MIT Open Access Articles

Recovering from Critical Incidents: Evidence from Paramedic Performance

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation	Bavafa, Hessam and J #onas Oddur J #onasson. "Recovering from Critical Incidents: Evidence from Paramedic Performance." Manufacturing and Service Operations Management (April 2020) © 2020 The Author(s)
As Published	10.1287/MSOM.2019.0863
Publisher	Institute for Operations Research and the Management Sciences (INFORMS)
Version	Author's final manuscript
Citable link	https://hdl.handle.net/1721.1/130369
Terms of Use	Creative Commons Attribution-Noncommercial-Share Alike
Detailed Terms	http://creativecommons.org/licenses/by-nc-sa/4.0/

Massachusetts Institute of Technology

DSpace@MIT

Recovering from Critical Incidents: Evidence from Paramedic Performance

Hessam Bavafa
Wisconsin School of Business, University of Wisconsin-Madison, hessam.bavafa@wisc.edu

Jónas Oddur Jónasson
MIT Sloan School of Management, joj@mit.edu

Problem Definition: In service operations settings where the difficulty of jobs is unpredictable, workers can encounter critical incidents (CIs)—jobs which are sufficiently disturbing to challenge workers' coping mechanisms. We examine the impact of encountering CIs on subsequent operational performance of workers.
Academic / Practical Relevance: Prior work has examined the effects of CIs on the long-term psychological health of workers. We demonstrate that encountering CIs has a practically meaningful impact on operational performance. We also examine the time-dependency and process-dependency of the effect, and analyze whether it is mitigated by individual characteristics such as age or experience.
Methodology: We use data on 902,002 ambulance activations conducted by paramedics at the London Ambulance Service (LAS). We define CIs as incidents where patients have a high probability of dying at the scene, and examine the impact of such events on the paramedics' performance for the remainder of their shifts. Our outcomes are the completion time of the ambulance activation and each of its five sub-processes. The exogenous assignment of CIs to paramedic crews allows a clean identification of our effect using a shift-level difference-in-differences specification.
Results: Crews who have encountered one prior CI (two prior CIs) spend on average 2.6% (7.5%) more time completing each remaining ambulance activation in the shift. The impact is strongest for the jobs immediately following a CI but persists throughout the shift. The largest effects come from the sub-processes which are least standardized and where paramedics cannot rely on standard operating procedures. The duration effect is larger for teams of older paramedics, but is simultaneously mitigated by experience.
Managerial Implications: Our results show that CIs increase subsequent job duration and that more than one CIs have a compounding, negative effect on operational performance. As a result, managers in settings where performance consistency is key would be advised not to assign new jobs to teams with recent CI experiences.

Key words: Critical Incidents, Team Performance, Service Operations, Ambulance Operations
History: October 6, 2019

1. Introduction

In many service operations settings, the difficulty of jobs is unpredictable, not only in complexity, but also in the emotional or psychological impact on workers. Examples include police officers or paramedics routinely responding to calls with only limited understanding of the situations awaiting them. In such settings, a worker may encounter a *critical incident* (CI), defined as a task or situation which is sufficiently disturbing to challenge or overwhelm the workers' usual coping mechanisms (Alexander and Klein 2001). While the impact of CIs on long-term stress (Marmar et al. 1999) and burnout (Halpern et al. 2014) have been widely discussed from a psychological perspective, it is unknown to what extent such events impact (immediate) operational task performance.

We study the impact of CIs on subsequent operational performance, measured by job completion times in a paramedic setting. Specifically, we test four hypotheses. First, we expect paramedic crews who encounter a CI to spend more time completing subsequent jobs as a result of the impact of CIs on short-term functioning (Halpern et al. 2012). Second, we expect the largest performance impact on the jobs immediately following a CI, since the recovery takes time. Third, with each job comprising multiple sub-processes, we expect the impact of a prior CI to be largest on those processes for which paramedics have no external decision-making support and must rely on their own judgment. Fourth, we expect the performance impact of prior CIs to be mitigated by the age and work experience of the paramedic crew.

We test our hypotheses using a dataset from the London Ambulance Service (LAS). The data describes every *ambulance activation* (the term used for a single job, including driving, patient pick-up at the scene, and patient handover at the hospital) undertaken by the service during 2011. Our dataset contains information on 902,002 ambulance activations, including time-stamps (e.g., dispatch time, arrival at scene, and arrival at hospital), crew information, patient characteristics, and the receiving hospital. The setting is ideal for our analysis as the probability of any given paramedic crew being assigned to a CI is effectively random and exogenous to performance. This allows a clean identification of the effects hypothesized above.

We measure operational performance by the completion time of an ambulance activation, and sub-processes thereof. The completion time is defined as the duration from the ambulance dispatch to a given incident until the ambulance becomes available again. This comprises driving times as well as the time spent at the scene picking up the patient, at the hospital handing a patient over to the emergency department, and post-handover preparation for the next dispatch. Shortening completion times is an important objective for the LAS. Also, response times, handover times, and ambulance preparation times are official key performance indicators for paramedics, and for most serious health conditions there are significant patient-health benefits of getting to a hospital as quickly as possible (e.g., Sacco et al. 2005). Shorter completion times result in lower ambulance utilization which is key to enable swift response times.

In the context of emergency medicine, prior work on the psychological well-being of paramedics has identified job characteristics common to CIs. Broadly, these capture whether the incident was *troubling* or *disturbing*, for instance because the experience was beyond the control of the paramedics or the paramedics felt helpless or overwhelmed (Halpern et al. 2012). For our analysis, we define a CI as an ambulance activation where the patient has a high probability of dying at the scene. Specifically, three illness codes, as recorded on the crew's patient report form, have been associated with such outcomes; *Hanging*, *Cardiac Arrest*, and *[Patient] Purple* (an informal LAS term used for patients who are known to have died) (Halter and Ellison 2008). Our sample includes 8,404 such CIs (0.9% of all ambulance activations) with at least one CI occurring as part of 5.5% of shifts. (§6.5 includes robustness checks with broader CI definitions.)

Our analysis is focused on the short-term operational performance impact of CIs within a given shift, and the results are consistent with the above hypotheses. First, we find that encountering a CI does significantly impact subsequent completion times. Specifically, crews who have encountered one prior CI (two prior CIs) as part of their shift spend 2.6% (7.5%) longer to complete the remaining ambulance activations of the shift. Second, we find that the effect is diminishing as the paramedics conduct more ambulance activations following the CI. However, the magnitude of the tapering-off is small, so the overall effect persists throughout the shift. Third, we find that the effect is not uniform across the sub-processes of the ambulance activation; there are greater effects on the completion times of processes which rely on paramedic decision-making in the absence of standardized operating procedures or external support. Finally, we observe that while more work experience mitigates the impact of prior CIs on performance, this effect is outweighed by the fact that paramedic crews of older age are more susceptible to the detrimental effects of CIs.

Our findings contribute to the literature on operational productivity by demonstrating that jobs which are emotionally or psychologically taxing can affect operational performance on subsequent jobs. An important aspect of our setting is that paramedics, including at the LAS, receive extensive training on dealing with CIs and enter the job knowing that they are likely to face disturbing events. As such, any effects of CIs on operational performance should be muted. Our finding, that CIs have an effect on paramedic performance, highlights the people-centric nature of service operations; despite training, workers' emotional state matters.

Our results also contribute to the vast applied psychology literature concerned with occupational stress, CIs, and workers' psychological well-being. This literature has so far mostly focused on long-term emotional and psychological outcomes (see §2.2) using worker surveys. Our setting allows for a careful analysis of consistently measured outcomes in the short-term, revealing a significant impact on operational performance. Our results imply that operations managers in settings where performance consistency is key would be advised to not assign new jobs to teams with recent CI experiences. These implications are particularly relevant for operational managers in settings where performance depends substantially on workers' own decision-making skills and where they are likely to encounter CIs.

In the rest of the paper, we provide a literature review in §2, introduce our empirical setting in §3, develop our hypotheses in §4, detail our empirical strategy in §5, and report our results in §6. We then conclude with a discussion of our contributions in §7.

2. Literature Review

Our work relates to two distinct streams of literature. First, the empirical literature on drivers of operational performance, discussed in §2.1. Second, the theoretical and empirical literature in behavioral psychology on the CI impact on worker wellbeing and functioning, summarized in §2.2.

2.1. Drivers of Worker Performance in Operational Systems

In many service settings, system outcomes are determined by the joint performance of many individual workers. Since worker performance is highly variable across individuals and settings, a growing stream of literature in operations management is devoted to understanding drivers of operational performance in such settings.

Experience. Substantial work has been done on the positive performance impact of the cumulative experience of individuals, teams, or organizations (e.g., Argote and Epple 1990, Reagans et al. 2005, Huckman et al. 2009, KC and Staats 2012, Staats 2012). Building on this work, it has been found that some of the performance improvements due to individual experience are only realized at the particular organization (Huckman and Pisano 2006), for the particular task (KC and Staats 2012), or for the particular customer (Clark et al. 2013), with which the experience was acquired. Similarly, studies have found that diversity in prior tasks (Boh et al. 2007), customers (Clark et al. 2013, Huckman and Staats 2011), and partners (Akşin et al. 2018, Kim et al. 2018) can impact performance.

Workload. In addition, researchers have established a range of behavioral factors affecting performance in human operated systems. Delasay et al. (2018) provide an overview of how workers adjust service rates due to system workload. These adjustments include early-task initiation (Batt and Terwiesch 2016), rationing service provision (Freeman et al. 2016), and adjusting admission decisions (Kim et al. 2014). However, adjusting service rates has been shown to only relieve workload up to a saturation point and to have a detrimental quality impact in certain settings (KC and Terwiesch 2009, Kuntz et al. 2014, Berry-Jaeker and Tucker 2016).

Shift effects. A set of related papers examines factors affecting worker performance within shift. Focusing on compliance to protocols, Dai et al. (2015) find that compliance to hand hygiene standards drops throughout workers' shifts but improves as a result of longer breaks between shifts. Staats et al. (2016) find that this effect can be mitigated through electronic monitoring of worker compliance, which substantially improves compliance. However, they find that the effect is reversed as soon as the monitoring is discontinued. In the context of an outpatient department, Deo and Jain (2018) show that workers start slow but then increase their service rates throughout a shift. This effect is partly achieved through doctors ordering relatively fewer tests for patients later in a shift.

Task selection. In settings where workers have autonomy to select which tasks to work on next, such decisions can affect productivity and outcomes. Ibanez et al. (2018) find that in a setting where doctors can order their individual list of tasks, they tend to prioritize similar tasks and tasks that can be completed faster. They find that this task ordering erodes productivity but that with increased doctor experience some of the negative productivity impact is mitigated. In addition, KC et al. (2017) find that workers who select their next task from a common queue, tend to choose easier tasks first (*task completion bias*) during periods of high workload. This lightens workload in the short-term but hurts long-term performance.

Broadly, we contribute to the above stream of literature by introducing the notion of CIs and establishing that encountering emotionally or psychologically challenging tasks can negatively affect subsequent performance, at least for the remainder of a given shift. Furthermore, our findings regarding the time-dependency and process-dependency of this effect have managerial implications for scheduling managers, who benefit from understanding how persistent the effects of CIs are, and for which process types they are most pronounced.

More specifically, the papers most related to our work consider how fatigue and emotional reactions affect operational performance.

Fatigue. Our work relates to the literature on fatigue, which, usually defined as exposure to "high load for an extended period of time" (Delasay et al. 2018) or "sustained load in the immediate past" (KC and Terwiesch 2009). As such, fatigue is used to describe the effect of performance deteriorating gradually as the worker conducts more and more jobs. From a theoretical standpoint, the impact of CIs is therefore clearly distinct from that of fatigue. A CI can happen at any time and has an instant effect on job performance, which decreases as time passes (or jobs accumulate) since the incident.

Emotional effects. Finally, two recent papers examine psychological or emotional mechanisms affecting worker performance and decision making. Altman et al. (2018) complements our analysis by focusing on agent-customer interactions, finding that customer sentiment affects the duration of such interactions, both through an increased number of turns (messages, back and forth) and the time it takes an agent to answer each message. While they, like us, focus on a psychological reaction of workers to a specific job, there are clear distinctions between their work and ours. First, they focus on interpersonal communication effects between the agent and the customer whereas we focus on exogenous critical incidents. Second, they investigate the impact of emotional load of the task at hand on its own completion time. In contrast, we control for the CI itself but investigate its impact on the duration of subsequent jobs.

Similar to our results, Ibanez and Toffel (2018) find that characteristics of prior jobs affect subsequent job outcomes. Specifically, their results indicate that food safety inspectors cite more violations at a given establishment if prior inspections at other establishments resulted in many citations. Similarly, they find that observing deteriorating conditions in one establishment increases the inspectors' stringency in the following one. There are three main differences between their work and ours. First, our outcome variable is the completion time of an operational process, whereas their outcome is a decision (number of violations cited) made by the worker. Second, their theory builds on various behavioral biases whereas the theory for CIs suggests that our effects stress-induced. Third, Ibanez and Toffel (2018) find that it is the prior decisions of the worker themselves that affects subsequent decisions. In our case the catalytic event is an exogenous CI.

2.2. Stress in the Workplace

A rich body of work in applied psychology studies the impact of stress on worker wellbeing and performance (see Daniels and Harris 2000, Jex 1998, for reviews). While there is no clear consensus on the impact of various types of stress on work performance (Jex 1998), researchers have suggested that there can be good and bad stress (Selye 1982): *hindrance* stressors—bad for performance—and *challenge* stressors—good for performance (LePine et al. 2005). We contribute to this discussion by estimating the short-term impact of CIs on operational work performance.

Most related to our work is a stream of literature on the impact of stressful or traumatic work-related events, often labeled as CIs, on the psychological wellbeing of workers. First, a subset of this literature is concerned with whether or not CIs affect workers' long-run wellbeing. In an overview of the literature on anesthesiologists' reactions to the death or serious injury of their patients, Gazoni et al. (2008) find that the average anesthesiologist experiences at least one patient death during their career and that at least one-third of them are profoundly affected by the experience. Furthermore, Gazoni et al. (2012) conduct a study among anesthesiologists to find that 84% of them had experienced an unanticipated death or serious injury of a perioperative patient, most of them experiencing a negative emotional effect and some never fully recovering. 67% of those encountering such an event believe that their ability to provide care was temporarily compromised following the event. Similarly, Pinto et al. (2013) find that surgeons report feeling emotionally affected by complications they experience in the operating room, which may affect their subsequent decision-making. In the context of forensic doctors, van der Ploeg et al. (2003) find that cases involving the death or suffering of children are considered to be the most disturbing and that experiencing more such events is associated with problems in the doctors' long-term coping with the associated trauma.

Second, a stream of literature has started examining to what extent encountering CIs affects subsequent performance or decision-making. Goldstone et al. (2004) survey consultant cardiac surgeons and anesthesiologists from UK cardiac surgery centers who report that average mortality rates do not increase following intraoperative patient deaths. However, the study's results suggest that morbidity *does* increase following such events and that mortality increases following intraoperative deaths if those were emergency or high risk cases. The mismatch between the reported and observed effects highlights that many of the operational performance inputs are difficult to observe, even for the affected workers. In related work, Hemmerich et al. (2012) show, through experiments with an abdominal surgery simulator, that physicians alter their decision-making after observing a bad outcome.

Some studies have been conducted in application domains which are similar to ours, i.e., in the context of first responders (fire fighters, police officers, and paramedics). Wilson et al. (1997) reached out to police officers who had been involved in CIs (defined as terrorist attacks in Northern-Ireland) 7-10 months earlier and find that 5% of survey participants fulfill criteria for post-traumatic stress disorder and 25% report symptoms consistent with mild to moderate depression. Similarly, routine stress has been associated with long-term distress. Liberman et al. (2002) analyze survey responses from police officers and find that exposure to routine occupational stress (excluding CIs) to be associated with psychological distress. Furthermore, they find that routine occupational stress to be a stronger predictor of psychological distress than cumulative exposure to CIs. Brunet et al. (2001) propose a measure for the peritraumatic stress (initial response to a CI), capturing whether the police officers experienced a range of negative emotions, bodily arousal, or a perceived life threat. Such immediate reaction to CIs was associated with subsequent symptoms of post traumatic stress disorder (PTSD). In the context of paramedics, Alexander and Klein (2001) collected data using a questionnaire and find that 90% of their sample of 110 ambulance workers had experienced a particularly disturbing incident (nominated by the subjects themselves) in the past six months. In addition, burnout and cases of general psychopathology were more likely in this group than their counterparts who did not experience a disturbing incident. Finally, for fire emergency responders, Monnier et al. (2002) develop a measure for CIs and find that workers' exposure to CIs is associated with their reporting of depressive symptoms and anger.

The above literature is distinct from our work on three key dimensions. First, it is mostly concerned with long-term effects (order of months or years) while our work is focused on the immediate short-term effects (for the rest of the shift) of encountering CIs. Second, the outcomes of interest for the above literature are usually measures of psychological well-being (e.g., symptoms of PTSD or depression), while we focus on the impact of CIs on operational field performance. Third, most of the above analysis is based on self-reported survey data, often well after the CI takes place, whereas our dataset provides a consistent, standardized, and contemporaneous outcome measure.

An open and often debated question in the existing literature is which factors might mitigate the potentially negative effects of encountering CIs. Gazoni et al. (2008) find that anesthesiologists who experience a CI often feel that they should stop for the day. Similarly, Smith and Jones (2001) write a note summarizing a questionnaire sent to surgeons, collecting data on whether they had experienced intraoperative deaths (53%) and whether they continued operating for the rest of their shift following the adverse event (81%). They conclude that there is no general consensus about how best to cope with intraoperative deaths, e.g., whether or not surgeons should discontinue operating for a period of time. In terms of other mitigating factors, (Pinto et al. 2013) interview surgeons and suggest that more experienced surgeons are less affected by—or may have developed better coping mechanisms for—complications in the operating room than their junior colleagues.

Our work is different from this literature in that using operational performance data allows us to trace how the impact of a prior CI on subsequent performance is diminishing as time passes from the CI itself, an important contributing factor when deciding whether to give workers time off following CIs. Similarly, the field data allows us to examine whether paramedic team characteristics, such as age and experience, have a mitigating effect on the impact of CIs on subsequent performance.

3. Empirical Setting

We conduct our empirical analysis on data describing each ambulance activation in the city of London during 2011, provided by the LAS. (The description of the LAS' organizational structure, operating procedures, and ambulance activations in this section are based on extensive interviews with the LAS operations managers, dispatchers, base managers, and paramedics as well as observational shifts in the LAS control room and ambulances.)

3.1. Organization

The LAS, a part of UK's National Health Service, is the busiest ambulance service in UK, its role being to respond to emergency calls and get medical help to patients with serious injuries or illnesses as quickly as possible. The LAS handles almost 2 million emergency calls and dispatches emergency vehicles to over one million incidents per year (LAS 2017).

3.2. Crews

While the LAS operates different types of vehicles (bicycles, cars, helicopters), their most common mode of emergency response is with ambulances. Ambulance crews generally (97% in our sample) consist of two paramedics. There are occasional exceptions to this, for example if a team leader from a base accompanies a crew to the scene to coordinate multiple resources for a large incident, but these are rare in our data. The paramedic crews of two are generally stable pairs who work together for multiple years. The only exception to this is new paramedic recruits, who are scheduled on a relief roster for the first year or two on the job (Akşin et al. 2018). To match the supply of ambulance resources with the highly variable demand for emergency response in London, the LAS shift roster consists of shifts of varying start times and varying durations (8, 10, or 12 hours). For a consistent set of observations, our analysis sample includes data from ambulance crews of two paramedics working together exclusively for an entire shift.

3.3. Ambulance Activation

For an emergency call which results in an ambulance activation the process works as follows.

Emergency call. The call is received by the control room staff at the LAS headquarters, who collect information about the patient, the incident, and the caller. Based on this information, the incident is triaged and passed on to one of eighteen dispatchers, who assign emergency vehicles to incidents, based on their geographical proximity.

The start of an ambulance activation is the dispatch notification from the control room, which appears as a notification on a screen in the ambulance, along with a severity estimate and location of the incident. Each activation then consists of five sub-processes.

Response. Upon dispatch, the crew must drive to the scene as quickly as possible. Usually the crew members alternate, with one of the two being responsible for the driving component for an entire shift.

Scene. Once the crew arrives at the scene they must locate, stabilize, and prepare the patient for transportation to a hospital. This is the most divergent and unpredictable sub-process of the activation, in the sense that the process steps themselves are unpredictable. Ambulance crews encounter a host of clinical and non-clinical issues which the crew must resolve together, usually under challenging circumstances and without any external support (Akşin et al. 2018, Shostack 1987). As a result the crew cannot follow a standard operating procedure at the scene, but must rely on their decision-making abilities in choosing the best course of action for each incident.

Transport. The crew must then transport the patient to the hospital. A screen in the ambulance lists the five geographically closest hospitals and the crew will usually choose the closest one, unless they have reasons to believe they can reach one of the other four sooner. While the driving paramedic operates the vehicle, the second paramedic stays in the back of the ambulance with the patient and completes the patient report form, noting down the patient details, the chief illness, and whether medicine was administered at the scene.

Handover. Upon arriving at the hospital, the crew joins the queue in front of the triage nurse station. As soon as the triage nurse becomes available, the crew communicates the information on the patient report form to him, who in turn decides whether to admit the patient to the resuscitation unit, the major incident unit, or the minor incident unit of the Accident & Emergency (hereafter A&E) department. This first part of the process is referred to as the clinical handover of the patient. Then the crew must physically move the patient to an assigned bed in the receiving unit. This second part of the process is referred to as the patient handover. In general, the handover process at the hospital is the most standardized sub-process of the ambulance activation, since the patient is already on a gurney and the paramedic crew can follow a standard operating procedure and rely on assistance from the hospital staff in case of any complications (Akşin et al. 2018).

Preparation. After the handover to the A&E department the crew are no longer responsible for the patient. They then do necessary preparations for the next ambulance activation (replenishing medicine, sheets, and other ambulance supplies as well as cleaning the ambulance) before they signal to the dispatchers in the LAS control room that they are available to be dispatched again—a step referred to as *going green*.

3.4. Dispatch Types

Not all ambulance dispatches result in the completion of all of the steps described above. A subset of the observations in our raw data describe ambulance dispatches which, for various reasons, are completed earlier (we will refer to these observations as *dispatches*, to distinguish from a full *activation*). Reasons for early completion vary, but include a dispatch being canceled before the ambulance arrives at the scene because another vehicle got there earlier or the paramedics deciding that the patient does not require A&E treatment (false alarms).

We define an activation as an observation which includes an entry for the hospital code (the ID of the receiving A&E) or an illness code (entered on the patient report form, which is later shared with the triage nurse at the A&E), indicating that the crew interacted with (diagnosed or treated) a patient and/or brought them to an A&E. We conduct our performance analysis on activations only, to ensure that our final sample includes observations describing a consistent set of tasks but excludes observations in which the crew may either have never reached the scene or not interacted with a patient. Making this distinction, the qualitative differences between dispatches and activations become evident in the data. The average completion time of activations (included in our performance analysis) is 74 minutes (interquartile range (IQR) 59-89, Std 25), but the average completion time of dispatches (excluded from our performance analysis) is 4.1 minutes (IQR 1.3-4.7, Std 6.8).

3.5. Performance Measurement

From an operational perspective, a key outcome for the LAS is the completion time of an ambulance activation. Shortening completion times corresponds to reducing ambulance utilization, which, in turn, enables swift response times. From a clinical perspective, obtaining hospital care as quickly as possible is beneficial for most health conditions. This is reflected in the targets the LAS sets for their paramedic crews. First, the LAS sets response time targets (i.e., for driving to the scene) which depend on the severity of the incident as determined during the call triaging process. Second, the objective for the handover process is that 85% (95%) of the time they should be completed in 15 (30) minutes. Third, the ambulance preparation step, before going green, is not supposed to take longer than 15 minutes. While the lack of explicit targets for the time spent at the scene reflects the unpredictable nature of the paramedics' job at the scene, they are strongly encouraged not to waste any time at the scene for the operational and clinical reasons mentioned above.

4. Critical Incidents and Operational Performance

In this section we develop our four sets of hypotheses, building on the existing literature summarized in §2 and the empirical setting described in §3.

4.1. CI impact on Subsequent Performance

We first consider the main research question of whether encountering CIs affects subsequent operational performance. Our setting of ambulance operations is an ideal setting to examine whether this is the case since it is well documented that paramedics are exposed to acutely stressful situations on a regular basis, which often has long-term consequences for psychological and emotional health (LeBlanc et al. 2011). In addition, they are presented with these unpredictable and challenging situations with little or no advance notice (Akşin et al. 2018). This unpredictability prevents psychological or behavioral preparation for the CI, which could otherwise mitigate the effects of acutely stressful events (Inzana et al. 1996).

Theory predicts that encountering a CI affects paramedics in a multitude of ways. By definition, a CI will overwhelm the paramedic's usual methods for coping with the stress of the job (Alexander and Klein 2001). In other words, paramedics have an emotionally distressed or anxious reaction to a situation due to their experiencing the demands of a given situation outweighing the resources they bring to bear (Weiss et al. 1995). In the context of PTSD diagnosis, encountering such events has been associated with impaired social, interpersonal, or occupational functioning (American Psychiatric Association 2013).

Specifically, the stress associated with CIs can affect different aspects of paramedics cognitive abilities, which are necessary for good job performance. First, multiple theories regarding the impact of stress on people's selective attention agree that stress depletes attentional resources due to cognitive system overload (Chajut and Algom 2003, LeBlanc 2009). This is likely to affect work performance, particularly in settings such as emergency medicine where an important component of the job is problem-solving based on the idiosyncratic characteristics of the job at hand, as described in §3.

Second, acutely stressful events such as CIs are likely to affect a worker's ability to store and retrieve relevant information. Elevated stress has been associated with reduced short-term ability to process and manipulate information in real-time (i.e., reduced working memory, see LeBlanc 2009). Similarly, encountering psychological stressors has been shown to negatively affect the retrieval of previously learned information (Domes et al. 2004, Kuhlmann et al. 2005). Both of these effects on a subject's memory abilities have been found more likely to occur if the subject experiences a threat response (elevated cortisol levels) to a stressor (Elzinga and Roelofs 2005, Buchanan et al. 2006). We expect the reduced ability to process and retrieve information to reduce a paramedic's capacity to choose the right operational or clinical action in ambulance activations immediately following a CI.

Third, research on the impact of acute stressors on team collaboration has found that team members interact less, due to a loss of team perspective, under acute stress (Driskell et al. 1999). In addition, team performance has been shown to deteriorate in stressful conditions, due to negative effects on mental models and transactive memory (Ellis 2006). Therefore, the completion times of ambulance activations following a CI are likely to be affected as much of the actions and decision-making of a paramedic crew during an ambulance activation is collaborative in nature.

In summary, various results in applied psychology predict reduced cognitive and decision-making abilities as a result of encountering acutely stressful situations such as CIs. We hypothesize that operational performance is affected for the remainder of the paramedic crew's shift, as follows;

H1a: Completion times increase for ambulance activations following a CI.

The theory that we summarize above predicts that operational performance is affected by CIs since acute stress overwhelms the paramedic's usual coping mechanisms as well as affecting their decision-making and collaboration. While this literature is not granular enough to specifically predict what happens if a worker encounters two such events in quick succession, we believe it is reasonable to expect that a second event of acute stress will worsen the impact of the first one;

H1b: The marginal impact of additional prior CIs, on completion times, is increasing.

4.2. Time-dependency of CI Impact

Above, we hypothesize about an average effect of encountering a CI on completion times of subsequent ambulance activations for the remainder of a shift. Although there is not a clear consensus about the persistence of the effects of acute stress, both survey results and biobehavioral experiments indicate that even short-term effects might be time-dependent.

In a survey of anesthesiologists, Gazoni et al. (2012) find that subjects often feel like CIs temporarily compromise their ability to perform on-the-job, with 84% of respondents feeling that time-off should be either offered or required. The exact duration of this temporary impact is unclear with 33% of responders favoring time off suggesting that the duration should be decided on a case by case basis but 35% (26%) suggesting that the rest of the day (and the following day) was appropriate. This indicates that anesthesiologists who have experienced CIs believe the effects to gradually taper off in the short term. In addition, from a biobehavioral point of view, neurological experiments have found the delay from the stressful event until a task is started to be an important moderating factor on the stress effects on task performance (Henckens et al. 2012, Shields et al. 2016).

Based on the above discussion we arrive at our second hypothesis;

H2: The impact of prior CIs on subsequent ambulance activation completion times is diminishing in the number of dispatches since the CI.

4.3. Process-dependency of CI Impact

Beyond the average impact on completion times, we are interested to understand which of the five sub-processes (see §3) of an ambulance activation are most affected. Evidence from prior literature suggests that the impact of various dimensions of prior experience can be process-type dependent. The dimensions on which processes differ include whether the process requires active reflection among team members (Arrow and McGrath 1993); the extent to which the process relies on transfer of non-codified knowledge (Lapré and Van Wassenhove 2001); the level of process compliance (Ton and Huckman 2008); and the level of process standardization (Akşin et al. 2018).

We examine the impact of prior CIs on the duration of each sub-process of ambulance activations. We expect some directional patterns based on the process types. In particular, the sub-processes differ on two main dimensions; the level of paramedic collaboration needed to complete the task and the level of external support or autonomy in decision-making. The starkest difference is between the patient pick-up at the scene and the patient handover at the hospital (these two sub-processes make up an average of 60% of the ambulance activation time). The process of locating, stabilizing, and preparing the patient for ambulance transport at the scene is a very collaborative effort. The paramedics must agree on a course of action to solve both clinical and operational challenges. In addition, they cannot rely on standard operating procedures (the process is highly divergent (Shostack 1987) and unpredictable) or external support. In contrast, the handover process at the hospital is more standardized, within and across hospitals. The paramedics can follow this process step-by-step and rely on assistance from the hospital staff in case of unanticipated complications.

This leads us to the following hypothesis describing the heterogeneous impact of CIs as a function of process type;

H3: The impact of prior CIs on subsequent ambulance activations is highest for sub-processes which rely on collaborative paramedic decision-making (particularly patient pick-up at scene) than sub-processes with substantial external support (particularly patient handover at the hospital).

The level of collaboration needed, and external support available, for the other sub-processes (driving to scene, driving to hospital, and ambulance preparation) can be argued to be intermediate. The driving sub-processes do not require the same level of collaboration as the patient pick-up at the scene. While they do certainly rely on the driving paramedic's skill and ability, the decision-making is supported by GPS navigation (and sometimes the emergency lights). Similarly, although the ambulance preparation is a somewhat collaborative task, it is fairly routine and requires little problem-solving. We therefore do not develop explicit hypotheses for these sub-process, but do interpret the results in §6.

4.4. Mitigation factors

Finally, we expect the performance impact of prior CI encounters to be heterogeneous across paramedics. Understanding which workers are likely to be affected by CIs is of importance to scheduling managers. We explore two potential mitigating factors in our analysis: average crew experience (with the LAS), and average crew age. The appeal of these variables is that they are observable and a scheduling manager could feasibly use them to inform task allocation or team formation.

We hypothesize that experienced workers recover from CIs more quickly, as shown in prior literature from different settings. For example, prior work has found that inexperienced surgeons experience higher stress levels during surgeries, i.e., are less capable of maintaining composure while performing the task at hand (Arora et al. 2010). Additionally, survey evidence from surgeons has revealed that more experienced surgeons are less affected by—or may have developed better coping mechanisms for—complications in the operating room than their junior colleagues (Pinto et al. 2013).

A priori, it is not clear how age, conditional on work experience, interacts with recovery from CIs. On the one hand, older workers may have a greater wealth of life experience to help them cope with stress. On the other hand, younger paramedics are likely to be better physically equipped to recover from stressful incidents such as CIs, and their paramedic training on building coping mechanisms is likely to be have been more recent.

This leads us to our final set of hypotheses;

H4a: The impact of CIs on subsequent completion times is diminishing in average crew experience.

H4b: The impact of CIs on subsequent completion times varies by average crew age.

5. Data and Empirical Strategy

Our dataset consists of all ambulance activations undertaken by the LAS in 2011. Variables include time-stamps for each of the activation sub-processes described in §3, the anonymized identity of the crew members, the illness code assigned by the paramedics once they have interacted with the patient, and the receiving hospital. The time-stamps we use to generate the dependent variables are either collected automatically or entered manually on the patient report form. Specifically, the dispatch, arrival at scene, departure from scene, and *going green* time-stamps are all recorded electronically through the press of a button in the ambulance. The handover time-stamps at the hospital are recorded manually on A&E forms and confirmed by A&E staff. The patient data and hospital information is collected on the patient report form while the crew information comes from an LAS human resources database.

We remove observations with obvious data-entry errors, specifically those for which the time-stamps indicate the wrong order of events or the ambulance activation duration exceeds 8 hours. In addition, we remove outliers by deleting observations for which one or more of the sub-process completion times exceed the 99.5 percentile in duration. Finally, we focus our performance analysis on ambulance activations (see §3.4) in which the crew interacted with a patient and/or brought them to an A&E. This ensures that each observation consists of the same set of sub-processes for which we can measure completion times. Our dataset for the main analysis describes 902,002 activations, undertaken by 4,040 paramedics, during 149,421 individual shifts. Our analysis of sub-processes relies on fewer observations as occasionally some of the within-activation time-stamps are missing. We include an overview of our data cleaning and exclusion in Table A1 and a summary of pairwise correlation between all variables in Table A4.

5.1. Main Independent Variables

Critical incident, as a theoretical construct, has a fairly consistent definition in the literature, as "*an incident that is sufficiently disturbing to overwhelm or threaten to overwhelm the individual's usual method of coping*" (Alexander and Klein 2001); as an event "*that involved actual or threatened death or serious injury or a threat to the physical integrity of self or others*" (Liberman et al. 2002), or as severe acute stressors (van der Ploeg et al. 2003). These definitions are in line with the American Psychiatric Association's definition of events that might subsequently result in PTSD, namely "exposure to actual or threatened death, serious injury, or sexual violence" in different forms, including through witnessing the event as it occurs or by being a first responder (see diagnostic criteria 309.81 and 308.3 in American Psychiatric Association 2013).

For our context, we define a CI as an ambulance activation where the patient has a high probability of dying at the scene—while in the paramedics' care. Specifically, prior analysis of the LAS data has revealed that three particular illness codes (as recorded by the crew on the patient report form) are associated with patients dying at the scene; *Purple* (an informal LAS term used for patients who are known to have died), *Cardiac Arrest*, and *Hanging* (i.e., a person who has hanged themselves) (Halter and Ellison 2008). This results in 8,404 (0.9%) ambulance activations being classified as CIs. Paramedic crews encounter at least one CI as part of 8,173 (5.5%) shifts, with 235 (0.1%) shifts involving two CIs. (We limit our analysis to two prior CIs, since only one shift during 2011 involved three CIs.) Based on this definition we define our main independent variables, below. (A robustness check using a broader definition of CIs is included in §6.5.)

One Prior CI. We first generate an indicator of a prior CI within a given paramedic crew's shift, denoted by $One\_Prior\_CI_{a,c}$. This variable is assigned a value of 0 at the start of every shift. It takes the value 1 for all ambulance activations a of paramedic crew c following a CI encounter, until the end of the shift.

Two Prior CIs. We generate a second indicator variable for two prior CIs within a given shift, denoted by $Two\_Prior\_CIs_{a,c}$. Analogously, it takes the value 1 for all activations once the paramedic crew has encountered two CIs during the shift, and 0 otherwise.

Dispatches Since CI. Part of our analysis focuses on whether the effects of CIs on paramedic crew performance diminish as the crew conducts more ambulance activations following the CI. To this end we generate two variables; $Dispatches\_Since\_First\_CI_{a,c}$ and $Dispatches\_Since\_Second\_CI_{a,c}$. The former takes the value 0 for all activations until (and including) the crew encounters the first CI of the shift and then counts how many dispatches the crew has been assigned to since the CI, for the remainder of the shift's activations. The second is defined analogously, but for the second CI that the crew encounters during the shift. Therefore, both variables will be non-zero following a second CI since they measure the number of dispatches since the first and second CI, respectively.

Team Experience. For our analysis of factors that may mitigate the negative impact of CI on paramedic crew performance, we generate a variable denoting the average experience of the crew, $Avg\_Team\_Experience_c$. This variable is defined as the average tenure at the LAS (years since joining the service) of the two members of the crew.

Team Age. Similarly, we define a variable denoting the average age of the paramedic crew, Avg_Age_c. For ease of interpretation, we demean Avg_Age_c in our analysis (see Table 5).

5.2. Main Dependent Variables

5.2.1. Completion Times. A main objective of our analysis is examining the impact of CIs on operational performance. As mentioned, we focus on CIs because of the people-centric mechanism through which they may impact operational performance; disturbing events are likely to challenge a worker's coping mechanisms, which has productivity impacts.

In line with this goal, our main outcome variable is $Completion\_Time_{a,c}$, defined as the time duration from dispatch until the crew becomes available again, for ambulance activation a, conducted by paramedic crew c. This is an aggregate measure of how long it takes a crew to reach a patient; to stabilize and prepare the patient for transport, at the scene; to drive the patient to the hospital; to hand the patient over to the receiving A&E department; and to prepare the ambulance vehicle for the next activation. Recall that this is an important outcome in the context of emergency medicine, for operational and clinical reasons. From an operational point of view, shortening completion time reduces utilization of resources, which in turn enables swift response times—an important objective for any ambulance service. From a clinical point of view, for most serious health conditions there are significant patient health benefits of getting to a hospital as quickly as possible (e.g., Sacco et al. 2005).

Note that we take higher completion times to be problematic because speed is an important outcome in ambulance services. That being said, one could argue that increased completion times following a CI could signal that the paramedic team is being more careful in their work. We think that this is unlikely for two reasons. First, we find that the completion time effect fades slightly over time, suggesting more of a coping mechanism reaction. Second, we also find effects on the sub-process of preparation time, which is the time *after* a job is completed and before the paramedic team is available for another job; this outcome does not directly affect the quality of care delivered. Still, we cannot fully rule out this potential alternative and view it as a limitation. An ideal marker of quality would be improved health outcomes for the patients in subsequent activations, but unfortunately we are unable to observe these health outcomes of interest because our data do not track patients after they are delivered to the hospital.

5.2.2. Activation Sub-processes. As described in §3, each ambulance activation comprises five distinct sub-processes. Completing each of those requires different skills and levels of collaboration among the paramedic crew. We, therefore, examine the impact of CIs on the completion times of each of those sub-processes. As before, we define each outcome for activation a by paramedic crew c.

First, we define outcomes for the pre-hospital sub-processes, which are solely the responsibility of the paramedic crew, with no external support. We denote the time from dispatch until arrival

Table 1 Summary statistics

	Mean	SD	N
Activations			903,787
Shifts			149,421
Critical Incidents			8,404
94.4% occurring in shifts with one CI			
5.5% occurring in shifts with two CIs			
Activation Characteristics			
Completion Time (Minutes)	73.52	25.01	903,787
Response Time	8.57	4.96	887,116
Scene Time	27.75	12.43	689,928
Transport Time	12.41	6.50	686,462
Handover Time	15.66	7.93	622,851
Preparation Time	15.72	8.25	626,281
Team Characteristics			
Avg Crew Experience (decades)	0.83	0.61	903,787
Avg Crew Age (decades)	3.77	0.76	903,787

at the scene by $Response\_Time_{a,c}$; the time spent at the scene by $Scene\_Time_{a,c}$; and the time spent driving to the hospital by $Transport\_Time_{a,c}$. The driving sub-processes can be thought of as individual tasks, as only one paramedic is at the wheel. However, the clinical and operational problem solving at the scene is a collaborative process which the parametric crew contributes to, jointly.

Second, we define the outcomes of the hospital sub-processes. The time spent handing the patient over to the A&E staff at the hospital is denoted by $Handover\_Time_{a,c}$, and the time spent preparing the ambulance for becoming available for dispatch again is denoted by $Preparation\_Time_{a,c}$. The former sub-process is an important outcome and a key performance indicator for both the LAS and the receiving hospital (as discussed in §3). Completing the handover is a standardized process which is completed jointly by the paramedic crew and the nursing staff at the A&E department. By contrast, the time spent preparing the ambulance for the next dispatch is determined by the paramedic crew only, since it does not require hospital resources, and has no implications for the patient.

5.3. Control Variables

In our analysis, we control for various shift (§5.3.1) and activation (§5.3.2) factors which may impact completion times, as well as seasonality (§5.3.3).

5.3.1. Shift Controls.

Crew Shift. We include a fixed effect for each shift of each paramedic crew. As a result, the focus of our analysis is strictly on short-term effects of CIs on productivity, within a given crew's shift.

In addition, the shift fixed effects control for any idiosyncrasies across individual paramedics or crews, due to their experience, general aptitude, or abilities. They also control for any changes in their motivation or effort across shifts.

Shift Fatigue. We include a linear and quadratic term of a variable measuring the time passed since the start of the shift. This is to control for a possible change in worker productivity as a result of fatigue, as the shift progresses (Danziger et al. 2011).

Shift Workload. To control for the well-known effects that workload can have on worker service rates (Delasay et al. 2018), we include a variable describing the workload experienced by the paramedic crew during the current shift up until the ambulance activation at hand. This is defined as the proportion of time the crew has been responding to calls since the start of the shift.

Shift Dispatch Count. In addition to the time-varying controls for experienced workload and fatigue, we include fixed effects for the number of ambulance dispatches which the crew has conducted as part of the shift at hand. This accounts for the possibility that having completed many previous dispatches could impact subsequent completion times of the paramedic crew.

5.3.2. Activation Controls.

Main illness. As the crew fills out the patient report form, they must note down a primary illness code (out of 98 distinct codes), describing the main illness of the patient. Since this code is recorded after the crew has been at the scene and observed the situation, it is a good descriptor of the type of incident. By including fixed effects for each illness code we account for the fact that different types of conditions require different activities at the scene.

Blue calls. A sub-sample (4.72%) of ambulance activations results in a *blue call*, by which the paramedic crew alerts the receiving hospital that they are on the way to the hospital, carrying a patient who needs treatment as soon as they arrive at the hospital. This action is reserved for patients who must be fast-tracked through the patient handover process at the hospital, which affects the overall completion time of the activation. We include a fixed effect to control for such cases. (We note that while these are, by definition, cases involving seriously ill patients, they do not necessarily constitute a CI from the perspective of the paramedic crew.)

Receiving hospital. We include fixed effects of the receiving hospital to control for idiosyncratic factors which may affect the handover process at the hospital.

5.3.3. Seasonality Controls. Finally, we control for seasonality using fixed effects for the hour-of-day, day-of-week, and month-of-year at the time of the ambulance activation. (Note that day-of-week and month-of-year fixed effects are identified using variation generated by shifts overlapping two calendar days or two calendar months, both of which occur in the data. The results are robust to excluding these two sets of fixed effects.)

5.4. Empirical Strategy

Here, we develop our regression specifications to test the set of hypotheses introduced in §4. The urban ambulance setting is empirically ideal for several reasons beyond being important in and of itself. First, CIs are assigned to crews only based on geographical proximity and availability, so there is effectively random assignment of paramedic crews to CIs. Figures A2(a) and A2(b) show the frequency of CIs for all levels of crew age and experience. Second, there is natural variation in the timing of CIs throughout the day. Figure A2(c) shows the distribution of CIs throughout the course of a shift. This is important for identification given our controls (in addition to shift fixed effects, we control for hour of day, day of week, and month of year effects, in addition to job order within the day); controlling carefully for the timing of incidents and job order is necessary to separate out the effects of CIs from potential confounds such as job fatigue. Third, the urban setting guarantees sufficient observations to identify the effects of CIs on subsequent job performance. Fourth, the ambulance setting generally is helpful because of the timestamp data available for not only overall job duration, but also for sub-processes (e.g., driving to the scene, time spent at the scene); this will allow us to estimate whether the effects of CIs differ for more versus less standardized components of each activation. Finally, ambulance services exist almost everywhere and paramedics are generally trained extensively in dealing with CIs; so to the extent that CIs affect performance among even these expert professionals, our results highlight the importance of people-centric factors on operational productivity.

Our analysis is concerned with the short-term (within-shift) performance impact of encountering a CI. We employ a shift-level difference-in-differences strategy in which the full dataset is used to estimate all control variables (§5.3) and the exogenous occurrence of CIs allows us a clean identification of the performance impact of the paramedic team having had one or two prior CIs (defined in §5.1). In Table A3 we provide a comparison of the characteristics of shifts with and without CIs. We find no meaningful difference between the two types of shifts.

We first explore our hypothesized main effect (**H1**), that CIs affect the average performance of subsequent jobs. We test this using the following specification;

$$Completion\_Time_{a,c} = \alpha + \beta_1 \; One\_Prior\_CI_{a,c} + \beta_2 \; Two\_Prior\_CIs_{a,c} + \gamma \; X_{a,c} + \epsilon_{a,c}, \qquad (1)$$

where $X_{a,c}$ denotes a vector of the shift, activation, and seasonality controls for activation a of paramedic crew c, described in §5.3. We note that the activation controls include illness codes, which control for the CIs themselves (making a specific CI control unnecessary). The errors are heteroskedasticity consistent and clustered by the crew shift. Note that β_2 captures the marginal performance impact of paramedic crews having encountered the second CIs within the shift. Therefore, the aggregate impact of two prior CIs is given by $\beta_1 + \beta_2$. In all our analysis we estimate models with and without Two_Prior_CIs. Since One_Prior_CI retains the value of 1 following the first CI (regardless of whether more CIs occur as part of the shift) it captures the average effect of one or more prior CIs in models where Two_Prior_CIs is not included.

Second, we examine whether there is evidence of a tapering-off of the impact of CIs on completion times, in accordance with **H2**. We use the following specification, adding the two measures for how many dispatches the paramedic crew has been assigned to since the first and second CIs of the shift.

$$\begin{aligned} Completion\_Time_{a,c} = {} & \alpha + \beta_1 \; One\_Prior\_CI_{a,c} + \beta_2 \; Two\_Prior\_CIs_{a,c} \\ & + \beta_3 \; Dispatches\_Since\_First\_CI_{a,c} + \beta_4 \; Dispatches\_Since\_Second\_CI_{a,c} \\ & + \gamma \; X_{a,c} + \epsilon_{a,c}. \end{aligned} \tag{2}$$

The above specification allows for a linear change in the impact of CIs on subsequent completion times. Since a paramedic crew will be dispatched on average 4.4 (2.9) times following the first (second) CI of the shift, we believe a linear term is sufficient to identify the first order time-dependency of the effect. As mentioned in §5.1, $Dispatches\_Since\_First\_CI_{a,c}$ ($Dispatches\_Since\_Second\_CI_{a,c}$) is only assigned non-zero values following a first (second) CI until the end of the shift.

Third, in line with **H3**, we examine how completion times of various sub-processes of the ambulance activation are affected by a prior CI. We re-estimate (1) after replacing the outcome variable by the completion times of each of the five ambulance activation sub-processes described in §3.3.

Using specification (1) in this manner will capture the average impact of a prior CI on subsequent completion times. However, as mentioned in §3, the sub-processes at the hospital (patient handover and ambulance preparation) have an explicit performance target of 15 minutes. Therefore, we also conduct analysis using a binary outcome of whether these targets were met. Specifically, we use a linear probability model (Greene 2002, Hellevik 2009), by re-estimating the coefficients of specification (1) with $\mathbf{1}_{\{Handover\_Time_{a,c}>15\}}$ and $\mathbf{1}_{\{Preparation\_Time_{a,c}>15\}}$ as the dependent variables, where $\mathbf{1}_A$ is an indicator variable for event A.

Finally, we examine whether there is evidence of the average age or experience of the paramedic crew having a mitigating effect on the impact of prior CIs on performance.

$$\begin{aligned} Completion\_Time_{a,c} = {} & \alpha + \beta_1 \; One\_Prior\_CI_{a,c} + \beta_2 \; Two\_Prior\_CIs_{a,c} \\ & + \beta_3 \; Avg\_Age_{a,c} \times One\_Prior\_CI_{a,c} + \beta_4 \; Avg\_Experience_{a,c} \times One\_Prior\_CI_{a,c} \\ & + \beta_5 \; Avg\_Age_{a,c} \times Two\_Prior\_CIs_{a,c} + \beta_6 \; Avg\_Experience_{a,c} \times Two\_Prior\_CIs_{a,c} \\ & + \gamma \; X_{a,c} + \epsilon_{a,c}, \end{aligned} \tag{3}$$

Table 2 The impact of critical incidents on activation completion time

	Completion Time (Minutes)			
	(1)	(2)	(3)	(4)
One Prior CI	1.953^{***}	1.930^{***}	2.282^{***}	2.285^{***}
	(0.242)	(0.242)	(0.283)	(0.284)
Two Prior CIs		3.583^{**}		6.345^{***}
		(1.409)		(1.869)
Dispatches Since the First CI			-0.097^{**}	-0.106^{**}
			(0.048)	(0.049)
Dispatches Since the Second CI				-0.797^{*}
				(0.417)
Shift Controls	✓	✓	✓	✓
Activation Controls	✓	✓	✓	✓
Seasonality Controls	✓	✓	✓	✓
Observations	902,002	902,002	902,002	902,002
Adjusted R^2	0.513	0.513	0.513	0.513

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

with $X_{a,c}$ and $\epsilon_{a,c}$ defined as before. The above specification uses the variation in the average crew age (standard deviation 7.6 years) and the average crew experience (standard deviation 6.1 years) to estimate the crew-type heterogeneity of the impact of CI on subsequent performance, as hypothesized in **H4**. Figure A1 describes the variation in average crew experience for each value of average crew age, among all the crew combinations that we observe in the sample. Although there is a clear relationship between crew age and experience, we observe substantial variation in experience, relative to age, which allows us to separately identify the two effects. Note that the impact of crew age and experience (both measured in decades) are time invariant. Therefore, while we can estimate their interactions with $One\_Prior\_CI_{a,c}$ and $Two\_Prior\_CIs_{a,c}$ (β_3 and β_4), their direct effects on completion time are not separately identified from the shift fixed effects.

6. Results

We now discuss our regression results, testing the hypotheses developed in §4 using the empirical specifications introduced in §5.4. In terms of interpretation, all our tables report the marginal effect of one (two) prior CIs on subsequent performance. Therefore, the aggregate impact of two prior CIs is the sum of the coefficient estimates for $One\_Prior\_CI_{a,c}$ and $Two\_Prior\_CIs_{a,c}$. (Note that following Correia (2015), we drop about 0.2% of singleton observations—e.g., shifts with only one activation—from the regression analysis to ensure proper inference. Therefore, there are small deviations from the sample sizes in Table 1 and the regression analyses.)

6.1. CI Impact on Subsequent Performance

Columns (1) and (2) in Table 2 report our main results for **H1**. We observe that paramedic crews who have encountered one prior CI spend on average 1.9 minutes (2.6%) longer completing each ambulance activation, for the remainder of the shift. The marginal impact of encountering a second CI is larger, with those crews spending an additional 3.6 minutes on subsequent activations. Taking the two effects together, crews who have encountered two prior CIs spend 7.5% (5.5 minutes, jointly significant at $p < 0.01$) longer completing ambulance activations than paramedic crews who have not encountered a CI. These results support **H1a** and **H1b**. Since our data includes fewer shifts with two CIs (233) than one CI (7,935), we run an extensive check to rule out concerns about influential points driving the results. Specifically, we re-estimate our model 233 times, removing one of the shifts with two CIs for each estimation. The difference in our coefficient estimates for One_Prior_CI and Two_Prior_CIs is less than 1% and 5%, respectively, as compared to our main estimates in Table 2. Both coefficients are statistically significant in each of those robustness checks.

These estimates represent an average impact on all subsequent ambulance activations for the remainder of the shift. On average a paramedic crew will conduct 2.9 and 2.0 (see Table A2) activations following a first and second CI, respectively—all of which will be affected. The paramedic utilization in our data is on average 80%, which is high for an emergency medicine setting relying on swift response times, so a 2.6%–7.5% increase in completion times for multiple ambulance activations is practically relevant.

6.2. Time-dependency of CI Impact

Columns (3) and (4) in Table 2 include our regression results for specification (2), which includes the $Dispatches\_Since\_First\_CI_{a,c}$ and $Dispatches\_Since\_Second\_CI_{a,c}$ variables. The coefficient estimates for these variables are negative and statistically significant, while the coefficients of $One\_Prior\_CI_{a,c}$ and $Two\_Prior\_CIs_{a,c}$ remain positive and significant. These results indicate that the impact of prior CIs on subsequent ambulance activations is not equal in magnitude for the remainder of the shift. Specifically, the effect is diminishing as the paramedics conduct more dispatches following the CI.

However, the magnitude of the demise is relatively small. Figure A3 shows the estimated impact for each dispatch, following one and two CIs. In summary, for each additional dispatch following the first or second CI the impact on activation completion times is reduced by 0.1 and 0.8 minutes, respectively. As mentioned above (see Table A2), paramedic crews conduct on average around 4.4 (2.9) dispatches following a first (second) CI until the end of the shift. In those cases, the impact of the first CI would reduce from 2.2 minutes to 1.8 minutes and the additional impact of the second CI, from 5.5 minutes to 4.0 minutes.

Table 3 The impact of critical incidents on sub-process completion time

	Ambulance Activation Sub-Processes (Minutes)				
	(1) Response Time	(2) Scene Time	(3) Transport Time	(4) Handover Time	(5) Preparation Time
Panel A					
One Prior CI	0.126^{**}	0.627^{***}	0.134	0.144	0.466^{***}
	(0.053)	(0.177)	(0.095)	(0.116)	(0.128)
Observations	885,167	685,107	681,505	614,910	618,572
Adjusted R^2	0.206	0.222	0.174	0.250	0.319
Panel B					
One Prior CI	0.122^{**}	0.612^{***}	0.131	0.137	0.459^{***}
	(0.053)	(0.178)	(0.095)	(0.116)	(0.128)
Two Prior CIs	0.640^{**}	1.359	0.323	0.531	0.577
	(0.303)	(0.975)	(0.544)	(0.622)	(0.716)
Observations	885,167	685,107	681,505	614,910	618,572
Adjusted R^2	0.206	0.222	0.174	0.250	0.319
Controls for Both Panels					
Shift Controls	✓	✓	✓	✓	✓
Activation Controls	✓	✓	✓	✓	✓
Seasonality Controls	✓	✓	✓	✓	✓

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3. Process-dependency of CI Impact

To examine the process-dependency of the impact of CIs on subsequent performance, we first discuss our analysis of the average outcomes of each of the five sub-processes introduced in §3 (Table 3). We then discuss the results of our analysis of the binary outcomes for handover time and preparation time (Table 4).

Average effects. Table 3 provides coefficient estimates for the impact of encountering a CI on subsequent performance for each of the five sub-processes of an ambulance activation. Panel A summarizes the average impact of a single CI on subsequent activations, while Panel B allows for a marginal impact of a second CI, as well. The results across panels are consistent in coefficient size and significance for the impact of one prior CI, so we focus our attention on Panel B.

We note that the coefficient estimates are positive for all sub-processes for both one or two CIs, indicating longer completion times following CIs. In line with **H3**, the highest coefficient estimate is for scene time, whereas we find no significant effect of either one or two prior CIs on handover time at the hospital. In relative terms, paramedics who have encountered a single prior CI spend on average 2.2% longer at the scene. While the coefficient estimate of the marginal impact of a second CI on scene time is larger (1.4 minutes or 4.9%), the standard errors are large and the estimate is not statistically significant.

We observe a meaningful impact of one prior CI on subsequent preparation times. The coefficient estimate in column (5) of Table 3 indicates that the paramedic crews take, on average, half a minute longer to prepare the ambulance for the next activation if they encountered one prior CI. Similar to scene times, the coefficient estimate for the second prior CI is twice as large but with the smaller sample the standard errors also increase so the effect is not statistically significant.

Interestingly, we observe a statistically significant effect on one of the driving times (response time) but not the other (transport time). The effect of one or two CIs is a relative 1.4% and 7.5% (8.9% in aggregate) increase in response time, a key performance indicator (for serious cases) at the LAS. The coefficient estimates in column (3) indicate a 1.1% and 3.7% increase in transport times for one and two prior CIs, respectively, but these effects are not statistically significant. A potential explanation is that usually ambulances drive to the scene using emergency lights to get through traffic. In contrast, the paramedic crew will only use the emergency lights when driving to the hospital if the patient is in a critical condition. Hence, response times may be more affected by the skill and ability of the driving crew member whereas the transport times are more noisy and dependent on external factors such as traffic.

As mentioned, the primary effect of a CI was a 2.6% increase in activation completion times for the remainder of the shift (column (1) in Table 2). Rounding the effect to 3%, we next investigate whether such an effect size can be ruled out given the variability and sample size of each sub-process. Using the estimates reported in Panel A of Table 3, we conduct a Wald test for a hypothesized effect of a relative 3% for each sub-process. A 3% impact on response time, transport time, and handover time would correspond to an effect of 0.26 min, 0.37 min, and 0.46 min, respectively. The estimates in columns (1), (3), and (4) of Table 3 suggest a relative impact of less than 3% for each of those processes. The p-values of a Wald test for a 3% impact are 0.014, 0.013, and 0.01, respectively. We, therefore, feel comfortable ruling out a 3% impact for each of those three sub-processes. In contrast, for scene time and preparation time, the p-values corresponding to a test for a 3% impact (0.83 min and 0.47 min) are 0.25 and 0.97, respectively. Therefore, we cannot reject a hypothesis of a 3% impact.

In summary, for transport time and handover time we cannot reject a hypothesis of no effect but can confidently reject an alternative hypothesis of a 3% effect. For scene time and preparation time we can reject a hypothesis of no effect but cannot reject a hypothesis of a 3% impact. For the response time we can reject both a hypothesis of no effect and a hypothesis of a 3% impact, indicating a statistically significant but relatively smaller effect than for scene time and preparation time. These results are in line with our **H3**.

Table 4 Linear probability model outcomes for missing sub-process performance targets

	(Handover Time > 15min)		(Preparation Time > 15min)	
	(1)	(2)	(3)	(4)
One Prior CI	0.014*	0.013*	0.029***	0.029***
	(0.008)	(0.008)	(0.008)	(0.008)
Two Prior CIs		0.030		0.058
		(0.043)		(0.044)
Shift Controls	✓	✓	✓	✓
Activation Controls	✓	✓	✓	✓
Seasonality Controls	✓	✓	✓	✓
Observations	614,910	614,910	618,572	618,572
Adjusted R^2	0.195	0.195	0.216	0.216

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Binary outcomes. All our main analysis uses the average completion time of ambulance activations (and sub-processes thereof) as a dependent variable, since this is an important system outcome which determines utilization. However, given that the LAS has explicit performance targets for handover time and preparation time (see §3), we also run a linear probability model where the outcome variable is an indicator for missing these targets and report our results in Table 4. These results are consistent with our main analysis. On average, 46.2% and 47.8% of the activations meet their target of 15 minutes for handover time and preparation time, respectively. While we found no significant impact of prior CIs on handover times in Table 2, we observe a marginally significant effect ($p < 0.1$) on the probability of missing the performance target of keeping handover times below 15 minutes. Specifically, the results indicate a 2.8% (1.3 percentage points) increase in the probability of missing the handover time target following CIs (column (2)). While this suggests there is some effect of prior CIs on handover performance, it seems small in magnitude and only marginally significant. This is in line with **H3**.

Regarding preparation time, the results of the linear probability model are consistent with those in Table 2. In particular, we observe a 6.1% increase ($p < 0.01$) in the probability of missing the performance target of 15 minutes, following a prior CI. Also, while the coefficient estimate for the marginal impact of a second CI is not statistically significant, the joint impact of two prior CIs on performance is significant ($p < 0.05$) and predicts an 18.2% increase in the probability of missing the performance target.

6.4. Mitigation Factors

Table 5 includes coefficient estimates for our analysis of the mitigation effects of age and experience, using specification (3). As discussed in §5, $Avg\_Crew\_Age_c$ (mean 3.77) and

Table 5 The moderation of impacts by paramedic experience and age

	Completion Time (Minutes)					
	(1)	(2)	(3)	(4)	(5)	(6)
One Prior CI	1.981^{***}	1.940^{***}	2.028^{***}	1.958^{***}	1.918^{***}	2.006^{***}
	(0.242)	(0.242)	(0.244)	(0.242)	(0.243)	(0.244)
Two Prior CIs				3.570^{**}	3.500^{**}	3.521^{**}
				(1.405)	(1.405)	(1.466)
Avg Crew Age × One Prior CI	1.064^{***}		1.656^{***}	1.055^{***}		1.657^{***}
	(0.299)		(0.433)	(0.301)		(0.435)
Avg Crew Age × Two Prior CIs				0.469		0.029
				(1.811)		(3.070)
Avg Crew Experience × One Prior CI		0.429	-1.019^{**}		0.410	-1.038^{**}
		(0.357)	(0.517)		(0.358)	(0.518)
Avg Crew Experience × Two Prior CIs					0.870	0.767
					(2.426)	(4.071)
Shift Controls	✓	✓	✓	✓	✓	✓
Activation Controls	✓	✓	✓	✓	✓	✓
Seasonality Controls	✓	✓	✓	✓	✓	✓
Observations	902,002	902,002	902,002	902,002	902,002	902,002
Adjusted R^2	0.513	0.513	0.513	0.513	0.513	0.513

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$Avg\_Crew\_Experience_c$ (mean 0.83) are defined as the average age and experience of the paramedics on the team, in decades.

For completeness, Table 5 includes estimates of the mitigating effects of average crew age and average crew experience on the performance effect of prior CIs. We recognize that these variables have a strong correlation (0.668), but there remains a considerable degree of variation in crew experience conditional on age as shown in Figure A1. Given the potential managerial insights related to crew age and experience, we focus our interpretation of the mitigation effects on columns (3) and (6) which include both variables. We do note, however, that the interpretation of these coefficients warrants some caution given their correlation.

We observe that the interaction of $One\_Prior\_CI_{a,c}$ and $Avg\_Crew\_Age_c$ is positive and significant, indicating that crews of older average age are more affected by CIs. However, the interaction of $One\_Prior\_CI_{a,c}$ and $Avg\_Crew\_Experience_c$ is negative and significant. Interpreting the entries of column (6) we observe that conditional on job experience, older paramedic crews have worse performance following a CI. In contrast, conditional on their age, paramedic crews with more on-the-job experience are less affected by a prior CI. Since the coefficient of the age interaction is higher than the coefficient of the experience interaction, the aggregate impact of a prior CI increases as paramedics gain experience on the job. In particular, as a crew gains a decade of experience (and simultaneously a decade of added age) the impact of a prior CI increases by 0.6 minutes. For example, the completion time of paramedics who joined the service at age 25 (45) increases from 0.8 (4.1) to 1.4 (4.7) minutes following a CI as they gain a decade of experience.

While the coefficients for the second CI and the corresponding interactions are all directionally the same as for the first CI, the results are not significant. This is likely attributable to the lack of statistical power, as there are many fewer shifts with two CIs versus one.

6.5. Robustness

6.5.1. Broader Definition of CIs. For our main analysis, we define CIs based on the patient report form illness codes which have been associated with a high probability of patients dying at the scene (Halter and Ellison 2008). As described in §5.1, these illness codes correspond to ambulance activations in which a patient has died, hanged themselves, or is in cardiac arrest. We conduct two additional sets of analysis to examine the robustness of our results to the definition of CIs.

First, there are other illness codes which could reasonably be considered as emotionally and psychologically demanding. In particular, cases where the patient is suffering from a *major head injury*, *multiple injuries*, or *a weapon (knife or gun) wound*, can be argued to fit the definition of a CI. We repeat our main analysis (specification (1)) using a broader definition of CIs which includes these illness codes. This results in 14,136 ambulance activations being classified as CIs (compared to the 8,404 in our main analysis). As before, the majority of shifts involving CIs have either one (95.0%) or two (4.8%) such events. Therefore, we continue to focus on the impact of one or two prior CIs, using specification (1). Using this broader definition of CIs, we find that our main results hold, with a positive and significant coefficients of $One\_Prior\_CI_{a,c}$ and $Two\_Prior\_CIs_{a,c}$. The coefficients are slightly smaller in magnitude, indicating that the cases we have added to the definition of CIs do not have as much impact on subsequent performance as those in the original definition.

Second, we repeat our analysis by using blue calls as the definition of CIs. This results in 4.72% of our observations being classified as CIs. With this definition, 19.8% (3.7%) of shifts include one (two) CIs. As before, we estimate equation (1) and report the results in Table A5. We find that with this definition our results continue to hold, with a positive and significant coefficients of $One\_Prior\_CI_{a,c}$ and $Two\_Prior\_CIs_{a,c}$, although the magnitude of the effects are slightly lower than in our main analysis.

6.5.2. Log-linear Specification. We repeat our main analysis with a natural log transformation of completion time as the dependent variable. These estimates are in columns (3) and (4) of Table A5. The estimated marginal effects from this analysis are consistent with our main estimates with slightly larger magnitudes. In particular, the impacts of one or two prior CIs on subsequent performance are statistically significant. We also find that paramedic crews encountering a single CI spend 3% longer completing their subsequent activations, and the estimated marginal impact of a second CI on subsequent completion times is 6.3%.

6.5.3. Matching Analysis. In addition to our main analysis we estimate a matched analysis. We match shifts with a CI with ones without a CI by the same crew in the same day of week. This matching provides us with shifts that are comparable in term of the crew and the weekly schedules and limits our sample to 8,894 shifts where exactly half of them (4,447) have a CI. We run our difference-in-differences estimation using only this sample. The results are presented in column 2 of Table A6. The estimates are similar to the ones in the main analysis with slightly larger magnitudes.

6.5.4. Placebo Analysis. In column 1 of Table A6, we include a placebo test to rule out the possibility that the significant coefficients of our variables of interest are due to spurious correlations. To this end, we randomly designate a subset of our observations to be faux-CIs (with the same frequency as the real CIs) and repeat the analysis. Our coefficients of interest are not statistically significant in this analysis, which further suggests that our analysis is identifying the impact of actual CIs on subsequent performance.

6.5.5. Other Robustness Checks. Finally, we conduct three additional robustness checks and report the results in columns (3)-(5) of Table A6. First, we repeat the analysis without our set of activation controls to alleviate concerns regarding our model specification. Second, we remove the first activation of each shift since, by definition, these activations can never follow a CI. Third, we add a dummy variable indicating whether a given activation was the last dispatch of the shift, to rule out the possibility of end-of-shift effects driving our results. In all cases the coefficients of our variables of interest are positive and statistically significant ($p < 0.01$) with the coefficient sizes comparable to the main analysis.

7. Conclusion

Operations management scholars and practitioners are increasingly interested in understanding and managing people-centric determinants of work performance. These include psychological, social, and environmental factors which can impact operational outcomes through various mechanisms. An important aspect of organizational performance in many service systems is the task completion time of workers. This outcome is not only a key performance indicator in service operations settings, such as ambulance services, but also an empirical signature of workers' emotional states.

In this paper, we focus on one such people-centric factor, CIs, which are defined as jobs which are likely to be disturbing enough to overwhelm the workers' usual coping mechanisms (Alexander and

Klein 2001). It is well documented that such events can have a lasting impact on the psychological well-being of workers (Wilson et al. 1997, Liberman et al. 2002, Monnier et al. 2002), but their impact on operational performance has not been studied previously. Our work makes a first attempt at understanding how such emotionally and psychologically difficult events impact task completion times. Our setting of emergency response medicine is ideal for this analysis as the assignment of these events to specific paramedic crews is effectively random. Furthermore, the completion time of tasks is an important performance indicator, both from clinical and operational perspectives.

Our work makes four main contributions. First, we find that encountering a CI has a significant immediate impact on subsequent operational performance. Specifically, completion times increase by 1.9 (5.5) minutes following one CI (two CIs), which corresponds to a 2.6% (7.5%) increase over shifts with no CIs. When interpreting the effect size it is worth noting that while CIs, by definition, are events which have the potential to overwhelm workers coping mechanisms, encountering such events is part of a paramedic's job (occurring at least once in 5% of shifts). In other words, we observe a practically and statistically significant impact of CIs on the performance of precisely the workers who should be best equipped (through training and experience) to deal with such events.

This main finding also contributes to the psychology literature (see §2) on CIs and occupational stress by showing that the impact of critical or traumatic events is not only an emotional or psychological matter (through symptoms of burnout, PTSD, or depression) for the workers involved. Rather, encountering CIs has an immediate short-term impact on operational performance and therefore has important implications for organizational performance.

Second, we find evidence of a short-term time-dependency of the effect. Specifically, the impact of a prior CI is highest on the ambulance activations immediately following the event and then diminishing for the remainder of the shift. However, this tapering-off effect is relatively small in magnitude indicating the performance is affected throughout the shift.

The magnitude of this diminishing effect is a key contribution of our work. Most of the prior research on the impact of CIs on worker's wellbeing has been conducted via surveys administered long after events may have taken place (Wilson et al. 1997, Liberman et al. 2002, Monnier et al. 2002). It is, therefore, difficult to evaluate the duration of the short-term effects of prior CIs. The consistent operational outcome measures which are contemporaneously collected in our setting allow us to quantify this diminishing effect. Therefore, these findings have managerial implications for scheduling managers, who benefit from understanding how persistent the effects of CIs are when deciding which jobs to assign to workers who have recently encountered a CI.

Third, we find evidence for process-dependency of the impact of CIs. Specifically, we observe a strong impact of CIs on paramedic performance in sub-processes in which they have substantial decision-making autonomy and completion times are determined exclusively by their skill and ability, as opposed to processes where they can rely on more external support. This is a somewhat intuitive result, but it has implications for the generalizability of our findings. This result suggests that operations managers in settings where processes are standardized and workers have access to operational support can be less worried about the operational performance impact of CIs than those in settings where workers operate more independently. Of course, in both settings, managers might consider taking action following CIs for reasons beyond operational performance, such as burnout or long-term occupational stress.

Fourth, we observe that conditional on the average age of a paramedic crew, more experience helps mitigate the detrimental performance impact of prior CIs. However, we find that paramedic crews of higher age are more affected by prior CIs and that this effect outweighs the benefits associated with gaining experience. This indicates that while experience is beneficial the lowest impact of CIs is on younger workers. This is in contrast with prior research which has found younger doctors to be more likely to suffer from symptoms of stress and depression, although in that profession the reasons sometimes have to do with age-related differences in working conditions (Rashid and Talib 2015).

Given that CIs are unpredictable in nature, the actionable managerial implications of our results relate to how operations managers should react in the aftermath of CIs. First, the fact that we observe a persistent negative impact on performance following CIs can present operations managers with a challenge. On the one hand, in settings where consistency in performance is key and managers have ample capacity, they would be advised to not assign new jobs to teams with recent CI experiences. As such our findings provide some empirical, operational grounding for the stated preferences of anesthesiologists and surgeons who often choose to discontinue their shifts following CIs (Gazoni et al. 2008, Smith and Jones 2001). On the other hand, in high utilization settings where discontinuing the shift of some workers might have a negative externality on system performance, operations managers would be advised to assign teams with recent CI experiences to lower priority tasks, where the impact of diminished operational performance is minimized. Future research is needed to establish whether the effects of CIs on subsequent performance are limited to operational outcomes or whether quality (e.g., health outcomes) is also affected. If such a quality effect is established for important outcomes, the former approach of not assigning any new jobs to teams who have encountered CIs would be recommended.

Second, while our setting does not allow for a comprehensive analysis of different process types (since we only have 5 sub-processes), our results suggest a process-dependency of the CI impact, which has potential managerial implications. Intuitively, our results (in support of **H3**) suggest that worker performance is more affected by a prior CI if the operational process at hand requires workers to engage in collaborative decision-making without external support. This indicates that operational managers should try, to the extent possible, to increase the external decision-making support available to workers, once they have encountered a CI.

We believe our findings present a number of future avenues for research. First, we find a significant effect of CIs on operational performance. As we mention above, an important next question is to understand whether CIs also affect other quality metrics (such as health outcomes). Second, much of the psychology literature on CIs and occupational stress is focused on long-term effects on the emotional and psychological well-being of workers. Since we find that the operational impact of CIs is persistent until the end of the shift an important next step is to evaluate the long-term operational effects. Third, while CIs are (fortunately) fairly infrequent, workers in many fields (paramedics included) can be expected to encounter a number of them during their tenure. Understanding the short- and long-term performance impact of cumulative exposure to CIs over time would be useful for operations managers in settings where the exposure to CIs over time can be managed. Fourth, theory predicts that the effects we find are driven by psychological, cognitive, and emotional mechanisms. However, another potential mechanism is deterioration in physical capacities. Unfortunately, our data does not allow us to clearly distinguish between those mechanisms. Therefore, further research (using data from other settings) is required to separate psychological effects from physical ones. Finally, additional mitigation factors should be explored. For example, there are suggestions in the literature that implicit coordination between team members reduces the detrimental effects of stress in the workplace (LeBlanc 2009). Therefore, it would be valuable to understand which well-known constructs from the operational productivity literature, such as team familiarity, can mitigate the performance impact of CIs.

Acknowledgments

We thank the London Ambulance Service for providing data, information about the study setting, and multiple opportunities to observe ambulance crews, dispatchers, and hospital staff in action. In particular, we thank Leanne Smith, Michael Damiani, and Patrick Brooks. We are grateful for excellent research assistance from Atikha Rizwan and for financial research support from the Wisconsin Alumni Research Foundation at the University of Wisconsin—Madison.

References

Akşin, Zeynep, Sarang Deo, Jónas Oddur Jónasson, Kamalini Ramdas. 2018. Learning from many: Partner exposure and team familiarity in fluid teams. *Working Paper.*

Alexander, David A, Susan Klein. 2001. Ambulance personnel and critical incidents: impact of accident and emergency work on mental health and emotional well-being. *The British Journal of Psychiatry* **178**(1) 76–81.

Altman, Daniel, Galit B. Yom-Tov, Marcelo Olivares. 2018. Do customer emotions affect worker productivity? an empirical study of emotional load in online customer contact centers. *Working Paper*.

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. Fifth ed. American Psychiatric Association, Arlington, VA.

Argote, Linda, Dennis Epple. 1990. Learning curves in manufacturing. *Science* **247**(4945) 920–924.

Arora, Sonal, Nick Sevdalis, Debra Nestel, Maria Woloshynowych, Ara Darzi, Roger Kneebone. 2010. The impact of stress on surgical performance: a systematic review of the literature. *Surgery* **147**(3) 318–330.

Arrow, Holly, Joseph E McGrath. 1993. Membership matters: How member change and continuity affect small group structure, process, and performance. *Small Group Research* **24**(3) 334–361.

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

Berry-Jaeker, Jillian A, Anita L Tucker. 2016. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Science* **63**(4) 1042–1062.

Boh, Wai Fong, Sandra A Slaughter, J Alberto Espinosa. 2007. Learning from experience in software development: A multilevel analysis. *Management Science* **53**(8) 1315–1331.

Brunet, Alain, Daniel S Weiss, Thomas J Metzler, Suzanne R Best, Thomas C Neylan, Cynthia Rogers, Jeffrey Fagan, Charles R Marmar. 2001. The peritraumatic distress inventory: a proposed measure of ptsd criterion a2. *American Journal of Psychiatry* **158**(9) 1480–1485.

Buchanan, Tony W, Daniel Tranel, Ralph Adolphs. 2006. Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & Memory* **13**(3) 382–387.

Chajut, Eran, Daniel Algom. 2003. Selective attention improves under stress: implications for theories of social cognition. *Journal of Personality and Social Psychology* **85**(2) 231.

Clark, Jonathan R, Robert S Huckman, Bradley R Staats. 2013. Learning from customers: Individual and organizational effects in outsourced radiological services. *Organization Science* **24**(5) 1539–1557.

Correia, Sergio. 2015. Singletons, cluster-robust standard errors and fixed effects: A bad mix. URL `http://scorreia.com/research/singletons.pdf`.

Dai, H., K.L. Milkman, D.A. Hofmann, B.R. Staats. 2015. The impact of time at work and time off from work on rule compliance: The case of hand hygiene in health care. *Journal of Applied Psychology* **100**(3) 846.

Daniels, K, C Harris. 2000. Work, psychological well-being and performance. *Occupational Medicine* **50**(5) 304–309.

Danziger, Shai, Jonathan Levav, Liora Avnaim-Pesso. 2011. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences* **108**(17) 6889–6892.

Delasay, Mohammad, Armann Ingolfsson, Bora Kolfal, Kenneth Schultz. 2018. Load effect on service times. *European Journal of Operational Research* .

Deo, Sarang, Aditya Jain. 2018. Slow first, fast later: Emprical evidence of speed-up in service episodes of finite duration. *Working Paper*.

Domes, Gregor, Markus Heinrichs, Ulrike Rimmele, Ursula Reichwald, Martin Hautzinger. 2004. Acute stress impairs recognition for positive words-association with stress-induced cortisol secretion. *Stress* **7**(3) 173–181.

Driskell, James E, Eduardo Salas, Joan Johnston. 1999. Does stress lead to a loss of team perspective? *Group Dynamics: Theory, Research, and Practice* **3**(4) 291.

Ellis, Aleksander PJ. 2006. System breakdown: The role of mental models and transactive memory in the relationship between acute stress and team performance. *Academy of Management Journal* **49**(3) 576–589.

Elzinga, Bernet M, Karin Roelofs. 2005. Cortisol-induced impairments of working memory require acute sympathetic activation. *Behavioral Neuroscience* **119**(1) 98.

Freeman, Michael, Nicos Savva, Stefan Scholtes. 2016. Gatekeepers at work: An empirical analysis of a maternity unit. *Management Science* **63**(10) 3147–3167.

Gazoni, Farnaz M, Peter E Amato, Zahra M Malik, Marcel E Durieux. 2012. The impact of perioperative catastrophes on anesthesiologists: results of a national survey. *Anesthesia & Analgesia* **114**(3) 596–603.

Gazoni, Farnaz M, Marcel E Durieux, Lynda Wells. 2008. Life after death: the aftermath of perioperative catastrophes. *Anesthesia & Analgesia* **107**(2) 591–600.

Goldstone, Antony R, Christopher J Callaghan, Jon Mackay, Susan Charman, Samer AM Nashef. 2004. Should surgeons take a break after an intraoperative death? attitude survey and outcome evaluation. *BMJ* **328**(7436) 379.

Greene, William H. 2002. The behavior of the fixed effects estimator in nonlinear models .

Halpern, Janice, Robert G. Maunder, Brian Schwartz, Maria Gurevich. 2012. The critical incident inventory: characteristics of incidents which affect emergency medical technicians and paramedics. *BMC Emergency Medicine* .

Halpern, Janice, Robert G Maunder, Brian Schwartz, Maria Gurevich. 2014. Downtime after critical incidents in emergency medical technicians/paramedics. *BioMed Research International* **2014**.

Halter, Mary, George Ellison. 2008. *Evaluation of the emergency care practitioner role in London: a study of the processes and outcomes of clinical decision making*. Faculty of Health and Social Care Sciences, Kingston University and St George's, University of London.

Hellevik, Ottar. 2009. Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity* **43**(1) 59–74.

Hemmerich, Joshua A, Arthur S Elstein, Margaret L Schwarze, Elizabeth Ghini Moliski, William Dale. 2012. Risk as feelings in the effect of patient outcomes on physicians' future treatment decisions: A randomized trial and manipulation validation. *Social Science & Medicine* **75**(2) 367–376.

Henckens, Marloes JAG, Guido A van Wingen, Marian Joëls, Guillén Fernández. 2012. Time-dependent effects of cortisol on selective attention and emotional interference: a functional mri study. *Frontiers in Integrative Neuroscience* **6** 66.

Huckman, Robert S, Gary P Pisano. 2006. The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science* **52**(4) 473–488.

Huckman, Robert S, Bradley R Staats. 2011. Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *Manufacturing & Service Operations Management* **13**(3) 310–328.

Huckman, Robert S, Bradley R Staats, David M Upton. 2009. Team familiarity, role experience, and performance: Evidence from indian software services. *Management Science* **55**(1) 85–100.

Ibanez, Maria, Michael W Toffel. 2018. How scheduling can bias quality assessment: Evidence from food safety inspections .

Ibanez, Maria R., Jonathan R. Clark, Robert S. Huckman, Bradley R. Staats. 2018. Discretionary task ordering: Queue management in radiological services. *Management Science* **64**(9) 4389–4407.

Inzana, Carolyn M, James E Driskell, Eduardo Salas, Joan H Johnston. 1996. Effects of preparatory information on enhancing performance under stress. *Journal of Applied Psychology* **81**(4) 429.

Jex, Steve M. 1998. *Stress and job performance: Theory, research, and implications for managerial practice*. Sage Publications Ltd.

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

KC, Diwas Singh, Bradley R Staats. 2012. Accumulating a portfolio of experience: The effect of focal and related experience on surgeon performance. *Manufacturing & Service Operations Management* **14**(4) 618–633.

KC, Diwas Singh, Bradley R Staats, Maryam Kouchaki, Francesca Gino. 2017. Task selection and workload: A focus on completing easy tasks hurts long-term performance *Working Paper*.

Kim, Song-Hee, Carri W Chan, Marcelo Olivares, Gabriel Escobar. 2014. Icu admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* **61**(1) 19–38.

Kim, Song-Hee, Hummy Song, Melissa Valentine. 2018. Cross-disciplinary temporary teams in health care organizations: The role of partner variety and shared experience. *Working Paper*.

Kuhlmann, Sabrina, Marcel Piel, Oliver T Wolf. 2005. Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience* **25**(11) 2977–2982.

Kuntz, Ludwig, Roman Mennicken, Stefan Scholtes. 2014. Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science* **61**(4) 754–771.

Lapré, Michael A, Luk N Van Wassenhove. 2001. Creating and transferring knowledge for productivity improvement in factories. *Management Science* **47**(10) 1311–1325.

LAS. 2017. Annual Report 2016-2017. Tech. rep., London Ambulance Service.

LeBlanc, Vicki R. 2009. The effects of acute stress on performance: Implications for health professions education. *Academic Medicine* **84**(10) S25–S33.

LeBlanc, Vicki R, Cheryl Regehr, Arija Birze, Kevin King, Aristathemos K Scott, Russell MacDonald, Walter Tavares. 2011. The association between posttraumatic stress, coping, and acute stress responses in paramedics. *Traumatology* **17**(4) 10–16.

LePine, Jeffery A, Nathan P Podsakoff, Marcie A LePine. 2005. A meta-analytic test of the challenge stressor–hindrance stressor framework: An explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal* **48**(5) 764–775.

Liberman, Akiva M, Suzanne R Best, Thomas J Metzler, Jeffrey A Fagan, Daniel S Weiss, Charles R Marmar. 2002. Routine occupational stress and psychological distress in police. *Policing: An International Journal of Police Strategies & Management* **25**(2) 421–441.

Marmar, Charles R, Daniel S Weiss, Thomas J Metzler, Kevin L Delucchi, Suzanne R Best, Kathryn A Wentworth. 1999. Longitudinal course and predictors of continuing distress following critical incident exposure in emergency services personnel. *The Journal of Nervous and Mental Disease* **187**(1) 15–22.

Monnier, Jeannine, Rebecca P Cameron, Stevan E Hobfoll, J Robert Gribble. 2002. The impact of resource loss and critical incidents on psychological functioning in fire-emergency workers: A pilot study. *International Journal of Stress Management* **9**(1) 11–29.

Pinto, A, O Faiz, C Bicknell, C Vincent. 2013. Surgical complications and their implications for surgeons' well-being. *British Journal of Surgery* **100**(13) 1748–1755.

Rashid, Irfana, Parvaiz Talib. 2015. Occupational stress and coping styles among doctors: role of demographic and environment variables. *Vision* **19**(3) 263–275.

Reagans, Ray, Linda Argote, Daria Brooks. 2005. Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science* **51**(6) 869–881.

Sacco, William J, D Michael Navin, Katherine E Fiedler, II Waddell, K Robert, William B Long, Robert F Buckman. 2005. Precise formulation and evidence-based application of resource-constrained triage. *Academic Emergency Medicine* **12**(8) 759–770.

Selye, Hans. 1982. History and present status of the stress concept. handbook of stress: Theoretical and clinical aspects.

Shields, Grant S, Matthew A Sazma, Andrew P Yonelinas. 2016. The effects of acute stress on core executive functions: a meta-analysis and comparison with cortisol. *Neuroscience & Biobehavioral Reviews* **68** 651–668.

Shostack, G Lynn. 1987. Service positioning through structural change. *The Journal of Marketing* 34–43.

Smith, Ian C, MW Jones. 2001. Surgeons' attitudes to intraoperative death: Questionnaire survey. *BMJ* **322**(7291) 896–897.

Staats, Bradley R. 2012. Unpacking team familiarity: The effects of geographic location and hierarchical role. *Production and Operations Management* **21**(3) 619–635.

Staats, Bradley R, Hengchen Dai, David Hofmann, Katherine L Milkman. 2016. Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science* **63**(5) 1563–1585.

Ton, Zeynep, Robert S Huckman. 2008. Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science* **19**(1) 56–68.

van der Ploeg, Eleonore, Sasja M Dorresteijn, Rolf J Kleber. 2003. Critical incidents and chronic stressors at work: Their impact on forensic doctors. *Journal of Occupational Health Psychology* **8**(2) 157.

Weiss, Daniel S, Charles R Marmar, Thomas J Metzler, Heidi M Ronfeldt. 1995. Predicting symptomatic distress in emergency services personnel. *Journal of Consulting and Clinical Psychology* **63**(3) 361.

Wilson, FC, AD Poola, Karen Trew. 1997. Psychological distress in police officers following critical incidents. *The Irish Journal of Psychology* **18**(3) 321–340.

Appendix for "Recovering from Critical Incidents: Evidence from Paramedic Performance"

A1. Descriptive statistics

Table A1 Summary of data cleaning and inclusion/exclusion criteria.

	N	Delta N	% dropped
Data cleaning and exclusion:			
1. All dispatch records of paramedic teams	1,537,498		
2. Drop if going green time stamp is missing	1,534,122	3,376	0.2%
3. Drop if activation takes over 8 hours	1,529,295	4,827	0.3%
4. Drop if time-stamps are in wrong order	1,508,975	20,320	1.3%
5. Drop if Completion Time is higher than 99.5th percentile	1,501,424	7,551	0.5%
6. Drop if any sub-component duration is higher than 99.5th percentile	1,441,261	60,163	3.9%
7. Include only crews of two paramedics	1,417,743	23,518	1.5%
8. Drop if crew did not work together for the entire shift	1,385,692	32,051	2.1%
Exclusion for performance analysis:			
9. Drop if no hospital code or illness code is entered on patient report form	903,787	481,905	31.3%

Table A2 Summary statistics for (patient transporting) activations and (all) dispatches for shifts involving one or two CIs.

	Shifts with 1 CI		Shifts with 2 CIs	
	Mean	Std	Mean	Std
Number of dispatches	10.2	3.7	10.6	3.2
Number of dispatches after 1st CI	4.4	3.7	6.9	3.2
Number of dispatches after 2nd CI	N/A	N/A	2.9	3.0
Number of patient transporting activations	6.5	2.0	6.9	1.7
Number of patient transporting activations after 1st CI	2.9	2.3	4.6	1.9
Number of patient transporting activations after 2nd CI	N/A	N/A	2.0	1.9
N	7,935		233	

Table A3 Comparison of shifts with and without CIs

	Shifts without CIs		Shifts with CIs	
	Mean	Std.	Mean	Std.
Number of Dispatches	9.3	3.9	10.2	3.6
Crew Workload	0.80	0.17	0.81	0.14
Blue Calls (Non CI)	0.28	0.55	0.26	0.53
Average Crew Experience (Years)	8.3	6.4	8.6	6.1
Average Crew Age (Decades)	3.8	0.8	3.8	0.8

Figure A1 **The variation in crew experience for each age.**

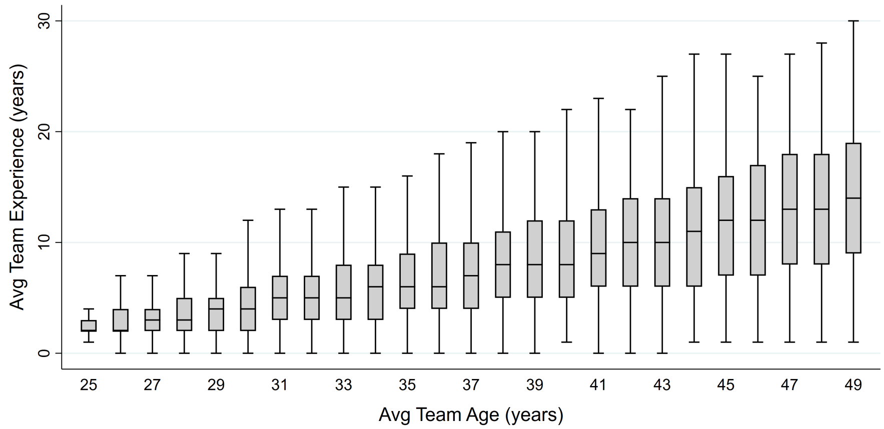
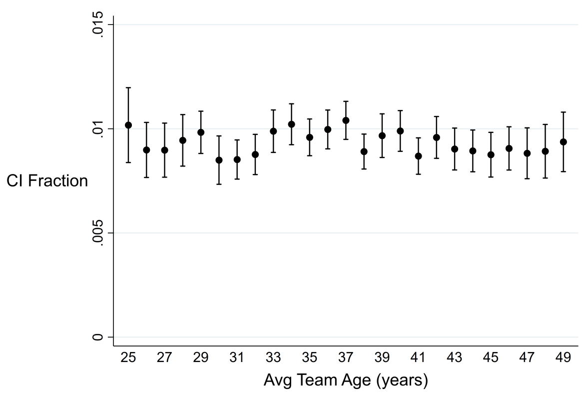
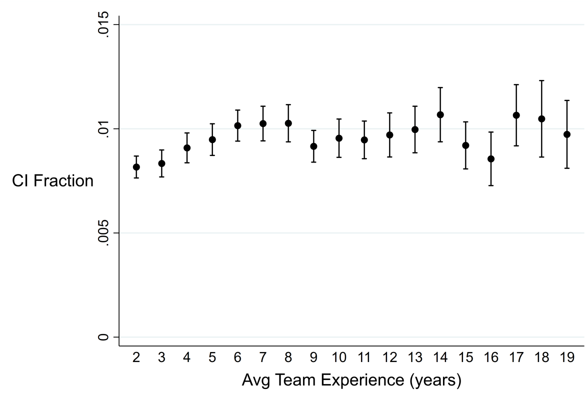


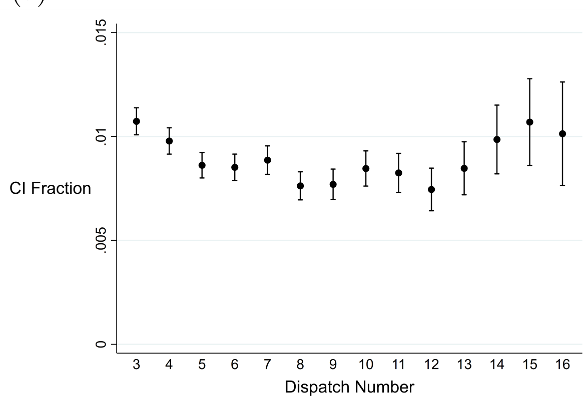
Figure A2 **Fraction of activations that are CIs by (a) average crew age, (b) average crew experience, and (c) cumulative number of dispatches in the day.**



(a)



(b)



(c)

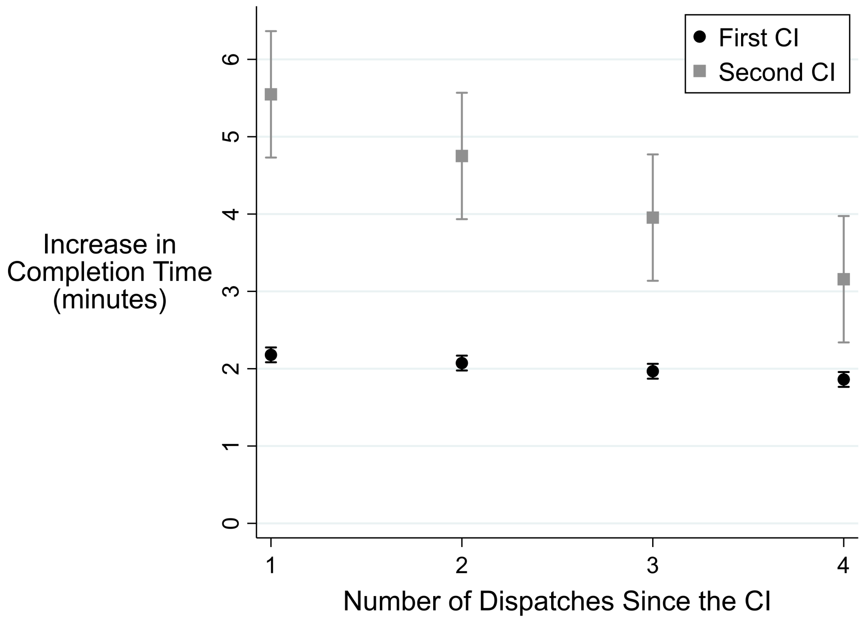
Table A4 Correlation table

	Completion Time	To Scene	At Scene	To Hospital	At Hospital	To Green	One Prior CI	Two Prior CIs	Utilization	Blue Call	Time Since Start	Dispatch Number	Avg Team Experience	Avg Team Age
Completion Time	1.000													
To Scene	0.328	1.000												
At Scene	0.724	0.004	1.000											
To Hospital	0.387	0.193	0.018	1.000										
At Hospital	0.446	0.060	0.142	0.047	1.000									
To Green	0.345	-0.055	0.061	-0.061	-0.196	1.000								
One Prior CI	-0.021	-0.020	-0.013	-0.002	-0.006	-0.010	1.000							
Two Prior CIs	-0.003	0.000	-0.003	0.003	-0.002	-0.002	0.137	1.000						
Utilization	-0.083	-0.137	-0.041	-0.079	0.011	0.002	0.051	0.008	1.000					
Blue Call	0.102	-0.047	0.077	-0.099	-0.102	0.333	0.001	-0.001	0.009	1.000				
Time Since Start	-0.151	-0.077	-0.061	-0.068	-0.051	-0.116	0.109	0.022	0.441	-0.001	1.000			
Dispatch Number	-0.199	-0.147	-0.095	-0.112	-0.078	-0.074	0.097	0.019	0.469	0.006	0.830	1.000		
Avg Team Experience	0.018	-0.036	-0.066	0.020	0.045	0.106	0.007	0.004	0.007	-0.019	-0.010	-0.002	1.000	
Avg Team Age	0.077	0.075	-0.009	0.046	0.058	0.058	-0.008	-0.000	-0.040	-0.028	-0.025	-0.072	0.668	1.000

All absolute values over 0.003 are significant at $p < 0.001$

A2. Additional results and robustness checks

Figure A3 The estimated impact of a prior CI for each dispatch following one and two CIs.



Note. The figure demonstrates the average magnitude of the performance impact (in minutes) of encountering CIs and the time-dependency of the effect, discussed in §6.2. All values correspond to the coefficient estimates in column (4) of Table 2 in the main text. Standard errors are those of the $Dispatches\_Since\_First\_CI$ and $Dispatches\_Since\_Second\_CI$. The black circles demonstrate the average effect of a prior CI as the crews conduct more activations (regardless of whether they encounter a second CI). The gray squares demonstrate the average effect of a second CI as the crew conducts more dispatches (regardless of when, during the shift, the first CI happened).

Table A5 Robustness checks

	Completion Time				Log(Completion Time)	
	(1)	(2)	(3)	(4)	(5)	(6)
One Prior CI (broader definition)	1.740***	1.713***				
	(0.181)	(0.181)				
Two Prior CIs (broader definition)		2.106***				
		(0.790)				
One Prior CI (blue call)			1.222***	1.150***		
			(0.108)	(0.108)		
Two Prior CIs (blue call)				1.847***		
				(0.244)		
One Prior CI					0.030***	0.030***
					(0.004)	(0.004)
Two Prior CIs						0.063***
						(0.023)
Shift Controls	✓	✓	✓	✓	✓	✓
Activation Controls	✓	✓	✓	✓	✓	✓
Seasonality Controls	✓	✓	✓	✓	✓	✓
Observations	902,002	902,002	902,002	902,002	902,002	902,002
Adjusted R^2	0.513	0.513	0.513	0.513	0.652	0.652

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6 Robustness checks

	Completion Time				
	Placebo Analysis (1)	Matching Analysis (2)	No Activation Controls (3)	Drop 1st Activation (4)	Dummy for Last Dispatch (5)
One Prior CI	-0.031	2.697***	3.829***	2.565***	1.911***
	(0.160)	(0.379)	(0.313)	(0.302)	(0.241)
Two Prior CIs	-0.025	5.338***	6.374***	4.090***	3.601***
	(0.721)	(1.905)	(1.800)	(1.514)	(1.391)
Shift Controls	✓	✓	✓	✓	✓
Activation Controls	✓	✓		✓	✓
Seasonality Controls	✓	✓	✓	✓	✓
Observations	902,002	50,894	902,078	748,626	902,002
Adjusted R^2	0.513	0.512	0.166	0.516	0.515

Notes: Activation controls include illness code FEs, whether it was a blue call, minutes since first dispatch (linear and quadratic terms), and workload since the first dispatch. Standard errors are robust and clustered at the shift level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

View publication stats