Discussion Paper Series – CRC TR 224

Discussion Paper No. 285
Project B 07

The Habit-Forming Effects of Feedback: Evidence From a Large-Scale Field Experiment

David P. Byrne [1]
Lorenz Goette [2]
Leslie A. Martin [3]
Lucy Delahey [4]
Alana Jones [4]
Amy Miles [4]
Samuel Schob [5]
Thorsten Staake [5]
Verena Tiefenbeck [6]

April 2021

[1] Department of Economics, University of Melbourne, byrned@unimelb.edu.au
[2] Institute for Applied Micro Economics, University of Bonn, lorenz.goette@uni-bonn.de
[3] Department of Economics, University of Melbourne, leslie.martin@unimelb.edu.au
[4] South East Water Corporation, Melbourne, Australia
[5] Department of Information Systems and Applied Computer Sciences, University of Bamberg
[6] Department of Management, Technology and Economics, ETH Zurich

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

The habit-forming effects of feedback: evidence from a large-scale field experiment*

David P. Byrne[a] Lorenz Goette[b] Leslie A. Martin[c]
Lucy Delahey[d] Alana Jones[d] Amy Miles[d]
Samuel Schob[e] Thorsten Staake[e] Verena Tiefenbeck[f]

March 30, 2021

Abstract

We provide a unique test of competing models of persistence in behavior. We propose a new attention-based behavioral mechanism for habit formation and contrast its predictions with the Stigler and Becker (1977) consumption-based mechanism. We test both mechanisms using a large-scale field experiment in shower water consumption. Our experiment varies cycles of household-level real-time feedback that temporarily draws attention to individuals' water consumption. Combining this design with real-time consumption data, we test the mechanism for persistence in behavior that our experiment generates. Our results strongly support a dynamic attention-based model of habit over the workhorse habit stock model used in economics.

JEL Classification: C93, D12, D83, Q25
Keywords: Habit formation; Attention; Realtime feedback; Water usage; Randomized Control Trial

---

*Disclosures: LD is the household support program manager at South East Water. AJ is the household solutions program manager at South East Water. AM is project manager in South East Water's strategic and household program. Amphiro AG, Switzerland, developed the feedback device used in this study. SS is chief technology officer of Amphiro. TS is chairman of the board of Amphiro. VT is a scientific advisor to Amphiro. **Author Contributions:** DB, LG, and LM developed the final experimental design, implemented the study, analyzed the data, and wrote the paper. LD, AJ, and AM managed recruitment and data collection at South East Water. SS implemented the design in the feedback device and developed a cloud service for data handling. TS and VT with LG proposed an initial experimental design. All authors contributed to revisions of the paper. **Acknowledgements:** We have received helpful comments and suggestions from seminar participants at the University of Melbourne, UTS, NUS, 2019 Banff Empirical Microeconomics Workshop, and the 2020 NBER Summer Institute (EEE). We also thank the University of Melbourne FBE Faculty Research Grant scheme and the Centre for Market Design for funding. Funding by the German Research Foundation (DFG) through CRC TR 224 (Project B07) is gratefully acknowledged.

All errors are our own.

[a]Department of Economics, University of Melbourne, byrned@unimelb.edu.au.
[b]Institute for Applied Micro Economics, University of Bonn, lorenz.goette@uni-bonn.de
[c]Department of Economics, University of Melbourne, leslie.martin@unimelb.edu.au
[d]South East Water Corporation, Melbourne, Australia
[e]Department of Information Systems and Applied Computer Sciences, University of Bamberg
[f]Department of Management, Technology and Economics, ETH Zurich

1 Introduction

Many behaviors display remarkable persistence: how often we go to the gym, when we wash hands, how much energy and water we use, all appear to be highly correlated over time. In several studies, it has also been shown that the impact of interventions, ranging from financial incentives to providing feedback on one's behavior, on these behaviors lingers beyond their duration (Charness and Gneezy, 2009; Jessoe and Rapson, 2014; Allcott and Rogers, 2014; Royer et al., 2015; Acland and Levy, 2015; Loewenstein et al., 2016; Hussam et al., 2017; Yang and Lim, 2017; Ito et al., 2018).

Despite much recognition of the importance of what might be described as habits, little is known about the underlying channels through which they form. Stigler and Becker (1977) developed the workhorse model of persistence in economics. In their model, past consumption decisions help grow a habit stock that creates complementarity between past and current consumption by increasing the marginal utility of current consumption, hence generating persistence in consumption over time. It is hard to overstate the importance of their *consumption habit model*, given its central role in many areas of economics such as macroeconomics, industrial organization, and public economics.

However, other mechanisms of persistence might be relevant as well. Research in neuropsychology has shown, for example, that attention also exhibits persistence (Anderson, 2016; Jiang and Sisk, 2019).[1] This literature shows that a temporary intervention, e.g., using incentives to detect a pattern shape, affects attention to that pattern even weeks after the incentive has been removed.

In economics, it is widely recognized that limited attention affects behavior (see, e.g. Gabaix, 2019). As many products have attributes that are difficult to perceive (e.g, calories in food, energy use of daily behaviors, hidden fees in contracts), drawing attention to them can have strong effects on behavior (Chetty et al., 2009; Berkouwer and Dean, 2019; Blake et al., 2018; Tiefenbeck et al., 2018). The neuropsychology literature suggests that temporarily highlighting these attributes persistently affects how much attention individuals pay to them, and thus opens up a new mechanism for persistence in behavior. We call this the *attention habit model*.

Empirically testing microfoundations for habit persistence has proven challenging. State dependence, correlated treatment components, and individual-level unobserved heterogeneity all confound tests of habit formation (Auld and Grootendorst, 2004). In this paper, we use a large-scale field experiment in shower water use to test mechanisms for habit persistence. Our intervention provides real-time feedback on the amount of water used through a smart shower meter. Across seven experimental conditions, we vary the intensity and frequency of feedback that households reveal. This allows us to examine the persistent effects of the intervention on behavior. Combining our novel research design with real-time data on water usage

[1] In psychology, habits are typically characterized by three features 1) frequency: habits are formed and sustained by frequently repeated patterns of activity, 2) automaticity: habits, like motor skills, no longer require active thought, and 3) triggers: habits are activated by contextual cues. Thus habits are a "shortcut" to decision making (Wood and Neal, 2007; Wood and Runger, 2016). The intuition behind this research, recently formalized in (Camerer et al., 2020), is that habit allow individuals to economize to decision making costs. While somewhat related to our model, they imply discrete jumps in behavior, whereas Stigler and Becker (1977) and our model predict continuous changes in behavior. We test for jumps predicted by this class of models in section 4.3, but find little evidence thereof in our particular setting.

allows us to test the implications of the Stigler and Becker (1977) consumption habit model, and contrast its predictions with a model in which attention is habit forming.

The empirical context – showering – is highly relevant for the study of habit for both research- and policy-related reasons. Water usage is difficult to gauge while showering, thus raising the possibility that limited attention plays a role. Further, the context avoids an important confound in identifying habit-induced behavioral change, namely households making technological investments in response to an intervention that, like habit formation, yields persistent changes in behavior over time. With our real-time consumption data, we can rule out the existence of such investments and credibly identify habit persistence due to behavioral change alone.

In terms of policy, water utilities worldwide are currently exploring digitally-enabled behavioral interventions for promoting water conservation. Understanding how habits form in the presence of real-time feedback is fundamental to informing investment decisions into these technologies and strategies for water conservation, and to understand how persistence affects optimal targeting of feedback more generally. This relevance is accentuated for places already seeing widespread drought due to climate change.

Overview

We develop our study over four sections, starting in Section 2 where we develop a model of habit persistence which nests consumption habit and attention habit. These mechanisms for habit persistence generate different predictions for the dynamics of consumption when feedback that makes water consumption salient is introduced and subsequently removed. We show that following a change in salience, the habit stock mechanism predicts an initial jump in consumption followed by a gradual reinforcing convergence to the new steady-state. When feedback is removed, the consumption habit model predicts a jump in the opposite direction, and a *symmetric* gradual convergence back to the original steady-state.

In contrast, our attention habit mechanism predicts that feedback has an immediate and stable effect on the consumption. It also raises an individual's attention stock. When feedback is removed, our model also predicts a gradual convergence back to the original steady-state as an individuals' attention wanes and attention reverts to its original level. Importantly, these attention-driven consumption dynamics imply *asymmetric* responses to the introduction of feedback and its subsequent removal.

We designed a large-scale field experiment to test these predictions empirically. The experiment, which we describe in Section 3, leverages a smart shower meter – the Amphiro B1 pictured in Figure 2 – that provides real-time feedback on water use to an individual. This technology, combined with our experimental design, delivers a uniquely well-suited setting for testing behavioral mechanisms for habit. Prior research has shown that real-time feedback from this shower meter yields large conservation effects in water usage (Tiefenbeck et al., 2018). This result, combined with the fact that our experimental design allows us to observe real-time individual-level consumption responses when feedback is introduced and removed, implies that we can observe at high-frequency the build up and subsequent decay of large-magnitude consumption responses to feedback. This high-frequency feedback provides the statistical power to test the habit stock and attention stock mechanisms, most notably to test the degree of symmetry in how habits build-up and decay.

We teamed with Amphiro, the manufacturer of the devices, to re-engineer the smart meters to cycle feedback through on and off modes, thereby allowing us to randomly expose households to different length spells of consecutive showers with and without feedback. In doing so, we created independent exogenous variation in the build-up of both habit and attention stocks. This design also allows us to further isolate and test for a one-time learning effect from the effect of salience, experimentally identify the rate at which each stock builds, and generate habit persistence in consumption when habit-forming stimulus is removed.

Section 4 presents a reduced-form analysis of feedback effects and persistence induced by our experiment. We first show that we are able to replicate previous real-time feedback interventions involving smart shower meters.[2] In our sample, individuals immediately reduce water use by approximately 15 percent in response to real-time feedback. Thus, the intervention powerfully shifts behavior, and provides us with the necessary impulse to be able to detect persistence effects due to the consumption habit and attention habit mechanisms.

Exploiting a within-subject experimental design, we use difference-in-difference regressions to characterize the degree of symmetry in the build-up and decay of treatment effects when feedback is introduced and subsequently removed. Through both visual evidence and a series of econometric tests, we establish that feedback induces highly asymmetric build-up and decay of treatment effects that in all dimensions directly align with an attention habit based mechanism for habit formation.

Moving beyond reduced-form treatment effects, in Section 5 we estimate a non-linear dynamic model of consumption and attention habit formation and decay. Our empirical framework borrows from Malmendier and Nagel (2011), and allows for arbitrary non-linear time discounting of feedback in the way a household's attention stock evolves with and without feedback, and its impact on consumption. The estimates from this habit model echo those of our time-varying treatment effects, namely that feedback effects yield asymmetric consumption responses over time that contradict a Stigler and Becker (1977) consumption habit model and align with attentional habit formation. Moreover, we find quantitatively feedback gives rise to a high degree of persistence. For example, we find that six weeks of real-time feedback exhibits a persistent impact on consumption even ten weeks after feedback is removed. Attentional habit formation has large persistence effects quantitatively.

Moreover, through our empirical analyses Sections 4 and 5, we further rule out other explanations for the consumption dynamics that our experiment induces. These include habit formation arising from automatic/default decision rules (such as, e.g., Wood and Runger, 2016; Camerer et al., 2020), as well as the role of household learning in response to being provided salient feedback over consumption.

Our results are important for water conservation for several reasons. First, we document that there is indeed behavioral persistence independent of any technology adoption. Second, the importance of the attention model in forming habits makes a case for pulsing information interventions if feedback is costly to the provider or the receiver. Finally, the direction of the effect observed in the attention model assuages

[2] See Tiefenbeck et al. (2018) and Agarwal et al. (2018). As in these earlier studies, we do not find that the treatments affect the number of showers taken (Tiefenbeck et al., 2018; Agarwal et al., 2018; Tiefenbeck et al., 2019), justifying an analysis at the shower level.

fears that too long a signal will cause receivers to tune out the message – we find no evidence of inattention driven by sustained personalized feedback.

Related literature

Our paper builds upon the recent use of field and natural experiments that temporarily create incentives for randomly-selected individuals to try new behaviors and form new habits, such as (Charness and Gneezy, 2009; Allcott and Rogers, 2014; Loewenstein et al., 2016; Hussam et al., 2017; Larcom et al., 2017; Ito et al., 2018). Our research complements recent work in behavioral economics that is starting to go beyond assuming a habit stock model in interpreting persistence in treatment effects from behavioral interventions, and instead testing the behavioral microfoundations for habit. This nascent area of research includes Camerer et al. (2020) who propose a neuroeconomic model of habit persistence motivated by evidence of dual systems thinking from neuroscience, and (Hussam et al., 2017) who propose a field experiment to test for the existence of rational habit formation (Becker and Murphy, 1988). We discuss the implications of the (Camerer et al., 2020) in our context in Section 4.3 and show that we find little evidence of the discrete behavioural breaks relevant to their model in our particular setting.

Understanding how habits form informs policy debates well-beyond resource conservation, such as those in monetary policy, antitrust, environmental conservation, and public health. By exploiting real-time data and a novel experimental design, we bridge economics, psychology and neuroscience in providing a unique test of their habit stock model, and in establishing a new attention-based behavioral mechanism for habit formation.

2 Model

This section describes a model of persistence that allows for both consumption-based and attention-based habit formation. Our base model is the canonical consumption-based habit stock model of Stigler and Becker (1977). Within this framework, we incorporate salience bias, whereby households may be inattentive to resource use, leading them to only perceive a fraction of the use, and, hence, only a fraction of its cost. We allow for interventions that change the experience of salience, via what we call "attention." Our model allows individual attentiveness to evolve over time: it rises when feedback is provided and wanes when it is taken away. On other words, the effects of salience may persist beyond the presence of active interventions.

2.1 Utility function

In period t a household realizes utility U_t as a function of their current consumption level c_t, habit stock h_t, the attention parameter θ_t, and exogenously–given price p:

$$U_t = u(c_t, h_t) - \theta_t p c_t, \tag{1}$$

The two non-standard elements in this utility function are limited attention and potentially two persistence mechanisms. We now discuss each in turn.

Limited attention. The parameter $\theta_t \in [\underline{\theta}, 1]$ is the household's level of attention to resource consumption, and hence cost, as in Chetty et al. (2009) or Della Vigna (2009). The intuition in our research context is that while the pleasant sensation of a shower is immediately and correctly felt, the associated resource use is difficult to perceive. As in Chetty et al. (2009) and Chetty (2009), we assume that individuals only give weight θ_t to the water use, and hence also its cost, due to limited attention. As $\theta_t \to 1$, the quantity and cost of consumption is correctly perceived. Attention is at its lower bound if $\theta_t = \underline{\theta}$.

This specification for attention in a demand model is reduced-form in the sense that it does not specify a deeper microfoundation. A plausible interpretation, along the lines of Enke and Graeber (2019), is that individuals need to pay attention to perceive their true water use: they observe a signal $z = x + u$, where $x \approx N(x^D, \sigma_x^2)$ is the distribution of their perceived water use, and $u \approx N(0, \sigma_{u,t}^2)$ is a perception error due to limited attention. Given a signal z, the individual rationally infers that her water use x is $E(x|z) = \underbrace{\theta_t x + \theta_t u}_{\equiv \theta_t z} + (1-\theta_t)x^D$. Thus, the attention parameter can be thought of as the signal-to-noise ratio $\theta_t = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_{u,t}^2}$, arising from this signal-extraction problem under limited attention. It naturally implies that a one-liter increase in actual water consumption is perceived as only a $\theta_t \leq 1$ liter increase.[3]

Persistence mechanisms. We model the consumption habit in utility by specifying a quadratic utility function:

$$u(c_t, h_t) = (a + \gamma h_t)c_t - \frac{1}{2}bc_t^2, \tag{2}$$

where a, b and γ are parameters. The latter parameter, γ, governs the impact of the household's consumption habit stock described below.

Both the consumption habit and attention habit can vary over time. The key intuition behind the Stigler and Becker (1977) model is that past consumption increases the marginal utility of current consumption. Thus, this model views habits as arising from long-term consumption complementarities. Past consumption is summarized as a "habit stock" akin to a capital stock:

$$h_t = (1-\delta)c_{t-1} + \delta h_{t-1}, \tag{3}$$

where δ governs the persistence of the habit stock over time. The habit stock changes more slowly as $\delta \to 1$. The parameter γ indicates by how much a one-unit increase in the habit stock changes the marginal utility of consumption.

In specifying the persistence process on attention, we explicitly incorporate the notion that feedback

[3] In the context of shower water usage, this interpretation is consistent with evidence from Tiefenbeck et al. (2018). In particular, they show that individuals with below-average water use tend to overestimate their water use, while individuals with above-average water use tend to underestimate their use. Under this interpretation, the regression coefficient on actual water use in explaining perceived water use can directly be interpreted at θ. Their estimates suggest $\theta \approxeq 0.4$.

affects the attention to consumption in the model, as our interest lies in understanding how this gives rise to changes in attention over time. With this in mind, we specify attention in period t as follows:

$$\theta_t = \begin{cases} 1 & \text{if } \mathrm{FB}_t \text{ is on} \\ \omega_t & \text{if } \mathrm{FB}_t \text{ is off,} \end{cases} \tag{4}$$

where FB_t is a binary variable equaling 1 if real-time feedback on consumption is "turned on" in period t, and 0 otherwise. The variable $\omega_t \in [\theta, 1]$ is the households' attention-stock in period t. Notice that in equation (4), we assume that real-time feedback makes consumption fully salient in period t: $\theta_t = 1$ if FB_t=1.[4]

We build on evidence from neuroscience, specifically Anderson et al. (2011) and Anderson (2016), to specify how recurrent feedback affects attention. This evidence shows that past exposure to a useful stimulus of a feature creates persistent attention to that feature even after the stimulus has been removed. Longer exposure leads to stronger persistence, though it tends to fade out with time, much in the spirit of a habit stock as for consumption habits. This motivates us to formulate the attention stock ω_t build-up and decay process as follows:

$$\omega_t = \begin{cases} 1 \cdot \alpha + (1-\alpha)\omega_{t-1} & \text{if } \mathrm{FB}_t \text{ is on} \\ \theta\alpha + (1-\alpha)\omega_{t-1} & \text{if } \mathrm{FB}_t \text{ is off,} \end{cases} \tag{5}$$

where $\alpha \in (0,1)$ governs the rate at which the attention stock changes, trending towards 1 when feedback is on, and decaying towards the lower bound $\theta < 1$ when feedback is off.

2.2 Optimal consumption and steady-state

We collect the parameters of the model with the vector $\phi = [\alpha, \delta, \theta, \gamma, a, b]$. Under the assumption that households are myopic in making consumption decisions each period, the first order condition that determines the optimal level of consumption in period t is:

$$\frac{\partial U}{\partial c_t} = a + \gamma h_t - b c_t - \theta_t p = 0 \Rightarrow c_t = \frac{a + \gamma h_t - \theta_t p}{b}. \tag{6}$$

Steady state consumption, habit stock, and attention stock are then defined where:

$$c_t = c_{t-1} = c^*; \qquad h_t = h_{t-1} = h^*; \qquad \omega_t = \omega_{t-1},$$

where $\omega^* \in \{\theta, 1\}$.

The model has two steady-states of interest. The first we label the **noFB** steady-state, which corresponds to a setting where the household makes consumption decisions in the absence of feedback for a long time. In our research context, we envisage households being in this steady-state at the start of the trial, before smart shower meters are installed. At this steady-state, salience-bias is at its maximal level with $\theta_t = \theta$. From the

[4] Consistent with this interpretation, Tiefenbeck et al. (2018) find that with real-time feedback, the regression coefficient of actual water use on estimated water use is approximately 1, implying that $\theta = 1$ when $\mathrm{FB}_t = 1$ under our structural interpretation.

first order condition and the habit and attention stock processes, the steady values for consumption, habit stock, and attention stock are given by:

$$c^*_{noFB} = \frac{a - \theta p}{b - \gamma}; \qquad h^*_{noFB} = c^*_{noFB}; \qquad \omega^*_{noFB} = \theta.$$

The second **FB** steady-state corresponds to the opposite scenario, namely when the household makes consumption decisions in the presence of feedback for a long-time. In our field experiment, this corresponds to a household who has been showering with a smart shower meter for many weeks. Here, the household and consumption costs are fully salient such that $\theta_t = 1$. In this steady-state, consumption, habit stock, and attention stock are given by:

$$c^*_{FB} = \frac{a - p}{b - \gamma}; \qquad h^*_{FB} = c^*_{FB}; \qquad \omega^*_{FB} = 1$$

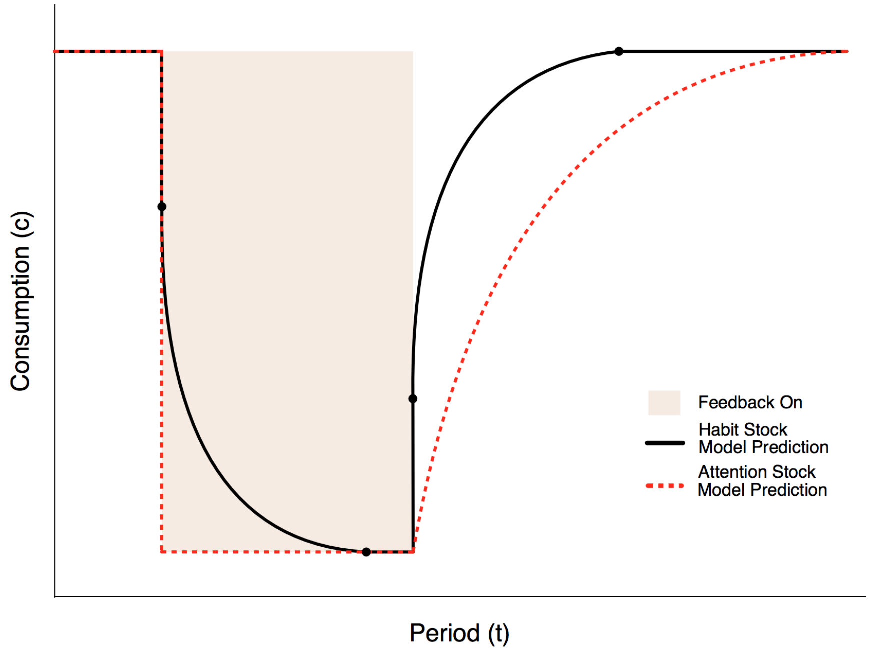
2.3 Steady-state transitions

How does consumption evolve over time when feedback is introduced and removed? How do these dynamics depend on whether consumption habit versus attention habit is the main driver of persistence in behavior? To investigate these transitions, we start with the **noFB** steady-state and turn FB_t on until we reach the **FB** steady-state. We then turn FB_t back off and allow consumption to return to the **noFB** steady-state. In characterizing these transitions, we consider two extreme specifications of the model: one where consumption dynamics are governed solely by consumption habits, and another where they are governed solely by attention habits.

In the first specification, we assume away any transitions in attention habits, fixing $\theta_{t-1} = \theta$ before the feedback-on phase, and θ_t=1 during the feedback-on phase. This environment corresponds to the Stigler and Becker (1977) model of consumption habits combined with the (non-dynamic) limited attention model of Chetty et al. (2009) or Della Vigna (2009). Figure 1 presents the predicted consumption path for this model setup with the solid navy blue line. Starting from the **noFB** steady-state, when feedback is turned on there is an immediate downward jump in consumption due to a salience effect as θ goes from $\theta_{t-1} = \theta$ to $\theta_t = 1$. However, due to the initial decrease in consumption, the habit stock starts falling, thus reinforcing the initial drop by further depressing the marginal utility of consumption. This generates a gradual decay in consumption levels until the **FB** steady-state is eventually reached. When feedback is then subsequently turned off, starting from the **FB** steady-state we see the mirror image: an immediate upward jump in consumption as salience bias returns with $\theta_t = \theta$ followed by a gradual growth in consumption levels as the habit stock evolves until consumption returns to the baseline **noFB** steady-state level.

What is highlighted from the figure, and proven in Appendix A, is that the immediate downward jump and subsequent gradual fall in consumption between the **noFB** and **FB** steady-states when feedback is turned on mirrors the immediate upward jump and subsequent gradual rise in consumption between the **FB** and **noFB** steady states when feedback is turned off. Foreshadowing our empirical analysis below, suppose we were to interpret the path of consumption in Figure 1 between the **noFB** and **FB** as the time-varying

Figure 1: Predicted Transitions Between No Feedback and Feedback Steady States from the Habit Stock and Attention Stock Models



Notes: Consumption paths from between steady states under the habit stock and attention stock model set-up from the model in Section 2.

treatment effect of feedback. Empirically, the consumption habit model predicts *symmetry* between treatment effect build-up when feedback is turned on and treatment effect decay when feedback is subsequently turned off.

To characterize steady-state transitions for a pure attention habit model, we shut down the influence of habit stock on consumption by setting $\gamma = 0$ and allow for dynamics in attention as characterized in equations (4) and (5). The dashed orange line in Figure 1 describes the **noFB** $\rightarrow$ **FB** and **FB** $\rightarrow$ **noFB** steady-state transitions under this model set-up. Starting with the former transition, when feedback is turned on, there is an immediate downward jump to the **FB** steady-state level of consumption with no subsequent second-order decline in consumption as we found with the habit stock model. However, while feedback is on and consumption is at its **FB** steady-state level, the household's latent attention stock is accumulating with ω_t trending upwards towards 1 as per equation (4).

When feedback is subsequently turned off and the second transition from the **FB** steady-state to **noFB** steady-state occurs, it is possible for there to be no upward jump in consumption, and instead consumption gradually rises. This can occur if the attention habit stock ω_t reaches 1, which implies that $\theta_t \approx 1$ in the period immediately after feedback is turned off. For shorter feedback phases, $\theta < \omega_t < 1$, and there would also be a jump upwards in the first period after feedback is turned off, followed by a gradual transition.

Figure 2: Amphiro B1 Smart Shower Meter

3 The experiment

We use a natural field experiment (Harrison and List, 2004) to test for consumption habit and attention habit stocks as underlying mechanisms for habit formation in the context of a feedback intervention. This sections describes the experiment, its implementation, and the data that it creates. We also present summary statistics to characterize the study's internal and external validity.

3.1 Design

Our experiment leverages a smart shower meter – the Amphiro B1 – which provides an individual real-time feedback on total water used and water temperature during a shower. The device, shown in Figure 2, is mounted between the shower hose and a hand-held showerhead and is powered water flow and not batteries. Once water flow stops, the device automatically shuts off after three minutes.

During our experiment, households' Amphiro B1s are in one of two feedback modes, which we present in Figure 3. Panel (a) depicts the *feedback-on* mode in which real-time feedback on shower water volume and temperature is displayed. Panel (b) depicts the *feedback-off* mode in which only feedback on temperature is displayed. We refer to the former mode as the feedback "treatment" mode, while the latter is the "control" mode.[5] By comparing shower water usage within and across households over time under the two feedback modes, we can identify the causal impact of providing real-time feedback on shower water consumption.

Our experimental design is presented in Figure 4. There are seven experimental conditions which we label T1 to T7. In each condition, the Amphiro B1 begins in the no feedback mode and collects baseline shower usage data for 10 showers per person in each household.[6] This is important for establishing balance on pre-feedback shower usage across our seven differential experimental conditions. It also permits a within-

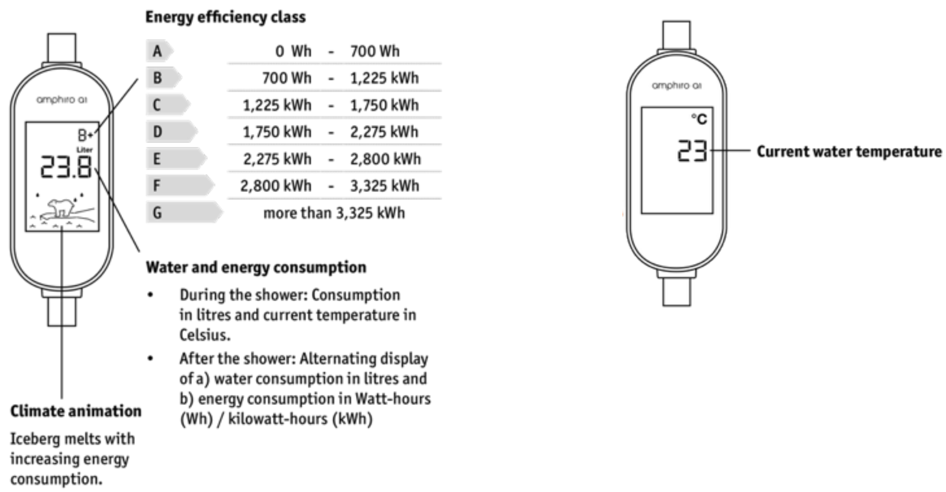
[5] We elected not to have a "pure" control where the Amphiro B1 provides no feedback at all as households might consider seeing no information as meaning their device is not working. By displaying temperature, households in the feedback-off mode can see that their Amphiro B1 is working.

[6] In Section 3.2, we describe in detail on how data is collected for households with one person versus two or more people.

Figure 3: Realtime Feedback Modes for the Amphiro B1

(a) Realtime Feedback On

(b) Realtime Feedback Off



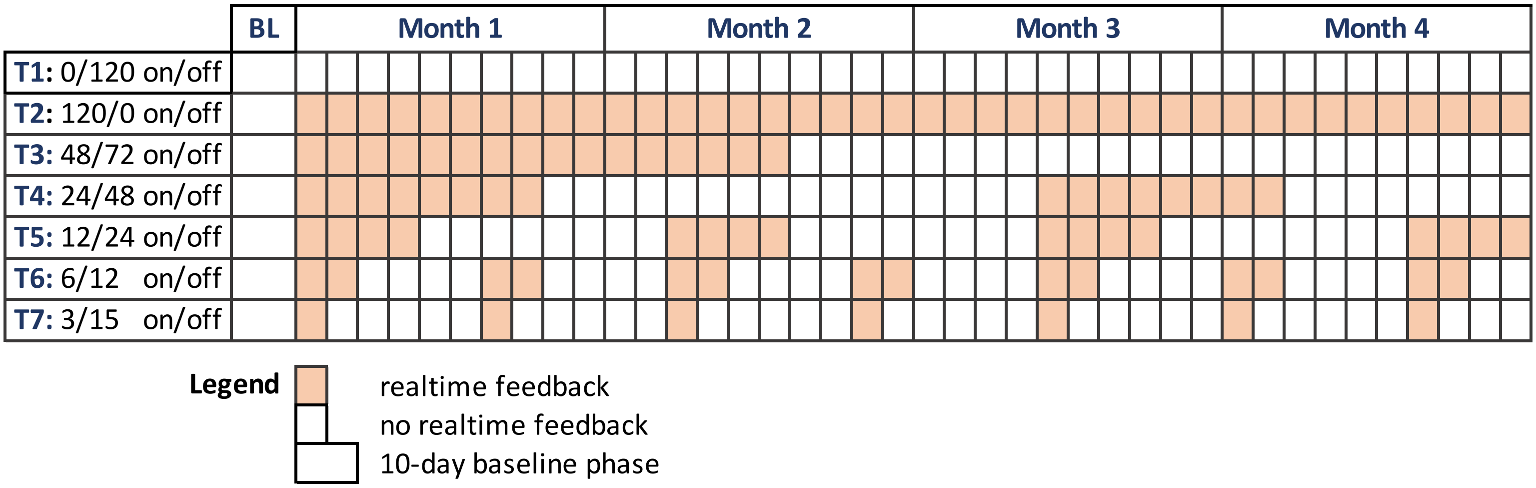
subject experimental design that allows us to employ household fixed effects in identifying consumption responses to real-time feedback, which we discuss in Section 4 below.

After this baseline data collection period, the Amphiro B1s start to cycle between the feedback-on and -off modes at different frequencies across T1 to T7. The first two conditions are important benchmarks. In T1, the Amphiro B1 is in feedback-off mode the entire time. Condition T2, in contrast, always has the device in the feedback-on mode.

Conditions T3 to T7 cycle the Amphiro B1 between the feedback-on and -off modes. For instance, in T3 we turn feedback on for 48 showers after which feedback is turned off for the remaining 72 showers of the experiment. In stark contrast, T7 cycles between turning feedback on for just 3 showers, followed by 15 showers of feedback-off, and then 3 more days of feedback-on, and so on. Conditions T4 to T6 contain intermediate levels of feedback intensity with feedback-on periods of 24, 12, and 6 showers, respectively.

Within our four-month study period, this research design trades off our ability to provide households with long periods of feedback against our ability to see how consumption evolves in the absence of feedback. Having a condition with a sufficiently long feedback cycle helps to ensure that, *a priori*, we observe households reaching a new steady-state level of consumption when feedback is turned on. At the same time, we also need to observe consumption for a sufficiently long period where feedback is turned off to observe households return to their original steady-state, if they do at all. Observing these respective transitions is critical for testing whether steady-state transitions with the introduction and subsequent removal of feedback are symmetric or asymmetric, which is important for testing the consumption habit mechanism. Condition T3 attempts to strike this balance between having a long-feedback cycle followed by a long no-feedback period to test for symmetry in steady-state transitions. Conditions T4 to T7 experimentally vary the duration feedback, thereby enabling us to further study how different levels of feedback intensity affect the rate of habit formation and the related persistent impact of past feedback on consumption once feedback is

Figure 4: Experimental Design



removed.

3.2 Context, recruitment, and implementation

We ran the experiment in 2017 with a large water utility, South East Water, based in Melbourne, Australia. Between April and May 2017 we recruited 700 South East Water customers as follows:

1. From the 700,000-household South East Water residential customer base, we identified 140,407 households that registered email addresses with the utility.

2. A year prior to the experiment, we emailed an online survey to 45,685 households randomly-selected from the 140,407. The survey asked questions about household characteristics, water usage, and shower type.[7]

3. We received 19,449 survey responses. Of these households, 4,999 households reported having a handheld shower nozzle in their primary bathroom, which is necessary for installing an Amphiro B1.

4. We sent a follow-up email to these 4,999 households asking if they were: (1) interested in participating in a trial involving the Amphiro B1; and (2) intending to be at their current address for the duration of the study period. 1,200 households expressed interest and availability. This represents our eligible sample.

5. From the 1,200 eligible households, we randomly selected 700 for the experiment and randomly-allocating 100 households to each experimental group T1-T7. We stratified allocation by household

[7]The list of survey questions and answers is presented in Appendix C. The survey was part of a larger research project at South East Water on household behavior. As part of this project, we emailed our survey to three mutually-exclusive randomly-selected groups of households: 25,685 households in September 2016, 10,000 in December 2016, and 10,000 in March 2017. As an incentive for responding to the survey, households were entered into three different lotteries for $1,000 of their water bills for the year, an iPad valued at $1,000, and a $1,000 charitable donation. Survey winners were excluded from the subsequent experiment.

size to prioritize all single-user households.

At the end of May 2017, we mailed the 700 Amphiro B1s to our experimental sample. Households were provided both paper and online tailor-made instructions and were asked to install their device upon receiving it. At the start of October 2017, we emailed households asking them to mail back their Amphiro B1 to The University of Melbourne for data extraction using a self-addressed stamped envelope that was sent with the device.[8] In total, 555 of the 700 devices were returned for data extraction.

3.3 Data

The Amphiro B1 records shower number, total water used, the water flow rate of each shower taken, and water temperature. Because the device does not have a battery, it does not have an internal clock, which means that the calendar date and time of the day a shower is taken is not recorded.

The meter's internal memory saves a maximum of 245 consecutive showers-worth of data. If we assume, for example, two showers per person per day in a two-person household, over a four month period we can expect the household to take approximately 240 showers. We therefore work with a four-month experimental period, as this corresponds to the upper bound of the amount of data that can be stored for a two-person household. Such households characterize 34% of the households in our experiment.

Working with the engineers at Amphiro, we programmed the devices to create the feedback-on and -off cycles across the seven different experimental conditions T1-T7 in Figure 4. This ensures that we also know the exact shower where feedback is turned on or off for a given household. The feedback cycles were calibrated depending on whether a household had one or two or more individuals. Single-person households had feedback cycles programmed on their Amphiro B1's exactly as described in Figure 4. Households with two or more residents had their Amphiro B1's programmed such that the feedback on and off were twice as long as described in Figure 4, collecting a maximum of 245 showers-worth of data.[9] The doubling of feedback-on and off cycle lengths for multi-person households aims to approximate, per person, the feedback cycles for one-person households. A two-person household where individuals alternate in showering once per day perfectly aligns with Figure 4 on a per-person basis.

Randomization of households into experimental conditions ensures that the household size distribution across conditions is balanced. Therefore, differences in household size across conditions will not be a source of bias in generating differences in households' consumption responses to real-time feedback across conditions. The main concern with multi-person households instead relates to the precision of our treatment effect estimates. To the extent that multi-person households create a mismatch between the programmed feedback cycles on the Amphiro B1s and the realized feedback cycles on a per-person basis within a household, our estimates of the relationship between consumption and real-time feedback cycle length will be potentially

[8] Households were informed that their Amphiro B1 would be mailed back to them once the data were extracted with the device reset to "factory" mode with full-functionality. During our experiment, Amphiro prevented households in our experiment from downloading the Amphiro B1 app from all app stores where it is offered. After data extraction and verification, we reset households' Amphiro B1 and informed households how to pair their device with the app, which enables them to view with their historical shower data.

[9] 17% of households have one-person, 34% have two people, 22% have three people, and 27% have four or more people.

Table 1: Sample Means for Household Water Usage and Household Account Characteristics Across Different Sub-Samples

	Sub-Sample				
	All Households with Email (1)	Emailed Survey (2)	Answered Survey (3)	Sent Amphiro B1 (4)	Returned Amphiro B1 (5)
Jul-Sep 2016 Water Usage (L)	31.45	31.59	29.56	29.31	28.86
Oct-Dec 2016 Water Usage (L)	37.90	37.88	36.06	34.07	33.75
Jan-Mar 2017 Water Usage (L)	45.41	45.43	44.07	38.68	38.44
Apr-Jun 2017 Water Usage (L)	39.35	39.34	37.52	34.99	34.46
Annual HH Income (1000s)	53.26	53.22	52.55	51.75	51.86
Average Age	37.67	37.62	37.82	36.52	36.42
Share of High School Graduates	0.46	0.46	0.45	0.45	0.45
Number of Bedrooms in Home	2.94	2.95	2.97	3.04	3.05
Share of Tenants	0.34	0.33	0.19	0.19	0.19
Share of HHs with Electronic Billing	0.47	0.49	0.68	0.75	0.79
Share of HHs Registerd with Web Portal	0.38	0.39	0.52	0.64	0.66
Number of People Living at Home			2.67	2.64	2.60
Self-Reported Shower Time			6.47	6.90	6.88
Number of Leaks Checks per Year			2.30	2.26	2.21
Number of Households	140407	45685	19449	700	555

attenuated and less precise.[10]

Other data sources

We use data from three other sources. The first is baseline survey data on household characteristics and water usage that we collected from each household prior to the experiment; see Appendix C for the survey. Second, we obtain pre-experiment billing and household account data including quarterly water usage and bills, electronic-billing status, hardship status and tenant status. Finally, we anonymously match households to their Statistical Area 1 (SA1) 200-person census block from the Australian Bureau of Statistics to obtain demographics such as average household income, age, education, and home size.[11]

3.4 Summary statistics

Table 1 presents sample means for households' water usage and characteristics across various samples. Comparing columns (2) and (3), we find that households who answered the baseline survey are 14% less likely to be tenants, 19% more likely to have electronic billing, and 13% more likely to have registered with South East Water's online web portal for managing their bills. There are no other statistically-significant differences between baseline survey respondents and non-respondents in terms of water usage or other demographics.

[10] Households with more people tend to have more than one shower in their home. Our main issue is, more precisely, whether more than two people use the shower in which they installed their Amphiro B1. We expect this to be the "main" shower in the home where the home owners / rent payers live as these are people who have their emails registered with South East Water, and is thus who we corresponded with in recruiting our experimental sample.

[11] SA1's contain 150 households on average. They are the most narrow census block that the Australian Bureau of Statistics makes publicly available. See http://www.abs.gov.au/ for details on SA1s.

Comparing columns (3) and (4) allows us to see whether survey respondents differ from households who were eligible to be part of our trial and who were randomly-chosen to be sent an Amphiro B1. We again see statistically significant differences related to electronic billing and web portal registration: households in the trial tend to be more likely to exhibit these account characteristics. Otherwise, there are minimal differences in household quarterly water usage and other characteristics between our 19,407 survey respondents and the 700 households in our trial.

Finally, by comparing columns (4) and (5), we can see if there is evidence of selection into returning the Amphiro B1. Here, we find virtually no differences between the group of households who were originally sent Amphiro B1s, and the 80% of households who cooperated and eventually returned their device for data extraction.

Table 2 presents an analogous set of mean characteristics for households across each of the seven different experimental conditions. The first three rows also add sample means for shower water usage volume, flow rate, and shower length. These variables are constructed by first computing means household-by-household using their initial 10-shower per person baseline phase. In the table, we report the sample mean of these mean characteristics across households in each experimental condition.

Looking across the bottom row of Table 2, we see that the number of households in each group ranges from 75 to 86, highlighting a similar Amphiro B1 return rate in each group. In terms of baseline showering data, we find very similar baseline levels of water usage, shower flow rates, and shower length across all conditions. Indeed, none of the differences are jointly statistically significant nor do any pairwise comparisons across groups yield statistically significant differences. We further find households exhibit similar quarterly water usage levels, have similar demographics, and similar household account characteristics. In sum, our randomization achieves balance on observables across our seven experimental conditions.

4 Testing for persistence in feedback effects

In this section, we study the treatment effects induced by our feedback intervention. We first graphically describe these treatment effects across each of our seven experimental conditions. Motivated by these figures, we then develop a regression analysis that allows us to test for persistence in the data.

4.1 Graphical analysis

Figure 5 graphically describes time-varying treatment effects from our experiment. To construct these figures, we run regressions of the following form:

$$y_{is} = \alpha_i + \sum_{b=1}^{B} \beta_b \left(\mathrm{T}g \times 1\{s \in b\}\right) + \delta_s + \varepsilon_{is} \tag{7}$$

where y_{is} is shower volume for household i in shower s, Tg equals one if household i is in treatment group Tg, $g = 2, \ldots, T$, $1\{s \in b\}$ is a dummy equaling one if shower s is within three-shower block b, α_i and

Table 2: Sample Means for Household Water Usage and Household Account Characteristics by Treatment Condition

	Experimental Condition (Realtime Feedback On/Off Cycles)						
	T1 (0/120)	T2 (120/0)	T3 (48/72)	T4 (24/48)	T5 (12/24)	T6 (6/12)	T7 (3/15)
Shower Water Usage Volume (L)	55.12	55.76	55.73	55.97	54.19	54.92	56.83
Shower Flow Rate (L/sec)	8.35	8.29	8.58	8.25	8.65	7.88	8.36
Shower Length (min)	6.70	6.78	6.79	6.81	6.45	7.21	7.08
Jul-Sep 2016 Water Usage (L)	28.85	27.06	30.72	29.69	29.63	25.49	30.58
Oct-Dec 2016 Water Usage (L)	33.20	34.14	34.89	35.18	31.85	30.50	36.33
Jan-Mar 2017 Water Usage (L)	34.76	37.07	39.58	41.58	35.77	39.18	41.05
Apr-Jun 2017 Water Usage (L)	31.10	33.28	34.89	35.85	33.73	32.77	39.67
Annual HH Income (1000s)	49.17	50.47	54.69	50.40	54.06	52.70	51.72
Average Age	35.77	37.09	37.60	35.23	35.74	36.83	36.75
Share of High School Graduates	0.43	0.45	0.45	0.44	0.45	0.46	0.47
Number of Bedrooms in Home	3.06	3.00	3.11	3.11	3.15	2.98	2.94
Share of Tenants	0.19	0.24	0.21	0.13	0.18	0.15	0.20
Share of HHs with Electronic Billing	0.77	0.81	0.79	0.80	0.83	0.77	0.78
Share of HHs Registerd with Web Portal	0.70	0.64	0.63	0.79	0.64	0.58	0.64
Number of People Living at Home	2.47	2.53	2.77	2.64	2.54	2.66	2.61
Self-Reported Shower Time	6.47	6.92	6.42	7.39	6.46	6.68	7.77
Number of Leaks Checks per Year	2.08	2.22	2.24	2.38	2.17	2.22	2.18
Number of Households	77	84	79	86	78	76	75

Notes: See Figure 4 for definitions of experimental conditions T1–T7.

δ_s are household and shower fixed effects, and ε_{it} is the regression error. Estimating β_b based on three-shower per-person blocks rather than for each individual shower reduces the noisiness of our time-varying treatment effects due to idiosyncratically long or short showers, thereby allowing us to better visualize trends in treatment effects when feedback is turned on and off. Given a maximum of 120 showers per-person in our sample, we plot β_b coefficient estimates for $B = 40$ three-shower per-person blocks in total.[12]

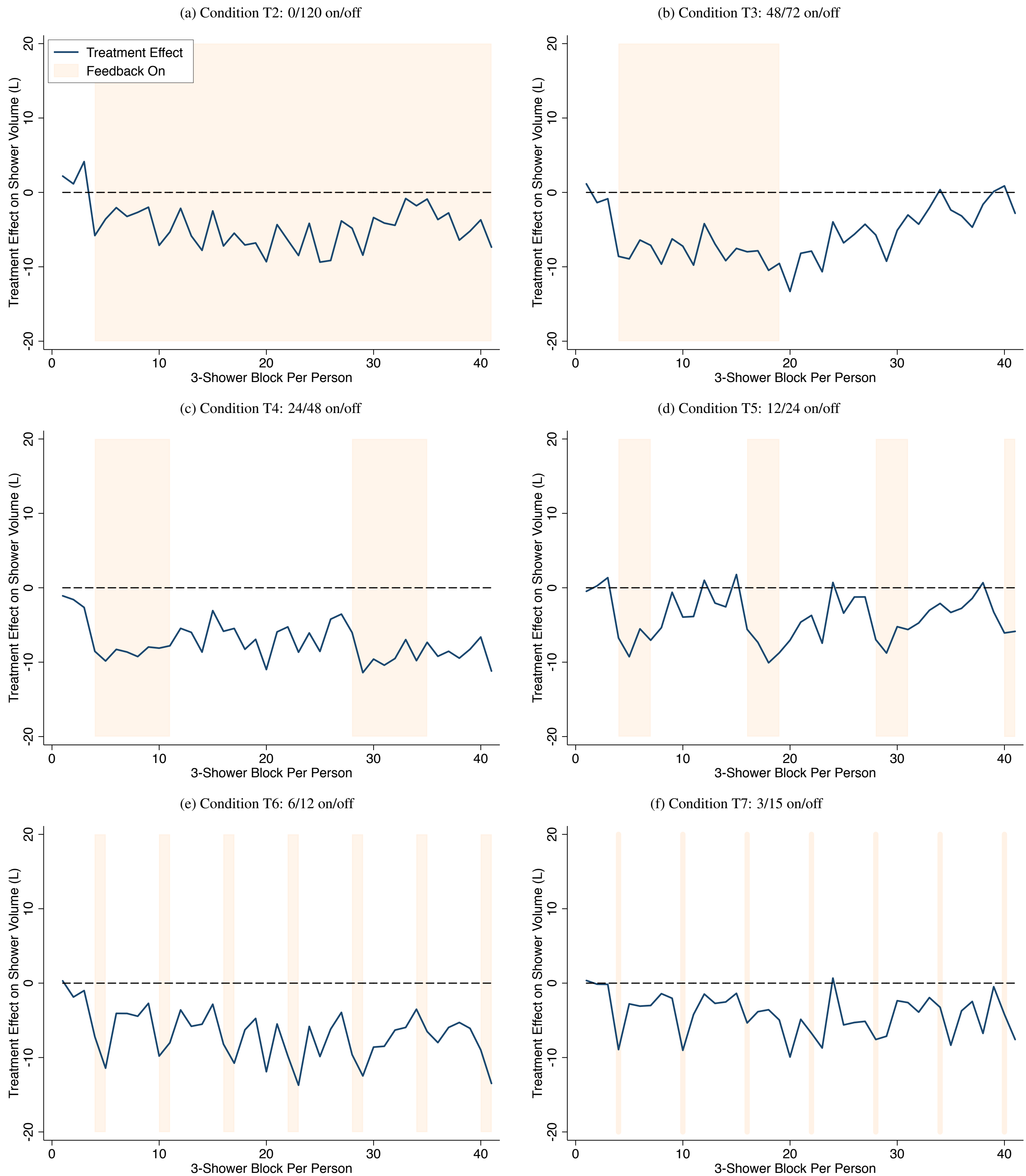
We estimate (7) separately for each group $g = 2, \ldots, 7$ where for a given group we use households in T1 (control) and Tg to estimate (7). By plotting the coefficients estimates $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_B$, we can visualize the time path of treatment effect build-up and decay when feedback is turned on and off across conditions T2 to T7. In this way, the coefficients let the "data speak" to the persistence in feedback effects induced by our experiment.

Panels (a)-(f) of Figure 5 present time-series plots of our coefficient estimates for conditions T2–T7.[13] Four notable patterns emerge. First, all panels reveal an immediate drop in shower water usage when feedback is turned on following the baseline phase. Second, there is little evidence of a subsequent downward trend in water usage following the initial drop in usage after feedback is turned on. This is particularly clear in panels (a)-(c) with longer cycles of real-time feedback. Third, water usage does not immediately jump up after feedback is turned off, and instead gradually trends back to baseline levels

[12] We break up the baseline phase into three per-person shower blocks for showers one to four, five to seven, and seven to ten.

[13] For the sake of brevity in the figures, we do not report confidence intervals. We defer formally testing for jumps and trends in consumption with and without feedback induced by our experiment to our structural analysis in Section 5 below.

Figure 5: Time-Varying Treatment Effects by Experimental Condition



Notes: See Figure 4 for details on the experimental design and equation (9) and associated discussion in the text for the regression equation used to generate these plots. For brevity, confidence intervals are not displayed.

after feedback is turned off. Interestingly, this persistence in feedback effects exists even in panels (e) and (f) under very short feedback periods.

The final notable result in panel (b) is that it shows that consumption eventually returns towards baseline levels if feedback remains off for a sufficiently long period. This is an important benchmark as it suggests that under our highest intensity feedback condition, our experimental design entails a sufficiently long window without feedback to reveal the transition back to baseline steady-state levels of consumption.

4.2 Treatment effects

The visual evidence from Figure 5 shows considerable persistence in each of the treatments. In a next step, we proceed to formally test the pattern of persistence in the different conditions by estimating the following two equations:

$$y_{is} = \alpha_i + \beta_1 ON_{is} + \beta_2 OFF_{is} + \delta_s + \varepsilon_{is}, \tag{8}$$

and

$$y_{is} = \alpha_i + \beta_1 ON_{is} + \beta_2 PostON_{is} + \beta_3 OFF_{is} + \beta_4 PostOFF_{is} + \delta_s + \varepsilon_{is}, \tag{9}$$

where y_{is} is again shower volume for household i in shower s, ON_{is} is a dummy equaling one if feedback is on for household i in shower s, $PostON_{is}$ is the number of showers since feedback was first turned within a current feedback-on spell, OFF_{is} is a dummy equaling one if feedback is off for household i in shower s and where s is after the baseline phase, and $PostOFF_{is}$ is the number of showers since feedback was first turned off within a current feedback-off spell.[14] All of our regressions include household and shower fixed effects, α_i and δ_s. We therefore identify feedback treatment effects on water usage using within-household variation in consumption, while simultaneously accounting for confounding factors such as seasonality in shower water usage through the time fixed effects. The econometric error term, ε_{it}, is clustered at the household–level.

Equation (8) provides the most basic test of persistence, as we can test whether consumption returns to baseline after feedback is turned off using our β_2 estimate. Equation (9) provides a first look at the dynamics of how feedback affects behavior. Specifically, it allows us to see feedback effects build-up while feedback is turned on using our β_1 and β_2 estimates, and how they decay after feedback is turned off using our β_3 and β_4 estimates.

Results

Table 3 presents our results in two panels. The top panel presents benchmark estimates from equation (8), while the bottom panel shows the estimation results from equation (9). In both panels the first column

[14] To take a concrete example, consider treatment group T4 with a 24/48 on/off feedback cycle. As depicted in panel (c) of Figure 5, this condition has two feedback-on and two feedback-off spells. Feedback is first turned on at shower 11, after the baseline phase. During the first feedback-on spell between showers 11 and 34, $PostON_{is}$ counts up from $1,2,\ldots,24$. After a 48-shower feedback-off spell between showers 35 and 83, a second feedback-on spell starts at shower 84. During this second feedback-on spell , $PostON_{is}$ once again counts up from $1,2,\ldots,24$ during showers 84 to 108. $PostOFF_{is}$ similarly counts up during the feedback-off spells.

Table 3: Regression Results by Experimental Condition

	Experimental Conditions Included in the Sample						
	T1-T7	T1,T2 0/120 on/off	T1,T2,T3 48/72 on/off	T1,T2,T4 24/48 on/off	T1,T2,T5 12/24 on/off	T1,T2,T6 6/12 on/off	T1,T2,T7 3/15 on/off
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ON_{is}	-7.33***	-7.14***	-7.52***	-7.20***	-7.24***	-7.48***	-7.02***
	(0.70)	(1.39)	(1.10)	(1.06)	(1.05)	(1.09)	(1.07)
OFF_{is}	-3.82***		-4.95***	-4.21***	-2.99**	-4.00***	-3.50***
	(0.72)		(1.41)	(1.23)	(1.22)	(1.05)	(1.04)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.46	0.44
Observations	87861	25134	38366	38984	37396	36783	36868
ON_{is}	-7.36***	-6.57***	-7.26***	-7.26***	-7.12***	-7.47***	-6.90***
	(0.70)	(1.39)	(1.09)	(1.06)	(1.04)	(1.08)	(1.05)
$PostON_{is}$	0.01	-0.02	-0.01	0.00	-0.01	-0.00	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
OFF_{is}	-4.86***		-7.89***	-5.36***	-5.25***	-4.98***	-3.69***
	(0.73)		(1.43)	(1.29)	(1.24)	(1.24)	(1.18)
$PostOFF_{is}$	0.08***		0.11***	0.06*	0.19***	0.14	0.03
	(0.02)		(0.03)	(0.03)	(0.06)	(0.09)	(0.07)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.46	0.44
Observations	87861	25134	38366	38984	37396	36783	36868

Notes: Dependent variable is shower water usage volume with baseline mean of 57 L (s.d.=42 L). All regressions include household and shower fixed effects. Figure 4 provides a visual representation of our experimental design, conditions T1–T7 and their feedback on/off spells. Standard errors are clustered at household level. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

presents estimates from pooling all 7 experimental conditions. Column (2) shows the estimate of the feedback effect using observations from the control group (T1) and feedback always-on condition (T2). Columns (3) to (7) estimate persistent effects by including one of the persistence conditions T3 to T7 in the sample along with T1 and T2 households to reliably identify the impact of feedback on consumption.

Turning to the first panel, column (1) shows that providing real-time feedback sharply lowers water use during showers. The point estimate of -7.33 L/shower is comparable to earlier studies using similar interventions (Tiefenbeck et al., 2018). The estimates also show that, across all conditions, there is a clear persistence effect that is highly statistically significant. It is remarkable that these persistence effects arise even in condition T7, where feedback is only given for 1/6th of the time in a 3/15 feedback on/off cycle. Nevertheless, this leads to an estimated persistence effect of -3.50L/shower, i.e. half of the corresponding -7.02L/shower treatment effect from providing feedback. Columns (1) to (7) also show that the treatment effect from feedback is nearly identical across conditions. However, the persistence effects show a tendency of being larger in conditions with a higher intensity of feedback, particularly in T3.

To further explore persistence effects, we turn to the estimates of equation (9) in the bottom panel of

Table 3. The estimation results in column (1) reveal a general asymmetry between real-time feedback effects and persistence effects: the point estimate of the *ON* coefficient now reflects the impact of feedback in the first episode of a feedback period. It is virtually identical to the estimate in the top panel. The interaction effect with the duration of exposure to feedback is a precisely estimated zero. That is, real-time feedback has a stable effect on behavior. By contrast, the estimates of the coefficient on $PostOFF$ show that the persistence effects slowly erode over time: the column (1) estimates suggest that with every shower during a feedback-off phase, the persistence effect decreases by 80 milliliters. The estimate is statistically significant at the 1% level.[15]

Turning to the results in columns (3) to (7) of Table 3, we find nearly identical stable feedback effects in each of the columns. Here, controlling for the duration of off-periods allows a clearer interpretation of the persistence effects as the persistence effect in the first episode of an off-period. We find a monotonicity in the point estimates. In particular, in column (3), the point estimate of β_3 is -7.89L after 48 periods of feedback and statistically indistinguishable from the feedback effect itself. The estimates of β_3 on OFF_{is} in equation (9) monotonically decline as the duration of feedback phases becomes shorter. Even in the treatment with the shortest feedback phase in column (7), there is a significant persistence effect. Notice also that each of the point estimates of β_4 on $PostOFF_{is}$ is positive, though not always significant. In part this is due to the shorter feedback-off phases, which reduces the precision with which the parameters can be identified.

Overall, the pattern that emerges from our reduced-form analysis is that there is clear evidence of persistence following feedback. However, consumption dynamics during the feedback-on and -off phases are different. When feedback is turned on, feedback effects emerge immediately and are stable thereafter. In contrast, persistence effects when feedback is turned off scale with feedback intensity and gradually erode over time.

4.3 Continuous vs. discrete persistence effects

Our empirics thus far point toward attention habit as the source of persistence effects and are not necessarily supportive of consumption habits. However, before moving onto a more formal analysis of these two models, we highlight that the decay in persistence effects that we find is consistent with a completely different mechanism: it could reflect consumers jumping back to baseline consumption at different points in time after feedback is turned off.

Indeed, a burgeoning area of research in neuroscience on habit (e.g., Wood and Runger, 2016) suggests this is an important potential confound. This research emphasizes "automatic control" models of habit, where habit is an automatic or default decision-making process. In these models of decision-making, individuals' automatic choices give rise to persistence in decision-making, as well as infrequent and discrete changes in decisions if an individual's external environment changes sufficiently causing them to update their automatic/default choice.[16] In our case, the introduction and removal of feedback could represent such

[15]It is natural to ask how many showers does it take until the water-conserving habit in consumption induced by feedback fully decays? We formally quantify this using our structural model in Section 5 below.

[16]Camerer et al. (2020) have recently developed a dual-systems neuroeconomic of habit involving automatic control and default decision-making when one's decision-making environment is stable, and deliberation and infrequent updating of decision rules

a sufficiently large change in an individual's decision-making environment that feedback-induced changes in consumption over time could reflect changes in automatic choices at the individual-level.

If such habit–as–automatic–control governs behavior at the individual level, it is possible that households' water consumption exhibits discrete jumps at different points in time after feedback is removed. Thus, it is possible that automatic control generates the smooth decay that we find after feedback is turned off even though behavior at the individual level is discrete. In this case, it would be inappropriate to interpret the estimated decay in treatment effects when feedback is turned off as being generated by an attention habit mechanism for persistence.

To address this potential concern, we provide an extensive analysis of habit–as–automatic–control in Appendix D. In this appendix, we adapt our regression models in (8) and (9) to allow for household-specific jumps in consumption after feedback is turned off, where both the timing and magnitude of jumps are allowed to vary household-by-household. We correct for the jumps in an iterative, two-step procedure. First, we determine the timing and magnitude of the jumps by running structural break tests a household at a time, and picking a jump date for every feedback episode that best explains the data. Having determined dates for each of the households, we then re-estimate models (8) and (9), but allow for household-specific jumps in consumption after feedback is turned off. We iterate on this two-step procedure until the regression coefficients and estimated jumps jointly converge. They do so quickly.

Through this analysis, we find little evidence to support a habit–as–automatic–control model of decision-making in our setting. Controlling for the household-specific jumps in consumption when feedback is turned off leaves our treatment effect estimates from Table 3 largely unchanged both in terms of their magnitude and precision.

5 Consumption habits versus attention habits

Our reduced-form tests establish persistence effects across each of our experimental conditions. They do not, however, leverage the additional structure that different models of habit impose. In this section, we develop an econometric model that captures key features of persistence implied by consumption habits and attention habits, and nests them in one model.

5.1 An empirical model of persistence

We define a non-linear regression model of the form:

$$y_{is} = \alpha_i + \beta_1 ON_{is} + \beta_2 ON_{is} \times A_{is}(\lambda) + \beta_3 OFF_{is} + \beta_4 OFF_{is} \times A_{is}(\lambda) + \delta_s + \varepsilon_{is}, \tag{10}$$

when an environment is sufficiently altered.

where $A_{is}(\lambda)$ captures the habit stock of past feedback for individual i in shower s. It is defined as a weighted sum over the history of i's feedback:

$$A_{is}(\lambda) = \sum_{k=1}^{s-1} w(k,\lambda) ON_{is-k}. \tag{11}$$

The parameter λ shapes the weight of past periods. Our approach closely parallels Malmendier and Nagel (2011) by specifying the weights as:[17]

$$w(k,\lambda) = \frac{k^{\lambda}}{\sum_{m=1}^{T} m^{\lambda}}. \tag{12}$$

Panel (a) of Figure 6 plots the weights $w(k,\lambda)$ for different values of λ for lags $k = 1,...20$. As the figure shows, when $\lambda \approx 0$, the weights become nearly uniform over time, giving each period the same weight on $A_{is}(\lambda)$, no matter how far in the past. As λ becomes negative, more weight is given to more recent periods.[18]

Panel (b) of the figure shows how habit stocks $A_{is}(\lambda)$ evolve for different values of λ. In this particular example, we plot $A_{is}(\lambda)$ over time for treatment condition T3, which starts with 48 periods of feedback, followed by 72 periods without. All habit stocks start out at zero in the 10-shower baseline period. For $\lambda = -1$, the habit stock accumulates rapidly and begins to plateau. Similarly, it decreases rapidly when feedback is subsequently turned off. Values of λ closer to zero produce a slower build-up, but also show more persistence in the decay. Thus, the function $A_{is}(\lambda)$ has the same properties as the habit stocks in the consumption habit and attention habit models from Section 2.

Notice that with our choice of the denominator in equation (12), all $A_{is}(\lambda)$'s reach 1 after T periods of feedback. We choose $T = 130$, the number of showers that are part of our study design, including the baseline phase. With this normalization, the coefficients β_2 and β_4 thus represent the persistence effect after 130 periods of feedback.

5.2 Hypotheses

The two models of habit formation make different predictions with regard to how the feedback stock, operationalized as $A_{is}(\lambda)$ in our empirical model, affects behavior.

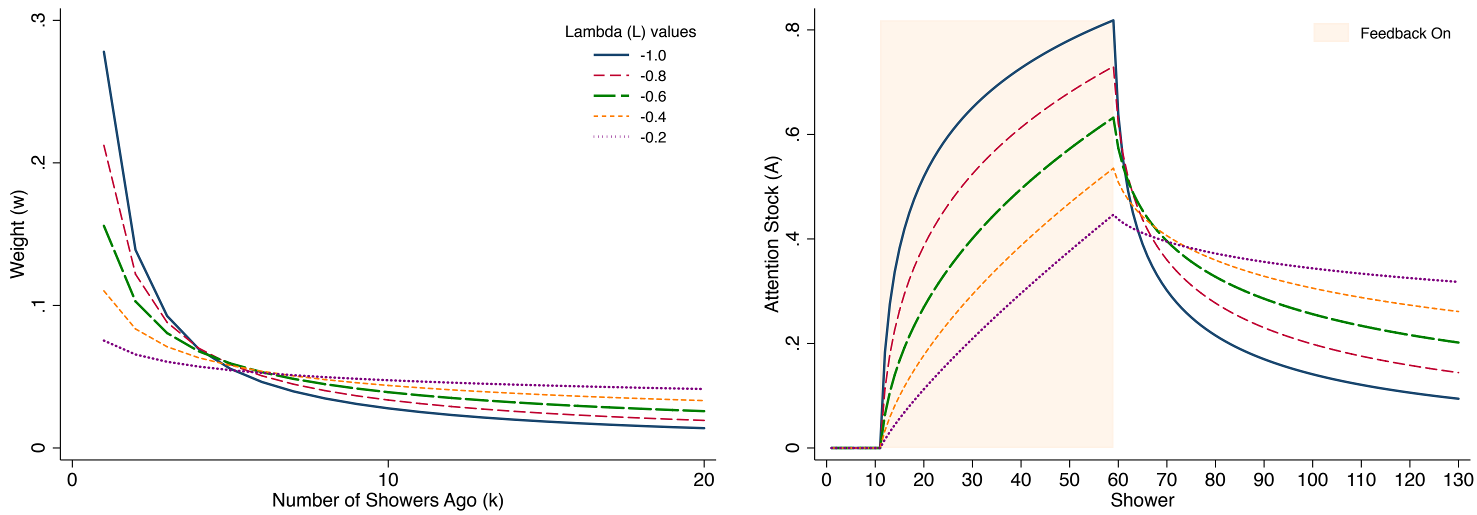
The consumption habit model predicts a jump in behavior with the onset of feedback. As feedback continues, the consumption habit stock starts to fall, thus reinforcing the initial drop and leading to a gradual convergence to a new steady-state. Similarly, when feedback is removed, water use increases immediately. Because of that increase, the habit stock will also increase over time, thus reinforcing the initial increase and leading to a gradual convergence back to the initial steady-state. Finally, recall that our model predicts that treatment effect build-up when transitioning from the **noFB** to the **FB** steady-state and treatment effect

[17] Our denominator differs from the one chosen in Malmendier and Nagel (2011), where the aim was to create a moving average with varying weights over time. Here, we choose a fixed denominator in order to be able to model how a stock variable increases or decreases over time.

[18] Our non-linear least squares estimator also λ to take on any value, including positive ones which instead imply more weight is put on $A_{is}(\lambda)$ for larger values of k. That is, earlier feedback has more influence on current consumption than more recent feedback.

Figure 6: Evolution of the Weighting Function and Attention Stock for Different λ Values

(a) Weighting Function $w(k,\lambda)$ (b) Attention Stock $A_{is}(\lambda)$



decay when transitioning from the **FB** to the **noFB** steady-state are symmetric.

These predictions imply that the feedback stock $A_{is}(\lambda)$ is the sole source of persistence and acts symmetrically, independently of whether feedback is on or off. This implies the testable restriction $\beta_2 = \beta_4 < 0$. Our formulation of $A_{is}(\lambda)$ also captures the essential features of the model that longer phases of feedback lead to a larger build-up of the habit stock, and that longer off-phases lead to more depreciation of feedback. Furthermore, the habit stock captures all the persistence effects under this model. There should not be a separate feedback effect during feedback-off phases, implying that $\beta_3 = 0$. Summarizing,

Hypothesis H1 [Consumption-habit]. Under the consumption-habit model (Stigler and Becker, 1977), $\beta_2 = \beta_4 < 0$ and $\beta_3 = 0$.

The attention habit model predicts that feedback induces full attention ($\omega_t = 1$) to water use. In addition, the attention habit starts building up with every period that feedback is on, which also fits the parametrization of $A_{is}(\lambda)$. However, because feedback already induces full attention, the attention habit stock has no additional impact on behavior during the feedback-on phases, which implies $\beta_2 = 0$.

When feedback is turned off, ω_t falls from 1 to the attention habit stock that has been built up by the end of the feedback-on phase, thus inducing higher water use. With every period during the feedback-off phase, attention wanes and ω_t continues to decline toward its baseline value θ. Thus, the model implies that $\beta_4 < 0$. Finally, just like the consumption habit model, the attention habit model predicts no persistence beyond the pattern captured by $A_{is}(\lambda)$. Thus, the model also predicts that $\beta_3 = 0$. This can be summarized in our second hypothesis,

Hypothesis H2 [Attention-habit]. Under the attention-habit model, $\beta_4 < 0$, but $\beta_2 = 0$ and $\beta_3 = 0$.

Learning

Another potential source of persistence that we have not yet ruled out is learning: individuals may be imperfectly informed about water use. Prior evidence suggests a bias towards the mean as individuals tend to underestimate the resource intensity of behaviors with above-average intensities, and overestimate the intensity of below-average behaviors (Attari et al., 2010; Attari, 2014). Showering is, both, water and energy intensive and thus likely underestimated. Thus, even a few periods of feedback could correct such biased perceptions and lead to a *permanent* reduction even when feedback ends.

Since we identify consumption- and attention-based channels for habit formation entirely through the non-linear build-up and decay of feedback effects, we can capture potential learning effects through the feedback-off phase main effect and its associated coefficient, β_3.

As a further check, we also introduce an interaction effect for the first instance of feedback in equation (10). If learning plays an important role in creating persistence effects, then the first phase of feedback should have the combined effect of learning plus the increased attention toward water use. That interaction effect should therefore be significantly negative under the hypothesis that feedback-induced learning leads to a permanent change in consumption behavior. Habit build-up and decay in subsequent feedback-on and -off cycles after the first cycle will be driven solely by the habit stock as the household would have already learned about their water usage in the first cycle. In this way, having multiple feedback-on and -off cycles in our experimental design, combined with our high-frequency consumption data, plays an important role in enabling tests for learning as a separate channel for persistence effects.

5.3 Estimation

We estimate the model parameters $\theta = [\beta_1, \beta_2, \beta_3, \beta_4, \lambda]'$ by least squares. Since λ enters the regression equation non-linearly, we perform a grid search over a fine grid. Conditional on a candidate λ value, the $\beta_1, \ldots, \beta_4$ coefficients are estimated by OLS. The non-linear least squares estimate of θ from our grid search over λ yields the best-fitting conditional OLS estimate. We bootstrap this two-step estimation procedure to obtain standard errors and conduct significance tests. We follow Cameron et al. (2008) and cluster at our level of randomization, the individual level, to account for persistence in unobserved consumption shocks.

5.4 Results

Table 4 contains our parameter estimates. Column (1) is our main specification of interest as it flexibly allows for asymmetry (or lack thereof) in the build-up and decay of habits when feedback is turned on and off. Our point estimate of the persistence parameter λ in column (2) is -0.60, with a standard error of 0.15. This estimate implies an intermediate degree of persistence building up in response to feedback, with more recent feedback having a large effect on current consumption. Panel (a) of Figure 7 shows the implied weights for the specification. The point estimate implies that the half life of a persistence effect is about 30 showers. The estimated standard error sets it apart from very short-lived persistence such as, e.g. $\lambda = 1$, where only a quarter of the effect persists after 30 days. Nevertheless, the estimate also clearly rejects a

Table 4: Non-linear Regression Results

	(1)	(2)	(3)	(4)
λ	-0.60***	-0.59***	-0.57***	-1.04***
	(0.15)	(0.15)	(0.13)	(0.37)
A_{is}				-2.90**
				(1.27)
ON_{is}	-7.39***	-6.95***	-6.51***	-2.97***
	(0.66)	(0.85)	(0.62)	(0.43)
$ON_{is} \times A_{is}$	-0.50	-0.02	-0.52	
	(1.52)	(1.64)	(1.53)	
$ON_{is} \times FIRST_{is}$		-0.70		
		(0.67)		
OFF_{is}	-1.38	-1.06		
	(0.78)	(0.90)		
$OFF_{is} \times A_{is}$	-10.91***	-11.07***	-12.40***	
	(2.36)	(2.36)	(2.18)	
R-Squared	0.43	0.43	0.43	0.43
Observations	87861	87861	87861	87861

Notes: Dependent variabale is shower water usage volume with baseline mean of 57 L (s.d.=42 L). All regressions include household i and shower s fixed effects. Bootstrap standard errors clustered at household level are presented.

model of permanent change following feedback where, for example, $\lambda = 0$.

Turning to the parameter estimates governing the impact of the habit stock $A(\lambda)$ on behavior, we see a clear pattern in Table 4. Again, focusing on column (1), the point estimate of β_2 is equal to -0.50 with a standard error of 1.52. This small and statistically insignificant estimate implies that during feedback phases, the habit stock $A(\lambda)$ has no impact on behavior. Visually, this can be clearly seen in panel (b) of Figure 7 as the predicted consumption profile from specification (2) in a 48/72 feedback on/off cycle is flat during the feedback-on phase.

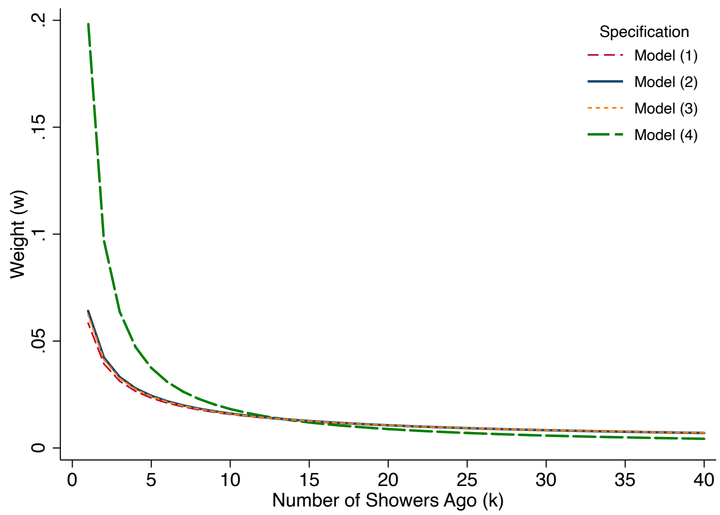
By contrast, the point estimate of β_4 is -10.91 in column (1) of Table 4 is negative and significant. Thus, the impact of previous feedback as summarized in $A(\lambda)$ has a strong and highly significant impact on behavior during off-phases that gradually declines over time. Again, this empirical result is clear from the predicted consumption profiles during the feedback-off phase in panel (b) of Figure 7.

Testing our hypotheses from the column (1) estimates, we clearly reject **H1** (consumption habit) with $p < 0.01$. The reason for this is that β_2 is not significantly different from zero, and is much smaller than β_4. In words, the asymmetry in feedback effect build-up and decay, as illustrated in Figure 7, leads us to reject the consumption habit model.

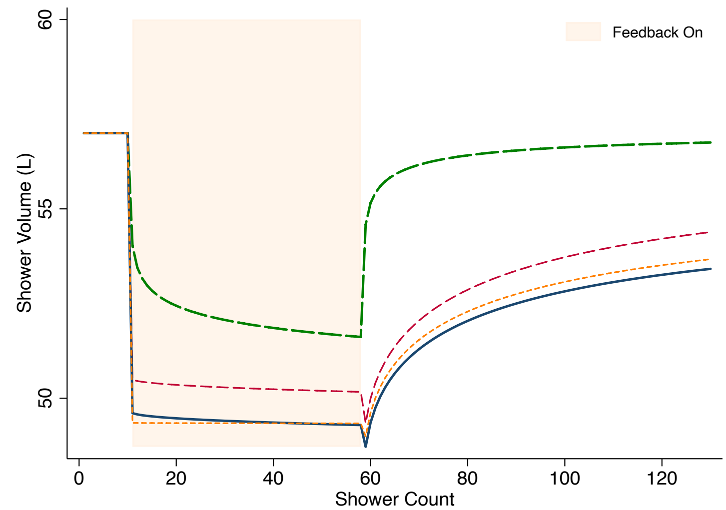
In contrast, our point estimates and their statistical significance in column (1), collectively, directly align with the predictions under **H2** (attention habit). Specifically, in-line with **H2**, we obtain a large and statistically significant negative β_4 coefficient and small-magnitude coefficient estimates for β_2 and β_3 that are statistically indistinguishable from 0. Visually, the tight correspondence between the attention-habit

Figure 7: Estimated Weighting Function and Model Predictions

(a) Weighting Function $w(k,\lambda)$



(b) Predictions from a 48/72 Feedback On/Off Cycle



theoretical predictions from Figure 1 and empirical results in panel (b) of 7 underscore how our experimental results clearly favor the attention-habit model over the consumption-habit model.

Additional results

The column (1) estimates yield little evidence of permanent learning: the point estimate of β_3 is small, and not significantly different from zero. Thus, conditional on the post-feedback effects captured by $A(\lambda)$, there is no evidence of other forms of persistence. As a further robustness check, the specification in column (2) of the table includes an interaction term $ON_{is} \times FIRST_{is}$ in the model for the first feedback episode. As explained above, this is to capture learning more fully. We find that the coefficient estimate on this term is also small and statistically indistinguishable from zero. Thus, individuals react to feedback in the first cycle exactly the same as in subsequent feedback cycle, in clear contradiction to the learning channel for persistence effects.

Column (3) of Table 4 examines the sensitivity of the parameter estimates by excluding the variables related to the learning channel. As can be seen, the point estimates remain virtually the same as those in the first two columns. Quantitatively, we can see in Figure 7 that the weighting functions and predicted consumption paths of a 48/72 feedback on/off cycle are very similar from the model specifications in columns (1) to (3) of Table 4. They all yield the same qualitative patterns that we observed from our model-free time-varying treatment effect plots in Figure 5.

Finally, column (4) of Table 4 imposes the structure of the Stigler and Becker (1977) consumption-habit model. In particular, it imposes the constraint $\beta_2 = \beta_4$ by including $A_{is}(\lambda)$ as an explanatory variable without any interactions. As can be seen from the estimates, the constraints do not fit the data well. Because the feedback effects in the data are very stable, the model needs to strike a compromise between the persistence observed in the feedback-off phases, and the stable behavior in the feedback-on phases. This leads to a much smaller estimate of the feedback effect, and a lower estimate of the persistence parameters. The model settles on a smaller (more negative) λ, but with a considerable loss in precision, with the standard error more than doubling. The poor fit can also be seen by plotting the predicted values for treatment T3 in Figure 7. The model fails to generate a stable feedback effect we see in Figure 5 above, and misses much of the persistence phase.

6 Conclusion

In this paper, we present the first evidence from a field experiment to examine the mechanisms underlying persistence in behavior. Over seven experimental conditions, we vary the intensity of frequency of feedback on showering, using a smart shower meter. Together with granular behavioral data, this allows us to examine mechanisms underlying behavioral persistence in the most detailed way to date.

We find strong evidence of persistence of the feedback intervention. Our most intense feedback treatment, providing feedback over 1.5 months, lead to measurable persistence in behavior for at least two months. However, even feedback episodes as short as three days of feedback induce significant, albeit weaker, persistence.

We test the predictions of two competing models of persistence. The Stigler and Becker (1977) model explains persistence as a complementarity between a stock of past consumption and the current consumption. It broadly predicts sluggish adjustment to changes in the environment, as the changes in the habit stock induced by the initial behavioral change gradually affect behavior. We contrast this with a model in which persistence is due to habit formation in attention (Anderson, 2016; Jiang and Sisk, 2019). In this model, feedback removes inattention while applied, but also builds up an attention habit stock that leads to persistence.

We find that there is and immediate and stable change in behavior when individuals receive feedback, but a gradual drift back to baseline behavior when feedback is turned off. These results clearly favor the attention-habit model in, both, the reduced-form evidence as well as a structural model. In fact, our structural model, borrowing functional form from Malmendier and Nagel (2011), fully captures the extent of persistence in the data, leaving little form for other mechanisms to explain our findings.

Distinguishing mechanisms of persistence is important for policy, as optimal policies may differ between mechanisms. Suppose the policy goal is to reduce a behavior, but that the policy maker is facing a tradeoff between coverage of individuals vs. intensity of the treatment for a treated individual. In many prominent formulations of the consumption-habit model (e.g., in Becker et al., 1991; Becker and Murphy, 1988), as well as in ours, the impact of the habit stock on the marginal utility is linear. Hence, to a first approximation, there is no advantage from either concentrating or spreading out the intervention. By contrast, the attention-habit model runs into diminishing returns as attention approaches its maximal level. Under this model, there exists an incentive for broader, but less intensive coverage of interventions.[19]

[19]The mechanisms also differ in their impact on consumer surplus, as pointed out by Aronsson and Löfgren (2008): in a

The distinction may be particularly relevant for interventions through the Internet of Things. With increased availability of feedback, our results raise the possibility that these interventions lead to persistent changes in behavior, even after relatively short exposure. However, attention habits are likely not confined to feedback interventions. The research in neurospychology emphasizes the role of incentives in shaping attention (Anderson et al., 2011; Anderson, 2016). Thus, temporary changes in prices, such as sales, may have a persistent effect on demand though changes in the attention stock. Future research should explore these questions further.

consumption-habit model, the persistence effects are welfare-neutral, since they are fully anticipated. By contrast, in an attention-habit model, they have first-order effects on consumer surplus.

References

Acland, Dan and Matthew R Levy, “Naiveté, projection bias, and habit formation in gym attendance,” *Management Science*, 2015, *61* (1), 146–160.

Agarwal, Sumit, Ximeng Fang, Lorenz Goette, Tien Foo Sing, Thorsten Staake, Verena Tiefenbeck, and Davin Wang, “The Role of Goals and Real-Time Feedback in Resource Conservation: Evidence from a Large-Scale Field Experiment,” Technical Report, National University of Singapore 2018.

Allcott, Hunt and Todd Rogers, “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation,” *American Economic Review*, 2014, *104* (10), 3003–3037.

Anderson, Brian A, “The attention habit: How reward learning shapes attentional selection,” *Annals of the new York Academy of Sciences*, 2016, *1369* (1), 24–39.

_ , Patryk A Laurent, and Steven Yantis, “Value-driven attentional capture,” *Proceedings of the National Academy of Sciences*, 2011, *108* (25), 10367–10371.

Andrews, Donald, “Tests for Parameter Instability and Structural Change with Unknown Change Point,” *Econometrica*, 1993, *61* (4), 821 – 856.

Aronsson, Thomas and Karl-Gustaf Löfgren, “Welfare equivalent NNP and habit formation,” *Economics Letters*, 2008, *98* (1), 84–88.

Attari, Shahzeen Z, “Perceptions of water use,” *Proceedings of the National Academy of Sciences*, 2014, *111* (14), 5129–5134.

_ , Michael L DeKay, Cliff I Davidson, and Wändi Bruine De Bruin, “Public perceptions of energy consumption and savings,” *Proceedings of the National Academy of sciences*, 2010, *107* (37), 16054–16059.

Auld, M. Christopher and Paul Grootendorst, “An Empirical Analysis of Milk Addiction,” *Journal of Health Economics*, 2004, *26* (3), 1117–1133.

Becker, Gary S and Kevin M Murphy, “A theory of rational addiction,” *Journal of political Economy*, 1988, *96* (4), 675–700.

Becker, Gary S., Michael Grossman, and Kevin M. Murphy, “Rational addiction and the effect of price on consumption,” *The American Economic Review*, 1991, *81* (2), 237–241.

Berkouwer, Susanna B and Joshua T Dean, “Credit and attention in the adoption of profitable energy efficient technologies in Kenya,” 2019.

Blake, Thomas, Sarah Moshary, Kane Sweeney, and Steven Tadelis, “Price salience and product choice,” Technical Report, National Bureau of Economic Research 2018.

Camerer, Colin F., Peter Landry, and Ryan Webb, "The Neuroeconomics of Habit," in Alan Kirman and M. Tedeschi, eds., *The State of Mind in Economics*, Oxford: Oxford University Press, 2020. forthcoming.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustere Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414 – 427.

Charness, Gary and Uri Gneezy, "Incentives to exercise," *Econometrica*, 2009, *77* (3), 909–931.

Chetty, Raj, "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

_ , Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *The American Economic Review*, 2009, *99* (4), 1145–1177.

Della Vigna, Stefano, "Psychology and Economics : Evidence from the Field," *Journal of Economic Literature*, 2009, *472*, 315–372.

Enke, Benjamin and Thomas Graeber, "Cognitive Uncertainty," Technical Report 26518, National Bureau of Economic Research 2019.

Gabaix, Xavier, "Behavioral inattention," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 261–343.

Harrison, Glenn W. and John A. List, "Field Experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

Hussam, Reshmaan, Atonus Rabbani, Giovanni Reggiani, and Natalia Rigol, "Habit Formation and Rational Addiction: A Field Experiment in Handwashing," 2017. HBS Working Paper #18-030.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka, "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand," *American Economic Journal: Economic Policy*, 2018, *10* (1), 3003–3037.

Jessoe, Katrina and David Rapson, "Knowledge is (less) power: Experimental evidence from residential energy use," *American Economic Review*, 2014, *104* (4), 1417–38.

Jiang, Yuhong V and Caitlin A Sisk, "Habit-like attention," *Current opinion in psychology*, 2019, *29*, 65–70.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems, "The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network," *Quarterly Journal of Economics*, 2017, *132* (4), 2019–2055.

Loewenstein, George, Joseph Price, and Kevin Volpp, "Habit Formation in Children: Evidence from Incentives for Health Eating," *Journal of Health Economics*, 2016, *45* (1), 47–54.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *Quarterly Journal of Economics*, 2011, *126* (1), 373 –416.

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Stigler, George J and Gary S Becker, "De Gustibus Non Est Disputandum," *American Economic Review*, 1977, *67* (2), 76–90.

Tiefenbeck, Verena, Anselma Wörner, Samuel Schöb, Elgar Fleisch, and Thorsten Staake, "Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives," *Nature Energy*, 2019, *4* (1), 35–41.

_ , Lorenz Goette, Kathrin Degen, Vojkan Tasic, Elgar Fleisch, Rafael Lalive, and Thorsten Staake, "Overcoming Salience Bias: How Real-Time Feedback Foster Resource Conservatioj," *Management Science*, 2018, *64* (3), 1458–1476.

Wood, Wendy and David T Neal, "A new look at habits and the habit-goal interface.," *Psychological review*, 2007, *114* (4), 843.

_ and Dennis Runger, "Psychology of Habit," *Annual Review of Psychology*, 2016, *67*, 289 – 314.

Yang, Nan and Long Yong Lim, "Temporary Incentives Change Daily Routines: Evidence from a Field Experiment on Singapore's Subways," *Management Science*, 2017.

Appendix

A Transitions Between Steady States

A.1 Transitions predicted by the habit stock model

To compute these transitions, we assume away transitions in salience levels and simply fix $\theta_t = \theta$ if FB_t is on, and θ_t=1 if FB_t is off. This environment is simply the Becker and Murphy (1988) habit stock model, combined with the (non-dynamic) price salience model of Chetty et al. (2009).

Consumption habit stock h_{t+1} is given by

$$h_{t+1} = (1-\delta)c_t + \delta h_t,$$

so Δh_{t+1} is given by

$$\Delta h_{t+1} = (1-\delta)\Delta c_t + \delta \Delta h_t. \tag{13}$$

Also, the general form for optimal consumption choice $c\star_t$ is

$$c_t\star = \frac{a}{b} + \frac{\gamma}{b}h_t - \frac{p}{b}\theta_t,$$

so the general form for Δc_{t+1} is

$$\Delta c_{t+1} = \frac{\gamma}{b}\Delta h_{t+1} - \frac{p}{b}\Delta\theta_{t+1}. \tag{14}$$

A.1.1 Jump in c_t when FB is first turned on and we initially depart from the noFB steady state

Suppose at $t = 0$ we are at noFB steady state, i.e., $\theta_0 = w_0 = \theta$ and $h_0 = c_0$. Suppose then that at $t = 1$, we turn FB on. Then by (13),

$$\Delta h_1 = (1-\delta)0 + \delta 0 = 0 \tag{15}$$

and θ_1 changes from θ to 1. By (14),

$$\Delta c_1 = 0 - \frac{p}{b}(1-\theta) = \frac{\theta - 1}{b}p. \tag{16}$$

A.1.2 Path of c_t when FB is left on and we converge to the FB steady state

Suppose we leave FB on from $t = 2$ to $t = T$, at which point the FB steady state is reached. As FB is on from period 1 onwards and $\theta_t = 1$ so long as FB is on, $\Delta\theta_t = 0$ for all $2 \le t \le T$. So by (14), Δc_t is driven only by Δh_t. So we will first focus on the transition dynamics of Δh_t.

When $t = 2$, $\Delta h_2 = (1-\delta)\Delta c_1$ by (13) and (15), and $\Delta c_2 = \frac{\gamma}{b}\Delta h_2$ by (14).

When $t = 3$, $\Delta h_3 = (1-\delta)\Delta c_2 + \delta\Delta h_2 = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]\Delta c_1$, and $\Delta c_3 = \frac{\gamma}{b}\Delta h_3$.

When $t = 4$, $\Delta h_4 = (1-\delta)\Delta c_3 + \delta\Delta h_3 = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^2\Delta c_1$, and $\Delta c_4 = \frac{\gamma}{b}\Delta h_4$.

We thus show by induction that for $2 \le t \le T$,

$$\Delta h_t = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^{t-2} \Delta c_1. \tag{17}$$

By (13) and (15), $\Delta h_2 = (1-\delta)\Delta c_1$, which satisfies (17). Now suppose

$$\Delta h_k = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^{k-2} \Delta c_1,$$

$2 \le k < T$. By (14), $\Delta c_k = \frac{\gamma}{b}\Delta h_k$. Then

$$\begin{aligned}
\Delta h_{k+1} &= (1-\delta)\Delta c_k + \delta \Delta h_k \\
&= \left[\delta + \frac{\gamma}{b}(1-\delta)\right] \Delta h_k \\
&= (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^{(k+1)-2} \Delta c_1
\end{aligned}$$

as desired. By (14), $\Delta c_t = \frac{\gamma}{b}\Delta h_t$.

A.1.3 Jump in c_t when FB is first turned off and we initially leave the FB steady state

Suppose instead that at $t = 0$ we are at FB steady state, i.e. $\theta_0 = w_0 = 1$ and $h_0 = c_0$. Suppose then that at $t = 1$, we turn FB off. Then by (13),

$$\Delta h_1 = (1-\delta)0 + \delta 0 = 0. \tag{18}$$

Also, $w_1 = \theta\alpha + (1-\alpha)\theta = \theta$, implying that θ_1 changes from 1 to θ. By (14),

$$\Delta c_1 = 0 - \frac{p}{b}(\theta - 1) = \frac{1-\theta}{b} p. \tag{19}$$

A.1.4 Path of c_t when FB is left off and we converge back to the noFB steady state

Suppose we leave FB off from $t = 2$ to $t = T$, at which point the noFB steady state is reached. As $\theta_t = \theta$ so long as FB is off, $\Delta\theta_t = 0$ for all $2 \le t \le T$. By (13) and (18), $\Delta h_2 = (1-\delta)\Delta c_1$, which satisifies (17). Through an identical induction step, we find that Δh_t follows the same transition path described in (5). Again by (14), $\Delta c_t = \frac{\gamma}{b}\Delta h_t$.

A.1.5 The magnitude of the jump in A.1.1 equals that of the jump in A.1.3

By (16) the magnitude of the first jump is $\frac{\theta-1}{b}p$. By (19), the magnitude of the second jump is $\frac{1-\theta}{b}p$. Hence the two jumps have equal magnitudes in opposite directions.

A.1.6 The transition path in A.1.2 and the transition path in A.1.4 are symmetric

We see in A.1.2 and A.1.4 that the transition path of consumption habit stock is described by (17):

$$\Delta h_t = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^{t-2}\Delta c_1$$

whether the transition is from noFB to FB steady state or vice versa. By A.1.5, Δc_1 has a different sign depending on the direction of the transition, but the same magnitude either way. Hence the two transition paths are symmetric.

A.2 Infinite sum of Δc_t

The transition path of Δc_t is described by a geometric sequence that takes the form xr^k, where $x = \frac{\gamma}{b}(1-\delta)\Delta c_1, r = \delta + \frac{\gamma}{b}(1-\delta)$, and $k = t-2$. The general formula for an infinite sum of a geometric sequence is

$$\sum_{k=0}^{\infty} xr^k = \frac{x}{1-r}.$$

Hence we can express the infinite sum of Δc_t from $t = 2$ to infinity as:

$$\begin{aligned}\sum_{t=2}^{\infty} \Delta c_t &= \frac{\frac{\gamma}{b}(1-\delta)}{1-\delta-\frac{\gamma}{b}(1-\delta)}\Delta c_1 \\ &= \frac{\frac{\gamma}{b}}{1-\frac{\gamma}{b}}\Delta c_1 \\ &= \frac{\gamma}{b-\gamma}\Delta c_1.\end{aligned}$$

We can then add Δc_1 to obtain the infinite sum from $t = 1$ onwards. Suppose we are transitioning from noFB to FB, i.e., Δc_1 is given by (16):

$$\begin{aligned}\sum_{t=1}^{\infty} \Delta c_t &= \frac{\gamma}{b-\gamma}\Delta c_1 + \Delta c_1 \\ &= \left(1+\frac{\gamma}{b-\gamma}\right)\frac{\theta-1}{b}p \\ &= \frac{\theta p - p}{b-\gamma},\end{aligned}$$

which when added to $c\star_{noFB} = \frac{a-\theta p}{b-\gamma}$ yields $c\star_{FB} = \frac{a-p}{b-\gamma}$. Similarly, adding the sum of Δc_t in the opposite direction to the FB steady state indeed yields the noFB steady state.

A.3 Finite sum of Δc_t

The general formula for a finite sum of a geometric sequence is

$$\sum_{k=0}^{n} xr^k = (1-r^n)\frac{x}{1-r} = (1-r^n)\sum_{k}^{\infty} xr^k.$$

Hence we can derive the finite sum of Δc_t over $t = 1 \ldots T$ from the infinite sum in A.2:

$$\sum_{t=1}^{T} \Delta c_t = \left(1 - \left(\delta + \frac{\gamma}{b}(1-\delta)\right)^{T-2}\right)\frac{\gamma}{b-\gamma}\Delta c_1 + \Delta c_1.$$

A.4 Transitions predicted by the attention stock model

In the attention stock model, we shut down the influence of consumption habit stock by setting $\gamma = 0$. Thus, by (14),

$$\Delta c_{t+1} = -\frac{p}{b}\Delta\theta_{t+1}. \tag{20}$$

A.4.1 Jump in c_t when FB is first turned on and we initially depart from the noFB steady state

Suppose at $t = 0$ we are at noFB steady state, i.e., $\theta_0 = w_0 = \theta$. Suppose that at $t = 1$ we turn FB on. Then θ_1 changes from θ to 1. By (20),

$$\Delta c_1 = -\frac{p}{b}(1-\theta) = \frac{\theta-1}{b}p.$$

Indeed,

$$\Delta c_1 + c\star_{noFB} = \frac{a-\theta p}{b} + \frac{\theta-1}{b}p = \frac{a-p}{b} = c\star_{FB},$$

implying that c_t reaches its new steady state in one time period, without any further transition path.

A.4.2 No jump in c_t when FB is first turned off and we initially leave the FB steady state

Suppose instead that at $t = 0$ we are at FB steady state, i.e., $\theta_0 = w_0 = 1$. Suppose that at $t = 1$ we turn FB off. Then θ_1 changes from 1 to w_1. When FB is off, w_t is defined recursively as

$$w_t = \theta\alpha + (1-\alpha)w_{t-1}$$

and $\theta_t = w_t$. This implies that θ_t, and by extension c_t, transitions smoothly away from FB steady state without any discrete jump.

A.4.3 Path of c_t when FB is left off and we converge to the noFB steady state

Suppose we leave FB off from $t = 2$ to $t = T$, at which point the noFB steady state is reached. Recall that $\theta_t = w_t$ so long as FB is off. At $t = 1$, $\theta_1 = \theta\alpha + (1-\alpha)$. At $t = 2$, $\theta_2 = \theta\alpha + (1-\alpha)[\theta\alpha + (1-\alpha)] = \theta + (1-\theta)(1-\alpha)^2$. Similarly, at $t = 3$, $\theta_3 = \theta + (1-\theta)(1-\alpha)^3$. Thus we show by induction that for $1 \leq t \leq T$,

$$\theta_t = \theta + (1-\theta)(1-\alpha)^t. \tag{21}$$

As shown above, $\theta_1 = \theta\alpha + (1-\alpha)$, satisfying (21). Now suppose

$$\theta_k = \theta + (1-\theta)(1-\alpha)^k,$$

$1 \leq k < T$. Then

$$\begin{aligned}
\theta_{k+1} &= \theta\alpha + (1-\alpha)\theta_k \\
&= \theta\alpha + (1-\alpha)\left[\theta + (1-\theta)(1-\alpha)^k\right] \\
&= \theta\alpha + \theta + (1-\theta)(1-\alpha)^k - \theta\alpha - \alpha(1-\theta)(1-\alpha)^k \\
&= \theta + (1-\theta)(1-\alpha)^{k+1}
\end{aligned}$$

as desired. With this general form for θ_t, we can express

$$\begin{aligned}
\Delta\theta_t &= \theta + (1-\theta)(1-\alpha)^t - \theta - (1-\theta)(1-\alpha)^{t-1} \\
&= -a(1-\theta)(1-\alpha)^{t-1}
\end{aligned}$$

for $t \geq 1$. By (14), $\Delta c_t = -\frac{p}{b}\Delta\theta_t$.

A.4.4 Infinite sum of Δc_t toward noFB

The transition path of Δc_t from the FB to noFB steady state is described by a geometric sequence that takes the form xr^k, where $x = \frac{ap}{b}(1-\theta), r = 1-\alpha$, and $k = t$. Again, the general formula for an infinite sum of a geometric sequence is

$$\sum_{k=0}^{\infty} xr^k = \frac{x}{1-r}.$$

Hence we can express the infinite sum of Δc_t from $t = 1$ to infinity as:

$$\begin{aligned}
\sum_{t=1}^{\infty} \Delta c_t &= \frac{\frac{ap}{b}(1-\theta)}{1-1+\alpha} \\
&= \frac{p}{b}(1-\theta),
\end{aligned}$$

which when added to $c\star_{FB} = \frac{a-p}{b}$ yields $c\star_{noFB} = \frac{a-\theta p}{b}$.

A.4.5 Finite sum of Δc_t toward noFB

Again, the general formula for a finite sum of a geometric sequence is

$$\sum_{k=0}^{n} xr^k = (1-r^n)\frac{x}{1-r} = (1-r^n)\sum_{k}^{\infty} xr^k.$$

Hence we can derive the finite sum of Δc_t over $t = 1 \ldots T$ from the infinite sum in A.2:

$$\sum_{t=1}^{T} \Delta c_t = \frac{\alpha^T p}{b}(1-\theta).$$

B Analytic expression for the accumulation of w_t

These are more general expressions of the accumulation of w_t from any starting point (not necessarily a steady state) when FB is turned on or off.

B.1 When FB is Turned On

When FB is on, $w_t = \alpha + (1-\alpha)w_{t-1}$. Suppose we turn FB on at time $t = 0$, at which point w takes an initial value of w_0. Without loss of generality, let $w_t = u_t + k$, where u_t has some initial value u_0, and k is the same in every time period t. Then

$$\begin{aligned} w_t &= \alpha + (1-\alpha)w_{t-1} \\ \implies u_t + k &= \alpha + (1-\alpha)u_{t-1} + (1-\alpha)k. \end{aligned}$$

Let us set $k = \alpha + (1-\alpha)k$, such that $k = 1$. We can then write

$$\begin{aligned} u_t + 1 &= \alpha + (1-\alpha)u_{t-1} + 1 - \alpha \\ \implies u_t &= (1-\alpha)u_{t-1} \\ \implies u_t &= (1-\alpha)^t u_0. \end{aligned}$$

It follows that $w_t = u_t + k = (1-\alpha)^t u_0 + 1$. And since by definition $u_0 = w_0 - 1$, we can rewrite

$$w_t = (1-\alpha)^t (w_0 - 1) + 1,$$

which guarantees that as $t \to \infty$, w_t goes to its FB steady-state level of 1.

B.2 When FB is Turned Off

When FB is off, $w_t = \theta\alpha + (1-\alpha)w_{t-1}$. The proof proceeds similarly to the previous case. Suppose we turn FB off at time $t = 0$, at which point w takes on an initial value of w_0. As before, let $w_t = u_t + k$, where

u_t has some initial value u_0, and k is the same in every time period t. Then

$$\begin{aligned} w_t &= \theta\alpha + (1-\alpha)w_{t-1} \\ \implies u_t + k &= \theta\alpha + (1-\alpha)u_{t-1} + (1-\alpha)k. \end{aligned}$$

Let us set $k = \theta\alpha + (1-\alpha)k$, such that $k = \theta$. We can then write

$$\begin{aligned} u_t + \theta &= \alpha + (1-\alpha)u_{t-1} + \theta - \alpha \\ \implies u_t &= (1-\alpha)^t u_0 \\ \implies w_t &= (1-\alpha)^t (w_0 - \theta) + \theta, \end{aligned}$$

which guarantees that as $t \to \infty$, w_t goes to its noFB steady-state level of θ.

B.3 Turning feedback on and off

These non-recursive expressions for simplify the process of solving for w_t after alternating periods of FB being on and off. Let w be in the noFB steady state at $t = 0$, i.e., $w_0 = \theta$. Suppose we turn feedback on for j periods, then turn feedback off for k periods. Then in time period $t = j$,

$$w_j = (1-\alpha)^j (\theta - 1) + 1.$$

After feedback is switched off for k more periods, w_{j+k} can be expressed as

$$\begin{aligned} w_{j+k} &= (1-\alpha)^k [w_j - \theta] + \theta \\ &= (1-\alpha)^k \left[(1-\alpha)^j (\theta - 1) + 1 - \theta\right] + \theta \\ &= \left[(1-\alpha)^{j+k} - (1-\alpha)^k\right] (\theta - 1) + \theta. \end{aligned}$$

C Household survey for sample recruitment

Survey Invitation Email

Subject line: Your next water bill may be on us

At South East Water, we're looking for new ways to help you better manage your water usage. To do this, it would be useful to know a little more about your household's water usage and lifestyle.

Please take a few moments to complete our short survey with 25 multiple-choice questions. As a thank you, you will be automatically entered into the draw for a chance to win one of the following prizes:

- One prize of $1000 off your next water bills,
- One iPad valued approximately at $1000, and
- One prize of $1000 to be donated to a choice of charities in your name.

"Complete Survey" button here

Link to Terms and Conditions included at the bottom of email.

Survey Questions

Suggested wording for the top of the survey
Thank you for taking the time to complete our survey. Your answers will help shape the development of tools and resources to better support customers.

Customer Questions

1. How many people live in your home?
   [1, 2, 3, 4, 5, 6, 7, 8, 9+]

2. How many household members are babies or toddlers (under age 5)?
   [0, 1, 2, 3, 4+]

3. How many household members are children between the ages of 5 and 12?
   [0, 1, 2, 3, 4+]

4. How many household members are teenagers (ages 13-19)?
   [0, 1, 2, 3, 4+]

5. How many showers are in your home?
   [1, 2, 3, 4+]

6. What best describes the shower that you use most of the time?
   [Hand-held shower, Wall or overhead shower, Combination]

7. What best describes the showerhead that you use most of the time?
   [Low-flow or restricted-flow, Power or high-pressure, Traditional, Don't know]

8. How many minutes long is a typical shower in your home?
   [Less than 4, 5-7, 8-9, 10-11, 12-13, 14-15, 16+]

9. What is your best guess of much water is used (in litres) during a typical shower in your home?
   [less than 10 L, 10-25 L, 25-50 L, 50-75 L, 75-100 L, 100-125 L, 125-150 L, more than 150 L]

10. How do you heat your hot water?
    [Electricity, Gas, Don't know]

11. Are any of the toilets in your house dual-flush?
    [Yes, No, Don't Know]

12. How often do you run the dishwasher with a less-than-full load?
    [Every time, Often, Occasionally, Never, I don't have a dishwasher]

13. How often do you run your clothes washing machine with a less-than-full load?
    [Every time, Often, Occasionally, Never, I don't have a washing machine]

14. What best describes your clothes washing machine?
    [Top-loading, front-loading, I don't have a washing machine]

15. How long has it been since someone has checked for leaking taps or toilets in your home?
    [We have never checked, Several years, Several months, Several days]

16. In your home, how much time typically goes by between first noticing and then fixing a leaking tap or toilet?
    [I have never had a leaky faucet or toilet, Several hours, Several days, Several weeks, Several months, Several years]

17. How long has it been since the last major remodel of your home?
    [My home is brand new, 2-5 years, 5-10 years, 10-15 years, 15+ years or never been remodeled, I don't know]

18. Which of the following do you have? [You can tick more than one answer.]
    Balcony garden
    Lawn grass
    Vegetable garden
    Only native or drought-tolerant plants
    Rainwater tank
    Drip irrigation system
    Swimming pool
    Spa pool

19. How many minutes a week do you water your plants or garden in the summer?
    [0, 1-10, 10-15, 15-20, 20-30, 30+, Does Not Apply]

20. How do you usually wash your car in the summer?
    [I don't own a car, At a paid/commercial car wash, At home with a hose, At home with only a bucket, Other]

21. Compared to water usage in homes with the same number of people as yours, what statement best describes your household's water use?

[High (top 20%), Above average (top 40%), Average, Below Average (bottom 40%), Low (bottom 20%)]

22. What do you expect your next quarterly water bill to be?
    [$25, $50, $75, $100, $125, $150, $175, $200, $225, $250, $275, $300, $325, $350, $375, $400, $425, $450, $475, $500]

23. Have you had any unexpectedly high water bills in the past year?
    [Yes, No]

24. If yes, on average how much higher were the water bills than what you expected?
    [$25, $50, $75, $100, $125, $150, $175, $200, more than $200, Does not apply]

25. If yes, can you recall which bills were unexpectedly high? Please check all that apply.
    [2015: Jul, Aug, Sep, Oct, Nov, Dec, 2016: Jan, Feb, Mar, Apr, May, Jun, Does not apply]

Environment/Health/Social Donation Question

Thank you for completing our survey. You are now in the draw to win the chance to donate $1000 in your name to a selection of charities. We have one donation prize to give away. Please indicate how you would like to split the money amongst the following charities should you win the prize:

Australian Red Cross [$0, $250, $500, $750, $1000]

World Wildlife Foundation [$0, $250, $500, $750, $1000]

National Breast Cancer Foundation [$0, $250, $500, $750, $1000]

Starlight Children’s Foundation [$0, $250, $500, $750, $1000]

Message once survey is completed

Thank you for completing our survey. You are now in the draw. The draw will take place on XX date. Winners will be notified via email by YY date.

D Testing for habit-as-automatic-control

Exploiting the richenss of our data and within-subject experimental design, we directly look for discrete jumps in households' consumption after feedback is turned off, and see whether such jumps explain post–feedback treatment effect decay. Specifically, we augment our baseline regression equation (9) as follows:

$$y_{it} = \alpha_i + \beta_1 ON_{it} + \beta_2 PostON_{it} + \beta_3 OFF_{it} + \beta_4 PostOFF_{it} + \sum_{i}^{K} \eta_i PostOFF_{it_i^*} + \alpha_i + \delta_t + \varepsilon_{it} \quad (22)$$

Through the inclusion of the new $PostOFF_{it_i^*}$ regressors, we estimate a household–specific post–feedback jump in water usage η_i, and take a data-driven approach to identify when a household's post–feedback shower the jump occurs, which we denote t_i^*. We iteratively construct the $PostOFF_{it_i^*}$ regressors and estimate η_i and t_i^* for all households as follows:

1. Initial (22) for household i by setting $PostOFF_{it_i^*} = PostOFF_{it} \times \alpha_i$, where recall α_i is a household i dummy variable/fixed effect. We construct these initialized household–specific $PostOFF_{it_i^*}$ variables for all K households in experimental conditions T3–T7 for whom we observe post–feedback showers.

2. With the initialized $PostOFF_{it_i^*}$ variables, run the regression in (22). In effect, this yields a distribution of household–specific post–feedback treatment effects $\hat{\eta}_1, \hat{\eta}_2, \ldots, \hat{\eta}_K$ above and beyond the common post–feedback treatment effect β_4. Without loss of generality, we enumerate $i = 1$ as the household with the lowest device number in T3, and $i = K$ as the household with the largest device number in T7.[20]

3. Iteratively test for household–specific post–feedback jumps in water usage. Starting with household $i = 1$, test the following hypothesis based on the regression results from step 2.:

$$H_0 : \eta_1 = 0 \text{ vs. } H_1 : \eta_1 \neq 0 \quad (23)$$

   Denote the F-statistic from this test as $F_{1,1}$, where the first "1" in the subscript corresponds to household $i = 1$, and the second "1" in the subscript corresponds to $\tau = 1$ showers since feedback was turned off for household 1.

4. Increment τ by 1 to $\tau = 2$, and update $PostOFF_{1t^*}$ such that it equals 0 if it has been less than $\tau = 2$ showers since feedback was turned off for household 1.

5. Run the regression in (22) again, test the hypothesis in (23), and denote the F-statistic for $i = 1$ and $\tau = 2$ from this test as $F_{1,2}$.

6. Iterate between steps 4 and 5 for household 1, each time incrementing τ by 1 and re-defining $PostOFF_{1t^*}$ such that it equals 0 if it has been less than τ showers since feedback was turned off for household 1. Denote the F-statistic for the hypothesis test in (23) at iteration $\tau = j$ by $F_{1,j}$.

[20] To avoid a perfect collinearity problem with $PostOFF_{it}$, we drop one of the households in conditions T3–T7 in estimating the coefficients in (22).

7. Find the value of j that corresponds to the maximum F-statistic from $F_{1,1}, F_{1,1}, \ldots, F_{1,J_1}$, where J_1 is the maximum number of consecutive post–feedback showers for household 1. Define t_i^* to be the shower that corresponds to this maximum F-statistic for household 1. This is our initial estimate of the timing of the post-feedback jump for household 1.

8. Move to household $i = 2$ in condition T3, and repeat steps 1–7, holding fixed t_i^* at their current values for all other households $i \neq 2$, including t_1^* at the value previously found from steps 1–7 for household 1.

   ...

   Repeat steps 1–8 for all households $k = 3, \ldots, K$ in conditions T3–T7, each time holding fixed t_i^* for all other households $m \neq k$.

Once we have looped through all K households in conditions T3–T7, we obtain estimates of the *timing* of the post–feedback jumps $t_1^*, t_2^*, \ldots, t_K^*$ and their *magnitudes* $\hat{\eta}_1, \hat{\eta}_2, \ldots, \hat{\eta}_K$.[21]

This iterative approach to computing F-statistics for each possible post–feedback break point t_i^* for each household corresponds to the Andrews (1993) supF test for finding structural breaks with an unknown break point. As with the supF test, we take a data-driven approach to finding the unknown break in consumption levels after feedback is turned off by searching over all possible post–feedback jumps, and finding the one that delivers the largest F-statistic from a test of a null that a break in the level of consumption exists at a τ showers after feedback is turned off versus the alternative that no break exists for that shower. The maximum of these F-statistics corresponds to the break point that best explains the timing and magnitude of the jump in a given household's consumption profile after feedback is turned off. Notably, the estimated jump may be positive, near–zero, or negative, whatever best fits the data.

D.1 Implementation

There are various practical considerations in implementing this search for household–specific post–feedback jumps. First, households must have sufficiently many observations without feedback to be included. We restrict households to having a minimum of 20 showers without feedback to implement our test, which leaves us with 489 of 555 total households from conditions T1–T7.

Second, in-line the suggestion of Andrews (1993), we check for t_i^* values for a given household up until the last 20% of observations during the post–fedback phase. That is, we search for t_i^* from showers $\tau = 1, \ldots, 0.8 \times J_i$ after feedback is turned off. This restriction ensures that we have sufficient data after a candidate t_i^* to stably implement the F-test for testing for structural break in the level of consumption at each candidate τ value.

Third, in principle, we could iterate on steps 1–8 again, starting back at household $i = 1$, holding fixed the break points found for all other households, and iterating through all households again to find their break points. In practice, however, when we iterate through all of the households a second time, we find virtually

[21] The latter coefficients are found by running the regression in (22) where each of the $PostOFF_{it_i^*}$ are defined according to the t_i^* values found by the routine described in steps 1 to 7.

no difference in our results, both in terms of our coefficients of interest β_4, and the timing and magnitude of consumption jumps across households. Therefore, our results below reflect just one iteration of steps 1–8 across all households.

Fourth, there are variations on the regression specification in (22) one might consider in estimating the timing and magnitude of households' post–feedback jumps. After presenting our main results, we discuss robustness checks where we vary the underlying regression specification.

Fifth, we use the cluster bootstrap from Cameron et al. (2008) to compute standard errors and confidence intervals that are clustered at the household level. To compute these, we construct bootstrap samples by randomly drawing households from our sample with replacement, and running steps 1-7 for each sample. Doing so yields bootstrap distributions of regression coefficients from (22) and household–specific jump timings $t_1^*, t_2^*, \ldots, t_K^*$ from which we compute standard errors and confidence intervals. In total, we use $B = 100$ bootstrap samples.

Finally, for the interested reader, Appendix ## provides detailed, visual examples from hand–selected households to illustrate how the routine described in steps 1–8 is effective in identifying the timing and magnitude of post–feedback jumps in consumption.

D.2 Results

Table D.1 and Figures D.1–D.3 presents our findings. The table is formatted as in Tables **??** and **??** above, with column (1) containing pooled results for all experimental conditions, while columns (2)–(7) contain results for experimental conditions T2–T7. The top panel of the table replicates our results from the bottom panel Table **??** based on the subsample of 489 of 555 household for which we estimate the regression in (22). Comparing these two sets of results, we obtain similar coefficient estimates for all coefficients in all colums. Most importantly, the coefficient estimates of interest for $PostOFF_{it}$ in Table D.1 are numerical identical to what we find in Table **??**. There is no evidence of sample–selection effects arising from our conditioning on households who have sufficient post–feedback data in testing for heterogeneous post–feedback jumps.

The bottom panel of Table D.1 presents our results from estimating (22). Comparing the top and bottom panels of the table yields our main result regarding the importance of post–feedback jumps: we obtain very similar magnitude estimates on the $PostOFF_{it}$ coefficients. We estimate a similar rate of decay in treatment effects when feedback is turned off, even if we allow household–specific post–feedback jumps of arbitrary timing and magnitude. In this way, the results in Table D.1 support our attention–based theory of habit in favor of a theory of habit–as–automatic–control.

In Figure D.1 we examine the distribution of post-feedback jumps magnitude and timing. Panel (a) plots the density of jump magnitudes for each experimental condition.[22] The distribution of jumps is centered around 0 for each condition, with non-negligible dispersion. While the majority of jumps are positive, we also find a non-negible share (35%) of jumps are negative. We find little systematic evidence of habit–as–

[22]Our within-subject design and sample size yields very precise jump estimates. Indeed, more than 90% of the estimated household–specific jumps are statistically significant at the 5% level, including very small–magnitude jumps. For this reason, we focus our analysis on post–feedback jump magnitudes and not their statistical significance.

Table D.1: Regression Results by Experimental Condition, Allowing for Household–specific Differences in the Magnitude and Timing of Jumps in Consumption When Feedback is Turned On and Off

	Experimental Conditions Included in the Sample						
	T1-T7	T1,T2	T1,T2,T3	T1,T2,T4	T1,T2,T5	T1,T2,T6	T1,T2,T7
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ON	-7.25***	-6.57***	-8.25***	-7.28***	-7.05***	-6.70***	-6.61***
	(0.72)	(1.39)	(1.10)	(1.05)	(1.06)	(1.07)	(1.10)
PostON	0.00	-0.02	-0.01	0.00	-0.01	-0.01	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
OFF	-5.07***		-8.71***	-5.47***	-5.13***	-5.49***	-3.75***
	(0.78)		(1.45)	(1.28)	(1.25)	(1.26)	(1.33)
PostOFF	0.08***		0.11***	0.06*	0.19***	0.19**	0.01
	(0.02)		(0.03)	(0.03)	(0.06)	(0.09)	(0.07)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.46	0.45
Observations	81280	25134	36899	38547	37216	35356	33798
ON	-7.30***	-6.57***	-8.09***	-7.11***	-6.92***	-6.93***	-6.68***
	(0.71)	(1.39)	(1.08)	(1.04)	(1.07)	(1.07)	(1.09)
PostON	-0.00	-0.02	-0.01	-0.00	-0.01	-0.00	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
OFF	-5.30***		-7.99***	-5.13***	-5.84***	-4.72***	-4.27***
	(0.79)		(1.29)	(1.18)	(1.22)	(1.25)	(1.34)
PostOFF	0.10***		0.08***	0.15***	0.11*	0.29***	0.11*
	(0.02)		(0.03)	(0.03)	(0.06)	(0.08)	(0.06)
R-Squared	0.44	0.44	0.42	0.45	0.43	0.46	0.45
Observations	81280	25134	36899	38547	37216	35356	33798

Notes: Dependent variable is shower water usage volume with baseline mean of 57 L (s.d.=42 L). See Figure 4 for the definitions of experimental conditions T1–T7. See the text for a description of the steps taken to estimate household–specific jumps in consumption when feedback is turned on and off. For brevity, we do not report household–specific ON_{it} and OFF_{it} coefficient estimates that we obtain. All regressions include household i and shower s fixed effects. Standard errors clustered at household level. P-values reported for the hypothesis test of ON=OFF and PostON=-PostOFF. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

automatic–control in the post–feedback phase at the household–level.[23]

Panel (b) of Figure D.1 describes the distribution of the timing of jumps. Specifically, the figure plots, by experimental condition, the survivor function for the event where a jump occurs. Across all conditions, a large 40% share of jumps occur quickly, within 10 showers after feedback is turned off. By 20 showers post–feedback, 50% or more of post-feedback jumps have occurred. Beyond shower 20, conditions T3 and T4 with longer feedback–off phases reveal significant heterogeneity in timing of the remaining jumps.

[23]This is not to say we find no evidence of delayed permanent upward shifts after feedback is removed. Indeed, we can find examples of shifts for a handful of households. For example, household 3507 in Appendix D exhibits a delayed post–feedback jump in consumption which is consistent with habit–as–automatic–control behavior. For completeness, Appendix D presents household–specific graphs of consumption profiles, highlighting where the supF test identifies the timing of the jump, as well as the jump's magnitude.

Figure D.1: Distributions of the Size and Timing of Post–Feedback Jumps in Water Usage

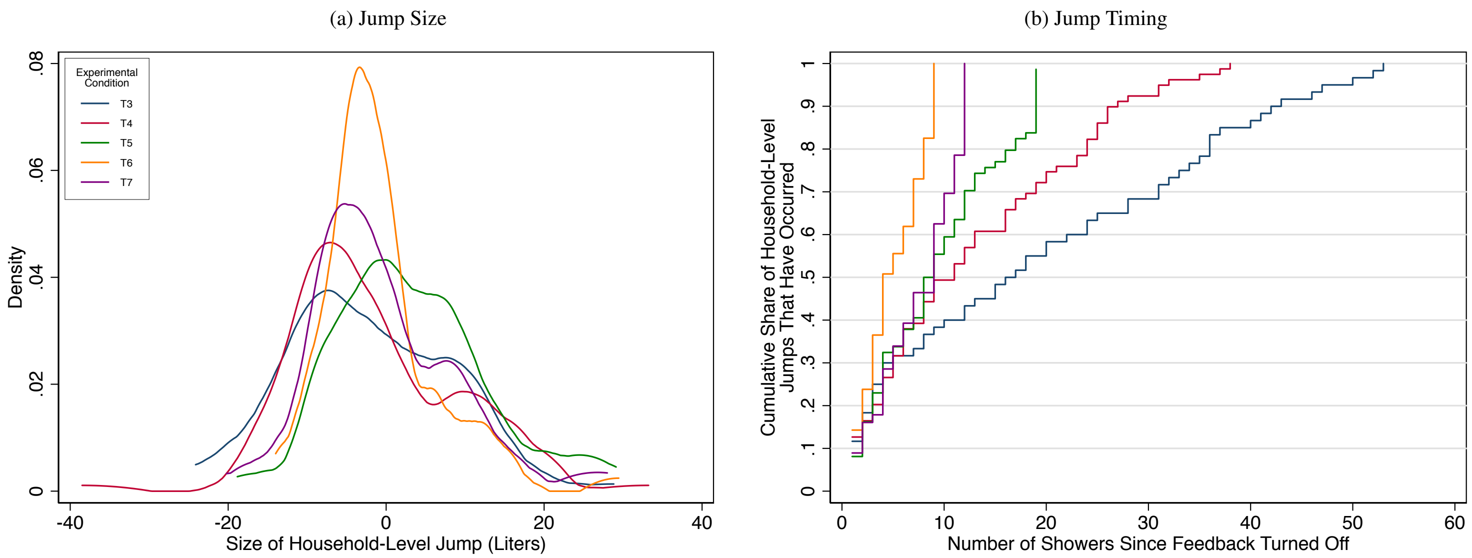
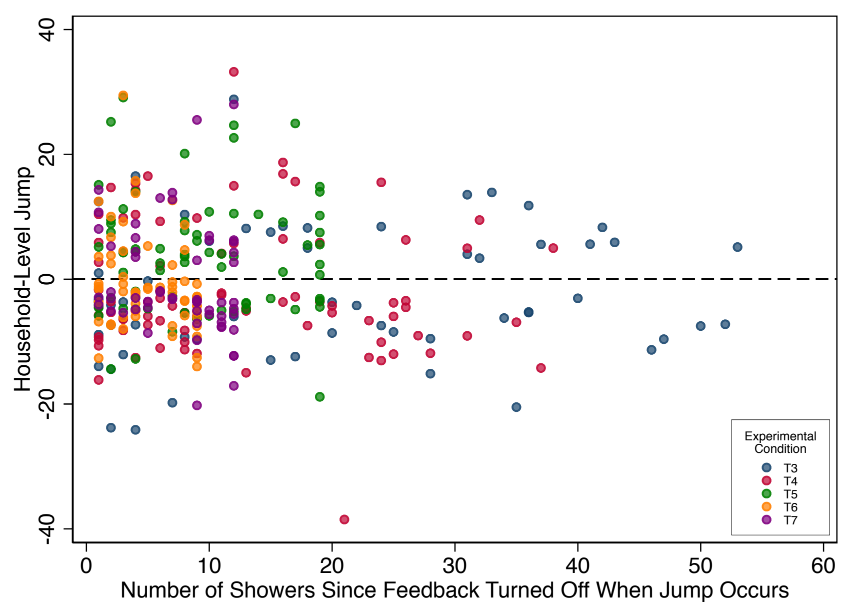


Figure D.2: Size vs. Timing of Post–Feedback Jumps in Water Usage



The heterogeneity in jumps' timing and magnitude are further highlighted in Figure D.2. Here, we provide a scatter plot of jump magnitude versus jump timing. As the figure shows, there's heterogeneous positive and negative jumps in consumption in the post–feedback phase in each experimental condition. We do not see that positive jumps tend to be more delayed than negative ones, which might otherwise suggest habit–as–automatic–control–type behavior tends to emerge over longer time horizons.

Figure D.3: Predictions of Consumption Responses to Feedback by Consumption and attention–habit Models

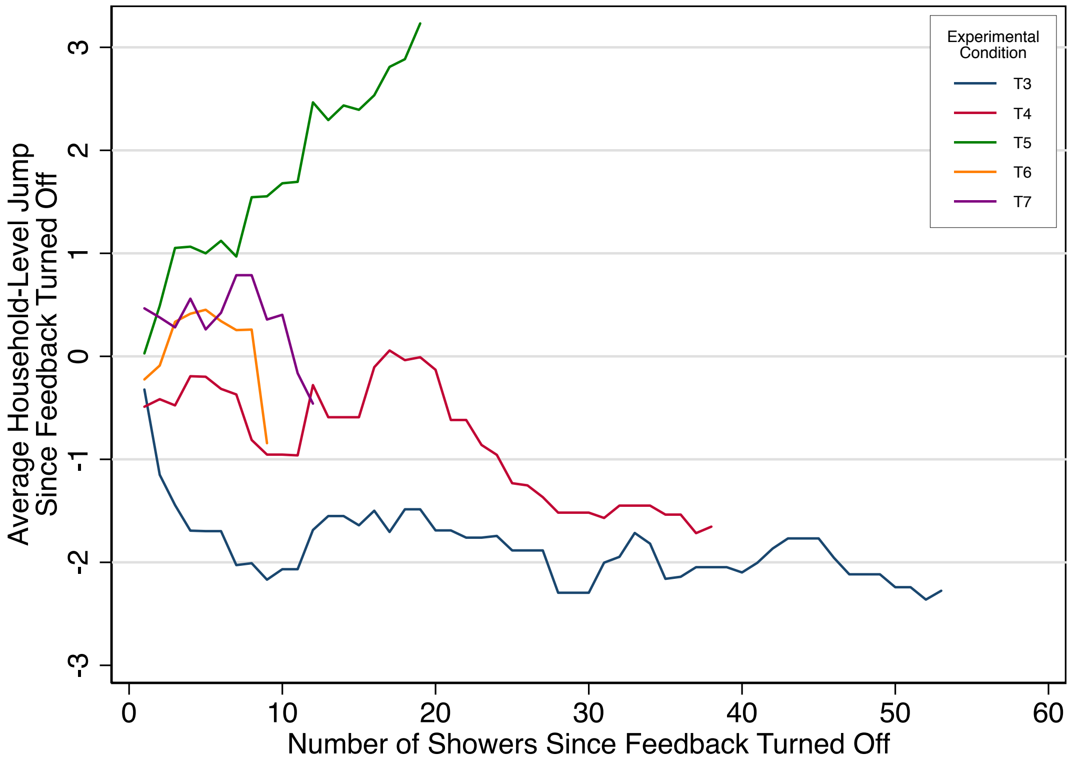


Figure D.3 contains our final set of results regarding post–feedback results. In this figure, we further study why the inclusion of household–specific jumps in (22) does not eliminate our estimates of post–feedback decay in treatment effects. To construct the figure, we compute the average post–feedback jump across households for each shower after feedback is turned off. In computing this average, we assign a household a jump equal to 0 if their jump had not yet occurred, otherwise they are assigned the jump estimated using the supF test. This is the time-varying average in jumps that, *a priori*, could have explained post–feedback treatment effect decay.

While the figure masks the underlying heterogeneity in positive and negative post–feedback jumps, it does reveal that the average jump does not have a systematic positive trend across all treatment conditions, particularly for conditions T3 and T4 with longer feedback-off cycles. This is precisely what underlies the modest fall in magnitude of our $PostOFF_{it}$ coefficient in the bottom panel of Table D.1 compared to the top panel in the table.

E Supplemental tables and figures

Figure E.1: Predicted Consumption Paths Under Experimental Condition T3 (Including 95% Confidence Intervals)