Learning from Many: Partner Exposure and Team Familiarity in Fluid Teams

Zeynep Akşin
Koç University, Istanbul, Turkey.
zaksin@ku.edu.tr

Sarang Deo
Indian School of Business, Hyderabad, India.
sarang_deo@isb.edu

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

Kamalini Ramdas
London Business School, London, UK.
kramdas@london.edu

In services where teams come together for short collaborations, managers are often advised to strive for high team familiarity so as to improve coordination and, consequently, performance. However, inducing high team familiarity, by keeping team membership intact, can limit workers' opportunities to acquire useful knowledge and alternative practices from exposure to a broader set of partners. We introduce an empirical measure for prior partner exposure and estimate its impact (along with that of team familiarity) on operational performance using data from the London Ambulance Service. Our analysis focuses on ambulance transports involving new paramedic recruits, where exogenous changes in team membership enable clean identification of the performance effect. Specifically, we investigate the impact of prior partner exposure on time spent during patient pick-up at the scene and patient handover at the hospital. We find that the effect varies with the process characteristics. For the patient pick-up process, which is *less standardized*, greater partner exposure directly improves performance. For the *more standardized* patient handover process, this beneficial effect is triggered beyond a threshold of sufficient individual experience. In addition, we find that the beneficial performance impact of prior partner exposure is amplified during periods of high workload, for both processes. Finally, a counterfactual analysis based on our estimates shows that a team formation strategy emphasizing partner exposure outperforms one that emphasizes team familiarity by about 9.2% in our empirical context.

Key words: Fluid teams, membership change, partner exposure, team familiarity
History: August 16, 2018

1. Introduction

In a wide variety of operational environments such as in-flight hospitality, police patrol, ambulance transport, and surgical procedures, teams are assembled for short collaborations and then

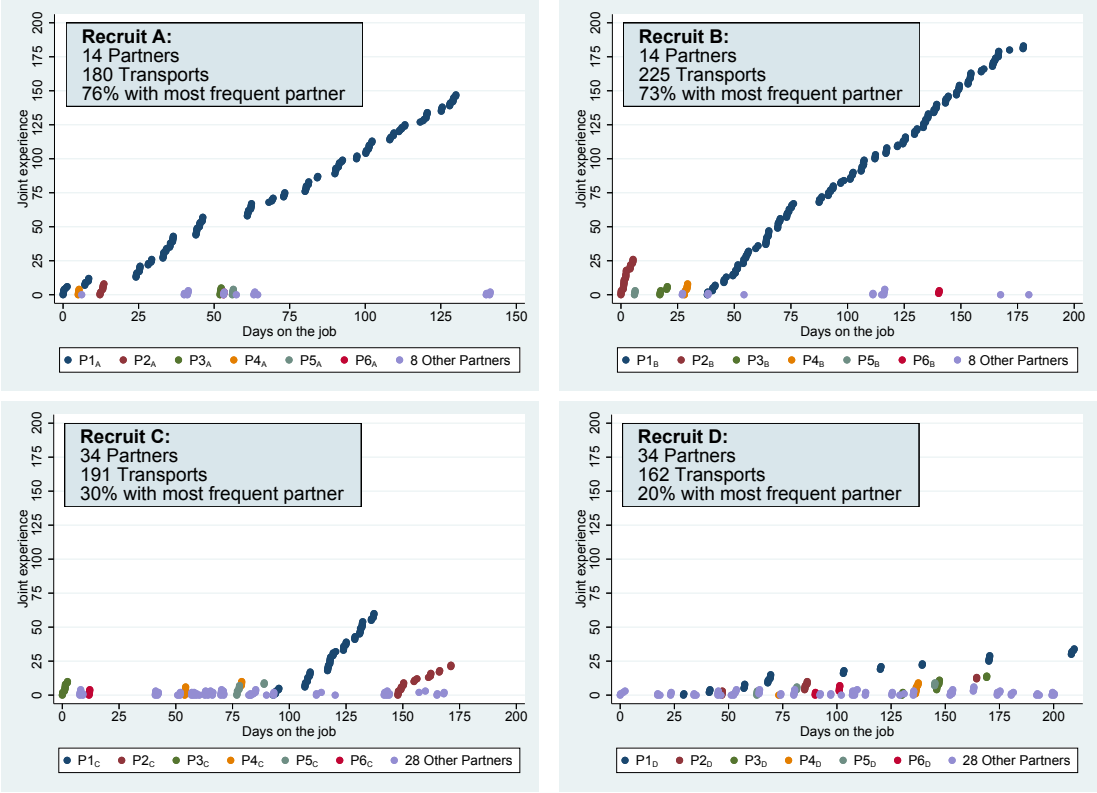
disbanded. In such fluid settings, more familiar teams, i.e., teams comprising individuals who have worked more often with each other, yield significantly better performance due to improved coordination and shared knowledge about the task at hand (Reagans et al. 2005, Huckman et al. 2009). As a result, operations managers are advised to keep teams intact to the extent possible (Huckman and Staats 2013). In establishing these findings the literature has thus far not considered the potential benefits of being exposed to the working practices of multiple partners over time. However, existing theoretical arguments and experimental findings suggest that changes in team membership can enhance creativity and problem solving (March 1991, Arrow and McGrath 1993, Choi and Thompson 2005, Gruenfeld et al. 2000, Lewis et al. 2007, Kane et al. 2005). Knowing whether being exposed to a diverse set of partners translates into improved operational performance in field settings and how these benefits compare with those of team familiarity can help managers decide on optimal team formation strategies (Huckman and Staats 2013).

We address this question using data on ambulance transports from the London Ambulance Service (LAS) involving new paramedic recruits. Generally, ambulance transports are conducted by teams of two paramedics. In contrast to senior paramedics, who usually have stable, long-term partners, new recruits are scheduled on a *relief roster* for the first two years of their service, which provides exogenous variation in their partner assignments[1]. As an example, Figure 1 depicts how the individual experience of 4 particular new recruits is distributed across partners in their first year. We observe that *Recruit A* encountered only 14 partners during the year and worked with one specific partner for 76% of her ambulance transports. In contrast *Recruit D* worked with a specific partner for at most 20% of her transports and encountered 34 distinct partners in total. The figure also highlights the distinction between team familiarity—the prior joint experience with only the partner at hand (on the Y-axis)—and the distribution of experience across all prior partners, introduced in this paper as an individual's *prior partner exposure*.

We expect the operational benefits of prior partner exposure to depend on the level of standardization of the process in which a worker is engaged. For instance, the patient pick-up process at the scene involves on-scene diagnosis, treatment and transfer to ambulance in a wide variety of physical, social, and emotional situations under limited guidelines. In a less standardized or *divergent* (Shostack 1987) process such as this, workers are likely to rely on tacit knowledge and have considerable executional latitude in applying it. Consequently, they are likely to benefit from prior partner exposure as it offers them the opportunity to observe different ways to perform process

[1] See more detail about scheduling in §3.1. In addition, we conduct empirical checks to verify these aspects of relief roster scheduling in Appendix A. Based on our analysis, we do not find any evidence that new recruits are more likely to be assigned a stable senior partner or are more likely to spend longer time with him/her as their experience grows. This is consistent with the LAS managers' statements that the scheduling of new recruits on the relief roster is driven by operational convenience rather than considerations of the new recruits' experience or performance.

Figure 1 **Partner experience of 4 new recruits during 2011.**



Note. The x-axis is the timeline from when the new recruit joins LAS until the end of the year. Each dot is a transport, its color corresponding to a particular partner (with the most infrequent partners displayed as one color). Its y-value shows the Team Familiarity (number of prior joint transports) with the partner. The most frequent partner of new recruit A during the year is denoted by $\mathrm{P1}_A$, the second most frequent by $\mathrm{P2}_A$, etc.

tasks and choose the best among them. In contrast, the patient handover process at the Accident & Emergency (A&E) department of a hospital involves communication of clinical information to a triage nurse using well-defined patient report forms and physical transport of the patient to an available bed across relatively uniform physical layouts. In a more standardized or routine process such as this, team members can rely on the standard operating procedures (SOPs). They are likely to overrule those only if they expect the useful knowledge they have gained by observing a diverse set of partners to substantially improve performance. As a result, the beneficial effects of prior partner exposure may start to manifest themselves only after individual team members have accumulated sufficient experience. In addition, the magnitude of performance benefits due to prior partner exposure is likely to depend on the paramedic workload. Based on prior work demonstrating that workers adjust service rates at times of high workload (Delasay et al. 2017), we expect increased access to useful knowledge through partner exposure to result in enhanced performance improvements during periods of higher workload.

We conduct our empirical analysis on a dataset comprising operational and clinical information on 5,773 ambulance transports staffed by one of the 81 new recruits (recruited in 2011) along with one or more of 702 senior paramedics, during the 2011. We measure prior partner exposure using a Herfindahl index of a new recruit's distribution of cumulative experience over prior partners. Since the variation in partner exposure across new recruits, demonstrated in Figure 1, is due to the administrative role of the relief roster and is exogenous to performance, our data allows for a clean identification of our main effects.

In line with the above hypotheses, we find that increased partner exposure of new recruits has a direct impact on scene time: increasing the partner exposure of a new recruit by two standard deviations results in his crew spending on average 7.3% less time (2.2 minutes relative to an average of 30.5 minutes) picking up patients at the scene. The effect of partner exposure on handover time is moderated by cumulative individual experience of the new recruit: greater partner exposure is beneficial only once the total number of transports by the new recruit is higher than a certain threshold. Beyond this level of sufficient experience (equivalent to approximately 4 months on the job), a two standard deviation increase in prior partner exposure reduces handover time by 6.7% (1.1 minutes relative to an average of 16.5 minutes). Additionally, we show that the beneficial impact of partner exposure is amplified at times of high individual workload, for both processes. Specifically, we observe that operational performance is generally improved during periods when paramedics are experiencing high workload—and the performance of new recruits with higher prior partner exposure improves significantly more than that of their counterparts. Our analysis also confirms the benefits of team familiarity that have been identified in other empirical contexts. Thus, a manager striving to optimize team formation and crew scheduling strategies must trade-off the beneficial effects of these two mechanisms. To obtain insight into this trade-off, we construct and analyze two illustrative counterfactual scenarios corresponding to two contrasting team formation strategies. The first, a *partner focus strategy*, assumes that new recruits are assigned a single stable partner for all transports whereas the second, a *partner exposure strategy*, assumes that new recruits take turns working with every senior paramedic at their location. Our results indicate that the partner exposure strategy outperforms the partner focused strategy for both patient pick-up and patient handover processes, resulting in about 9.2% improvement in aggregate. Apart from improved efficiency, transport duration reductions of this magnitude can also lead to significant improvement in health outcomes for time-sensitive conditions such as heart attacks, strokes and trauma injuries (Sacco et al. 2005, Bradley et al. 2006), which comprise approximately 20% of the ambulance transports in our dataset.

To summarize, our work contributes towards understanding the joint performance impact of team familiarity and partner exposure. We develop and introduce the notion of prior partner exposure

into the team productivity literature and operationalize it using an HHI measure. We jointly estimate the impact of prior partner exposure and team familiarity on operational performance, using data from a setting which allows for a clean identification of both effects. We find that the impact of prior partner exposure depends on the type of process at hand; observing a direct performance effect on a less standardized process (patient pick-up at the scene) but an experience-moderated effect on a more standardized process (patient handover at the hospital). In addition, we find that the beneficial performance impact of prior partner exposure is amplified during periods of high workload, for both processes. Finally, we use our model estimates to show that a scheduling strategy favoring partner exposure would outperform one favoring team familiarity in our setting.

The rest of the paper is organized as follows. We discuss the academic literature in §2, describe the operational details of our study setting in §3 and combine the two to construct testable hypotheses in §4. We describe our dataset and construction of empirical measures in §5, present and discuss the main empirical results in §6, elaborate on managerial implications in §7, and conduct extensive robustness checks in §8. We provide concluding remarks in §9.

2. Literature review

In this section, we first review the recent literature on team productivity and highlight how prior partner exposure is conceptually distinct from team familiarity as well as from task and customer diversity, all of which have previously been studied in that literature. Next, we differentiate our empirical focus on the impact of prior partner exposure on operational field performance from that of extant experimental studies on the impact of new team members on group creativity. We then review the evidence on how process characteristics and team composition interact in affecting operational performance and describe how our results contribute to this literature. Finally, we discuss existing results on the impact of workload on worker service rates.

2.1. Operational team productivity

Recent literature has identified team familiarity—measured by the average number of times team members have worked with each other in the past—as one of the key drivers of operational performance of fluid teams[2] (Reagans et al. 2005, Boh et al. 2007, Huckman et al. 2009, Huckman and Staats 2011). It is argued that greater experience of working with each other enables team members to learn "who knows what" thereby resulting in more effective division of labor and coordination of individual tasks and activities. Reagans et al. (2005) find that surgical teams whose members have greater experience of working with each other have shorter surgical completion times. Boh et al. (2007) and Huckman et al. (2009) find similar results in the context of software development.

[2] A team is considered *fluid* if team members temporarily work together to generate output, before the team is disassembled and its members move on to other tasks (Huckman and Staats 2011).

In contrast, in the context of stable teams, excessive team familiarity has been shown to be detrimental to team performance over the long run (Huckman and Staats 2013). Katz (1982) reports that R&D teams that remained together for more than 3.5 years on average showed a decline in performance due to reduced communication among team members and reduced interaction with external knowledge sources. Berman et al. (2002) find a similar relationship between shared team experience and team performance for professional basketball teams in the NBA. They attribute this effect to a phenomenon termed as knowledge ossification, where the body of tacit knowledge held jointly by the team starts depleting if the team is not infused with new ideas from external sources.

Our contribution to this literature is twofold. On the conceptual front, we develop the notion of prior partner exposure that is distinct from team familiarity; this is also reflected in their respective quantitative definitions. On the empirical front, we show that prior partner exposure has a beneficial effect on the performance of fluid teams in the short run, which is different from the adverse effect of excessive team familiarity on performance of stable teams over the long run. Thus, our results indicate the existence of two simultaneous effects in the short run: team familiarity improves coordination among team members on a given task whereas prior partner exposure broadens the knowledge base of the team members.

A related stream of literature has documented how diversity in tasks and customers served can help teams build a diverse knowledge base and thus improve performance over time (Boh et al. 2007, Narayanan et al. 2009, Staats and Gino 2012, Narayanan et al. 2013, Huckman and Staats 2011, Clark et al. 2013). Similar questions of scope and scale have been studied at the organizational level (Clark and Huckman 2012, Freeman et al. 2018). Our results show that the benefits of a diverse knowledge base can also be accrued through diversity of prior partners for individual workers. This effect is especially relevant if there is minimal task or customer diversity but processes required to complete the tasks are divergent (Shostack 1987), i.e., the same task for the same customer can be (and is) carried out in multiple different ways.

2.2. Experimental studies of team membership changes

Our concept of prior partner exposure and the proposed mechanism of its impact on team performance are partly motivated by prior literature on membership changes. March (1991) theorizes that employee turnover and interaction of new team members with existing members create diversity in the knowledge base, which facilitates *exploration* of new ideas. This has been validated by experimental studies that find improvement in team creativity following exogenous membership changes. Gruenfeld et al. (2000) make one member of an essay-writing group rotate through other groups and find that his return to the original group stimulates significantly more ideas among

members who did not rotate. Similarly, Choi and Thompson (2005) compare two types of groups—open (i.e., experiencing membership changes) and closed (i.e., no membership changes)—and find that open groups generate a larger number of ideas as well as more diverse ideas.

Our work differs from these experimental studies on three key dimensions. First, they follow teams that retain their identity throughout the study and are classified as open or closed depending on whether they undergo partial membership change. In contrast, our context comprises fluid teams, which are continually assembled and disbanded. We therefore introduce a continuous measure for their individual prior partner exposure. Second, the outcome of interest in those studies is team creativity, i.e., teams' ability to generate ideas. We focus on teams' operational performance on the ground, which requires the ability to generate ideas as well as implement them. Third, those studies do not consider the effect of membership changes on team familiarity, which, in turn, can also affect operational performance. We simultaneously estimate the effects of partner exposure and team familiarity, both of which vary exogenously because of the allocation of new recruits to a relief roster.

2.3. Process characteristics and operational performance

A growing body of literature documents how underlying characteristics of the process and associated knowledge moderate the effect of team characteristics on performance. Huckman and Staats (2011) show that greater familiarity enables teams to better cope with changes in task description. Furthermore, they find that higher intrapersonal diversity—in terms of customer experience—is beneficial to teams in coping with task changes whereas higher interpersonal team diversity is detrimental. Ton and Huckman (2008) find that a greater level of process compliance reduces the negative impact of changes in team membership (i.e., staff turnover). Arrow and McGrath (1993) find that student teams that experience membership change perform better on tasks requiring more active reflection by team members (e.g. essay writing). Similarly, Argote et al. (1995) show experimental evidence that the negative effects of membership change on team performance are less pronounced for complex tasks. On the flip side, Espinosa et al. (2007) find that beneficial effects of team familiarity decline for more complex tasks. Lapré and Van Wassenhove (2001) analyze productivity data from model lines in three manufacturing plants and suggest that teams with a more diverse knowledge base are more effective in understanding causal mechanisms, which are necessary for transferring non-codified knowledge. Hansen (1999) finds that knowledge transfer between organizational units who have infrequent contact (also called weak ties, see Granovetter (1973)) is effective for codified knowledge but less so for complex or tacit knowledge.

Our results contribute to this literature by suggesting that the effects of partner exposure may also be contingent on the characteristics of the underlying process and the type of knowledge

required to execute it. In our research setting we find that partner exposure has a direct effect on performance for a non-standardized process (requiring non-codified knowledge) while its effect is moderated by cumulative individual experience for a standardized process (requiring codified knowledge).

2.4. Workload and operational performance

Finally, a stream of research has demonstrated how workload can impact processing rates in various service systems (see Delasay et al. (2017), for an overview). This literature has identified ways in which workers can react at times of high workload, in the absence of formal process changes, in order to minimize disruptions—e.g., in the form of long waits—due to the high utilization. Batt and Terwiesch (2016) consider a multistage service operation and identify early-task-initiation (proactively initiating tasks to shorten delays) as a response mechanism to high workload in an emergency department. Kim et al. (2014) show that ICU admission decisions become more conservative. Freeman et al. (2016) find that midwives manage their workload through the rationing of services and by adjusting specialist referral rates. Berry-Jaeker and Tucker (2016) show that speeding up (e.g. through early discharge) can only relieve workload up to a saturation point, beyond which it no longer mitigates the workload. Kuntz et al. (2014) find evidence of health outcomes deteriorating beyond a tipping point in utilization.

We contribute to this literature by showing that prior partner exposure allows workers to adjust (speed up) to a greater extent than their counterparts, during times of high workload. This finding is consistent with the prior research which finds that workers have the ability to 'work harder' or 'speed up' during times of high utilization (Schultz et al. 1998, Kc and Terwiesch 2009, Staats and Gino 2012); we demonstrate that exposure to useful knowledge (e.g., different ways to complete a complex task) by working with multiple partners can be one way through which workers gain this ability.

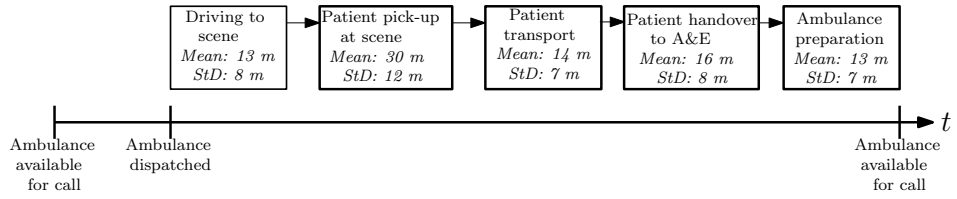
3. Empirical setting: London Ambulance Service

This section comprises two parts. In the first part, we provide a brief overview of the operating structure of LAS and its crew scheduling policies, based on extensive interviews with its management and staff (§3.1). In the second part, we provide a detailed description of the processes involved in ambulance transport based on field observations on ambulance runs and at the dispatch center (§3.2) as well as interviews with ambulance crews.

3.1. Operating structure and policies

The LAS operates over 100 ambulance stations spread across the city of London. A central control center at the LAS headquarters receives and triages all emergency calls and transfers them to one

Figure 2 Ambulance transport process diagram



Ambulance utilized		✓	✓	✓	✓	✓
Ambulance crew with patient			✓	✓	✓	
Time duration largely under crew control			✓		✓	✓

of the three dispatch desks (West, East, and South) depending on the geographical location of the incident. Each dispatch desk comprises six dispatch sectors, each of which has a designated dispatcher, responsible for assigning ambulances to emergency calls as the need arises.

Ambulance crew members with more than two years of experience are permanently assigned to one of the ambulance stations and typically have a stable partner and shift pattern. In contrast, during the first two years of tenure at LAS, new recruits are scheduled on a *relief roster*, which is used in one of the following three situations: (i) staffing of normal shifts when stable partners of senior paramedics go on leave, (ii) scheduling extra capacity for special events, and (iii) filling unpopular shifts. Thus, new recruits on the relief roster are frequently assigned different partners from different stations for a short period (e.g., a single shift). This assignment of new recruits on the relief roster to shifts and partners is made on the basis of administrative convenience rather than previous performance or prior partners. Appendix A provides a more quantitative analysis of this scheduling practice, showing that we find no evidence that new recruits are more likely to be assigned a stable senior partner or are more likely to spend longer time with him/her as their experience grows

On a daily basis, paramedics are scheduled to shifts of various lengths (specifically, 8, 9, 10, or 12 hours in duration) and with varying start times throughout the day and night (LAS 2017). This is the result of the LAS aiming to match the variable demand for ambulance transport by varying their capacity.

3.2. Ambulance transport processes

Figure 2 depicts the process flow for a typical ambulance transport from the moment an ambulance is dispatched for an emergency call until its crew becomes available again for the next dispatch after handing over the patient to the A&E. Shortening the duration of the processes in Figure 2 is an important operational objective for LAS because it reduces response times and increases ambulance availability thereby improving patient outcomes.

Table 1 **Illustration of the variety of scene scenarios encountered on a single day, during one of the field visits conducted by the authors**

Patient	Description	Clinical challenges	Non-clinical challenges
1: Male (90)	Lives with daughter; severe diarrhoea	Patient senile; had a serious diarrhoea episode; cause unclear.	Calming hysterical daughter; cleaning; and carrying patient down stairs.
2: Male (42)	Drug addict; haemorrhage	No clear diagnosis; medical history unclear.	Getting the patient (stuck in a motel room) to ambulance; and rationing 'gas and air'.
3: Male (50)	French tourist; dizzy, shortness of breath	No clear diagnosis; unclear need for tests; some simple tests conducted.	Communication; getting a clear description of the symptoms.
4: Female (80)	Senile; lit hair on fire	Understanding which medicine the woman had taken.	Explaining situation to the senile woman.
5: Female (50)	Speaks no English; rapid heartbeat	No clear diagnosis; multiple tests conducted, in apartment and ambulance.	Entire family observing; only one English speaker.

3.2.1. Dispatch. Once an emergency call has been triaged by the call center, the incident location, along with its severity score, appears on the computer terminal of the dispatcher responsible for that dispatch sector. The dispatcher also observes the location and status of the emergency vehicles (ambulances, cars, or bikes) closest to the scene and makes the dispatch decision based on this information. When called upon to be dispatched, an ambulance crew drives to the scene as quickly as possible.

3.2.2. Patient pick-up. A simplified description of the work performed by an ambulance crew at the scene is that it must locate and stabilize the patient, conduct a rudimentary diagnosis of his condition, and bring the patient into the ambulance for transport to an A&E. However, situations at the scene differ substantially along clinical and non-clinical dimensions making it impossible to fully specify the process or knowledge required to guarantee good operational (scene time) or clinical (patient outcome) performance (LAS 2015). Table 1 demonstrates the variety of possible challenges met by an ambulance crew during a single shift, on clinical as well as non-clinical dimensions.

On the clinical dimension, the paramedics must make diagnosis and treatment decisions regarding pain management and drug administration. They must immediately ensure that a patient's airway is clear and breathing is normal, and conduct a circulatory and disability assessment, while simultaneously taking into account the legal and ethical considerations of their actions (Fisher et al. 2006). On some occasions the clinical diagnosis is straightforward, given the symptoms (e.g., Patient 1 in Table 1 had diarrhoea whereas Patient 4 had lit her hair on fire). However, on other occasions, the clinical task is difficult if symptoms cannot be directly linked to a particular condition (Patients 3 and 5). In such cases, the ambulance crew members must rely on their experience and knowledge to decide whether to conduct a battery of tests at the scene (Patient 5) or to rush to the A&E after only a couple of simple tests (Patient 3).

On the non-clinical dimension, the physical and social surroundings at the scene can create challenges for paramedics. Carrying a 90 year old down a narrow staircase in a high-rise building or trying to move a wheelchair-bound drug addict through a narrow motel room door that does not accommodate a wheelchair (Patients 1 and 2, respectively) are two specific examples that capture the nature of these challenges. Interacting with family members of patients, bystanders, or patients themselves may also require varied skills and tact depending on their emotional state (Patient 1). Similarly, the non-clinical dimension can interact with the clinical dimension. For instance, the assessment of symptoms can be difficult because of language barriers (e.g., Patients 3 and 5).

In summary, the variety of medical and non-medical considerations that the paramedics must account for and deal with as part of their job make the patient pick-up process at the scene a divergent one (Shostack 1987) thereby precluding the creation of explicit SOPs[3].

3.2.3. Patient transport. Once the ambulance crew has administered the appropriate level of care at the scene and safely brought the patient into the ambulance, the patient is taken to an A&E as quickly as possible. The ambulance crew chooses the geographically closest A&E based on information available on a computer screen and thereafter follows the shortest driving route to it.

3.2.4. Patient handover. In the UK, barring extremely serious cases[4], ambulance patients arrive to the A&E through a special entrance but are not prioritized over other A&E patient arrivals as part of a deliberate policy to encourage self-transportation and reduce ambulance utilization. Upon their arrival at the A&E, the ambulance crew places the patient at the end of the queue in front of the triage nurse's station. Thereafter, the crew assigns the responsibility of the patient to the A&E staff in two steps. First, when the triage nurse becomes available, the crew communicates the main clinical information about the patient using a standardized patient report form. The triage nurse, aided by a standard checklist and results of diagnostic tests, if any, assigns the patient to one of the three A&E units: minor incidents, major incidents, or resuscitation. Second, upon assignment, the crew physically moves the patient to the corresponding unit and places the patient safely in an empty bed as suggested by the head nurse of that unit. Finally, the crew obtains the signature of the head nurse to confirm transfer of patient responsibility.

3.2.5. Ambulance preparation. After transferring the patient to the A&E, the crew cleans the ambulance, picks up necessary materials from the A&E store, and restocks the ambulance to get it ready for the next call. Occasionally, especially following a strenuous case, the crew members also use this time to recuperate and prepare themselves for the next dispatch.

[3] When asked about the variety of non-clinical issues, a crew member replied that the best characterization of their job at the scene was, in fact, 'problem solving'.

[4] For extremely serious cases, the crew gives advance notice to the resuscitation unit on its way to the A&E using a "blue call" and bypasses the triage nurse upon arrival. We control for such cases in our analysis.

For the remainder of the paper, we focus on the patient pick-up process at the scene and the patient handover process at the A&E because these two processes affect ambulance utilization, involve the patient, and are substantially under the control of the crew. They also constitute a large portion of the prehospital delay (see Figure 2 and histograms in Appendix B for the duration distribution for each step of a patient transport).

4. Partner exposure and operational performance

In this section, we discuss how partner exposure might affect the ambulance crew's performance during the patient pick-up process at the scene (§4.1) and the patient handover at the A&E (§4.2). We also discuss how the crew workload might amplify these effects (§4.3). To this end, we build on the theoretical arguments from the prior literature (§2) and operational details of the empirical setting from our field observations and interviews (§3).

4.1. Patient pick-up at scene

As explained in §3.2.2, the patient pick-up process at the scene comprises both clinical and non-clinical components. Although some guidelines are provided to the crew for the clinical component, there is substantial leeway in interpreting and executing them during routine operations (Fisher et al. 2006). Furthermore, no SOPs exist for the non-clinical component demonstrated in Table 1. Therefore, good operational performance often hinges on the expertise, judgment, and tacit knowledge of the ambulance crew, which in turn can be affected by partner exposure in several ways, as explained below.

First, increased partner exposure can enable new recruits to identify and locate existing, useful knowledge from their partners. Given the divergent nature of the patient pick-up process at the scene, senior crew members are likely to have developed their own methods and techniques, especially for the non-clinical component of the job for which SOPs do not exist. Consequently, working with a diverse set of partners is likely to create more opportunities for the new recruits to observe and internalize a variety of methods and techniques (also called non-redundant useful knowledge; Granovetter 1973) of executing a task.

Second, partner exposure is likely to increase the amount and the novelty of new knowledge generated. Since new recruits are continually exposed to short collaborations with different partners, crews that have a new recruit on board must implicitly or explicitly agree on certain operational practices as a team. This induces active reflection and rethinking of the practices that team members have learned or used before, thus helping them to generate new knowledge and techniques (Arrow and McGrath 1993, Nonaka 1994, Delaney et al. 1998). For instance, before carrying an old lady strapped to an ambulance chair down a narrow staircase the crew must discuss which method best ensures her safety and decide on the appropriate method, e.g., for both paramedics to lean

against the staircase wall for stabilization while carrying the patient down the stairs. Furthermore, generated ideas are likely to be more creative and diverse if team members have experience of working in other teams (Gruenfeld et al. 2000, Choi and Thompson 2005). A new recruit who has observed a variety of methods can contribute more ideas, which can be creatively combined with those of his partner to further improve their effectiveness.

Third, because the process is divergent, i.e., it does not have rigidly defined SOPs, crew members possess executional latitude in applying the new knowledge and techniques to perform the current task at hand (Shostack 1987). Hence, the ideas contributed by the new recruits are also likely to be actually implemented thereby resulting in improved operational performance.

Based on the above arguments, the effect of partner exposure on the crew's performance in the patient pick-up process at the scene can be formalized in the following hypothesis:

H1*: Greater prior partner exposure reduces average scene pick-up time.*

4.2. Patient handover at hospital

The patient handover process at the A&E is much more structured, with well defined SOPs, compared to the patient pick-up process at the scene (see §3.2.4). Furthermore, the environment at A&Es across London is much more uniform and controlled than that at the scene. Consider, for example, a patient in a car crash. At the scene, even approaching the patient might be difficult; he might be unsafe to move; and the environment might induce clinical and non-clinical complications. In contrast, at the A&E the patient is already on a gurney, access is straightforward, and the environment is specifically designed to deal with complications. As a result of these differences in process characteristics, we expect the impact of prior partner exposure on performance during patient handover to be different from that during patient pick-up.

First, due to the standardized nature of the patient handover process, workers are likely to employ similar techniques. Hence, despite prior exposure to multiple partners, a new recruit has fewer opportunities to observe and acquire non-redundant knowledge during patient handover than during patient pick-up. Second, even if a new recruit has acquired useful knowledge for the task at hand, it is unlikely to be utilized because crew members are more likely to revert to the rigidly defined SOPs when in doubt about the next steps (Amabile 1997) rather than actively reflect upon the task (Arrow and McGrath 1993). Third, SOPs are typically designed to guarantee good performance on average. Hence, crew members are unlikely to overrule the SOP unless they are quite sure that doing so is likely to positively impact performance. Developing such insights is possible only if new recruits have sufficient experience in both the clinical and the operational components of the patient handover process. For instance, consider the use of the patient report form during the patient handover process. A crew member minimizes the probability of accidentally

leaving out critical clinical information by explicitly communicating all the information on the form to the triage nurses. However, a paramedic with substantial experience can shorten the process considerably by overriding the SOP and highlighting only the most important clinical information needed to decide the next course of action.

Summarizing the above discussion regarding the impact of partner exposure on performance in the patient handover process, we obtain the following hypothesis:

H2*: Greater prior partner exposure reduces average handover times, for new recruits with sufficiently high individual experience.*

4.3. The impact of workload

As we highlight in §2, the existing literature has established how service rates are often endogenous to system utilization since workers can adjust their practices in reaction to their environment. Different mechanisms have been suggested for how service rates are adjusted, ranging from speeding up (Delasay et al. 2016), to altering operating procedures (Freeman et al. 2016, Batt and Terwiesch 2016, Berry-Jaeker and Tucker 2016, Tan and Netessine 2014). Usually, these changes in service rates are not the result of a protocol but represent an effort on the worker's part to adjust to high workloads.

In order to make this adjustment, workers must have the required skill or knowledge to adjust how they finish the task at hand. One way to acquire that skill set is by getting the opportunity to observe the methods and problem-solving techniques of multiple partners. Hence, based on our earlier reasoning, we expect workers with high prior partner exposure to be better equipped to make adjustments to their practice, in reaction to high workload.

As with many service systems the paramedics in the LAS system experience workload through how quickly they are dispatched to a new incident after their ambulance goes green following a previous dispatch, i.e. through their individual workload. We therefore expect paramedics to adjust their service rates (to the extent that they can) at times of high individual workload, as measured by current shift utilization (i.e. the proportion of time, during the current shift, that they have been on call).

In summary, we expect that the new recruits who have been exposed to useful knowledge and different methods through multiple partners are more likely to be able to adjust their operating procedures to a high workload situation, to get the job done quicker. We have argued how the acquisition and application of such useful knowledge depends on the process type. Combining these arguments, we arrive at the two hypotheses below, complementing **H1** and **H2**, respectively.

H3*: Greater prior partner exposure leads to greater reduction of average scene time during periods of high workload.*

H4*: Greater prior partner exposure leads to greater reduction of average handover times during periods of high workload, for new recruits with sufficiently high individual experience.*

5. Data and variables

Our main dataset comprises operational and clinical information on 5,773 ambulance transports staffed by one of the 81 new recruits (recruited in 2011) along with one or more of 702 senior paramedics, during the entire calendar year 2011[5]. The operational information includes arrival and departure times at the scene and A&E that are recorded electronically by the ambulance crew. The clinical information—including the *Advanced Medical Priority Dispatch System* (AMPDS) primary condition classification, used for call triaging by the emergency call handler—is collected manually through A&E forms and later collated and digitized by LAS. We supplement this operational dataset with a de-identified crew dataset that contains information on the job tenure of the senior crew members. Table 2 provides summary statistics and correlation values for all main variables.

5.1. Outcome variables

We measure the performance of the ambulance crew for the patient pick-up and patient handover processes by measuring the time spent at the scene and at the A&E, respectively. $Scene_Time_{rt}$ for the t^{th} ambulance transport of new recruit r is defined as the time interval between the arrival of the ambulance at the scene and its departure towards an A&E. Similarly, $Handover_Time_{rt}$ for the t^{th} transport of new recruit r is defined as the time interval between arrival of the ambulance at the A&E and completion of both the clinical and physical handover of the patient to the ambulance staff .

5.2. Main predictor variables

Prior partner exposure. To quantify the extent to which a new recruit has been exposed to useful knowledge through interacting with multiple partners, we calculate a new recruit's dispersion of experience across all prior partners. We measure this using the familiar Hirschman-Herfindahl index (HHI), which for our purposes is defined as follows:

$$Partner_HHI_{rt} = \sum_{p \in \mathcal{P}_{rt}} \left(\frac{Joint_Exp_t^{rp}}{Exp_{rt}} \right)^2, \tag{1}$$

where $Joint_Exp_t^{rp}$ denotes the cumulative number of transports undertaken by the new recruit r with partner p prior to his t^{th} ambulance transport, the set $\mathcal{P}_{rt}$ denotes all partners that the new recruit r has worked with prior to his t^{th} transport, and Exp_{rt} denotes the total number of ambulance transports undertaken by the new recruit prior to his t^{th} transport.

[5] Before conducting our analysis we refined and cleaned the data to focus on standard ambulance transports that conveyed patients to an A&E. We excluded patient transports on other vehicles such as motorcycles, helicopters, etc. and to other destinations such as long-term care facilities and specialty departments. We excluded observations where key variables were missing or had extreme outlier values. Finally, we also excluded transports involving one of the six paramedics with substantial prior experience, who were laterally recruited by LAS during 2011.

Table 2 Summary statistics and correlation table

	Summary statistics		Correlation														
Dependent variables	Mean	StD	*Scene_Time*	*Handover_Time*	*Partner_HHI*	*Team_Familiarity*	*Exp*	*Task_Exp*	*A&E_Exp*	*Crew_Workload*	*Partner_Tenure*	*Number_Of_Partners*	*Blue_Call*	*Recent_Base_Arrivals*	*Average_Base_Arrivals*	*Recent_A&E_Arrivals*	*Average_A&E_Arrivals*
Scene_Time	30.5	12.7	1.00														
Handover_Time	16.5	7.9	0.15	1.00													
Prior Partner Exposure & Team Familiarity																	
Partner_HHI	0.30	0.21	0.02	-0.03	1.00												
Team_Familiarity	26.9	42.4	-0.01	-0.08	0.10	1.00											
Controls																	
Exp	108.4	88.3	-0.01	-0.03	-0.35	0.36	1.00										
Task_Exp	19.3	27.4	0.09	0.06	-0.19	0.20	0.60	1.00									
A&E_Exp	21.1	27.4	-0.04	-0.01	-0.26	0.16	0.58	0.34	1.00								
Crew_Workload	0.60	0.34	-0.03	0.01	0.06	0.07	-0.08	-0.03	-0.09	1.00							
Partner_Tenure	2.79	3.15	-0.09	0.00	-0.02	-0.05	-0.09	-0.09	0.07	0.05	1.00						
Number_Of_Partners	1.2	0.4	-0.03	0.00	0.11	0.06	-0.44	-0.29	-0.25	0.14	0.20	1.00					
Blue_Call	0.03	0.16	0.04	-0.10	0.01	-0.01	-0.05	-0.04	-0.02	0.03	0.03	0.07	1.00				
Recent_Base_Dispatches	11.4	6.5	0.02	0.04	0.09	0.01	-0.05	0.01	-0.05	0.15	-0.05	-0.13	-0.02	1.00			
Average_Base_Dispatches	81.2	40.9	0.02	0.01	0.09	0.09	0.05	0.06	0.04	0.04	-0.03	-0.14	-0.02	0.58	1.00		
Recent_A&E_Arrivals	7.4	3.8	-0.01	0.16	-0.09	-0.09	-0.03	-0.01	0.17	0.08	0.06	-0.04	0.01	0.16	0.08	1.00	
Average_A&E_Arrivals	7.0	2.3	-0.02	0.14	-0.15	-0.15	0.00	0.01	0.29	-0.02	0.09	-0.11	0.02	0.13	0.09	0.66	1.00

All correlation entries of absolute values over 0.03 are statistically significant at the 5% level.

For a new recruit with minimal partner exposure, i.e., if she has worked with a single partner for all prior transports, $Partner_HHI_{rt}$ takes the value of 1, whereas for a new recruit with maximal partner exposure, i.e., if her prior experience is distributed equally among all her partners, $Partner_HHI_{rt}$ takes the value of $\frac{1}{|\mathcal{P}_{rt}|}$. More generally, a smaller value of $Partner_HHI_{rt}$ denotes greater prior partner exposure.

An HHI measure is ideal for our purposes as it reflects both the number of prior partners and the depth of each new recruit's exposure to these partners (Narayanan et al. 2009). For instance, a new recruit who has worked with two senior partners on 50 ambulance transports each ($Partner_HHI$ = 0.5) is likely to have broader base of knowledge as compared to one who has worked on 1 and 99 transports with the two partners, respectively ($Partner_HHI$= 0.98). Additionally, the marginal impact on $Partner_HHI$ is diminishing in the number of ambulance transports conducted with the same partner, reflecting the fact that a new recruit is less likely to be exposed to new knowledge the more often they have worked with the same partner. Similarly, the marginal impact on the value of $Partner_HHI$ is diminishing in the number of new partners, since a new recruit is more likely to acquire new and useful knowledge from the first few senior paramedics they interact with than from subsequent ones.

To examine how individual experience moderates the impact of partner exposure on the handover time, we define an indicator variable for sufficient individual experience as follows:

$$Suff_Exp_{rt} = \begin{cases} 1 & if \quad t > Exp_Threshold \\ 0 & otherwise. \end{cases} \tag{2}$$

and interact it with $Partner_HHI_{rt}$. For our main analysis we define $Suff_Exp_{rt}$ using an experience threshold ($Exp_Threshold$) corresponding to the median number of ambulance transports undertaken by new recruits during their first four months at LAS. In Appendix C, we check the robustness of our results by varying $Exp_Threshold$ around the median.

5.3. Control variables

Next, we describe four groups of control variables that are included in our empirical analysis, to control for various effects which have previously been found to impact team performance.

5.3.1. Cumulative experience. We include the following variables to control for various dimensions of the cumulative experience of the new recruit. Those have been shown to impact operational performance in the prior literature.

Individual experience. As defined above, Exp_{rt} denotes the cumulative individual experience of new recruit r prior to ambulance transport t. This variable controls for any individual learning effect of the new recruit (Reagans et al. 2005, Ramdas et al. 2018, Staats and Gino 2012).

Team familiarity. Consistent with the team productivity literature (Reagans et al. 2005, Huckman et al. 2009, Staats 2012), we measure team familiarity using the average number of prior ambulance transports undertaken jointly over all pairs of crew members. More formally, we define $Team_Familiarity_{rt}$ for the t^{th} ambulance transport of new recruit r as follows:

$$Team_Familiarity_{rt} = \frac{\sum_{p \in \mathcal{T}_{rt}} \sum_{q \in \mathcal{T}_{rt}} Joint_Exp_t^{pq}}{\frac{N_{rt}(N_{rt}-1)}{2}},$$

where $\mathcal{T}_{rt}$ denotes the set of individuals on the crew including the focal new recruit, $N_{rt} = |\mathcal{T}_{rt}|$ denotes the the size of the crew, and $Joint_Exp_t^{pq}$, defined as before, captures the prior joint experience of any pair $p, q \in \mathcal{T}_{rt}$, $p \neq q$. For a typical ambulance crew with one new recruit and one senior partner, this measure is simply the number of prior ambulance transports the two have jointly undertaken.

Task experience. Let C_{rt} denote the primary clinical condition of the patient in the t^{th} ambulance transport of new recruit r. The condition C_{rt} is defined as an element of the set $\mathcal{C}$ of AMPDS codes for different clinical conditions. $Task_Exp_{rt}$ denotes new recruits r's cumulative experience of transporting patients with primary condition C_{rt} to the A&E, prior to his t^{th} transport. This variable controls for task specific knowledge gained on the job, which has been shown to positively impact performance (Clark et al. 2013, Staats and Gino 2012).

A&E experience. Prior research has demonstrated that the performance benefits of cumulative experience are not easily transferable across organizations (Clark et al. 2013, Huckman and Pisano 2006). Hence, while analyzing the operational performance of the patient handover process, we also control for the cumulative experience of the new recruit r at the focal A&E prior to the current transport t, denoted by $A\&E_Exp_{rt}$.

5.3.2. Ambulance transport controls. We control for three characteristics of the ambulance transport and partners of the new recruit which can impact performance.

Team tenure. Greater individual experience of the partners of the new recruit is also likely to have beneficial impact on the operational performance of the crew. However, many partners have a tenure at LAS spanning several years and ambulance transport data is not available to us over that period. Therefore, controlling explicitly for their ambulance transport experience is infeasible given our data. Hence, we use the average tenure (in years) of the partners of new recruit r on ambulance transport t as a proxy for their cumulative experience:

$$Partner_Tenure_{rt} = \sum_{n=1}^{N_{rt}} \frac{Tenure_{rt}^{n}}{N_{rt}}.$$

Team size. $Number_Of_Partners_{rt}$ denotes the total number of senior crew members on the crew along with the new recruit r for his t^{th} ambulance transport. The typical crew size is two paramedics (one new recruit and one senior partner in our data), but occasionally one or two more seniors are present.

Blue call. For extremely critical patients, the ambulance crew can alert the A&E staff in advance to expedite the patient handover process on arrival (see §3.2.4). We use an indicator variable $Blue_Call$ to control for such cases.

5.3.3. A&E workload. The handover times at the A&E can be affected by the workload of the A&E staff at the time of arrival of the focal ambulance transport. Although we do not have access to the actual workload data for the 37 A&Es appearing in our data, we include two variables, defined below, as proxies for the predictable and unpredictable workload, respectively, of the receiving A&E.

Average workload. $Average_A\&E_Arrivals_{rt}$ denotes the average number of ambulance arrivals at the receiving hospital's A&E during the time of day (in 2 hour periods) of the arrival of the ambulance, for the year of 2011.

Recent workload. $Recent_A\&E_Arrivals_{rt}$ denotes the number of arrivals at the receiving hospital's A&E during the last two hours and eight minutes, which corresponds to the median length of stay of A&E patients (Health & Social Care Information Centre 2014).

5.3.4. Base workload. Similarly, we control for base workload by calculating the average and recent ambulance dispatches from the home base of the paramedic crew at the time of the transport at hand.

Average workload. $Average_Base_Dispatches_{rt}$ denotes the average number of ambulance dispatches from the home base of the paramedic crew during the time of day (in 2 hour periods) of the arrival of the ambulance, for the year of 2011.

Recent workload. $Recent_Base_Dispatches_{rt}$ denotes the actual number of dispatches from the home base of the paramedic crew during the last two hours.

5.3.5. Individual workload. Furthermore, we control for the shift utilization of the paramedic crew. The LAS data does not include explicit indicators for when individual shifts start and end. In addition, according to LAS managers the shift durations can vary between 8 hours and 12 hours. We extract the shift structure by assigning a new shift following very long idle times[6], and calculate the shift utilization as follows.

Crew workload. The $Crew_Workload_{rt}$ variable measures the utilization of the paramedic crew on the current shift, prior to the current transport. To calculate the utilization, we consider the time from each dispatch until the crew 'goes green' (see Figure 2) to be active time and the period from going green until they receive the next dispatch to be idle time.

High workload. To test our hypotheses regarding the moderating effect of workload on the positive impact of partner exposure, we define the binary variable $High_Workload_{rt}$ to take the value of 1 when $Crew_Workload_{rt}$ exceeds 0.88 (75th percentile of all transports) and 0 otherwise. We conduct extensive sensitivity analysis for this threshold and report those results in Appendix D.

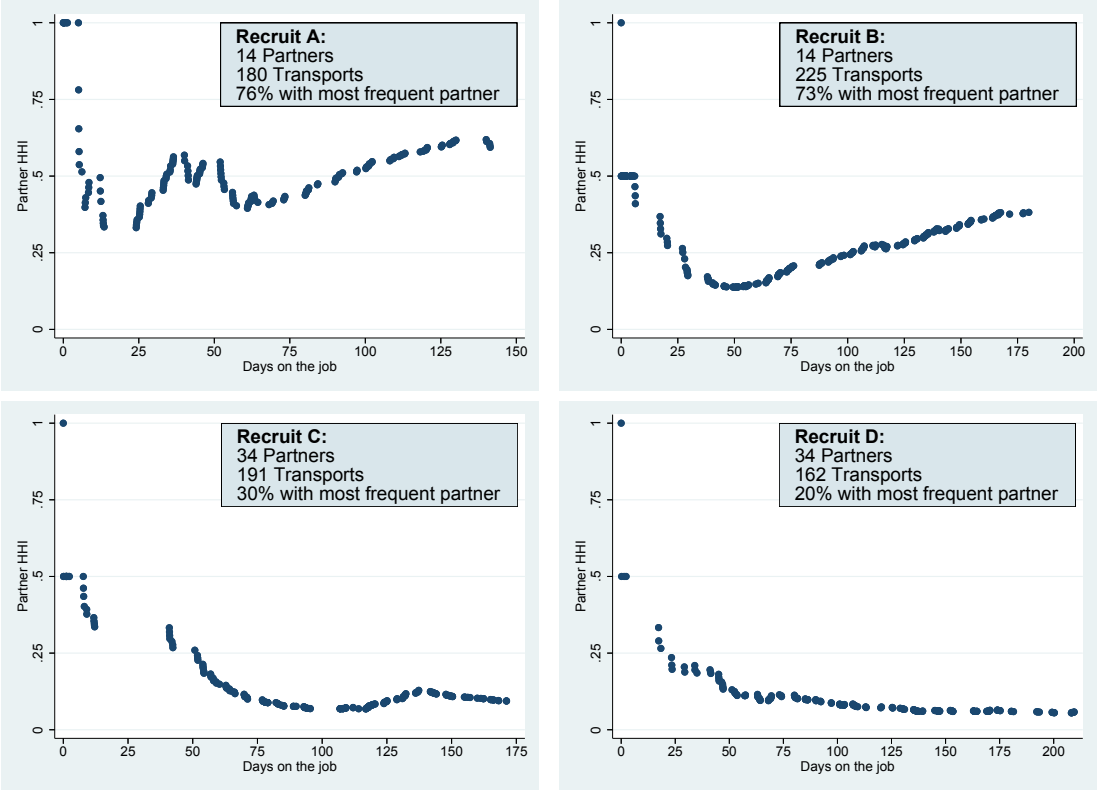
5.3.6. Fixed effects. Finally, we include a large set of indicator variables to account for unobserved heterogeneity across individual recruits, hospital A&Es, tasks, and time periods.

Inherent individual capability. New recruits may differ in their innate ability and aptitude to handle ambulance transports. There can be differences in idiosyncratic time invariant traits (such as their intelligence and dexterity) as well as in the average effect of time-varying factors (such as their motivation to perform well). These can create systematic differences in the performance of the paramedics even if all other factors are controlled for. Hence, we include a fixed effect for each new recruit.

A&E factors. Similarly, some hospital A&Es might systematically be more or less efficient in completing the tasks associated with the patient handover process, resulting in lower or higher

[6] We assume that if there is five hours or more of idle time between consecutive transports, a new shift was started. This results in a very reasonable shift structure in terms of number of shifts per month, the length of an average shift, and vehicle usage patterns (paramedics almost never change vehicles within shifts but frequently start a new shift in a new vehicle).

Figure 3 **Partner exposure ($Partner_HHI$) of the 4 new recruits from Figure 1.**



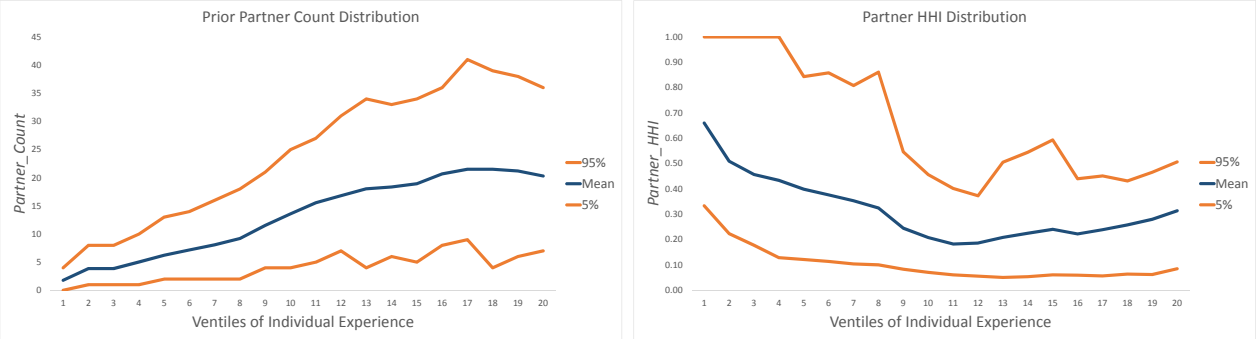
handover times even if other factors were kept constant. To control for this variation, we include a fixed effect for each A&E. These capture time-invariant differences across A&Es along dimensions such as staffing levels, information systems and patient mix.

Task type. The scene and handover times can differ systematically based on the clinical condition of the patient. To control for this source of variation, we include a fixed effect for the AMPDS code assigned to each patient based on their clinical condition.

Task severity. In addition to the clinical codes described above, the LAS also assigns a severity score for the clinical status of the patient in each ambulance transport. The data includes eleven distinct ordinal scores. We include fixed effects for each of these.

Seasonality and weather. The working conditions of the ambulance crew, and consequently the scene and handover times, can be affected by external factors. We include fixed effects for the weather conditions during the day of the transport (average wind, rainfall, and temperature). Furthermore, we include indicator variables for the time of day (morning, afternoon, evening, night), the day of the week, and the month of the year to control for other temporal variations (e.g. outdoor lighting conditions and traffic).

Figure 4 **The distribution of prior partner count and partner exposure for each ventile of individual experience.**



6. Econometric models and main results

We discuss our identification strategy in §6.1, our empirical models and results for the impact of partner exposure on performance in the patient pick-up process in §6.2, and those for the patient handover process in §6.3.

6.1. Identification strategy

To identify the impact of prior partner exposure and team familiarity on team performance we rely on the fact that the assignment of partners to new recruits is exogenous to performance and that this generates adequate variation in both prior partner exposure and team familiarity for new recruits at all levels of individual experience.

We have confirmed that the manager overseeing the relief roster does not have access to prior performance data. Figure 1 illustrates how 4 new recruits have very different experience profiles in terms of how many partners they work with during their first year and how often. This translates directly into a variety of sample paths for our partner exposure variable, as illustrated in Figure 3.

Figure 4 provides a more systematic view of the variability used to identify the effect of partner exposure, demonstrating the variation in partner count and partner exposure for our entire sample of new recruits as they gain individual experience in the first year. We observe that for any level of individual experience our data includes new recruits who have had considerable prior partner exposure ($Partner_HHI$ close to 0) and new recruits who have had less prior partner exposure. Appendix A further demonstrates that we find no evidence in the data that new recruits are more likely to be assigned a stable senior partner or are more likely to spend longer time with him/her as their experience grows during their first year on the job.

6.2. Partner exposure and scene time (H1 & H3)

We use the following panel specification to identify the impact of $Partner_HHI$ on $Scene_Time$:

$$\begin{aligned} Scene_Time_{rt} = \alpha_0 + \alpha_1\ Partner_HHI_{rt} + \alpha_2\ Partner_HHI_{rt} * High_Workload_{rt} \\ + \alpha_3\ Team_Familiarity_{rt} + \alpha_4\ Exp_{rt} + \alpha_5\ Exp_{rt}^2 + \alpha_6\ Task_Exp_{rt} \end{aligned} \tag{3}$$

$$+\alpha_7\ Crew_Workload_{rt} + \alpha_8\ Partner_Tenure_{rt} + \alpha_9\ Number_of_Partners_{rt}$$
$$+\alpha_{10}\ Blue_Call_{rt} + \alpha_{11}\ Recent_Base_Dispatches_{rt} + \alpha_{12}\ Average_Base_Dispatches_{rt}$$
$$+\alpha_{\mathbf{13}}\ Controls_{rt} + \epsilon^S_{rt}.$$

The main explanatory variable of interest is our measure for partner exposure, $Partner_HHI_{rt}$. A positive and significant value of α_1 would indicate that a greater partner exposure of new recruits reduces the scene time, supporting **H1**. Similarly, a positive and significant value of α_2 would indicate that the effect is stronger at times of high workload, providing support for **H3**.

In the model above we include all the controls listed in §5.3 that might affect performance at the scene. Specifically, we control for the experience of the new recruit along dimensions other than prior partners, including their total prior individual experience and total prior task experience. We include the squared term of individual experience to allow for diminishing returns (Reagans et al. 2005) as the new recruit gains more experience (including a squared term for task experience results in an insignificant coefficient and does not change any of the results so we omit this variable). Additionally we control for team and system measures of workload. At the team level, we include $Crew_Workload$ to capture the prior utilization of the team during the current shift. At a system level, we control for the workload at the home base of the team by including $Average_Base_Dispatches$, the average number of dispatches for the time of day (predictable variability), as well as $Recent_Base_Dispatches$, the number of dispatches during the previous 2 hours (unpredictable variability). Finally, $Controls$ includes weather at the time of the transport (wind, rain, temperature) and fixed effects for the new recruit, task, severity, and seasonality (time of day, day of week, and month).

Table 3 reports the coefficient estimates and heteroscedasticity-consistent standard errors of model (3). We cluster the standard errors by new recruit, to account for autocorrelation in unobservable factors across ambulance transports for the same new recruit. We observe a strong association of both team familiarity and partner exposure with how long the crew spends at the scene. The positive coefficient of $Partner_HHI$ in our base model (column (A)) indicates that teams including new recruits with a prior experience focused on few partners spend longer at the scene than their counterparts with more partner exposure, which is consistent with **H1**. A strict interpretation of the magnitude of the coefficient implies that crews involving new recruits with a fully focused partner experience ($Partner_HHI = 1$) spend on average 5.3 minutes (p =0.009) longer at the scene as compared to their counterparts involving new recruits with a fully diverse partner experience ($Partner_HHI = 0$). Alternatively, increasing the partner exposure of new recruits by one standard deviation would decrease scene times by over one minute[7].

[7] It is worth noting that including a $Partner_HHI * Suff_Exp$ interaction (which we use in our analysis of handover times) does not impact our results and the coefficient of that interaction is insignificant, see Appendix E.

Table 3 OLS coefficient estimates for pick-up time at scene regressions

	Dependent variable: Scene time	
	(A)	(B)
Partner exposure & Team familiarity	**Base model**	**Workload interaction**
$Partner_HHI$	5.348**	4.475*
	(2.048)	(2.069)
$Partner_HHI$ * $High_Workload$		3.613**
		(1.338)
$Team_Familiarity$	−0.014*	−0.014*
	(0.006)	(0.006)
Experience controls		
Exp	−0.018	−0.016
	(0.017)	(0.017)
Exp^2	0.000	0.000
	(0.000)	(0.000)
$Task_Exp$	−0.019	−0.022
	(0.030)	(0.030)
Other time-varying controls		
$Crew_Workload$	-1.054^{+}	−1.895**
	(0.543)	(0.595)
$Partner_Tenure$	−0.015	−0.015
	(0.074)	(0.074)
$Number_of_Partners$	−1.429	-1.586^{+}
	(0.901)	(0.900)
$Blue_Call$	4.385***	4.385***
	(1.199)	(1.197)
$Recent_Base_Dispatches$	−0.001	0.002
	(0.035)	(0.035)
$Average_Base_Dispatches$	0.001	0.001
	(0.006)	(0.006)
Observations	5483	5483
Adjusted R^2	0.096	0.097

All models include fixed effects for new recruits, main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions, which are not shown here. The standard errors reported are heteroscedasticity-consistent and clustered by new recruit. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

The model in Column (B) includes the interaction between high workload during the shift and partner exposure. The coefficient (and significance) for the direct effect of partner exposure is slightly lower ($\alpha_1 = 4.5$ minutes, p =0.031) during periods of high utilization but the aggregate effect of partner exposure on scene performance is higher and more significant (with $\alpha_2 = 3.6$ minutes, p =0.007) during periods of high utilization. This result indicates that teams involving new recruits who have had exposure to a more diverse set of partners are able to perform faster during periods of high workload than teams whose new recruits have less partner exposure, thus providing support for **H3**.

The effect of team familiarity is negative and significant, indicating that higher team familiarity of paramedic teams results in shorter scene times. Specifically, with each additional joint ambulance transport the scene time is shortened by about 0.84 seconds ($p = 0.023$). According to our data a new recruit can accumulate a $Team_Familiarity$ of up to 248 jobs with the same partner in the

first year. This would correspond to about a 3.5 minute (11%) reduction of average scene times for that paramedic pair, throughout the course of the year[8].

In terms of control variables, we observe that the coefficient of the linear term for individual experience is always negative and the coefficient of the squared term is always positive albeit very small, indicating slightly diminishing returns. While both coefficients are insignificant (p-values around 0.3) the coefficient estimates indicate that scene time shorten by over a second with each additional ambulance transport the new recruit conducts in the first year.

Similarly, the coefficient of task experience is negative, indicating a shorter time at the scene for new recruits who have more experience dealing with patients with the same primary condition as the patient at hand. However, this coefficient is not significant ($p = 0.699$)[9]. The results for the individual experience and the task-specific experience, taken together, suggest that while new recruits seem to become more efficient at the scene as they gain experience, their prior partner exposure and experience with the partner at hand have a much stronger effect on performance than overall experience and experience with the specific complaint of the patient. One possible explanation, provided by LAS paramedics in our discussions with them, is that the non-medical problem solving and coordination challenges faced by new recruits at a scene have a more significant impact on performance than disease-specific challenges.

In terms of workload controls we observe that team workload ($Crew_Workload$) is significant but system workload ($Recent_Base_Dispatches$ and $Average_Base_Dispatches$) is not. Specifically, the coefficient of $Crew_Workload$ is negative indicating that if the crew has experienced high prior utilization during the shift they spend less time at the scene. In other words it seems that teams tend to speed up at times of high utilization. While shorter scene times are desirable in general (both from a patient and network perspective) we do not have outcome data to analyze whether speeding up at the scene is associated with better outcomes. What our results show, however, is that all teams speed up (on average) at times of high utilization ($\alpha_7 = -1.895,\ p = 0.002$) but teams including new recruits with more partner exposure are (on average) able to speed up even more than their counterparts (as illustrated by the previously discussed coefficient of the $Partner_HHI * High_Workload$ interaction).

Finally we observe that patient pick-ups which result in a blue call to the A&E due to the seriousness of the case at hand require significantly longer times at the scene than pick-ups of patients with less serious conditions.

[8] We note that our sample includes teams in which a single new recruit works with one or more senior paramedics. Therefore the measure for team familiarity is likely to be underestimated as the senior paramedics may have worked with each other in the past, but those data are not available to us. Our robustness section (§8) includes results for the same models run on the subsection of data which includes only teams of one new recruit and one senior partner.

[9] We have conducted robustness checks allowing for diminishing returns of $Team_Familiarity$ and $Task_Exp$ as well. This has no effect on our results (regression tables omitted).

6.3. Partner exposure and handover time (H2 & H4)

Next, we estimate the impact of partner exposure on the performance of the patient handover process at the A&E using the following panel specification:

$$\begin{aligned} Handover_Time_{rt} = {} & \beta_0 + \beta_1\ Partner_HHI_{rt} + \beta_2\ Partner_HHI_{rt} * Suff_Exp_{rt} \\ & + \beta_3\ Partner_HHI_{rt} * Suff_Exp_{rt} * High_Workload_{rt} \\ & + \beta_4\ Team_Familiarity_{rt} + \beta_5\ Exp_{rt} + \beta_6\ Exp_{rt}^2 + \beta_7\ Task_Exp_{rt} + \beta_8\ A\&E_Exp_{rt} \\ & + \beta_9\ Crew_Workload_{rt} + \beta_{10}\ Partner_Tenure_{rt} + \beta_{11}\ Number_of_Partners_{rt} \\ & + \beta_{12}\ Blue_Call_{rt} + \beta_{13}\ Recent_A\&E_Arrivals_{rt} + \beta_{14}\ Average_A\&E_Arrivals_{rt} \\ & + \beta_{\mathbf{15}}\ Controls_{rt} + \epsilon_{rt}^H, \end{aligned} \tag{4}$$

which includes the same variables as (3) with four exceptions. First, we include an interaction between the $Partner_HHI$ variable and $Suff_Exp$. A positive and significant β_2 would support **H2**. To test **H4** we include a triple interaction to evaluate the moderating effect of high workload on the impact of $Partner_HHI$ for sufficiently experienced new recruits. Second, we include the cumulative prior experience of new recruits at the specific A&E, to account for any hospital specific learning effects. Third, we control for A&E workload instead of base workload as the workload in an A&E is likely to directly affect the handover of patients. Fourth, we include additional A&E fixed effects to account for any time-invariant characteristics which might impact handover times. For the handover we calculate two-way standard errors clustered by A&E and new recruits (Cameron et al. 2011), which allows for the errors of two observations to be correlated as long as either the new recruit or the A&E department is the same.

Table 4 reports results for model (4). First, from our base model we note that while the coefficients of $Partner_HHI$ and the interaction $Partner_HHI * Suff_Exp$ are both positive—indicating shorter handover times for teams whose new recruit have higher parter exposure—the effect is statistically significant only for teams including new recruits with higher levels of individual experience. Specifically, this implies that increase in partner exposure reduces handover times significantly only when cumulative experience of the new recruit is greater than the median number of transports conducted in the first four months on the job. This provides support for **H2**. In particular, the interpretation of the β_2 coefficient is that having a fully diverse partner experience ($Partner_HHI = 0$), as opposed to a fully focused partner experience ($Partner_HHI = 1$), reduces the handover time of teams involving new recruits by 2.6 minutes (p =0.046) on average when cumulative experience of the new recruit is greater than the median number of transports conducted by a new recruit in their first four months (i.e. 157 transports). Alternatively, increasing the partner exposure of new recruits such that $Partner_HHI$ is reduced by one standard deviation would shorten handover

Table 4 OLS coefficient estimates for handover time regressions

	Dependent variable: Handover time	
	(C)	(D)
Partner exposure & Team familiarity	**Base model**	**Workload interaction**
$Partner_HHI$	1.200	1.206
	(1.191)	(1.191)
$Partner_HHI * Suff_Exp$	2.643*	1.159
	(1.325)	(1.476)
$Partner_HHI * Suff_Exp * High_Workload$		5.201**
		(1.932)
$Team_Familiarity$	-0.008^{+}	-0.008^{+}
	(0.005)	(0.005)
Experience controls		
Exp	−0.014	−0.014
	(0.012)	(0.012)
Exp^2	0.000	0.000
	(0.000)	(0.000)
$Task_Exp$	0.002	−0.001
	(0.018)	(0.018)
$A\&E_Experience$	-0.011^{+}	-0.011^{+}
	(0.006)	(0.005)
Other time-varying controls		
$Crew_Workload$	−0.091	−0.355
	(0.270)	(0.291)
$Partner_Tenure$	0.002	0.001
	(0.050)	(0.050)
$Number_of_Partners$	0.578	0.581
	(0.642)	(0.644)
$Recent_A\&E_Arrivals$	0.221***	0.222***
	(0.040)	(0.040)
$Average_A\&E_Arrivals$	0.356**	0.348**
	(0.133)	(0.130)
$Blue_Call$	−4.875***	−4.840***
	(0.732)	(0.724)
Observations	5568	5568
Adjusted R^2	0.114	0.116

All models include fixed effects for new recruits, A&E, main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions, which are not shown here.. The standard errors reported are heteroscedasticity-consistent and clustered by new recruit. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

times by 0.6 minutes (3.3%) once the new recruits have accumulated sufficient experience. Our results are robust to changes in this threshold (Appendix C)[10].

Column (D) lists the estimates for the handover time regression with the interaction of prior partner exposure with both sufficient experience and high workload. We observe that the triple interaction of $Partner_HHI * Suff_Exp * High_Workload$ has a positive and significant coefficient ($\beta_3 = 5.2$, $p = 0.007$), providing support for **H4**. Interestingly, once the high workload interaction is included in the regression, the coefficient of $Partner_HHI * Suff_Exp$ becomes insignificant

[10] Specifically, the coefficients of the two main interactions ($Partner_HHI * Suff_Exp$ and $Partner_HHI * Suff_Exp * High_Workload$) remain consistent in size and significance for values of $Exp_Threshold$ ranging from 156-171.

(while still positive). This indicates that the benefits of partner exposure for the handover process are focused on periods of high workload[11].

As before, the coefficients of team familiarity are negative, indicating that the higher the team familiarity of a paramedic crew conducting an A&E handover, the more efficient they are. The coefficient size indicates that the average reduction in handover time associated with each joint transport undertaken by the same team is about half a second ($p = 0.087$). In practical terms, if a new recruit accumulates 248 joint transports with the same partner (the maximum we observe in the first year for new recruits at LAS), that would result in a 2 (12%) minute shortening of handover times on average. Since we are using the entire sample, which includes teams of new recruits working with more than one senior paramedic, we may be underestimating the extent of team familiarity. By running the analysis on the subset of the data in which a new recruit works with a single senior paramedic we capture team familiarity exactly and its significance in all regressions becomes significant at the 5% level (see robustness checks in §8) with the coefficient sizes remaining comparable.

Similar to the scene regression results, the coefficient of the individual experience is negative ($\beta_5 = -0.014,\ p = 0.248$) but not significant and the squared term is positive but very small. However, we do observe that the prior experience of the new recruit at the particular hospital ($A\&E_Experience$) is more significant ($p = 0.054$). Specifically, each prior handover to a specific A&E results in a 0.67 second reduction in subsequent handover times at that same A&E. In our data the maximum number of handovers a new recruit conducts at the same hospital in their first year is 193, which would correspond to a shortening of handover times at that hospital of over 2 minutes on average, throughout the first year on the job. These results demonstrate that in the 'joint production process' between the triage nurses and the ambulance crew, it is more beneficial for the new recruit to have had some prior experience at the hospital than having had prior experience with the medical condition of the patient. As with the scene, the coefficients of task experience are highly non-significant (the estimated coefficients turn out to be positive but the standard errors are relatively large).

Interestingly, while at the scene the system workload and individual workload were insignificant and significant, respectively, the result is reversed during handover. The negative coefficient of the individual prior workload indicates that there might be some tendency for new recruits to speed up during times of high prior workload but this effect is not statistically significant. However, we observe a positive and very significant impact of the number of recent and average arrivals to the

[11] We also note that including an interaction of $Partner_HHI$ with $High_Workload$ does not result in a significant coefficient (regression table omitted), indicating that even at times of high workload it is indeed necessary for new recruits to have gained a sufficient level of individual experience for the benefit from partner exposure to accrue.

A&E on handover times—making them longer. This is most likely due to congestion effects at the A&E for the incoming crews.

Finally, we observe a negative and significant effect of blue calls on handover times. This is to be expected as the reason for making a blue call in the first place is to alert the A&E of a serious incoming case to expedite the handover process.

7. Implications for team formation strategies

Our results from the previous section confirm our hypotheses regarding the benefits of prior partner exposure and also reconfirm the benefits of team familiarity reported earlier in the literature. From a managerial perspective, these results raise a question about team formation strategies: Should managers aim to keep teams intact—to induce high team familiarity (Huckman and Staats 2013)—or change partner assignments more frequently—to induce high partner exposure? In this section, we address this question through counterfactual analysis based on our model predictions.

We consider two contrasting hypothetical team formation strategies for new recruits at the LAS. The first, a *partner focus strategy*, assumes that new recruits are assigned a stable partner from the beginning thus yielding $Partner_HHI = 1$ for all ambulance transports. The second, a *partner exposure strategy*, assumes that new recruits take turns working with every senior paramedic at their ambulance base thus yielding the lowest practically feasible value of $Partner_HHI$ for each transport. We generate counterfactual values for partner exposure and team familiarity based on these definitions. We then use this counterfactual data and the coefficient estimates of columns (B) and (D) in Tables 3 and 4 to predict the hypothetical performance for each ambulance transport for the two strategies.

The left-hand side of Table 5 includes the predicted values of scene time and handover time for each of the two partner assignment strategies[12]. Note that the partner exposure strategy outperforms the partner focus strategy at the scene by about 11% (29.5 minutes vs. 33.2 minutes) and by about 5.5% (16.2 minutes vs. 17.2 minutes) for handover. Consequently, these results suggest that emphasizing exposure to multiple partners is preferable to emphasizing stable teams in the context of new recruits at the LAS, and particularly for the divergent process of patient pick-up at the scene.

Finally, to highlight the importance of incorporating partner exposure into performance predictions, we note that not including partner exposure in the analysis would result in the wrong managerial recommendation. We compare the predicted performance of the two team formation strategies, with and without including the impact of partner exposure. We first re-estimate the regression coefficients of (3) and (4), ignoring the effect of partner exposure (i.e. imposing $\alpha_1 = \alpha_2 =$

[12] The current performance is 30.5 min and 16.5 min on average for Scene and Handover, respectively.

Table 5 Performance predictions for Partner Focus and Partner Exposure team formation strategies, with and without the knowledge of partner exposure

	Prediction in knowledge of Partner Exposure		*Prediction in absence of Partner Exposure*	
	Scene time	Handover time	Scene time	Handover time
Partner focus strategy	33.2	17.2	29.7	16.0
Partner exposure strategy	29.5	16.2	30.9	16.6

The difference in performance for the two hypothetical team formation strategies (or the current performance) is statistically significant with $P < 0.000$.

$\beta_1 = \beta_2 = \beta_3 = 0$). We then use the new coefficient estimates to predict the scene and handover times for the two team formation strategies described above. As expected, the partner focus strategy is erroneously (due to the incomplete model, ignoring partner exposure) predicted to outperform the partner exposure strategy for both patient pick-up and handover processes. These findings emphasize that managers responsible for designing team formation strategies must simultaneously account for the benefits of partner exposure and team familiarity.

8. Robustness analysis

In this section, we investigate the robustness of our results to changes in the econometric specification or definitions of key variables (§8.1) and data inclusion/exclusion criteria (§8.2). Since each robustness check involves 4 separate regressions (corresponding to columns (A), (B), (C), and (D) from Tables 3 and 4) we relegate all the regression tables (Tables 9 to 18) from this section to Appendix E.

8.1. Robustness to variable definitions and specification

8.1.1. Log specification of dependent variable. In some of the prior literature on learning and operational performance, researchers have used a log-transformation of the dependent variable (e.g., Huckman and Staats 2011, Narayanan et al. 2013). We re-estimate a log-linear version of the specifications (3) and (4) and find that all our results are qualitatively the same. The coefficients of Partner HHI and Team Familiarity are slightly less significant for the scene regression (p-values 0.090 and 0.078, respectively), while the coefficients for all our main variables during handover have increased significance (see Table 9).

8.1.2. Alternate definition of partner exposure variable. As described in §5 we define our partner exposure based on individual partners, i.e. the HHI measure captures the concentration of prior experience on individuals. About 18% of our sample comprises ambulance transports with more than one partner. We therefore generate an alternative to the $Partner_HHI$ variable based on unique team compositions, instead of unique individuals, and denote it by $Team_HHI$. We repeat our analysis and find that all our results continue to hold. In particular we observe a positive and significant impact of $Team_HHI$ and $(Team_HHI * High_Workload)$ on scene time

and of ($Team_HHI * Suff_Exp$) and ($Team_HHI * Suff_Exp * High_Workload$) on handover times. Other results are qualitatively similar except that the coefficient of $Team_Familiarity$ becomes marginally insignificant in the handover regressions ($p = 0.104$ and $p = 0.125$—no change in significance for the scene regressions). Details are listed in Table 10.

8.1.3. Piece-wise linear experience controls. To allow for diminishing returns of the individual learning curve we re-run the analysis with piece-wise linear experience controls. For this analysis we replace the Exp and Exp^2 variables with an interaction between Exp and dummies indicating the median number of ambulance transports undertaken by new recruits in their first 1, 2, 3,..., 6 (and more) months. All the main results remain the same (see Table 11).

8.1.4. Including intrapersonal and interpersonal task diversity. Prior work has demonstrated the impact of interpersonal and intrapersonal task diversity on team performance (Bunderson and Sutcliffe 2002, Huckman and Staats 2011, Narayanan et al. 2013). In line with this literature we capture the difference in task experience across the members of an ambulance crew (interpersonal task diversity) using the average pairwise Euclidean distance between their prior task experience levels as follows:

$$Task_Diversity_{rt}^{Inter_P} = \frac{\sum_{n=1}^{N_{rt}} \sum_{m=1}^{N_{rt}} \sqrt{\sum_{c \in \mathcal{C}} \left(Task_Share_{rt}^{c,n} - Task_Share_{rt}^{c,m}\right)^2}}{\frac{N_{rt}(N_{rt}-1)}{2}},$$

where $Task_Share_{rt}^{c,n}$ denotes the fraction of prior transports of crew member n that involved patients with primary condition c. The above measure takes the value of 0 if all members of a particular ambulance crew have spent an equal proportion of their prior ambulance transports on each patient condition and increases as these proportions become different from each other.

Similarly, we calculate the intrapersonal task diversity of an ambulance crew by first calculating Herfindahl's index over prior tasks of each individual crew member and then averaging over these different indices as follows:

$$Task_Diversity_{rt}^{Intra_P} = \sum_{n=1}^{N_{rt}} \frac{\left(1 - \sum_{c \in \mathcal{C}} \left(Task_Share_{rt}^{c,n}\right)^2\right)}{N_{rt}}. \tag{5}$$

Including these variables in our regressions does not change any of the results. Furthermore, the coefficient of neither variable is significant. We do not believe this is because these are not important predictors for performance in general. Rather, a plausible explanation is that since tasks (in terms of the condition of the patients) are completely randomly assigned to crews (not only exogenous but more or less uniformly distributed) in our setting, these variables have very little variation for a given level of experience, in contrast to the variation we have for $Partner_HHI$ as illustrated in Figure 4. The coefficient estimates for this analysis are included in Table 12.

8.1.5. Nonlinear controls for the number of senior partners. For simplicity our main models control for the number of senior partners on a crew in a linear way through the actual count. We repeat the analysis using dummies for each value of the number of senior partners and report coefficient estimates in Table 13. Our results remain the same as before in magnitude and statistical significance, with the exception of a slight increase in the p-values of $Team_Familiarity$ for the handover regressions ($p = 0.115$ and $p = 0.073$).

8.1.6. Broader classification of the clinical condition. In the main analysis, $Task_Exp_{rt}$ is defined on the basis of the most detailed AMPDS classification (332 in total) of a patient's condition. This implicitly assumes that there is minimal spillover of knowledge across subgroups that affects the operating performance. To acknowledge the possibility of such spillover, we repeat our analysis using a broader definition of task based on 32 main groups of AMPDS classification. Our results do not change for either patient pick-up at the scene or patient handover performance. Results are included in Table 14.

8.1.7. Including sufficient experience interaction in scene regression. Including $Suff_Exp$ interactions in our scene regressions does not affect any of our results. Table 15 shows that including a ($Partner_HHI$ * $Suff_Exp$) interaction in our base model results in an insignificant coefficient of the interaction, whereas our main effect of $Partner_HHI$ remains similar in size and significance. Similarly, adding a triple interaction ($Partner_HHI$ * $Suff_Exp$ * $High_Workload$) to our high workload analysis does not result in a statistically significant effect whereas our main workload interaction ($Partner_HHI$ * $High_Workload$) remains statistically significant.

8.2. Robustness to data inclusion/exclusion

8.2.1. Exclusion of blue calls. Since the patient condition in ambulance transports involving a blue call is usually critical, the process for handing over such patients at the A&E is expedited. Although we have included fixed effects to account for this in all our analyses we also re-estimate our models excluding blue calls. All our results remain unchanged for this subset of the data (see Table 16).

8.2.2. Excluding the first 10 ambulance transports for each new recruit. By the definition of $Partner_HHI$, most new recruits are assigned a $Partner_HHI$ value of 1 following their first ambulance transport, as most likely they would have encountered only one senior partner. We therefore exclude the first two transports of each new recruit from our analysis. For robustness, we repeat our analysis excluding the first 10 transports of each new recruit, eliminating about 8% of our data. Despite the smaller sample size all our results continue to hold at the same significance levels as before (see Table 17).

8.2.3. Excluding transports with more than one senior. As we mention in the results section, our measure for Team Familiarity is likely to be underestimated for ambulance transports which include multiple senior partners, as they may have worked together prior to 2011. We repeat the analysis focusing on crews with only one senior. In this case, the effect of Team Familiarity is even stronger than in our main results. All other results remain unchanged (see Table 18).

9. Conclusion

The impact of team member interactions on team performance in operational systems has long been of interest to management scholars (Mas and Moretti 2009, Tan and Netessine 2018, Schultz et al. 1998, Valentine and Edmondson 2015). One of the dominant paradigms in team formation emerging out of the recent operations management literature is that of a beneficial effect of team familiarity on team productivity (Reagans et al. 2005, Huckman et al. 2009, Huckman and Staats 2011, Staats 2012). On the other hand, research in psychology and sociology has documented the benefits of occasional contacts and new team members on knowledge acquisition and creativity (Granovetter 1973, Gruenfeld et al. 2000, Choi and Thompson 2005). Our paper unifies these perspectives and provides empirical evidence for the simultaneous effect of both these aspects of team member experience (team familiarity and prior partner exposure) on the operational performance of teams using field data.

These findings have important managerial implications at two levels. First, in the specific empirical setting of the LAS, our results show that a team formation strategy that encourages partner exposure of new recruits can outperform one that encourages team familiarity. The magnitude of improvement (almost 5 minutes or 9.2%) is likely to have a significant impact on patient outcomes for approximately 20% of the transports in our data that are classified as "*Category A – Immediately life-threatening*" (Sacco et al. 2005, Bradley et al. 2006). Second, for more general service settings, our results highlight the need for managers to balance the beneficial effects of team familiarity and partner exposure by analyzing the work history of their employees (e.g. who has worked with whom in the past and on what tasks) and their performance. Such data is routinely captured by enterprise IT systems (Huckman and Staats 2013) and could be used to inform improved team formation strategies. Our work also highlights that the appropriate balance between team familiarity and partner exposure depends on the type of task at hand. Furthermore, the existence of fluid teams often indicates a turbulent setting where stable partnerships might be managerially difficult to ensure. In such settings, understanding the potential benefits of partner exposure can be particularly valuable.

Our work opens up several avenues for further investigation of the effects of partner exposure and its implications for team formation strategies. Although our results are indicative of how the

standardization (or lack thereof) of a process interacts with team formation, we are unable to draw definitive conclusions because we examine only two processes. Therefore, external validation of these results in other empirical settings would help to strengthen the knowledge base for practicing managers. Second, we focus our analysis on new recruits to exploit exogenous variation in the main independent variable of interest. Future studies should investigate whether the beneficial effects of partner exposure extend to workers with longer tenure. Third, while shortening scene and handover delays is considered a key objective at LAS—many settings are characterized by a general trade-off between speed and quality. Future work should develop knowledge to distinguish which properties of prior experience improve outcomes on each (or possibly both) of those fronts. Finally, given the multiple ways through which team composition affects operations performance simultaneously, future research could focus on embedding these effects into modern crew scheduling models, which have hitherto often focused on operational feasibility.

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 Lizzy Bovil, Patrick Brooks, Kevin Brown, Michael Damiani, Mark Gatfield, Brian Jordan, Jason Killens, Peter McKenna, and Martin McTigue. We are also grateful to John Murray and Neil Storey for valuable insights. The third author is grateful for the financial support provided by the Deloitte Institute of Innovation and Entrepreneurship.

References

T. Amabile. How to kill creativity. *Harvard Business Review*, 76(5):76–87, 1997.

L. Argote, C. A. Insko, N. Yovetich, and A. A. Romero. Group learning curves: The effects of turnover and task complexity on group performance1. *Journal of Applied Social Psychology*, 25(6):512–529, 1995.

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

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

S. L. Berman, J. Down, and C. W. Hill. Tacit knowledge as a source of competitive advantage in the national basketball association. *Academy of Management Journal*, 45(1):13–31, 2002.

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

W. F. Boh, S. A. Slaughter, and J. A. Espinosa. Learning from experience in software development: A multilevel analysis. *Management Science*, 53(8):1315–1331, 2007.

E. H. Bradley, J. Herrin, Y. Wang, B. A. Barton, T. R. Webster, J. A. Mattera, S. A. Roumanis, J. P. Curtis, B. K. Nallamothu, D. J. Magid, et al. Strategies for reducing the door-to-balloon time in acute myocardial infarction. *New England Journal of Medicine*, 355(22):2308–2320, 2006.

J. S. Bunderson and K. M. Sutcliffe. Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of management journal*, 45(5):875–893, 2002.

A. C. Cameron, J. B. Gelbach, and D. L. Miller. Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2), 2011.

H.-S. Choi and L. Thompson. Old wine in a new bottle: Impact of membership change on group creativity. *Organizational Behavior and Human Decision Processes*, 98(2):121–132, 2005.

J. R. Clark and R. S. Huckman. Broadening focus: Spillovers, complementarities, and specialization in the hospital industry. *Management Science*, 58(4):708–722, 2012.

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

P. F. Delaney, L. M. Reder, J. J. Staszewski, and F. E. Ritter. The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, 9(1):1–7, 1998.

M. Delasay, A. Ingolfsson, and K. Schultz. Inventory is people: How load affects service times in emergency response. In *Cross-Functional Inventory Research*, pages 21–49. World Scientific, 2016.

M. Delasay, A. Ingolfsson, B. Kolfal, and K. Schultz. Load effect on service times. Working paper, 2017.

J. A. Espinosa, S. A. Slaughter, R. E. Kraut, and J. D. Herbsleb. Familiarity, complexity, and team performance in geographically distributed software development. *Organization Science*, 18(4):613–630, 2007.

J. D. Fisher, S. N. Brown, and M. W. Cooke. *UK Ambulance Service Clinical Practice Guidelines (2003)*. Joint Royal Colleges Ambulance Liaison Committee, London, 2006.

M. Freeman, N. Savva, and S. Scholtes. Gatekeepers at work: An empirical analysis of a maternity unit. *Management Science*, 63(10):3147–3167, 2016.

M. Freeman, N. Savva, and S. Scholtes. Economies of scale and scope in hospitals. *Working Paper*, 2018.

M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):l, 1973.

D. H. Gruenfeld, P. V. Martorana, and E. T. Fan. What do groups learn from their worldliest members? Direct and indirect influence in dynamic teams. *Organizational Behavior and Human Decision Processes*, 82 (1):45–59, 2000.

M. T. Hansen. The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44(1):82–111, 1999.

Health & Social Care Information Centre. Provisional Accident and Emergency Quality Indicators for England - By provider for July 2013. http://www.hscic.gov.uk/catalogue/PUB11898, 2014. Accessed on 2015-03-20.

R. Huckman and B. Staats. The hidden benefits of keeping teams intact. *Harvard Business Review*, 91(12): 27, 2013.

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

R. S. Huckman and B. R. Staats. 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, 2011.

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

A. A. Kane, L. Argote, and J. M. Levine. Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organizational Behavior and Human Decision Processes*, 96 (1):56–71, 2005.

R. Katz. The effects of group longevity on project communication and performance. *Administrative Science Quarterly*, 27(1), 1982.

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

S.-H. Kim, C. W. Chan, M. Olivares, and G. Escobar. Icu admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science*, 61(1):19–38, 2014.

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

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

LAS. Personal communication with Jason Killens, Director of Operations., 2015. Project meeting on February 27th.

LAS. Personal communication with Patrick Brooks., 2017. E-mail exchange in September.

K. Lewis, M. Belliveau, B. Herndon, and J. Keller. Group cognition, membership change, and performance: Investigating the benefits and detriments of collective knowledge. *Organizational Behavior and Human Decision Processes*, 103(2):159–178, 2007.

J. G. March. Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87, 1991.

A. Mas and E. Moretti. Peers at work. *The American economic review*, 99(1):112–145, 2009.

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

S. Narayanan, J. M. Swaminathan, and S. Talluri. Knowledge diversity, turnover, and organizational-unit productivity: An empirical analysis in a knowledge-intensive context. *Production and Operations Management*, 2013.

I. Nonaka. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1):14–37, 1994.

K. Ramdas, K. Saleh, S. Stern, and H. Liu. New joints more hip? learning in the use of new components. Working paper, March 2018.

R. Reagans, L. Argote, and D. Brooks. 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, 2005.

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

K. L. Schultz, D. C. Juran, J. W. Boudreau, J. O. McClain, and L. J. Thomas. Modeling and worker motivation in jit production systems. *Management Science*, 44(12-part-1):1595–1607, 1998.

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

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

B. R. Staats and F. Gino. Specialization and variety in repetitive tasks: Evidence from a japanese bank. *Management Science*, 58(6):1141–1159, 2012.

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

T. F. Tan and S. Netessine. When You Work with a Superman, Will You Also Fly? An Empirical Study of the Impact of Coworkers on Performance. *Management Science*, Forthcoming, 2018.

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

M. A. Valentine and A. C. Edmondson. Team scaffolds: How mesolevel structures enable role-based coordination in temporary groups. *Organization Science*, 26(2):405–422, 2015.

Appendix A: Partner assignment of new recruits on the relief roster

According to various operations managers at LAS the assignment of new recruits on the relief roster to shifts and partners is driven by operational convenience without consideration of their previous experience, tenure, or performance. To better understand this assignment we perform a number of checks using the data.

First, one might suspect that new recruits may be randomly assigned to partners at first but become more likely to be assigned a stable partner as their experience grows. For example, it could be the case that a stable partner assignment occurs if the new recruit and a particular partner work well together and ask to be teamed up, which would be an omitted variable that would impact performance. Using the data, we examine if it is the case that stable partners emerge over time as follows. For each new recruit, define each period of his or her consecutive transports with the same team as a *stint*, and the number of transports in a stint as its duration. Note that a new recruit's experience at any point in time is the sum durations of all his or her prior stints. If new recruits are more likely to pair off with stable partners over time, then stint duration should be positively correlated with experience. However, in our data, this correlation is close to zero and is insignificant (-0.0026, p-value 0.9155).

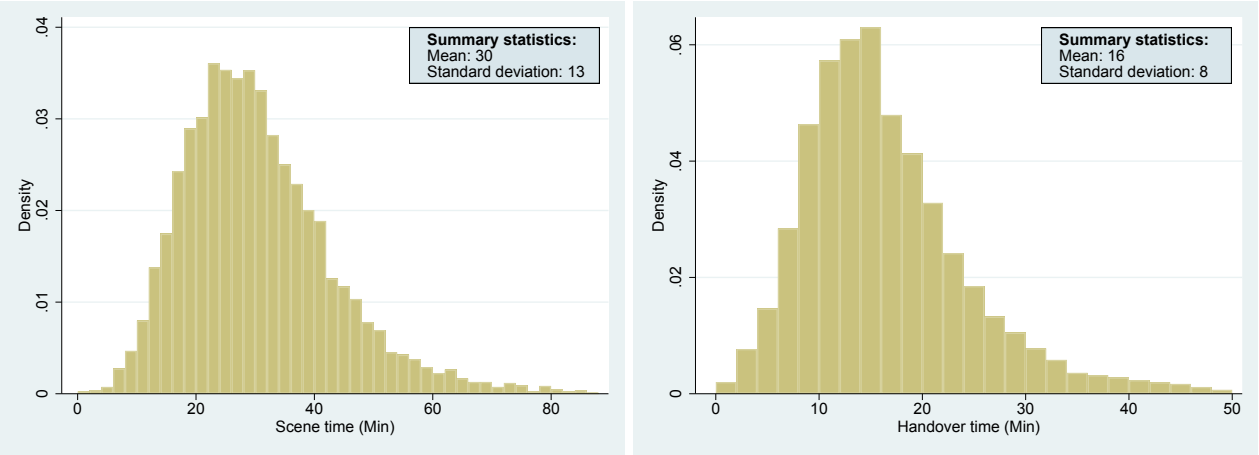
Second, the above analysis does not rule out the possibility that a new recruit mostly works with the same partner but occasionally does shifts with other partners. In such cases there would be little correlation between the total number of partners during the year and the total number of stints during the year for a new recruit. In contrast, we find that in our data there is a high correlation between the number of distinct partners and the number of stints (0.6343, p-value 0.0000) for the new recruits in 2011. Furthermore, the average ratio of the number of partners to the number of stints is 0.8, where this average is taken across all new recruits. This indicates that on average eighty percent of a new recruit's stints are with different partners.

Finally, none of the above results change by defining *stints* as consecutive transports undertaken by a new recruit with a single partner, regardless of whether other senior partners join the team for some transports.

Appendix B: Distribution of outcome variables

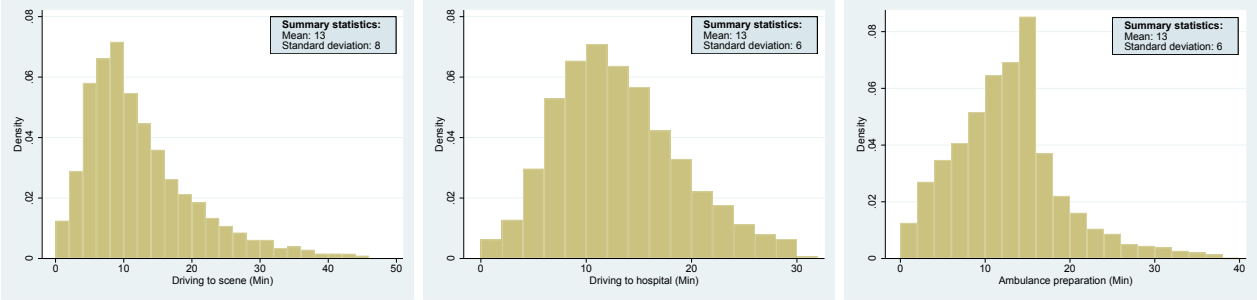
The main outcome variables in our analysis are scene time and handover time. The distribution of those is plotted below.

Figure 5 Descriptive statistics for the main outcome variables.



For comparison we also include the other three delay components from the time-line in Figure 2, below.

Figure 6 Descriptive statistics for the other delay components of an ambulance transport.



Appendix C: High individual experience threshold sensitivity analysis

Table 6 below summarizes the results of 62 robustness checks. Each line reports the coefficient estimates corresponding to the changing the threshold value of the indicator $Suff_Exp$ for model (4). The first column lists the new threshold value, the second and third columns contain the coefficient estimate and standard errors for the β_2 coefficient of the $Partner_HHI * Suff_Exp$ interaction (when $\beta_3 = 0$), respectively. The fourth and fifth column contain the coefficient estimates and standard errors for the β_3 coefficient of the $Partner_HHI * Suff_Exp * High_Workload$ interaction in the full regression.

Table 6 Result sensitivity to the high experience threshold of Table 4

	Sensitivity analysis for model (C) in Table 4		Sensitivity analysis for model (D) in Table 4	
Threshold for High Exp.	Estimate of β_2	Robust std. err.	Estimate of β_3	Robust std. err.
145	0.872	(1.172)	4.411*	(1.890)
146	1.519	(1.216)	4.517*	(1.920)
147	1.928	(1.178)	4.659*	(1.950)
148	2.027$^+$	(1.214)	4.664*	(1.977)
149	2.361$^+$	(1.264)	4.714*	(1.878)
150	2.184$^+$	(1.230)	4.713*	(1.852)
151	1.810	(1.232)	4.744*	(1.876)
152	2.248$^+$	(1.271)	4.819*	(1.896)
153	2.238$^+$	(1.270)	5.205**	(1.919)
154	2.060	(1.326)	5.177**	(1.915)
155	2.105	(1.319)	5.217**	(1.917)
156	2.386$^+$	(1.356)	5.221**	(1.943)
157	2.643*	(1.325)	5.201**	(1.932)
158	2.755*	(1.316)	5.280**	(1.973)
159	3.066*	(1.344)	5.307**	(1.994)
160	3.137*	(1.330)	5.364**	(2.002)
161	3.265*	(1.373)	5.609**	(1.915)
162	3.181*	(1.527)	5.688**	(1.993)
163	3.072*	(1.465)	5.759**	(1.973)
164	3.434*	(1.455)	5.623**	(1.979)
165	3.228*	(1.434)	5.775**	(1.962)
166	2.998*	(1.466)	5.789**	(2.002)
167	2.683$^+$	(1.532)	6.109**	(1.990)
168	3.016*	(1.522)	6.182**	(2.046)
169	3.180*	(1.548)	6.243**	(2.104)
170	2.989*	(1.458)	5.485**	(1.931)
171	2.461$^+$	(1.347)	5.512**	(1.993)
172	2.296	(1.444)	5.392**	(2.034)
173	2.106	(1.393)	5.601**	(2.051)
174	2.071	(1.473)	5.983**	(2.067)
175	1.966	(1.537)	5.867**	(2.084)

Appendix D: High individual workload threshold sensitivity analysis

Tables 7 and 8 summarize the results of repeating the main analysis (scene and handover regressions, respectively) for alternate definitions of high workload. The tables demonstrate that the results are not sensitive to the precise threshold definition. For the scene, we observe that for a range of reasonable values defining high workload the direct effect of partner exposure is still statistically and economically significant. This is in line with the results in Table 3. Similarly, we can replicate the handover results from Table 4 for a range of thresholds. In particular, the benefits of partner exposure for the handover operation seem to be concentrated to times of high workload.

Table 7 Sensitivity of scene results (Table 3) to threshold definition of $High_Workload$.

Shift utilization threshold	Percentile of transports	$Partner_HHI$ estimates		$Partner_HHI * High_Workload$ estimates	
		α_1	Robust SE	α_2	Robust SE
0.78	0.41	4.689*	(2.048)	1.637	(1.235)
0.79	0.39	4.680*	(2.056)	1.756	(1.133)
0.80	0.38	4.593*	(2.066)	2.152*	(1.074)
0.81	0.36	4.423*	(2.059)	2.739*	(1.133)
0.82	0.35	4.621*	(2.052)	2.302^{+}	(1.192)
0.83	0.33	4.487*	(2.047)	2.733*	(1.232)
0.84	0.32	4.526*	(2.053)	2.710*	(1.212)
0.85	0.30	4.409*	(2.051)	3.128*	(1.249)
0.86	0.28	4.509*	(2.051)	2.991*	(1.223)
0.87	0.27	4.458*	(2.067)	3.309*	(1.311)
0.88	0.25	4.475*	(2.069)	3.613**	(1.338)
0.89	0.23	4.599*	(2.062)	3.316*	(1.300)
0.90	0.22	4.513*	(2.055)	3.875**	(1.304)
0.91	0.20	4.657*	(2.046)	4.099**	(1.313)
0.92	0.18	4.741*	(2.056)	3.626**	(1.315)

Table 8 Sensitivity of handover results (Table 4) to threshold definition of $High_Workload$.

Shift utilization threshold	Percentile of transports	$Partner_HHI$ estimates		$Partner_HHI * High_Workload$ estimates		$Partner_HHI * Suff_Exp * High_Workload$ estimates	
		β_1	Robust SE	β_2	Robust SE	β_3	Robust SE
0.78	0.41	1.165	(1.191)	1.536	(1.696)	2.580	(1.910)
0.79	0.39	1.161	(1.191)	1.462	(1.712)	2.787	(1.949)
0.80	0.38	1.169	(1.192)	1.510	(1.736)	2.731	(2.020)
0.81	0.36	1.174	(1.192)	1.543	(1.767)	2.775	(2.264)
0.82	0.35	1.177	(1.189)	1.359	(1.704)	3.379	(2.182)
0.83	0.33	1.173	(1.189)	1.041	(1.698)	4.460*	(2.242)
0.84	0.32	1.181	(1.188)	1.060	(1.648)	4.590*	(2.172)
0.85	0.30	1.185	(1.189)	1.077	(1.564)	4.658*	(2.063)
0.86	0.28	1.203	(1.190)	1.009	(1.587)	5.017*	(2.026)
0.87	0.27	1.205	(1.191)	1.296	(1.482)	4.521*	(1.959)
0.88	0.25	1.206	(1.191)	1.159	(1.476)	5.201**	(1.932)
0.89	0.23	1.212	(1.188)	1.021	(1.488)	6.385**	(2.092)
0.90	0.22	1.229	(1.191)	0.933	(1.448)	7.139***	(2.099)
0.91	0.20	1.215	(1.195)	1.463	(1.304)	6.175**	(2.211)
0.92	0.18	1.173	(1.185)	1.607	(1.360)	7.183**	(2.786)

Appendix E: Regression tables for robustness results

E.1. Log specification of dependent variable

Table 9 OLS coefficient estimates for robustness checks

	DV: LN(Scene time)		DV: LN(Handover time)	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
$Partner_HHI$	0.111^{+}	0.082	0.040	0.040
	(0.066)	(0.066)	(0.070)	(0.077)
$Partner_HHI$ * $High_Workload$		0.121**		
		(0.046)		
$Partner_HHI$ * $Suff_Exp$			0.215**	0.111
			(0.083)	(0.101)
$Partner_HHI$ * $Suff_Exp$ * $High_Workload$				0.367***
				(0.096)
$Team_Familiarity$	0.000^{+}	0.000^{+}	-0.001^{+}	−0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Experience controls				
Exp	−0.001	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Task_Exp$	−0.001	−0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$A\&E_Experience$			0.000	0.000
			(0.000)	(0.000)
Other time-varying controls				
$Crew_Workload$	−0.042*	−0.070***	−0.017	−0.036*
	(0.017)	(0.019)	(0.017)	(0.017)
$Recent_Base_Dispatches$	0.000	0.000		
	(0.001)	(0.001)		
$Average_Base_Dispatches$	0.000	0.000		
	(0.000)	(0.000)		
$Recent_A\&E_Arrivals$			0.013***	0.013***
			(0.002)	(0.003)
$Average_A\&E_Arrivals$			0.026**	0.025**
			(0.009)	(0.008)
$Number_of_Partners$	−0.047	-0.052^{+}	0.042	0.042
	(0.031)	(0.031)	(0.040)	(0.048)
$Partner_Tenure$	−0.002	−0.002	0.002	0.002
	(0.003)	(0.003)	(0.004)	(0.004)
$Blue_Call$	0.123**	0.123**	−0.536***	−0.534***
	(0.039)	(0.039)	(0.077)	(0.078)
Observations	5483	5483	5568	5568
Adjusted R^2	0.108	0.109	0.143	0.145

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.2. Alternate definition of partner exposure variable

Table 10 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
$Team_HHI$	3.116*	2.206^{+}	0.154	0.210
	(1.264)	(1.330)	(0.733)	(0.729)
$Team_HHI$ * $High_Workload$		2.506**		
		(0.922)		
$Team_HHI$ * $Suff_Exp$			2.931*	1.834
			(1.162)	(1.369)
$Team_HHI$ * $Suff_Exp$ * $High_Workload$				3.892*
				(1.716)
$Team_Familiarity$	−0.013*	−0.013*	−0.008	−0.008
	(0.006)	(0.006)	(0.005)	(0.005)
Experience controls				
Exp	−0.020	−0.019	−0.018	−0.018
	(0.017)	(0.017)	(0.012)	(0.012)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Task_Exp$	−0.019	−0.022	0.002	−0.001
	(0.030)	(0.030)	(0.018)	(0.018)
$A\&E_Experience$			−0.011^{+}	−0.011^{+}
			(0.006)	(0.005)
Other time-varying controls				
$Crew_Workload$	−1.054^{+}	−1.779**	−0.091	−0.331
	(0.543)	(0.584)	(0.271)	(0.294)
$Recent_Base_Dispatches$	0.001	0.002		
	(0.035)	(0.035)		
$Average_Base_Dispatches$	0.001	0.001		
	(0.006)	(0.006)		
$Recent_A\&E_Arrivals$			0.220***	0.221***
			(0.040)	(0.040)
$Average_A\&E_Arrivals$			0.356**	0.351**
			(0.132)	(0.129)
$Number_of_Partners$	−1.770^{+}	−1.940*	0.584	0.572
	(0.904)	(0.908)	(0.667)	(0.670)
$Partner_Tenure$	−0.013	−0.011	0.001	0.001
	(0.075)	(0.074)	(0.050)	(0.050)
$Blue_Call$	4.362***	4.344***	−4.864***	−4.839***
	(1.196)	(1.194)	(0.728)	(0.720)
Observations	5483	5483	5568	5568
Adjusted R^2	0.096	0.097	0.114	0.115

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.3. Piece-wise linear experience controls

Table 11 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
Partner_HHI	4.528*	3.722^{+}	0.991	0.992
	(2.080)	(2.116)	(1.123)	(1.117)
*Partner_HHI * High_Workload*		3.543**		
		(1.318)		
*Partner_HHI * Suff_Exp*			2.750*	1.223
			(1.333)	(1.500)
*Partner_HHI * Suff_Exp * High_Workload*				5.179**
				(1.942)
Team_Familiarity	−0.014*	−0.014*	−0.008^{+}	−0.008^{+}
	(0.006)	(0.006)	(0.004)	(0.005)
Experience controls				
*Exp * Under_1m_Exp*	−0.008	−0.004	−0.001	−0.001
	(0.036)	(0.036)	(0.021)	(0.020)
*Exp * Over_1m_Exp* (33 transports)	−0.029	−0.025	−0.014	−0.014
	(0.022)	(0.022)	(0.013)	(0.013)
*Exp * Over_2m_Exp* (64 transports)	−0.040*	−0.037*	−0.013	−0.013
	(0.018)	(0.018)	(0.011)	(0.011)
*Exp * Over_3m_Exp* (93 transports)	−0.010	−0.007	−0.018*	−0.018*
	(0.015)	(0.015)	(0.008)	(0.008)
*Exp * Over_4m_Exp* (122 transports)	−0.017	−0.016	−0.020*	−0.019*
	(0.013)	(0.013)	(0.008)	(0.008)
*Exp * Over_5m_Exp* (150 transports)	−0.012	−0.010	−0.021*	−0.020*
	(0.012)	(0.012)	(0.010)	(0.010)
*Exp * Over_6m_Exp* (184 transports)	−0.007	−0.006	−0.018*	−0.018*
	(0.011)	(0.011)	(0.008)	(0.008)
Task_Exp	−0.021	−0.024	0.003	0.000
	(0.030)	(0.030)	(0.018)	(0.018)
A&E_Experience			−0.011*	−0.011*
			(0.005)	(0.005)
Other time-varying controls				
Crew_Workload	−1.044^{+}	−1.865**	−0.070	−0.334
	(0.541)	(0.594)	(0.265)	(0.284)
Recent_Base_Dispatches	−0.003	0.000		
	(0.035)	(0.035)		
Average_Base_Dispatches	0.000	0.000		
	(0.006)	(0.006)		
Recent_A&E_Arrivals			0.219***	0.220***
			(0.040)	(0.040)
Average_A&E_Arrivals			0.354**	0.347**
			(0.136)	(0.133)
Number_of_Partners	−1.784*	−1.927*	0.521	0.523
	(0.905)	(0.905)	(0.630)	(0.632)
Partner_Tenure	−0.031	−0.031	0.002	0.001
	(0.074)	(0.073)	(0.049)	(0.049)
Blue_Call	4.316***	4.319***	−4.855***	−4.822***
	(1.197)	(1.195)	(0.732)	(0.723)
Observations	5483	5483	5568	5568
Adjusted R^2	0.098	0.099	0.114	0.115

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.4. Including intrapersonal and interpersonal task diversity

Table 12 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
$Partner_HHI$	5.417**	4.512*	1.090	1.089
	(2.077)	(2.095)	(1.202)	(1.353)
$Partner_HHI * High_Workload$		3.628**		
		(1.341)		
$Partner_HHI * Suff_Exp$			2.780*	1.293
			(1.299)	(1.637)
$Partner_HHI * Suff_Exp * High_Workload$				5.249**
				(1.640)
$Team_Familiarity$	−0.013*	−0.013*	−0.009^{+}	−0.009*
	(0.006)	(0.006)	(0.004)	(0.004)
Experience controls				
Exp	−0.015	−0.013	−0.018	−0.017
	(0.017)	(0.017)	(0.013)	(0.015)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Task_Exp$	−0.018	−0.022	0.001	−0.002
	(0.030)	(0.030)	(0.018)	(0.015)
$A\&E_Experience$			−0.010^{+}	−0.010^{+}
			(0.006)	(0.005)
$Task_Diversity^{Intra_P}$	−1.120	−1.567	−0.034	−0.008
	(3.742)	(3.719)	(3.250)	(3.186)
$Task_Diversity^{Inter_P}$	1.458	1.330	−1.424	−1.528
	(2.609)	(2.610)	(1.726)	(3.186)
Other time-varying controls				
$Crew_Workload$	−1.045^{+}	−1.888**	−0.094	−0.361
	(0.544)	(0.596)	(0.274)	(0.285)
$Recent_Base_Dispatches$	−0.001	0.002		
	(0.035)	(0.035)		
$Average_Base_Dispatches$	0.001	0.001		
	(0.006)	(0.006)		
$Recent_A\&E_Arrivals$			0.222***	0.223***
			(0.040)	(0.042)
$Average_A\&E_Arrivals$			0.356**	0.349**
			(0.133)	(0.117)
$Number_of_Partners$	−1.358	−1.497	0.551	0.551
	(0.921)	(0.920)	(0.692)	(0.755)
$Partner_Tenure$	−0.012	−0.011	0.003	0.003
	(0.075)	(0.075)	(0.053)	(0.056)
$Blue_Call$	4.381***	4.386***	−4.864***	−4.829***
	(0.921)	(1.198)	(0.731)	(0.741)
Observations	5483	5483	5568	5568
Adjusted R^2	0.096	0.097	0.114	0.116

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.5. Nonlinear controls for the number of senior partners

Table 13 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
Partner_HHI	5.379**	4.511*	1.116	1.122
	(2.049)	(2.070)	(1.190)	(1.342)
*Partner_HHI * High_Workload*		3.599**		
		(1.336)		
*Partner_HHI * Suff_Exp*			2.571*	1.095
			(1.312)	(1.651)
*Partner_HHI * Suff_Exp * High_Workload*				5.176**
				(1.608)
Team_Familiarity	−0.014*	−0.014*	−0.007	−0.007^{+}
	(0.006)	(0.006)	(0.005)	(0.004)
Experience controls				
Exp	−0.018	−0.015	−0.017	−0.016
	(0.017)	(0.017)	(0.013)	(0.014)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Task_Exp	−0.019	−0.023	0.003	0.000
	(0.030)	(0.030)	(0.018)	(0.015)
A&E_Experience			−0.010^{+}	−0.010^{+}
			(0.006)	(0.005)
Other time-varying controls				
Crew_Workload	−1.061^{+}	−1.899**	−0.074	−0.336
	(0.543)	(0.595)	(0.272)	(0.277)
Recent_Base_Dispatches	0.000	0.003		
	(0.035)	(0.035)		
Average_Base_Dispatches	0.001	0.001		
	(0.006)	(0.006)		
Recent_A&E_Arrivals			0.221***	0.222***
			(0.040)	(0.041)
Average_A&E_Arrivals			0.359**	0.352**
			(0.133)	(0.118)
Two_Senior_Partners	−1.296	−1.443	−0.050	−0.043
	(0.971)	(0.969)	(0.688)	(0.810)
Three_Senior_Partners	−4.307	−4.628	6.795***	6.746***
	(3.391)	(3.408)	(1.273)	(1.304)
Partner_Tenure	−0.016	−0.016	0.008	0.008
	(0.074)	(0.074)	(0.050)	(0.052)
Blue_Call	4.386***	4.387***	−4.868***	−4.834***
	(1.199)	(1.197)	(0.733)	(0.737)
Observations	5483	5483	5568	5568
Adjusted R^2	0.096	0.097	0.115	0.116

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.6. Broader classification of the clinical condition

Table 14 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
$Partner_HHI$	5.237*	4.389*	1.183	1.189
	(2.048)	(2.070)	(1.185)	(1.333)
$Partner_HHI * High_Workload$		3.519**		
		(1.330)		
$Partner_HHI * Suff_Exp$			2.641*	1.162
			(1.321)	(1.658)
$Partner_HHI * Suff_Exp * High_Workload$				5.179**
				(1.602)
$Team_Familiarity$	−0.014*	−0.013*	−0.008^{+}	−0.008*
	(0.006)	(0.006)	(0.005)	(0.004)
Experience controls				
Exp	−0.022	−0.020	−0.015	−0.014
	(0.017)	(0.017)	(0.012)	(0.014)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Broad_Task_Exp$	0.019	0.018	0.004	0.003
	(0.012)	(0.012)	(0.005)	(0.004)
$A\&E_Experience$			−0.011^{+}	−0.011*
			(0.006)	(0.005)
Other time-varying controls				
$Crew_Workload$	−1.075*	−1.894**	−0.094	−0.357
	(0.542)	(0.594)	(0.270)	(0.277)
$Recent_Base_Dispatches$	−0.001	0.002		
	(0.035)	(0.035)		
$Average_Base_Dispatches$	0.001	0.001		
	(0.006)	(0.006)		
$Recent_A\&E_Arrivals$			0.221***	0.222***
			(0.040)	(0.041)
$Average_A\&E_Arrivals$			0.354**	0.347**
			(0.133)	(0.117)
$Number_of_Partners$	−1.533^{+}	−1.683^{+}	0.564	0.566
	(0.899)	(0.899)	(0.648)	(0.734)
$Partner_Tenure$	−0.019	−0.019	0.001	0.001
	(0.074)	(0.074)	(0.050)	(0.051)
$Blue_Call$	4.358***	4.356***	−4.875***	−4.843***
	(1.195)	(1.193)	(0.734)	(0.741)
Observations	5483	5483	5568	5568
Adjusted R^2	0.097	0.098	0.114	0.116

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.7. Including $Suff_Exp$ interactions in scene regressions

Table 15 OLS coefficient estimates for robustness checks

	DV: Scene time			
Partner exposure & Team familiarity	Main effect	Experience interaction	Workload interaction	All interactions
$Partner_HHI$	5.348**	4.901*	4.051^{+}	4.077^{+}
	(2.048)	(2.173)	(2.199)	(2.196)
$Partner_HHI$ * $Suff_Exp$		2.206	2.108	1.871
		(2.676)	(2.659)	(2.758)
$Partner_HHI$ * $High_Workload$			3.598**	3.464*
			(1.334)	(1.390)
$Partner_HHI$ * $Suff_Exp$ * $High_Workload$				0.933
				(3.334)
$Team_Familiarity$	−0.014*	−0.015*	−0.014*	−0.014*
	(0.006)	(0.006)	(0.006)	(0.006)
Experience controls				
Exp	−0.018	−0.023	−0.020	−0.020
	(0.017)	(0.019)	(0.019)	(0.019)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Task_Exp$	−0.019	−0.019	−0.023	0.000
	(0.030)	(0.030)	(0.030)	(0.000)
Other time-varying controls				
$Crew_Workload$	−1.054^{+}	−1.047^{+}	−1.885**	−1.901**
	(0.543)	(0.543)	(0.593)	(0.597)
$Recent_Base_Dispatches$	−0.001	−0.001	0.002	0.001
	(0.035)	(0.035)	(0.035)	(0.035)
$Average_Base_Dispatches$	0.001	0.001	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.006)
$Number_of_Partners$	−1.429	−1.432	−1.588^{+}	−1.581^{+}
	(0.901)	(0.901)	(0.900)	(0.898)
$Partner_Tenure$	−0.015	−0.013	−0.013	−0.013
	(0.074)	(0.074)	(0.074)	(0.074)
$Blue_Call$	4.385***	4.391***	4.391***	4.396***
	(1.199)	(1.199)	(1.198)	(1.197)
Observations	5483	5483	5483	5483
Adjusted R^2	0.096	0.096	0.097	0.097

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.8. Exclusion of blue calls

Table 16 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
Partner_HHI	5.926**	5.034*	1.444	1.450
	(1.970)	(1.996)	(1.233)	(1.464)
*Partner_HHI * High_Workload*		3.699**		
		(1.347)		
*Partner_HHI * Suff_Exp*			2.480^{+}	1.090
			(1.370)	(1.793)
*Partner_HHI * Suff_Exp * High_Workload*				4.826**
				(1.864)
Team_Familiarity	−0.014*	−0.014*	−0.007	−0.007*
	(0.006)	(0.006)	(0.004)	(0.003)
Experience controls				
Exp	−0.013	−0.010	−0.013	−0.012
	(0.018)	(0.017)	(0.012)	(0.013)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Task_Exp	−0.035	−0.038	0.000	−0.003
	(0.030)	(0.030)	(0.019)	(0.016)
A&E_Experience			−0.011*	−0.011*
			(0.005)	(0.005)
Other time-varying controls				
Crew_Workload	−0.985^{+}	−1.840**	−0.178	−0.425
	(0.543)	(0.592)	(0.254)	(0.276)
Recent_Base_Dispatches	−0.008	−0.006		
	(0.035)	(0.035)		
Average_Base_Dispatches	0.001	0.001		
	(0.006)	(0.006)		
Recent_A&E_Arrivals			0.224***	0.225***
			(0.042)	(0.044)
Average_A&E_Arrivals			0.386**	0.379**
			(0.137)	(0.123)
Number_of_Partners	−1.573^{+}	−1.721^{+}	0.642	0.644
	(0.904)	(0.904)	(0.622)	(0.697)
Partner_Tenure	0.004	0.003	−0.002	−0.003
	(0.076)	(0.076)	(0.053)	(0.054)
Observations	5332	5332	5417	5417
Adjusted R^2	0.096	0.098	0.111	0.112

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.9. Excluding first 10 transports for each new recruit

Table 17 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
Partner_HHI	5.308*	4.338^{+}	2.316	2.321
	(2.245)	(2.300)	(1.580)	(1.584)
*Partner_HHI * High_Workload*		4.412***		
		(1.333)		
*Partner_HHI * Suff_Exp*			3.015*	1.571
			(1.423)	(1.562)
*Partner_HHI * Suff_Exp * High_Workload*				5.047**
				(1.927)
Team_Familiarity	−0.014*	−0.013*	−0.011*	−0.011*
	(0.006)	(0.007)	(0.005)	(0.005)
Experience controls				
Exp	−0.020	−0.018	−0.020	−0.020
	(0.019)	(0.018)	(0.013)	(0.013)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Task_Exp	−0.013	−0.017	0.002	−0.001
	(0.030)	(0.030)	(0.018)	(0.018)
A&E_Experience			−0.009	−0.009
			(0.006)	(0.006)
Other time-varying controls				
Crew_Workload	−0.944^{+}	−1.908**	0.062	−0.211
	(0.555)	(0.606)	(0.298)	(0.326)
Recent_Base_Dispatches	0.011	0.015		
	(0.036)	(0.036)		
Average_Base_Dispatches	0.001	0.001		
	(0.006)	(0.006)		
Recent_A&E_Arrivals			0.231***	0.232***
			(0.045)	(0.045)
Average_A&E_Arrivals			0.336*	0.328*
			(0.139)	(0.136)
Number_of_Partners	−1.026	−1.194	0.638	0.640
	(0.913)	(0.909)	(0.679)	(0.679)
Partner_Tenure	−0.050	−0.052	0.018	0.017
	(0.074)	(0.073)	(0.051)	(0.051)
Blue_Call	4.073**	4.039**	−4.935***	−4.900***
	(1.293)	(1.301)	(0.767)	(0.755)
Observations	5123	5123	5197	5197
Adjusted R^2	0.098	0.100	0.116	0.117

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.

E.10. Excluding transports with more than one senior

Table 18 OLS coefficient estimates for robustness checks

	DV: Scene time		DV: Handover time	
Partner exposure & Team familiarity	Main effect	Workload interaction	Main effect	Workload interaction
$Team_HHI$	4.491*	3.931^{+}	1.923	1.900
	(2.068)	(2.076)	(1.291)	(1.450)
$Team_HHI$ * $High_Workload$		3.151*		
		(1.448)		
$Team_HHI$ * $Suff_Exp$			2.897*	1.873
			(1.162)	(1.473)
$Team_HHI$ * $Suff_Exp$ * $High_Workload$				3.699*
				(1.499)
$Team_Familiarity$	−0.017*	−0.017*	−0.011*	−0.011*
	(0.007)	(0.007)	(0.005)	(0.004)
Experience controls				
Exp	0.005	0.006	−0.017	−0.016
	(0.019)	(0.019)	(0.012)	(0.012)
Exp^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Task_Exp$	−0.028	−0.030	−0.005	−0.007
	(0.031)	(0.031)	(0.019)	(0.015)
$A\&E_Experience$			−0.009	−0.009
			(0.006)	(0.006)
Other time-varying controls				
$Crew_Workload$	−1.336*	−2.081**	0.080	−0.186
	(0.587)	(0.645)	(0.333)	(0.330)
$Recent_Base_Dispatches$	0.032	0.034		
	(0.039)	(0.039)		
$Average_Base_Dispatches$	0.003	0.003		
	(0.007)	(0.007)		
$Recent_A\&E_Arrivals$			0.233***	0.234***
			(0.041)	(0.045)
$Average_A\&E_Arrivals$			0.508***	0.501***
			(0.096)	(0.083)
$Number_of_Partners$				
$Partner_Tenure$	−0.042	−0.043	0.001	0.001
	(0.083)	(0.083)	(0.063)	(0.069)
$Blue_Call$	3.881*	3.911*	−4.217***	−4.189***
	(1.539)	(1.542)	(0.702)	(0.801)
Observations	4555	4555	4634	4634
Adjusted R^2	0.101	0.102	0.127	0.128

All models include fixed effects for new recruits, A&E (for handover), main conditions, severity scores, time of day, day of week, and month of year, as well as controls for weather conditions (for scene time). The standard errors are clustered as before, for scene and handover, respectively. $^{+}$, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% level, respectively.