Do Incentivized Behaviors Become Habits?

Susanna Gallani

Harvard Business School
Accounting and Management Unit

gallani@hbs.edu

October 2016
Preliminary Draft

Please do not cite without consent by the author

Acknowledgments:
This paper has benefited from the useful comments and suggestions offered by Katie Coffman, Lauren Cohen, Shane Greenstein, Ayelet Israeli, Bob Kaplan, Liz Keenan, Rem Koning, Greg Sabin. I would like to especially thank Robert (Bob) Kaplan for facilitating the connection with the Hospital. This research project could never be performed without the support, availability, and collaboration of several staff members of the Hospital, to whom I am sincerely grateful.
Do Incentivized Behaviors Become Habits?

Abstract:

Organizations often introduce temporary incentive programs with a view of establishing long lasting behaviors. While economic theory assumes that monetary incentives facilitate the implementation of desired practices by increasing the marginal value of performing the preferred tasks, crowding-out theory predicts that monetary award programs will reduce the motivation to perform incentivized behaviors, especially after the removal of the incentive. Using performance data on hand hygiene from a California hospital, I find that a focused incentive program, implemented to motivate a behavior that is strongly aligned with the mission of the organization, produces performance improvements that last beyond the end of the incentive period. These results are in line with theoretical predictions of habit formation, where the transformation of a deliberate behavior into an automated cognitive process decouples it from the monetary value of the award. Additionally, the probability and magnitude of habit formation, as well as its persistence are positively influenced by the initial response to the incentive program.

Keywords: Habits, Persistence of Performance Improvements, Crowding Out, Incentives, Healthcare

Data availability: The data used in this study is subject to a confidentiality agreement and cannot be shared without express consent of the hospital’s legal representatives.
1 - Introduction

Organizations use monetary incentives to motivate the adoption of new behaviors, or to stimulate performance improvements of existing practices. These initiatives often identify specific targets and deadlines, as well as a time when the monetary incentive will be paid. While the success of the incentive program is generally measured by the achievement of the goals by the set deadlines, incentive programs are often implemented with a spirit of introducing or improving certain behaviors in a sustainable way, beyond the deadline for the incentive payment. Whether and how incentivized behaviors persist after the associated explicit monetary incentive is removed is an open empirical question that I address in this study.

Monetary payoffs conditional on the achievement of identified targets serve the role of aligning the objectives of principal and agent by remunerating the agent for the cost of performing activities that would, otherwise, reduce their utility, as well as for the risk incurred by the agent in performing those actions (Stiglitz 1974; Sappington 1991). A large body of literature provides evidence of a positive relation between incentives and performance, and posits that incentives are a necessary condition for the observation of such activities. However, one of the downsides of monetary incentive programs is that they create explicit links between behaviors that serve the interests of the organization and monetary values (pay for performance). The exchange of actions for money becomes transactional. Researchers have observed that linking monetary incentives to behaviors that would be, otherwise, adopted or performed on a voluntary basis, has the undesired effect of reducing the performance of those behaviors. This phenomenon, known in the literature as the crowding-out effect of explicit incentives on intrinsic motivation (Kreps 1997; Deci 1971; Fehr and Gachter 2001) also implies that, once the
transactional link has been established, in absence of the monetary incentive, the desired behavior will no longer be observed.

In many cases, performance improvement initiatives depend on processes entailing repetitive actions (e.g. compliance to safety protocols, utilization of a new ERP system, adherence to professional standards of practice, etc.). Research in psychology and in neuroscience shows that learning mechanisms reduce the difficulty experienced by the individual in performing novel tasks. Feedback processes provide information to the agent about the mapping between their actions and outcomes, thus inducing Bayesian updates that inform subsequent choices of actions (van Houwelingen and van Raaij 1989). Rewards associated with the performance of desired behaviors cause neural representations that bolster individual motivation toward pursuing goals, thus reinforcing habit formation (Kelley and Berridge 2002). The repetition of actions yielding positive feedback over time can generate a habit, an acquired behavior pattern that becomes almost automatic (Aarts et al. 1997; Dunlap 1928; Lally et al. 2010). Being the result of an automated cognitive process, habits require significantly lower levels of effort and are, therefore, less costly to the individual. Decisions about habitual activities are made in a quasi-mindless way (Aarts et al. 1997). Therefore, the disutility associated with a habit is likely lower than the disutility associated with a required action, and might require a lower – if any – level of explicit incentives.

This study is based on the analysis of hand hygiene performance data from a California hospital specialized in joint replacements (herein: Hospital). Hand hygiene is a critical factor in the prevention of infections (Haas and Larson 2007; Sax et al. 2009; Gould et al. 2008). To incentivize and monitor hand hygiene performance, the Center for Medicare and Medicaid Services (CMS) requires hospitals to report hand hygiene data as one of the factors that
determine the rating of individual hospitals. 1 In addition, CMS performs periodic audits of hand hygiene performance as part of its Quality Assessment and Improvement Program (QAPI). Hospital ratings are expressed with a number of stars from 1 to 5, where 5 stars indicate excellent performance, based on the relative evaluation of a composite of up to 64 performance measures assessed at the hospital level. Hand hygiene is measured by the degree of compliance with the practice of sanitizing one’s hands when entering and exiting a patient care area (e.g. a hospital room).

Hand sanitizing is a widely adopted practice in healthcare facilities. Hospitals’ management stresses the importance of good hand hygiene practices sharing research findings and statistical results about the relation between hand hygiene and infection prevention, as well as publishing periodic reports on internal hand hygiene performance. A reduced incidence of infections is also correlated with fewer complications and lower readmission rates. Members of healthcare organizations are generally aware of the importance of this practice, which is in line with general organizational goals, and facilitated by the presence of sanitizer dispensers in many locations within the facility. The expectation of adherence to hand hygiene practices is, in general, very high.2

In 2015 the management team at Hospital introduced a temporary incentive program focused on the improvement of three key metrics related to the CMS star rating, with a purpose of “boosting” the level of performance on such metrics. These included hand hygiene, quietness of the environment, and effectiveness of communication about medications. The management team set specific goals for each measure and established a lump sum monetary incentive payable

---

1 Source: [http://www.hcahpsonline.org/Star Ratings.aspx](http://www.hcahpsonline.org/StarRatings.aspx)
2 At the time of this study, hospitals would need to achieve at least 95% compliance with respect to hand hygiene to receive a 5-star rating by CMS.
to each employee at the end of the incentive period upon achievement of the stated goals. Performance on all three metrics was reported at the organization level. Individual level data were collected only with respect to hands hygiene.

Results of statistical tests reveal that the incentive program achieved the goal of improving hands hygiene performance not only during the incentive period, but also in the quarter subsequent to the end of the incentive period, when no additional bonus would be paid for hand hygiene performance. Based on the evidence in the data, I conclude that the practice of sanitizing hands upon entry and exit from locations of contact with patients was converted from an incentivized behavior into a habit. Additionally, the data show that, at the individual level, both the likelihood of habit formation, as well as the magnitude of the habit effect depend on the direction and magnitude of the response to the incentive program (incentive effect). Similarly, the incentive effect influences the persistence of the habit over the subsequent quarter (i.e. two quarters after the end of the incentive period).

This study offers several contributions to the literature on compensation and incentives. First, it provides evidence of conditions under which crowding-out effects are muted, and incentive programs can generate long term positive performance results. Second, it contributes to the literature on incentive design by examining a situation where the incentive program is limited in time, but the expectation is that the incentivized behavior continues further. Extant literature on compensation, especially on executive compensation, has yet to study the persistence of performance of behaviors linked to bonus plans beyond the vesting date of the incentive. Third, this study contributes to the literature on habit formation and habit persistence by examining the influence of external incentives on the establishment of habitual practices, and by linking the individual sensitivity to incentives with the likelihood and intensity of habit formation. Finally,
this study offers a contribution to practice, by highlighting how focused temporary incentive programs can generate persistent performance improvements.

The remainder of this paper is organized as follows. Section 2 presents a review of the extant literature relevant for this study, and develops the main hypotheses. Section 3 describes the field settings and the data, as well as the research design. Section 4 presents the results of statistical testing and develops inferences. The last section concludes.

2 - Prior literature and hypotheses development

Despite a long tradition of academic research on incentives, consensus is still lacking as to whether external incentives are effective motivators for desired organizational behaviors. Standard agency theory posits that, when ownership and control are separated in an organization, explicit incentives, together with monitoring activities, are essential to align goals between the agent and the principal, and to influence the agent to make choices in the best interest of the principal (Jensen and Meckling 1976; Sappington 1991; Stiglitz 1974). The agent’s utility function is typically strictly concave and monotonically increasing in income, and decreasing in effort (Fehr and Gachter 2001). This conjecture is based on the assumption that an agent (homo economicus) will act primarily to contribute to their own economic self-interest. Monetary incentives will therefore increase the attractiveness of the required actions and reduce the disutility associated with making choices in the best interest of someone else. Neoclassical economic theory, however, does not share this view, and posits that agents derive utility not only from maximizing their wealth, but also from contributing to organizational goals (Simon 1991). Psychologists and behavioral economist, on the other hand, provide arguments and empirical evidence of detrimental effects of tangible incentives on intrinsic motivation and on performance
(see Deci et al. (1999) for a review). Nonetheless, the institution of monetary incentive programs in organizations is a very common practice.

Monetary incentive programs observed in practice present a large variation in terms of structure, duration, amount and frequency of the reward, and goal specificity. This study focuses on incentive programs that are temporary and focused on a specific, quantifiable goal. That is, a one-time monetary reward is paid upon accomplishment of a stated goal. It is not uncommon for organizations to institute temporary focused incentive programs to facilitate the implementation of new practices or to boost performance on existing ones. In these cases, the role of the incentive program is to motivate desired behaviors, as well as enhancing the salience of the required tasks, especially when communication initiatives are insufficient to affect behavioral patterns (Verplanken and Wood 2006). Temporary monetary incentive programs are subject to two major sources of criticism. First, the monetary nature of the incentive raises concerns in terms of crowding-out effects on intrinsic motivation (Deci 1971; Gneezy et al. 2011). Second, the temporary nature of the reward appears to be at odds with the expectation of maintaining the new practices or new levels of performance beyond the duration of the incentive period (Greene and Podsakoff 1978; Gneezy et al. 2011).

The introduction of an incentive program changes the nature of the contract between the employee and the organization. Monetary incentives generate two effects. On the one hand, the promised award imposes a higher marginal cost on shirking, or increases the marginal benefit of performing (Frey and Jegen 2001). On the other hand, the introduction of extrinsic rewards generates psychological effects that might crowd-out agents’ intrinsic motivation, leading to lower levels of effort and worse performance (Kreps 1997). The notion of crowding-out refers to the phenomenon by which the introduction of extrinsic rewards linked to the performance of a
certain activity reduces the “attractiveness” of such activity in the eyes of the agent. Researchers supporting this theory propose two main explanations. On the one hand, the introduction of the reward changes the individual’s perception of locus of control from the individual to an external entity. Consequently, the individual experiences a reduction in their freedom of choice, as the behavior becomes driven by the reward instead of being driven by intrinsic motivation (Deci 1971). On the other hand, the association of the behavior with the reward introduces a monetary valuation of the behavior. If the valuation of the required effort is considered to be lower than the expected threshold (i.e. the value that the individual attributes to their voluntary cooperation), the exchange will be considered to be unfair, thus impacting negatively the probability for the principal to receive the desired behavior (Fehr and Gachter 2001). The effectiveness (i.e. obtaining the desired behavior) of an explicit incentive program depends on whether the relative price effect prevails over the crowding-out effect.

When incentive programs are temporary, whether the incentivized behavior will persist after the end of the incentive period is an open question. Gneezy et al. (2011) posit that the removal of an incentive is likely to exacerbate the crowding out effect. This is because, once agents update their beliefs with respect to the drivers of their behaviors and consider the exchange as a market transaction, they are likely to perceive the removal of the incentive as breach of contract, thus justifying their withholding of the behavior previously incentivized.

Often times, incentive programs aim at facilitating the introduction of new practices or at boosting performance for existing ones. In many cases these initiatives rely on the repetition of desired behaviors that are in line with the overall organizational goals. Research studies about learning processes posit that a sufficient number of repetitions of a certain behavior in consistent environmental settings transfers the control of the behavior to external environmental stimuli that
trigger an automated response (Lally et al. 2010). The behavioral pattern resulting from such automated cognitive process is a *habit* (Aarts et al. 1997). The cognitive process underlying a habitual behavior is significantly less costly than that driving a deliberate choice. Once established, automated processes are less likely to require external incentives, and the link between the behavior and the reward might be broken, thus making the crowding-out mechanism ineffective. I therefore formulate the following hypothesis:

**H1: An incentivized repetitive behavior persists as a habit after the removal of the specific incentive**

The prevalence between the relative price effect and the crowding-out effect in response to the introduction of an incentive program is likely to exhibit significant variation across members of an organization. Individuals who are sensitive to the crowding-out effect might revise their beliefs with respect to the locus of control of the incentivized behavior and may not improve (or even deteriorate) their performance during the incentive period. Others might be responsive to the relative price effect and adopt the desired behavior during the incentive period with a view of earning the promised reward. These individuals, however, are less likely to continue to perform the incentivized behavior once the incentive is removed, as a result of a perceived breach of contract by the principal. Focused incentive programs serve the role of raising awareness of the importance of the targeted behavior, and are often introduced as a means to improve performance beyond the duration of the incentive period. The formation of a habit is consistent with a long lasting effect of the intervention (Verplanken and Wood 2006). I posit that a positive response to the incentive is a necessary condition for the repetition of the behavior and sustained improved performance. The following hypothesis formulates the conjecture that the formation of a habit is facilitated by a positive response to the incentive program:
**H2: The formation of a habit is positively associated with the effectiveness of the specific incentive**

Persistence is a key characteristic of a habit, as well as a key indicator of the success of a focused incentive program. In his model of consumer behavior, Brown (1952) allows for the possibility that individuals react to changes in their economic situation with a lag. Building on Brown’s theoretical prediction, I question whether the persistence of a behavior immediately after the removal of the incentive might be the result of the formation of a habit, or due to a lagged response to the removal of the explicit link between behavior and reward. However, if the incentive program was effective in its goal of raising the salience of the desired behavior and generating the long lasting habit, the behavior should persist beyond a reasonable lag. I therefore test the following hypothesis:

**H3: The persistence of a habit is positively associated with the effectiveness of the incentive program**

Field data from a California for profit orthopedic hospital (herein: Hospital) were used to test the hypotheses formulated above. The next section describes the field settings, the data, and presents the research design.

3 – Data and Research Design

3.1 – Field settings

The study involves a natural quasi-experiment in a healthcare setting. In the fall of 2015 the management team at Hospital introduced a temporary incentive program to improve performance on a selected set of measures that were both aligned with the overall mission of the hospital to provide quality healthcare services, and included in the metrics reported as part of their star-rating, a public assessment of healthcare providers instituted by the Center for
Medicare and Medicaid Services (CMS) as part of the Affordable Care Act. The culture at Hospital had always been one of excellence, quality, and continuous improvement. Hospital had been rated with 5 stars (the highest possible rating) in the past, but recent assessments were showing a decline in their performance relative to their peers. Hospital’s management identified metrics needing improvement: hand hygiene, quietness of the environment, and communication about medication.

Performance on quietness of the environment and communication about medication would be assessed based on HCAHPS survey results, but presented important limitations with respect to the purpose of this study. First, HCAHPS would be subject to the limitations typical of any survey instrument, including respondent bias, partial recall, low respondent rate, etc. Second, data would refer uniquely to the organization as a whole. Hand hygiene data, in contrast, would be measured at the individual level and then aggregated to the organization level for reporting to CMS. Additionally, hand hygiene was not a new practice for the members of Hospital. A robust body of research provides evidence of important causal relations between hand hygiene and infection prevention (Sax et al. 2009; Haas and Larson 2007; Gould et al. 2008; Pessoa-Silva et al. 2007; Lankford et al. 2003; Pittet et al. 1999), and an official document issued by the Joint Commission in 2009 details best practices for hand hygiene adherence to recommended

---

3 The CMS star-rating is calculated based on data collected via HCAHPS surveys (http://www.hcahpsonline.org/home.aspx), as well as hospitals’ self reported data. Cluster analyses are performed on a composite measure including up to 64 indicators. Star ratings are assigned to each cluster, where 5 stars represent the best performing hospitals, while 1 star is assigned to the hospitals in the lower end of the distribution. Additional information at https://www.medicare.gov/hospitalcompare/Data/Data-Updated.html#.

4 See http://www.hcahpsonline.org/home.aspx for detailed information on HCAHPS survey methodology, scope, and availability of data to the public.
standards. Furthermore, the performance on this particular behavior is completely under control of the organization members and not subject to stochastic factors that could influence its measurement.

Hand hygiene performance is measured as the degree of compliance with the practice of sanitizing hands when entering and exiting a patient care area. That is, hospital members are supposed to *gel-in* and *gel-out* using the sanitizer dispensers available throughout the facility every time they enter or exit a location where they might have contact with a patient. Performance data were collected via randomized audits performed by independent secret shoppers undisclosed to the hospital members. Baseline data referring to the two months prior to the incentive period indicated an average performance of 91% overall, with some departments rating as low as 86%. Goals were set at a threshold level of 92% and a stretch level of 95%. Performance was measured at the Hospital level, including Hospital employees and Hospital contractors (e.g. physicians cannot be employees of the hospital by California law). While all members of the organization were accountable for the improvement in hand hygiene performance, only Hospital employees were eligible for the payment of the incentive (see Table 1). A one-time bonus of $2,500 would be paid to each employee of Hospital upon achievement of the stretch goal by the end of the next fiscal quarter (Q2).

As part of their regular operations, Hospital management produced a quarterly scorecard including multiple measures of performance, including the three metrics that were chosen as the focus of the incentive program. To enhance the awareness about the incentive program and to

5 Source: https://www.jointcommission.org/sustaining_and_spreading_improvement_in_hand_hygiene_compliance/
6 Management clearly communicated that the incentive program would be of temporary nature, so not to set the expectations of future payments of monetary rewards with respect to the three metrics included in the program.
sustain improvements toward the stated goals, during the incentive period additional
communication about the measures included in the incentive plan was provided to all employees
via by-weekly emails reporting updates on the performance levels and distances to the goals.
After the end of the incentive period, communication returned to the original practice of
reporting performance via quarterly scorecard.

During our conversations, members of the management team described some important
concerns related to the structure of the incentive program. First, because of the asymmetry
between accountability and reward eligibility across the members of the organization there was a
concern with respect of the efficacy of the incentive program. Second, setting the goal at the
organization level while paying the incentive at the individual level would generate opportunities
of free riding. Third, the team was unsure as to whether the boost in performance would persist
after the end of the incentive period.

3.2 - Research design

Hospital provided individual level data of hand hygiene over 4 quarters. Observations
reported in Q1 relate to the period preceding the introduction of the incentive program (pre-
period), Q2 data relate to the quarter during which the incentive initiative took place (incentive
period), performance data reported in Q3 and Q4 related to the two quarters subsequent to the
removal of the incentive (post-period and post-post period). Individuals’ identities were
disguised and replaced with alphanumerical ID’s. Data included the quarterly number of
assessments of gel-in and gel-out for each employee, the number of successful gel-in’s and gel-
out’s, the department, and organizational role. Following the methodology employed by
Hospital, hand hygiene performance ($HH$) was calculated as:

$$HH_{it} = \frac{gel_{in_{it}} + gel_{out_{it}}}{gel_{in_{assess_{it}}} + gel_{out_{assess_{it}}}}$$ (1)
where \( i \) represents the individual member of the organization, and \( t \) represents the quarter of assessment, \( gel_{in} (gel_{out}) \) is the number of observations of individuals sanitizing their hands upon entry (exit) of the patient care area, and \( gel_{in\_assess} (gel_{out\_assess}) \) is the number of observations of hand hygiene behavior upon entry (exit) of the patient care area\(^7\). Because the focus of this study is to explore the effect of temporary incentive programs on the formation of habits, the sample was limited to the employees that were eligible to receive the award upon achievement of the goal. This decision was driven by the awareness that changes in the performance of organization members subject to the initiative but ineligible to receive the bonus is likely driven by different dynamics, which should be the subject of a separate study. Table 1 reports the criteria for sample selection\(^8\).

--- Insert Table 1 here ---

In order to assess whether the incentivized behavior persists as a habit after the removal of the incentive, I compare values of HH reported in Q3 (quarter after the removal of the incentive) to those reported in Q1 (baseline, or pre-period). The measure of hand hygiene habit is formulated as follows:

\[
HH_{\_Habit_i} = HH_{i3} - HH_{i1}
\]

where \( HH_{i3} \) is the individual performance of individual \( i \) reported at the end of Q3 and \( HH_{i1} \) is the individual’s performance prior to the incentive program. A positive value of \( HH_{\_Habit} \)

\(^7\) Because of the presence of multiple secret shoppers and the performance of daily randomized audits, members of the organization can be subject to multiple assessments in each period. The data shared by Hospital did not include the identity of the auditor or the exact data of the assessment (see table 1 – panel B)

\(^8\) Of 464 total members of the organization only 351 were subject to the assessment of hand hygiene at least once during the four quarters. The remaining 113 were excluded from the sample as there would be no information about their hand hygiene performance. Based on conversations with Hospital’s representatives, these individuals server roles that do not involve entering patient care areas as part of their regular duties.
indicates that hand hygiene performance after the removal of the incentive is higher than the performance measured before the incentive program. That is, individuals continue to perform the required behavior with a higher degree of compliance even after the removal of the promise of a monetary reward. A negative value for $HH\_Habit$ would be consistent with crowding-out theories, whereby the removal of the incentive reduces the motivation to perform the behavior to a lower level than when the contract between organization and employees did not include any explicit reward.

Whether the habit formation is a result of the incentive program and not just an unrelated trend in hand hygiene performance can be assessed by relating the likelihood of forming a habit with the likelihood of responding positively to the incentive, as represented by the following logit model:

$$\text{prob}(HH\_Habit > 0)_i = \alpha + \beta_1 HH\_Pos\_Inci + \beta_2 HH_i + \beta_3 Gender_i + \beta_4 Tenure_i + \epsilon \quad (3)$$

where $HH\_Pos\_Inc$ is an indicator variable that assumes the value of one if individuals respond positively to the incentive program, $Gender_i$ is an indicator variable assuming the value of 1 if the employee is female, and $Tenure_i$ measures the number of years the employee has been part of the organization, measured at the end of the observation period (Q4). The incentive response is measured by:

$$HH\_Incentive_i = HH_{i2} - HH_{i1} \quad (4)$$

where $HH_{i2}$ is the hand hygiene performance reported for individual $i$ at the end of the incentive period and $HH_{i1}$ is the hand hygiene performance during the pre-period. In other words, $HH\_Pos\_Inc$ will assume the value of one if $HH\_Incentive$ assumes a positive value.

Additionally, the individual reaction to the incentive is likely to drive the magnitude of the habit effect, which I measure with the following model:
\[ HH\_Habit_i = \alpha_i + \beta_1 HH\_Incentive_i + \beta_2 HH_{i1} + \beta_3 Gender_i + \beta_4 Tenure_i + \epsilon \]  

(5)

In both models (3) and (4) I include the initial hand hygiene performance as a control variable, as it is likely that the level of initial performance might drive the reaction to the incentive program.

An important objective for the management team at Hospital is that the incentivized performance persists in the long run. The following metric approximates the persistence of the hand hygiene habit:

\[ HH\_Persist_i = HH_{i4} - HH_{i1} \]  

(6)

where \( HH_{i4} \) is the hand hygiene reported for individual I at the end of Q4 and \( HH_{i1} \) is the performance reported for individual I in the quarter prior to the incentive program. A positive value for \( HH\_Persist \) indicates that the hand hygiene performance after 2 quarters subsequent to the removal of the incentive is still at higher levels than what was reported prior to the introduction of the incentive. With a similar logic to that described above with reference to the habit effect, I test the likelihood and the magnitude of the persistence of the habit with the following models:

\[ prob(HH\_Persist > 0)_i = \alpha + \beta_1 HH\_Pos\_Inci + \beta_2 HH_{i1} + \beta_3 Gender_i + \beta_4 Tenure_i + \epsilon \]  

(7)

\[ HH\_Persist_i = \alpha_i + \beta_1 HH\_Incentive_i + \beta_2 HH_{i1} + \beta_3 Gender_i + \beta_4 Tenure_i + \epsilon \]  

(8)

Descriptive statistics for the main variables of interest are reported in Table 2.

--- Insert Table 2 here ---

The following section describes the results of the estimation of the above models and proposes inferential conclusions based on statistical tests.

4 – Results of Statistical Tests of Hypotheses

The first hypothesis (H1) predicts that the incentive program succeeds in converting the incentivized behavior into a habit. The data utilized in this study support the prediction. The
means reported in Table 2, Panel A indicate that the average hand hygiene performance reported at the end of Q3 is higher than it was reported in Q1. A paired t-test between the two quarters (Table 3, Panel A) shows that the average performance in Q3 is significantly higher than that in Q1, which indicates a positive and significantly different than zero value (p<0.05, two-tailed) for $HH_{Habit}$, defined as the difference between the performance at Q3 and that at Q1. Note that, because the paired t-test compares means from a within-subjects test group, this test considers only subjects for whom hand hygiene performance was observed in both Q1 and Q3. This is the reason why the number of observations and performance score means reported in this panel differ from those reported in the descriptive statistics (Table 2, Panel A).

The second hypothesis (H2) explores the relation between the individual reaction to the introduction of the incentive program and habit formation. Table 3, Panel B reports the results of the estimation of the logistic regression described by model (3), which show that a positive response to the incentive program (that is, a positive difference between performance in Q2 and performance on Q1) is positively associated with the probability of forming a habit ($\beta_1 = 2.194$, $p<0.01$ two-tailed), thus confirming the prediction in H2. Additionally, the estimation of model (3) shows that the probability of forming a habit subsequent to the end of the incentive program is inversely related to the level of performance prior to the incentive program ($\beta_2 = -9.946$, $p<0.01$ two-tailed). That is, worse initial performers tend to maintain a higher level of hand hygiene behavior after the incentive program is removed. Finally, results of OLS estimation of model (5) with heteroskedasticity robust errors, reported in Table 3, Panel C, show a positive association between the magnitude of the response to the incentive program and the magnitude to the habit effect ($\beta_1 = 0.514$, $p<0.05$ two-tailed), as well as a further confirmation that the habit
effect is negatively correlated with the initial performance ($\beta_2 = -0.405$, $p<0.10$ two-tailed).\(^9\)

Taken together these results support H2, by which the transformation of the incentivized behavior into a habit is influenced by the degree of positive reaction by the individual to the introduction of an incentive program.

To examine the persistence of the habit beyond the first quarter subsequent to the incentive removal, I first compare the mean hand hygiene performance in Q4 with the average performance in Q1. Univariate analyses (Table 2, Panel A and B), as well as the results of the paired t-test reported in Table 4, Panel A, are inconclusive with respect to the expectation that performance improvements will persist in Q4. The difference between hand hygiene performance means in Q4 and Q1 is not significantly different than zero (Table 4, Panel A).

However, the analysis of the relation between incentive effect and persistence of the habit effect supports the prediction formulated with H3. In fact, the estimation of model (7) using logistic regression with heteroskedasticity robust standard errors (Table 4, Panel B) shows that the likelihood of habit persistence is positively and significantly correlated with the probability of reacting positively to the incentive program ($\beta_1 = 1.634$, $p<0.01$ two-tailed). Also, the analysis of the coefficients estimated for model (8) via OLS regression with robust errors shows a positive and significant correlation between the size of the performance improvement due to the incentive program and its persistence ($\beta_1 = 0.517$, $p<0.05$ two-tailed). The initial performance is also a predictor of habit persistence, both in terms of likelihood ($\beta_2 = -11.416$, $p<0.01$ two-tailed for

\(^9\) In order to take into consideration the ceiling effect for the improvement and the fact that employees that start at a lower level of performance have greater room for improvement, I executed the same analyses scaling the DV (magnitude of habit effect or persistence effect) and the incentive effect by the potential improvement available in Q1. The results of the estimation of the models with this specification are equivalent (in direction, magnitude, and significance) to those reported above.
model 7), and in terms of magnitude ($\beta_2 = -0.543$, $p<0.10$ two-tailed). Taken together, these results support H3.

Because of the high historical hand hygiene performance at Hospital, a significant number of employees were already performing above the stated goal at the time the focused incentive was introduced. It is likely that those individuals might not perceive the need to improve their performance. Their behavior might, therefore, be impacted less by the incentive program. On the other hand, those individuals whose performance is below the stated goal at the beginning of the incentive period might be more responsive to the increased awareness and motivational effect of the incentive program. To validate this assumption, I restrict the sample to employees whose reported hand hygiene performance in Q1 is lower than 95% and repeat the estimations of the above models. Tables 5 and 6 report the estimation results. Results in Table 5, Panel A, further support H1, by showing that the incentivized behavior persists past the removal of the incentive ($HH_{Habit} > 0$, $p<0.01$ two-tailed). Additionally, these analyses confirm that the probability of habit formation (Table 5, Panel B) and the magnitude of the habit effect (Table 5, Panel C) are directly influenced by the response to the incentive effect (respectively $\beta_1 = 5.572$, $p<0.01$ two-tailed, and $\beta_1 = 0.783$, $p<0.01$ two-tailed). Table 6 reports the results of the estimations performed with respect to the persistence of the habit effect. The paired t-test provides evidence supporting H1 with respect to the subsample of employees that performed below the goal in Q1 ($HH_{Persist} > 0$, $p<0.01$ two-tailed). Panel B and C of Table 6 provide further support for the positive relation between individual response to the incentive program and
persistence of the habit (respectively $\beta_1 = 4.083$, $p<0.01$ two-tailed, and $\beta_1 = 0.533$, $p<0.01$ two-tailed).\textsuperscript{10}

In summary, the results of the statistical analyses performed in this study suggest that a temporary monetary incentive program can generate performance improvements that last beyond the end of the incentive period. These findings provide empirical evidence contrary to the predictions documented by numerous studies on the crowding-out influence of extrinsic incentives on intrinsic motivation, by which monetary incentives would deteriorate performance. In particular, my results show that responsiveness to the incentive program exhibit variation across individuals, and that a positive response to the incentive program is more likely to engage the formation of habitual behaviors persisting in the long run. That is, individuals that do not exhibit behaviors corresponding to the crowding-out prediction are more likely to convert the incentivized behavior into a sustainable habit.

V – Conclusions

This study explores the effect of temporary incentive programs on the formation of habitual behaviors. When organizations implement initiatives aiming at the establishment of new practices or at improving performance of existing ones, links between successful implementations and monetary rewards are common practice. In many cases management teams use these temporary initiatives to boost performance, with the expectations that the new behaviors and performance levels will persist beyond the incentive period.

An extensive body of academic literature provides arguments against the usefulness of extrinsic incentives for performance improvement purposes, a phenomenon known as the crowding-out effect of extrinsic rewards on intrinsic motivation. The contract modification

\textsuperscript{10} Using scaled specifications of the persistence effect and the incentive effect yields equivalent results.
introduced by the extrinsic incentive changes the interpretation of the relationship between the employee and the organization, transforming behaviors that might have been driven by voluntary cooperation into a transactional exchange. If the relative price established by the new contractual agreement is not in line with the employee’s expectations, or if the employee perceives the contractual change as a transfer of the locus of control for the behavior to the manager, the motivation for performing the behavior might reduce to a point where the contract without the incentive would be more efficient than the one with the monetary reward. The removal of the incentive, once the change in the contract has been internalized by the employee, generates further detrimental effects on intrinsic motivation and performance. This study provides evidence contrary to these predictions.

Using panel data on hand hygiene performance from a California for-profit orthopedic hospital, I show that a temporary performance improvement program involving the payment of a monetary reward upon achievement of the stated goals succeeds in establishing favorable behaviors that persist beyond the end of the incentive period. Additionally, my results indicate that the formation of a habit as a result of an incentive program depends on the individual responsiveness to the incentive program itself. In other words, for individuals whose intrinsic motivation is less subject to the detrimental effects of extrinsic rewards, habit formation is more likely and more sustainable over time.

This study is subject to several limitations. First, the particular nature of the incentivized behavior might lend itself to a higher likelihood of habit formation. Habits are automated behaviors that are triggered by environmental cues and require less cognitive effort than deliberate choices. The action targeted by the particular incentive program in these settings exhibits low complexity and requires little cognitive processing. Nonetheless, hand hygiene is a
critical factor influencing the incidence of infections, which is an important quality metric for every healthcare provider. Additionally, many other performance improvement initiatives might similarly depend on the performance of repetitive actions (e.g. compliance with safety procedures, utilization of standard protocols for data recording, implementation of quality assurance procedures, etc.), therefore the generalization of these results should not be considered to be excessively limited. Second, the design of the particular incentive program at Hospital allows for the occurrence of free riding. Further examination of the influence of such dysfunctional behaviors might shed more light on the potential consequences of such design choices. Third, the particular settings for this study did not allow to relate the amount of the monetary award to the observed effects of the initiative. A different design, where the amount of the bonus would depend on the individual, instead of collective, performance improvement would likely provide further interesting insights.

Despite its limitations, this study contributes to the debate about the usefulness of monetary incentives for performance improvement purposes by challenging the findings of crowding-out research and proposing that the influence of extrinsic incentives on performance improvement might exhibit a larger degree of variation than proposed by many esteemed scholars. Additionally, this study contributes to the research on incentive design and performance measurement by analyzing the effects of a temporary incentive program applied to rank-and-file members of the organization with a clear and objectively measurable objective for an activity that is completely under the control of the agent. Finally, while prior research on habit formation has been carried out primarily within health studies, education research, and psychology, this study addresses the habit formation phenomenon in a business setting, highlighting its relevance for academics and practitioners in the field of management accounting and compensation design.
References:


Table 1: Sample selection and number of hand hygiene assessments

<table>
<thead>
<tr>
<th>Panel A: sample selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total members of the organization 464</td>
</tr>
<tr>
<td>Less: members that were not assessed during the 4 quarters 114</td>
</tr>
<tr>
<td>Total members assessed during the 4 quarters 350</td>
</tr>
<tr>
<td>Less: non-employees - non eligible for incentive payment 86</td>
</tr>
<tr>
<td>Assessed Employees - eligible for incentive payment 264</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: number of individual hand hygiene assessments per quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand hygiene assessments per employee</td>
</tr>
<tr>
<td>Q1 = pre period</td>
</tr>
<tr>
<td>Q2 = incentive period</td>
</tr>
<tr>
<td>Q3 = post period</td>
</tr>
<tr>
<td>Q4 = post-post period</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>Employees’ demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>N</td>
</tr>
<tr>
<td>225</td>
<td>0.764</td>
</tr>
<tr>
<td>Tenure</td>
<td>N</td>
</tr>
<tr>
<td>225</td>
<td>7.338</td>
</tr>
</tbody>
</table>

Notes: Panel A reports the criteria used in selecting the sample for this study. Although hand hygiene performance was measured at the aggregate organizational level, Hospital’s members that were not employees (e.g. physicians) would not be eligible to receive the award payment, despite being accountable for their contribution to the overall result, and were excluded from the sample. Additionally, employees that, due to their organizational role not requiring direct contact with patients, were not subject to any assessment in any of the four quarters were also excluded. Panel B: due to the random audits performed by multiple secret shoppers, the number of assessment varied significantly across individual members. Panel C reports descriptive statistics for the characteristics of the employees. Gender is an indicator variable assuming the value of 1 if the employee is a female, and zero otherwise. Tenure is a variable indicating the number of years the employee has been part of the organization, calculated at the end of the observation period (end of Q4).
Table 2: Descriptive Statistics

Panel A  Hand hygiene performance by quarter

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHit</td>
<td>Q1</td>
<td>148</td>
<td>0.928</td>
<td>0.156</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>202</td>
<td>0.943</td>
<td>0.163</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>197</td>
<td>0.942</td>
<td>0.166</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>177</td>
<td>0.913</td>
<td>0.203</td>
<td>1.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Panel B  Hand hygiene incentive effect, habit effect and habit persistence

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH_Incentive_i</td>
<td>123</td>
<td>0.013</td>
<td>0.118</td>
<td>0.000</td>
<td>-0.500</td>
<td>0.385</td>
</tr>
<tr>
<td>HH_Habit_i</td>
<td>117</td>
<td>0.033</td>
<td>0.155</td>
<td>0.000</td>
<td>-0.400</td>
<td>1.000</td>
</tr>
<tr>
<td>HH_Persist_i</td>
<td>117</td>
<td>-0.009</td>
<td>0.169</td>
<td>0.000</td>
<td>-1.000</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Notes: Panel A reports descriptive statistics for the measure of hand hygiene performance for each of the four quarters. HHit is defined as $HH_{it} = \frac{gel_{in, it} + gel_{out, it}}{gel_{in, assess, it} + gel_{out, assess, it}}$. Panel B reports descriptive statistics for three measures of effectiveness of the incentive program, respectively defined as $HH\_Incentive_i = HH_{i2} - HH_{i1}$, $HH\_Habit_i = HH_{i3} - HH_{i1}$, and $HH\_Persist_i = HH_{i4} - HH_{i1}$. 
Table 3: Habit Effect – Pooled Sample

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Estimation of the average habit effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired t-test analysis</td>
<td></td>
</tr>
<tr>
<td><strong>Variable</strong></td>
<td><strong>N</strong></td>
</tr>
<tr>
<td>$HH_1$</td>
<td>117</td>
</tr>
<tr>
<td>$HH_3$</td>
<td>117</td>
</tr>
<tr>
<td>$HH_{Habit}$</td>
<td>117</td>
</tr>
<tr>
<td>$df$</td>
<td>116</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>2.302</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>Determinants of the probability of habit formation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logit regression with heteroskedasticity robust standard errors</td>
<td></td>
</tr>
<tr>
<td>$DV = \text{Prob}(HH_{Habit} &gt; 0)$</td>
<td></td>
</tr>
<tr>
<td><strong>Predictors</strong></td>
<td><strong>Coefficient</strong></td>
</tr>
<tr>
<td>$HH _ Pos _ Inc$</td>
<td>2.194</td>
</tr>
<tr>
<td>$HH_1$</td>
<td>-9.946</td>
</tr>
<tr>
<td>Gender</td>
<td>-1.541</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.025</td>
</tr>
<tr>
<td>Intercept</td>
<td>9.482</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>127</td>
</tr>
<tr>
<td>Pseudo R$^2$</td>
<td>0.399</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>Determinants of the magnitude of the habit formed</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS regression with heteroskedasticity robust standard errors</td>
<td></td>
</tr>
<tr>
<td>$DV = HH_ Habit$</td>
<td></td>
</tr>
<tr>
<td><strong>Predictors</strong></td>
<td><strong>Coefficient</strong></td>
</tr>
<tr>
<td>$HH _ Incentive$</td>
<td>0.514</td>
</tr>
<tr>
<td>$HH_1$</td>
<td>-0.405</td>
</tr>
<tr>
<td>Gender</td>
<td>0.007</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.002</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.366</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>95</td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.513</td>
</tr>
</tbody>
</table>

**Notes:** This table reports three analyses of habit formation. The variable $HH\_ Habit_i$ is calculated as the difference between $HH_{i3}$ and $HH_{i1}$. Panel A: the paired t-test compares means from a within-subjects test group; that is, this test considers only subjects for whom hand hygiene performance was observed in both Q1 and Q3; Panel B estimates the model $\text{prob}(HH\_ Habit > 0) = \alpha + \beta_1 HH\_ Pos\_ Inc + \beta_2 HH_1 + \varepsilon$ using logistic regression with robust standard errors; Panel C reports the OLS estimation of the model

$HH\_ Habit_i = \alpha_i + \beta_1 HH\_ Incentive_i + \beta_2 HH_{i1} + \varepsilon$ (robust standard errors). In all panels: * = p<0.10; ** = p<0.05; *** = p<0.01
Table 4: Habit Persistence – Pooled Sample

### Panel A Estimation of the average persistence effect
Paired t-test analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std dev</th>
<th>Std err</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH1</td>
<td>117</td>
<td>0.946</td>
<td>0.105</td>
<td>0.010</td>
</tr>
<tr>
<td>HH4</td>
<td>117</td>
<td>0.936</td>
<td>0.151</td>
<td>0.014</td>
</tr>
<tr>
<td>HH_Persist</td>
<td>117</td>
<td>-0.009</td>
<td>0.169</td>
<td>0.016</td>
</tr>
</tbody>
</table>

df | 116  

| t  | -0.594  |
| n.s. |

### Panel B Determinants of the probability of individual habit persistence
Logit regression with heteroskedasticity robust standard errors

\[ DV = \text{Prob}(HH_{Persist} > 0) \]

**Predictors** | **Coefficient** | **Std. error** | **z**  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HH_Pos_Inc</td>
<td>1.634</td>
<td>0.602</td>
<td>2.71  ***</td>
</tr>
<tr>
<td>HH1</td>
<td>-11.416</td>
<td>2.819</td>
<td>-4.05 ***</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.999</td>
<td>0.752</td>
<td>-1.33</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.004</td>
<td>0.039</td>
<td>-0.11</td>
</tr>
<tr>
<td>Intercept</td>
<td>9.844</td>
<td>2.938</td>
<td>3.35  ***</td>
</tr>
</tbody>
</table>

| N  | 127 |
| Pseudo R² | 0.408 |

### Panel C Determinants of the magnitude of the habit persistence
OLS regression with heteroskedasticity robust standard errors

\[ HH_{Persist} = \alpha_i + \beta_1 HH_{Incentive_i} + \beta_2 HH_{1i} + \epsilon_i \]

**Predictors** | **Coefficient** | **Std. error** | **t**  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HH_Incentive</td>
<td>0.517</td>
<td>0.270</td>
<td>1.91  **</td>
</tr>
<tr>
<td>HH1</td>
<td>-0.543</td>
<td>0.283</td>
<td>-1.92 *</td>
</tr>
<tr>
<td>Gender</td>
<td>0.103</td>
<td>0.071</td>
<td>1.45</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.002</td>
<td>0.002</td>
<td>1.07</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.391</td>
<td>0.272</td>
<td>1.43</td>
</tr>
</tbody>
</table>

| N  | 102 |
| R² | 0.326 |

Notes: This table reports three analyses of habit persistence. The variable HH_Persist is calculated as the difference between HH4 and HH1. Panel A: paired t-test compares means from a within-subjects test group; that is, this test considers only subjects for whom hand hygiene performance was observed in both Q1 and Q4; Panel B estimates the model \( \text{prob}(HH_{Persist} > 0) = \alpha + \beta_1 HH_{Pos\_Inc} + \beta_2 HH_{1} + \epsilon \) using logistic regression with robust standard errors; Panel C reports the OLS estimation of the model \( HH_{Persist} = \alpha_i + \beta_1 HH_{Incentive_i} + \beta_2 HH_{1i} + \epsilon_i \) (robust standard errors). In all panels: * = p<0.10; ** = p<0.05; *** = p<0.01
Table 5: Habit Effect – Subsample Analyses

Panel A  
**Estimation of the average habit effect**

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std dev</th>
<th>Std err</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH1</td>
<td>34</td>
<td>0.758</td>
<td>0.142</td>
<td>0.029</td>
</tr>
<tr>
<td>HH3</td>
<td>34</td>
<td>0.914</td>
<td>0.173</td>
<td>0.024</td>
</tr>
<tr>
<td>HH Habit</td>
<td>34</td>
<td>0.156</td>
<td>0.227</td>
<td>0.039</td>
</tr>
<tr>
<td>df</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>4.011</td>
<td></td>
<td>***</td>
</tr>
</tbody>
</table>

Panel B  
**Determinants of the probability of habit formation**

Logit regression with heteroskedasticity robust standard errors

\[ DV = \text{Prob}(HH\_Habit>0) \]

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Coefficient</th>
<th>Std. error</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH Pos Inc</td>
<td>5.572</td>
<td>1.941</td>
<td>2.87</td>
</tr>
<tr>
<td>HH1</td>
<td>-11.940</td>
<td>8.147</td>
<td>-1.47</td>
</tr>
<tr>
<td>Gender</td>
<td>1.727</td>
<td>1.570</td>
<td>1.10</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.097</td>
<td>0.114</td>
<td>0.86</td>
</tr>
<tr>
<td>Intercept</td>
<td>5.989</td>
<td>4.963</td>
<td>1.21</td>
</tr>
<tr>
<td>N</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.505</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C  
**Determinants of the magnitude of the habit formed**

OLS regression with heteroskedasticity robust standard errors

\[ DV = HH\_Habit \]

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Coefficient</th>
<th>Std. error</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH Incentive</td>
<td>0.783</td>
<td>0.268</td>
<td>2.92</td>
</tr>
<tr>
<td>HH1</td>
<td>-0.520</td>
<td>0.430</td>
<td>-1.21</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.019</td>
<td>0.003</td>
<td>0.19</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.001</td>
<td>0.003</td>
<td>0.19</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.447</td>
<td>0.377</td>
<td>1.18</td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.588</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports three analyses of habit formation. The variable \( HH\_Habit \) is calculated as the difference between \( HH_{i3} \) and \( HH_{i1} \). **Panel A**: the paired t-test compares means from a within-subjects test group; that is, this test considers only subjects for whom hand hygiene performance was observed in both Q1 and Q3; **Panel B** estimates the model \( \text{prob}(HH\_Habit>0) = \alpha + \beta_1 HH\_Pos\_Inc + \beta_2 HH1 + \epsilon \) using logistic regression with robust standard errors; **Panel C** reports the OLS estimation of the model \( HH\_Habit = \alpha_i + \beta_1 HH\_Incentive_i + \beta_2 HH1_i + \epsilon \) (robust standard errors). In all panels: * = p<0.10; ** = p<0.05; *** = p<0.01
Table 6: Habit Persistence – Subsample Analyses

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Estimation of the average persistence effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paired t-test analysis</td>
</tr>
<tr>
<td>Variable</td>
<td>N</td>
</tr>
<tr>
<td>HH1</td>
<td>31</td>
</tr>
<tr>
<td>HH4</td>
<td>31</td>
</tr>
<tr>
<td>HH Persist</td>
<td>31</td>
</tr>
<tr>
<td>df</td>
<td>30</td>
</tr>
<tr>
<td>t</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>Determinants of the probability of individual habit persistence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Logit regression with heteroskedasticity robust standard errors</td>
</tr>
<tr>
<td>DV = Prob(HH_Persist&gt;0)</td>
<td></td>
</tr>
<tr>
<td>Predictors</td>
<td>Coefficient</td>
</tr>
<tr>
<td>HH Pos Inc</td>
<td>4.083</td>
</tr>
<tr>
<td>HH1</td>
<td>-13.577</td>
</tr>
<tr>
<td>Gender</td>
<td>1.308</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.134</td>
</tr>
<tr>
<td>Intercept</td>
<td>7.182</td>
</tr>
<tr>
<td>N</td>
<td>34</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.356</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>Determinants of the magnitude of the habit persistence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS regression with heteroskedasticity robust standard errors</td>
</tr>
<tr>
<td>DV = HH_Persist</td>
<td></td>
</tr>
<tr>
<td>Predictors</td>
<td>Coefficient</td>
</tr>
<tr>
<td>HH Incentive</td>
<td>0.533</td>
</tr>
<tr>
<td>HH1</td>
<td>-0.827</td>
</tr>
<tr>
<td>Gender</td>
<td>0.002</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.001</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.694</td>
</tr>
<tr>
<td>N</td>
<td>28</td>
</tr>
<tr>
<td>R²</td>
<td>0.563</td>
</tr>
</tbody>
</table>

Notes: This table reports three analyses of habit persistence. The variable HH_Persist is calculated as the difference between HH_{i4} and HH_{i1}. Panel A: paired t-test compares means from a within-subjects test group; that is, this test considers only subjects for whom hand hygiene performance was observed in both Q1 and Q4; Panel B estimates the model \( \text{prob}(HH\_Persist > 0) = \alpha + \beta_1 HH\_Pos\_Inc + \beta_2 HH_1 + \epsilon \) using logistic regression with robust standard errors; Panel C reports the OLS estimation of the model \( HH\_Persist = \alpha + \beta_1 HH\_Incentive + \beta_2 HH_1 + \epsilon \) (robust standard errors). In all panels: * = p<0.10; ** = p<0.05; *** = p<0.01.