Staffing Temporary Teams:
Understanding the Effects of Team Familiarity and Partner Variety

Song-Hee Kim
University of Southern California

Hummy Song
University of Pennsylvania

Melissa A. Valentine
Stanford University

Abstract

In the emergency department (ED), temporary teams of physicians and nurses work together to deliver patient care. We examine how team familiarity and partner variety affect the performance of ED teams and whether these effects vary depending on how long ago the experiences were accumulated. Addressing these questions are important for equipping managers with actionable recommendations, as the existing literature provides limited insight regarding the temporal nature of these effects. We use visit-level data from a hospital ED and leverage the *ad hoc* team assignment and the round-robin assignment of patients to these teams as our identification strategy. To measure team familiarity and partner variety, respectively, we vary the length of the lookback window from the past 1 to 12 months. We find that while there is a significant positive impact of team familiarity on performance, there is a significant negative impact of nurses' partner variety on performance and there is no meaningful impact of physicians' partner variety on performance. The magnitudes of these effects do not differ based on whether the experiences were accumulated longer ago as opposed to more recently, which suggests there is not a measurable forgetting effect. As such, ED managers should staff physician-nurse teams to maximize team familiarity, which will in turn minimize physicians' and nurses' partner variety. In addition, team familiarity should be measured using a relatively long lookback window. Using counterfactual analyses, we illustrate that doing so would allow the study ED to treat 7 percent more patients without increasing staffing levels.

Key words: temporary teams, team staffing, team familiarity, partner variety, health care

1. Introduction

In recent years, many industries have come to rely on temporary teams to accomplish complex, high-value work (Edmondson and Nembhard 2009, Huckman et al. 2009, Kerrissey et al. 2020, Mortensen and Haas 2018). This trend has been attributed to several factors, including new internet communication technologies that facilitate the assembly of temporary teams, as well as new macro-employment models where people pursue career paths that span projects, organizations, and industries (Benkler 2017, Cappelli 1999, Klein et al. 2006) . These temporary teams are often composed of members who are relative strangers to each other, and who are assembled on demand for short-term engagements that require them to coordinate tightly-coupled and complex work. Examples of temporary teams used in different industry settings include innovation project teams (Dugan and Gabriel 2013), crowdsourced "flash teams" and "flash organizations" (Retelny et al. 2014, Valentine et al. 2017), "tour of duty" start-up teams (Hoffman et al. 2013), "fluid" project teams (Staats and Upton 2011), and *ad hoc* virtual teams (Crisp and Jarvenpaa 2013). Temporary teams are also now common in professional service industries including health care, consulting, and law (Gardner and Valentine 2015, Weinberg et al. 2011). These industry trends prompt the need for research about the effectiveness of temporary teams that operate with extremely fluctuating membership.

A substantial research literature explores the ways in which various team characteristics contribute to effective teamwork and improved team performance. This research examines topics spanning from collective intelligence (Riedl and Woolley 2017, Woolley et al. 2010, 2015), conflict (Jehn 1995, De Wit et al. 2012), network structure (Balkundi and Harrison 2006, Cummings and Cross 2003, Reagans et al. 2004, Sackett and Cummings 2018), and psychological safety (Edmondson 1999, Siemsen et al. 2009), among others. One of the most consistent results in this literature is that the amount of team members' prior shared experiences (sometimes called team familiarity) is associated with the team's performance (Espinosa et al. 2007, Huckman et al. 2009, Huckman and Staats 2011, Ramachandran et al. 2018, Staats 2012). This result is particularly relevant for temporary teams, whose members are likely to have varying levels of familiarity or shared experience.

Despite the consistency of this result, and despite its relevance for temporary teams, the research literature linking team familiarity and team performance is incomplete in several fundamental ways. Few studies have explicitly recognized team familiarity as an issue related to team staffing that is rife with tradeoffs (e.g., Salehi et al. 2017). Limited research has looked beyond substantiating the relationship between team familiarity and performance, to explore models and research questions on how managers can understand these inherent tradeoffs in staffing temporary teams. For example, should managers staff teams such that members who have worked together before continue to work together (i.e., prioritize high levels of team familiarity) or such that they gain more experience working with a diverse set of partners,

even if that means they have not worked together before (i.e., prioritize high levels of partner variety)? When considering members' prior experiences working with one another, should the manager focus on more recently-accumulated experiences (e.g., a lookback window of the past 1 month), since experiences from longer ago may be forgotten? Or, should she capture experiences accumulated over a longer period of time (e.g., a lookback window of the past 12 months), since this forgetting effect may be quite minimal? These gaps in the literature pose a barrier to implementing research insights into practice, thereby leading managers to staff teams *ad hoc* rather than with a particular staffing objective to balance tradeoffs and optimize the actual impact of familiarity. Our observation is that more understanding about staffing temporary teams is needed—and is also now more feasible than ever because of the increasing prevalence of data stored in organizations' information technology systems available to help managers understand and optimize team staffing processes (Contractor 2013, Kerrissey et al. 2020, Pentland 2012).

In this paper, we leverage the archival trace data in the electronic health records (EHR) of a hospital emergency department (ED) to address these gaps in the research. In hospital EDs, temporary teams of physicians and nurses are formed *ad hoc* every shift, with no particular staffing policy or objective. We leverage this *ad hoc* assignment of physicians and nurses to teams and the round-robin assignment of patients to these physician-nurse teams to cleanly identify the effects of team familiarity and partner variety on team performance. We focus on time to disposition as the main measure of team performance in the ED, and use revisit within 48 hours as a secondary measure of interest. Together, these measures capture the speed and quality dimensions of team performance in this setting (Song and Veeraraghavan 2018). We use these data and this empirical setting to answer the following research questions:

(1) Across various lookback windows, what are the effects of team familiarity and partner variety on the performance of physician-nurse teams in the ED?

(2) How do the effects of team familiarity and partner variety vary depending on how long ago these experiences were accumulated?

(3) Given these findings, what should be the policy used to staff temporary teams of physicians and nurses in our study setting, and what is the potential impact of implementing such a policy?

In measuring team familiarity and partner variety, we allow the length of the lookback window to vary from 1 month to 12 months. To examine the extent to which the effects of team familiarity and partner variety vary depending on how long ago these experiences were accumulated, we analyze marginal effects at 1-month intervals over a 12-month time horizon. To assess the potential impact of implementing our recommended staffing policy, we conduct a counterfactual analysis in which we change the membership of the physician-nurse teams at our study ED, subject to various scheduling constraints.

We obtain the following results. First, we show that while team familiarity significantly improves team performance, partner variety does not. Specifically, we find that, regardless of the lookback window used to measure team familiarity and partner variety, cases assigned to a physician-nurse dyad with a high level of team familiarity (i.e., the physician and the nurse have previously worked together on many cases) exhibit a substantially shorter time to disposition. But when the nurse on the case has a high level of partner variety (i.e., has previously worked with a large number of distinct physicians), the time to disposition *increases*, regardless of the lookback window used. When the physician on the case has a high level of partner variety (i.e., has previously worked with a large number of distinct nurses), there is generally no difference in the time to disposition.

Second, we find that the effects of team familiarity and partner variety do not meaningfully differ based on whether the experiences were accumulated longer ago as opposed to more recently. In other words, we do not find evidence to support a forgetting effect, at least over the period of the past 12 months. These findings suggest that, when developing staffing policies with the goal of improving team performance, there may be gains to employing longer lookback windows (e.g., 12 months instead of 1 month).

Third, using counterfactual analyses, we illustrate that staffing ED teams to maximize team familiarity accumulated over a relatively long lookback window would result in an average reduction in time to disposition of up to 6.8 percent, which would allow the study ED to treat 7.3 percent more patients (or approximately 4,000 more patients annually). This effect size compares favorably to other effects that have been identified as managerial levers, such as workload effects (KC and Terwiesch 2009) and work design effects (Ibanez et al. 2018). As such, this paper has important practical implications for ED managers, offering a concrete recommendation that can be implemented to improve performance without increasing staffing levels or sacrificing the quality of care.

In addition to its practical contribution, this paper contributes to the literature on team performance. Our findings help further the understanding around how team characteristics, such as team familiarity and partner variety, impact the performance of temporary teams in complex service settings. To our knowledge, our work is among the first to explicitly examine the *temporal* nature of the effects of team familiarity and partner variety on performance (Ancona et al. 2001, Harrison et al. 2003). In addition, we add to the nascent literature on partner variety by identifying conditions under which it may work to hinder rather than bolster team performance.

2. Related Literature and Hypothesis Development

In this paper, we focus on two factors relating to the staffing of temporary physician-nurse teams in the ED: team familiarity and partner variety. Both relate to the composition of teams based on team

members' prior experiences with one another. While our paper focuses on physician-nurse dyads specifically, we discuss in this section a broader prior literature on teams of any size.

2.1 Team Familiarity and Performance

The first experience-related characteristic of team staffing we consider is *team familiarity*, which is defined as the amount of shared experience that team members have accumulated in working together on a focal task or project (Espinosa et al. 2007, Huckman et al. 2009). This is a structural property that emerges from prior team staffing patterns, and is particularly relevant to temporary teams, which are teams that are typically short-lived and characterized by highly fluid membership (Tannenbaum et al. 2012, Valentine 2018, Wageman et al. 2012).

A well-documented empirical result from the prior literature is that team familiarity has a significant impact on team performance—i.e., how well a team performs new work depends on how much work experience members have accumulated together in the past (e.g., Reagans et al. 2005). Much of this prior work identifies a positive impact. In the context of software development teams, Huckman et al. (2009) finds that an increase in team familiarity leads to a reduction in defects and an improvement in schedule adherence. Espinosa et al. (2007) also use the context of software teams to illustrate that when it is challenging to coordinate the team (e.g., due to size or geographic dispersion), team familiarity is even more helpful in improving performance. In the context of the film industry, Cattani et al. (2013) illustrates that team familiarity can mitigate the negative effects of having too many stars on the same team. In contrast, other work suggests that higher levels of team familiarity can have a negative impact on team performance. Regarding creative tasks, Ramachandran et al. (2018) find that team familiarity leads to standardization, which in turn negatively impacts team performance in terms of creativity. In the context of new product development teams in the video game industry, Ramachandran et al. (2018) also show that while team familiarity improves coordination, it deters creativity and, thus, can negatively impact the new product's market performance.

Albeit across a variety of settings, much of this prior literature has examined team familiarity in the context of relatively stable teams. In contrast, we focus on temporary teams, in which team members frequently and quickly come together for a discrete task and then disband and reassemble with a different set of team members for the next task. This is highly relevant to how most health care settings operate. Specifically in this setting, prior work has shown team familiarity to be associated with shorter surgical operative time (Avgerinos and Gokpinar 2017, Reagans et al. 2005, Xu et al. 2013), shorter lengths of stay (Patterson et al. 2015, Valentine and Edmondson 2015), and lower health care costs (Agha et al. 2018).

In our context of physician-nurse dyads in the ED, we expect team familiarity to positively impact team performance. This is because work in the ED is thought to benefit from evidence-based

recommendations and guidelines (Institute of Medicine 2001). Physicians and nurses who have had more experience working together in the past are more likely to have developed similar mental models of their respective roles and expertise (Lewis 2004, Ren and Argote 2011, Zhang et al. 2007), a similar language for talking about problems and solutions (Thompson and Fine 1999), shared experiences that enable them to quickly synchronize their understanding of and responses to changing situations (Mathieu et al. 2000, Yuan et al. 2018), and psychological safety to facilitate knowledge sharing (Edmondson 1999, Siemsen et al. 2009). All of this would enable the physician and nurse to work together faster, and without sacrificing on the quality of care. As such, we hypothesize the following:

H1: Team familiarity is positively associated with team performance.

In understanding the effect of team familiarity on performance, it is also important to consider whether the *recency* of the shared experience matters (e.g., Ancona et al. 2001, Harrison et al. 2003), a factor that to our knowledge has not yet been considered in the research literature. This lack of evidence is problematic because competing hypotheses are plausible. On the one hand, a shared work experience between two team members that occurred a year ago may have little or no effect on performance whereas one that occurred a month ago might have a stronger impact on performance. This differential relationship over time may be because the team members' shared mental models, shared language, and synchronized responses may be disrupted by different factors over longer periods of time (Anderson Jr. and Lewis 2014, Froehle and White 2014), including experiences with other teams (Kane et al. 2005), changes to the tasks or context (Leonard-Barton 1992), or simply the passing of time whereby experiences become less salient and retrievable (Ramdas et al. 2018). As team members encounter these various disruptions, they may see fewer benefits to performance from their shared experience together. If this were the case, managers should employ a short lookback window to measure team familiarity. If they were to instead use a long lookback window, they would over-credit the shared experiences that were accumulated longer ago and, thus, overestimate the stock of team familiarity from which team members stand to benefit. To examine this possibility, we test the following hypothesis:

H2a: The magnitude of the effect of team familiarity on team performance is smaller when the team familiarity was accumulated longer ago as opposed to more recently.

On the other hand, it may be the case that the aforementioned disruptions do not alter the benefits to performance from a shared work experience that occurred a long time ago. In the context of temporary teams, it is possible that the rapid pace of learning, and the gains to performance thereof, outpaces the rate of forgetting. This may be the case if the performance gains from higher levels of team familiarity stem primarily from intangible know-how rather than from concrete know-what (Paiva et al. 2008). In this case, managers may be better off employing a reasonably long lookback window to measure team familiarity, as team familiarity that was accumulated longer ago may still meaningfully contribute to the

stock of overall team familiarity from which team members stand to benefit. As such, we test the following competing hypothesis:

H2b: The magnitude of the effect of team familiarity on team performance is not significantly different whether the team familiarity was accumulated longer ago as opposed to more recently.

In the research to date on the effects of team familiarity, this temporal aspect has largely been ignored. In addition, much of the prior literature does not explicitly define a lookback window. Instead, researchers have typically calculated team familiarity by counting the number of cases on which team members have worked together since the beginning of the dataset being used for the study (Avgerinos and Gokpinar 2017, Espinosa et al. 2007, Reagans et al. 2005, Xu et al. 2013). This approach often results in lookback windows that vary in length by each observation, where observations near the beginning of the dataset have a shorter lookback window and observations near the end of the dataset have a longer lookback window (e.g., Espinosa et al. 2007). In contrast, some other studies do use a consistent lookback window across all observations by constraining the lookback window to a fixed length of time. For example, Huckman et al. (2009) uses a 2-year lookback window, Huckman and Staats (2011) uses a 3-year lookback window, Cattani et al. (2013) uses a 4-year lookback window, and Ramachandran et al. (2018) uses a 10-year lookback window, respectively, for all observations. Nevertheless, all of these works effectively assume that the impact of team familiarity is cumulatively additive rather than potentially diminishing over time, in which shared experiences accumulated longer ago have less of an effect on team performance.

2.2 Partner Variety and Performance

Relative to the prior research on the effects of team familiarity on performance, the literature on the effects of *partner variety* on performance is much more nascent, especially with regards to temporary teams. Nevertheless, partner variety is an important structural property of teams, and one that managers may need to trade off against team familiarity. Take, for example, an extreme case, in which a manager chooses to always keep one team together. In this scenario, these team members will have a high level of team familiarity but a low level of partner variety, because they will not have been exposed to other partners, who could have been a source of learning. Thus, it is important to understand the effects on performance of partner variety alongside that of team familiarity.

To our knowledge, Aksin et al. (2018) is the only work to date that explicitly considers the performance effects of partner variety together with team familiarity among temporary teams. The authors study ambulance transport teams (each of which is comprised of two paramedics) to examine how partner variety and team familiarity, respectively, affect team performance in terms of speed-related measures. They find that both partner variety and team familiarity positively impact the speed at which the transport teams work, and the benefits of partner variety outweigh that of the benefits of team familiarity. As such,

they recommend that managers in this setting should prioritize maximizing partner variety when staffing teams.

There are two key dimensions along which this literature should be extended. First is with regards to the solo-disciplinary versus cross-disciplinary nature of the individual members comprising the team. In the ambulance transport teams studied by Aksin et al. (2018), the two paramedics on the team differ in their level of individual experience but not in their discipline or in their roles and responsibilities as a member of the team; i.e., they are *solo-disciplinary* teams. However, in practice, many teams are comprised of individuals that vary in their disciplinary backgrounds and the functions for which they are responsible as a member of the team; i.e., they are *cross-disciplinary* teams.

A priori, it is not obvious whether cross-disciplinary teams would benefit from higher levels of partner variety in the same way solo-disciplinary teams do as discussed in Aksin et al. (2018). Similar to what is observed in the context of solo-disciplinary teams, partner variety could improve team performance by broadening the collective knowledge base of the team (Aksin et al. 2018, Cummings 2004, Harrison and Klein 2007, Reagans and Zuckerman 2001). On the other hand, if the cross-disciplinary nature of the team introduces differences in hierarchy or the level of discretion that team members have in making decisions, this may impact the extent to which they are able to work together effectively, which in turn could negatively impact team performance (Alexander et al. 2005, Dougherty 1992, Greer et al. 2018, Greer and Chu 2020, Lemieux-Charles and McGuire 2006). When this is the case, higher levels of partner variety may have a differential impact on team performance depending on the specific disciplinary role of the team member.

In the context of physician-nurse dyads in the ED, which are cross-disciplinary teams, we expect that the effect of partner variety on team performance may indeed be different for physicians and nurses. In this setting, the main role of the physician is to make plans and decisions (e.g., requesting clinical tests, performing clinical procedures, creating treatment plans), whereas the main role of the nurse is to execute these plans and decisions (e.g., obtaining patient samples for the clinical tests, monitoring and administering medications and intravenous infusions). As such, there is a difference in the level of discretion that physicians and nurses exercise with regards to decision making, which in turn may impact how higher levels of partner variety for physicians versus for nurses may impact team performance. For physicians, having previously worked with many different nurses may primarily serve to increase their exposure to new ways of thinking about patient care plans and prioritizing across patients. In other words, higher levels of partner variety may expose physicians to ideas and practices in a way that leads to learning and performance benefits through experimentation and exploration (Cummings 2004, Harrison and Klein 2007, O'Leary et al. 2011, Reagans and Zuckerman 2001). This leads us to the following hypothesis:

H3: Physicians' partner variety is positively associated with team performance.

For nurses, having previously worked with many different physicians may have a very different effect on team performance. In the ED physician-nurse dyad relationship, nurses need to regularly interact with physicians to carry out the care plan, which requires adapting to different practice preferences and styles of the physician. When nurses are faced with higher levels of partner variety, this may hinder how fast they can work because it takes more time and effort to adjust to working with new partners (Edmondson et al. 2007). Furthermore, in contrast to nurses vis-à-vis physicians, physicians may have less motivation to actively share knowledge and their thought processes with nurses, which may pose an impediment to nurses from benefitting from higher levels of partner variety (Siemsen et al. 2008). As such, we hypothesize the following:

H4: Nurses' partner variety is negatively associated with team performance.

The second dimension along which we seek to contribute to the literature on partner variety is by considering whether and the extent to which it matters how recently the partner variety was accumulated. Like much of the previous work on team familiarity, Aksin et al. (2018) calculate the measures of both team familiarity and partner variety by assuming cumulative additivity and by using a lookback window that varies in length across observations (i.e., it extends to the beginning of the study's dataset rather than spanning a fixed period of time). For the same reasons as discussed in section 2.1, it is important to examine the differential effects of this temporal dimension on the relationship between partner variety and team performance. As another experience-related team characteristic, partner variety may also be subject to forgetting effects (Anderson Jr. and Lewis 2014, Froehle and White 2014, Ramdas et al. 2018). Thus, we test the following set of competing hypotheses for physicians and nurses, respectively:

H5a: The magnitude of the effect of partner variety on team performance is smaller when the partner variety was accumulated longer ago as opposed to more recently.

H5b: The magnitude of the effect of partner variety on team performance is not significantly different whether the partner variety was accumulated longer ago as opposed to more recently.

3. Empirical Setting: Temporary Teams in Hospital Emergency Departments

Many emergency departments (EDs) have been facing an increase in patient volumes and higher levels of patient complexity without a corresponding increase in staffing levels (GAO 2009, Pitts et al. 2012). This has resulted in longer wait times and higher rates of patients leaving without being seen (GAO 2009), both of which are linked to worse patient outcomes in the form of higher rates of admission to the hospital and higher mortality rates, among others (Bernstein et al. 2009, Singer et al. 2011). To ensure access to emergency care with reasonable wait times, ED administrators have been looking for ways to improve patient throughput using existing resources without sacrificing the quality of care.

Recent reviews of the literature in health care management have highlighted that "operational characteristics play an important role in influencing patient outcomes and warrant just as much attention as patient-level clinical characteristics" (KC et al. 2020, p.75). A multitude of different operational levers have been studied in recent years, such as appointment scheduling (e.g., Gupta and Denton 2008, Liu et al. 2010, White et al. 2011, Zacharias and Pinedo 2014), bed utilization in hospitals (e.g., Allon et al. 2013, Berry Jaeker and Tucker 2017, KC and Terwiesch 2009, Kuntz et al. 2015, Roth et al. 2019), and the time of treatment (e.g., Anderson et al. 2014, Batt et al. 2019, Deo and Jain 2019, KC 2019). In the setting of the ED specifically, others have documented the impact of emphasis framing in clinical information systems (Laker et al. 2018), multitasking (e.g., KC 2014), operational flexibility (Laker et al. 2014, Ward et al. 2015), patient streaming and queueing (Saghafian et al. 2012, 2014, Song et al. 2015), peer influence (e.g., Song et al. 2018, Yuan et al. 2018), and staff quality and training (e.g., Kuntz and Sülz 2013). Yet, there exists an important and pervasive aspect of ED operations that has received little attention: how to staff physician-nurse teams, and the performance implications thereof. In this paper, we examine how the staffing of ED teams can be improved as a way to increase patient throughput. We focus specifically on the dyads of a physician and a nurse working together to deliver patient care, as they are the most commonly employed team unit in this setting.

To conduct our analyses, we use data from a hospital ED, which we anonymize as Metro ED. Metro ED is part of a high-volume, academic hospital in a large metropolitan area of the United States. As is typical of most EDs, temporary teams comprised of physicians and nurses deliver care to patients at Metro ED.

3.1 Physician-Nurse Dyad Work Processes at Metro ED

Work in the ED is highly interdependent between physicians and nurses, each of whom brings expertise from different disciplines. Metro ED organizes temporary teams of physicians and nurses into 3 pods, typically with 1 physician and 2 nurses in each pod at a given time. Physicians and nurses are assigned to a single pod for the duration of their shifts. Start times of shifts are staggered across providers in order to ensure that at least 1 provider stays on the team who knows about the patient. Each patient is assigned to a cross-disciplinary team within the pod, comprised of 1 physician and 1 nurse. Based on scheduling and training needs that are exogenous to patient conditions, some patients also have a resident physician or a second nurse as part of the care delivery team.[1] The physician assesses the condition of the patient, orders tests or imaging studies, makes a diagnosis, creates a treatment plan, and decides on the ultimate disposition of the patient (i.e., admission to the hospital or discharge to home or an outside facility). The nurse interacts more continuously with the patient, from placement into a treatment bed to information

[1] In section 6.2, we consider alternate sample definitions in which we exclude visits that included a resident physician or a second nurse (or both) as part of the care delivery team.

collection about the patient's medical status that the physician then uses in his/her diagnosing and decision-making. The nurse also carries out the physician's clinical orders (e.g., drawing blood from the patient and sending it to the lab, monitoring and administering medications and intravenous infusions). As such, there is clear distinction in the roles and responsibilities that each of the team members is expected to fulfill.

3.2 Staffing and Patient Assignment at Metro ED

A feature of the staffing process at Metro ED that is important for our analysis is that the physician manager and the nurse manager each staff the ED separately (with physicians and nurses, respectively), with no attempt to preferentially pair specific physicians with specific nurses. This staffing process seeks to accommodate individual preferences for the type and number of shifts (e.g., if a nurse wants to work three 12-hour shifts or five 8-hour shifts in a particular week), but the process of staffing each physician-nurse team is by random assignment. In other words, there is no attempt to pair specific physicians with specific nurses in creating the teams. This random assignment of physicians and nurses onto patient care teams is critical for our analysis, because it ensures that variation in team familiarity and partner variety is exogenous, rather than based on team member preferences or capabilities. In section A of the Online Supplement, we provide evidence from the data and several simulation analyses to show that the assignment of physicians and nurses onto teams is indeed random, subject to scheduling constraints.

In addition, patients are exogenously assigned to the physician-nurse teams. This is the result of a department policy of using round-robin assignment for fairness reasons. In Figure A.5 of the Online Supplement, we use the data to illustrate that the average Emergency Severity Index (ESI) level is tightly concentrated around 3 for all physician-nurse dyads, which suggests that Metro ED was adhering to the round-robin assignment policy in assigning patients to physician-nurse teams. Several papers have explored the performance implications of round-robin patient assignment (Chan 2016, Song et al. 2015, Valentine 2018). In this paper, we leverage this round-robin assignment to cleanly identify variation in performance that is exogenous and not based on differential task assignment.

4. Data

We collected data for every adult patient who received care at Metro ED from January 2008 to December 2011. For each ED visit, the data include the patient-level information including age, gender, a 5-level ESI (level 1 is the most urgent and level 5 is the least urgent) (Gilboy et al. 2011), discharge disposition, physician identifier, nurse identifier(s), resident identifier and whether the patient returned to the ED within 48 hours after discharge. The data also include several time stamps related to patient flow through the ED: arrival (time of patient arriving at the ED), nurse start (nurse first signing up for the patient), physician start (physician first signing up for the patient), disposition order, and departure (patient leaving

the ED). In our data, the physician of record is the physician who is initially assigned to the patient, even if the patient is handed off to an oncoming physician at the end of the shift. Note, the initial physician is responsible for completing a care plan for each patient who is handed off to an oncoming physician.

4.1 Sample Selection

Table B.1 of the Online Supplement describes our sample selection process. In order to utilize a consistent lookback window across all observations when constructing our variables of interest (team familiarity and partner variety), we exclude from the analysis sample those visits that occurred in the first 12 months of our 3-year data collection period. We also exclude visits by pediatric patients and those with a missing value for age, gender, physician identifier, or nurse identifier. We further limit our analysis sample to patients seen by physicians and nurses who treated at least 50 cases in the 3-year data collection period. Using discharge disposition information, we also exclude patients who died in the ED, were transferred to another hospital, or left without being seen by a care provider. We exclude patients whose ESI level is 1 (most urgent); these patients comprise less than 1 percent of our sample. Finally, we exclude patients whose time to disposition (difference between disposition order time stamp and patient arrival time stamp; see section 4.2 for details) is shorter than the 1st percentile value (28 minutes) or longer than the 99th percentile value (652 minutes) to remove outliers in the time to disposition.

The resulting final sample consists of 111,491 ED visits, with 71 unique attending physicians and 100 unique nurses who worked in 4,572 unique physician-nurse teams in our 2-year study period. During the 2-year study period, each physician worked with 64 nurses (s.d.=26) on 1,570 cases (s.d.=1,413), on average. Each nurse worked with 46 physicians (s.d.=15) on 1,115 cases (s.d.=916), on average.

4.2 Variables

4.2.1 Measures of Team Performance

We employ two key measures of team performance to capture the speed and quality dimensions of service in the ED: time to disposition and revisit within 48 hours. Time to disposition has been used frequently as a proxy for provider productivity in ED settings (e.g., Batt et al. 2019, Pourmand et al. 2013, Saghafian et al. 2014, Song et al. 2015). Discussions with several ED managers and clinicians also point to this as a first-order productivity measure of interest. In contrast to time in ED, which is defined as the time from patient arrival to departure from the ED, time to disposition focuses specifically on the time from a patient's arrival to the ED to the time a disposition order was signed (which indicates that a patient is ready to be discharged or admitted), thus excluding any time spent boarding in the ED or in an inpatient unit.[2]

[2] In section 6.1, we consider time in the ED and time from first provider to disposition as alternate measures of performance.

It is possible that the reduction in time to disposition is associated with lower quality of care, due to a speed-quality tradeoff (Anand et al. 2011, Bartel et al. 2019, Song et al. 2020). To examine this possibility, we consider a second-order measure of whether the patient returned to the ED within 48 hours after discharge—a proxy for quality (e.g., Batt et al. 2019, Han et al. 2015, Lerman and Kobernick 1987, Rising et al. 2014, Song et al. 2015, 2018).

The first panel of Table 1 shows summary statistics of these two measures of team performance. Patients' time to disposition was 197.59 minutes or 3.3 hours, on average. In our sample, 4 percent of the patients returned to the ED within 48 hours. Note that patients whose final disposition was admission to the hospital are excluded from this calculation.

---------- Insert Table 1 About Here ----------

4.2.2 Measures of Team Familiarity and Partner Variety

The key variables of interest are team familiarity and partner variety, respectively. For each observation, we define *team familiarity* to be the number of prior cases a given physician and a given nurse worked on together within a defined lookback window. The lookback windows we employ range from 1 month to 12 months, which were determined based on 20 interviews that we conducted with physician and nurse managers at 8 EDs. Note, these measures of team familiarity capture the number of times a given physician-nurse dyad has worked together, and therefore is not dependent on a focal role.

Unlike team familiarity, *partner variety* is dependent on a focal role; i.e., it captures the variety of partners a given physician or a given nurse, respectively, has had prior to the focal observation. We operationalize partner variety as the prior partner count. As such, a *physician's partner variety* is captured as the number of distinct nurses a given physician has worked with within the given lookback window, whereas a *nurse's partner variety* is captured as the number of distinct physicians a given nurse has worked with within the given lookback window. As we do with team familiarity, we employ a range of lookback windows from 1 month to 12 months.

In calculating team familiarity and partner variety measures, we count a case towards each of the measures only if the case was completed at least 12 hours prior to the focal patient's arrival. We impose this restriction in order to exclude experiences that were accumulated within the same shift. For example, a case would count towards the measure of team familiarity accumulated over the past 1 month only if the physician and nurse worked on the case together within the past 1 month *and* if they had completed the case at least 12 hours prior to the time the focal patient arrived in the ED.

In Table 1, we provide summary statistics of team familiarity and partner variety measures over select lookback windows (1 month and 12 months). We provide summary statistics for the full range of lookback windows we employ in Table B.2 of the Online Supplement. As we would expect, team

familiarity, the physician's prior partner count, and the nurse's prior partner count all increase as the lookback window becomes longer.

Table 2 presents correlation values between the team performance measures and each of the variables of interest. We find there is a negative correlation between the main variables of interest and time to disposition; this negative correlation is stronger for team familiarity than it is for the physician's prior partner count and the nurse's prior partner count, respectively. Additionally, given a 1-month lookback window, the correlation between team familiarity and each of the prior partner count variables is 0.26 (physician's prior partner count) and 0.24 (nurse's prior partner count), respectively. Table B.2 shows that the correlation values are similar for longer lookback windows.

---------- Insert Table 2 About Here ----------

4.2.3 Control Variables

Our data allow us to control for several patient-, provider- and ED-level covariates that can potentially affect our performance measures. To adjust for heterogeneity across patient types, we control for patient age, gender, and ESI level. We also control for seasonality by including dummies for patient arrival hour, day-of-week, and month. Because prior studies have shown that workload can affect worker performance (KC and Terwiesch 2009, Tan and Netessine 2014), we control for the number of cases the assigned physician is concurrently working on (i.e., physician current workload) and the number of cases the assigned nurse is concurrently working on (i.e., nurse current workload), as well as their squared terms. We also control for the number of cases currently in the ED (i.e., ED census) as a proxy for ED congestion. To account for the within-shift variation in service rates (Batt et al. 2019, Deo and Jain 2019), we control for the time since the start of the physician's shift. We also control for the presence of a resident and/or a second nurse on the case. Finally, we control for the number of prior cases the assigned physician and the assigned nurse, respectively, worked on since the beginning of the study period as a proxy for experience. For brevity, we refer to these measures as physician experience and nurse experience, respectively. Table 1 provides summary statistics of all control variables, and Table 2 shows their correlations.

5. Effects on Logged Time to Disposition

In this section, we address our first two research questions: (1) Across various lookback windows, what are the effects of team familiarity and partner variety on the performance of physician-nurse teams in the ED? (2) How do the effects of team familiarity and partner variety vary depending on how long ago these experiences were accumulated? For our analyses, we leverage the exogenous variation in team and task assignment that comes from the random assignment of physicians and nurses to teams and the round-robin assignment of patients to these teams.

5.1. Estimation Models

We begin by estimating the effects of team familiarity and partner variety on time to disposition, for each of the 12 lookback windows we consider. To account for the right-skewed distribution of the time to disposition variable, we employ a log transformation. Then, we estimate the following log-linear model at the encounter level for each lookback window *l*:

$$\log(TimetoDispo_i) = \gamma_0 + \gamma_1 TF_{MD_i,Nur_i,l} + \gamma_2 MDPartnerCount_{MD_i,l} + \gamma_3 NurPartnerCount_{Nur_i,l} + \delta \boldsymbol{X}_i + \rho_{MD_i} + \eta_{Nur_i} + \varepsilon_i \quad (1)$$

Here, $\log(TimetoDispo_i)$ represents the logged number of minutes from patient arrival to disposition for patient encounter i. $TF_{MD_i,Nur_i,l}$ denotes the number of prior cases that physician MD_i and nurse Nur_i worked on together during the lookback window *l* preceding encounter i. $MDPartnerCount_{MD_i,l}$ captures the number of distinct nurses that physician MD_i worked with during the lookback window *l* preceding encounter i. $NurPartnerCount_{Nur_i,l}$ captures the number of distinct physicians that nurse Nur_i worked with during the lookback window *l* preceding encounter i. $\boldsymbol{X}_i$ is a vector of control variables described in section 4.2.3. ρ_{MD_i} are physician fixed effects where MD_i is the physician for patient encounter *i*. Similarly, η_{Nur_i} are nurse fixed effects where Nur_i is the nurse for patient encounter *i*. Collectively, ρ_{MD_i} and η_{Nur_i} allow us to control for time-invariant aspects of physicians and nurses, respectively. Thus, our model assesses within-physician and within-nurse variance in measuring productivity. ε_i captures standard errors clustered by physician-nurse teams. The main coefficients of interest are γ_1, γ_2, and γ_3, which capture the effects of team familiarity, physician's partner variety, and nurse's partner variety, respectively, on time to disposition.[3]

Next, to examine whether the effects of team familiarity and partner variety vary depending on how long ago these experiences were accumulated, we estimate a model using a 12-month lookback window, allowing each of the effects of team familiarity and partner variety to vary depending on how many months ago they were accumulated. We do this using constructing measures of team familiarity, physician's prior partner count, and nurse's prior partner count over each 1-month interval from 1 month ago to 12 months ago. Specifically, we let $TF_{MD_i,Nur_i,(l-1,l]}$ denote the number of prior cases that physician MD_i and nurse Nur_i worked on together from *l* months ago to $l-1$ month(s) ago preceding encounter *i*. We similarly define $MDPartnerCount_{MD_i,(l-1,l]}$ and $NurPartnerCount_{Nur_i,(l-1,l]}$, respectively. Then, we estimate the following log-linear model at the encounter level:

[3] We check for multicollinearity by calculating variance inflation factors (VIF). When using a 1-month lookback window, the mean VIF in our empirical model is 3.0, which falls below the conventional threshold of 10 (Hair et al. 1998). We obtain similar results when using lookback windows ranging from 2 months to 12 months. This suggests that multicollinearity is not a concern in our model (Wooldridge 2012).

$$\log(TimetoDispo_i) = \gamma_0 + \sum_{l=1}^{12} \gamma_{1,l} TF_{MD_i,Nur_i,(l-1,l]} + \sum_{l=1}^{12} \gamma_{2,l} MDPartnerCount_{MD_i,(l-1,l]} + \sum_{l=1}^{12} \gamma_{3,l} NurPartnerCount_{Nur_i,(l-1,l]} + \delta \boldsymbol{X}_i + \rho_{MD_i} + \eta_{Nur_i} + \varepsilon_i \quad (2)$$

The main coefficients of interest are $\gamma_{1,l}$, $\gamma_{2,l}$, and $\gamma_{3,l}$ for $l = 1, \ldots, 12$, which capture the effects of team familiarity, physician's partner variety, and nurse's partner variety, respectively, across each of the 12 1-month intervals on time to disposition.

5.2. Results

Table 3 presents the results of estimating Equation 1 over 12 lookback windows ranging from 1 month to 12 months (see Table B.3 of the Online Supplement for full results tables with coefficients for control variables). Table 4 presents the results of estimating Equation 2. Below, we discuss whether the results yield support for each of our five hypotheses about the effects of team familiarity and partner variety on team performance.

---------- Insert Tables 3 and 4 About Here ----------

First, we find that higher levels of team familiarity are strongly associated with faster speed—more specifically a shorter time to disposition. This is evidenced by the fact that the each of the coefficients on team familiarity (γ_1 in Equation 1) in Table 3 is negative and statistically significant across all 12 specifications using lookback windows ranging from 1 month to 12 months. For example, in column (1) of Table 3, we find that a 1-unit increase in team familiarity accumulated over the past 1 month is associated with a 0.28-percent decrease in time to disposition (p<0.001). In other words, patients under the care of a physician-nurse team that has worked on 1 additional case together over the past month experience a 0.28-percent decrease in the time to disposition on the focal case. Another example, in column (12) of Table 3, shows that a 1-unit increase in team familiarity accumulated over the past 12 months is associated with a 0.1-percent decrease in time to disposition (p<0.001). As such, we find strong evidence in support of Hypothesis 1 that team familiarity is positively associated with faster speed.

Second, we find that the magnitude of the effect of team familiarity on team performance is *not* significantly different based on whether the team familiarity was accumulated longer ago as opposed to more recently; this finding offers support for Hypothesis 2b. This is evidenced by the fact that the coefficients for team familiarity in each of the 12 1-month intervals ($\gamma_{1,l}$for $l = 1, \ldots, 12$ in Equation 2) are not statistically different from one another (column (1) of Table 4). Specifically, the test of equality of coefficients fails to reject the null hypothesis that the coefficients are the same across all 12 intervals (p=0.1253). A similar test of equality comparing the coefficients from the most recent interval (between 0 to 1 month ago) and the least recent interval (between 11 to 12 months ago) also fails to reject the null that the two coefficients are the same (p=0.1349).

Third, we find that the effects of partner variety on time to disposition depend on the role of the focal provider. Specifically, we find that higher levels of physicians' partner variety are not statistically significantly associated with time to disposition (γ_2 in Equation 1); thus, Hypothesis 3 is not supported. On the other hand, higher levels of nurse's partner variety are associated with a longer time to disposition (γ_3 in Equation 1); thus, Hypothesis 4 is supported. As we see in Table 3, the coefficients on MD's prior partner count are not generally statistically significantly different from zero; when they are (e.g., column (1) with a lookback window of 1 month), the magnitude of this coefficient is quite small. On the other hand, the coefficients on RN's prior partner count tend to be positive and significant at the 0.01 level. For example, column (12) of Table 3 shows that patients under the care of a nurse who has worked with 1 additional distinct physician over the past 12 months experiences a 0.34-percent increase in time to disposition (p<0.001). As such, our findings suggest that partner variety as experienced by nurses is associated with slower speeds but partner variety as experienced by physicians does not have a meaningful impact on speed.

Fourth, we find that the magnitude of the effect of partner variety on team performance is *not* significantly different whether the partner variety was accumulated longer ago as opposed to more recently; this finding supports Hypothesis 5b. We see evidence of this in columns (2) and (3) of Table 4, which show that the coefficients for MD's prior partner count ($\gamma_{2,l}$for $l = 1, \dots, 12$ in Equation 2) and those for RN's prior partner count ($\gamma_{3,l}$for $l = 1, \dots, 12$ in Equation 2) are, respectively, not statistically different from one another. For both MD's prior partner count and RN's prior partner count, respectively, the test of equality of coefficients fails to reject the null hypothesis that the coefficients are the same across all 12 intervals (p=0.2143 for MD's prior partner count; p=0.4370 for RN's prior partner count). In addition, a similar test of equality comparing the coefficients from the 1-month intervals that are the farthest apart also fails to reject the null hypothesis that the two coefficients from the most recent and the least recent months are the same (p=0.2264 for MD's prior partner count; p=0.1361 for RN's prior partner count).

Taken together, these findings suggest that there is not a measurable forgetting effect. In other words, the effects of team familiarity and partner variety on team performance do not meaningfully depend on how recently the experiences were accumulated (at least over a period of the past 12 months). This suggests that experiences accumulated longer ago should contribute to the stock of overall team familiarity and partner variety just as much as experiences that were accumulated more recently. As such, managers should employ a reasonably long lookback window in measuring these experience-related team characteristics, so that they can more fully capture the true stock of team familiarity and partner variety.

Before adopting such a recommendation, it is important to first understand how the magnitude of the effects of team familiarity and partner variety compare across the shorter versus longer lookback

windows. Simply comparing the coefficients in Table 3 across the 12 lookback windows would be misleading, since each of the coefficients reports the magnitude of the effect that is associated with a 1-unit change in team familiarity and partner variety, respectively, whereas the average levels of these characteristics vary substantially across the 12 lookback windows (Table B.2 shows that each of these measures is generally increasing in the lookback window used to calculate it). To account for the differences across lookback windows, we compute the effect associated with a 1-standard deviation increase in team familiarity, physician's prior partner count, and nurse's prior partner count, respectively. For ease of interpretation, we present this visually in Figure 1. This shows that the percent change associated with a 1-standard deviation increase in team familiarity gradually *increases* in magnitude as the lookback window increases (Figure 1). For example, a 1-standard deviation increase in team familiarity over the past 1 month corresponds to a 1.27-percent decrease in time to disposition while a 1-standard deviation increase in team familiarity over the past 12 months corresponds to a 2.55-percent decrease in time to disposition. Again, this points to the idea that when developing staffing policies with the goal of decreasing time to disposition, there may be gains to employing longer lookback windows. We examine this in further detail in our counterfactual analyses presented in section 7.

---------- Insert Figure 1 About Here ----------

6. Additional Analyses and Robustness Checks

We conduct several additional analyses to assess the robustness of our findings and the sensitivity of our main results.

6.1. Alternate Measures of Performance

To check whether team familiarity and partner variety have any measurable effect on quality-related performance, we examine the effects of team familiarity and partner variety on the likelihood of returning to the ED within 48 hours of discharge from the ED. We estimate a logit model at the encounter level for each lookback window *l* that retains the same form on the right-hand side as Equation 1 but accounts for the binary nature of the dependent variable, $Revisit_i$. $Revisit_i$ equals 1 if patient encounter i resulted in a revisit to the ED within 48 hours after discharge, and 0 otherwise. For estimating this model, we restrict our sample to only those encounters that ended with a discharge from the ED (as opposed to an admission to the hospital).

We present the results in Table B.4 of the Online Supplement. Unlike our findings for time to disposition, here we find little evidence of statistically or economically significant changes in the likelihood of revisit within 48 hours associated with an increase in team familiarity or partner variety. When using lookback windows that are 8 months or longer, we do find results that are statistically significant at the 0.05 or 0.10 level. However, the magnitude of these results are very small and do not

rise to the level of economic significance. Specifically, using a 10-month lookback window, we find that a 1-unit increase in MD prior partner count is associated with a 1.0056 increase in the odds of the patient returning to the ED within 48 hours. Collectively, these results suggest that the decrease in time to disposition that we find to be associated with an increase in team familiarity and the increase in time to disposition that we find to be associated with an increase in RN's prior partner count do not seem to be simultaneously associated with a decrease in the quality of care.

Next, we also examine whether our results are sensitive to different ways of measuring speed in the context of the ED. To do this, we use two alternate measures of performance that are also commonly used: logged time in the ED and logged time from first provider to disposition. For each of these dependent variables, we estimate a log-linear model at the encounter level for each lookback window *l* that retains the same form on the right-hand side as Equation 1.

We begin by assessing the effects of team familiarity and partner variety on logged time in the ED. This measure is different from logged time to disposition in that it includes discharge processing times and boarding times. These portions of time are typically beyond the control of the ED providers, but nevertheless have a significant impact on ED operations. When using time to disposition as the speed-related team performance measure, we find our main results to be robust to the use of this alternate measure in terms of magnitude, directionality, and statistical significance (see Table B.5).

In Table B.6, we repeat the estimation for logged time from first provider to disposition. This captures the time elapsed between when the first provider (either a physician or a nurse) started interacting with the patient's EHR and the time when the disposition order was signed. With this model as well, we find our main results to be highly robust in terms of magnitude, sign, and statistical significance.

6.2. Alternate Sample Definitions

We consider alternate ways of defining the analysis sample, first by relaxing some of our exclusion criteria. Specifically, we include in our sample visits whose time to disposition is shorter than the 1st percentile value or longer than the 99th percentile value (Table B.7). We also include visits of patients whose ESI level is 1 (Table B.8). We find our results to be robust to this expanded sample.

We also consider imposing additional exclusion criteria. We exclude visits that included a second nurse as part of the care team (Table B.9), visits that included a resident physician as part of the care team (Table B.10), and visits that included a second nurse or a resident physician (or both) as part of the care team (Table B.11). Again, our results are highly robust to these additional sample exclusions.

6.3. Alternate Model Specification

We use quadratic specifications for team familiarity and partner variety to test whether a non-linear model that allows for differences in effects across varying levels of team familiarity and partner variety fits the data better. Specifically, for each lookback window *l*, we estimate the same model as in Equation 1 but

add three quadratic terms for team familiarity, MD's prior partner count, and RN's prior partner count. We find that the change in each of the model fit statistics is negligible, which suggests that using a non-linear modeling approach does not lead to an improvement in model fit. For example, Table 3 shows that estimating Equation (1) results in R^2 values ranging from 0.1686 to 0.1694 (depending on the length of the lookback window). For this same specification, we obtain AIC values ranging from 192,963 to 193,074 and BIC values ranging from 195,272 to 195,383. When we add quadratic terms, we obtain R^2 values ranging from 0.1686 to 0.1696, AIC values ranging from 192,935 to 193,072, and BIC values ranging from 195,273 to 195,410. We thus conclude that the quadratic specifications do not offer better model fit than the linear specifications.

7. Managerial Recommendation for Staffing Teams

To put into perspective the possible benefits of staffing ED teams based on recommendations using our findings, we conduct a set of counterfactual analyses in which we change the membership of the temporary teams at Metro ED and assess the potential impact of the interventions on team performance. This allows us to address our third research question: Given our empirical findings, what should be the policy used to staff temporary teams of physicians and nurses in the ED, and what is the potential impact of implementing such a policy?

Our results so far have shown that (a) there is a significant positive impact of team familiarity on performance, (b) there is not a meaningful impact of the physician's prior partner count on performance, (c) there is a significant negative impact of RN's prior partner count on performance, and (d) the magnitudes of these effects do not differ based on whether the experiences were accumulated longer ago as opposed to more recently. As such, we propose that ED managers should staff physician-nurse dyads with the objective of maximizing team familiarity, which will simultaneously minimize RN's prior partner counts because of the way the two measures are defined. In addition, team familiarity should be measured using a relatively long lookback window (e.g., 12 months). Nevertheless, for completeness, we conduct our counterfactual analyses for each of the 12 lookback windows previously considered in our empirical analyses.

7.1. Setup of Counterfactual Analyses

To obtain an order-of-magnitude estimate of the potential impact of implementing such a policy, we use the following nurse shift swap approach.[4] Suppose a given week *t* has a total of *N* nurse shifts. For each shift *i*, we swap its nurse with the nurse for another shift *j* (among the *N* shifts) if and only if (a) the average level of team familiarity of either shift *i* or shift *j* increases after the swap, and (b) both shift *i* and

[4] In section C of the Online Supplement, we use an example to provide a more in-depth illustration of how we implement this procedure.

shift *j* have an equal or higher average level of team familiarity after the swap. If there is more than 1 shift *j* that meets the criterion for a swap, we select the one that increases the average levels of team familiarity the most.

In this nurse shift swap approach, we account for scheduling constraints by keeping the physician staffing as observed in the data and by restricting nurse shift swaps only to within a given week. By doing so, we are able to preserve the number of shifts that each nurse works within each week, which we believe is an important constraint as nurses typically maintain a regular weekly schedule (e.g., 3 shifts per week).

7.2. Results of Counterfactual Analyses

Table 6 provides the results of the counterfactual analyses. By swapping nurse shifts within the same week to maximize team familiarity, the average level of team familiarity given a 1-month lookback window increases from 3.1[5] to 23.5. Meanwhile, the average physician's prior partner count given a 1-month lookback window decreases from 26.4 to 20.6, and the average nurse prior partner count given a 1-week lookback window decreases from 18.2 to 13.4. When employing a 12-month lookback window, the average level of team familiarity increases from 28.1 to 175.0, while the average physician's prior partner count decreases from 71.3 to 70.0 and the average nurse's prior partner count decreases from 43.9 to 43.1, respectively.

---------- Insert Table 6 About Here ----------

We predict the counterfactual time to disposition for each individual patient using the counterfactual values of team familiarity, prior partner counts, nurse fixed effect, and nurse experience that result from the nurse swaps. For all other variables, we use the observed values. We find that the reduction in the average time to disposition ranges from 7.0 minutes (when using a 1-month lookback window) to 13.4 minutes (when using a 11-month lookback window). This corresponds to a 3.5-percent decrease and a 6.8-percent decrease, respectively. In general, we find that using a longer lookback window yields more substantial reductions in the predicted time to disposition as a result of the nurse swaps. This suggests that there are benefits to employing longer lookback windows for the measurement of team familiarity and partner variety.

In the last column of Table 6, we show the number of additional patients per year (and the corresponding percent increase) that Metro ED could treat using the estimated savings in time to disposition from the existing encounters. We find that Metro ED could handle anywhere from 2,000

[5] Table B.2 shows levels of team familiarity, physician's prior partner count, and nurse's prior partner count given no swaps (i.e., at baseline) given each of the 12 lookback windows.

(when using a 1-month lookback window) to 4,000 (when using a 11-month lookback window) additional patient encounters.[6] This corresponds to a 3.7-percent or 7.3-percent increase in patient volume.

The results of our counterfactual analyses demonstrate that staffing ED physician-nurse teams to maximize team familiarity can have significant benefits, especially when employing lookback windows that are 8 months or longer. We note that the size of this benefit compares favorably to other effects that have been identified as managerial levers, such as workload effects (KC and Terwiesch 2009) and work design effects (Ibanez et al. 2018).

8. Discussion and Conclusions

In this paper, we investigate how ED managers can improve the staffing of temporary teams of physicians and nurses to facilitate team performance. We leverage the random team assignment and round-robin task assignment at Metro ED to cleanly identify the performance effects of team familiarity and partner variety as they are accumulated over varying lookback windows ranging from 1 month to 12 months. We find that while there is a significant positive impact of team familiarity on performance, there is a significant negative impact of nurses' partner variety on performance and there is no meaningful impact of physicians' partner variety on performance. Furthermore, we find that the magnitudes of these effects do not differ based on whether the experiences were accumulated longer ago as opposed to more recently. Hence, we recommend that ED managers staff physician-nurse dyads with the objective of maximizing long-term team familiarity. Using counterfactual analyses, we illustrate that implementing such a staffing policy would result in enough time savings to allow Metro ED to increase their patient census by approximately 7 percent (which corresponds to approximately 4,000 patients per year).

As with many field-based studies, a limitation of this work is that our investigation was limited to a single setting—a high-volume hospital ED. Future work should explore the extent to which these findings hold not only in other EDs, but also in a variety of other service settings. We were also constrained by some limitations in data availability, such as more detailed information about patient conditions. That said, this limitation should not bias our estimations, given the round-robin assignment of patients to physician-nurse dyads. Another limitation is that we restricted our focus to physician-nurse dyads rather than also considering larger teams that may include more individuals across more disciplines. Though beyond the scope of this paper, this would be a fruitful avenue for future research.

[6] For this estimation, we use the following approach: average change in time to disposition x total encounters over 2 years / 2 years) / average predicted time to disposition. Given a 1-month lookback window, (7.0 minutes x (111,491 encounters over 2 years / 2 years) / 190.6 minutes = 2048.4; given a 11-month lookback window, (13.4 minutes x (111,491 encounters over 2 years / 2 years) / 184.2 minutes = 4044.8.

This paper makes several contributions to the literature on the performance of temporary and cross-disciplinary teams. First, we show that the effects of partner variety on team performance may be setting specific and role specific. Our findings suggest that, in the context of cross-disciplinary temporary teams, partner variety tends to hinder rather than bolster team performance. This is in contrast to the findings by Aksin et al. (2018), where the authors document a positive impact of partner variety on team performance in the context of solo-disciplinary temporary teams. While beyond the scope of this paper, future work should examine the mechanisms underlying the different effects emerging from solo-disciplinary versus cross-disciplinary teams.

Second, we show that despite the absence of substantial forgetting effects, the definition of the lookback window plays an important role in the estimation of the effects of team familiarity and partner variety. Prior work seldom discusses whether and how the lookback window was defined in measuring various experience-related team characteristics, such as team familiarity and partner variety. In fact, much of the existing literature to date uses a rolling lookback window that is effectively shorter for observations occurring closer to the beginning of the data set and longer for observations occurring later in the data set. In future work, we recommend that researchers use pre-determined and consistent lookback windows across all observations when measuring these types of team characteristics.

Most importantly, this paper shows how managers can improve operational performance within the constraints of a *fixed* staffing level. Prior work in operations management has focused primarily on the question of setting *optimal* staffing levels rather than *optimizing within* a fixed staffing level (Atlason et al. 2008, Green et al. 2013, Kamalahmadi et al. 2019). Some work has also examined how the quality of staff and their leadership orientation relate to optimal staffing levels in health care settings (Kuntz and Scholtes 2013, Kuntz and Sülz 2013). To our knowledge, limited work exists on how managers might leverage team characteristics like team familiarity and partner variety to optimize team staffing given a fixed staffing level.

In today's world, a growing number of organizations—within and beyond health care settings—are utilizing temporary teams that are comprised of cross-disciplinary team members. These teams are essential to the execution of the complex service operations that these firms carry out daily. Given the many automated scheduling systems available for use, some of which are home-grown and relatively easy to program, our recommendation that organizations prioritize the maximization of team familiarity when assigning individuals to temporary teams is a feasible one to implement.

References

Agha L, Ericson KM, Geissler KH, Rebitzer JB (2018) *Team Formation and Performance: Evidence from Healthcare Referral Networks*

Aksin Z, Deo S, Jónasson JO, Ramdas K (2018) *Learning from Many: Partner Diversity and Team Familiarity in Fluid Teams*

Alexander JA, Lichtenstein R, Jinnett K, Wells R, Zazzali J, Liu D (2005) Cross-functional team processes and patient functional improvement. *Health Serv. Res.* 40(5):1335–1355.

Allon G, Deo S, Lin W (2013) The impact of size and occupancy of hospital on the extent of ambulance diversion: Theory and evidence. *Oper. Res.* 61(3):544–562.

Anand KS, Pac MF, Veeraraghavan SK (2011) Quality-Speed Conundrum: Trade-offs in Customer-Intensive Services. *Manage. Sci.* 57(1):40–56.

Ancona DG, Okhuysen GA, Perlow LA (2001) Taking Time to Integrate Temporal Research. *Acad. Manag. Rev.* 26(4):512–529.

Anderson D, Gao GG, Golden B (2014) Life Is All about Timing: An Examination of Differences in Treatment Quality for Trauma Patients Based on Hospital Arrival Time. *Prod. Oper. Manag.* 23(12):2178–2190.

Anderson Jr. EG, Lewis K (2014) A Dynamic Model of Individual and Collective Learning Amid Disruption. *Organ. Sci.* 25(2):356–376.

Atlason J, Epelman MA, Henderson SG (2008) Optimizing Call Center Staffing Using Simulation and Analytic Center Cutting-Plane Methods. *Manage. Sci.* 54(2):295–309.

Avgerinos E, Gokpinar B (2017) Team Familiarity and Productivity in Cardiac Surgery Operations: The Effect of Dispersion, Bottlenecks, and Task Complexity. *Manuf. Serv. Oper. Manag.* 19(1):19–35.

Balkundi P, Harrison DA (2006) Ties, Leaders, and Time in Teams: Strong Inference about Network Structure's Effects on Team Viability and Performance. *Acad. Manag. J.* 49(1):49–68.

Bartel AP, Chan CW, Kim SH (2019) Should Hospitals Keep Their Patients Longer? The Role of Inpatient Care in Reducing Postdischarge Mortality. *Manage. Sci.*:mnsc.2019.3325.

Batt RJ, KC DS, Staats BR, Patterson BW (2019) The Effects of Discrete Work Shifts on a Nonterminating Service System. *Prod. Oper. Manag.* 28(6):1528–1544.

Benkler Y (2017) Peer production, the commons, and the future of the firm. *Strateg. Organ.* 15(2):264–274.

Bernstein SL, Aronsky D, Duseja R, Epstein S, Handel D, Hwang U, McCarthy M, et al. (2009) The effect of emergency department crowding on clinically oriented outcomes. *Acad. Emerg. Med.* 16(1):1–10.

Berry Jaeker JA, Tucker AL (2017) Past the Point of Speeding Up: The Negative Effects of Workload Saturation on Efficiency and Patient Severity. *Manage. Sci.* 63(4):1042–1062.

Cappelli P (1999) Career jobs are dead. *Calif. Manage. Rev.* 42(1):146–167.

Cattani G, Ferriani S, Mariani MM, Mengoli S (2013) Tackling the "Galacticos" effect: team familiarity and the performance of star-studded projects. *Ind. Corp. Chang.* 22(6):1629–1662.

Chan DC (2016) Teamwork and moral hazard: evidence from the emergency department. *J. Polit. Econ.* 124(3):734–770.

Contractor N (2013) Some assembly required: leveraging Web science to understand and enable team assembly. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 371(1987):20120385.

Crisp CB, Jarvenpaa SL (2013) Swift Trust in Global Virtual Teams. *J. Pers. Psychol.* 12(1):45–56.

Cummings JN (2004) Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization. *Manage. Sci.* 50(3):352–364.

Cummings JN, Cross R (2003) Structural properties of work groups and their consequences for performance. *Soc. Networks* 25(3):197–210.

Deo S, Jain A (2019) Slow First, Fast Later: Temporal Speed-Up in Service Episodes of Finite Duration. *Prod. Oper. Manag.* 28(5):1061–1081.

Dougherty D (1992) Interpretive Barriers to Successful Product Innovation in Large Firms. *Organ. Sci.*

3(2):179–202.
Dugan RE, Gabriel KJ (2013) “Special Forces” Innovation: How DARPA Attacks Problems. *Harv. Bus. Rev.* 91(10):74–84.
Edmondson AC (1999) Psychological Safety and Learning Behavior in Work Teams. *Adm. Sci. Q.* 44(2):350–383.
Edmondson AC, Dillon JR, Roloff KS (2007) Three Perspectives on Team Learning. *Acad. Manag. Ann.* 1(1):269–314.
Edmondson AC, Nembhard IM (2009) Product Development and Learning in Project Teams: The Challenges Are the Benefits. *J. Prod. Innov. Manag.* 26(2):123–138.
Espinosa JA, Slaughter SA, Kraut RE, Herbsleb JD (2007) Familiarity, Complexity, and Team Performance in Geographically Distributed Software Development. *Organ. Sci.* 18(4):613–630.
Froehle CM, White DL (2014) Interruption and Forgetting in Knowledge-Intensive Service Environments. *Prod. Oper. Manag.* 23(4):704–722.
GAO (2009) *Hospital emergency departments: Crowding continues to occur, and some patients wait longer than recommended time frames*
Gardner HK, Valentine M (2015) Collaboration among Highly Autonomous Professionals: Costs, Benefits, and Future Research Directions. *Adv. Gr. Process.* (Emerald Group Publishing Limited), 209–242.
Gilboy N, Tanabe T, Travers D, Rosenau A (2011) *Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4, Implementation Handbook* (Agency for Healthcare Research and Quality, Rockville, MD, MD).
Green L V, Savin S, Savva N (2013) “Nursevendor Problem”: Personnel Staffing in the Presence of Endogenous Absenteeism. *Manage. Sci.* 59(10):2237–2256.
Greer LL, Chu C (2020) Power struggles: when and why the benefits of power for individuals paradoxically harm groups. *Curr. Opin. Psychol.* 33:162–166.
Greer LL, de Jong BA, Schouten ME, Dannals JE (2018) Why and when hierarchy impacts team effectiveness: A meta-analytic integration. *J. Appl. Psychol.* 103(6):591–613.
Gupta D, Denton B (2008) Appointment scheduling in health care: Challenges and opportunities. *IIE Trans.* 40(9):800–819.
Hair JF, Anderson RE, Tatham RL, Black WC (1998) *Multivariate data analysis* 5th ed. (Prentice Hall, Upper Saddle River, NJ).
Han CY, Chen LC, Barnard A, Lin CC, Hsiao YC, Liu HE, Chang W (2015) Early Revisit to the Emergency Department: An Integrative Review. *J. Emerg. Nurs.* 41(4):285–295.
Harrison DA, Klein KJ (2007) What’s the difference? Diversity constructs as separation, variety, or disparity in organizations. *Acad. Manag. Rev.* 32(4):1199–1228.
Harrison DA, Mohammed S, McGrath JE, Florey AT, Vanderstoep SW (2003) Time matters in team performance: Effects of member familiarity, entrainment and task discontinuity on speed and quality. *Pers. Psychol.* 56(3):633–669.
Hoffman R, Casnocha B, Yeh C (2013) Tours of duty: the new employer-employee contract. *Harv. Bus. Rev.* 91(6):48–58.
Huckman RS, Staats BR (2011) Fluid Tasks and Fluid Teams: The Impact of Diversity in Experience and Team Familiarity on Team Performance. *Manuf. Serv. Oper. Manag.* 13(3):310–328.
Huckman RS, Staats BR, Upton DM (2009) Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services. *Manage. Sci.* 55(1):85–100.
Ibanez MR, Clark JR, Huckman RS, Staats BR (2018) Discretionary Task Ordering: Queue Management in Radiological Services. *Manage. Sci.* 64(9):4389–4407.
Institute of Medicine (2001) Applying Evidence to Health Care Delivery. *Crossing Qual. Chasm A New Heal. Syst. 21st Century*.
Jehn KA (1995) A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. *Adm. Sci. Q.* 40(2):256–282.
Kamalahmadi M, Yu Q, Zhou YP (2019) *Call to Duty: Just-in-Time Scheduling in a Restaurant Chain*

Kane AA, Argote L, Levine JM (2005) Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organ. Behav. Hum. Decis. Process.* 96(1):56–71.
KC DS (2014) Does Multitasking Improve Performance? Evidence from the Emergency Department. *Manuf. Serv. Oper. Manag.* 16(2):168–183.
KC DS (2019) Heuristic Thinking in Patient Care. *Manage. Sci.*
KC DS, Scholtes S, Terwiesch C (2020) Empirical research in healthcare operations: Past research, present understanding, and future opportunities. *Manuf. Serv. Oper. Manag.* 22(1):73–83.
KC DS, Terwiesch C (2009) Impact of Workload on Service Time and Patient Safety: An Econometric Analysis of Hospital Operations. *Manage. Sci.* 55(9):1486–1498.
Kerrissey MJ, Satterstrom P, Edmondson AC (2020) Into the fray: Adaptive approaches to studying novel teamwork forms. *Organ. Psychol. Rev.* (April 2019).
Klein KJ, Ziegert JC, Knight AP, Xiao Y (2006) Dynamic Delegation: Shared, Hierarchical, and Deindividualized Leadership in Extreme Action Teams. *Adm. Sci. Q.* 51(4):590–621.
Kuntz L, Mennicken R, Scholtes S (2015) Stress on the Ward: Evidence of Safety Tipping Points in Hospitals. *Manage. Sci.* 61(4):754–771.
Kuntz L, Scholtes S (2013) Physicians in leadership: The association between medical director involvement and staff-to-patient ratios. *Health Care Manag. Sci.* 16(2):129–138.
Kuntz L, Sülz S (2013) Treatment speed and high load in the Emergency Department—does staff quality matter? *Health Care Manag. Sci.* 16(4):366–376.
Laker LF, Froehle CM, Lindsell CJ, Ward MJ (2014) The Flex Track: Flexible Partitioning Between Low- and High-Acuity Areas of an Emergency Department. *Ann. Emerg. Med.* 64(6):591–603.
Laker LF, Froehle CM, Windeler JB, Lindsell CJ (2018) Quality and Efficiency of the Clinical Decision-Making Process: Information Overload and Emphasis Framing. *Prod. Oper. Manag.* 27(12):2213–2225.
Lemieux-Charles L, McGuire WL (2006) What do we know about health care team effectiveness? A review of the literature. *Med. Care Res. Rev.* 63(3):263–300.
Leonard-Barton D (1992) Core capabilities and core rigidities: A paradox in managing new product development. *Strateg. Manag. J.* 13(S1):111–125.
Lerman B, Kobernick MS (1987) Return visits to the emergency department. *J. Emerg. Med.* 5(5):359–362.
Lewis K (2004) Knowledge and Performance in Knowledge-Worker Teams: A Longitudinal Study of Transactive Memory Systems. *Manage. Sci.* 50(11):1519–1533.
Liu N, Ziya S, Kulkarni VG (2010) Dynamic Scheduling of Outpatient Appointments Under Patient No-Shows and Cancellations. *Manuf. Serv. Oper. Manag.* 12(2):347–364.
Mathieu JE, Heffner TS, Goodwin GF, Salas E, Cannon-Bowers JA (2000) The influence of shared mental models on team process and performance. *J. Appl. Psychol.* 85(2):273–283.
Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, Berns SD (2002) Error Reduction and Performance Improvement in the Emergency Department through Formal Teamwork Training: Evaluation Results of the MedTeams Project. *Health Serv. Res.* 37(6):1553–1581.
Mortensen M, Haas MR (2018) Rethinking teams: From bounded membership to dynamic participation. *Organ. Sci.* 29(2):341–355.
O'Leary MB, Mortensen M, Woolley AW (2011) Multiple Team Membership: a Theoretical Model of its Effects on Productivity and Learning for Individuals and Teams. *Acad. Manag. Rev.* 36(3):461–478.
Paiva EL, Roth A V., Fensterseifer JE (2008) Organizational knowledge and the manufacturing strategy process: A resource-based view analysis. *J. Oper. Manag.* 26(1):115–132.
Patterson PD, Pfeiffer AJ, Lave JR, Weaver MD, Abebe K, Krackhardt D, Arnold RM, Yealy DM (2015) How familiar are clinician teammates in the emergency department? *Emerg. Med. J.* 32(4):258–262.
Pentland A "Sandy" (2012) The New Science of Building Great Teams. *Harvard Business Review* (April).
Pitts SR, Pines JM, Handrigan MT, Kellermann AL (2012) National trends in emergency department occupancy, 2001 to 2008: Effect of inpatient admissions versus emergency department practice

intensity. *Ann. Emerg. Med.* 60(6):679-686.e3.
Pourmand A, Lucas R, Pines J, Shokoohi H, Yadav K (2013) Bedside Teaching on Time to Disposition Improves Length of Stay for Critically-ill Emergency Departments Patients. *West. J. Emerg. Med.* 14(2):137–140.
Ramachandran K, Tereyagoglu N, Unal M (2018) *Help or Hindrance? The Role of Familiarity in Collaborative Product Development*
Ramdas K, Saleh K, Stern S, Liu H (2018) Variety and experience: Learning and forgetting in the use of surgical devices. *Manage. Sci.* 64(6):2590–2608.
Reagans R, Argote L, Brooks D (2005) Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together. *Manage. Sci.* 51(6):869–881.
Reagans R, Zuckerman EW (2001) Networks, Diversity, and Productivity. *Organ. Sci.* 12(4):502–517.
Reagans R, Zuckerman EW, McEvily B (2004) How to Make the Team: Social Networks vs . as Criteria Demography for Designing Effective Teams. *Adm. Sci. Q.* 49(1):101–133.
Ren Y, Argote L (2011) Transactive Memory Systems 1985–2010: An Integrative Framework of Key Dimensions, Antecedents, and Consequences. *Acad. Manag. Ann.* 5(1):189–229.
Retelny D, Robaszkiewicz S, To A, Lasecki WS, Patel J, Rahmati N, Doshi T, Valentine M, Bernstein MS (2014) Expert crowdsourcing with flash teams. *Proc. 27th Annu. ACM Symp. User interface Softw. Technol. - UIST '14*. (ACM Press, New York, New York, USA), 75–85.
Riedl C, Woolley AW (2017) Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Emergent Collaboration to Crowd-Based Problem Solving Performance. *Acad. Manag. Discov.* 3(4):382–403.
Rising KL, Victor TW, Hollander JE, Carr BG (2014) Patient Returns to the Emergency Department: The Time-to-return Curve. *Acad. Emerg. Med.* 21(8):864–871.
Roth A, Tucker AL, Venkataraman S, Chilingerian J (2019) Being on the Productivity Frontier: Identifying "Triple Aim Performance" Hospitals. *Prod. Oper. Manag.* 28(9):2165–2183.
Sackett E, Cummings JN (2018) When team members perceive task interdependence differently: Exploring centrality asymmetry and team success. *Gr. Dyn. Theory, Res. Pract.* 22(1):16–31.
Saghafian S, Hopp WJ, Van Oyen MP, Desmond JS, Kronick SL (2012) Patient Streaming as a Mechanism for Improving Responsiveness in Emergency Departments. *Oper. Res.* 60(5):1080–1097.
Saghafian S, Hopp WJ, Oyen MP Van, Desmond JS, Kronick SL (2014) Complexity-Augmented Triage: A Tool for Improving Patient Safety and Operational Efficiency. *Manuf. Serv. Oper. Manag.* 16(3):329–345.
Salehi N, McCabe A, Valentine M, Bernstein M (2017) Huddler: Convening Stable and Familiar Crowd Teams Despite Unpredictable Availability. *Proc. 2017 ACM Conf. Comput. Support. Coop. Work Soc. Comput.* (ACM), 1700–1713.
Siemsen E, Roth A V., Balasubramanian S (2008) How motivation, opportunity, and ability drive knowledge sharing: The constraining-factor model. *J. Oper. Manag.* 26(3):426–445.
Siemsen E, Roth A V., Balasubramanian S, Anand G (2009) The Influence of Psychological Safety and Confidence in Knowledge on Employee Knowledge Sharing. *Manuf. Serv. Oper. Manag.* 11(3):429–447.
Singer AJ, Thode HC, Viccellio P, Pines JM (2011) The association between length of emergency department boarding and mortality. *Acad. Emerg. Med.* 18(12):1324–1329.
Song H, Andreyeva E, David G (2020) *Time is the Wisest Counselor of All: The Value of Provider-Patient Engagement Length in Home Health Care*
Song H, Tucker AL, Murrell KL (2015) The Diseconomies of Queue Pooling: An Empirical Investigation of Emergency Department Length of Stay. *Manage. Sci.* 61(12):3032–3053.
Song H, Tucker AL, Murrell KL, Vinson DR (2018) Closing the Productivity Gap: Improving Worker Productivity Through Public Relative Performance Feedback and Validation of Best Practices. *Manage. Sci.* 64(6):2628–2649.

Song H, Veeraraghavan S (2018) Quality of Care. Dai T, Tayur S, eds. *Handb. Healthc. Anal.* (John Wiley & Sons, Inc., Hoboken, NJ), 79–108.
Staats BR (2012) Unpacking team familiarity: The effects of geographic location and hierarchical role. *Prod. Oper. Manag.* 21(3):619–635.
Staats BR, Upton DM (2011) Lean Knowledge Work. *Harv. Bus. Rev.* 89(10):100–110.
Tan TF, Netessine S (2014) When Does the Devil Make Work? An Empirical Study of the Impact of Workload on Worker Productivity. *Manage. Sci.* 60(6):1574–1593.
Tannenbaum SI, Mathieu JE, Salas E, Cohen D (2012) Teams Are Changing: Are Research and Practice Evolving Fast Enough? *Ind. Organ. Psychol.* 5:2–24.
Thompson L, Fine GA (1999) Socially Shared Cognition, Affect, and Behavior: A Review and Integration. *Personal. Soc. Psychol. Rev.* 3(4):278–302.
Valentine M (2018) When Equity Seems Unfair: The Role of Justice Enforceability in Temporary Team Coordination. *Acad. Manag. J.* 61(6):2081–2105.
Valentine MA, Edmondson AC (2015) Team Scaffolds: How Mesolevel Structures Enable Role-Based Coordination in Temporary Groups. *Organ. Sci.* 26(2):405–422.
Valentine MA, Retelny D, To A, Rahmati N, Doshi T, Bernstein MS (2017) Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds As Organizations. *Proc. 2017 CHI Conf. Hum. Factors Comput. Syst.* (ACM), 3523–3537.
Wageman R, Gardner H, Mortensen M (2012) The changing ecology of teams: New directions for teams research. *J. Organ. Behav.* 33(3):301–315.
Ward MJ, Ferrand YB, Laker LF, Froehle CM, Vogus TJ, Dittus RS, Kripalani S, Pines JM (2015) The Nature and Necessity of Operational Flexibility in the Emergency Department. *Ann. Emerg. Med.* 65(2):156–161.
Weinberg DB, Cooney-Miner D, Perloff JN, Babington L, Avgar AC (2011) Building Collaborative Capacity Promoting Interdisciplinary Teamwork in the Absence of Formal Teams. *Med. Care* 49(8):716–723.
White DL, Froehle CM, Klassen KJ (2011) The Effect of Integrated Scheduling and Capacity Policies on Clinical Efficiency. *Prod. Oper. Manag.* 20(3):442–455.
De Wit FRC, Greer LL, Jehn KA (2012) The paradox of intragroup conflict: A meta-analysis. *J. Appl. Psychol.* 97(2):360–390.
Wooldridge JM (2012) *Introductory Econometrics: A Modern Approach* 5th ed. (South-Western Cengage Learning, Mason, OH).
Woolley AW, Aggarwal I, Malone TW (2015) Collective Intelligence and Group Performance. *Curr. Dir. Psychol. Sci.* 24(6):420–424.
Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science (80-.).* 330(6004):686–688.
Xu R, Carty MJ, Orgill DP, Lipsitz SR, Duclos A (2013) The Teaming Curve A Longitudinal Study of the Influence of Surgical Team Familiarity on Operative Time. *Ann. Surg.* 258(6):953–957.
Yuan CT, Pronovost PJ, Marsteller JA (2018) Harnessing the Power of Peer Influence to Improve Quality. *Am. J. Med. Qual.* 33(5):549–551.
Zacharias C, Pinedo M (2014) Appointment Scheduling with No-Shows and Overbooking. *Prod. Oper. Manag.* 23(5):788–801.
Zhang ZX, Hempel PS, Han YL, Tjosvold D (2007) Transactive memory system links work team characteristics and performance. *J. Appl. Psychol.* 92(6):1722–1730.