

Team Familiarity and Productivity in Cardiac Surgery Operations: The Effect of Dispersion, Bottlenecks and Task Complexity

Emmanouil Avgerinos

IE Business School, Maria de Molina 12, 28006 Madrid, Spain, emmanouil.avgerinos@ie.edu

Bilal Gokpinar

UCL School of Management, University College London, 1 Canada Square, London E14 5AB, United Kingdom, b.gokpinar@ucl.ac.uk

Fluid teams are commonly used by a variety of organizations to perform similar and repetitive yet highly critical and knowledge-intensive tasks. Such teams operate for a limited time, after which they dissolve and some of their members may work together again as part of another team. Using a granular dataset of 6,206 cardiac surgeries from a private hospital in Europe over seven years, our study offers a new and detailed account of how team familiarity (i.e., shared work experience) influences team productivity. We highlight the role of nuanced team composition dynamics beyond average team familiarity. We observe that teams with high dispersion of pairwise familiarity exhibit lower team productivity, and the existence of a “bottleneck-pair” may significantly hinder overall knowledge transfer capability, thus, productivity of fluid teams. In addition, we find that the higher the percentage of familiarity gained from complex tasks, the higher the productivity of the team. Finally, our results suggest that the positive effect of average team familiarity on productivity is enhanced when performing more complicated tasks. Our study provides new operational insights to improve productivity of fluid teams with better team composition strategies.

Key words: Empirical Research, Health Care Management, OM-Organizational Behavior Interface

1. Introduction

The modus operandi of many of today’s organizations in knowledge-intensive environments is to employ fluid teams (Bushe and Chu 2011, Huckman et al. 2009, Edmondson and Nembhard 2009). One can observe fluid teams in a variety of settings, from scientific research collaborations and management consulting teams to surgical teams, flight crews, and product development teams. Fluid teams have a number of unique characteristics: They have no permanent memberships, they operate as a team only for a limited time or for a specific task, and they normally have clear hierarchies, roles, and task responsibilities. After their job is over, these teams dissolve, but many members may work again with each other on a new task as part of another team (Cohen and Bailey 1997, Paletz and Schunn 2011).

A key management challenge that is particularly pertinent to fluid team operations is team composition. If individuals operate as part of a team only temporarily, how should they be assigned to various teams to make sure that these teams can perform their tasks more efficiently and hence achieve higher productivity over time? Our study seeks to address this question by focusing on a major element of team composition: team familiarity (i.e., shared work experience among team members).

Previous research has identified team familiarity as an important performance driver for teams with a primary focus on the role of average team familiarity (e.g., average of pairs' familiarity). This line of work reported mostly beneficial effects (Edmondson 1999, Reagans et al. 2005, Huckman et al. 2009, Huckman and Staats 2011) with the exception of two studies which suggested diminishing returns in the long run (Katz 1982, Berman et al. 2002).

Despite these studies, several important questions regarding the effect of team familiarity on team productivity have remained unaddressed. First, previous research has mainly focused on the role of the first moment of familiarity distribution among team members (i.e., average familiarity) and has not accounted for two important distributional characteristics that go beyond team averages: (i) dispersion of familiarity among team members (e.g., the second moment), and (ii) existence of a bottleneck pair in the team (e.g., similar to the third moment capturing the skewness of the lowest familiar pair). Fluid teams are composed of members who have various levels of shared work experience in different pairs, and hence teams' familiarity distributions might be substantially different despite having similar average familiarity levels. Indeed, a key characteristic of a distribution such as dispersion constitutes an inherent property of every team-level mechanism (De Jong and Dirks 2012). As a result, any model that solely focuses on team averages and neglects important distributional properties can lead to biased and ambiguous results for the examined phenomenon (Cole et al. 2011). Consequently, we first examine how team familiarity dispersion and the existence of bottleneck pairs might influence productivity of fluid teams. Team familiarity dispersion and bottleneck enrich our understanding of familiarity because they capture distinct team level characteristics that are undefined at the individual or pair level, and they only emerge at the team level. Next, inspired by the concept of "bottleneck" in traditional operations settings which limits a system's capacity; we define and identify an analogous construct "bottleneck-pair" for fluid teams. We argue that the existence of a pair with very low shared experience-compared to the average familiarity of the team, may act as a bottleneck for the team. It is worth noting that although the most famous example of a bottleneck in the operations management literature is a person (i.e., Herbie in the hiking track) (Goldratt and Cox 1984), there is little work that relates system bottleneck with persons and teams (Boudreau et al. 2003).

Second, although previous studies have examined various contingent factors about the effect of team familiarity on team performance such as geographic dispersion and team size (Espinosa et al. 2007), interpersonal diversity in terms of customer experience (Huckman and Staats 2011), and hierarchical roles (Staats 2012), this line of research has not examined the role of task type (i.e., complexity) on the relationship between team familiarity and productivity. On this front, we examine two important questions. First, does familiarity gained from a higher ratio of complex tasks result in

higher productivity? Second, how does complexity of the focal task moderate the relationship between team familiarity and productivity? These questions are also important from a practical viewpoint because fluid teams essentially perform related but not identical jobs over time which may significantly differ in terms of complexity (e.g., consider flight crews or surgical teams where each surgery or flight is different in complexity depending on patient or weather conditions).

Overall, in this study, we focus on four questions related to familiarity in fluid teams. We first look at the distribution of familiarity among team members by examining (i) the effect of team familiarity dispersion, and (ii) the existence of a bottleneck pair on productivity. We then focus on task complexity and investigate (iii) whether familiarity gained from a higher percentage of complex tasks in the past provide higher productivity benefits, and (iv) whether a more complex focal task is more conducive for the positive effects of team familiarity on productivity.

We test our hypotheses using a unique dataset of 6,206 cardiac surgeries from the cardiac unit of a private hospital in Europe over 87 months. With clear team member roles and responsibilities, limited team durations, and repeated interactions among individuals as part of different teams, our setting is ideal to investigate fluid teams (Edmondson et al. 2001). In addition, the presence of similar and repetitive, yet highly critical and complex tasks make cardiac operations a very suitable context to examine the role of team composition on productivity. Our results indicate that teams with high dispersion of pairwise familiarity exhibit lower team productivity (i.e., higher surgery duration), and the existence of a “bottleneck-pair” may significantly reduce productivity. In addition, we observe that the higher the percentage of familiarity gained from complex tasks, the higher the productivity of the team. Finally, our results suggest that the positive effect of average team familiarity on productivity is amplified when performing more complex tasks.

Our paper offers important contributions to operations management literature. Identifying team composition as a major apparatus with which managers can influence productivity; our study provides insights on effective management of fluid team operations. By examining nuanced factors beyond average team familiarity; our study unravels when and how team familiarity influences productivity of fluid teams. An important focus in the operations management literature has been to increase productivity at the individual, team and organizational level, though for mostly repetitive and routine tasks in contexts such as manufacturing or blue-collar work (Hopp and Spearman 2000, Hopp et al. 2009). A natural extension and focus for this line of work is to study productivity in more knowledge intensive and more complex settings (Hopp et al. 2009). As Peter Drucker famously noted (1999, p. 79) “the most important, and indeed the truly unique, contribution of management in the 20th century was the 50-fold increase in the productivity of the manual worker in manufacturing. The most important

contribution management needs to make in the 21st century is similarly to increase the productivity of knowledge work and knowledge workers.”

In the next section we present the existing literature and develop our hypotheses. We then present our data and results, robustness checks, and finally discuss our findings and conclusions.

2. Literature Review and Theory Development

Team familiarity (i.e., prior shared work experience) and its performance implications have been studied in several literatures. A number of theoretical studies have suggested positive effects of past collaborations (i.e. team familiarity) on team performance by highlighting the role of transactive memory system (Wegner 1986), collective mind (Weick and Roberts 1993), team human capital (Chillemi and Gui 1997), team psychological safety (Edmondson 1999), among others. In healthcare research, focusing on cardiac surgery team members, Fleming et al. (2006) suggested that limited shared experience may lead to breakdowns in communication, decision making and leadership which in turn may threaten surgical safety. Using data from pediatric, cardiac, and orthopedic operations, Catchpole et al. (2007) also showed that shared experience and effective teamwork during surgery can prevent minor problems from escalating and becoming serious threats for patients. Finally, examining general/vascular surgery services, Davenport et al. (2007) suggested that past collaborations can help decrease patient morbidity. Broader management literature has generally highlighted beneficial effects of team familiarity on productivity in various empirical settings including software teams (Faraj and Sproull 2000, Huckman et al. 2009) and surgery teams (Reagans et al. 2005); although some scholars have highlighted diminishing returns of team familiarity (Berman et al. 2002, Katz 1982).

Previous studies have also explored more subtle effects of team familiarity. Espinosa et al. (2007) showed that in software development, team size and geographic dispersion interact positively with team familiarity resulting in increased team productivity. Huckman and Staats (2011) found that the interaction between team familiarity and interpersonal diversity in terms of customer experience results in significant improvement in effort and schedule deviation. Also in software teams, Staats (2012) observed that horizontal familiarity increase quality whereas hierarchical familiarity promotes productivity.

Familiarity Dispersion

While average team familiarity is an appropriate construct to characterize average levels of prior shared work experience in a team, it may not give us the complete picture with regards to team composition and familiarity dynamics in fluid teams. In such settings in which teams are not stable, there may be a high degree of variation both across and within teams in terms of team members' familiarity between

each other. Moreover, two teams with the same levels of average team familiarity may differ substantially in their composition dynamics (e.g., one team with all members moderately familiar with each other, another team in which some pairs highly familiar and other pairs barely familiar with each other).

In fact, previous research has pointed out the importance of including dispersion (e.g., diversity) while examining team-level effects (Harrison and Klein 2007, Jackson et al. 2003). Klein and Kozlowski (2000) suggested that most group constructs come from a combination of mean and dispersion mechanisms. Indeed, mean and dispersion have been treated as equal properties for the formation of team-level constructs such as trust and peer monitoring (De Jong and Dirks 2012), leader-member exchange (Liao et al. 2010) and satisfaction (Dineen et al. 2007). Building on this literature, we argue that team familiarity dispersion may hinder team productivity through two main mechanisms.

First, while common experience makes individuals aware of each other's expertise and therefore leads to more efficient task allocation and knowledge sharing (Faraj and Sproull 2000, Lewis 2003, Reagans et al. 2005), when similar levels of familiarity are not shared by all individuals within the team, they will face difficulties in knowledge sharing and effective integration. This is due to different beliefs and lack of past collaboration among some members, which will in turn lower productivity (Gardner et al. 2012). Moreover, when relational resources are concentrated within parts of the same team, individuals will develop stronger bonds only with some team members, creating distance from others (Hornsey and Hogg 2000, Van Knippenberg et al. 2004). This uneven distribution may lead to opposition in task-related goals (Bezrukova et al. 2007), and conflict and distrust among team members (Choi and Sy 2010). Consequently, familiarity dispersion will result in lower productivity because of the unevenly distributed relational resources within the team.

Second, trust among co-workers has been recognized as an important performance driver for healthcare personnel (Cook 2013), and its significance has been highlighted for individuals (e.g. nurses, Altuntas and Baykal 2010) as well as pairs in teams (such as between a surgeon and a nurse, Pullon 2008). Trust is particularly critical for performance in high pressure settings (Colquitt et al. 2011) such as cardiac operations (Tucker and Edmondson 2003, Nembhard and Edmondson 2006), and a key enabler of trust among team members is prior shared experience (Gruenfeld et al. 1996, Edmondson 1999). As a result, high familiarity dispersion indicates high trust variance and trust asymmetry within the team, which will have negative effects on team productivity due to unbalanced social exchange structures that inhibit reciprocation of resources (De Jong and Dirks 2012). Within a team, while individuals are likely to share information and adopt shared information with highly familiar and hence trusted members, they are less likely to do so with less familiar members (Kane et al. 2005, Gardner et

al. 2012). This will in turn result in inefficient knowledge sharing within the team. Individuals may refrain from seeking help and advice from less familiar members of the team who may have the right information and who may be more capable than highly familiar members (Hoffman et al. 2009, Gardner et al. 2012), resulting in decreased productivity. We therefore predict that:

Hypothesis 1: *Team familiarity dispersion has a negative impact on team productivity.*

Familiarity and Bottlenecks

In production settings, bottlenecks are highly important as they constrain the system capacity. They are also fundamental in determining throughput, cycle time, customer service, and other performance metrics (Hopp et al. 2009). Despite their significant influence on performance, and their obvious connotations in knowledge intensive environments, only a small number of studies have highlighted “bottlenecks” in white-collar settings. Taking the concept of bottleneck to team settings and employing social network analysis, Cross et al. (2001) suggested that some team members may be acting as ‘bottlenecks’ in sharing information, resulting in the whole team to operate less efficiently. Siemsen et al. (2008) used the motivation-opportunity-ability (MOA) framework to show that one of these three factors can act as a bottleneck and therefore determine the degree of knowledge sharing among members of a workgroup.

Because the smallest unit in a team that exchange information is a pair, and our conceptualization of team familiarity is based on pairwise prior shared work experience, we suggest that the existence of a pair with very low levels of shared work experience (i.e., familiarity) compared to average familiarity of the team may act as a bottleneck for the whole team’s operation, and thus reduce productivity. There are several reasons for such an effect. First, emphasizing the role of pairs in a team, Bion (1962) suggested that cohesion will develop easier between two individuals than among all group members. One of the initial consequences of prior shared work experience is the development of cohesion (Hackman 1987, Evans and Dion 1991). We argue that the existence of a bottleneck-pair in terms of team familiarity will result in a lack of basic levels of cohesion in that pair, hence constraining some of the members’ effective willingness to engage in a relationship and share information within the team, thus negatively influencing team’s productivity.

Second, groups and pairs, similar to individuals, can develop habitual routines (Gersick and Hackman 1990) and norms (Feldman 1984). When members have previous common experience, they share a priori premise about the way they should proceed in a given situation (Gersick and Hackman 1990). Bottleneck-pairs which are in significant shortage of habitual routines would need explicit

communication about the task itself (Gersick and Hackman 1990, Espinosa et al. 2007) which will in turn reduce productivity of the team.

In addition, an important consequence of team familiarity is to reduce communication errors by reducing uncertainty among them (Harrison et al. 2003). Members who form the bottleneck pair (i.e., with limited familiarity with compare to others), however, will have uncertainty about each other, producing higher anxiety (Gruenfeld et al. 1996). This may in turn reduce the fluency of individuals (Nemeth 1986), resulting in higher likelihood of communication errors. Because communication errors among surgical team members are the most common cause of operative problems (Makary et al. 2006), and such errors in one pair can be sufficient to cause operative problems and ineffectiveness for the whole surgery, we suggest that the existence of a bottleneck pair will increase the likelihood of communication errors, thus, reduce team productivity.

Finally, cardiac operations consist of a set of steps performed by pairs of the surgical team (KC and Staats 2012) in which tacit knowledge plays a critical role (Edmondson et al. 2003) in smooth performance of these tasks. Because familiarity promotes formation of tacit knowledge among individuals and pairs (Edmondson et al. 2003, Huckman and Pisano 2006) and transferring tacit knowledge is also inherently difficult due to the contextualized nature of the knowledge itself (Joshi et al. 2007), the bottleneck pair will have difficulty in information sharing. Since the entire surgery is a sequence of steps and operations, and the outcome of the whole team depends on the effort of each pair and individual in the team (Hollingshead 2001); the lack of efficient information sharing in one step (i.e, by the bottleneck pair) will have a cascading effect on other steps, which will result in reduced productivity for the whole team. As a result, we expect that:

Hypothesis 2: The existence of a bottleneck-pair has a negative impact on team productivity.

Notice that, although team familiarity dispersion and bottleneck-pair are somewhat related, we contend that they are two distinct theoretical concepts. While the existence of a bottleneck pair (i.e., a pair with very low familiarity compared to the average familiarity of the team) may indicate high dispersion, the opposite is not true. A high dispersion can be caused by one or two pairs that are highly familiar compared to the other pairs in the team. This does not necessarily mean that there will be a bottleneck pair. They are therefore distinct concepts that capture two important characteristic of familiarity distribution within a team.

Gaining Familiarity

We next consider productivity implications of the way team familiarity is gained. Members in a fluid team inevitably gain familiarity as they perform tasks and work together as part of a team, but in

developing this familiarity, the nature of tasks that they perform together could be quite different. This may lead to differentiated effects in future task productivity. Consider two similar surgical teams both performing a single operation. Members of both teams will develop familiarity within their teams, but if one team has performed a simple routine operation, whereas the other one has performed a high risk complicated operation, the nature of gained familiarity as a result of performing these operations could be quite different. Consequently, their impact on subsequent operations' productivity will be different, too.

Development of a transactive memory system, which involves a representation of who knows what in a team (Lewis 2003) enhances task performance (Wegner 1986). Furthermore, experience from more complex tasks promote more efficient development of a transactive memory resulting in better encoding, storage, retrieval, and communication of information (Brandon and Hollingshead 2004). As a result, we expect team members to develop better transactive memory systems through working on more complex tasks, leading to higher productivity in subsequent operations.

In addition, team members working on challenging, non-routine tasks which require increased coordination and communication, will relate better to each other by forming cognitive interdependence (Levine and Moreland 1999, Moreland 1999). As a result, team members will realize that the outcome of their effort does not depend solely on them but it also depends on other group members (Hollingshead 2001). This realization will bring about better information sharing and knowledge transfer among group members (Brandon and Hollingshead 2004) in subsequent tasks, and therefore, improve productivity.

Finally, cardiac surgeries are characterized by high levels of uncertainty and pressure because the stakes are high (Tucker and Edmondson 2003). Under conditions of uncertainty, working on simple tasks with low information processing requirements can be decreasingly effective (Tushman and Nadler 1978). As such, if familiarity is gained through mostly simple tasks, we expect limited productivity benefits. In contrast, if familiarity is gained by performing a high ratio of complex tasks which inherently require better information sharing and coordination (Gittel 2002), we expect higher cohesiveness and productivity in succeeding operations. Considering all of the above points, we expect:

Hypothesis 3: The higher the ratio of familiarity which is gained from complex tasks in the past, the higher the productivity of the team.

Familiarity and Task Nature

Team members who are highly familiar with each other can communicate more effectively since they can refer to the same technical terminology (Cramton 2001), identify and access expertise more effectively (Faraj and Sproull 2000), and generally know what to expect from each individual, and how

that can be used to promote team productivity. As a result, when a setting makes it harder for a team to locate specialized knowledge and access expertise, team familiarity can have an increased benefit for the team (Espinosa et al. 2007). That is, we expect the positive relationship between team familiarity and productivity to be amplified when performing more complicated tasks.

We argue that benefits of team familiarity become more prominent during more complex and uncertain tasks. Such benefits include reducing the need for explicit interactions (Espinosa et al. 2007), reducing uncertainty about team members (Harrison et al. 2003), and enabling them to figure out more efficient ways to work with each other without the need of extensive interactions (Campbell 1988). Team familiarity also enables team members to deal with variation by applying knowledge gained through their past collaborations (Banker et al. 2001, Field and Sinha 2005), and it may help develop adaptability (Van de Ven et al. 1976, Gittel 2002). We therefore expect team familiarity to have a stronger effect on productivity in environments with high uncertainty. Complex operations are characterized by higher levels of uncertainty and they have an increased probability of adverse events. We therefore expect team familiarity to have stronger effects on productivity during these operations. Especially in the operating room, where “you look at the surgeon and you know the body language and you act” (Edmondson et al. 2003, p.221), the influence of team familiarity on productivity will be more prominent during more complex tasks. Also, team members demonstrate higher cognitive alertness when performing complex tasks (Levine and Moreland 1999). This leads to better motivation and utilization of tacit knowledge, resulting in increased benefits of team familiarity on productivity. Considering these, we predict that:

Hypothesis 4: Team familiarity and task complexity interact positively in their effect on team productivity, such that the positive effect of team familiarity on team productivity is stronger when the nature of the focal task is more complex.

3. Setting, Data and Variables

We test our hypotheses using a dataset of all cardiac surgery operations in a hospital over seven years. Our setting is the cardiac unit of a 300-bed private hospital in Europe, which is property of an American non-profit organization. The hospital serves more than 2,000 patients per year and hosts around 70 cardiac surgeries per month. Our data set consists of archival data of all cardiac surgeries during the period from 01/01/2004 to 03/31/2011 and includes 6,206 operations. For each surgery, the dataset includes, among others, information about the duration of the surgery, specific surgery type, members of the surgical team and patient characteristics.

Each surgery team has three to eight members including the Lead Surgeon, zero to four Assistant Surgeons, the Anesthesiologist, the Perfusionist (a technician who runs the heart-lung machine) and zero to three Scrub Nurses. Our dataset consists of 115 individuals in total: 44 surgeons who appear as Lead or Assistant Surgeon, 12 anesthesiologists, 10 perfusionists, and 49 nurses. 51 of these individuals (19 surgeons, 9 anesthesiologists, 3 perfusionists and 20 nurses) started working after the beginning of our dataset. 37 of the 115 individuals (13 surgeons, 6 anesthesiologists, 4 perfusionists and 14 nurses) do not appear during the last year of our dataset. We also have data regarding in-hospital mortality and information related to the patient's condition prior to operation. Specifically, all patients are divided into three groups by the hospital according to the severity of their case and their characteristics after an initial clinical assessment. These groups are: "Severe", "Medium" and "Mild".

After removing operations for which there are missing data, we are left with 6,171 operations. We also dropped operations during which the patient died since the duration of these operations can be misleading with regards to team productivity. As a result, we are left with 6,129 operations. Using the hospital's classification system, and conducting interviews with several surgeons, we categorize each surgery into one of the following nine major sets: Coronary Artery Bypass Graft (CABG), Valve Repair/Replacement, Congenital Surgery, Heart Failure, Tumor Removal, Routine Cardiac Surgery, Other Normal Surgery (where we include operations that are not characterized by any of the previous categories), Double Surgery (including two types of the previous categories during the same operation) and Triple Surgery (including three previous categories during the same operation). Our dataset consists of 3,273 CABG operations, 1,324 Valve Repair/Replacements, 86 Congenital Surgeries, 70 Heart Failure procedures, 20 Tumor Removals, 185 Routine Cardiac Surgeries, 78 Other Normal Surgeries, 965 Double Surgeries and 128 Triple Surgeries. It is worth noting that all operations that belong to the last two categories (Double and Triple Surgeries) are characterized as "severe" cases.

Finally, we conduct a limited number of interviews with several staff members to get a better understanding of the cardiac surgery procedures, and to acquire information related to management practice and organizational policies of the hospital.

Variables

Dependent Variable. We use surgery duration in minutes as our dependent variable. Consistent with prior research (Pisano et al. 2001, Edmondson et al. 2003, Reagans et al. 2005) we contend that shorter surgery durations represent an increase in the productivity of the surgical team. The duration is calculated by the time the patient enters the operating room until the time he/she is taken out of the operating room. The duration of the procedure is considered to be one of the most important factors

affecting the total costs for treating patients undergoing a cardiac surgery (Pisano et al. 2001). In our dataset, the average duration of an operation is equal to 294 minutes, with a minimum of 29 minutes and a maximum of 900 minutes.

Independent Variables

Team Familiarity Dispersion. In order to capture team familiarity dispersion, we first calculate the number of past collaborations for all pairs within the team. This represents pairwise familiarity for each pair in a team. We then calculate the standard deviation of these pairwise familiarity scores. This way, we characterize familiarity dispersion within the team, and we distinguish between teams which may have similar levels of average team familiarity, but with a difference in the distribution of past collaborations among the pairs in a team.

Notice that dispersion can be operationalized in several ways depending on the theoretical construct. Indeed, Harrison and Klein (2007) provide a related discussion of three major types of diversity constructs (i.e., ‘separation’, ‘variety’ and ‘disparity’) and their appropriate operationalizations. Our conceptualization of familiarity dispersion within fluid teams where pairs differ one another in their position along a continuous attribute (e.g., team familiarity) is very similar to a ‘separation’ type construct which is typically operationalized by standard deviation. Notice that we have also considered alternative measures (e.g., coefficient of variation, range) for team familiarity dispersion in the robustness checks section.

Bottleneck-Pair. For each team performing a surgery, we first identify the pair with the minimum number of past collaborations (i.e., the least familiar pair), and get its pairwise familiarity score. We divide this score with the team’s average familiarity score, and call the resulting number the ‘bottleneck score’ of the team. That is, the ‘bottleneck score’ captures the degree of familiarity of the least familiar pair relative to the team’s average. We then use four bins to create a discrete variable representing bottleneck scores. Specifically, we create a new variable which is equal to 4 if the bottleneck score is lower than the 25th percentile, 3 if it is between the 25th percentile and the median, 2 if it is between the median and the 75th percentile and 1 if it is higher than the 75th percentile. This approach helps us identify teams who are more susceptible to bottleneck effects (i.e., those with very low bottleneck scores -lower than the 25th percentile) with compare to teams who are less susceptible to such bottleneck effects (i.e., those with bottleneck scores in other bins).

Another way of capturing this effect could have been to take the pair with the lowest familiarity score and use their pairwise familiarity score directly, without dividing by the average team familiarity. Notice that similar to a factory setting, we expect pairs with the lowest familiarity scores to act as

bottlenecks relative to other pairs in the same team. That is, we argue that existing levels of information sharing and collaboration within a team will be hampered by a pair whose familiarity is much lower than the team's average. Hence, taking the pair with the lowest familiarity in every team and using their familiarity score directly would not have been appropriate, because a low score may merely indicate a newly formed team with many unfamiliar pairs rather than an actual bottleneck.

One could have also used bottleneck scores (lowest score divided by the team average) directly as a continuous variable. Notice that, instead of doing so, we assign each bottleneck score to one of the four bins described above. This is for two reasons. First, bottleneck scores have a highly skewed distribution (see Figure S1) whereby even transforming the variable (e.g., log transformation) does not alleviate the skewness problem. Second, our conceptual model suggests that the existence of a bottleneck pair in a team would hinder overall knowledge transfer capability of a team. Specifically, it suggests that certain teams are more susceptible to bottleneck effects than others. For example, consider two teams with very high bottleneck scores (i.e., their least familiar pairs do have a good level of familiarity compared to team's average). Then, even if these two teams differ in their actual bottleneck scores, we expect neither of them to demonstrate any bottleneck related hindrance, that is, we do not expect to see any difference in their productivity as a result of the small difference in their actual bottleneck scores.

Finally, notice that we have also considered alternative measures of this variable in the robustness checks section.

Severe Familiarity Percentage. Recall that our third hypothesis asserts that the higher the ratio of familiarity gained from complex tasks, the higher the productivity. In order to measure the ratio of familiarity gained from complex tasks, we first need to identify complex tasks in our setting. As mentioned earlier, each procedure is characterized as “severe”, “medium” or “mild” according to the condition of the patient. In healthcare settings, variation in patient conditions is suggested to be the leading driver of ‘input’ uncertainty (Argote 1982, Gittel 2002). In addition, “severe” cases are associated with both higher output uncertainty as well as higher likelihood of adverse events. Combined together, we contend that ‘severe’ operations constitute more complex tasks due to higher uncertainty associated with them (Schroder et al. 1967, Campbell 1988).

After identifying “severe” operations as complex tasks, we then break down team familiarity into two components depending on the way it is gained: Familiarity gained while performing a complex (i.e., “severe”) task, and familiarity gained while performing either a “medium” or a “mild” task. Our independent variable Severe Familiarity Percentage, then captures the ratio of team familiarity that is gained through working together in complex (i.e., severe) cases.

We compute this variable by first calculating the sum of all pairwise familiarity scores in the team considering shared experience in all three (mild, medium, severe) types of cases. We then calculate the same sum, but this time by only considering the severe cases. We divide the latter with the former to get Severe Familiarity Percentage. This way, we capture the percentage of past collaborations that took place during a complex operation and expect it to have a beneficial effect on team productivity.

Control Variables

Average Team Familiarity. Similar to previous works (Reagans et al. 2005), we count the number of times every pair in the team has worked together before (without including the current operation), take the sum for all pairs in the team and then divide this number by $(N(N-1)/2)$, where N is the team size. This variable represents the average level of familiarity within the team. Similar to previous empirical studies (Reagans et al. 2005, Huckman et al. 2009, Huckman and Staats 2011, Staats 2012) we expect team familiarity to promote team productivity.

Team Size. We control for the number of all team members N . Larger teams may have difficulty in coordination, which could result in decreased productivity (Hackman 2002, Reagans et al. 2005).

Individual Average Direct Experience. We control for individuals' prior experience in the same type of procedures. For each team member, we calculate the number of times they appear in a same type procedure prior to the current procedure (not including the current one). We then take the sum, and divide by the number of team members. This way, we control for individual experience of team members in a same type of operation. Instead of considering individuals' direct experience only in the same type of operations in the past, we could have also considered their experience on other types of operations (i.e., indirect experience), and combine them in a variable called Individual Average Experience which captures the total number of operations participated by each team member prior to the focal one. Our results remained the same when our present variable is replaced with this one. In addition, we control separately for Lead Surgeon's experience and the other team members (since one can claim that Lead Surgeon's experience is more important than the one of the other team members) and the results remain the same.

Indicator of Quarter. In order to capture any effects of potential technological advances that may influence operation times, or potential changes in the hospital policy that may have an impact on operation durations (e.g., additional paperwork required, etc.), we include an indicator variable for each quarter of our sample indicating the quarter since the beginning of our dataset.

Indicator for Patient Condition. In order to control for the patient's condition which is characterized as "mild", "medium", or "severe", we include two dummy variables in all models.

Considering “medium” as the reference category, “severe” variable is equal to one if the case is characterized as severe and zero otherwise, whereas “mild” variable is equal to one if the case is characterized as mild and zero otherwise. Because doctors already take into account a good number of critical factors such as age, clinical history, past complications, other existing health problems, etc. when they classify patients into these three categories prior to the operation, we believe it is a good indicator of overall patient characteristic.

Indicator for Procedure Type. As described above, each operation can be characterized and considered in one of the nine categories. Hence, in our model we include eight indicators (we consider the most common type CABG as the reference category) that are equal to one if the procedure is in the respective category and zero otherwise.

Indicator for the Lead Surgeon. In order to control for the skill of the Lead Surgeon, which may be one of the most influential factors in surgery performance, we control for the Lead Surgeon fixed effects by including a dummy variable indicating the Lead Surgeon in each operation.

Indicator for the Day of the Week. Day of the week may influence productivity. So we include a dummy variable indicating the day of the week that each operation was performed.

4. Results

Given the structure of our data and the level of our analysis, to test our hypotheses, we use Ordinary Least Square Regression with AR(1) covariance structure to control for serial correlation among operations performed close in time (see Reagans et al. 2005 for similar structure). We also checked for normality of residuals and for heteroscedasticity. Also, consistent with the previous research which introduced team familiarity using a learning curve model (Reagans et al. 2005) and the learning literature (Pisano et al. 2001, Edmondson et al. 2003), we similarly take the logarithm of our dependent variable (i.e., surgery duration) and our learning related continuous variables. These variables also demonstrate positive right skewed distributions. However, for other continuous variables which are not learning related (i.e., team familiarity dispersion and team size), we keep their original forms. All of our transformations were also consistent with Tukey’s (1977) ladder of powers which reveals the most suitable transformations. Our model is the following:

$$\begin{aligned}
 \ln(\text{surgery duration}_t) = & \beta_0 + \beta_1 (\text{team familiarity dispersion}_t) + \\
 & \beta_2 (\text{bottleneck-pair}_t) + \\
 & \beta_3 \ln(\text{severe familiarity percentage}_t) + \\
 & \beta_4 (\text{severe}_t) \times \ln(\text{average familiarity}_t) + \\
 & \beta_5 \text{controls} + u_t, \\
 \text{where } u_t = & \rho u_{t-1} + e_t
 \end{aligned}$$

Table 1 shows descriptive statistics and correlations among the variables. It reports the values for the logged and normalized variables as described earlier. Table 2 shows the results for hypotheses 1, 2, 3 and 4. Notice that we report robust standard errors in all our models. Also, while our final sample includes 6,129 operations, notice that the number of observations in Table 2 is equal to 6,128 due to the AR(1) covariance structure. In order to check for any potential multicollinearity issues, we introduce our variables of interest one at a time in our models. In addition, we checked the Condition Index (CI) and it is below the suggested threshold of 30 (Cohen 2003). In model 1, we include only the control variables. In model 2, we include *Team Familiarity Dispersion* and we observe a 0.84% increase (significant at 1% level) in the adjusted R². *Team Familiarity Dispersion* has a positive and significant coefficient at 1% level, providing support for our first hypothesis. In model 3 we remove *Team Familiarity Dispersion* and include *Bottleneck-Pair*. We observe a 0.84% increase (significant at 1% level) in the adjusted R² compared to model 1. The coefficient of the *Bottleneck-Pair* is significant at 1% level and positive providing support for H2. In model 4 we include both *Team Familiarity Dispersion* and *Bottleneck-Pair* and still find support for both H1 and H2. We observe a 1.49% increase (significant at 1% level) in the adjusted R² compared to model 1. In model 5 we add *Severe Familiarity Percentage* and we observe a 0.41% increase (significant at 1% level) in the adjusted R² compared to model 4. *Severe Familiarity Percentage* is significant at 1% level with a negative coefficient. That is, Hypothesis 3 is also supported.

While the increase in the adjusted R² between the initial model (with only control variables) and the full model may be small (i.e., 1.89%), this is not surprising. This is because of the dominance of clinical factors in explaining variation in most healthcare performance metrics (see for example Reagans et al. (2005), which report similarly small changes in the adjusted R²).

An increase of one standard deviation on *Team Familiarity Dispersion* will increase the duration of the operation by 2.14% (assuming all other variables remain constant). In addition, the coefficient of our variable *Bottleneck-Pair* is equal to 0.0176, which suggests that moving from one category to another increases the duration of an operation by 1.76% (assuming all other variables remain constant). Furthermore, an increase of one standard deviation on *Severe Familiarity Percentage* will decrease duration by 16.84% (assuming all other variables remain constant).

In order to test our fourth hypothesis, we include an interaction term, which is equal to the product of the dummy variable *Severe* and the continuous control variable *Average Team Familiarity*. Model 6 of Table 2 presents the results for our fourth hypothesis. We observe that the results for our previous hypotheses remain the same and the adjusted R² is increased by 0.21% (significant at 5% level). The

interaction term is significant at 1% level and negative, providing support for H4. Finally, we conduct a post hoc plot for Hypothesis 4 in order to examine the moderating effects as described by Aiken and West (1991) and Dawson and Richter (2006). Figure 1 reveals the moderating relationship as suggested in Hypothesis 4.

5. Robustness Checks

We perform a number of robustness checks to analyze the sensitivity of our results, to address endogeneity concerns and to rule out potential alternative explanations. We group these robustness checks into four sections as listed below:

Potential Biases related to Data

While we observe team familiarity over a long time period (i.e. seven years) and in large number of operations, we do not have information about team assignments prior to our dataset, that is, there is some unobserved familiarity before the start of our dataset. Like Reagans et al. (2005) which use a very similar data structure to calculate team experience, we expect our results to have no systematic bias. Nevertheless, we investigate potential effects of missing team familiarity by repeating our analysis after excluding different time intervals from the beginning of our dataset. That is, we re-run our analyses in subsamples by removing the first 12 months and 24 months respectively from our overall dataset of 87 months. In each subsample, the observations include only surgeries in that subsample and we also calculate team familiarity by using that subsample (e.g., in the removal of 24 months case, our observations include only those surgeries between month 24 and month 87 and when calculating team familiarity we also use surgeries between month 24 and month 87). This way, we investigate the robustness of our results when there is missing data at the beginning. Our results are the same in terms of significance (with different magnitudes as expected) both in the full sample and in the subsample (where there is simulated missing experience), providing further support for our hypotheses.

We also check the sensitivity of our results by repeating our analysis differently. In this case we remove the first 12 and 24 months, and while the observations include only surgeries in the remaining subsample, we calculate team familiarity by using the whole dataset. (e.g., when removing the first 24 months, our observations include only those surgeries between month 24 and 87, but when calculating team familiarity we use all previous surgeries from month 0). Again our results are the same in terms of significance with slightly different coefficients.

Potential Biases related to Methods and Variables Choice

While our log transformation of the dependent variable is consistent with the literature, this is equivalent to a duration or survival model with a lognormal distributional assumption (Cameron and Trivedi 2005,

Cleves et al. 2010). This may put a strong functional form assumption on the shape of the hazard function which may influence our results. Therefore, in order to check that our results are not sensitive to linear or log specification of our dependent variable, we have re-run all our analyses using linear form of the dependent variable. All of our results remained the same in terms of significance with different magnitudes in this specification.

A potential issue with our use of AR (1) covariance structure is that our sample does not satisfy the equal-spacing assumption. We believe this does not pose a threat to our results because we have a large sample (i.e., 6,206 observations) representing a long time interval (i.e., 87 months). Nonetheless, we repeated our analysis after controlling for the unequal spacing in our dataset. Specifically, for each operation we calculate the time difference between subsequent operations in hours, and include this new variable in our analysis. Our results remained the same for all of our hypotheses.

We also consider alternative measures of our main independent variables. First, instead of using *standard deviation* to capture team familiarity dispersion, we consider *coefficient of variation* (i.e., standard deviation/mean) and *range* (i.e., max-min) of pairwise familiarity scores in our teams. Our results with these alternative measures remained the same. Second, we consider an alternative construction of our bottleneck-pair variable. Specifically, after calculating the bottleneck score as before (i.e., familiarity of the least familiar pair divided by the average familiarity of the team), we create four bins depending on the bottleneck score: (1) if the bottleneck score is lower than the 25th percentile, (2) if it is between the 25th percentile and the median, (3) if it is between the median and the 75th percentile, and (4) if it is higher than the 75th percentile. We then used three indicator variables for these categories taking the last category as the reference and re-run our models. Indicator for the first category (i.e., 1) is significant at 1% with a positive coefficient, indicator for the second category (i.e., 2) is significant at 5% and positive with a lower coefficient than the first category (a t-test also indicated that the two coefficients are different), and the indicator for the third category (i.e., 3) is insignificant. This confirms our earlier results for the second hypothesis. In addition, we use an alternative continuous variable for Bottleneck-Pair, which is equal to the absolute distance between the least familiar pair and the team's average familiarity. Consistent with our previous results for H2, this variable is significant and positive ($p < 0.01$). Finally, we also use alternative variables for both dispersion and bottleneck-pair in the same model (*coefficient of variation* and absolute distance between the least familiar pair and team's average familiarity, and *range* and absolute distance between the least familiar pair and team's average familiarity) and get the same qualitative results for all of our hypotheses.

Our analyses indicate that while every team has a least familiar pair, bottleneck influence come into effect at a certain point, after which it becomes a performance hampering constraint for the entire

team. Similar to Berry Jaeker and Tucker (2016) which identifies an inflection point in utilization and performance relationship in healthcare, we next try to identify the threshold for bottleneck effects. To do so, we create ten indicator variables corresponding to each decile of the bottleneck scores, using the top decile (above 90%) as the reference. All indicators up to the 5th decile are significant at either 1% or 5% level and positive, whereas the 5th decile indicator is significant at 10%, and the remaining indicators for higher deciles are insignificant. This suggests that the threshold lies between the 5th and the 6th decile of the bottleneck score at which point productivity hindrance of bottleneck effects come into effect.

Potential Selection Biases

Selection of team members in surgery operations could be a significant concern. Although according to the hospital management the assignment of surgeons to operations is random, one may argue that Lead Surgeons may prefer working with specific Assistants Surgeons, and this may bias our results related to Hypothesis 1. In order to address this, we repeated our analysis by calculating team familiarity dispersion without taking into account any pairs that include only surgeons (i.e., surgeon-surgeon pair) in a team. Our results with the modified team familiarity dispersion remained the same in terms of significance with slightly smaller effect sizes. In addition, nurses, anaesthesiologists and perfusionists are randomly assigned to operations according to their shifts. Nonetheless, in order to eliminate any potential bias in our results, we examined this issue further: We calculated for each Lead Surgeon the percentages of the operations that she has worked with each anaesthesiologist, perfusionist and first scrub nurse (i.e. lead nurse). With anaesthesiologists, the highest percentage is 33.10% (but the Lead Surgeon has worked with 9 out of the 12 anaesthesiologists in our dataset), with perfusionists the highest percentage is equal to 36.19% (and the Lead Surgeon has worked with 6 out of the 10 perfusionists appearing in our dataset) and with lead nurses it is equal to 29.81% (the lead surgeon has worked with 38 out of the 41 nurses appearing as lead nurses in our sample). Considering these percentages and the fact that out of the 115 individuals in our sample, 51 of them (19 surgeons, 9 anesthesiologists, 3 perfusionists and 20 nurses) started working after the beginning of our dataset and 37 (13 surgeons, 6 anesthesiologists, 4 perfusionists and 14 nurses) do not appear during the last year of our dataset, we believe that there is no systematic selection of lead nurses, anesthesiologists, or perfusionists by the Lead Surgeon in the hospital.

One may also argue that some surgeons could avoid operations with high chance of complications, more severe cases might be assigned to more experienced Lead Surgeons, or easier cases might be assigned to junior Lead Surgeons. First, according to our interviews, official hospital policy

is clear in not allowing surgeons to choose or avoid certain operations. Second, we investigate the spread of “severe”, “moderate” and “mild” operations among surgeons with a chi-square test of independence where the null hypothesis is that each patient’s assigned doctor is independent of her condition. We do not reject the null hypothesis (with chi-square statistics of 13.271, 16.887 and 14.674 for each operation type respectively), and conclude that surgeries are indeed evenly spread among lead surgeons. Third, in our analysis we have excluded all highly complex cases that resulted in the death of the patient during surgery (42 in total). We repeat our analysis after including them and get the same results in terms of significance. Finally, we repeat our analysis after excluding the severe and all the mild cases in our sample, and left with only the medium cases. Our conclusions remained the same for all of our hypotheses.

Finally, if lead surgeons prefer to work with more familiar people for severe cases (since there is an increased likelihood of major adverse events), this may create bias in our results for H4. In order to rule out this alternative explanation, we examined an alternative specification of task complexity. We created a new variable called *Patient*, which reflects the condition of the patient before the operation. Recall that severe cases represent the most complex operations. In the new specification, *Patient* is equal to 1 if the condition of the patient characterized by the hospital is “mild”; 2 if it is “medium”; and 3 if it is “severe”. Next, we created a new interaction term by multiplying *Average Team Familiarity* and *Patient* and we ran our analysis for H4 using the new variables. Our results for hypotheses 1, 2 and 3 are the same in terms of significance and very similar in terms of coefficients, and the new interaction term is also significant and negative at 5% level providing support for the fourth hypothesis. We therefore believe that our results for H4 are not biased by potential selection of more familiar members by the lead surgeons in severe cases. This is because we observe an increasing beneficial effect of team familiarity on productivity with increasing task complexity (i.e., from mild-1 to severe-3).

Other Alternative Explanations

An alternative explanation for H1 and H2 could be that the detrimental effects from familiarity dispersion and bottleneck-pair may actually be created not by the mechanisms we suggested, but rather by the sole presence of a new team member or a new pair (i.e., new to other team members), or by a team member with very little prior individual experience (i.e., new to the hospital). In order to address this, we conducted two analyses. First, we ran our analysis after removing 780 operations from our sample in which there is a new member, or a pair working together for the first time. Second, we repeated our analysis after introducing a new variable which is equal to the direct experience of the

team member with the lowest direct experience in the team. Our conclusions remained the same in both analyses.

An alternative explanation for H3 could be that it is not familiarity, but individual experience gained in highly complex operations that may make the team more productive subsequently. In order to test this, we repeated our analysis by replacing our variable Individual Average Direct Experience with a new variable Individual Average Severe Experience, which controls for individual experience of team members from severe operations in the past. Introducing this new variable did not change our conclusions for H3. Moreover, we repeat our analysis after controlling for the Lead Surgeon Direct Experience (while including the same variables for other team members) and find consistent results with our earlier findings.

6. Policy Implications of Familiarity Dispersion and Bottleneck Pair

We believe the new refinements of team familiarity proposed in this paper are not only statistically significant but also managerially relevant, as our results suggest new and potentially different managerial actions than those previously suggested in the literature. For example, finding a positive effect of average team familiarity on performance, Huckman et al. (2009) and Huckman and Staats (2011) discussed “the managerial advice to keep teams together” (p.325) to increase the average familiarity of the teams. However, they also acknowledged the practical difficulty of such strategy as “all team members cannot work on teams with high team familiarity” (p.325). In our study, our key findings highlight the negative effects of team familiarity dispersion and bottleneck pairs on productivity. Therefore, rather than trying to keep team members together by allocating the same team members to the same teams repeatedly, we suggest managers to actively focus on the least familiar pairs and individuals, and try to increase their familiarity levels.

In order to further assess the practical implications of our proposals, we next run a basic policy simulation with team assignment rules based on our results. Specifically, we compare productivity implications of two policies: Based on our findings, *Policy_dispersion* is concerned with the detrimental effects of dispersion and bottlenecks, and aims to increase the familiarity levels of the least familiar pairs in the hospital. *Policy_stable*, on the other hand is based on previous findings in the team familiarity literature which highlight the beneficial effects of average team familiarity on productivity. That is, the objective of *Policy_dispersion* is to reduce dispersion and bottleneck effects whereas the objective of *Policy_stable* is to increase average team familiarity by keeping teams as stable as possible. We then compare the estimated productivity benefits of these two policies.

We use real data from our study and run our simulations in three different time intervals, each

with duration of three months, so that we can observe diverse set of familiarity scores. The simulations take place between the 21st - 24th, 43rd - 47th, and 65th – 68th months in our sample. These time intervals represent the 25th percentile, the median and the 75th percentile of our longitudinal sample. Also, during these three month time intervals, there were 1, 2 and 1 new hires respectively. Considering that 51 individuals joined the hospital during our entire sample period of 87 months (i.e., 0.586 new hires per month on average), these intervals are indeed representative of our sample in terms of employee hires. In each of these quarters, we use existing pairwise familiarity levels (i.e., those from the beginning of our dataset to the focal month), and form a new team for each upcoming operation based on the proposed policy. We then calculate predicted productivity improvements (i.e., reduction in duration) from the proposed team assignment rules based on our econometric model by focusing on average team familiarity, dispersion, and bottleneck pair, and by keeping all other variables constant.

Policy_dispersion: First, we calculate the familiarity score of each individual with all other individuals in our sample using the data from the beginning of our sample to the current operation. Next, following our hospital's existing practice, we assign Lead Surgeons to individual operations with a rotating schedule. After assigning Lead Surgeons, we then assign the remaining team members based on their familiarity levels with each other. Specifically, in line with this policy, we assign individuals with the lowest pairwise familiarity scores to the focal operation. In making such assignments, we also impose multiple constraints to reflect realities of our setting as much as possible. These include taking into account the required roles needed in surgery teams and also making sure that no individual conducts more than two operations per day. Overall, the objective of this policy is to reduce familiarity dispersion and bottleneck effects by actively assigning the least familiar individuals to the same teams.

Policy_stable: For this policy, similar to the previous policy, we first calculate the familiarity score of each individual with all other individuals and assign Lead Surgeons to operations on a rotating basis. Next, after identifying the Lead Surgeon for the focal operation, we calculate his/her existing familiarity scores with all individuals (using the entire sample up to the focal operation), and then assign those individuals to whom the Lead Surgeon has the highest familiarity scores to the same team to perform the focal operation. We then keep these team members together with the Lead Surgeon for the rest of the simulation period. This allows us to create highly stable teams around Lead Surgeons with high average team familiarity scores. In addition, when a newly hired individual appears, we replace him/her with the member who has the lowest familiarity score in our sample. Similar to the previous policy, we also impose multiple constraints such as considering the particular role of individuals with replacement decisions (e.g., a nurse to nurse replacement), and not allowing individuals to conduct more than two

operations per day. Overall, this policy aims to promote average team familiarity by keeping teams as stable as possible.

We next compare the two policies with each other and with the existing team assignment practice as observed in our sample.

Comparison of the policies: As expected, both policies result in lower operations durations with compare to the current case. However, notice that in all simulations, Policy_dispersion performs better than Policy_stable with a higher reduction in overall expected duration. Table 3 demonstrates the changes in team familiarity dispersion, average team familiarity, bottleneck pair and the expected duration in the two policies with compare to the current practice. For dispersion and average familiarity, we show the percentage change with respect to the current policy. For bottleneck-pair, we show the actual difference between the proposed policy and the current policy as this is an ordinal discrete variable ranging from 1 to 4 (a score of 1 in the Table indicates a one unit change, e.g., from 2 to 3). We also indicate in parenthesis the percent change in the bottleneck score, from which bottleneck-pair is calculated.

With Policy_dispersion, expected duration is reduced by 3.63% translating into a reduction of 10.42 minutes over three months on average. This reduction is due to a decrease in dispersion and in bottleneck pair. Despite the fact that average team familiarity also decreases, which leads to an increase in duration in this policy, the expected gains from dispersion and bottleneck-pair are higher than the detrimental effects of reduced average team familiarity. On the other hand, by employing Policy_stable, expected duration is decreased by 0.47% which is translated into 1.35 minutes on average. Although dispersion and bottleneck pair increase as a result of this policy (which would have increased duration), the gains from average team familiarity are higher than the losses associated with increased dispersion and bottleneck-pair, therefore resulting in overall slightly decreased duration. In addition to our original policy period of three months, we also checked the robustness of our comparisons by considering two-month and four-month periods respectively. In all cases, consistent with our original analysis, Policy_dispersion results in higher productivity improvement than Policy_stable.

While numerous other team assignment policies could be considered with similar objectives, our goal in these simulations is to capture the essentials of the proposed mechanisms in a simple and practical way. Also, because we want to demonstrate that Policy_dispersion may perform better than Policy_stable, in making team assignment policies, we incorporate favorable scenarios for Policy_stable (e.g., by replacing a newcomer with the least familiar individual) to make a conservative comparison between the two policies.

Previous work on team familiarity highlighted the beneficial effects of team familiarity on productivity and suggested managers to keep team members together as much as possible to maximize average team familiarity. While this seems like an obvious recommendation, our study points out to a detrimental role of dispersion and bottleneck pairs, two elements that are not considered in the previous literature. Indeed, we demonstrate that stable team recommendations based entirely on average team familiarity may be myopic. Our simple policy analyses reveal that a dispersion and bottleneck minimizing policy perform better than a team familiarity maximizing policy.

Many fluid team settings are characterized by high employee turnover, and therefore new team members will inevitably need to adjust to those team members who are already familiar with each other. Noticing the potential detrimental effects of dispersion and bottlenecks created by these individuals, managers can mitigate this problem by adopting rotation-based schedules with a special emphasis on the least familiar pairs. Such practice will lead to long-term increased productivity. In addition, in such a policy, average team familiarity will also eventually increase in the long-run. In contrast, a policy aiming to keep the same teams together might create a short-term beneficial effect for team productivity, but will eventually lead to decreased performance in the long-run due to detrimental effects of dispersion and bottleneck-pairs.

7. Conclusion and Discussion

Although organizations increasingly rely on fluid teams to perform essential tasks, our understanding of formation and operations of fluid teams remain limited. In this study, we observe the relationship between composition dynamics of fluid teams and subsequent productivity implications over a long time period. Studying fluid teams at a granular level, we offer a new account of how shared work experience among team members influence team productivity.

Most of the previous literature in operations and management has focused on mean levels of familiarity within teams (e.g., Edmondson 1999, Faraj and Sproull 2000, Huckman and Staats 2011) and has largely ignored the distribution of pairwise familiarity scores. This is a significant limitation for understating the role of team familiarity in fluid settings, because fluid teams have a dynamic nature in terms of team memberships, and there is likely significant variation in terms of pairwise familiarity within these teams even if they have similar average levels of team familiarity. Consequently, the present study improves our conceptual understanding of team familiarity by considering not only the mean, but also the shape of the familiarity distribution, and by demonstrating the importance of dispersion and bottlenecks in explaining team productivity. Considering the two levels of relationships in fluid team settings: at the pair and at the team level, while the emergence of average team familiarity is a straightforward extension of pairwise familiarity, dispersion and bottleneck capture distinct team

level characteristics. In addition, although these two are somewhat related, we contend that dispersion and bottleneck are two distinct theoretical concepts.

An important observation we had is that a pair within the team with a familiarity level much less than the average familiarity of the team may slow down the whole team, function as a bottleneck, and cause delays and eventually decrease the productivity of the whole team. Our conceptualization of bottleneck-pair could also be viewed with the lens of recent research about the introduction of surgical checklists and their benefits (Haynes et al. 2009, Gawande 2010). A central feature of surgical checklists is preoperative team briefings where each person in the team speaks and introduces herself to other team members. Surgical checklists are suggested to be highly effective in improving teamwork and communication in the operating room (Gawande 2010), and team briefings have been associated with better surgical outcomes (Lingard et al. 2008). We suggest that, a key benefit of team briefings could be to reduce or eliminate the potentially negative effect of a bottleneck pair in team. While team briefings may not contribute much to team members who are already quite familiar with each other, it may provide substantial benefits in reducing potential communication and information sharing problems between team members with very little prior shared work experience (i.e., bottleneck pairs). Our results indicate that the existence of such a bottleneck pair can increase the duration of an average operation by 5.17 minutes. So eliminating such a bottleneck pair (moving a pair from the highest category to the next one) will result in about 6 hours savings in total surgery time per month. Also, when there is a new staff member, hence the effect is inevitable; managers can moderate the detrimental effect by initially assigning her and other relatively new employees to simple cases together.

Another important finding which may help managers to develop better team composition strategies is that team members that have worked together previously in more complex tasks could develop a transactive memory system more effectively. They may also create stronger bonds and relationships when their shared experience comes from complex cases, which helps them better collaborate in subsequent tasks. When assigning individuals to teams, managers could consider not only the amount of shared work experience, but also the nature of shared experience in previous operations (e.g., complex vs non-complex tasks). One suggestion for hospital managers could be that if an occasional operation requires very high levels of efficiency (e.g., a patient in critical condition), then for this operation, assigning individuals who worked together in many complex tasks in the past could be a preferable strategy to enhance their productivity.

Finally, teams with high levels of team familiarity seem to be ideal for performing more complex tasks because the benefits from being familiar with each other increase as the level of task complexity increases. That is, when the hospital faces a highly complex case, it may be a preferable strategy to

assign individuals who are highly familiar with each other to these cases, as the benefit of team familiarity will be maximized during this complex operation.

As in all empirical studies, our results come with limitations. First, our data comes from a single hospital, so one has to be cautious when interpreting our results. While cardiac surgery operations are an appealing setting with a mix of repetitive and highly critical knowledge-based tasks to study fluid teams, generalizing our results to other fluid team settings requires a careful approach. Moreover, our dataset includes limited information regarding patients' condition before the operation, as well as doctors' characteristics. One would ideally like to have more detailed information about patients' condition and characteristics such as Higgins Score or EuroScore, but unfortunately this was not available from our data source. Similarly, more detailed information on doctors' experience, performance, skill sets, and other characteristics would have been useful to control in predicting team productivity. In addition, in this study we focus on productivity implications of team familiarity with the help of our highly granular surgery duration data (in minutes), and used in-hospital mortality (a crude measure of quality) as a robustness check to make sure productivity improvements do not come at the expense of increased mortality rates. Future research can investigate how team familiarity can simultaneously affect both productivity and quality in healthcare settings by obtaining more granular quality data such as hospital revisits, clinical and patient reported improvements, among others. Finally, our study joins to the stream of work which highlights the importance of team familiarity in fluid teams with a large scale empirical study. As fluid teams become more of a rule than exception in many organizational settings, future research could provide a more in-depth and vivid account of how team familiarity influences team operations and task dynamics with observational or experimental studies.

Despite these limitations our results provide important insights into operations management research and practice. Our study suggests that managers could practice better team composition strategies that may allow them to make optimal use of past shared experience of team members, and thus improve productivity of fluid team operations without making use of any additional resources. In addition, in a critical and costly setting like healthcare, any productivity gain (e.g., reduction in operation time) will not only reduce total costs of treating patients, but it may also translate into better clinical outcomes for patients by reducing infection risks or adverse event probabilities. Also, answering the call to decompose team familiarity and further investigate its detailed effects on performance (Huckman et al. 2009), our study introduces two new concepts: team familiarity dispersion and bottleneck pairs. Further disentangling team familiarity by concentrating on how much of the shared work experience is gained during complex tasks in the past; we demonstrate that not all experiences are equivalent, and the higher the ratio of familiarity which is gained from complex tasks, the higher the

productivity of the team. Finally, we identify focal task complexity as an important moderator between average team familiarity and productivity, and find that team familiarity is particularly helpful when performing more complex tasks.

References

- Aiken, L. S., S. G. West. 1991. *Multiple regression: Testing and interpreting interactions*. Sage, Thousand Oaks, CA.
- Altuntas, S., U. Baykal. 2010. Relationship between nurses' organizational trust levels and their organizational citizenship behaviors. *J. of Nursing Scholarship*. **42**(2) 186-194.
- Argote, L. 1982. Input uncertainty and organizational coordination in hospital emergency units. *Admin. Sci. Quart.* **27**(3) 420-434.
- Banker, R. D., J. M. Field, K. K. Sinha. 2001. Work-team implementation and trajectories of manufacturing quality: A longitudinal field study. *Manufacturing & Service Oper. Management*. **3**(1): 25-42.
- Berman, S. L., J. Down, C. W. L. Hill. 2002. Tacit knowledge as a source of competitive advantage in the National Basketball Association. *Acad. of Management J.* **45**(1) 13-31.
- Berry Jaeker, J. A., A. Tucker. 2016. Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Sci.* Published online in Articles in Advance 04 Apr 2016: 1-21.
- BR. Staats, D.M. Upton. 2009. Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Sci.* **55**(1) 85-100.
- Bezrukova, K., S. M. B. Thatcher, K. A. Jehn. 2007. Group heterogeneity and faultlines: Comparing alignment and dispersion theories of group composition. *Conflict in organizational groups: New directions in theory and practice*: 57-92.
- Bion, W. R. 1962. *Experience in groups*. Basic Books, NY.
- Boudreau, J, W. Hopp, J. O. McClain, L. J. Thomas. 2003. On the interface between operations and human resources management. *Manufacturing & Service Oper. Management*. **5**(3): 179-202.
- Brandon, D. P., A. B. Hollingshead. 2004. Transactive memory systems in organizations: Matching tasks, expertise, and people. *Organ. Sci.* **15**(6) 633-644.
- Bushe, G. R., A. Chu. 2011. Fluid teams: Solutions to the problems of unstable team membership. *Organ. Dynam.* **40**(3) 181-188.
- Campbell, D. J. 1988. Task complexity: A review and analysis. *Acad. of Management Rev.* **13**(1) 40-52.
- Catchpole, K. R., A. E. B. Giddings, M. Wilkinson, G. Hirst, T. Dale, M.R. De Leval. 2007. Improving patient safety by identifying latent failures in successful operations. *Surgery*. **142**(1) 102-110.
- Chillemi, O., B. Gui. 1997. Team human capital and worker mobility. *J. of Labor Econom.* **15**(4) 567-585.
- Choi, J. N., T. Sy. 2010. Group-level organizational citizenship behavior: Effects of demographic faultlines and conflict in small work groups. *J. of Organ. Behav.* **31**(7): 1032-1054.
- Cohen, S. G., D. E. Bailey. 1997. What makes teams work: Group effectiveness research from the shop floor to the executive suite. *J. of Management*. **23**(3) 239-290.
- Cohen, J., P. Cohen, L. S. Aiken, S. G. West. 2003. *Applied Multiple Regression—Correlation Analysis for the Behavioral Sciences*, 3rd eds. Lawrence Erlbaum, Mahwah, NJ.
- Cole, M. S., A. G. Bedeian, R. R. Hirschfeld, B. Vogel. 2011. Dispersion-Composition Models in Multilevel Research A Data-Analytic Framework. *Organ. Res. Methods*. **14**(4): 718-734.
- Coleman, J. S. 1988. Social capital in the creation of human capital. *Amer. J. of Sociology*. 94 95-120.
- Colquitt, J.A, J. A. Lepine, C. P. Zapata, R. E. Wild. 2011. Trust in typical and high-reliability contexts: Building and reacting to trust among firefighters. *Academy of Management J.* **54**(5) 999-1015.
- Cook, L. M. 2013. Can nurses trust nurses in recovery reentering the workplace? *Nursing* **43**(3) 21-24.
- Cramton, C. D. 2001. The mutual knowledge problem and its consequences for dispersed collaboration. *Organ. Sci.* **12**(3) 346-371.

- Cross, R., A. Parker, L. Prusak, S. P. Borgatti. 2001. Knowing what we know: Supporting knowledge creation and sharing in social networks. *Organ. Dynam.* **30**(2) 100-120.
- Davenport, D. L., W. G. Henderson, C. L. Mosca, S. F. Khuri, R. M. Mentzer. 2007. Risk-adjusted morbidity in teaching hospitals correlates with reported levels of communication and collaboration on surgical teams but not with scale measures of teamwork climate, safety climate, or working conditions. *J. of the Amer. College of Surgeons.* **205**(6) 778-784.
- Dawson, J. F., A. W. Richter. 2006. Probing three-way interactions in moderated multiple regression: development and application of a slope difference test. *J. of Appl. Psych.* **91**(4) 917-926.
- De Jong, B. A., K. T. Dirks. 2012. Beyond shared perceptions of trust and monitoring in teams: implications of asymmetry and dissensus. *J. of Appl. Psych.* **97**(2): 391.
- Dineen, B. R., R. A. Noe, J. D. Shaw, M. K. Duffy, C. Wiethoff. 2007. Level and dispersion of satisfaction in teams: Using foci and social context to explain the satisfaction-absenteeism relationship. *Acad. of Management J.* **50**(3) 623-643.
- Drucker, P. F. 1999. Knowledge-worker productivity: The biggest challenge. *California Management Rev.* **41**(2) 79-94.
- Edmondson, A. C. 1999. Psychological safety and learning behavior in work teams. *Admin. Sci. Quart.* **44**(2) 350-383.
- Edmondson, A. C., R. M. Bohmer, G. P. Pisano. 2001. Disrupted routines: Team learning and new technology implementation in hospitals. *Admin. Sci. Quart.* **46**(4) 685-716.
- Edmondson, A. C., I. M. Nembhard. 2009. Product Development and Learning in Project Teams: The Challenges Are the Benefits. *J. of Product Innovation Management.* **26**(2) 123-138.
- Edmondson, A. C., A. B. Winslow, R. M. J. Bohmer, G. P. Pisano. 2003. Learning how and learning what: Effects of tacit and codified knowledge on performance improvement following technology adoption. *Decision Sci.* **34**(2) 197-224.
- Espinosa, J. A., S. A. Slaughter, R. E. Kraut, J. D. Herbsleb. 2007. Familiarity, complexity, and team performance in geographically distributed software development. *Organ. Sci.* **18**(4) 613-630.
- Evans, C. R., K. L. Dion. 1991. Group cohesion and performance a meta-analysis. *Small Group Res.* **22**(2) 175-186.
- Faraj, S., L. Sproull. 2000. Coordinating expertise in software development teams. *Management Sci.* **46**(12) 1554-1568.
- Feldman, D. C. 1984. The development and enforcement of group norms. *Acad. of Management Rev.* **9**(1) 47-53.
- Field, J. M., K. K. Sinha. 2005. Applying process knowledge for yield variation reduction: A longitudinal field study. *Decision Sciences.* **36**(1) 159-186.
- Fleming, M., S. Smith, J. Slaunwhite, J. Sullivan. 2006. Investigating interpersonal competencies of cardiac surgery teams. *Canadian J. of Surgery.* **49**(1) 22.
- Gardner, H. K., F. Gino, BR. Staats 2012. Dynamically integrating knowledge in teams: Transforming resources into performance. *Acad. of Management J.* **55**(4) 998-1022.
- Gawande, A. 2010. Checklist Manifesto, Vol. 200. New York: Metropolitan Books.
- Gaynes, R. P., D. H. Culver, T. C. Horan, J. R. Edwards, C. Richards, J. S. Tolson, and the National Nosocomial Infections Surveillance System. 2001. Surgical site infections (SSI) rates in the United States, 1992-1998: The national nosocomial infections surveillance system basic SSI risk index. *Clinical Infectious Diseases.* **33**(Suppl 2) 69-77.
- Gersick, C. J., J. R. Hackman. 1990. Habitual routines in task-performing groups. *Organ. Behav. and Human Decision Processes.* **47**(1) 65-97.
- Gibbons, C., J. Bruce, J. Carpenter, A. P. Wilson, J. Wilson, A. Pearson, D. L. Lamping, Z. H. Krukowski, B. C. Reeves. 2011. Identifications of risk factors by systematic review and development of risk-adjusted models for surgical site infection. *Health Technological Assessment.* **15**(30) 1-156.
- Gittel, J. H. 2002. Coordinating mechanisms in care provider groups: Relational coordination as a mediator and input uncertainty as a moderator of performance effects. *Management Sci.* **48**(11) 1408-1426.
- Goldratt, E. M., J. Cox. 1984. *The Goal: A Process of Ongoing Improvement.* North River Press, Croton-on-Hudson, NY.

- Gruenfeld, D. H., E. A. Mannix, K. Y. Williams, M. A. Neale. 1996. Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organ. Behav. and Human Decision Processes* **67**(1) 1-15.
- Hackman, J. R. 1987. The design of work teams. J. W. Lorsch eds. *Handbook of Organizational Behavior*. Prentice Hall, Englewood Cliffs, NJ, 315-342.
- Hackman, J. R. 2002. *Leading Teams: Setting the Stage for Great Performances*. Harvard Business Press, Boston, MA.
- Harrison, D. A., K. J. Klein. 2007. What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Acad. of Management Rev.* **32**(4) 1199-1228.
- Harrison, D. A., S. Mohammed, J. E. McGrath, A. T. Florey, S. W. Vanderstoep. 2003. Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality. *Personnel Psych.* **56**(3) 633-669.
- Haynes, A.B., Weiser, T.G., Berry, W.R., Lipsitz, S.R., Breizat, A.H., Dellinger, E.P., et al. 2009 A surgical safety checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine.* 360(5):491-499.
- Hoffmann, D. A., Lei, Z., & Grant, A. M. 2009. Seeking help in the shadow of a doubt: The sensemaking processes underlying how nurses decide whom to ask for advice. *J. of Appl. Psych.* **94**(5): 1261-1274.
- Hollingshead, A. B. 2001. Cognitive interdependence and convergent expectations in transactive memory. *J. of Personality and Soc. Psych.* **81**(6) 1080.
- Hornsey, M. J., M. A. Hogg. 2000. Assimilation and diversity: An integrative model of subgroup relations. *Personality and Social Psych. Rev.* **4**(2) 143-156.
- Hopp, W. J., S. M. R. Iravani, F. Liu. 2009. Managing white-collar work: An operations-oriented survey. *Production and Oper. Management.* **18**(1) 1-32.
- Hopp, W. J., M. L. Spearman. 2000. *Factory Physics: Foundations of Manufacturing Management*, 2nd ed. McGraw-Hill, Burr Ridge, IL.
- Huckman, R. S., BR. Staats. 2011. Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *Manufacturing & Service Oper. Management.* **13**(3) 310-328.
- Huckman, R. S., BR. Staats, D.M. Upton. 2009. Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Sci.* **55**(1) 85-100.
- Jackson, S. E., A. Joshi, N. L. Erhardt. 2003. Recent research on team and organizational diversity: SWOT analysis and implications. *J. of Management.* **29**(6) 801-830.
- Kane, A. A., L. Argote, J. M. Levine. 2005. Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organ. Behav. and Human Decision Processes* **96**(1) 56-71.
- Katz, R. 1982. The effects of group longevity on project communication and performance. *Admin. Sci. Quart.* **27**(1) 81-104.
- Klein, K. J., S. W. Kozlowski. 2000. Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions, Jossey-Bass.
- Levine, J. M., R. L. Moreland. 1999. Knowledge transmission in work groups: Helping newcomers to succeed. J. M. Levine, D. M. Messick, L. L. Thomson eds. *Shared Cognition in Organizations: The Management of Knowledge*, Psychology Press, New York, 267-296.
- Lewis, K. 2003. Measuring transactive memory systems in the field: Scale development and validation. *J. Appl. Psych.* **88**(4) 587-604
- Liao, H., D. Liu, R. Loi. 2010. Looking at both sides of the social exchange coin: A social cognitive perspective on the joint effects of relationship quality and differentiation on creativity. *Acad. of Management J.* **53**(5) 1090-1109.
- Lingard, L., G. Regehr, B. Orser, R. Reznick, G. R. Baker, D. Doran, S. Espin, J. Bohnen, S. Whyte. 2008. Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. *Arch. of Surgery.* **143**(1) 12.
- Makary, M. A., J. B. Sexton, J. A. Freischlag, C. G. Holzmueller, E. A. Millman, L. Rowen, P. J. Pronovost. 2006. Operating room teamwork among physicians and nurses: teamwork in the eye of the beholder. *J. of the Amer. College of Surgeons.* **202**(5) 746-752.

-
- Moreland, R. L. 1999. Transactive memory: Learning who knows what in work groups and organizations. L. Thomson, D. Messick, J. Levine, eds. *Sharing Knowledge in Organization*. Lawrence Erlbaum, Hillsdale, NJ.
- Nembhard, I. M., A. C. Edmondson. 2006. Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *J. of Organ. Behav.* **27**(7) 941-966.
- Nemeth, C. J. 1986. Differential contributions of majority and minority influence. *Psych. Rev.* **93**(1) 23-32.
- Paletz, S. B. F., C. D. Schunn. 2011. Assessing group-level participation in fluid teams: Testing a new metric. *Behav. Res. Methods.* **43**(2) 522-536.
- Pisano, G. P., R. M. J. Bohmer, A. C. Edmondson. 2001. Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Sci.* **47**(6) 752-768.
- Pullon S., Competence, respect and trust: Key features for successful interprofessional nurse-doctor relationships. 2008. *J. of Interprofessional Care.* **22**(2) 133-147.
- Reagans, R., L. Argote, D. Brooks. 2005. Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Sci.* **51**(6) 869-881.
- Schroder, H., M. Driver, S. Streuferi. 1967. *Human information processing*. Holt, Rinehart and Winston, New York.
- Siemsen, E., A.V. Roth, S. Balasubramanian. 2008. How motivation, opportunity, and ability drive knowledge sharing: The constraining-factor model. *J. of Oper. Management.* **26**(3) 426-445.
- Staats, BR. 2012. Unpacking Team Familiarity: The Effects of Geographic Location and Hierarchical Role. *Production and Oper. Management.* **21**(3) 619-635.
- Tucker, A. L., A. C. Edmondson. 2003. Why hospitals don't learn from failures: organizational and psychological dynamics that inhibit system change. *California Management Review.* **45**(2) 55-72.
- Tucker, A. L., I. M. Nembhard, A. C. Edmondson. 2007. Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management Sci.* **53**(6) 894-907.
- Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Tushman, M. L., D. A. Nadler. 1978. Information Processing as an Integrating Concept in Organizational Design. *Acad. of Management Rev.* **3**(3) 613-624.
- Van Knippenberg, D., C. K. W. De Dreu, A. C. Homan. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology* **89**(6) 1008-1022.
- Van de Ven, A. H., A. L. Delbecq, R. Koenig Jr. 1976. Determinants of coordination modes within organizations. *Amer. Sociological Rev.* **41**(2) 322-338.
- Wegner, D. M. 1986. Transactive memory: A contemporary analysis of the group mind. G. Mullen, G. Goethals, eds. *Theories of Group Behavior*. Springer-Verlag, New York, 185-208.
- Weick, K. E., K. H. Roberts. 1993. Collective mind in organizations: Heedful interrelating on flight decks. *Admin. Sci. Quart.* **38**(3) 357-381.

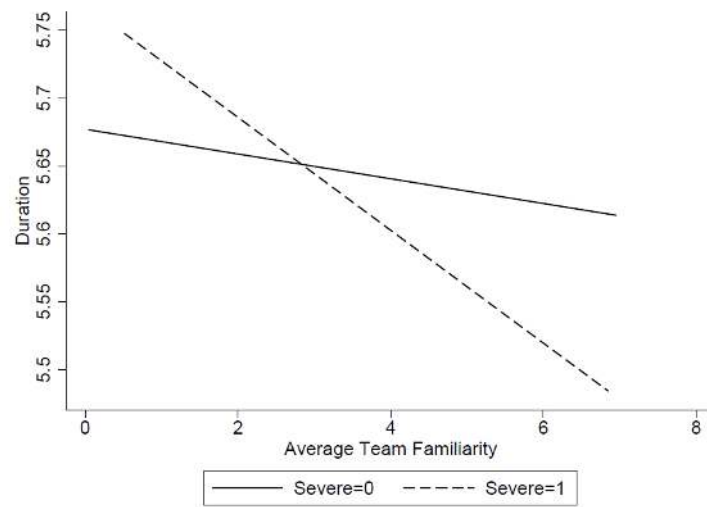
Figure 1. Moderating role of task complexity

Table 1. Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9
1. Duration	5.657	0.268	3.367	6.768	1								
2. Team Familiarity Dispersion	278.197	213.761	0	1097	0.0902**	1							
3. Bottleneck-Pair	2.501	1.118	1	4	0.0509**	-0.0267*	1						
4. Severe Familiarity Percentage	0.090	0.031	0	0.431	-0.0649**	-0.3273**	0.2116**	1					
5. Average Team Familiarity	4.990	1.030	0	6.928	0.0222	0.6274**	0.2910**	-0.4432**	1				
6. Individual Average Direct Experience	5.076	1.376	0	7.227	0.2419**	0.5939**	0.0906**	0.2458**	-0.2686**	1			
7. Team Size	5.414	0.785	3	8	0.0442**	0.0519**	0.3323**	0.0437**	0.0104*	-0.0734**	1		
8. Severe	0.283	0.224	0	1	-0.0925**	-0.0365	0.0536*	-0.0907**	0.0638**	-0.2034**	0.0670**	1	
9. Mild	0.221	0.421	0	1	-0.0185	-0.0136	0.0157	-0.0165	-0.0162	0.0419**	-0.1159**	-0.1793**	1

+, * and ** denote significance at 10%, 5% and 1% levels respectively.

Table 2. Regression of Team Familiarity on Surgery Duration

Variable	Duration					
	Model: (1)	(2)	(3)	(4)	(5)	(6)
Team Familiarity Dispersion		0.0001** (0.0000)		0.0001** (0.0000)	0.0001** (0.0000)	0.0001** (0.0000)
Bottleneck-Pair			0.0113** (0.0032)	0.0170** (0.0036)	0.0176** (0.0036)	0.0178** (0.0036)
Severe Familiarity Percentage					-0.4106** (0.1141)	-0.4087** (0.1152)
Average Team Familiarity	-0.0290** (0.0092)	-0.0354** (0.0120)	-0.0390** (0.0099)	-0.0413** (0.0141)	-0.0460** (0.0142)	-0.0425** (0.0142)
Average Team Familiarity x Severe						-0.0284* (0.0124)
Individual Average Direct Experience	-0.0676** (0.0114)	-0.0675** (0.0134)	-0.0703** (0.0124)	-0.0515** (0.0144)	-0.0470** (0.0146)	-0.0481** (0.0145)
Team Size	0.0386** (0.0060)	0.0379** (0.0061)	0.0418** (0.0064)	0.0458** (0.0065)	0.0462** (0.0065)	0.0469** (0.0066)
Severe	0.0243** (0.0043)	0.0237** (0.0043)	0.0250** (0.0043)	0.0243** (0.0043)	0.0243** (0.0043)	0.1614** (0.0600)
Mild	-0.0124* (0.0060)	-0.0129* (0.0060)	-0.0112+ (0.0060)	-0.0119* (0.0060)	-0.0118* (0.0060)	-0.0114+ (0.0060)
Constant	5.3009** (0.0812)	5.3158** (0.0806)	5.2591** (0.0822)	5.2637** (0.0819)	5.2985** (0.0825)	5.2812** (0.0827)
Observations (N)	6128	6128	6128	6128	6128	6128
Adjusted R ²	0.477	0.481	0.481	0.484	0.486	0.487
Increase in Adjusted R ²		0.004**	0.004**	0.003**	0.002**	0.001*
BIC	-1512.742	-1518.935	-1517.899	-1531.114	-1533.067	-1535.958
AIC	-1700.992	-1707.909	-1712.873	-1732.811	-1741.482	-1751.096
Quarter Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Lead Surgeon Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Procedure Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Day of the Week Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

+, * and ** denote significance at 10%, 5% and 1% levels respectively

Table 3. Policy Comparison

	Policy_dispersion				Policy_stable			
	Quarter 1	Quarter 2	Quarter 3	Average	Quarter 1	Quarter 2	Quarter 3	Average
Team Familiarity Dispersion	-22.43%	-64.25%	-68.56%	-51.75%	21.01%	32.82%	22.24%	25.36%
Average Team Familiarity	-19.35%	-61.05%	-54.91%	-45.10%	26.79%	71.43%	68.30%	55.51%
Bottleneck-Pair	-1 (31.35%)	-3 (93.17%)	-3 (108.77%)	-2.33 (77.77%)	0 (-8.93%)	1 (-41.04%)	0 (-22.85%)	0.33 (-24.27%)
Overall Duration	-1.06%	-4.11%	-5.73%	-3.63%	-1.00%	-0.27%	-0.14%	-0.47%