

Motivating Process Compliance through Individual Electronic Monitoring: An Empirical Examination of Hand Hygiene in Healthcare

Bradley R. Staats
University of North Carolina at Chapel Hill
Campus Box 3490, McColl Building
Chapel Hill, NC 27599-3490
Tel: 919.962.7343
bstaaats@unc.edu

Hengchen Dai
Washington University in St. Louis
Campus Box 1156, One Brookings Drive
St. Louis, MO 65130
Tel: 314.935.4198
hdai@wustl.edu

David Hofmann
University of North Carolina at Chapel Hill
Campus Box 3490, McColl Building
Chapel Hill, NC 27599-3490
Tel: 919.962.7731
David_Hofmann@kenan-flagler.unc.edu

Katherine L. Milkman
The Wharton School
Philadelphia, PA 19104
Tel: 215.573.9646
kmilkman@wharton.upenn.edu

Acknowledgments

We are grateful to many individuals at Proventix, including David Sellers, Harvey Nix, and Elizabeth Pembleton, whose assistance made this project possible. We thank Christian Terwiesch for his helpful comments as we developed the manuscript. We acknowledge the developmental assistance of the DE, AE, and reviewers in improving this paper throughout the review process. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG034532. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Any errors are the responsibility of the authors.

Abstract

The design and use of standard processes are foundational recommendations in many operations practices. Yet, given the demonstrated performance benefits of standardized processes, it is surprising that they are often not followed consistently. One way to ensure greater compliance is by electronically monitoring the activities of individuals, although such aggressive monitoring poses the risk of inducing backlash. In the setting of hand hygiene in healthcare – a context where compliance with standard processes is frequently less than 50% and where this lack of compliance can result in negative consequences – we investigated the effectiveness of electronic monitoring. We did so using a unique, RFID-based system deployed in 71 hospital units. We found that electronically monitoring individual compliance resulted in a large, positive increase in compliance. We also found that there was substantial variability in the effect across units and that units with higher levels of pre-activation compliance experienced increased benefits from monitoring relative to units with lower levels of pre-activation compliance. By observing compliance rates over three-and-a-half years, we investigated the persistent effects of individual monitoring and found that compliance rates initially increased before they gradually declined. Additionally, in multiple units, individual monitoring was discontinued, allowing for an investigation of the impact of removing the intervention on compliance. Surprisingly, we found that after removal, compliance rates declined to *below* pre-activation levels. Our findings suggest that although individual electronic monitoring can dramatically improve process compliance, it requires sustained managerial commitment.

Key Words: Process Compliance, Hand washing, Electronic Monitoring, Empirical Operations

1. Introduction

Since at least the publication of Frederick Taylor's (1911) seminal work, the creation of and compliance with standard processes have been key drivers of quality and productivity improvement within operations (Bohn 2005). Examples of standardized processes that have driven quality and productivity improvements include Toyota's approach to standardizing work and then enhancing it through continuous improvement (Shah et al. 2008; Staats, Brunner and Upton 2011), and the ISO 9000 system, which creates standard processes that can be adhered to and then improved over time (Corbett, Montes-Sancho and Kirsch 2005; Levine and Toffel 2010). Many other industries have seen operational performance improve as a result of standardized processes, such as software (Harter, Krishnan and Slaughter 2000), long-haul trucking (Baker and Hubbard 2004), and retail (DeHoratius and Raman 2008). Within healthcare, standard processes have become a major focus due to persistent patient safety issues and rising costs (IOM 1999). For example, the use of core processes in healthcare (e.g., for heart failure & pneumonia) commissioned by the Centers for Medicare and Medicaid Services in the US has been shown to yield higher quality and more efficient care (Senot et al. 2015; Andritsos and Tang 2014).

Given the demonstrated performance benefits frequently associated with adopting standard

processes, it is somewhat surprising that an ongoing challenge has been achieving *consistent use* of such processes. One way to encourage the consistent use of standard processes might be through monitoring individuals' behavior either by other individuals (e.g., managers) or with technology (Ouchi 1979; Nagin et al. 2002). Electronic monitoring – given its low cost and omnipresent nature – seems to offer significant benefits in this respect (Kallman 1993; Aiello and Kolb 1995). In this investigation, we explore the effectiveness of electronic monitoring in overcoming the challenge of process compliance using the context of hand hygiene in healthcare. Since the work of Ignaz Semmelweis in the 1800s, healthcare professionals have known that hand washing reduces hospital-acquired infections. Despite this widespread knowledge, hand hygiene compliance rates often fall below 50% (CDC 2002). Recently emerging technologies that electronically monitor this behavior provide a reliable and unobtrusive approach to measuring hand hygiene compliance (Boyce 2011).

For our investigation, we relied on data from Proventix, a company that uses a radio frequency identification (RFID)-based system to monitor individual healthcare workers' hand hygiene compliance in hospitals. Our data included over three-and-a-half years of compliance data from caregivers in 71 hospital units at 42 hospitals where individual electronic monitoring was deployed (encompassing approximately 20 million hand hygiene compliance opportunities). Our results showed that, on average, caregivers exhibited a large and significant increase in hand hygiene compliance after individual electronic monitoring was activated. We also found that there was significant variability in the activation effect across units and that units with higher levels of pre-activation compliance experienced larger gains from individual monitoring than did units with lower levels of pre-activation compliance. Our analysis of the persistence of compliance behavior over time revealed that the benefits of monitoring increased for nearly two years before they eventually started to degrade. Finally, we considered the implications of terminating individual electronic monitoring after an extended period of oversight. Our examination of termination effects in nine units at six hospitals where we were able to view compliance after monitoring was discontinued, showed that electronic monitoring did not lead to habit formation. Instead, surprisingly, we found that compliance after the removal of monitoring dropped *below* pre-activation levels.

Our results make several contributions to both theory and practice. First, we documented the initial, large positive effect of individual electronic monitoring across a substantial number of organizations using archival data, confirming that the benefits of individual monitoring dominate the risk of employee reactance (Frey 1993), at least on average. Second, we found that there is significant variability in the aforementioned effect, highlighting the need to understand what factors may drive this heterogeneity. Third, we showed that units with higher pre-activation hand hygiene compliance experienced a larger boost in compliance from monitoring than units with lower pre-activation compliance. Fourth, we provided tests of the long-term effects of monitoring on compliance, revealing a long, initial period of beneficial increases in hand hygiene compliance followed by a gradual tapering-off effect. Fifth, and perhaps most surprisingly, we found that compliance after the removal of individual monitoring was lower than compliance before monitoring was introduced. This result provides support for a crowding out perspective on motivation (Deci 1975) whereby employees' internal motivation for compliance was replaced by external forces associated with monitoring, highlighting the limitations of merely monitoring desired behavior as a means of producing lasting habits. More generally, the growing availability of monitoring technologies, and the trace data they create, makes understanding the link between process compliance and electronic monitoring one that is important theoretically and practically.

2. Process Compliance and Individual Electronic Monitoring

Despite the operational benefits that have been shown to result from standard processes (Taylor 1911; Smiddy and Naum 1954), compliance rates are often low (Martin et al. 2013). For example, within healthcare, standard processes have been developed to improve patient care. Yet, according to US government statistics (<http://www.medicare.gov/hospitalcompare>), compliance rates vary greatly. The compliance challenge is not new. In fact, Frederick Taylor referenced this problem noting that, "it is only through *enforced* standardization of methods, *enforced* adoption of the best implements and working conditions, and *enforced* cooperation that this faster work can be assured. And the duty of enforcing the adoption of standards and enforcing this cooperation rests with *management* alone" (Taylor 1911: 83).

One way to help ensure that employees adopt and comply with standard processes is monitoring

(Ouchi 1979). Electronic monitoring – due to its low cost and omnipresent nature – seems to offer an ideal way to monitor compliance (Kallman 1993; Aiello and Kolb 1995). Note that electronic monitoring can be done in aggregate for a unit or with respect to the individual. In this paper we are interested in understanding the effect of individual electronic monitoring.¹ Examples of individual electronic monitoring currently in wide use include electronic monitoring of ethical behavior through video cameras in casinos, electronic monitoring of trading activity in investment banks, electronic monitoring of safety steps in manufacturing, and systems to monitor employee movements and actions in numerous industries.

Individual electronic monitoring may improve compliance for several reasons. First, by making individuals feel accountable for their observable actions, the mere presence of electronic monitoring may change behavior. For example, prior work finds that simply making one's actions visible to others may influence different categories of behavior ranging from voter turnout to employee theft in restaurants (e.g., Gerber, Green and Larimer 2008; Yoeli et al. 2013; Pierce, Snow and McAfee 2014). Further, the data produced by electronic monitoring can be used in compliance improvement efforts by providing employees with ongoing feedback, by increasing their rewards for compliance (e.g., through more accurately calculated piece-rate payments), or by increasing non-compliance costs (e.g., termination of employment or social punishment through the enforcement of norms, Nagin et al. 2002). Additionally, the use of monitoring may signal management's commitment to process compliance. Notwithstanding the benefits of electronic monitoring, an important concern about individual electronic monitoring is that aggressive monitoring could be perceived as invasive and as a signal of distrust by management, which could result in reactance or reduced compliance (Frey 1993; Bernstein 2012).

Thus, monitoring creates the potential for competing mechanisms. The individual setting is likely to determine which of these forces from electronic monitoring—its ability to generate transparency, feedback, incentives, and a signal of management commitment, or its risk of arousing reactance—dominates. Factors determining employees' response to electronic monitoring may include the

¹ For the remainder of the paper we will use the terms individual electronic monitoring and electronic monitoring interchangeably. If we wish to refer to electronic monitoring at an aggregate-level we will explicitly do so.

intrusiveness of the monitoring and the desired action's consistency with employees' professional and prosocial goals. In a setting where electronic monitoring is not intrusive and the compliance behavior is in line with professional and prosocial goals (e.g., to provide better patient care), we would expect:

HYPOTHESIS 1: *Process compliance will increase, on average, after individual electronic monitoring is implemented.*

Prior work examining monitoring (e.g., Gerber et al. 2008; Yoeli et al. 2013; Pierce et al. 2014) finds that although there may be an average effect of deploying monitoring, there is also significant heterogeneity across contexts. We next explore one potential cause of such variability – pre-activation compliance. There are at least three reasons to predict that organizations with higher pre-activation compliance will see a greater increase after electronic monitoring implementation than will organizations with lower pre-activation compliance: (1) social norms; (2) organizational design; (3) management attention. First, social norms represent the behavior that is considered “typical” in a given context. Social norms have been shown to have a strong effect on individual behavior (Cialdini, Reno and Kallgren 1990; Pierce and Snyder 2008) such that individuals respond in kind to their environmental norms. Units with higher pre-activation compliance may have stronger compliance norms signaling the importance of the targeted behavior. In such units, individuals may view compliance as a key part of the job and therefore embrace tools designed to help with compliance (see, Klein and Knight 2005, for a similar argument for innovation norms and innovation use, p. 245). For units with lower compliance norms, individuals may be less committed to compliance. Thus, individuals may not be as engaged with electronic monitoring or not as strongly value the necessity of compliance, decreasing the intervention's effectiveness.

Further, Cialdini et al. (1990) noted that, “norms should motivate behavior primarily when they are activated (i.e., made salient or otherwise focused on)” (p. 1015). When individual monitoring is deployed, process compliance as well as the existing norms of compliance behavior is made more salient, prompting individuals to behave in line with compliance norms in their environment. This process may create a “magnetic middle” effect (Goldstein, Cialdini and Griskevicius 2008) whereby employees in an organization with strong compliance norms who previously complied below the norm level may improve

their compliance rates, whereas employees in an organization with weak compliance norms who previously complied above the norm level may decrease their compliance rates to behave consistently with the salient norm. Through this process, organizations with stronger compliance norms (potentially indicated by higher pre-activation compliance) may benefit more from activating electronic monitoring.

Organizational design is a second reason why higher pre-activation compliance may increase the impact of electronic monitoring on individual compliance. Organizational design may include the members that are selected for the organization, as well as the operating environment's physical design. With respect to the former, higher pre-activation compliance may indicate that organization members are more prosocially inclined (e.g., care more about patient's health) or have a greater appreciation of the general need for compliance. If this is the case then these individuals may view monitoring as creating an opportunity to learn and improve. This would, in turn, increase the positive effect of electronic monitoring on the targeted behavior (c.f. Song et al. 2015). Physical design may also influence the impact of electronic monitoring: organizations with higher pre-activation compliance may have better-placed equipment which allows them to maximize the benefits that may accrue from monitoring.

Finally, higher pre-activation compliance may indicate greater management attention to process compliance. Greater management attention may indicate a willingness and ability to use the information that is created from electronic monitoring more effectively. These managers may share the knowledge to motivate workers or help them learn. Moreover, with greater management attention then individuals may recognize the need to respond more strongly to the deployment of individual electronic monitoring.

Although the above reasons suggest that greater pre-activation compliance will lead to a larger effect from electronic monitoring deployment there is at least one boundary condition to note. If compliance rates are extremely high (e.g., close to 100%), then the degree of improvement following deployment could be limited due to ceiling effects. However, in cases where higher compliance is not sufficiently high to create such a ceiling effect, we hypothesize:

HYPOTHESIS 2: *The increase in process compliance due to the implementation of individual electronic monitoring will be greater among organizations with higher pre-activation*

compliance than organizations with lower pre-activation compliance.

2.2 Effects of Individual Electronic Monitoring over Time

Moving beyond the initial implementation of process compliance initiatives like electronic monitoring, it is imperative to understand what happens over the ensuing months and years. Does individual electronic monitoring maintain its effectiveness, or does it gradually lose its salience, with its impact diminishing over time? Despite the importance of understanding the long-term impact of operational interventions (Gino and Pisano 2008; Boyce 2011; Rogers and Frey 2014), the kind of longitudinal data necessary for examining such issues is rarely available to researchers.

When considering the potency of electronic monitoring effects over time, competing predictions are possible. One reason electronic monitoring may maintain or increase its potency is that it is always present. Its constant presence may suffice to maintain a constant treatment effect if employee expectations and management structures aligned with the deployment of monitoring remain unchanged. Moreover, electronic monitoring may increase compliance due to the development of habitual responses. Habits are psychological inclinations to engage in a past behavior (Neal et al. 2012). Research on habituation suggests that by continually engaging in the same set of activities, with high frequency, in similar contexts, without a change in stimuli, an individual is likely to build a habit such that a stimulus-response cycle is automatically repeated (Wood and Neal 2007). This implies that the continuous nature of monitoring could well result in the habituation of the monitored behavior. As the habitual responses become ingrained in the automatic routines of workers, compliance should increase over time, albeit at a decreasing rate as individuals approach 100% compliance. Supporting this perspective, single and dual unit studies of hospital units adopting hand hygiene monitoring technologies find that process compliance may improve and hold steady for up to a year and a half (Armellino et al. 2011; Walker et al. 2014).

Other streams of research, however, suggest that monitoring may lose its potency over time, resulting in a gradual decline in compliance after rollout. First, individuals may become desensitized to a stimulus over time (Thompson and Spencer 1966; Haselhuhn et al. 2012). Similarly, signal detection theory suggests that over time the salience of signals typically decays with the associated behavior

experiencing a commensurate decline (MacMillan 2002). Second, it is possible that individuals' expectations about compliance could change over time, perhaps due to management's changing enforcement or emphasis, and as a result compliance behaviors may gradually decrease.

Although the different perspectives suggest opposing effects in terms of the persistence of electronic monitoring on compliance, we propose that these effects may dominate over different time periods. Namely, we propose that initially, the positive effect dominates as habits are developed, but that as discussed previously, monitoring loses its saliency over time. Therefore, we propose:

HYPOTHESIS 3: *Ceteris paribus, process compliance will exhibit an inverted-U shaped relationship over time after individual electronic monitoring is implemented such that process compliance will increase first and then decline over time*

Hypothesis 3 pertains to the persistence of individual electronic monitoring on process compliance. But, another important and distinct question is what happens to compliance behavior if monitoring is discontinued altogether. On the one hand, if habits are formed, then the discontinuation of monitoring should have little or no impact on behavior. Although the strength of the habit could gradually weaken over time with the removal of monitoring, if the behavior were truly habituated then an immediate and significant degradation would not be expected after monitoring is turned off. This result would be consistent with Charness and Gneezy (2009) who found that incentives for exercise accrued long-term benefits even after the treatment was removed. Similarly, Allcott and Rogers (2014) found that informing people of neighbors' energy usage had a lasting effect for years after such mailings were discontinued – although the effect did dissipate over time. Further, if the initial investment in electronic monitoring by management signaled a commitment to behavior change or changed workers' beliefs about the importance of behavior change, higher levels of compliance should persist after monitoring removal.

On the other hand, two factors may lead individuals to decrease their compliance after electronic monitoring is removed. First, electronic monitoring removal may signal to workers a lack of managerial commitment to encouraging this particular behavior, resulting in an updating of employee expectancies (i.e., process compliance is no longer a valued behavior) and a subsequent decline in compliance behavior. Second, there is a risk of motivational crowding out due to the implementation of electronic

monitoring (Deci 1975; Bénabou and Tirole 2003). Crowding out refers to the substitution of an external motivation for what was once done for internal purposes. One primary mechanism leading to internal motivation is a desire to feel competent at one's work (Deci and Ryan 2002). Thus, an individual may choose to comply with a standard process, believing that it is appropriate in a given circumstance. The risk is that the introduction of electronic monitoring may change an individual's decision calculus. Now, instead of complying with an activity for internal reasons, the individual may have her motivation shifted externally and thus comply for external purposes (e.g., appearing responsible in front of managers). If the external motivator is removed, as is the case with the discontinuation of electronic monitoring, then both internal and external motivation would be gone, resulting in a compliance decline for the given behavior. In fact, if crowding out were to occur, then not only might compliance fall, as compared to the time period right before the deactivation of electronic monitoring, but also a strong form of the prediction would suggest that compliance may fall below the level before electronic monitoring deployment.

Although the different theoretical perspectives suggest opposite effects of terminating monitoring on compliance, we hypothesize that the negative effect from termination will dominate:

HYPOTHESIS 4: *Process compliance will decline below the pre-activation level of compliance after individual electronic monitoring is terminated.*

3. Data and Empirical Strategy

3.1 Setting

We investigate our hypotheses with data from Proventix, a company that uses RFID technology to help healthcare providers improve hand hygiene compliance. Proventix's system involves distributing RFID badges to hospital caregivers who then wear their personalized, active RFID badges along with their standard hospital identification. The company also installs RFID readers throughout the hospital unit. The RFID badges then communicate wirelessly with a network of sensors connected throughout the monitored area. In addition, communication units are installed on hand hygiene dispensers (both soap and sanitizer dispensers) within a focal hospital unit. As a result of the system, both the date and time when a caregiver enters the area monitored by a given sensor as well as whether or not the caregiver uses a dispenser are recorded. The system also records all dispenses by a dispenser, even if it is used by someone

without a badge. The information is then transmitted back to Proventix's central servers. Further, each communication unit contains an LCD display that delivers personalized messages to caregivers wearing RFID badges (e.g., personalized feedback on their hand hygiene compliance rate, up-to-date weather forecasts, and/or professional and educational messages).

Proventix follows a consistent process in rolling out their system. First, for each hospital unit, Proventix installs an RFID reader to each of the hand hygiene dispensers. Proventix strives to fit its system into the existing infrastructure of the hospital unit, so new dispensers are not typically deployed and the location of existing dispensers is not normally changed. The RFID readers, called communications units, collect dispensing data, measure caregiver movement through the care area, and provide point of care messaging intended to inform and engage caregivers as they cleanse their hands. Once the communication units have been installed, Proventix is able to track dispenser usage in a hospital unit. We refer to this date as the "installation date." At the installation date, aggregate electronic monitoring is possible, but individual electronic monitoring is not possible, as individuals are not yet badged and so dispenser usage data is not linked to individual caregivers. During the time period after communication units have been installed, Proventix seeks to gather baseline dispenser usage data in order to create a reference point for evaluating future compliance behavior. After the installation date, but prior to the full-scale rollout of individual monitoring, a small subset of caregivers, who are selected by the hospital units, are designated as pilot testers. These individuals receive an active RFID badge and wear their badge at work. At this stage, Proventix does not share pilot testers' compliance statistics with hospital management.

Proventix's next stage in the implementation process is to assign badges to caregivers and commence individual monitoring. This stage is typically rolled out several months after the "installation date". We refer to this date hereafter as the "activation date." From the activation date on, Proventix monitors *individual* caregivers' hand hygiene performance. Before the activation date, Proventix holds information sessions and distributes handouts to explain how the monitoring system works and how hand hygiene compliance rates are calculated.

Although Proventix recommends to hospitals that as many caregivers as possible receive RFID badges, it is ultimately the hospital's choice as to which caregivers receive a badge. Traditionally all nursing staff and patient care technicians within a unit receive a badge. However, that still leaves many caregivers who do not receive badges. According to Proventix, not providing badges to employees in certain roles is typical because those individuals spend only a part of their time in a given unit (e.g., temp nurses, nurses from other units, residents on rounds, dieticians, housekeeping staff, transporters, therapists, physicians). Proventix records roles for individuals who are assigned badges. We confirmed that most caregivers who received a badge were nurses (71.5%) or patient care technicians (12.08%).

During our study period from February 2010 to August 2013, 42 hospitals (including 71 hospital units) rolled out Proventix's monitoring system² and twelve of these hospital units at nine hospitals discontinued electronic monitoring. We note that, in our sample, a majority of the hospital units that removed individual monitoring did so because the funding for monitoring came from an outside grant and at the expiration of the grant they either did not have or chose not to deploy internal funding to continue the service. For three units, Proventix stopped tracking dispenser usage after reclaiming caregivers' badges. However, for the remaining nine hospital units, Proventix continued to record the date and time of every dispenser used for several months after RFID badges were collected. These nine hospital units provided us with an opportunity to examine whether and how deactivating electronic monitoring affects dispenser usage. The date when Proventix collected RFID badges from a given hospital unit is hereafter referred to as the "deactivation date." Figure 1 demarcates the three dates relevant to our analysis.

Proventix has developed a standardized measurement system to identify situations when hand hygiene is expected. When a caregiver stays in a patient's room for 20 seconds or longer – a length of time that Proventix has deemed from expert consultation to be sufficient for hand hygiene to be clinically relevant – the room entry and room exit associated with the stay in question are classified as hand hygiene opportunities. Guidelines from groups such as the WHO recommend that hand hygiene be conducted on

² Proventix provided information on one additional hospital unit. However, in this unit, all observations happened in cleaning rooms. We excluded this unit from all analyses, but our results were robust to including these observations.

both entry and exit. By default, a 90-second “window” is allocated for caregivers entering a patient’s room (60 seconds before entry and 30 seconds after entry are included in this window) and exiting a patient’s room (30 seconds before exit and 60 seconds after exit are included in this window) to sanitize their hands in order to be deemed compliant with a given hand hygiene opportunity. See Appendix A for more details about the hand hygiene measurement system.

An important question is what additional changes with respect to hand hygiene occur when Proventix activates individual monitoring. First, to accurately capture baseline hand hygiene compliance rates, Proventix recommends its clients *not* to make other practice changes at the time of system activation. As noted by a Proventix executive, “Hospital education is primarily technical at the time of badging and is intended to answer basic questions and establish caregiver confidence in the system.” The executive referred to technical training as an opportunity to reinforce the importance of wearing the badge properly to make sure that caregivers receive “credit” for their hand hygiene compliance. Proventix’s standard practice is to distribute to the client the first performance reports after a 30-day post-activation compliance baseline is established. Managers have the ability to review caregivers’ compliance behavior within these 30 days but are encouraged not to introduce interventions such as compliance-based rewards and goal-setting until after the baseline period has ended. Examining data provided by Proventix describing interventions eventually deployed by some hospital units (summarized in the next paragraph) we find that none of the units in our sample implemented incentives in the first 60 days post-activation and only five units implemented performance feedback posting within the first 30 days post-activation.

After a baseline of compliance use is established and technical training has been completed, Proventix works with its hospital clients to structure interventions that may further increase compliance rates. Proventix classifies these interventions into five categories: performance feedback (e.g., individual hand hygiene compliance rates are openly posted), goal-setting (e.g., a group compliance goal is posted in the unit), competition (e.g., unit staff are divided into teams that compete to improve their compliance), leadership (senior level leaders demonstrate active and visible participation in the program), and incentives (e.g., goal-based financial incentives are set based on hand hygiene compliance rates).

Retrospectively, Proventix sought to identify the presence and timing of these interventions and was able to do so for 50 hospital units. We do not include controls for these interventions in the main analyses we reported in Section 4; however, we explore them further in Section 5.1. We note that the results for all hypotheses remain meaningfully unchanged in direction and significance with their inclusion.

The setting we study has a number of advantages for an exploration of the impact of individual electronic monitoring on compliance. First, 71 hospital units rolled out electronic monitoring over the course of three years, and activation dates varied across hospital units. Activation dates were distributed between May 2010 and August 2013 (see Appendix B for more details on activation timing). This staggered roll-out allows us to isolate the effects of activating individual electronic monitoring on dispenser usage from the effects of other potentially confounding factors, such as general trends over time in hand hygiene compliance and the rollout of various public health campaigns. Second, the large number of hospital units involved in this study allows us to examine whether and why there is variability in the monitoring effect across hospital units. Third, the three-year longitudinal panel data allow us to explore whether the initial effects of activating electronic monitoring are strengthened over time or instead decay. The fourth advantage of this setting is that for some hospitals, electronic monitoring was discontinued, allowing us to evaluate the effects of the removal of monitoring on compliance behavior.

3.2 Data

To answer our research questions we would ideally have process compliance data for individuals working within a hospital during the three stages in Figure 1: (1) Pre-activation; (2) Post-activation; and (3) Post-deactivation. Unfortunately, it was not possible to precisely examine compliance for all caregivers at the first and third stages since it was the use of individual electronic monitoring that permitted us to examine all individuals' compliance decisions. To examine the effects of activating and deactivating individual electronic monitoring (Hypotheses 1, 2, and 4), we used data on dispenser usage and explored process compliance at the hospital unit-level. Since we were unable to distinguish usage generated by caregivers who were eventually badged from usage generated by others in the pre-activation period, we examined how total usage by all users changed following the activation and deactivation of

monitoring. Prior to badge activation and post badge deactivation, Proventix tracked use of hand hygiene dispensers. Therefore, after controlling for other factors, an increase in dispenser use within units following the rollout of individual electronic monitoring would indicate an increase in hand washing and thus process compliance. Similarly, examining how dispenser use changed after monitoring was removed provided insight into our hypothesis regarding the termination of individual electronic monitoring.

In addition to the dispenser-level data, we also had information on individual compliance about caregivers who ever received a badge during our study period. After individual monitoring was turned on, we investigated whether individuals complied each time they had a hand hygiene opportunity. We used these data to examine the persistence question (Hypothesis 3). Below we provide more detail about each data set (unit- and individual-level compliance data).

3.2.1 Unit-Level Process Compliance Data. The first dataset included records from each of the 71 hospital units on each day prior to August 31, 2013 (the end of our study period). We observed two data points per day per unit: (1) a record of how many times dispensers in this unit were used each day by someone wearing an RFID badge (“badged users”), and (2) a record of how many times unit dispensers were used on a given day by someone who was not wearing an RFID badge (“unbadged users”). Before each department’s activation date, only pilot testers wore badges, and the vast majority of users were unbadged. After the electronic monitoring activation date, some caregivers received badges and became badged users, and unbadged users include caregivers who do not have a badge, visitors, and patients.

Our final unit-level dataset included 35,552 observations for 71 hospital units in 42 hospitals. Twenty-six hospitals were located in Alabama; the remaining 16 were spread across eight other U.S. states (Arizona, California, New Jersey, Pennsylvania, Tennessee, Texas, Virginia, and Wisconsin). Thirty-one of these hospitals were located in urban areas; eleven were in rural areas. The number of staffed beds across hospitals ranged from 25 to 1,097 ($M = 270$, $SD = 233.16$).

3.2.2 Caregiver Compliance Data. Our second data set tracked each of the 5,247 unique caregivers working in the 71 units at 42 hospitals using the electronic monitoring system who received a RFID badge prior to August 31, 2013. It included one observation per hand hygiene opportunity per caregiver

including information about: (a) the date and time when a hand hygiene opportunity occurred, (b) whether or not the caregiver in question sanitized her hands, and (c) how many times the caregiver in question sanitized her hands during the 90-second compliance window associated with a hand hygiene opportunity.

We used these data to examine how compliance changed in the long term after the activation of individual monitoring for *all* users (Hypothesis 3). Our sample consists of observations that occurred after the activation date of the corresponding unit ($N_{hand\ hygiene\ opportunities} = 19,585,738$). Prior to releasing these data, Proventix assigned anonymous identification numbers to caregivers and dispensers.

3.3 Variables

3.3.1 Dependent Variables. *Total daily usage per unit bed.* We first calculated the total number of times all dispensers in a given hospital unit were used on a given day, including usage generated by both badged users and unbadged users. We divided this number by the number of beds in a unit to obtain our dependent variable, *total daily usage per unit bed*, in order to adjust dispenser usage for unit size. *Total daily usage per unit bed* ranged from 0 to 579 with an average value of 48 and a standard deviation of 39.

Compliance. We operationalized *compliance* with a dichotomous indicator variable recording whether a caregiver washed her hands during a given hand hygiene opportunity (either a patient room entry or exit). If a caregiver washed her hands, this indicator was set to one; otherwise, it was set to zero. The mean compliance rate in our hand hygiene compliance dataset was 45%, which is comparable to the 39% compliance rate across hospitals reported by the World Health Organization (WHO 2009).

3.3.2 Independent Variables. *Activation.* For each unit on each day, *activation* was coded using an indicator variable recording whether individual monitoring was active. *Deactivation.* For each unit on each day, the indicator variable *deactivation* was set equal to one if Proventix had reclaimed the RFID badges on or before the day in question from the hospital unit; otherwise, it equaled zero. For 62 hospital units that did not reclaim badges before their last observations in our data (including three units that discontinued individual monitoring during our study period but stopped tracking dispenser usage once badges were reclaimed), *deactivation* always equaled zero.

Months since activation: For each observation in our caregiver compliance data, we calculated the number of months elapsed since the focal hospital unit's activation date up to the day in question.

3.3.3. Moderator. Pre-activation usage. For each unit, we calculated the average value of its total daily usage *per unit bed* during the pre-activation period.

3.3.4 Control Variables. To control for time trends we included a set of indicators for each month-year pair in our data. We also controlled for the day of the week when an observation occurred. Based on past research examining how work environments affect caregivers' hand hygiene compliance (Dai et al. 2014), we included additional controls in our analysis of the caregiver compliance data; however, our results were robust if we did not control these variables. Specifically, we controlled for the length of the stay in a patient's room associated with a given hand hygiene opportunity, an indicator for whether the hand hygiene opportunity in question involved a room entry (as opposed to a room exit), the hour of the day when the hand hygiene opportunity occurred, the time (in hours) elapsed since the start of a given caregiver's shift leading up to a given hand hygiene opportunity, as well as the time (in hours) between the preceding shift and the current shift when a given hand hygiene opportunity occurred. In addition, past research indicates that individual behavior may change in the presence of others. The effect is unclear because others could provide either a distraction that lowers compliance or social pressure that encourages compliance (Mas and Moretti 2009; Chan, Li and Pierce 2014). For each room entry and room exit that occurred during the post-activation period, we were able to identify whether other caregivers were in the same patient room when the focal caregiver entered or exited the room. Thus, we controlled for the presence of other caregivers in regressions that predict individual compliance in the post-activation period.

Finally, we included an indicator variable for each hospital unit in the unit-level analyses and each caregiver in the caregiver analyses. Thus, all analyses controlled for time-invariant aspects of hospital units (e.g., medical services provided) or individual caregivers (e.g., individual propensity to comply). These indicator variables had the effect of controlling for all between-unit or between-caregiver

variance such that our models explored within-unit or within-caregiver variance. Table 1 provides information on each control variable and Table 2 provides summary statistics for all variables.

3.4 Empirical Analysis Strategy

We relied on ordinary least squares (OLS) regressions to analyze our data. Following Pierce et al. (2014), which examined an intervention that was rolled out at different times across multiple locations, we used a difference-in-differences design to test the effects of activating individual electronic monitoring on process compliance. Specifically, the activation of individual electronic monitoring was the treatment on each hospital unit. For a given unit that activated individual monitoring, other hospital units in the post-installation but pre-activation period served as the control group.³ The control groups, which had begun aggregate electronic monitoring, provided a counterfactual for how compliance would have changed in treatment units had they not activated individual electronic monitoring. For the unit-level analyses, our dependent variable was the total daily number of uses of all dispensers in a given hospital unit, i , on a given day, t , adjusted for the number of beds in the unit. The regression specification used to estimate the effects of activating monitoring on dispenser usage was:

$$(1) \quad \text{total daily usage per unit bed}_{it} = \alpha_0 + \alpha_1 * \text{activation}_{it} + \alpha * X_{it} + \lambda_i + \epsilon_{it}$$

where X_{it} represents the vector of the control variables described in Table 1, λ_i represents an indicator for unit i , and ϵ_{it} represents an error term. Hypothesis 1 predicted that α_1 would be positive. To test Hypothesis 2, whether the individual monitoring effect was larger among units with higher pre-activation usage, we added the interaction between *activation* and (mean centered) *pre-activation usage*. Hypothesis 2 predicted that the interaction term coefficient would be positive.

To examine how compliance changed over time after individual monitoring activation, we included all caregivers in our analysis and relied on the following OLS regression:

$$(2) \quad \text{compliance}_{jt} = \beta_0 + \beta_1 * \text{months since activation}_{jt} +$$

³ Seven hospital units' activation dates did not fall between the installation and activation dates of any other units. According to our design, those units did not have a control group. However, if we excluded those units from our analysis, our findings regarding the positive effect of activating monitoring on dispenser usage remained statistically significant ($p < .0001$) with virtually the same effect size as the effect size reported.

$$\beta_2 * \text{months since activation}_{jt}^2 + \boldsymbol{\beta} * X_{jt} + \tau_j + \varepsilon_{jt}$$

where X_{jt} represents the vector of the control variables described in Table 1, τ_j represents an indicator for caregiver j , and ε_{jt} again represents an error term. We mean-centered *months since activation* before we added a squared transformation of this variable to the regression model. Hypothesis 3 predicted that β_1 would be positive and β_2 would be negative. Since this sample only included observations that occurred after the unit's activation date, the indicator for activation was not included in this specification.

Building on specification (1), we added an additional variable, *deactivation*, to examine how dispenser usage changed in response to the deactivation of individual monitoring:

$$(3) \quad \text{total daily usage per unit bed}_{it} = \alpha_0 + \alpha_1 * \text{activation}_{it} + \alpha_2 * \text{deactivation}_{it} +$$

$\boldsymbol{\alpha} * X_{it} + \lambda_i + \varepsilon_{it}$, where X_{it} represents the vector of the control variables described in Table 1, λ_i represents an indicator for unit i , and ε_{it} again represents an error term. *Activation* equaled one when individual monitoring was active, and zero for both the pre-activation period of all hospital units and the post-deactivation period for the units that terminated individual monitoring. Thus, *deactivation* captured the regression-estimated differences in total daily usage per unit bed between the pre-activation period and the post-deactivation period. Hypothesis 4 predicted that α_2 would be negative.

We clustered standard errors at the hospital unit level, as activation and termination of monitoring was implemented at individual hospital units. Our results were robust when we clustered standard errors at the hospital or caregiver level (in the caregiver compliance data).

4. Results

We first examine the effect of initiating individual monitoring before turning in Section 4.2 to the persistence of compliance over time and then to the termination of individual monitoring in Section 4.3.

4.1 The Effects of Individual monitoring on Process Compliance

4.1.1 Unit-Level Analysis. Each unit was observed from at least 27 days before the activation date (-27) to 17 days after the activation date (+17). We therefore begin our examination of the overall effects of

individual monitoring by evaluating the symmetric time window surrounding the activation date from 17 days before the activation date (-17) to 17 days after the activation date (+17).

Figure 2 displays the average daily dispenser usage per unit bed across all of the hospital units in our data set as a function of days since activation. Total daily usage per unit bed increased as soon as individual monitoring was activated. The average total daily usage per unit bed on the hospital unit's date of activation and the 17 days following it was 54.97% higher ($M_{after} = 43.95$, $SD = 1.83$; including the activation date) than the average total daily usage per unit bed in the 17 days preceding the hospital unit's activation date ($M_{before} = 28.36$, $SD = 1.66$), a difference that is both meaningful and statistically significant ($t(33) = 29.67$, $p < 0.0001$). Thus, our results support Hypothesis 1 and do not show evidence of reactance in response to individual monitoring activation.

We conducted regression analyses to determine whether this effect was robust to controlling for other variables expected to affect hand hygiene in healthcare. Table 3 presents the results of a series of OLS regressions predicting daily dispenser usage per unit bed by hospital unit. In Model 1, we tested regression specification (1) from above. This model includes all observations in our final, unit-level daily usage dataset with the exception of observations that occurred after individual monitoring was deactivated (if a given unit terminated individual monitoring during our study period). We observed a large and significant positive effect of activating individual monitoring on daily dispenser usage per unit bed ($p < 0.001$). Compared with the regression-adjusted baseline total daily usage per unit bed of 36.64, activating individual monitoring increased daily dispenser usage by an estimated 44.61%, providing further support for Hypothesis 1 and highlighting the enormous benefits of individual monitoring on compliance.

4.1.2 Badged vs. Unbadged Usage. As noted above, not all caregivers in a hospital received badges. Moreover, non-caregivers (such as patients and visitors) also may have used dispensers. In the pre-activation period, we could not identify which dispenser uses came from caregivers who eventually received badges and were monitored at the individual level and which came from individuals who never received badges. As a result, the most appropriate analysis for testing Hypothesis 1 was to compare dispenser usage by all users before and after the activation of individual electronic monitoring, as

conducted above. This left open the question, however, of how monitoring affected badged versus unbadged users. To begin exploring this question, we examined how dispenser usage by people not wearing a badge (i.e., unbadged usage) changed pre- versus post-activation. If monitoring only influenced badged users' compliance behaviors, then *unbadged* usage would be expected to decline dramatically on the monitoring activation date. This is because Proventix converted a substantial portion of previously unbadged users to badged users upon activating its monitoring system, leaving a smaller number of people as unbadged users who continued to contribute to unbadged usage. To quantify how much unbadged usage should have decreased when monitoring activation shifted some caregivers were badged, we first estimated that the percentage of caregivers who became badged users ranged from 6% to 73% across 71 hospital units with the mean being 27%. Based on these estimates, for each hospital unit, we further estimated how much of total daily dispenser usage in the 17 days prior to monitoring activation came from caregivers who eventually were badged. This step led to our estimate that the decrease in unbadged usage ranged from 0.26 to 41.47 uses per unit bed per day across 71 hospital units (with the mean being 8.67). Figure 2 depicts the actual and counterfactual average daily unbadged usage 17 days before and 17 days after monitoring activation. Interestingly, unbadged usage did not drop precipitously right after the activation of individual monitoring. Further, we predicted unbadged usage using an OLS regression model similar to specification (1), and confirmed that unbadged usage did not significantly differ between the pre-activation and post-activation periods ($\beta = -0.68, p = 0.48$). This is surprising, suggesting that unbadged users also increased dispenser usage after electronic monitoring activation.

To make sure that unusual patient or visitor use of dispensers is not driving the effect of monitoring documented in Sections 4.1.1, we created a dataset that isolated time periods when visitors were unlikely to be in rooms and patients were less likely than usual to be washing their hands: the six hour window between midnight and 6am.⁴ We assumed that dispenser use during this 12am-6am time window came mostly from caregivers (with or without a badge). As Table 4 shows, we found that total

⁴ The results in Table 4 have the same pattern when we (a) use 9pm-6am to proxy for visitor-free hours, (b) examine all observations in the pre-activation and post-activation periods, or (c) estimate changes in total usage during visitor-free hours using regression specification (1).

usage during this time period increased, on average, by 68.9% in the 18-day period after monitoring was activated (including the activation date), compared with the 17-day period prior to monitoring activation. The size of this increase is similar to that observed in the full dataset during the same 17 day window before and after the activation date (i.e., a 62.2% increase, as explained in Section 4.1.1). In addition, after individual monitoring was activated, unbadged usage during the visitor-free window from 12am-6am (which we assume was generated primarily by caregivers without a badge) accounted for 49.3% of total usage; this usage breakdown is similar to what we observed during the entire day: 51.0% of usage was by unbadged users and 49% was by badged users. This likely indicates that after individual monitoring activation, a significant portion of unbadged use within a day in our data is from caregivers without a badge, as opposed to visitors and patients. We cannot be certain, but we have every reason to believe that visitors and patients account for a small proportion of dispenser usage before monitoring activation too.

In an additional analysis designed to gain traction on the extent to which our findings on total usage are driven by badged versus unbadged users, we classified the 71 hospital units into three groups based on the percentage of their total daily dispenser usage that was attributable to badged users in the 18 days following their activation date (including the activation date). For hospital units in each category, we calculated the change in total daily usage per bed between the 17 days prior to and the 18 days after a unit's activation date. We found that the 24 hospital units with the highest composition of badged usage experienced a larger increase in total daily usage per bed (112.0%) than the 24 units with mid-level composition of badged usage (47.5%) or the 23 units with the lowest composition of badged usage (24.3%).⁵ This analysis shows that hospitals with higher badged usage experienced greater increases in compliance behavior than hospitals with lower badged usage. This result suggests that badged users, rather than unbadged users, have primarily driven the effects of electronic monitoring on dispenser usage.

Finally, we examine one subset of caregivers for whom we had data both before and after the activation of individual electronic monitoring – pilot testers. Across 71 hospital units, 195 of the 5,247

⁵ We see the same patterns when we (a) examine all observations in the pre-activation and post-activation periods or (b) estimate changes using regression specification (1). These additional analyses are available from the authors upon request.

caregivers (who experienced a total of 1,159,338 hand hygiene opportunities) were designated as pilot testers, received an active RFID badge, and were instructed to wear their badge at work during the pre-activation period of their corresponding unit. These caregivers were informed that as pilot testers their hand hygiene compliance rates would be tracked by Proventix but *not* by their managers. Since pilot testers were tracked both before and after the activation of individual monitoring, we are able to identify a precise effect of activating individual monitoring on their compliance rates. Such an analysis is not subject to the same limitation surrounding badged versus unbadged use as our unit-level analysis above.

Using an OLS regression model similar to specification (2) described in Section 3.4, we predicted the dependent variable, *compliance*, indicating whether a pilot tester washed her hands at a given hand hygiene opportunity. We removed *months since activation* from regression specification (2) and added the indicator variable *activation* as our key predictor variable.⁶ The positive and significant coefficient on *activation* shows that pilot testers' hand hygiene compliance rates significantly increased following the activation of individual monitoring ($\beta = 0.11$; $p < 0.01$), supporting Hypothesis 1 (Appendix D provides more detail). Compared with the regression-adjusted baseline compliance rates among these pilot testers of 43.95%, compliance rates in the post-activation period increased to a rate of 54.56%, representing an improvement of 24.14%. Note that pilot testers were wearing badges that tracked their compliance and reported it to Proventix (just not their managers) prior to activation, so they likely already felt monitored individually prior to activation. Thus, this analysis is a conservative test of Hypothesis 1. We also note that pilot testers were not randomly selected. That said, we believe that this finding, combined with the unit-level results above, provide valuable support for our first hypothesis.

. All in all, unbadged usage does not appear to be driving our observed effect. However, it is interesting that even though they are not monitored, unbadged users increase their use with individual electronic monitoring deployment. Unbadged caregivers, patients and visitors may have increased their dispenser usage as they observed that badged caregivers who were being monitored increased their hand

⁶ We retained the same control variables in specification (2) except the indicator for other caregivers' presence at a hand hygiene opportunity because we did not have information about other caregivers' presence prior to individual monitoring activation.

hygiene activity. In addition, unbadged caregivers may have felt organizational pressure to comply once they knew that badged caregivers were being watched. As noted in a client prepared report shared with the researchers, “Increases in HH [hand hygiene] activity by non-monitored individuals suggest that the ability to influence HH behavior extended beyond the directly monitored healthcare workers.”

4.1.3 Pre-Activation Usage. To test Hypothesis 2 we first noted that, as expected, there is substantial variation in the effect size of activation on usage across units. Overall there is a clear, positive average effect of activation, consistent with Hypothesis 1. Notably, 55 out of 71 hospital units (i.e., 77.46% of our sample) showed a boost in usage after initiating electronic monitoring, with the effect being statistically significant at the 5% level for 44 individual hospitals units (i.e., 61.97% of all observed units). However, in the unit where monitoring was least effective, activating individual monitoring actually decreased daily dispenser usage significantly by 39.32 uses per unit bed per day (75.62% of this unit’s regression-adjusted baseline); in the unit where monitoring was most effective, it increased daily dispenser usage significantly by 92.23 uses per unit bed per day (180.88% of this unit’s regression-adjusted baseline). A Wald test indicated that the differences in the estimated effects of individual monitoring across 71 hospital units were significantly larger than would be expected by chance ($F(38, 70) = 3.9 \times 10^8, p < 0.0001$). This means that the differences between hospital units were not simply due to random noise but were substantive.

Turning to the effect of pre-activation usage on adoption, we examined Model 2 of Table 3 where we added an interaction between *activation* and (mean centered) *pre-activation usage* to our regression model to predict total daily usage per unit bed. We note that mean reversion would predict that units with higher pre-activation usage should show a smaller activation effect, in contrast to our Hypothesis 2, which predicted just the opposite of this. The significant and positive interaction effect shown in Model 2 of Table 3 suggests that usage improved more as a result of monitoring in hospital units with higher pre-activation dispenser usage than in hospital units with lower pre-activation dispenser usage ($p = 0.02$). Specifically, a one standard deviation increase in pre-activation dispenser usage (i.e., an increase of 15 uses per unit bed per day) is associated with a 7.34 use per unit bed per day increase in the positive effect

of individual monitoring on total daily usage (i.e., 21.58% boost in regression-adjusted baseline total daily usage). This result supports Hypothesis 2.

4.2 The Long-term Effects of Individual Monitoring on Compliance

So far we have demonstrated that when hospital units initiated electronic monitoring, process compliance increased significantly and this effect was stronger in units with more badged users and more pre-activation usage. We next turned to an investigation of how persistent these effects are over time. We first examined whether compliance shows an inverted U-shaped relationship over time, declining in efficacy after the initial boost when caregivers first learned they were being monitored (H3).

We used data on caregivers' compliance (rather than data tracking overall dispenser usage) for these analyses in order to examine how individuals responded to monitoring over the long-term. We used regression specification (2), described above, to explore this pattern rigorously. Model 3 in Table 3 showed a significant and positive coefficient on *months since activation* ($p = 0.01$) and a significant and negative coefficient on its squared term ($p = 0.02$). This indicates that caregivers' hand hygiene compliance increased at first, post-activation, before subsequently decreasing. To confirm that there was an inverted U-shaped relationship between *months since activation* and compliance, we conducted further analyses (Kesavan, Staats and Gilland 2014; Tan and Netessine 2014). First, we determined that the stationary point occurred 21.6 months after the activation of monitoring, which was well within the observation period. The 95% confidence interval surrounding the stationary point calculated based on the delta method, [12.5 months, 30.7 months], was also within the observation period (Muggeo 2003). Second, following Lind and Mehlum (2010), we confirmed that the slope of the curve was positive and significant at the minimum point of *months-since-activation* ($p = 0.003$), and the slope was negative and significant at $p = 0.05$ at the maximum value. Further, following Nelson and Simonsohn (2014), we ran two separate regression models using (a) observations that occurred before the stationary point (i.e., 21.6 months) and (b) observations that occurred after the stationary point. We found that compliance increased significantly over time before *months since activation* reached the stationary point ($\beta = 0.006, p < 0.0001$), whereas compliance decreased significantly over time after *months since activation* reached the

stationary point ($\beta = -0.016, p = 0.002$). Finally, instead of imposing a curvilinear relationship between compliance and months since activation by regressing compliance on the linear and quadratic terms of *months since activation*, we tested Hypothesis 2 using a more flexible specification. In a regression model similar to Model 3 in Table 3, we replaced the continuous measure of *months since activation* with 39 indicator variables for each of the months in the post-activation period for each hospital unit, and we treated the month of activation as the omitted, reference group. Based on the results of this regression model, we then calculated the fitted hand hygiene compliance for each month in the post-activation period. As shown in Figure 3, this analysis clearly reveals that fitted compliance increased first and then gradually decreased over time. The 20th month following the activation of electronic monitoring had the highest fitted compliance, which was significantly higher than both compliance in the month of activation and compliance in the last month during our observation period (both p-values < 0.05). These analyses based on a flexible regression specification produced the same conclusions as our curvilinear model and confirmed the existence of an inverted U-shaped relationship, supporting Hypothesis 3.

4.3 The Effect of Terminating Individual Monitoring

Finally, we turn to our termination hypothesis (H4), which predicted that compliance would decrease below the baseline pre-activation level when monitoring was discontinued. To investigate this, we first analyzed data on nine hospital units at six hospitals that terminated their monitoring program in the middle of 2013 ($N_{observations} = 4,520$). We compared post-deactivation dispenser usage to pre-activation dispenser usage using all observations prior to activation and all observations following deactivation among the nine units that deactivated individual monitoring. Doing so revealed that average, post-deactivation daily usage per bed ($M = 28.40, SD = 27.93$) was lower than the average, pre-activation daily usage per bed ($M = 35.23, SD = 17.26, t(933) = 4.40, p < 0.0001$), supporting H4. In fact, the average post-deactivation daily usage per unit bed was lower than the average pre-activation daily usage in eight out of nine units, with the decrease being statistically significant at the 5% level in six units.

As an additional test of Hypothesis 4 we used a regression model. Model 4 in Table 3 relies on regression specification (3), described previously, and included all observations from all of the units in

our data. We included two primary predictor variables: *activation* and *deactivation*, to simultaneously identify the effects of activating and deactivating individual monitoring. Consistent with previous support presented for Hypothesis 1 in Section 4.1.1, a positive and significant coefficient on *activation* indicated that total daily usage per unit bed increased following the activation of individual monitoring ($p < 0.001$). We next turned to *deactivation* to examine the differences between post-deactivation dispenser usage and pre-activation usage. In support of Hypothesis 4, the negative and significant coefficient ($p < 0.001$) on *deactivation* suggested that dispenser usage on average dropped below the pre-activation usage level following the removal of individual electronic monitoring. Specifically, compared with the regression-adjusted daily dispenser usage per unit bed in the pre-activation period (36.63), terminating individual electronic monitoring decreased daily dispenser usage per unit bed by 46.60% to 19.56.

4.4 Endogeneity of Monitoring Adoption and Termination

We conducted additional analyses to explore whether the timing of activating individual monitoring is correlated with key hospital characteristics. Specifically, we split the 71 hospital units into two groups based on when they adopted individual monitoring. The 36 hospital units with earlier adoption dates and the 35 hospital units with later adoption dates did not significantly differ in (i) the number of beds in a unit, (ii) average monthly inpatient admissions, and (iii) pre-activation daily usage (all p-values > 0.18).⁷ We also split the 42 hospitals based on when they adopted individual monitoring and confirmed that the 21 hospitals with earlier activation dates did not differ from the 21 hospitals with later activation dates along any of the observables that Proventix provided to us and collected from the American Hospital Directory database (e.g., total inpatient discharges; all p-values > 0.30).

To address the concern that dispenser usage was already increasing before individual monitoring was activated, we analyzed pre- and post-activation time trends in dispenser usage, following past research (Autor 2003). Specifically, we created two “lead” indicator variables to indicate 0-30 days before

⁷ We examined whether the monitoring effect was different between early and late activating hospital units. To estimate the monitoring effect size for every unit, we used regression specification (1) from Section 3.4 where the dependent variable is total daily usage per bed. We replaced the *activation* indicator with indicators for the post-activation period for each of the 71 units. We found that the *activation* effect was on average larger for the 35 hospital units that activated monitoring later, as compared to the 36 hospital units that activated monitoring earlier ($p = 0.002$). In particular, this effect is driven by the 17 units that activated in 2013, as compared to the 54 units that activated before 2013. We find no significant difference in activation effect between units that activated in 2010 vs. 2011 vs. 2012 (p-values for pairwise comparisons > 0.41).

activation and 30-60 days before activation, respectively. We also created 11 “lag” indicator variables to capture 0-30 days, 30-60 days, 60-90 days, up to 270-300 days after activation and beyond 300 days after activation, respectively. In a model similar to Model 1 from Table 3, we replaced our key predictor variable, *activation*, with these aforementioned indicator variables for the 30-day periods before (lead) and after (lag) activation. In this model, the reference group was the period of time preceding activation by more than 60 days. The coefficients on the two “lead” indicator variables were not statistically significant, suggesting there was no significant anticipatory response to activation among hospital units that were about to activate individual monitoring. See Appendix E for more details.

We also explored whether individual monitoring termination was correlated with key hospital characteristics. We found that the nine units that discontinued monitoring did not significantly differ from the remaining 62 units on observable dimensions: (i) unit beds, (ii) average monthly inpatient admissions, and (iii) pre-activation daily usage (all *p-values* > 0.11). Also, these two unit types were comparable in terms of the average size of the effect of activating monitoring on dispenser usage (*p-value* > 0.67).

4.5 Additional Robustness Checks

The regression results presented in Table 3 remained meaningfully unchanged in terms of magnitude and significance when we performed the robustness checks discussed in the results section. In addition, we performed checks detailed below. All reported results are available from the authors upon request.

First, we find that for the dispenser usage data described in Section 3.2.1, our findings were unchanged when we removed cases where total daily usage per unit bed was at least three standard deviations above the mean ($n = 542$) or when we log-transformed total daily usage per unit bed to address the concern that the variable was right-skewed. Further, for the caregiver compliance data described in Section 3.2.2, we obtained the same basic results with a daily-level model where the dependent variable was the average compliance rate of a given caregiver on a given day, instead of the hand hygiene opportunity model reported here. Also, our results remained unchanged if we excluded control variables in Table 1 that were only used for analyzing caregiver data (e.g., *hours at work*, *hours off work*, etc.).

Second, an additional concern pertains to potential problems with the standard errors in our difference-in-difference models. Bertrand et al., (2004) suggest that because of serial correlation, these models may underestimate the standard errors of the estimated treatment effects and so following Bertrand et al., (2004), we addressed the serial correlation problem by clustering standard errors at the hospital unit level in all of our models (Table 3). Further, we implemented placebo tests to demonstrate that the estimated effects of activating monitoring were not spuriously driven by our data structure (Pierce et al., 2014). Specifically, for each hospital unit, we randomly assigned an activation date by selecting a day that fell between the unit's activation date and its last day in our data (or its deactivation date if this unit terminated monitoring). Then we ran Model 1 in Table 3 by replacing actual activation dates with placebo activation dates. We repeated this placebo test 100 times. Appendix F presents the estimated effects of activating monitoring, together with the 95% confidence intervals, for each placebo test and the actual data (i.e., Model 1 in Table 3). Only three out of 100 placebo models produced positive coefficients significant at the 5% level, and all estimates were smaller than the true data's estimates.

Finally, a concern is whether monitoring termination signaled a decrease in patient volume that could have led to a reduction in dispenser usage, supporting Hypothesis 4. As a reminder, most units terminated because a grant expired. The importance of hand hygiene compliance, as captured by funding and accreditation body recommendations, did not materially change during this time. As a check, using information on monthly admissions in eight hospital units that terminated monitoring, we confirmed that average monthly admissions after activation did not decrease leading up to badge reclamation.

5. Discussion and Conclusion

Our empirical results demonstrated large and significant initial benefits from individual electronic monitoring on process compliance. First, we found that, on average, a large and significant increase in hand hygiene compliance ensued after the introduction of individual electronic monitoring. However, there was significant variability in the treatment effect across organizations. Second, we found that units with higher pre-activation hand hygiene compliance had a larger treatment effect than units with lower pre-activation hand hygiene compliance. Third, utilizing individual-level data, we found that the effect of

monitoring was persistent over time; however, the effect did eventually begin to degrade approximately two years after rollout. Although this initial increase in process compliance is consistent with a habit formation model, eventually compliance declined, which seems more consistent with a desensitization perspective. Within the 1,214-day follow-up period we studied, these declines were not sufficient to produce compliance levels below those observed prior to the electronic monitoring intervention; however, it is possible that a longer panel of data could reveal such a reversal. Based on our regression results, we extrapolated the compliance rates for the period beyond our observation period. We estimated that, if everything else remained the same, the average compliance rates approximately 43 months after activation would be the same as the compliance rates on the activation date. Further, it would take about ten years post-activation for the average compliance rates to drop significantly below compliance rates at the time of activation if the decline we detected continued unabated. Finally, we investigated termination effects within our sample and found that not only did levels of use decline when individual electronic monitoring was removed, but that strikingly, they declined below pre-activation levels.

Our findings make a number of contributions to academic theory. Achieving process compliance has been a foundational challenge in operations for over a century. Although Taylor noted as early as 1911 that managers needed to enforce process compliance, studies from the early 20th century (Mayo 1933) through the modern era (Bernstein 2012) indicate that process compliance often does not occur. In some cases, it may be appropriate to ignore standardized processes if they are poorly designed. However, many standardized processes yield better, faster, and safer outputs. This is particularly true with hand hygiene where the medical literature has convincingly showed the positive safety implications of compliance by caregivers with hand hygiene recommendations (WHO 2009; Boyce 2011).

The first contribution of this work is to demonstrate that individual electronic monitoring can have a large and significant positive effect on process compliance. We identified this effect across a substantial number of organizations using archival data. Providing such a finding across multiple organizations is an important contribution, since prior work notes that the limited research on the topic has considered only one or two units, typically, and that not all units exhibit a positive effect (Boyce

2011; Ward et al. 2014). Studying the deployment of the same system across multiple organizations also provides four additional benefits. First, it provides a constant measure of compliance to evaluate across organizations, as opposed to prior studies that measured compliance differently from each other (Srigley et al. 2015). Second, it allows us to eliminate concerns about correlated, external events (e.g., a CDC report emphasizing the importance of hand hygiene) that could be driving the effect. The organizations included in our study deployed electronic monitoring at different times during our three-and-a-half-year observation period. This diversity in timing aids in controlling for any concurrent events that might bias a typical one-off event study. The prevalent benefits of electronic monitoring observed in our study further suggest that our findings are unlikely to be explained by external, concurrent events. Third, studying multiple organizations enables us to overcome publication bias where a positive effect in a single organization is more likely to be submitted and published in the literature.

Documenting the significant variability that exists in the activation effect is the paper's second contribution. It not only responds to a call for exploration of heterogeneous treatment effects in healthcare (Boyce 2011), but also provides value to operations management by highlighting the substantial variance that can exist across organizations. With the heterogeneity identified, it becomes possible to explain why such differences exist. This leads to our third contribution: an exploration of the effect of pre-activation compliance on the adoption of individual electronic monitoring. Prior work highlights that organizational members' actions are affected by the organization's norms, organizational design, and management attention. We further demonstrate that those organizational factors can help explain the significant variation we observe in the effects of electronic monitoring on compliance behavior. In line with the results about monitoring termination that we subsequently report, this finding highlights that individual electronic monitoring is not a panacea, by itself. Organizations looking to build process compliance must think about how electronic monitoring fits within a broader system encompassing not only technology, but also norms, culture, and leadership, among other things.

Our fourth contribution is to shed light on the impact of individual electronic monitoring on compliance over time. We evaluated whether the positive initial effect of monitoring increased or

gradually decayed as time since the initial rollout passed. On one hand, if the initial introduction of monitoring creates a lasting change in employees' beliefs about the importance of monitoring or perceptions of management's commitment or enforcement of hand hygiene, we expect that the increase in compliance following monitoring should be sustained. Further, research on habituation (Wood and Neal 2007; Neal et al. 2012) suggests that the constant presence of electronic monitoring may help people eventually build hand hygiene habits whereby hand hygiene compliance becomes unconsciously integrated into their behavioral routines, leading to either constant or perhaps increasing process compliance over time. On the other hand, another stream of research suggests that interventions may slowly lose their effectiveness with the passage of time as they grow less salient (Thompson and Spencer 1966; Nahrgang, Morgeson and Hofmann 2011; Haselhuhn et al. 2012). This would predict a decline in process compliance over time. Our results showed that neither prediction alone is entirely accurate. Rather, we found that at first, compliance increased, albeit at a decreasing rate, and eventually it declined. Practically speaking, the good news is that we do not see compliance decline until almost two years after monitoring is deployed. Interestingly single and dual organization studies in the healthcare literature have identified the initial increasing portion of this relationship (Armellino et al. 2011; Walker et al. 2014), but not the eventual decreasing returns. Our results not only provide multiple organizations to explore the effect, but given our longer, longitudinal panel we are able to observe compliance begin to decrease.

This paper's fifth contribution, which is its most surprising one, comes from its examination of the implications of terminating electronic monitoring. Prior theory is unclear as to what impact monitoring termination will have, and we address this uncertainty with our analysis. We found that levels of compliance decreased significantly after electronic monitoring was terminated. Not only that, but remarkably, we found that levels of compliance decreased *below* the initial level of compliance observed before the original deployment of electronic monitoring. These findings suggest that electronic monitoring did not produce a habit that was sufficiently strong to withstand its removal. Our results are also consistent with a crowding out hypothesis whereby individuals who were previously complying due to internal motivation may have shifted their motivation to an external focus. When this external stimulus

was removed, the behavior declined. More work is needed to understand why process compliance decreased on average after the termination of monitoring, but the observation that removal of monitoring technology can leave compliance worse than it was before installation is an important, cautionary lesson.

The final contribution of this work is methodological. A continual challenge for researchers has been how to study behavioral responses to changes within organizations as the mere act of observing a behavior can change that behavior (Roethlisberger and Dickson 1934). Webb et al. (1966) suggested that researchers use trace data to track the actions of individuals without biasing their behavior. In 1966 this meant doing things such as measuring carpet wear to estimate traffic patterns. New technologies present novel sources of digital trace data. By using RFID badge data we introduce the small, but growing body of sociometric badge research to the operations field. Not only can RFID badges be used to study this paper's questions, but they offer the potential to answer many process questions. For example, RFID tags may make it possible to observe stockouts in inventory research. Time and motion studies may include precise data on where and when actions occurred. For example, in healthcare, a researcher could examine if time spent by caregivers in patient rooms predicts a decrease in length of stay or adverse outcomes. These studies will require partnerships between academia and industry, and they will offer an exciting, new glimpse into the behavioral drivers of operational efficiency and effectiveness.

5.1 Hospital Intervention Efforts

As mentioned in Section 3.1, Proventix management identified the presence and timing of management interventions. While acknowledging that the intervention count is small (Performance Feedback = 48 units; Goal Setting = 36; Leadership = 19; Competition = 8; Incentive = 25) and their introduction is endogenous, it is useful to examine their impact on process compliance to spur future research. It is rare to find a setting with data on such diverse interventions. Thus, we repeated Table 3's models with the addition of indicators for each of the five interventions types. The indicators equaled one when an intervention was introduced in a unit and thereafter and otherwise were equal to zero. We only included units for which we had intervention data, and we reported the regression results in Table 5.

First, we confirmed that all of our hypotheses were statistically significantly supported, as before. Next, to estimate the effects of the interventions, we focused on Model 3, as this model captured individual compliance rates over time. We found that indicators for the five types of interventions were jointly significant ($p < .0001$). The coefficients on goal setting ($p = 0.76$), incentives ($p = 0.07$), and performance feedback ($p=0.65$) were not statistically significant at conventional levels. The indicators for the leadership and competition interventions were, however, statistically significant at conventional levels. Specifically, in the cases where leadership learned about the initiative and showed commitment, compliance rates were on average 13 percentage points higher; in the cases where caregivers were divided into groups that were placed in competition to reach the highest compliance, compliance rates were, on average, 6 percentage points higher. Of course, leaders' demonstration of commitment and the introduction of team competition can occur contemporaneously with other changes in management practices (e.g., using monitoring for learning purposes, rather than just punishment). Nevertheless, these results highlight the importance of leadership commitment to change initiatives and the motivating effects of competition. Ideally, future work would include random assignment of interventions to understand the relative magnitudes of the benefits produced by these interventions and their interactive effects.

5.2 Limitations

It is important to note that our study had several significant limitations. First, we are unable to identify the effects of individual electronic monitoring on caregivers per se but can only estimate the overall effects of monitoring on all users of hand hygiene dispensers in hospitals. Though our analysis of pilot testers sheds some light on the effects of monitoring on individual caregivers, it would be valuable for future research to precisely calibrate the magnitude of the monitoring effect among average caregivers. Such a calibration would be tremendously useful for making more precise policy recommendations. Second, we were unable to precisely track all hand-hygiene-related interventions that were implemented in our sample during the study period. As described in Section 5.1, we replicate our results when we control for known management interventions that were layered on top of electronic

monitoring. However, to cleanly isolate the effect of monitoring in the field over an extended period time, future research would need to create a far more controlled work environment than our setting provided.

A third limitation pertains to the generalizability of our findings to the broader population of hospitals. Comparing the hospitals in our sample to the broader population, we find that our sample is slightly larger and slightly more urban than average, but the case mix index is comparable. Although we would theoretically expect to find similar effects of activating individual electronic monitoring across the population of hospitals, additional work with still larger samples and random assignment of monitoring would be valuable. In a similar vein, our investigation of the effect of terminating individual monitoring relies on a small subset of the hospital units in our overall sample. Though as detailed in Section 4.4, the hospital units that terminated monitoring were similar to other hospital units in our sample on important dimensions, we need to be cautious about generalizing our findings to a broad population of hospitals.

A fourth potential limitation of our work is that the monitoring technology we study requires caregivers to wear their badges for compliance to be tracked. It could be that caregivers only wore their badges when they planned to be compliant and removed them during periods of non-compliance. In analyses of total hand hygiene dispenser usage, this is not a concern because it would not alter caregivers' observability. We also do not think such "strategic behaviors" are likely to explain the long-term trend in caregivers' compliance that we observed either because when we controlled for the number of daily hand hygiene opportunities in Model 3 of Table 3, our findings remained virtually unchanged. Still, our inability to monitor badged caregivers if they elected to remove their badges is a limitation.

Further, we do not have direct evidence to explain why process compliance, on average, decreased below the pre-activation usage level after the termination of electronic monitoring. There is a concern that a post-termination decrease in the behavior could be the result of a response to a change in the estimated importance of that behavior. However, in the case of hand hygiene, the expectation from accreditation bodies, patients, and payers remained that hospitals should have high compliance. Termination of electronic monitoring was unlikely driven by an external or clinical decline in hand hygiene emphasis. That said, it is possible that management –implicitly or explicitly – did not place as

high of a value on hand hygiene post-termination, as illustrated by their decision not to deploy financial resources towards continued electronic monitoring of hand hygiene. Finally, although our study analyzes data across many hospitals, it looks at only one practice (hand hygiene) in one industry. Future work should extend this research to explore other industries and other compliance behaviors.

5.3 Directions for Future Research

There are a number of ways that future research could and should follow-up on our work. First, there is an opportunity to further explore the boundary conditions that influence the efficacy of electronic monitoring. As discussed earlier, the intrusiveness of monitoring and its consistency with employees' goals may impact the effect of monitoring. In our setting, monitoring is minimally intrusive and few caregivers are likely to object to the goal of hand hygiene compliance, at least publicly. Future work would ideally examine settings where not only are these factors are different, but where there is within-study variation on these dimensions. Second, there is an opportunity to investigate additional moderators of the effect of monitoring on compliance and termination. For example, work on operational transparency shows that there are benefits to consumers and workers that accrue from revealing these groups' actions to one other (Buell and Norton 2011; Buell, Kim and Tsay 2014). Transparency creates monitoring, by a customer instead of a manager, and so creates opportunities for extending research on monitoring. Seeing how operational transparency – as well as other factors – influence the effects of electronic monitoring on process compliance would be fruitful areas for future study.

In our setting, individuals were also able to receive personalized feedback on their compliance levels by looking at the display screen on a given dispenser. Although we cannot fully disentangle the individual feedback and monitoring effects, our robustness checks with pilot testers help address this concern. Pilot testers were able to receive personalized feedback during the pre-activation period but were told that their compliance was not being monitored by management. Although one can question if they completely believed this instruction, we still observed a significant, 19% increase in compliance among these pilot testers following the full rollout of individual electronic monitoring. This provides strong

evidence that electronic monitoring itself and not merely personalized feedback on compliance, meaningfully improves process compliance. However, future work on this topic would be valuable.

Finally, topics such as habit formation have been studied traditionally in the laboratory (Wood and Neal 2007). New work is beginning to unpack these effects over time in the field (Allcott and Rogers 2014; Pierce et al. 2014; Milkman et al. 2014). Examining repeated behaviors in their native context allows researchers to study them over longer periods of time than is practical in the lab, and to understand the underlying causal dynamics that drive changes through the use of detailed archival data and field experiments. For example, individuals' likelihood of complying with standard processes is not only a function of their choices and organizational environment (e.g., the deployment of electronic monitoring), but also the compliance activities of others around them (c.f., Huckman, Staats and Upton 2009; Chan et al. 2014). Future work could examine how persistence changes based on the actions of co-workers.

For example, in our regression specification (2) for individual compliance rates we included a control variable for the presence of another caregiver in the room. Examining the variable *presence* in Model 3 of Table 3 we find the coefficient is negative ($\beta = -0.04$) and statistically significant ($p < .0001$). This suggests that caregivers wash their hands *less* frequently both at room entries and room exits if their coworkers are in the same patient room. Future work should unpack this result further, but there are at least three possible reasons why this might be the case. First, having others in the room may prove distracting as opposed to creating social pressure to comply. It is uncommon for coworkers to be co-present in the same room (10% of the time), thus it is plausible that caregivers are distracted by the presence of coworkers and forget to wash their hands. It is also plausible that the presence of coworkers leads caregivers to focus on doing a good job at their primary tasks (e.g., handling an emergency) and shifts their attention away from secondary tasks (e.g., hand hygiene). Finally, cases where caregivers enter and/or exit a patient room with coworkers in the same location might be cases where caregivers perform a multi-person task together with their coworkers (e.g., restraining a patient, pulling a patient up in the bed). Caregivers might consider it clinically less important to wash their hands if they just come into the room to lend a hand to coworkers, although the fact that they have touched something in the room

means that they still should wash their hands. Building on work that has examined peer effects (Mas and Moretti 2009; Chan et al. 2014), future research should delve deeper to understand this effect.

5.4 Conclusion

Altogether our paper has important implications for managers and scholars. We found that electronic monitoring is one tool that managers can use to dramatically improve standardized process compliance. Given the low compliance rates observed in many contexts (e.g., checklists, standardized work, use of safety gear), this is encouraging news. However, we also found that managers cannot simply “monitor and forget.” Rather, the observed drop-off in compliance after a lengthy period of monitoring suggests that there is a need for ongoing managerial interventions to sustain the benefits of monitoring. Finally, our finding that after monitoring is terminated, compliance falls below its original, natural levels offers a cautionary note for managers. Deploying electronic monitoring without a long-term plan for supporting its retention may be short-sighted as it could end up actually *harming* process compliance.

Our paper also contributes to operations management literature. Significant attention is focused on creating standard practices to improve operational performance. However, without adoption and compliance, the benefits from such standardized practices cannot be realized. By studying electronic monitoring we examined one important way that the use of standardized practices can be encouraged. By shedding light on the effects of adoption, persistence, and termination, we not only gain insight into how to encourage process compliance, but we also learn how better to design operating systems.

6. References

- Aiello, J. R. and K. J. Kolb (1995). "Electronic performance monitoring and social context: Impact on productivity and stress." *J. Appl. Psych.* **80**(3): 339-353.
- Allcott, H. and T. Rogers (2014). "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation." *American Economic Review* **104**(10): 3003-3037.
- Andritsos, D. and C. S. Tang (Forthcoming). "Linking process quality and resource usage: An empirical analysis." *Production and Operations Management*.
- Armellino, D., E. Hussain, M. E. Schilling, W. Senicola, A. Eichorn, Y. Dlugacz and B. F. Farber (2011). "Using high-technology to enforce low-technology safety measures." *Clinical Infectious Diseases*.
- Autor, D. H. (2003). "Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing." *J. Labor Econ.* **21**(1): 1-42.
- Baker, G. and T. Hubbard (2004). "Contractibility and asset ownership: On board computers and governance in U.S. trucking." *The Quarterly Journal of Economics* **119**: 1443-1479.
- Bénabou, R. and J. Tirole (2003). "Intrinsic and extrinsic motivation." *The Review of Economic Studies* **70**(3): 489-520.
- Bernstein, E. S. (2012). "The Transparency Paradox: A role for privacy in organizational learning and operational control." *Administrative Sci. Q.* **57**(2): 181-216.
- Bertrand, M., E. Duflo and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* **119**(1): 249-275.
- Bohn, R. E. (2005). "From art to science in manufacturing: The evolution of technological knowledge." *Foundations and Trends in Technology, Information and Operations Management* **1**(2): 129-212.
- Boyce, J. M. M. D. (2011). "Measuring healthcare worker hand hygiene activity: Current practices and emerging technologies." *Infection Control and Hospital Epidemiology* **32**(10): 1016-1028.
- Buell, R. W., T. Kim and C.-J. Tsay (2014). "Cooks make tastier food when they can see their customers." *HBS Working Paper*.
- Buell, R. W. and M. I. Norton (2011). "The labor illusion: How operational transparency increases perceived value." *Management Sci.* **57**(9): 1564-1579.
- Centers for Disease Control and Prevention (2002). "Guideline for hand hygiene in health-care settings." *Morbidity and Mortality Weekly Report* **51**: RR-16.
- Chan, T. Y., J. Li and L. Pierce (2014). "Compensation and peer effects in competing sales teams." *Management Sci.* **60**(8): 1965-1984.
- Charness, G. and U. Gneezy (2009). "Incentives to exercise." *Econometrica* **77**(3): 909-931.
- Cialdini, R. B., R. R. Reno and C. A. Kallgren (1990). "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places." *J. Personality Soc. Psych.* **58**(6): 1015-1026.
- Corbett, C. J., M. J. Montes-Sancho and D. A. Kirsch (2005). "The financial impact of ISO 9000 certification in the United States: An empirical analysis." *Management Sci.* **51**(7): 1046-1059.
- Dai, H., K. L. Milkman, D. A. Hofmann and B. R. Staats (2015). "The impact of time at work and time off from work on rule compliance: The case of hand hygiene in healthcare." *J. Appl. Psych.* **100**(3): 846-862.
- Deci, E. L. (1975). *Intrinsic Motivation*. New York, Plenum Publishing.
- Deci, E. L. and R. M. Ryan (2002). Overview of self-determination theory: An organismic dialectical perspective. *Handbook of Self-Determination Research*. E. L. Deci and R. M. Ryan. Rochester, NY, Rochester University Press: 3-33.
- DeHoratius, N. and A. Raman (2008). "Inventory record inaccuracy: An empirical analysis." *Management Sci.* **54**(4): 627-641.
- Frey, B. S. (1993). "Shirking or work morale?" *European Economic Review* **37**(8): 1523-1532.

- Gerber, A. S., D. P. Green and C. W. Larimer (2008). "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* **102**(1): 33-48.
- Gino, F. and G. P. Pisano (2008). "Toward a theory of behavioral operations." *Manufacturing Service Oper. Management* **10**(4): 676-691.
- Goldstein, N. J., R. B. Cialdini and V. Griskevicius (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels.
- Harter, D. E., M. S. Krishnan and S. A. Slaughter (2000). "Effects of process maturity on quality, cycle time, and effort in software product development." *Management Sci.* **46**(4): 451-466.
- Haselhuhn, M. P., D. G. Pope, M. E. Schweitzer and P. Fishman (2012). "The impact of personal experience on behavior: Evidence from video-rental fines." *Management Sci.* **58**(1): 52-61.
- Huckman, R. S., B. R. Staats and D. M. Upton (2009). "Team familiarity, role experience, and performance: Evidence from Indian software services." *Management Sci.* **55**(1): 85-100.
- Kallman, E. (1993). Electronic monitoring of employees: Issues & guidelines. *Journal of Systems Management.* **44**: 17+.
- Kesavan, S., B. R. Staats and W. G. Gilland (2014). "Labor-mix and volume flexibility: Evidence from a retailer." *Management Sci.* **60**(8): 1884-1906.
- Klein, K. J. and A. P. Knight (2005). "Innovation implementation: Overcoming the challenge." *Current Directions in Psychological Science* **14**: 243-246.
- Levine, D. I. and M. W. Toffel (2010). "Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers." *Management Sci.* **56**(6): 978-996.
- Lind, J. T. and H. Mehlum (2010). "With or without U? The appropriate test for a U - shaped relationship*." *Oxford Bulletin of Economics and Statistics* **72**(1): 109-118.
- MacMillan, N. A. (2002). "Signal Detection Theory." *Stevens' Handbook of Experimental Psychology.*
- Martin, A. W., S. H. Lopez, V. J. Roscigno and R. Hodson (2013). "Against the rules: Synthesizing types and processes of bureaucratic rule-breaking." *Acad. Management Rev.* **38**(4): 550-574.
- Mas, A. and E. Moretti (2009). "Peers at work." *American Economic Review* **99**(1): 112-145.
- Mayo, E. (1933). The Human Problems of an Industrial Civilization. New York, Macmillan.
- Medicine, I. o. (1999). To err Is human: Building a safer health system. Washington, DC, National Academies Press.
- Milkman, K. L., J. A. Minson and K. G. M. Volpp (Forthcoming). "Holding the hunger games hostage at the gym: An evaluation of temptation bundling." *Management Sci.*
- Muggeo, V. M. R. (2003). "Estimating regression models with unknown break-points." *Statistics in Medicine* **22**(19): 3055-3071.
- Nagin, D. S., J. B. Rebitzer, S. Sanders and L. J. Taylor (2002). "Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment." *American Economic Review* **92**(4): 850-873.
- Nahrgang, J. D., F. P. Morgeson and D. A. Hofmann (2011). "Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes." *J. Appl. Psych.* **96**(1): 71-94.
- Neal, D. T., W. Wood, J. S. Labrecque and P. Lally (2012). "How do habits guide behavior? Perceived and actual triggers of habits in daily life." *J. of Experimental Soc. Psych.* **48**(2): 492-498.
- Nelson, L. and U. Simonsohn. (2014). "Thirty-somethings are shrinking and other U-shaped challenges." Data Colada Retrieved October 19, 2014, from <http://datacolada.org/2014/09/17/27-thirty-somethings-are-shrinking-and-other-u-shaped-challenges/>.
- Ouchi, W. G. (1979). "A conceptual framework for the design of organizational control mechanisms." *Management Sci.* **25**(9): 833-848.
- Pierce, L., D. Snow and A. McAfee (2014). "Cleaning house: The impact of information technology monitoring on employee theft and productivity." *Management Sci.*

- Pierce, L. and J. Snyder (2008). "Ethical spillovers in firms: Evidence from vehicle emissions testing." *Management Sci.* **54**(11): 1891-1903.
- Roethlisberger, F. J. and W. J. Dickson (1934). *Management and the Worker*. Boston, Harvard Univ.
- Rogers, T. and E. Frey (2014). "Changing behavior beyond the here and now." *HKS Working Paper*.
- Senot, C., A. Chandrasekaran, P. Ward, A. Tucker and S. Moffat-Bruce (2015). "The impact of combining conformance quality and experiential quality on readmissions and cost performance." *Management Sci.*
- Shah, R., S. M. Goldstein, B. T. Unger and T. D. Henry (2008). "Explaining anomalous high performance in a health care supply chain." *Decision Sci.* **39**(4): 759-789.
- Smiddy, H. F. and L. Naum (1954). "Evolution of a "Science of Managing" in America." *Management Sci.* **1**(1): 1-31.
- Song, H., A. Tucker, K. L. Murrell and D. R. Vinson (2015). "Learning from the best: The effects of public relative performance feedback on variability and productivity." *HBS Working Paper*.
- Srigley, J. A., M. Gardam, G. Fernie, D. Lightfoot, G. Lebovic and M. P. Muller (2015). "Hand hygiene monitoring technology: a systematic review of efficacy." *Journal of Hospital Infection* **89**(1): 51-60.
- Staats, B. R., D. J. Brunner and D. M. Upton (2011). "Lean principles, learning, and knowledge work: Evidence from a software services provider." *J. of Operations Management* **29**(5): 376-390.
- Tan, T. F. and S. Netessine (2014). "When does the devil make work? An empirical study of the impact of workload on server's performance." *Management Sci.* **60**(6): 1574-1593.
- Taylor, F. W. (1911). *The Principles of Scientific Management*. New York, Harper & Brothers.
- Thompson, R. F. and W. A. Spencer (1966). "Habituation: A model phenomenon for the study of neuronal substrates of behavior." *Psych. Rev.* **73**: 16-43.
- Walker, J., W. Sistrunk, M. A. Higginbotham, K. Burks, L. Halford, L. Goddard, . . . V. Sanchez (2014). "Hospital hand hygiene compliance improves with increased monitoring and immediate feedback." *American Journal of Infection Control* **40**(5): e104.
- Ward, M. A., M. L. Schweizer, P. M. Polgreen, K. Gupta, H. S. Reisinger and E. N. Perencevich (2014). "Automated and electronically assisted hand hygiene monitoring systems: A systematic review." *American Journal of Infection Control* **42**(5): 472-478.
- Webb, E. J., D. T. Campbell, R. D. Schwartz and L. Sechrest (1966). *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago, Rand McNally.
- Wood, W. and D. T. Neal (2007). "A new look at habits and the habit-goal interface." *Psych. Rev.* **114**(4): 843-863.
- World Health Organization (2009). *WHO Guidelines on Hand Hygiene in Healthcare*. Geneva, Switzerland, World Health Organization Press.
- Yoeli, E., M. Hoffman, D. G. Rand and M. A. Nowak (2013). "Powering up with indirect reciprocity in a large-scale field experiment." *Proceedings of the National Academy of Sciences* **110**(Supplement 2): 10424-10429.

Tables and Figures

Table 1. Description of control variables included in our primary regression analyses.

Name	Description
Month Indicator	To control for time trends, we constructed indicator variables for each month-year pair in our data. For the correlation tables (Table 2), we created a continuous variable, <i>months since start</i> , to indicate the distance in months between the first month of our data and a given observation.
Day of the Week	To control for the previous observation that workers wash their hands less frequently during the week than on weekends (WHO, 2009), we constructed indicator variables for each day of the week (Tuesday through Sunday; Monday is the omitted indicator in our analyses).
<i>Control Variables Listed Below Were Only Included in Our Analyses of Caregiver-Level Compliance</i>	
Hours at Work	Prior research shows that hand hygiene compliance rates decrease over the course of a single caregiver shift (Dai et al., 2014). Thus, we created a variable to indicate the number of hours elapsed since the start of a caregiver's shift.
Hours off Work	Past research has shown that the longer a caregiver's break between consecutive work shifts, the higher her compliance will be during her subsequent shift (Dai et al., 2014). Thus, we created a variable to capture the number of hours between a given caregiver's shift and her preceding shift.
Duration of a Patient Room Stay (in hours)	Prior research suggests that caregivers' duration of contact with patients matters: caregivers who are in contact with patients for more than two minutes are more likely to wash their hands than caregivers who are in contact with patients for less than two minutes (WHO, 2009). Thus, we controlled for the duration of a patient room stay.
Room Entry Indicator	Several studies have shown that compliance with hand hygiene guidelines is lower prior to patient care than following patient care (WHO, 2009). We constructed an indicator variable that was equal to one if a hand hygiene opportunity occurred at room entry; otherwise, this indicator was set equal to zero.
Hour of the Day	To control for the possibility that workflow differs at different times of the day or that people's daily circadian rhythms influence their energy levels and attentional resources, we constructed indicator variables for each hour in the 24 hour clock (00:00 hours through 23:00 hours; 01:00 hours is the omitted indicator in our analyses).
<i>Control Variables Listed below Were only Included When We Predicted Caregiver Compliance during the Post-activation Period</i>	
Presence of Other Caregivers	We relied on the time stamps of room entries and exits to identify whether other caregivers were in the same patient room when the caregiver in question entered or exited the room. We created an indicator variable which equaled one if at least one other caregiver was in the same patient room when the caregiver in question entered (or exited) the room and had a hand hygiene opportunity.

COMPLIANCE AND MONITORING

Table 2. Means, standard deviations, and correlations for all variables included in our analyses

Panel A: Unit-Level dispenser usage data ($N = 35,552$)

	Mean	Standard Deviation	Total Daily Usage Per Unit Bed	Activation	Deactivation	Pre-activation Usage	Months Since Start	Day of the Week
Total Daily Usage Per Unit Bed	48.05	38.60	1.00					
Activation	0.88	0.33	0.21*	1.00				
Deactivation	0.01	0.12	-0.06*	-0.32*	1.00			
Pre-activation Usage	26.13	14.95	0.55*	-0.02*	0.04*	1.00		
Months Since Start	30.87	8.33	0.25*	0.23*	0.12*	0.11*	1.00	
Day of the Week	4.00	2.00	-0.07*	0.00	0.00	0.00	0.00	1.00

* p-value < .05

Panel B: Caregiver compliance data ($N = 19,585,738$)

	Mean	Standard Deviation	Compliance	Months Since Activation	Months Since Start	Day of the Week	Hours at Work	Hours off Work	Duration of a Patient Room Stay	Room Entry Indicator	Hour of the Day	Presence of Other Caregivers
Compliance	0.45	0.50	1.00									
Months Since Activation	313.12	237.71	0.15*	1.00*								
Months Since Start	31.24	7.77	0.30*	0.38*	1.00*							
Day of the Week	3.88	1.95	0.00	0.00*	0.00*	1.00*						
Hours at Work	5.49	9.70	-0.01*	-0.04*	0.00*	0.01*	1.00*					
Hours off Work	65.85	184.84	-0.02*	0.02*	-0.01*	-0.03*	-0.01*	1.00*				
Duration of a Patient Room Stay	0.09	0.15	0.03*	0.02*	0.03*	0.00	0.01*	0.01*	1.00*			
Room Entry Indicator	0.50	0.50	-0.09*	0.00	0.00	0.00	0.00*	0.00	0.00	1.00		
Hour of the Day	11.95	6.27	-0.01*	-0.01*	0.00	-0.02*	-0.03*	0.00*	0.01*	-0.01*	1.00	
Presence of Other Caregivers	0.10	0.30	0.00*	-0.01*	0.02*	-0.01*	0.01*	0.00*	0.03*	0.00	0.00*	1.00

* p-value < .05

Note. In Table 2 Panels A-C, months since start, day of the week, and hour of the day are treated as continuous variables. Our regression models include dummy indicators for each month, day of the week, and hour of the day as described in Table 1.

COMPLIANCE AND MONITORING

Table 3. Model 1 examined the effects of activating individual electronic monitoring on total daily usage of hand hygiene dispensers per hospital unit per bed, and Model 2 examined whether a unit's pre-activation dispenser usage affected the monitoring effect. Model 3 examined the long-term effects of activating electronic monitoring on all caregivers' individual hand hygiene compliance decisions. Model 4 examined the effects of deactivating electronic monitoring on total daily usage of hand hygiene dispensers per hospital unit per bed.

Regression Outcome Variable	Total Daily Usage Per Unit Bed		Did Caregivers Sanitize Their Hands at a Given Hand Hygiene Opportunity? (Y=1, N=0)	Total Daily Usage Per Unit Bed
	All Units		All Caregivers	All Units
Sample	Model 1	Model 2	Model 3	Model 4
Activation	16.34*** (3.14)	16.26*** (3.01)		16.36*** (3.15)
Pre-activation Usage (Mean Centered) X Activation		0.49* (0.21)		
Months Since Activation (Mean Centered)			4.6e-03* (1.8e-03)	
Months Since Activation ²			-2.0e-04* (8.3e-05)	
Deactivation				-17.07** (6.28)
Timing Controls¹	Yes	Yes	Yes	Yes
Hand Hygiene Opportunity Controls²	N/A	N/A	Yes	N/A
Hospital Unit Fixed Effects	Yes	Yes	N/A	Yes
Caregiver Fixed Effects	N/A	N/A	Yes	N/A
Observations	35,048	35,048	19,585,738	35,552
Number of Hospital Units	71	71	71	71
Number of Caregivers	N/A	N/A	5,222	N/A
R²	0.65	0.65	0.26	0.64

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively. Standard errors are clustered at the hospital unit level.

¹ Timing controls include fixed effects for each month in our data and day-of-the-week fixed effects.

² Hand hygiene opportunity controls include hour-of-the-day fixed effects, the number of hours a caregiver had been at work since the start of a shift up to an hand hygiene opportunity, the length (in hours) of a work break from the preceding shift, the duration of a patient room stay, and whether or not a hand hygiene opportunity was associated with a room entry or a room exit. Models 4 and 5 also control for whether other caregivers were present at a given hand hygiene opportunity. Coefficients on all control variables are available from the authors upon request.

Table 4. Badged, Unbadged and Total Usage during visitor-free hours and throughout the entire day.

Time of the Day	During visitor-free hours (12am-6am)		Across the day	
	Pre-activation	Post-activation	Pre-activation	Post-activation
Badged Usage	0.4 (pilot testers)	3.6 (caregivers wearing a badge)	3.2 (pilot testers)	22.5 (caregivers wearing a badge)
Unbadged Usage	3.7 (caregivers who were not pilot testers)	3.4 (caregivers not wearing a badge)	25.1 (caregivers who were not pilot testers + non-caregivers)	23.4 (caregivers not wearing a badge + non-caregivers)
Total Usage	4.1 (all caregivers)	6.9 (all caregivers)	28.3 (all caregivers + non-caregivers)	45.9 (all caregivers + non-caregivers)

Notes: The numbers are mean values across all units for the 17 days before and 18 days (including the activation date) after monitoring activation. Usage refers to usage per unit per bed. Text in parentheses indicates the population of users. During visitor-free hours, we assume that caregivers (with or without a badge) are the primary users of dispensers.

COMPLIANCE AND MONITORING

Table 5. Models from Table 3 with the inclusion of intervention indicators.

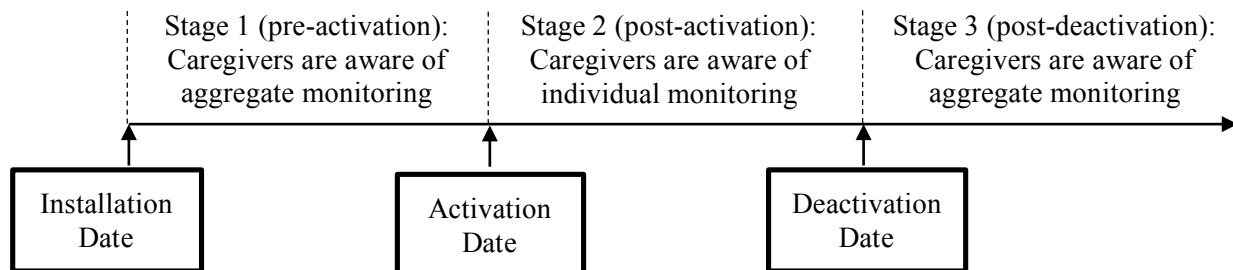
Sample	All Units		All Caregivers	All Units
	Model 1	Model 2	Model 3	Model 4
Activation	15.91*** (3.00)	15.23*** (2.59)		15.67*** (2.97)
Pre-activation Usage (Mean Centered) X Activation		0.54* (0.25)		
Days Since Activation (Mean Centered)			5.6e-03*** (1.5e-03)	
Days Since Activation ²			-2.5e-04*** (6.7e-05)	
Deactivation				-14.87* (7.08)
Indicators for additional interventions				
Performance feedback	-3.94 (3.67)	-3.73 (3.73)	-0.01 (0.02)	-3.96 (3.67)
Goal setting	5.05 (4.03)	4.81 (4.09)	0.01 (0.03)	5.84 (3.90)
Leadership	20.07*** (5.42)	19.79*** (5.29)	0.13*** (0.03)	20.27*** (5.48)
Competition	9.20 (8.55)	9.00 (8.13)	0.06* (0.02)	8.28 (7.98)
Incentive	7.73 (6.29)	7.91 (6.16)	0.04 [^] (0.02)	6.02 (5.97)
Time Controls ¹	Yes	Yes	Yes	Yes
Hand Hygiene Opportunity Controls ²	N/A	N/A	Yes	N/A
Hospital Unit Fixed Effects	Yes	Yes	N/A	Yes
Caregiver Fixed Effects	N/A	N/A	Yes	N/A
Observations	30,375	30,375	17,737,536	30,879
Number of Hospital Units	50	50	50	50
Number of Caregivers			4,211	
R²	0.66	0.67	0.26	0.66

[^], *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% levels, respectively. Standard errors are clustered at the hospital unit level.

¹ Timing controls include fixed effects for each month in our data and day-of-the-week fixed effects.

² Hand hygiene opportunity controls include hour-of-the-day fixed effects, the number of hours a caregiver had been at work since the start of a shift up to an hand hygiene opportunity, the length (in hours) of a work break from the preceding shift, the duration of a patient room stay, and whether or not a hand hygiene opportunity was associated with a room entry or a room exit. Model2 also controls for whether other caregivers were present at a given hand hygiene opportunity.

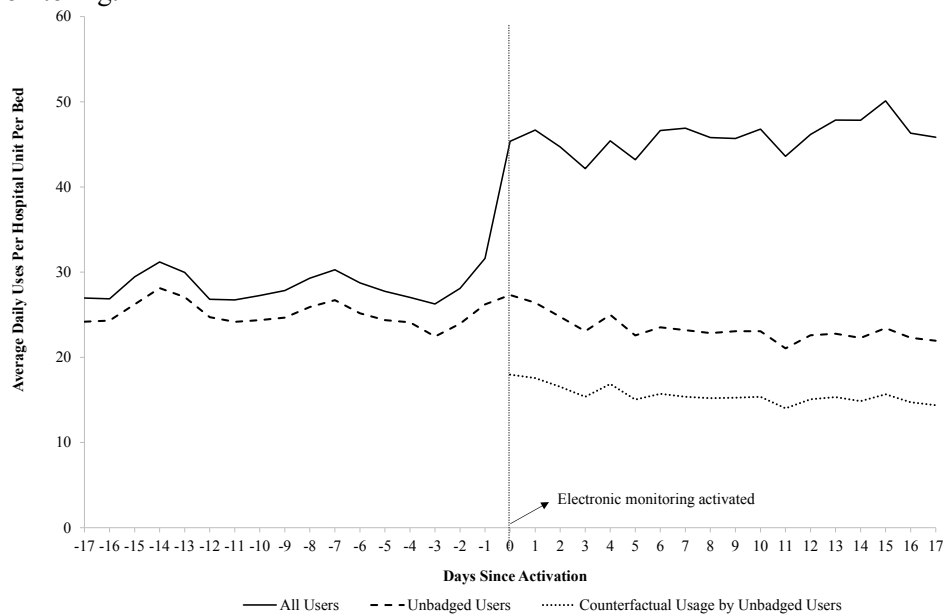
Figure 1. Timeline illustrating when the electronic monitoring system is installed, and when individual monitoring is activated and (in a small subset of cases) de-activated in a given hospital unit.



Note. Information sessions were held in advance of or at the time of when the majority of caregivers received an RFID badge prior to the activation date.

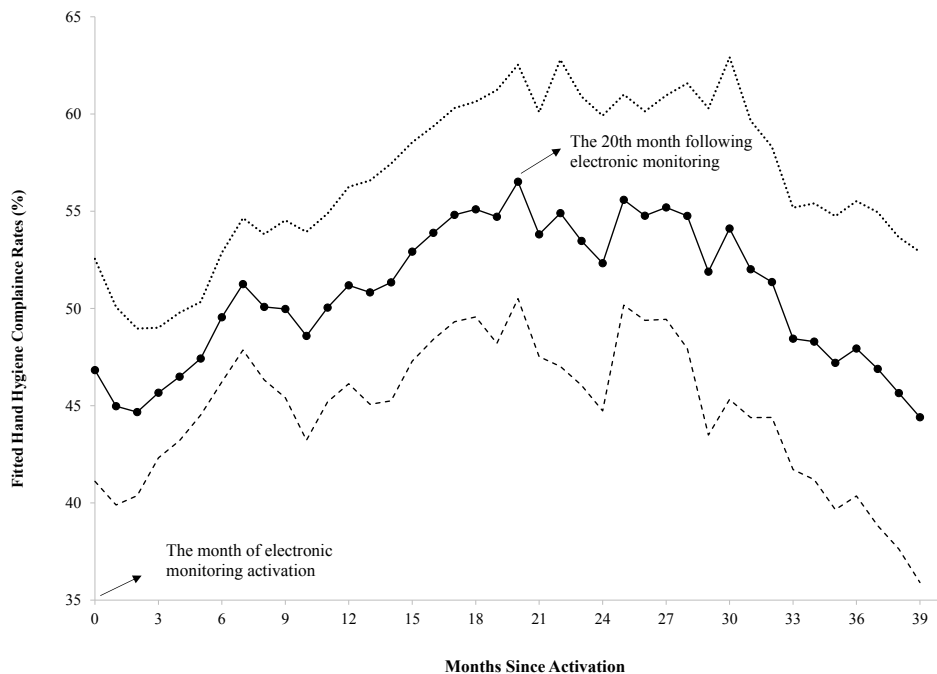
COMPLIANCE AND MONITORING

Figure 2. A plot of the average total daily dispenser usage per hospital unit per bed, average daily unbadged usage, and average daily counterfactual unbadged usage around the activation of individual electronic monitoring.



Note. Counterfactual unbadged usage reflects an estimate of what unbadged usage would have been after monitoring activation if monitoring activation did not alter unbadged users' compliance (see Appendix D for details).

Figure 3. Plot of fitted hand hygiene compliance rate as function of months since activation.



Note: Continuous control variables are assigned their mean value. Fixed effects for indicator variables are assigned the value of the relevant reference group omitted from our regressions.