) | −0.46***<br>(0.148) |
| *Physician multitasking square* | 0.0516**<br>(0.020) | 0.052**<br>(0.02) | 0.052**<br>(0.02) | 0.052**<br>(0.02) | 0.058***<br>(0.02) |
| Log-likelihood | −6,799 | −6,764 | −6,760 | −6,760 | −6,756 |
| Akaike information criterion | 17,854 | 17,833 | 17,858 | 17,860 | 17,853 |
| Likelihood ratio (Pr > Chi-square) | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |

*Notes.* Asymptotic standard errors are in parentheses. The RVU coefficient is divided by a factor of 1,000. F.E., fixed effects.

**5% statistical significance; ***1% statistical significance.

model to estimate the impact on revisits and number of diagnoses. We find that initially, the multitasking physician experiences fewer 24-hour patient revisits (Table 8, specification 5, coefficient $= -0.46$, $p < 0.01$). However, with increasing levels of multitasking, we begin to see more frequent 24-hour revisits (quadratic coefficient $= 0.058$, $p < 0.01$). This provides support for Hypothesis 2A. Based on these estimates, the optimal number of patients to multitask to reduce revisits is around 4. Table 9 (column 5) provides the results for the effect of multitasking on the total number of diagnoses. At low levels, an increase in the level of multitasking initially leads to more diagnoses (coefficient $= 0.0836$, $p < 0.01$). However, at increasing levels of multitasking, physicians make fewer diagnoses (quadratic coefficient $= -0.0065$, $p < 0.01$); thus physicians tend to make the most number of diagnoses when the level of multitasking is closer to six

**Table 9 Effect of Multitasking on Busy Period Diagnoses**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Physician F.E.* | No | Yes | Yes | Yes | Yes |
| *Number of patients F.E.* | Yes | Yes | Yes | Yes | Yes |
| *Triage severity* | — | — | 0.009***<br>(0.0005) | 0.009***<br>(0.0005) | 0.009***<br>(0.0005) |
| *Day* | Yes | Yes | Yes | Yes | Yes |
| *Month* | Yes | Yes | Yes | Yes | Yes |
| *RVU* | — | — | — | −0.1***<br>(0.0) | −0.1***<br>(0.0) |
| *Types of patients* | — | — | — | — | 0.0043***<br>(0.0016) |
| *Adjusted system load* | 0.0579***<br>(0.002) | 0.011***<br>(0.0021) | 0.0107***<br>(0.0021) | 0.0107***<br>(0.0021) | 0.0101***<br>(0.0021) |
| *Physician multitasking* | 0.365***<br>(0.001) | 0.0932***<br>(0.010) | 0.0839***<br>(0.010) | 0.0848***<br>(0.01) | 0.0836***<br>(0.01) |
| *Physician multitasking squared* | −0.0225***<br>(0.001) | −0.0069***<br>(0.0012) | −0.0065***<br>(0.001) | −0.0066***<br>(0.012) | −0.0065***<br>(0.0012) |
| Log-likelihood | −74,996 | −59,258 | −59,107 | −59,103 | −59,100 |
| Akaike information criterion | 150,211 | 118,784 | 118,484 | 118,479 | 118,474 |
| Likelihood ratio (Pr > Chi-square) | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |

*Notes.* Asymptotic standard errors are in parentheses. The RVU coefficient is divided by a factor of 1,000. F.E., fixed effects.

***1% statistical significance.

patients. This provides support for the Hypothesis 2B. Our results are robust to various specifications, which include workload, severity, and patient mix controls, indicating that sources of patient heterogeneity do not weaken the support for Hypotheses 2A and 2B.

### 6.2. Robustness Analysis

Physicians who are more adept at multitasking due to their innate cognitive abilities, skills, and experiences may generate greater gains from multitasking. To rule out unobserved physician heterogeneity as a source of potential confounding, our estimations include the physician fixed effect, and our panel data estimation is driven by the intraphysician variation in multitasking.

To further rule out spurious correlations between unobserved patient conditions and physician multitasking, we examined the patient–physician assignment process in several different ways. As described earlier, physicians pick patients with higher levels of acuity on a first come, first served basis; that is, if a more severe patient requires attention, the available physician chooses the urgent care patient from the top of the queue. If there are no urgent care patients, the physician picks up a nonurgent care patient at the top of the lower acuity queue. Physicians do not observe detailed information on the patient other than the triage severity level and the broad definition of medical condition at the time of pickup. The detailed medical needs of the patient are determined only after the initial review of the patient's case, face-to-face evaluation, and test results. In other words, given the assignment process, the lack of detailed clinical information, and the investigative nature of the critical care delivery, physicians are not able to cherry-pick patients. This assignment process was further confirmed during our interview with the chief of the emergency department.

Second, we further examine the data to verify the lack of cherry picking. Given that the arrival of patients is a random process, we expect the characteristics of the patient at the top of the pile (conditional on the physician's decision to pick up a new patient) to be independent of the level of physician multitasking. We empirically examined the relationship between the severity level of patient $j$ as proxied by the emergency severity designation ($EmergencySeverity_{ijt} = 1$) and the level of multitasking of physician $i$ at the time of pickup of patient $j$ ($M_{ij}$) using the following logistic specification:

$$logit(EmergencySeverity_{ijt}) = \alpha_i + \gamma_1 M_{ij} + \gamma_2 SystemLoad_t + \gamma_3 Month_t + \gamma_4 Day_t.$$

Our specification includes the physician fixed effect ($\alpha_i$) and controls for system load, as well as monthly and daily seasonality. In other words, conditional on picking up a new patient, is the patient severity correlated with the level of multitasking? We find the coefficient estimate on multitasking $\gamma_1$ to be close to zero and statistically insignificant (Table A-1 in the online supplement), which further suggests the lack of cherry-picking by physicians.

Third, and most importantly, by including the patient's clinical controls (which are determined after a complete diagnosis) in our estimators, we account for more detailed information of the patient's medical condition and level of severity beyond that available to the physician at the time of pickup. In particular, it is worth noting that the inclusion of these controls has no appreciable impact on the linear and quadratic coefficient estimates across specifications (2)–(5) in Table 5. This suggests that unobserved heterogeneity may not be significant.

## 7. Conclusion and Future Research

In this paper, we studied the multitasking behavior of physicians in an emergency department and looked at its implications for physician performance, including productivity and output quality. We measured physician productivity in two distinct ways: by examining the throughput rate of patients over a fixed amount of time and by examining the overall time taken to discharge a given number of patients. We find that some multitasking is initially beneficial, but that excessive multitasking is detrimental to productivity. The initial performance gain is a result of the physician's reduction of idle time by switching to other tasks; however, once she is already operating at high levels of multitasking (and has reduced much of the idle time), sources of waste and the creation of nonvalue-added activities, including more frequent setups, interruptions, and coordination costs, begin to play a more dominant role.

Although beneficial at lower levels, at higher levels, multitasking was also found to adversely impact patient quality of care: physicians who multitask excessively are likely to make fewer patient diagnoses, and their patients are more likely to revisit the ED within 24 hours. Taken together, these two results provide a more compelling picture of quality lapses due to multitasking.

Prior work in the operations management literature has developed analytical tools to maximize productivity in the presence of setups and changeover costs (Allahverdi et al. 1999, Krishnan et al. 1997). We contribute to this literature by demonstrating that the behavioral response of workers in a critical care setting also depends significantly on the overlap of underlying activities. Our empirical analyses also allow us to identify the optimal levels of multitasking for maximizing throughput rate and output quality.

Improving ED operations calls for a better understanding of physician behavior. In particular, capacity expansion decisions need to account for the fact that the marginal productivity of physicians depends on their level of multitasking. For example, Khare et al. (2009) find that adding more beds (which may allow a physician to multitask at a higher level) does not improve ED congestion. Our research shows that at the tactical level, for a given system load and a fixed number of physicians on staff, more informed patient-to-physician assignments can be made to improve the aggregate system throughput rate.

Future research on physician staffing can incorporate the relationship between multitasking and physician throughput. In deciding the ideal number of physicians to staff, some considerations might include (1) cost of staffing a physician, (2) cost of patient waiting (prior to physician assignment), (3) congestion in the waiting area, (4) throughput rate, and (5) quality of care, among others. For instance, physicians may reduce the time spent in the waiting area prior to the start of care by taking on multiple patients simultaneously. In our analysis, we do not observe the waiting time prior to the start of care. However, a shorter waiting time could enhance patient perception of service quality, and may be an important outcome measure. As such, a multiserver queuing model with an objective function that includes the costs of both waiting and staffing but also allows the individual server throughput rate to vary with multitasking could be used to recommend optimal staffing levels. Such a model would extend the literature on queuing theory, which has typically incorporated the two costs (waiting and server), but has not considered the setup costs due to multitasking in a multiserver environment.

For the average physician in the ED studied, a multitasking level of 5 maximizes physician shift throughput. However, the level of multitasking that maximizes throughput will likely vary across physicians, due to differences in individual working styles and inherent capabilities. In addition, some of the throughput gains that we observe are predicated on utilizing idle time between patient encounters. These idle times may vary across systems. For example, in an ED with longer test result turnaround times, a physician should ideally operate at higher levels of multitasking. The optimal level of multitasking should therefore be determined separately for each ED. Once the optimal level is determined, an ED can offer guidelines or recommend specific levels of multitasking to individual physicians. Alternatively, the ED could enforce specific levels of multitasking by reorganizing the care process to assign a fixed set of rooms to each physician on shift, while the triage nurse assumes responsibility for assigning patients to ED rooms.

In summary, multitasking has previously demonstrated impairment of decision making in the laboratory, as identified in the experimental psychology literature and the emerging work in cognitive neuroscience. This paper presents the first work in the healthcare operations management literature to empirically examine the effect of multitasking on worker performance using transactional data from the field. Although the relative effects of multitasking will vary across different settings, we believe that the theory and methods presented in the present study provide an important step toward documenting this important phenomenon as well as deepening our understanding of the behavioral underpinnings of worker productivity, particularly in a critical care setting. Future research needs to further explore the effects on other measures of quality of care and medical decision making due to multitasking. In our setting, physicians choose their desired level of multitasking. A randomized control study where physicians are assigned to varying levels of multitasking would allow us to examine the extent to which the physician decision to multitask drives performance. An exciting opportunity for future research may yet be opened by the use of microlevel data that include detailed time and motion observations.

## Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2013.0464.

## Acknowledgments

The author is grateful for comments received from the editor in chief, the associate editor, three anonymous referees, and from Rob Shumsky, Brad Staats, Marcelo Olivares, Nikolay Osadchiy, Armann Ingolfsson, Avi Mandelbaum, Steve Pitts, and seminar participants at the Wharton Empirical OM Workshop and at the Alberta School of Business.

## References

Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *J. Amer. Medical Assoc.* 288(16):1987–1993.

Allahverdi A, Gupta JND, Aldowaisan T (1999) A review of scheduling research involving setup considerations. *Omega* 27(2):219–239.

Allport A, Styles EA, Hsieh S (1994) Shifting intentional set: Exploring the dynamic control of tasks. Umiltà C, Moscovitch M, eds. *Attention and Performance XV* (MIT Press, Cambridge, MA), 421–452.

Aral S, Brynjolfsson E, Van Alstyne M (2012) Information, technology, and information worker productivity. *Inform. Systems Res.* 23(3):849–867.

Baddeley AD, Hitch GJ (1994) Developments in the concept of working memory. *Neuropsychology* 8(4):485.

Bendoly E, Croson R, Goncalves P, Schultz K (2009) Bodies of knowledge for research in behavioral operations. *Production Oper. Management* 19(4):434–452.

Charron S, Koechlin E (2010) Divided representation of concurrent goals in the human frontal lobes. *Science* 328(5976):360–363.

Chisholm CD, Collison EK, Nelson DR, Cordell WH (2000) Emergency department workplace interruptions are emergency physicians interrupt-driven?and multitasking? *Acad. Emergency Medicine* 7(11):1239–1243.

Chisholm CD, Collison EK, Nelson DR, Cordell WH (2001) Emergencywork interrupted. A comparison of workplace interruptions in emergency departments and primary care offices. *Ann. Emergency Medicine* 38(2):146–151.

Crocker J (2006) How to improve your revenue cycle processes in a clinic or physician practice. *AHIMA's 78th Natl. Convention and Exhibit Proc.*, October 2006.

Croskerry P, Shapiro M, Campbell S, LeBlanc C, Sinclair D, Wren P, Marcoux M (2004) Profiles in patient safety: Medication errors in the emergency department. *Acad. Emergency Medicine* 11(3):289–299.

Denning PJ (1971) Third generation computer systems. *ACM Comput. Surveys* 3(4):175–216.

Edwards M, Gronlund S (1998) Task interruption and its effects on memory. *Memory* 6(6):665–687.

Gino F, Pisano G (2008) Toward a theory of behavioral operations. *Manufacturing Service Oper. Management* 10(4):676–691.

Gladstones WH, Regan MA, Lee RB (1989) Division of attention: The single-channel hypothesis revisited. *Quart. J. Experiment. Psych.* 41(1):1–17.

Hackman JR, Oldham GR (1976) Motivation through the design of work: Test of a theory. *Organ. Behav. Human Performance* 16(2): 250–279.

Hallowell EM (2005) Overloaded circuits: Why smart people underperform. *Harvard Bus. Rev.* 83(1):54–62.

Hasija S, Pinker E, Shumsky RA (2010) OM practice—Work expands to fill the time available: Capacity estimation and staffing under Parkinson's law. *Manufacturing Service Oper. Management* 12(1):1–18.

Holmstrom B, Milgrom P (1991) Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *J. Law, Econom., Organ.* 7:24–52.

Hoshino LC, Goel V, Raymont V, Grafman J (2006) Assessing the problem-solving abilities of patients with frontal lobe lesions using a real world travel planning task. Working paper, York University, Toronto.

Institute of Medicine (2006) Hospital based emergency care: At the breaking point. Future of emergency care. Report, National Academies Press, Washington, DC.

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

KC DS, Terwiesch C (2012) An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing Service Oper. Management* 14(1):50–65.

Khare RK, Powell ES, Reinhardt G, Lucenti M (2009) Adding more beds to the emergency department or reducing admitted patient boarding times: Which has a more significant influence on emergency department congestion? *Ann. Emergency Medicine* 53(5):575–585.

Krishnan V, Eppinger SD, Whitney DE (1997) A model-based framework to overlap product development activities. *Management Sci.* 43(4):437–451.

Kuntz L, Mennicken R, Scholtes S (2013) Stress on the ward: Evidence of safety tipping points in hospitals. Working paper, University of Cologne, Cologne, Germany.

Levinthal D, Rerup C (2006) Crossing an apparent chasm: Bridging mindful and less-mindful perspectives on organizational learning. *Organ. Sci.* 17(4):502–513.

Lin L (2009) Breadth-biased versus focused cognitive control in media multitasking behaviors. *Proc. Natl. Acad. Sci. USA* 106(37):15521–15522.

Lupien SJ, Maheu F, Tu M, Fiocco A, Schramek TE (2007) The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition* 65(3):209–237.

Mayer RE, Moreno R (2003) Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist* 38(1):43–52.

McEwen BS (2002) *The End of Stress As We Know It* (Joseph Henry Press, Washington, DC).

Narayanan S, Balasubramanian S, Swaminathan JM (2009) A matter of balance: Specialization, task variety, and individual learning in a software maintenance environment. *Management Sci.* 55(11):1861–1876.

Needleman J, Buerhaus PI, Stewart M, Zelevinsky K, Mattke S (2006) Nurse staffing in hospitals: Is there a business case for quality? *Health Affairs* 25(1):204–211.

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

Ophir E, Nass C, Wagner AD (2009) Cognitive control in media multitaskers. *Proc. Natl. Acad. Sci. USA* 106(37):15583–15587.

Parekh AK, Gallager RG (1993) A generalized processor sharing approach to flow control in integrated services networks: The single-node case. *IEEE/ACM Transactions on Networking* 1(3): 344–357.

Pashler H (1989) Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psych.* 21(4):469–514.

Pashler H (1994) Dual-task interference in simple tasks: Data and theory. *Psych. Bull.; Psych. Bull.* 116(2):220–244.

Pashler H (2000) Task switching and multitask performance. Monsell S, Driver J, eds. *Control of Cognitive Processes*, Vol. XVIII (MIT Press, Cambridge, MA) 277–308.

Pennebaker R (2009) The mediocre multitasker. *New York Times* (August 29), http://www.nytimes.com/2009/08/30/weekinreview/30pennebaker.html.

Powell A, Savin S, Savva N (2012) Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing Service Oper. Management* 14(4): 512–528.

Powell SG, Schultz KL (2004) Throughput in serial lines with state-dependent behavior. *Management Sci.* 50(8):1095–1105.

Rubinstein JS, Meyer DE, Evans JE (2001) Executive control of cognitive processes in task switching. *J. Experiment. Psych.: Human Perception and Performance* 27(4):763–797.

Salamon J (2008) *Hospital: Man, Woman, Birth, Death, Infinity, Plus Red Tape, Bad Behavior, Money, God, and Diversity on Steroids* (Penguin Press HC, New York).

Schultz KL, Juran DC, Boudreau JW (1999) The effects of low inventory on the development of productivity norms. *Management Sci.* 45(12):1664–1678.

Schultz KL, McClain JO, Thomas LJ (2003) Overcoming the dark side of worker flexibility. *J. Oper. Management* 21(1):81–92.

Smith A (1776) *The Wealth of Nations*. [Reprint 1937, Modern Library, New York.]

Speier C, Valacich JS, Vessey I (1999) The influence of task interruption on individual decision making: An information overload perspective. *Decision Sci.* 30(2):337–360.

Staats BR, Gino F (2012) Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* 58(6):1141–1159.

Taylor FW (1914) *The Principles of Scientific Management* (Harper, New York).

Tucker AL, Edmondson AC (2003) Why hospitals don't learn from failures: Organizational and psychological dynamics that inhibit system change. *California Management Rev.* 45(2):55–72.

Tucker AL, Spear SJ (2006) Operational failures and interruptions in hospital nursing. *Health Services Res.* 41(3, part 1):643–662.

Vega V, McCracken K, Nass C (2008) Multitasking effects on visual working memory, working memory and executive control. Presentation, Annual Meeting of the International Communication Association, May 22, Montreal.