Pre-Analysis Plan: Handwashing and Habit Formation

Reshmaan Hussam, Giovanni Reggiani, and Natalia Rigol

March 21, 2016

Abstract

Diarrheal disease and acute respiratory infection result in two million child deaths per year worldwide (WHO 2013) and the physical and cognitive stunting of millions more. Caused by fecal and bacterial contamination, public health officials argue that handwashing with soap is the most cost-effective tool against such contamination. However, campaigns have met with consistent failure in engendering sustained behavioral change, suffering from two major weaknesses: (1) an inability to credibly measure handwashing; and (2) an incapacity to untangle the mechanisms behind the initial behavioral failure. In partnership with the MIT Media Lab, we develop a time-stamped sensor that unobtrusively measures household handsoap use. With this high frequency measurement tool, we examine three behavioral interventions targeted towards young children in an RCT setup in which control households receive a standard public health information campaign only: (1) a dispenser arm, in which the household is also provided with a foaming dispenser and soap; (2) a monitoring arm, in which the household is also informed that their behavior is being tracked by the dispenser and receives biweekly reports on performance; and (3) an incentive arm, in which the household is informed it is being tracked and receives discrete incentives for daily handwashing performance. The design is embedded within the classic habit loop of trigger, routine, and reward: the trigger is an agreed-upon mealtime against which performance is evaluated; the routine is handwashing with soap, and the reward is either clean hands only, a satisfied conscience (monitoring), or a monetary incentive. Within the two latter arms, some households are told that they will receive a future boost in their monitoring or incentive services; this allows us to examine the habit formation process against the standard model of rational addiction. External interventions (monitoring and incentives) will be removed after three months, but household dispenser use will continue to be tracked for one year to measure the precise patterns of habit formation. At the six month mark, households will be cross-randomized into an additional reminder arm, in which alarm clocks programmed for the daily meal time are attached to their dispensers, thereby enhancing the saliency of the trigger and examining limited attention. In conjunction with the high frequency dispenser use data, we will collect biweekly child health data on diarrhea and ARI incidence, as well as detailed anthropometric, blood, and stool data at endline.

"The acquisition of a new habit, or the leaving off of an old one, we must take care to launch ourselves with as strong and decided an initiative as possible. Accumulate all the possible circumstances which shall reinforce the right motives; put yourself assiduously in conditions that encourage the new way; make engagements incompatible with the old; take a public pledge, if the case allows; in short, envelop your resolution with every aid you know. This will give your new beginning such a momentum that the temptation to break down will not occur as soon as it otherwise might; and every day during which a breakdown is postponed adds to the chances of its not occurring at all.

- William James, *Habit*, 1914

1 Introduction

Bacterial and viral contamination, resulting in anemia, diarrheal disease, and acute respiratory infection, end the lives of nearly three million children per year and contribute to the severe stunting of millions more. Handwashing with soap is widely regarded as "the most cost effective vaccine" against such deaths, but campaigns to increase handwashing have failed to generate lasting changes in behavior or health. Only a handful find positive impacts of hygiene campaigns (ex. Luby et al. (2004), Haggerty et al. (1994), Han and Hliang (1989)), but these have limited capacity to untangle the mechanisms behind their results and therefore little power to generalize to alternative contexts.

The vaccum in knowledge on the barriers to preventative health practices extend well beyond handwashing: the use of latrines, water treatment systems, bednets, and vaccines are all high return preventative health behaviors that suffer low rates of takeup in the developing world despite their affordability (Dupas (2011)). Across the board, neither information provision nor materials and/or infrastructure provision generate sustained improvements in such practices (Clasen et al. (2014), Kremer and Zwane (2007), Banerjee et al. (2010)). In our setting in West Bengal, of the 94% of households in our sample region who own soap in the home, only 8% [claim to] wash their hands with soap before eating. Likewise, among the 23% who have access to sanitary latrines, 75% continue to defecate in the open. The vast majority know it is important to their health to use both.

This leaves behavioral mechanisms as the key determinant of preventative health and hygiene engagement. But current standards lack the precision in measurement (measures of behavior are biased, noisy, and scarce) and in design (given low statistical power in measurement, interventions are a fusion of many treatments) to tease apart the host of stories one can tell: from forgetting to engage (limited attention, top of mind), to absence of a social cost (group norms), to the presence of unobservable high mental or physical startup costs to the behavior. Without targeted interventions and high frequency, objective measures of behavior, it is difficult to identify the mechanisms behind increased takeup, a necessity to designing effective health policy. We employ a noninvasive sensor technology embedded in a liquid soap dispenser, developed in partnership with the MIT Media Lab, to precisely measure, monitor, and encourage handwashing behavior in households. Through an RCT evaluation in West Bengal, we pair this device with a series of treatments

targeting specific barriers to handwashing. The design allows us to separately test the roles of information and material provision, third party monitoring, high startup costs, and rational habit formation (described in detail below) on both short and long term takeup of handwashing and the resulting impact on child health.

Importantly, preventative health behaviors are habit-forming in nature . Handwashing with soap is no exception: in our conversations with mothers in our survey area, the most common reason for not handwashing is *"Obhash nai,"* or "I don't have the habit." We therefore embed each intervention into a classic habit loop with the evening mealtime as the trigger, handwashing as the routine, and various incentives as rewards. The measurement technology allows us to both reward households for daily units of behavior (essential to closing the loop) and unobtrusively track behavior after the withdrawal of extrinsic incentives to observe habit formation in the long run.

2 Habit Formation

2.1 Rational addiction theory

Our framework for habit formation builds upon the seminal work of Becker and Murphy (1988) on rational addiction. They and others in their spirit have focused on characterizing and testing the implications of rational addiction in the context of bad habits. We articulate the same and expand to the context of good habits, of which handwashing with soap before mealtime is my focus. Substantively, the shift from a bad habit to a good habit is equivalent to the shift from an activity in which the user experiences positive gains in the present but incurs costs in the future to an activity in which the user incurs costs in the present but experiences positive gains in the future.

Habitual behaviors share two defining properties: (1) reinforcement, or the development of a craving, such that the more one engages in the behavior, the more one wishes to engage in it; and (2) tolerance, such that the more one engages in the behavior, the easier it becomes (the lower the cost incurred in engaging).[1]

Reinforcement and tolerance are intrinsic properties of a habit, to be experienced by the user by nature of the activity. *Rational* habit formation (what Becker and Murphy (1988) term rational addiction) is the recognition of these properties: the user is aware of the habit forming nature of the activity, and is thus aware not only of the present cost and future return, but of the craving and tolerance developed through continual engagement, and chooses to engage conditional on this knowledge.

The key tradeoff that a rational individual engaging in a habit forming activity faces is between the drop in utility from consumption today and the increase in long-run utility from the accumulation of the stock in the addictive good.

[1]Note that for bad habits, reinforcement follows the same direction, but tolerance moves in the opposite direction: the more one engages in the bad behavior, the lower ones future utility given the amount of future consumption. In contrast, the more one engages in the good behavior, the less ones utility drops from each unit.

2.1.1 Empirical Evidence

Awareness of the habit forming nature of an activity implies that a user internalizes two components in the value of a good habit: the effect of current consumption on future return, and the additional multiplicative effect of current consumption on future consumption. Therefore, engagement in such behavior will depend both on past and future engagement. These intertemporal complementaries are the essence of the habit forming activity: a habit is reinforced through the fact that a larger stock of past consumption raises the marginal utility of current consumption; therefore, if an individual knows she wishes to engage in the behavior in the future, she will increase the marginal utility gained in the future by increasing her stock of the consumption today.

As mentioned earlier, existing empirical literature revolves around bad habits; most commonly smoking cigarettes, seconded by alcohol consumption. The standard empirical test involves regressing present consumption on past and future consumption and other demand shifters, instrumenting for the lag and lead of consumption using lag and lead of prices or tax rates.

$$c_t = \theta c_{t-1} + \beta\theta c_{t+1} + \delta p_t + \epsilon_t$$

Becker, Grossman, and Murphy (BGM 1994), the seminal empirical test of the rational addiction theory, interprets a positive coefficient θ as evidence of addictiveness, and a positive coefficient $\beta\theta$ as evidence of rational addiction. The ratio of the latter to the former yields the discount rate β.

The vast majority of the literature rests in favor of rational addiction. However, Auld and Grootendorst (2004) describe the implausible variation in discount rates, unstable demand, and low price elasticities implied by such literature. They go on to demonstrate that entirely non-addictive goods such as milk display the same positive and significant coefficient on future consumption as cigarettes under the standard empirical test, using this supposed rational addictiveness of milk as evidence for the abundance of false positives in the empirical literature. The authors demonstrate how high serial correlation in the commodity of interest, endogeneity in the price instruments, and overidentified IV estimators can all contribute to positive coefficient on future consumption incorrectly interpreted as evidence of rational addiction.

The field as such has advanced little in the last thirty years, as the datasets employed are nearly all aggregate time series data which suffer from the problems described above. Other problems that plague the literature are

1. A violation of the exclusion restriction on a tax instrument:

    - there may be underreporting of (eg.) smoking which is driven precisely by the antismoking sentiment that is motivating the tax increase;

- the types of regions in which cigarette prices rise are the types in which smoking is becoming less popular, resulting in confounding differential time trends

2. The implausibility of knowledge of future price changes: studies employ price changes so far in advance that it is unlikely consumers are aware of them when making their present consumption decisions;

3. The use of sales data as an instrument for consumption: sales data is confounded by the hoarding effect in which customers stock up on the commodity upon which they anticipate a tax increase (although this would yield an underestimate of the presence of rational addiction).

Gruber and Koszegi (2001) attempt to address the three problems outlined above by employing state specific time trends and using announced but as of yet unenforced tax rate increases (rather than far future sales data) as instruments for future consumption. However, they are still vulnerable to the endogeneity of prices to consumption yielding spurious results in favor of rational addiction. Furthermore, although the announced tax rate change is an improvement upon previous work, there is no way to verify whether consumers are aware of the future tax rate, and the likelihood is low given the year or more between the observed consumption decision and the tax enactment.

Gruber and Koszegi identify a further complication in the empirical test of rational addiction: given the nature of available consumption and price data, there is no way to empirically distinguish between the BGM model of rational addiction and an alternative model incorporating time inconsistency. Both generate the prediction that future prices matter for today's consumption; their difference lies only in the shape of discounting a consumer exhibits. The latter can only be deduced using detailed data on price changes over various periods of time, and no natural experiment to our knowledge exists to allow for such tests.

We find the literature well poised for the introduction of a field experiment. A field experiment can address nearly every difficulty that existing empirical work has faced: (1) prices are imposed exogeneously so there is no concern for endogeneity between prices and consumption; (2) future prices are explicitly announced so consumer knowledge of future price changes is confirmed; (3) differential time trends are no longer a concern given randomization; (4) endogenous misreporting is not a concern if we have an objective measurement device; serial correlation is no longer a concern as we are comparing outcomes across groups rather than over time.

Lastly, we can make considerable progress towards identifying the extent of time inconsistency in consumer behavior and how that impacts the empirical estimates of rational addiction, as we can examine heterogeneity in forward thinking behavior by baseline measures of discounting and myopia. We can further vary both prices and timing of price changes to trace out the shape of each consumer's discounting. Note that our current design is limited to a single price change, so the shape of discounting will not be deduced in the present experiment. We leave this to future rounds of the experiment.

Learning about a technology, being the accumulation of a stock which facilitates use, is empirically indistinguishable from building a 'craving' for a behavior: The rational addiction model and all other habit formation models in the literature cannot distinguish between the accumulation of habit stock from other forms of stock. such as learning. In fact, such learning can be subsumed in the intertemporal complementarities (craving) parameter of these models.

3 Dispenser and Soap Features

We use a standard wall-mounted dispenser as depicted in the left picture of Figure 1. The dispenser is opened with a specially designed key that was not supplied to the households during the course of the experiment. Soap is loaded in a one liter plastic container inside the dispenser. The sensor module is fitted between the container and the soap spout, as shown in the right picture of Figure 1. The module is encased in a waterproof mold, an essential feature for the oft wet environment of West Bengal and broadly for outdoor environments. Each push of the outer black button is registered in the sensor, which records the time of each push to the seconds unit. The sensor is a modular unit, easily removed and refitted into the dispenser; this design permits surveyors to replace the modules with fully charged versions on their biweekly visit with ease.

The dispenser was installed on the household premise near the dining space or water station as chosen by the household. Figure 2 depicts a typical setting for the dispenser: families usually eat on a mat on the verandah or just inside the front door. We chose a wall-mounted dispenser after repeated prototypes of sensor-embedded tabletop dispensers revealed that (1) the tabletop dispenser was at greater risk of being lost or stolen given its size and mobility; (2) creating a permanent 'handwashing station' through mounting in a prominent place made it easier for households to remember to wash, thereby enhancing the physical trigger in the habit loop[2] and (3) households regarded the mounted dispenser as a novelty to be cared for and were less impressed by a tabletop configuration, which they were familiar with. The dispenser was positioned at a height reachable by young children, as shown in Figure 3.

Identifying an appropriate soap likewise required extensive piloting. We experimented with several scents and consistencies which revealed that households preferred: (1) unscented or lightly scented soap that would not interfere with their eating experience; (2) soap of a thinner consistency; and (3) soap that lathered easily. We thus chose a foaming soap with a subtle scent approved by pilot households. We preserved some scent as the olfactory system is a powerful sensory source of both memory and pleasure and thus easily embedded into the habit loop.

[2]Pilot households motivated their valuation for the dispenser with the phrase "chokhe pore," literally meaning that it falls upon the eyes, making soap use easy to remember.

4 Experimental Design

4.1 Study sample

Our sample population is made up of 2943 peri-urban and rural households containing 3763 children below the age of ten across 105 villages in the Birbhum District of West Bengal, India.

4.2 Timeline and treatment groups

Figure 4 provides a map of all treatment arms and the time-contingent randomization process. All households received a basic information campaign regarding the importance of washing their hands with soap, especially prior to eating. They also received a calendar with the SHDS logo as a token for participation. They were notified that they would be visited biweekly for several months (time left unspecified) to collect information on child health and (for those who received dispensers) check soap supplies, which would be replenished as needed free of charge.

The randomization was conducted in five stages. First, the 105 villages were randomized into Monitoring Villages (MV) and Incentive Villages (IV). Households in MV were then randomized into two groups: (MV1) control and (MV2) dispenser. Households in IV were likewise randomized into two groups: (IV1) control and (IV2) dispenser + monitoring + incentive. These first two stages of randomization were determined after baseline, prior to the roll out of the dispensers.[3] During roll-out, MV2 households were notified about an upcoming lottery in which selected households would receive, along with the dispenser, a monitoring service (described in detail below). Willingness to pay for this monitoring service was elicited, with households informed that a higher willingness to pay would result in a higher chance of winning the lottery for the service (to ensure incentive compatibility). IV2 households were notified about an upcoming lottery in which selected households would receive a larger incentive for thirty days (described in detail below).

After roll-out was completed, the third stage of randomization was run, determining which households in MV would remain dispenser only (MV2) and which would receive the monitoring service (MV2a), and which households in IV would keep the standard incentive (IV2) and which would experience the thirty-day incentive bump (IV2a). Both the monitoring service (in the case of MV2a) and the incentive bump (in the case of IV2a) were scheduled to begin approximately 1.5 months after the day the household received the dispenser; since roll-out was staggered, the specific date varied by household and was clearly circled on the SHDS calendar by the surveyor. Importantly, households were notified of their future service or incentive immediately after the completion of roll-out. The staggered roll-out implied that some households were

[3]Households were first randomized at the village level in order to limit the scope for inter-household tension: surveyors expressed concern that control households would be angered if they had some neighboring households who received a dispenser and others who received a dispenser and incentives. It would be easier to justify the interventions through the limited resources lottery framework if all dispenser-receiving households within a village received a consistent package of goods (i.e. the dispenser either always came paired with incentives or never did).

told about their future reward two weeks after receiving the dispenser while others were told two days after receiving the dispenser. We embed this variation in timing of the announcement in an effort to disentangle the effects of learning and experimentation from those of habit formation.[4] [5]

The fourth randomized allocation occurred on the date of the monitoring or incentive change (1.5 months after dispenser distribution). At this point, half of MV2 households (MV2b) and half of IV2 households (IV2b) were randomly selected to also receive the monitoring service or incentive boost, respectively. These households, having been surprised with this additional service or incentive which was then effective immediately, could not have altered their behavior in anticipation of the change.

Finally, six months after roll-out, the fifth stage of randomization will be implemented, with households in all arms cross-randomized into a Reminders (R) intervention. We choose to integrate this final intervention three months after the withdrawal of all extrinsic incentives for three reasons: (1) we fear that the effectiveness of reminders might soak up power to detect the rational addiction, incentive, and monitoring effects, and wish to preserve power for tracking longer run effects; (2) reminders should have a smaller marginal effect in settings where habits are already deeply embedded, implying heterogeneity in response across treatment arms depending on the effectiveness of the earlier treatments; and (3) logistical ease.

The components of each arm are described in detail below.

(MV1 and IV1) Control: Households were given a simple informational lecture on the importance of washing hands with soap, with stress placed on the responsibility of the mother to do so and encourage her household to do so for the sake of her children's health. This speech was repeated twice: once at baseline and once again at roll-out.

(MV2) Dispenser: Households were given a liquid handsoap dispenser. They were informed that this was a high quality soap dispenser that would make it easier to wash hands. They were informed that there was a switch inside the dispenser that, if turned on, would allow their behavior to be tracked. A partner organization wanted to offer a monitoring service to the households (if they desired it) in which household handwashing behavior would be tracked and reported biweekly. If households

[4]This variation also speaks to the malleability of habit formation: households who have not yet determined a routine in dispenser use may be more able to respond to a future change in the value of the behavior than those who have already completed their experimentation and learning regarding the new dispenser.

[5]Because all households were notified about the lottery for the future price or service change, all households, regardless of assignment in the third stage, should have had the same expected valuation for the handwashing behavior. Upon being randomized into receiving the future price change or monitoring service, treatment households face an increased future return to the behavior but, in a world without rational addiction, identical current returns to the behavior. In a world of experimentation with risky technology, an increase in the return to future use of a technology should only affect my current experimentation with the technology if I believe that current experimentation affects my ability to use the technology in the future and thereby reap returns to future use. This is distinct from a world where subsidizing experimentation with a risky technology increases use: the subsidy is an adjustment to *current* (or constant) returns to the technology, *not* a time-varying adjustment to the return. In this latter case, returns to experimentation may be realized in the future, but this is different from returns to future behavior being higher than returns to current behavior. It is in this way that the learning and habit formation stories can be distinguished, and our experimental design identifies only the latter mechanism.

were interested, they were told to report the monthly fee they would be willing to pay for this service. Because resources were limited, the service would be administered in lottery fashion; however, the more the household valued the service, the more likely they were to receive it. If they did not get selected for the service, their switch would not be turned on and their behavior would not be monitored.

(MV2a) Announced monitoring: Households were informed that they were selected in the lottery for the monitoring service, and would in fact receive the service free of charge: the internal switch would be turned on, and the device would record the time and frequency with which they washed their hands with soap. The surveyor would be carefully observing this data every two weeks to track their behavior and would provide the household with a biweekly report of their daily behavior. This arm can therefore be regarded as a combination of a third party-monitoring and self-monitoring intervention. The service would begin 1.5 months after dispenser distribution on a date circled clearly by the surveyor on the calendar.

(MV2b) Unannounced monitoring: 1.5 months after dispenser distribution, these households were surprised with an identical monitoring service to those in MV2a, effective immediately.

(IV2) Incentives: Households were informed that their dispenser tracked the frequency and time of use and that their behavior would be tracked. They were then given a small coin purse and told that they would receive one ticket for every day in which the device was active prior to their stated dinnertime, which they should accumulate in their purse. These tickets could be exchanged for various prizes as detailed on an incentive catalog (see Appendix B).[6] These incentive payments would last for three months. They were also told that SHDS anticipated receiving additional funding from the government for the project in one month, at which point SHDS hoped to increase the reward for handwashing to three-fold for thirty days. Because the future funds were limited, households would be entered into a random lottery to see who would receive the future increase in reward. They would be notified of the results of this lottery within two weeks. It was important that we provide all households with an incentive from the beginning (prior to the increase in incentives) in order to establish an understanding of the nature of the incentives and trust between the surveyors and the households that the future increase would indeed by fulfilled.

(IV2a) Announced incentives with boost: Households were informed that they had been selected in the lottery for the incentive boost, and could anticipate receiving three tickets for every day in

[6]The ideal incentive requires three conditions: (1) the incentive must be divisible; (2) the daily amount offered must be sufficiently high to induce behavioral change on a daily basis, which is key to habit formation; and (3) the marginal value of the units accumulated as the process of habit formation continues must also remain sufficiently high to continue inducing behavioral change. Tickets exchanged for goods provides satisfies all three conditions while also offering flexibility in the types of goods that a household may find appealing. Prizes were selected to focus on child health and schooling and adult household goods. We chose a range of goods of various values to examine heterogeneity in purchase of type and price of goods.

which the device was active prior to their stated dinnertime for thirty days. The boost would begin 1.5 months after dispenser distribution on a date circled clearly by the surveyor on the calendar. For the remainder of the three months, they would receive the standard one ticket per day of activity.

(IV2b) Unannounced incentives with boost: 1.5 months after dispenser distribution, these households were surprised with an identical incentive boost to those in IV2a, effective immediately.

(R) Reminders: Households [will be] offered a miniature alarm clock which is, upon household consent, attached to the dispenser and set to beep at the household's reported dinner time daily. The alarm can only be turned off by the household upon beeping; it cannot be reprogrammed, as the method is complicated and known only to the surveyors.[7]

4.3 Identification of effects

The effect of receiving the dispenser alone is captured in the comparison of households in MV2 to MV1.

A higher take-up of handwashing behavior in MV2a relative to MV2 and IV2a relative to IV2 (*before* the price change) demonstrates the presence of rationally addictive behavior: households who increase take-up today due to an increase in the future value (or decrease in cost) of handwashing must recognize that higher take-up today will generate a greater accumulation of the positive internalities and craving stock over time, making it easier to reap the benefits of the future rewards to the behavior.

By maintaing the same incentive stream across both groups, a comparison of MV2a to MV2b and IV2a to IV2b (*after* the price change) allows us to identify the effect of forward looking, rationally addictive behavior on habit formation (conditional on finding evidence of rational addiction prior to the price change). In other words, a long term comparison of take-up between the 2a and 2b groups demonstrates whether forward-looking behavior in fact facilitated the formation of the handwashing habit.

Note that a zero difference in take-up between households in MV2 versus MV1 and IV2 versus IV1 prior to the price change could be due to two reasons: (1) households are not rationally addictive with regards to the handwashing habit, or (2) the future price change was not sufficiently compelling to induce behavioral change, even for forward-looking individuals. The second possibility is eliminated if households do indeed respond to the price change when it is enacted. This pure price change effect can be identified by comparing households in MV2 to those in MV2b and households in IV2 to those in IV2b after the price change, as the only difference between these sets of households is price change itself, with no behavioral response to anticipation. This comparison gives us the pure effect of the incentive boost or the monitoring service on handwashing behavior.

[7] We choose an alarm clock attached to the dispenser rather than test messages, the more common reminder intervention utilized in the literature, because it both (a) requires a zero cost of effort to receive the reminder, and (b) requires the household to expend effort to reach the handsoap dispenser (a large part of the startup cost to handwashing) in order to prevent the reminder from being a source of disutility from continual beeping.

This design precludes perfectly capturing the effect of incentives on top of monitoring. Monitoring was introduced (MV2a and MV2b) 45 days after rollout, while incentives (IV2) were introduced immediately after rollout. We were deliberate in this choice: monitoring was delayed in order to increase our sample size on the rational addiction test, with the tradeoff of a loss in the perfect comparison between monitoring only and monitoring+incentive households. Given the habitual nature of handwashing (or more broadly, dispenser use), the delay in introducing monitoring may have reduced the malleability of the behavior and therefore the potential effect of the treatment relative to that of incentives [8] [9].

That said, a comparison of MV2b to IV2 (or IV2, IV2a, and IV2b pooled) offers the first estimate in the literature of the marginal value of positive material rewards relative on top of (potentially negative) socially motivated feedback on daily behavior.

After the third month of incentives or monitoring, all extrinsic treatments will be removed. Three months later, households will be cross-randomized into the Reminders arm. Households will be told that they will continue to be visited biweekly for the next two months. Those in arms MV2 and IV2 will be reminded that data will continue to be collected. This portion of the study is ongoing in the field, and will shed light on the persistence of the behavioral change absent incentives, with and without reminders. Given the nature of the treatment, we cannot examine the persistence of the behavioral change absent monitoring, a topic of future study.

5 Beliefs and Behavior

Note that the rational habit formation model note only assumes that individuals are aware of the habit forming nature of an activity, but that they can *act* upon their current desires and beliefs regarding the habitual behavior. We make a distinction between beliefs and behavior because individuals may believe they will respond to a future price change by altering their current behavior (being forward looking and conscious of the nature of habit), but not react as such in reality (being naive about their present-biasedness or suffering from limited attention at the moment of action). Aside from a heterogeneity analysis using baseline measures of present biasedness, this study does not engage the issue of time inconsistency, as the predictions of the rational addiction model alone, with regards to the extensive margin of response to a future price change, are not empirically distinguishable from those of the rational addiction model incorporating

[8] Which is precisely why we did not delay the introduction of incentives to parallel the introduction of monitoring: this would mean a 75 day delay in the introduction of the future price change, reducing the likelihood of finding a rational addiction effect.

[9] Marshall, 1980: "There is, however, an implicit condition...which should be made clear. It is that we do not suppose time to be allowed for any alteration in the character or tastes of the man himself. It is, therefore, no exception to the law that the more good music a man hears, the stronger is his taste for it likely to become; that avarice and ambition are often insatiable; or that the virtue of cleanliness and the vice of drunkenness alike grow on what they feed upon. For in such cases our observations range over some period of time; and the man is not the same at the beginning as at the end of it."

time inconsistency (Gruber and Koszegi, 2001).[10]

We do, however, structure our experimental design to examine the role of limited attention in inhibiting the translation of belief into action. By cross randomizing arms (1) through (4) with reminders, we can observe whether a larger proportion of those who believe they will increase their handwashing behavior in fact do so when given reminders. We can further examine whether this effect is differentially higher for rational addicts - in other words, whether the effect is differentially higher for those who in effect have 'more to remember.'[11]

Note that the reminders also allow us to examine a key component of the habit formation process: the trigger. *A priori*, the effect of the reminder on long term habit formation (handwashing behavior after the removal of all interventions) is ambiguous: on one hand, the alarm clock can serve to enhance the saliency of the existing trigger (dinner time) by bringing attention to both the time and the place of handwashing, thereby making it easier to associate the desired behavior with the trigger. However, if households substitute the alarm sound with the time of day as the trigger, the removal of the alarm may disrupt the habit entirely.

6 Outcomes of Interest

The primary outcomes of interest encompass both beliefs and behavioral changes of households. We capture beliefs through (1) forecasts of own daily handwashing behavior and (2) willingness to pay for the liquid handsoap dispenser. We capture behavioral change through (3) recorded total daily handwashing rates and (4) recorded dinner time-specific daily handwashing rates. Beliefs and behavioral measures could only be collected for those households with dispensers, so we do not have data from the pure control households on these metrics. Finally, we collect child health data in the form of (5) self-reported biweekly incidence of child diarrhea and respiratory illness and (6) child blood and stool reports. Each is defined below.

A. Household Beliefs

Forecasts of own daily handwashing behavior are collected biweekly to elicit how each intervention impacted household beliefs about future behavior. Respondents are asked to forecast how many days in the coming week they anticipate themselves and their children washing their hands with soap prior to dinner time.[12]

[10] Multiple price changes over varying time periods would be required to trace out the discount curve of consumers, the intensive margin of response which can distinguish the behavior of time consistent and inconsistent rational addicts from one another. Although this is outside the scope of the current project, our high frequency data paired with the ability to vary prices and intervals between price changes exogenously offers an ideal setting to examine this question in future work.

[11] The rational addict considers not only the effect of the price/incentive today on behavior, but also that of the future; literature on reminders notes that activities that are further in the future are less salient, placing them further from the 'top of mind' and thus more likely to be forgotten absent reminders.

[12] Respondents may overstate their forecasts due to the experimenter effect, but this effect should be constant across the

Willingness to pay for the liquid handsoap device will be elicited around the eight month mark and again at endline amongst those households who receive a dispenser (all but pure control) using a version of the BDM mechanism suited for our context. We will present households with a series of choices between keeping the dispenser or trading it in for a prize of sequentially increasing monetary value. Each answer will be put into a bowl and the respondent will be told that she will be asked to randomly draw from the bowl after all questions are answered; whichever question-answer pair is drawn will be fulfilled.

B. Household Behavior

Handsoap dispenser data is collected every two weeks during surveyor visits. Although it is not possible to identify the identity of the user at any given press, we proxy for separate users by collapsing presses that happen two or fewer seconds apart into a single press. In other words, if the device is used in seconds 34, 35, 37,45, and 46, the first three presses are considered a single use by one household member, and the later two presses as a single use by another member. Though not exact, observations from pilots made clear that users press several times in quick succession and rarely return for more soap during a single handwashing event, since the water source (usually a bucket right outside the front porch) is not within reach of the dispenser, unlike the familiar setting of sink, soap, and running water common to more developed contexts.

Daily handwashing rates are calculated as the sum of all 'individual' uses over the course of each twenty-four hour period.

Mealtime-specific handwashing rates are calculated as the total number of 'individual' uses in the interval of 90 minutes before and after the household's reported start of the evening meal time. If a family reported eating dinner every day at 8:00 PM, for example, this outcome would be the sum of all individual presses observed between 7:00 PM and 8:30 PM.

Binary use at mealtime is derived from the above and is a binary variable which equals one if at least one 'individual' use was observed in the dinner time interval.

Evening handwashing rates are calculated as the total number of 'individual' uses any time 4pm or later. Use of the dispenser during this time of day is almost surely tied to use before eating, since all other potential uses (laundry, shampoo, or hand washing after defection or cleaning the house) have all been completed by this time. Since we rolled out in the winter season, children often ate earlier than their parents due to the cold, so the reported mealtime is not a consistent measure of child eating habits. This broader measure allows us to capture all handwashing-before-eating events in the evening.

monitoring and incentive arms (where households are aware that their true behavior can be verified by the surveyor). Those in the dispenser control arm, who were unaware of observation, may have been more likely to overstate in order to please the surveyor.

Weekly dinnertime use is calculated at the household level as the fraction of days per seven days that the household's dispenser was active during the evening mealtime.

Consistent-use households are defined as households who use the dispenser during their evening mealtime at least six out of every seven days. While our primary definition of habit formation revolves around the long-term persistence of handwashing behavior, this measure gives us a proxy for habit formation in the short run.

C. Child Health

Incidence of child diarrhea and respiratory illness is collected every two weeks by surveyors, and consist of self reports in which mothers are asked how many days each child had experienced diarrhea in the past two weeks, and likewise for respiratory illness (cough, cold). The latter was initially collected through a question asked to the mother: "Has your child had any coughs, runny nose, or fever in the last two weeks?" and recorded as ARI if the child had a fever paired with either of the first two symptoms. The method was adjusted at the four month mark to elicit any symptoms associated with respiratory infection, and surveyors were told to see the child in person rather than ask the mother only. This change was instituted as field visits made clear that mothers often did not know or did not report what symptoms their children suffered from, and children who had runny noses but could neither self report a fever or have such reported by their mother were still likely suffering from a respiratory infection.

Anthropometric outcomes were collected at baseline, and will be repeated at the four month and eight month mark. These include child weight, height, and mid-arm circumference.

Child blood and stool analysis will be conducted for a limited subset of children whose parents consent to the process fourteen months after baseline.

D. Time-Contingent Outcomes

Because various interventions were phased in at various times, below we define the time period for each effect of interest.

Baseline behavior is defined first through the baseline survey, which was conducted four months prior to rollout. To more precisely examine handwashing behavior at the beginning of the experiment, we also define baseline behavior as dispenser use in the first week after rollout. This comes with the caveat that IV households were told from the first day of rollout that they would be monitored and incentivized, while MV households were told neither on the first day. Thus this "baseline" check is used to establish balance across treatment arms within villages, not across MV and IV villages.

Rational addiction period is defined broadly as the time between rollout and the service/price change, but we intend to look specifically at the three week period prior to the price change. This is because (1) we showed a video to all dispenser-receiving households at this time in order to increase and standardize comprehension regarding which treatment group each household was in; and (2) any rational habit formation effect should increase as the date of the anticipated change approaches, so we focus on the weeks nearest the price change.

Medium run post-change period is defined as the three months following the price/service change.

Long run post-change period is defined as the period 8-18 months after the price/service change. This is equivalent to 6-16 months after the withdrawal of all extrinsic interventions (incentives and monitoring).

7 Primary Analysis

Our preferred specification for primary outcomes is as follows:

$$Y_{hvt} = \alpha_{hvt} + \beta Treatment_{hvt} + \gamma_t + \theta_v + \phi_s + \epsilon_{hvt}$$

in which Y_{hvt} represents the household behavior outcomes specified in Section 6, *Treatment* is the assigned treatment for each subset of comparisons described in Section 4.3, γ_t is day fixed effects, θ_v is village fixed effects, and ϕ_s is surveyor fixed effects. The latter two are included in all but those regressions comparing treatments across Monitoring and Incentive Villages (we omit both village and surveyor fixed effects in these regressions since many surveyors have only MV households or only IV households). Standard errors are clustered at the household level.

8 Secondary Analysis

8.1 Heterogeneity

Baseline evidence of experimentation with dispenser: Are households who experiment most with the dispenser in the first week of use (defined by 90th percentile of use in first week) more or less likely to respond to the future changes in the value of handwashing?

Female education, discount factors, bargaining power and aspirations: Do mothers with higher education, lower discount factors, higher aspiration scores (all likely correlated with awareness of future returns to investments) and higher bargaining power (correlated with the ability to act on

such knowledge) as determined by our baseline modules demonstrate greater responsiveness to changes in the future value of handwashing (i.e. more evidence of rational habit formation)?

Willingness to pay for monitoring service: MV2 households were asked their willingness to pay for the monitoring service at rollout. Do households with a higher WTP for the service respond more upon receiving the service?

Correlation of handwashing episodes over days can be used as a proxy for the formation of a handwashing habit. This is similar to our primary outcome of defining consistent-use households as those who use the dispenser in the evenings at least six days per week.

Quantile regressions on dispenser use can allow us to explore differences in response over the distribution of use: certain interventions may impact only a subset of the population that is behaviorally prepared to react to the treatments.

Distributional comparisons across treatment groups more broadly (eg. first and second order stochastic dominance in handwashing performance over time) can likewise describe how the populations in each treatment group evolve: do all incentivized households increase their handwashing behavior, or do we see a bifurcation in behavior between those households that fully establish a habit under incentives and those that decide it is not worth their effort and drop to low handwashing rates?

Spillover effects between treatment and control households can be captured using GPS data on household location. Spillovers will be investigated both along the handwashing dimension (control children in neighboring households who see treatment children wash may learn to do the same) and the health dimension (ARI, being highly communicable, may be reduced among control children who play with nearby treatment children).

9 Bibliography

Banerjee, A., Duflo, E., Glennerster, R., Kothari, D. (2010). Improving immunization coverage in rural India: clustered randomized controlled evaluation of immunization campaigns with and without incentives. BMJ. 340: c2220.

Becker, G.S., Grossman, M., Murphy, K.M. (1991). Rational Addiction and the Effect of Price on Consumption. The American Economic Review. 81(2):237-241.

Becker, G.S., Murphy, K.M. (1988). A Theory of Rational Addiction. The Journal of Political Economy. 96(4): 675-700.

Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., Ensink, J., Freeman, M., Jenkins, M., Odagiri, M., Ray, S., Sinha, A., Suar, M., Schmidt, W. (2014). Effectiveness of a rural sanitation programme on diarrhea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. The Lancet Global Health. 2(11): e645-e653.

Dupas, P. (2011). Health Behavior in Developing Countries. Annual Review of Economics, Vol. 3.

Gruber, J. Koszegi, B. (2001). Is Addiction ?Rational?? Theory and Evidence. The Quarterly Journal of Economics 116(4): 1261-1303.

Haggerty, P. A., Muladi, K., Kirkwood, B. R., Ashworth, A. Manunebo, M. (1994). Community-based hygiene education to reduce diarrhoeal disease in rural Zaire: impact of the intervention on diarrhoeal morbidity. Int J Epidemiol 23, 1050?1059.

Han, A. Hliang, T. (1989). Prevention of diarrhea and dysentery by hand washing. Trans R Soc Trop Med Hyg. 83(1):128-31.

Kremer, M. Zwane, A. (2007). Cost-Effective Prevention of Diarrheal Diseases: A Critical Review. Center for Global Development Working Paper Number 117.

Luby, S. P., Agboatwalla, M., Painter, J., Altaf, A., Billhimer, W. L. Hoekstra, R. M. (2004). Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan. JAMA 291, 2547?2554.

Figure 1: Soap dispenser

Figure 2: Typical dispenser location

Figure 3: Children using dispenser

Figure 4: Randomization Map