How nudges create habits: theory and evidence from a field experiment *

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

October 15, 2023

Abstract

We run a field experiment that monitors behavioral responses to nudges at a high frequency during and post-treatment and varies the duration of treatment cycles. We document asymmetry: treatment effects emerge immediately with nudges, neither grow nor wane with continued treatment, and gradually decay after nudges stop, taking longer to decay the longer the duration of treatment. To study the underlying behavioral mechanism for these dynamics, we extend the traditional consumption-based model of habit formation to incorporate salience and the possibility of state-dependent attention. We structurally estimate the model and find that a dynamic attention-based mechanism best predicts consumption responses to nudges in our context, both in and out of sample. Through counterfactual simulations, we illustrate the importance of identifying the underlying behavioral mechanism by contrasting implications of consumption- and attention-based habit formation when designing nudge interventions for sustained behavioral change.

Keywords: Habit formation; Salience nudges; Attention; Randomized Control Trial; Structural Estimation; Water Consumption

---

*Stefano DellaVigna, Ali Hortacsu, Steve Tadelis, Antonio Penta and participants at numerous conferences and seminars have provided helpful comments and suggestions. Edwin Chan, Zhi Hao, Rafi Kamsani, and Dorothy Ting provided outstanding research assistance. Funding from the University of Melbourne FBE Faculty Research Grant Scheme, Centre for Market Design, and German Research Foundation is gratefully acknowledged. **Disclosures:** 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 field experiment, built the structural model, analyzed the data, conducted the counterfactual policy analyses, and wrote the paper. 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. Human Research Ethics Approval ID 1544989 from the University of Melbourne governs this research. AEA RCT Registry ID AEARCTR-0008720. All errors are our own. Corresponding authors: byrned@unimelb.edu.au (Byrne), ecslfg@nus.edu.sg (Goette) and leslie.martin@unimelb.edu.au (Martin).

[a]Department of Economics, University of Melbourne

[b]Department of Economics, National University of Singapore

[c]South East Water Corporation

[d]Department of Information Systems and Applied Computer Science, University of Bamberg

[e]School of Business, Economics and Society, University of Erlangen-Nuremberg

1 Introduction

Habits are ubiquitous in everyday life, shaping our daily routines, from diet and exercise to our consumption of natural resources. Their importance underpins government interventions and digital technologies such as mobile apps and wearables that help us build "good habits" (see, e.g., Wood and Neal, 2016; World Bank, 2014) via feedback and reminders.

This approach to spurring habit-formation reflects substantive evidence from economics and psychology that individuals have limited attention and, as a consequence, fail to account for all the costs and benefits of consumption decisions (Thaler and Sunstein, 2009; Chetty et al., 2009; DellaVigna, 2009; Gabaix, 2019). By repeatedly drawing attention to the benefits of exercise or the environmental cost of electricity use, policymakers and companies attempt to create lasting changes in these behaviors. While existing research establishes nudges can have persistent effects, little is known about the underlying structure of persistence and its implications for the design of behavioral interventions.

This paper studies the micro dynamics of how nudges create habits. Specifically, we study the habit-forming effects of nudges that provide real-time feedback on consumption costs. We design and implement a field experiment that allows us to examine real-time behavioral responses to feedback, focusing primarily on what happens when treatment stops. We observe that behavioral responses to feedback occur immediately with they are provided, do not grow or wane as long as they are provided, and gradually decay when feedback stops, taking longer to decay the longer they were provided.

Inspired by these treatment effect dynamics, we develop a structural econometric framework that incorporates salience into models of habit formation. We begin by combining standard behavioral economic models of limited attention (Chetty et al., 2009) and habit formation (Stigler and Becker, 1977). In this model, persistence in feedback effects arises from complementarities between current and past consumption. We then consider an alternative behavioral mechanism for habit formation where we switch the source of persistence from consumption to attention. This mechanism follows the Stigler and Becker (1977) formulation by specifying a time-varying stock for attention that is motivated by extensive research in neuropsychology (Anderson, 2016). We develop a methodology to structurally estimate the model under consumption-based and attention-based habit formation mechanisms and find, in our context, that the attention-based mechanism best predicts and explains dynamic consumption responses to feedback.

Lastly, we use the structural model to study policy design for creating behavioral change through a series of counterfactual simulations. We show when feedback is costly to provide or receive, a consumption-based mechanism for habit formation favors a long front-loaded feedback intervention. In contrast, an attention-based mechanism that better explains our experimental data favors an intervention that initially provides a series of feedback to build an attention stock and then cycles feedback on and off to maintain attention to the costs of resource use thereafter.

The paper bridges two substantive areas of research in behavioral economics on limited attention and habit formation. The idea that attention is time-varying connects to research that finds that households exhibit short-term demand responses to intermittent bills (see, e.g., Ito et al., 2018; Allcott and Rogers, 2014; Gilbert and Graff Zivin, 2014).[1] We move beyond documenting persistence in salience-reducing interventions by developing and estimating a habit formation model of attention dynamics that can explain time-varying persistent feedback effects.

Research on habit formation is even more pervasive. Observational and experimental studies have documented persistence in smoking (Chaloupka, 1991; Becker et al., 1994; Cameron, 2000; Gruber and Kőszegi, 2001), coffee drinking (Olekalns and Bardsley, 1996), alcohol consumption (Baltagi and Griffin, 2002), exercising (Charness and Gneezy, 2009; Royer et al., 2015), water usage (Ferraro and Price, 2013), voting (Fujiwara et al., 2016), blood donations (Bruhin et al., 2020), electricity usage (Ito et al., 2018), handwashing (Hussam et al., 2022), hospital hygiene (Steiny Wellsjo, 2022), and social media usage (Allcott et al., 2022). Stigler and Becker (1977)'s or Becker and Murphy (1988)'s consumption-based habit stock models with consumption complementarities over time are often the lens through which such persistence is interpreted.[2]

Our paper more closely relates to recent papers that design experiments to investigate microfoundations of habit formation. Hussam et al. (2022) test for intentionality and anticipation: they randomize future rewards to show that people are more likely to form habits when they know that in the future those habits will be helpful to have. Allcott et al. (2022) test for awareness of habit formation, with a focus on self-control problems. And Camerer et al. (2020) and Steiny Wellsjo (2022) develop and test automatic control models of habit formation, whereby habitual decisions represent a "shortcut" to avoid costly deliberation in decision-making day-to-day.

In contrast, we focus on the interactions between limited attention and habit formation to explore *transitions* in behavior during and after cost salience treatments. A key feature of our study is high-frequency data on consumption and feedback and our ability to manipulate cycles of repeated feedback. We design our experiment to test two main hypotheses describing: (1) *symmetry* of behavioral responses to feedback and when feedback stops; and (2) the gradient between treatment duration and post-treatment *persistence*. The richness of our data enable highly powered, model-

[1] See DellaVigna (2009) and Gabaix (2019) for overviews of inattention and salience bias research.

[2] Early studies on habit formation focus on physically addictive behaviors such as smoking (e.g., Becker et al., 1994, 1991) that create such a complementarity through the metabolic properties of the good. However, other mechanisms can generate similar effects. Dal Bó and Terviö (2013) show that such complementarities can also arise in a model of self-signaling: good actions today signal to the individual that she is likely a good type, thus creating a complementarity with future good actions, in the spirit of Stigler and Becker (1977).

free tests of these symmetry and persistence hypotheses. Our ability to test these hypotheses, in turn, permit a quantitative evalution of competing mechanisms for habit formation in the presence of limited attention. In particular, through a series of out-of-sample validation exercises, we find that the attention stock mechanism best explains behavior in our setting, even when compared to a consumption-based mechanism that flexibly allows for asymmetries in behavioral responses to feedback.

The counterfactual policy simulations from our model also reveal a novel rule for optimal allocation of feedback over time for an intervention where feedback must be rationed. In particular, we uncover an (I,S,s) optimal feedback rule, whereby feedback is provided for an initial I periods, and behavioral change is then maintained for the remainder of the intervention through an (S,s) rule. This finding creates new linkages between Scarf's (1959) classic (S,s) rule for solving dynamic inventory management problems and the solution to dynamic attention management problems, such as feedback-based behavioral interventions that help create "good" habits.[3]

We develop our study in six parts. Section 2 describes our experimental design and implementation, focusing on how we vary the duration of feedback exposure across individuals. Section 3 provides evidence of time-varying consumption effects with feedback and when feedback stops, and how persistence in these effects depend on the duration of repeated feedback. Motivated by the reduced-form evidence, Section 4 develops and estimates a structural model for explaining persistence in treatment effects that allows for both a traditional consumption-based habit stock mechanism and attention to salience, where attention is also modeled as a stock that may evolve over time. We further exploit the richness of our data to rule out automatic control (Camerer et al., 2020) and experimentation and learning (Larcom et al., 2017) as alternative behavioral mechanisms. Section 5 then turns to policy examining how the design of interventions affects behavior change in the presence of time-varying attention stocks. Finally, in Section 6, we summarize our results and discuss avenues for future research in identifying new mechanisms for habit formation to inform policy design in behavioral interventions.

2 The field experiment

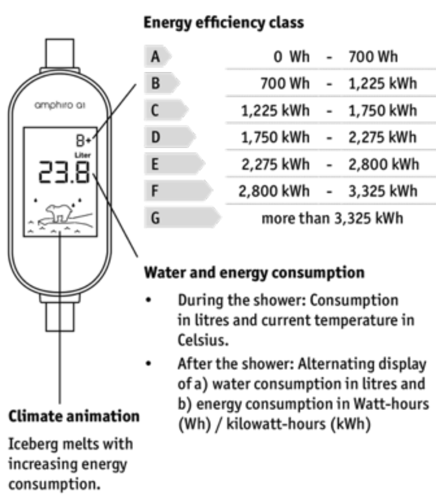
This section describes the experiment, its implementation, and the data that it creates. We also present summary statistics to characterize the study's internal and external validity.

[3] (S,s) rules emerge is various dynamic models that study nominal price rigidity (Sheshinkski and Weiss, 1977; Caballero and Engel, 2007), consumer demand for durables (Eberly, 1994; Attanasio, 2000), markup dynamics (Aguirregabiria, 1999), and trade frictions and import pricing (Alessandria et al., 2010).

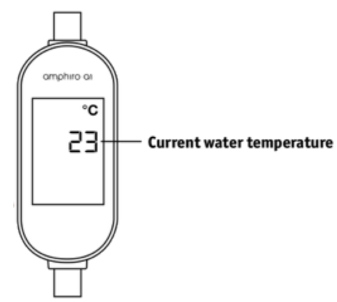
Figure 1: Amphiro B1 Smart Shower Meter

Figure 2: Real-time Feedback Modes for the Amphiro B1

(a) Real-time Feedback On



(b) Real-time Feedback Off

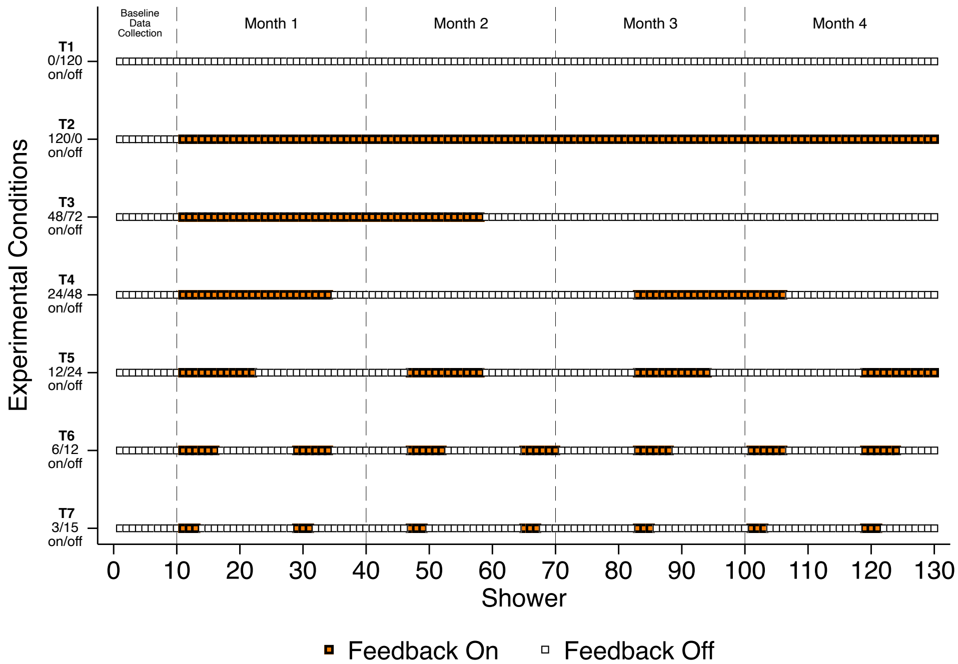


2.1 Design

The intervention we study in this paper provides real-time in-the-shower feedback on water consumption. The smart shower meter that provides this salient feedback is the Amphiro B1 shown in Figure 1. The meter is mounted between the shower hose and a hand-held showerhead and is powered by water flow.

During our experiment, participants' Amphiro B1s are either in a feedback-on or feedback-off mode, as illustrated in Figure 2. Panel (a) depicts the *feedback-on* mode where the meter displays real-time feedback on liters of water used since the beginning of each shower and current water temperature. The device also converts the cumulative water use into a melting-icecap visualization and, at the end of the shower, provides an energy efficiency class rating. We also refer to this as

Figure 3: Experimental Design



our "treated" mode. Panel (b) shows the *feedback-off* mode where the meter only displays current water temperature. This is our "control" mode. We display water temperature instead of a blank screen in the feedback-off mode so that individuals experiencing that mode can be confident that their Amphiro B1 is working. The meters collect data even when feedback is turned off.

We partnered with engineers that manufacture these meters to implement a research design that cycles device features between these feedback-on and feedback-off modes according to set schedules across various experimental conditions over a four-month study period. Our experimental design is presented in Figure 3. There are seven experimental conditions labeled T1 to T7. In each condition, the Amphiro B1 begins in the feedback-off mode and collects baseline shower usage data for ten showers. The baseline period allows us to confirm balance on baseline shower usage across our seven different experimental conditions and identify within-subject treatment effects.

After the baseline data collection period, the Amphiro B1s cycle between the feedback-on and feedback-off modes at different frequencies across conditions T1 to T7. The first two conditions provide benchmarks: in T1, feedback is always off (i.e., a pure "control" mode group), while in T2 after the baseline period feedback is always on (i.e., a pure "treatment" mode group). From T2 through T7, each subsequent treatment has progressively shorter feedback-on cycles: in T3 feedback is on for 48 showers, then off for 72 showers. In T4 it is on/off for 24/48 showers in multiple cycles, in T5 on/off for 12/24, in T6 on/off for 6/12, and in T7 on/off for 3/15. In T3 through T7, the periods of feedback-off are always longer than the periods of feedback-on, reflecting our interest in examining what happens when feedback stops.

2.2 Hypotheses

We designed the feedback cycles in the experiment to test two hypotheses that describe how consumption may respond dynamically to feedback. They are:

Hypothesis 1 (Symmetry): Consumption falls and rises *symmetrically* when feedback is turned on and off.

Hypothesis 2 (Persistence): The longer feedback is on, the longer it takes for consumption to return to baseline levels when feedback is turned off. In other words, there is a positive relationship between the duration of feedback being on and the *persistence* in treatment effects when feedback is subsequently turned off.

Experimental arm T3 is crucial for testing **Hypothesis 1** because it includes long cycles of both feedback-on and feedback-off. This feature of our design helps ensure, *a priori*, that participants have enough time to completely accumulate and subsequently depreciate any stock effects associated with feedback or its removal. In other words, this treatment arm allows us to observe transitions to and from long-run consumption levels with long durations of feedback being turned on and off.

The combination of conditions T3 through T7 allows us to test **Hypothesis 2**. These treatment arms provide variation across experimental conditions that allow us to check whether and, if applicable, measure the extent to which post-treatment persistence of feedback strengthens with the duration of feedback. For example, under this hypothesis, we expect a slower rebound in consumption during feedback-off periods in T3, after 48 showers of feedback, than in T7, with just 3 showers of feedback.

As we will see in Section 4 below, our ability to test these two main hypotheses is central to empirically discerning underlying behavioral mechanisms for habit formation.

2.3 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 ran an online survey of randomly-selected households and, from the respondents, identified those with the shower configuration required for the Amphiro B1. We invited these households to express interest and, of those who did, randomly selected 700 customers for the trial.

Appendix A.1 contains details on the recruitment process, including the full Plain Language Statement (PLS) provided to households. We framed the trial in terms of examining how real-time feedback affects shower water usage. The PLS described the Amphiro B1 shower water meters and explained that at different times during the trial period the meters would show different real-time indicators, including water temperature and, in some cases or at some times, water usage. The PLS also asked participants to commit to mailing back the devices at the end of the trial for researchers to extract anonymized water usage data and factory reset the devices to full-functionality. The meters were then returned to participants to keep.

We mailed the Amphiro B1s to our experimental sample in late May 2017. We included paper and online trial and installation instructions with the meters (reproduced in Appendix A.1) and postage-paid return envelopes. As with the PLS, the instructions maintained neutral language and, for treatment groups that would be observing different feedback on and off modes, did not explain when or why the meters would have different display modes.

In October 2017, we emailed participants asking them to return their Amphiro B1s for data extraction and factory reset. We explained that when reset to full functionality, the meters would display the full range of features and become pairable to an app that allows users to view their historical data.[4]

2.4 Data

We obtained anonymized pre-experiment billing and account data on quarterly household water usage and bills, electronic-billing, hardship, and tenant status for all utility customers. The utility also matched households to their Statistical Area 1 (SA1) 150-household census block from the Australian Bureau of Statistics to obtain block demographics such as average household income, age, education, and home size.[5]

Appendix A.1 describes the baseline survey sent to utility customers. This survey yields information on household characteristics, shower configurations, and shower habits, including estimated shower water usage.

For the participants in the experiment, the Amphiro B1 provides us a time series of actual shower data. The meter records shower number, total water used, the average water flow rate, and average water temperature for each shower taken. Because the device does not have a battery, it does not have an internal clock, which means it does not record a given shower's date or time. Therefore, in our reduced-form regressions below for testing **Hypotheses 1** and **2**, we can control for shower count, but not date.

Our experiment includes both single and multi-person households. Single-person households have feedback on and off cycles programmed on their Amphiro B1's exactly as described in Figure 3. Multi-person households, primarily households with two adults, or households with two adults and children who used a separate shower, have Amphiro B1's programmed with twice as long

[4]During the experiment, online stores for the Amphiro B1 apps were shut down in the country to ensure households could only access Amphiro B1 feedback shower-by-shower. The app stores were opened after the experiment ended.

[5]SA1's contain approximately 150 households on average and are the most narrow census block available.

baseline periods and feedback on and off cycles. In Appendix A.2 we demonstrate that all of our results below hold when restricting our analysis to single-person households.

The meter's internal memory saves a maximum of 245 consecutive showers-worth of data. Our four-month experimental period (June-September 2017) corresponds to the upper bound for data storage on an Amphiro B1 for a two-person household where each household member showers once a day. We confirm in Appendix A.3 that there are no differences in the cumulative number of showers taken by households, as recorded by the Amphiro B1, across conditions T1–T7 during the trial. This result mitigates concerns such as differential attrition across conditions or that adults in households with multiple showers selective switch to their child's shower in response to or in anticipation of experimental conditions.

Finally, we ran an endline survey with responses from 427 of the 555 households (77%) who returned their Amphiro B1's used with data at the end of the trial. Appendix A.1 also reproduces this survey. This survey helps us confirm how and where households installed their Amphiro B1 and the ease with which they could read feedback from the device. Although respondents overwhelming claimed to prefer the feedback-on mode, reported satisfaction with the Amphiro B1s was constant over the course of the experiment, regardless of feedback-on duration or cycle frequency. Respondents also stated that they mainly focused their attention on the real-time metrics of shower water temperature and water usage.

2.5 Summary statistics

Table 1 presents mean household water usage and characteristics across various samples from South East Water's database of 140,407 customers with email. We emailed our baseline survey to a random subset of 45,685 (33%) of these customers. Comparing columns (2) and (3), households who answered the 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 in water usage or demographics between survey respondents and non-respondents.

Among the 19,449 households that answered the baseline survey (43% response rate), 5,866 had a handheld shower that could mount the Amphiro B1 (30% eligible). We invited these households to the trial using the PLS described above. In total 1,201 households (20%) opted in, and of those, we randomly selected 700 households for the experiment, stratifying the sample to prioritize first single-person households, then two-person households, and then households where any extra household members used a separate shower.

Columns (3) and (4) of Table 1 shows to what extent trial participants differ from survey respondents. Trial households are more likely to have electronic billing and have registered for

Table 1: Mean Household Characteristics for Different Subsamples from South East Water's Customer Base and by Device End-of-Trial Return Status

	Sub-Sample					
	Households with Email (1)	Emailed Survey (2)	Answered Survey (3)	Sent Device (4)	Device Accounted For (5)	Returned Device with Data (6)
Jul-Sep 2016 Water Usage (L)	31.45	31.59	29.56	29.31	29.09	28.86
Oct-Dec 2016 Water Usage (L)	37.90	37.88	36.06	34.07	33.79	33.75
Jan-Mar 2017 Water Usage (L)	45.41	45.43	44.07	38.68	38.68	38.44
Apr-Jun 2017 Water Usage (L)	39.35	39.34	37.52	34.99	34.66	34.46
Annual HH Income (1000s)	53.26	53.22	52.55	51.75	51.93	51.86
Average Age	37.67	37.62	37.82	36.52	36.51	36.42
Share of High School Graduates	0.46	0.46	0.45	0.45	0.45	0.45
Number of Bedrooms in Home	2.94	2.95	2.97	3.04	3.04	3.05
Share of Tenants	0.34	0.33	0.19	0.19	0.18	0.19
Share of HHs with Electronic Billing	0.47	0.49	0.68	0.75	0.76	0.79
Share of HHs Registerd with Web Portal	0.38	0.39	0.52	0.64	0.65	0.66
Number of People Living at Home			2.67	2.64	2.60	2.60
Self-Reported Shower Time			6.47	6.90	6.84	6.88
Number of Leaks Checks per Year			2.30	2.26	2.25	2.21
Households	140407	45685	19449	700	653	555

Notes: Households with devices accounted for in column (5) correspond to households from Table 2 who returned their Amphiro B1 used, uninstalled and unused, or who experience postal or device error. Quarterly water usage and customer account information is from South East Water. Census block demographics are from the Australian Bureau of Statistics and correspond to Statistical Area 1 census block averages in which a customer lives.

South East Water's web portal. Otherwise, there are minimal differences in quarterly water usage and other characteristics between survey respondents and trial households. By comparing columns (4), (5) and (6) of Table 1, we check for selection effects in Amphiro B1 return status. We find no differences in pre-treatment water use or any other household characteristic between the 700 households to whom we mailed a device, the 653 households whose Amphiro B1 we can account for at the end of the trial, and the 555 households who successfully used and returned their device for data extraction.

Table 2 describes the device return statuses in more detail and tabulates the data by experimental condition. There are four mutually exclusive groups: (1) device returned with shower usage data; (2) devices returned unused, most frequently with the package unopened (i.e., having never been installed);[6] (3) postal error (e.g., return-to-sender) or technical error (e.g., the device had a blank-screen faulty display); or (4) no response. Overall, we have very low non-response rate. For instance, we obtain a minimum device return rate of 89% to 97% among working devices that were

[6] Some households included written reasons with their returned devices for non-installation. The most common reasons given in our endline survey are vacations, forgetting, and suspected incompatibility with their shower head.

Table 2: Amphiro B1 End-of-Trial Return Status by Experimental Condition

	Experimental Condition						
	T1	T2	T3	T4	T5	T6	T7
End-of-trial device return status							
Device returned used	77	84	79	86	78	76	75
Device returned uninstalled	12	8	10	9	12	10	18
Postal or device error	5	3	1	2	2	3	3
No response	6	5	10	3	8	11	4
Total devices sent	100	100	100	100	100	100	100
Percentage of devices returned used out of ...							
Total devices sent	77%	84%	79%	86%	78%	76%	75%
Total devices with no error and installed	93%	94%	89%	97%	91%	87%	95%

received and installed.[7]

Table 2 also shows that installation rates and device return rates do not vary significantly across treatment groups. Most importantly, there is no evidence that households experiencing more feedback-on or more frequent on/off cycles have lower device return rates. Appendix A.4 further confirms that there is no selection into returning devices based on households' pre-trial actual household water use or stated within-shower water use.

Table 3 presents mean characteristics for households across our experimental conditions. The first three rows display mean shower water usage volume, flow rate, and shower length as reported by the Amphiro B1. We construct these variables by computing household-specific means from their initial 10-shower per person baseline phase with feedback off. In the table, we report the sample mean of these characteristics across households within each condition.

Baseline shower water use, shower flow rates, and shower length are very similar across all conditions. Indeed, none of the differences is jointly statistically-significant, nor do any pairwise comparisons across groups yield statistically-significant differences. Likewise, all other household characteristics are statistically similar across all groups. In sum, our randomization achieves balance on observables across our seven experimental conditions.

3 Symmetry and persistence in feedback effects

In this section we explore and formally test **Hypotheses 1** and **2** to characterize empirically how treatment effects build up when feedback is turned on and decay after feedback stops.

[7] These estimates are a lower bound because they assume that all non-responsive households received and installed the devices.

Table 3: Mean Household Characteristics by Experimental Condition

	Experimental Condition						
	T1 0/120 on/off (1)	T2 120/0 on/off (2)	T3 48/72 on/off (3)	T4 24/48 on/off (4)	T5 12/24 on/off (5)	T6 6/12 on/off (6)	T7 3/15 on/off (7)
Baseline water usage per shower							
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
Quarterly household water usage							
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
Census block demographics							
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
Household billing account information							
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
Household survey information							
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
Households	77	84	79	86	78	76	75

Notes: Individual shower usage is the sample mean during the baseline period by individual in the sample; see the text for details. Quarterly water usage and customer account information is from South East Water. Census block demographics are from the Australian Bureau of Statistics and correspond to Statistical Area 1 census block averages in which a customer lives. See Figure 3 for details on the experimental conditions.

3.1 Graphical analysis

Figure 4 graphically describes time-varying, feedback-induced treatment effects from our experiment. To construct these figures, we run regressions of the following form

$$y_{is} = \eta_i + \sum_{b=1}^{B} \beta_b \left(\mathrm{T}_j \times 1\{s \in b\}\right) + \tau_k + \varepsilon_{is}, \tag{1}$$

where y_{is} is shower volume for household i in shower s, T_j equals one if household i is in experimental condition Tj ($j = 2, \ldots, 7$), $1\{s \in b\}$ is a dummy equaling one if shower s is within shower block b (defined momentarily), η_i is a household fixed effect, τ_k is a fixed effect it being k showers per person in the household since the start of the experiment,[8] and ε_{it} is the regression error.

We implement (1) such that β_b quantifies the within-household change in consumption in shower block b *relative* to household i's mean baseline shower usage. We estimate β_b for $B = 36$ shower blocks. For one-person households, these correspond to 3-shower blocks, while for multi-person households, they correspond to 6-shower blocks. Blocking in this way reflects our doubling of baseline and feedback on/off phases for multi-person households relative to single-person households, as discussed in Section 2.4 above. The 36 blocks reduce noisiness in the time-varying treatment effects, allowing us to visualize how they evolve. The precise alignment of the blocks with when feedback turns on and off under our experimental design allows us to visually inspect whether there are sharp or gradual changes in consumption immediately after feedback is turned on and off.[9]

We estimate equation (1) separately for conditions $j = 2, \ldots, 7$ where for a given condition we use households in T1 (control) and Tj in estimation. Plotting the coefficients estimates $\hat{\beta}_{1,j}, \hat{\beta}_{2,j}, \ldots, \hat{\beta}_{B,j}$ visualizes the time path of treatment effect build up and decay when feedback is turned on and off for condition j. In this way, the coefficients let the data speak to the symmetry and persistence in feedback effects induced by our experiment, per **Hypotheses 1** and **2**. Panels (a)-(f) of Figure 4 plot the estimates for each condition.[10]

Five patterns of interest emerge in the figure. First, real-time feedback has an immediate and stable effect on water use in showers. All panels reveal an immediate drop in shower water usage when feedback is turned on after the baseline phase. Second, there is no evidence of a subsequent downward trend in water usage following the initial drop in water usage after feedback is turned on and kept on. Panels (a)-(c), with longer cycles of real-time feedback, most clearly reveal this no-further-decrease pattern.

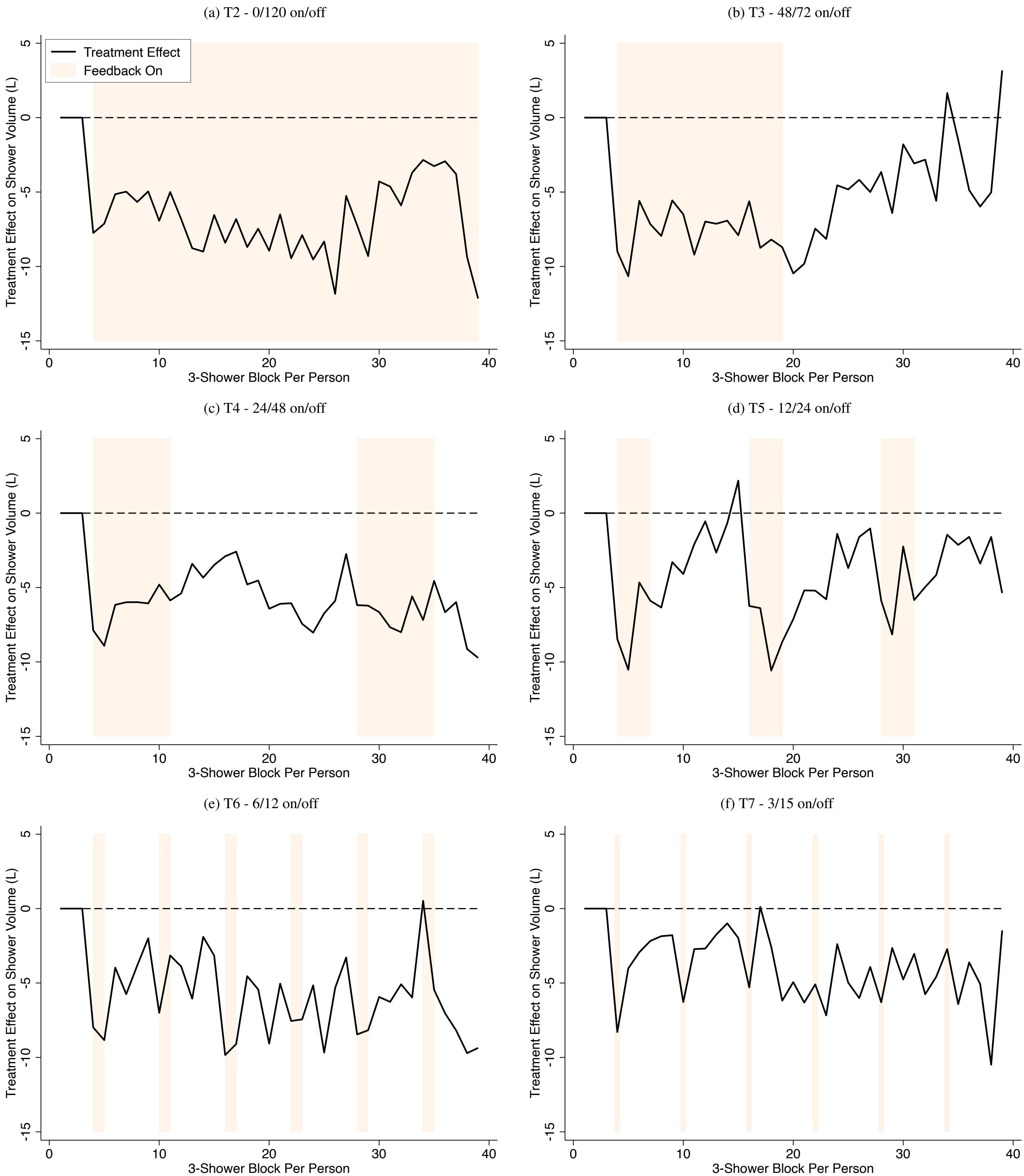
Third, there is clear evidence of persistence when real-time feedback is turned off. Water usage does not immediately return to baseline once the water usage cost salience is removed. Fourth, in all conditions, the persistence effect degrades over time: after feedback is turned off, water usage gradually trends back to baseline levels. Panel (b), in particular, shows that consumption eventually reaches baseline levels if feedback remains off for a sufficiently long period. This finding suggests that our experimental design provided a long enough window without feedback to indeed reveal a transition back to baseline water usage in the presence of decaying feedback effects. Together,

[8] For example τ_k enumerates $1,2,3$ for one-person households, $1,1,2,2,3,3$ for two-person households, and so on. The τ_k fixed effects help control for seasonality in shower water usage as long as our treatment and control groups are: (1) balanced on household size; and (2) install and use the devices at the same rate throughout our experiment. We confirm the former in Table 3 and the latter in Appendix A.3. We cannot include weather-related controls in (1) because the Amphiro B1 does not have an internal clock that records date of shower.

[9] Our results are unchanged if we plot β_b with smaller or larger shower blocks that do not align as well with our experimental design.

[10] For clarity, we do not report confidence intervals in Figure 4. We defer formally testing **Hypotheses 1** and **2** to Section 3.2 below.

Figure 4: Time-Varying Treatment Effects by Experimental Condition



Notes: See Figure 3 for details on the experimental design and equation (2) and associated discussion in the text for the regression equations used to generate these plots. For clarity, confidence intervals are not displayed.

the first four patterns of interest show that the build-up and decay of the treatment effect when feedback is turned on and off is asymmetric, which is evidence against **Hypothesis 1**.

Finally, although persistence in feedback effects exists even following short feedback periods (panels (e) and (f)) the strength of the persistence effect appears to scale with the duration of the feedback. Comparing panels (e) and (f) to panel (b) we find shorter duration feedback cycles are associated with less persistence. In panel (b), consumption gradually trends back to baseline when feedback is turned off. In constrast, in panels (e) and (f) consumption exhibits an initial upward jump when feedback is turned off and then begins trending back to baseline. These patterns provide preliminary support for **Hypothesis 2**.

3.2 Treatment effects

We now formally test **Hypotheses 1** and **2** using the following regression:

$$y_{is} = \eta_i + \beta_1 ON_{is} + \beta_2 PostON_{is} + \beta_3 OFF_{is} + \beta_4 PostOFF_{is} + \tau_k + \varepsilon_{is}, \tag{2}$$

where y_{is} is 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 on 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.[11] All of our regressions include household and shower fixed effects, η_i and τ_k. We 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 shower fixed effects. The econometric error term, ε_{it}, is clustered at the household level, which is our level of randomization.

Our analysis considers estimates from two regressions based on equation (2). The first restricts $\beta_2 = 0$ and $\beta_4 = 0$, allowing us to test whether consumption returns to baseline after feedback is turned off through the β_3 estimate. We then examine estimates from the unrestricted regression in (2), allowing us to examine how treatment effects build-up when feedback is turned on with β_1 and β_2 and decay after feedback is turned off with β_3 and β_4.

Baseline results

Table 4 presents our empirical results. The top panel presents benchmark estimates from the restricted regression while the bottom panel presents estimates from the unrestricted regression.

[11] To take a concrete example, consider treatment group T4 with a 24/48 on/off feedback cycle. As depicted in panel (c) of Figure 3, this condition has two feedback-on and two feedback-off spells. Feedback is initially 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 4: Treatment Effects by Experimental Condition

	Experimental Conditions Included in the Sample						
	T1-T7	T1,T2 120/0 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	-7.31***	-7.10***	-7.49***	-7.18***	-7.21***	-7.44***	-7.00***
	(0.70)	(1.39)	(1.10)	(1.06)	(1.05)	(1.09)	(1.07)
OFF	-3.81***		-4.89***	-4.19***	-2.91**	-4.07***	-3.45***
	(0.72)		(1.42)	(1.23)	(1.22)	(1.06)	(1.04)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.45	0.44
Observations	86376	24648	37798	38286	36885	35862	36137
ON	-7.39***	-6.65***	-7.31***	-7.31***	-7.13***	-7.53***	-6.96***
	(0.70)	(1.39)	(1.10)	(1.06)	(1.04)	(1.09)	(1.05)
$PostON$	0.01	-0.01	-0.01	0.00	-0.01	0.00	-0.00
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
OFF	-4.85***		-7.70***	-5.44***	-5.17***	-4.68***	-3.60***
	(0.73)		(1.44)	(1.28)	(1.24)	(1.22)	(1.19)
$PostOFF$	0.08***		0.11***	0.07*	0.19***	0.08	0.02
	(0.02)		(0.04)	(0.04)	(0.06)	(0.09)	(0.07)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.45	0.44
Observations	86376	24648	37798	38286	36885	35862	36137

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. Within R-Squared reported for each model. See Figure 3 for details on the experimental design and the feedback on/off spells for experimental conditions T1–T7. Standard errors are clustered at household level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

The column (1) estimates pool data for all seven 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 persistence effects from a sample that includes T1 and T2 households and one of the T3 to T7 groups.

Column (1) in the top panel of the table shows that providing real-time feedback sharply lowers water use during showers. The point estimate of -7.31L / shower represents a large 13% reduction in shower water usage compared to baseline mean usage of 57 L. The estimates further yield, across all conditions, statistically significant and economically meaningful β_2 estimates on OFF_{is}, thereby revealing persistence effects. Remarkably, persistence effects arise even in condition T7, where feedback is only given for $1/6^{th}$ of the time in a 3/15 feedback on/off cycle. Nevertheless, this leads to an estimated persistence effect of -3.45 when feedback is off, which is approximately half of the -7.00 treatment effect when feedback is on.

Symmetry in feedback effects

We now turn to the estimates in the bottom panel of Table 4. The results in column (1) strongly reject **Hypothesis 1** as they reveal stark asymmetry between feedback-on and feedback-off consumption dynamics. Real-time feedback has an immediate and stable effect on consumption behavior. The -7.39 L/shower point estimate of the ON_{is} coefficient reflects the impact of feedback in the first episode of a feedback period. It is virtually identical to its corresponding -7.31 point estimate in the top panel, suggesting that the entire effect occurs instantaneously. Moreover, the interaction effect with the duration of exposure to feedback is a precisely estimated zero, as shown by the 0.01 $PostON_{is}$ coefficient estimate, This finding implies stable treatment effects: they do not grow or wane while feedback is kept on. In contrast, the effect of treatment post-feedback slowly erodes, as revealed by the statistically significant 0.08 coefficient estimate on $PostOFF_{is}$ in column (1). This estimate implies that with every shower during a feedback-off phase, post-treatment water use increases by 80 milliliters, gradually returning towards pre-treatment levels.[12]

Based on the column (1) specification in the bottom panel of Table 4, the predicted cumulative water savings induced by T2 (120/0 on/off) over 120 showers is 814 liters.[13] Conditions T3, T4 and T5, comparable treatments that also have 48 of 120 showers with feedback on, have predicted cumulative water usage reductions of 482, 580, and 629 L, respectively. These back-of-the-envelope calculations highlight how, in the presence of asymmetric feedback effect build-up and decay, intermittent feedback throughout a trial (T4 and T5) can induce greater conservation than continuous feedback from the start of a trial (T3). We formalize and examine feedback intermittency and intervention design using our structural model in Section 5.3 below which, as we will see, better fits our experimental data than simpler linear regression models.

Feedback duration and post-feedback persistence

The results in columns (3) to (7) of the bottom panel in Table 4 yield nearly identical stable feedback effects. Indeed, the ON coefficients in columns (3)-(7) have similar magnitude, and the $PostON$ coefficients have precise zero estimates. Collectively, these results reaffirm that feedback effects emerge immediately and remain stable while feedback is on, irrespective of duration.

Turning to persistence, accounting for the duration of off-periods in columns (3)-(7) allows a clearer interpretation of the persistence effects as the effect in the first episode of an off-period. We find monotonicity in the point estimates. In particular, in column (3), the point estimate of β_3 is -7.70L / shower after 48 periods of feedback and statistically indistinguishable from the feedback effect itself. The estimates of β_3 on OFF_{is} in equation (2) monotonically decline as the duration of

[12] How many showers does it take until the water-conserving habit in consumption induced by feedback fully decays? We quantify the potentially non-lineary decay rate and half-life for feedback persistence effects using our structural model in Section 5.1 below.

[13] This water usage reduction equates to a \$2.17 (or 1%) reduction in the average household's water bill.

feedback phases becomes shorter, which again supports **Hypothesis 2**. Interestingly, column (7) reveals statistically and economically significant persistence effects even in the T7 experimental condition with the shortest feedback phase of just three showers.[14]

A particular cut of our data admits a direct test of **Hypothesis 2**. In particular, we estimate a variation on equation (2),

$$y_{is} = \eta_i + \beta_1(\mathrm{T}_1 \times ON_{is}) + \sum_{j=3}^{7} \left[\beta_{2j}(\mathrm{T}_j \times ON_{is}) + \beta_{3j}(\mathrm{T}_j \times OFF_{is})\right] + \tau_k + \varepsilon_{is}, \tag{3}$$

with a subsample that includes: (1) data for all households and showers in experimental conditions T1 and T2; (2) all baseline showers and the first three showers with feedback on immediately after the baseline for conditions T3–T7; and (3) the first 12 showers with feedback off following the *first* feedback cycles in conditions T3–T7.[15] Given our finding from Table 4 that feedback effects emerge immediately and do not evolve while feedback is on, β_{2j} in effect quantifies the feedback effect for condition Tj.

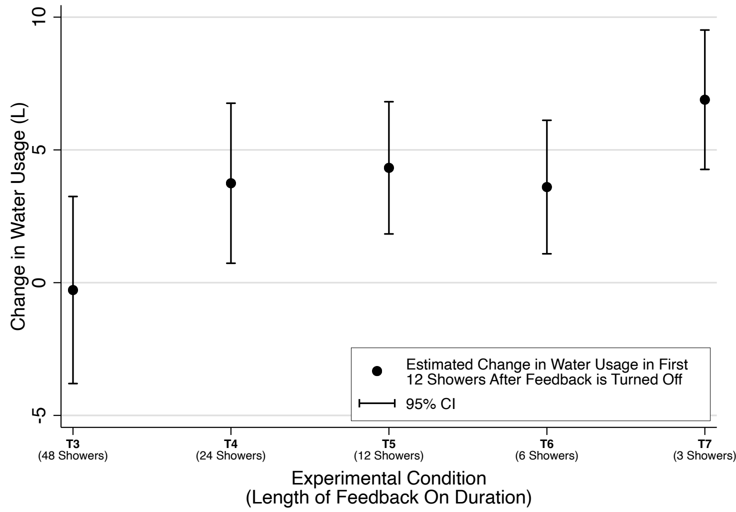
The β_{3j} coefficient in equation (3) quantifies the level of consumption in the first 12 showers with feedback off following condition Tj's first feedback cycle. Therefore, $\beta_{3j} - \beta_{2j}$ quantifies a persistence effect of feedback after the first feedback cycle ends. By examining this difference between conditions T3–T7, we can non-parametrically plot the relationship between persistence and feedback duration *after* 48 (T3), 24 (T4), 12 (T5), 6 (T6), and 3 (T7) showers of feedback. Importantly, focusing on persistence effects following the first feedback cycles allows us to abstract from spillovers in feedback effects across feedback cycles in conditions T4–T7. Examining persistence from three showers with feedback in the first 12 showers without feedback allows us to directly compare persistence effects across all five experimental conditions. In this way, our estimates from (3) provide a data-driven test of how feedback-on duration affects the persistence of treatment effects, leveraging our unique experimental design.

Figure 5 presents our results, plotting $\hat{\beta}_{3j} - \hat{\beta}_{2j}$ for T3–T7 and the 95% confidence interval for the difference in coefficient estimates. The results provide compelling evidence of the relationship between feedback duration and post-feedback persistence, and hence **Hypothesis 2**. After 48 showers in T3, the feedback effect fully persists, with a 0 L/shower increase in water usage in the first 12 showers with feedback off. In contrast, in conditions T4–T6, we estimate 3.6-4.3 L increases (from a $\approx$ 7.5 L feedback effect) after 24, 12, and 6 showers feedback. Following just 3 showers of feedback in T7, we find a much larger 6.9 L increase in usage, and hence a much smaller persistence effect.

[14]Notice also that each of the point estimates of β_4 on $PostOFF_{is}$ is positive, though not always significant. The relative imprecision in these estimates in columns (6) and (7) is partly due to the shorter feedback-off phases for identifying persistence effects.

[15]From experimental design in Figure 3, accounting for the 10-shower baseline phase, the first 12 showers with feedback off after the first feedback cycle are showers 59-70 (T3), 35-46 (T4), 23-34 (T5), 17-28 (T6), and 14-25 (T7).

Figure 5: Feedback Duration and Post-Feedback Persistence



Statistically, we reject the null that all post-feedback persistence effects in Figure 5 are jointly equal ($p = 0.02$). Moreover, all pairwise tests of equality in persistence effects involving either T2 or T7 with any other condition reject the null at the 5% level. That is, our 'extreme' experimental conditions, T3 and T7, with 48 and 3 showers of feedback, reveal substantial differences in persistence effects, which is directly in-line with **Hypothesis 2**. We cannot, however, distinguish persistence effects among conditions T4–T6 with intermediate levels of feedback duration.

Summary

In sum, our reduced-form analyses reveal: (1) asymmetry in feedback effect build-up and decay (rejecting **Hypothesis 1**); and (2) post-feedback persistence effects that become stronger with longer feedback duration (supporting **Hypothesis 2**). These reduced-form results motivate our structural analysis of feedback-based habit formation which seeks to identify the underlying behavioral mechanism (Section 4) and examine implications for the design of feedback-based feedback interventions (Section 5).

Importantly, Appendix A.2 shows our reduced-form results are robust to narrowing to subsamples containing households with just one person or one shower. Given this robustness, our structural model builds from an individual decision-maker and is estimated on our entire sample.

4 A model of feedback-based habit formation

Inspired by the reduced form evidence, our structural analysis of habit formation begins by combining two standard behavioral economic models: (1) Chetty et al. (2009)'s model of salience bias, which allows individuals to be inattentive to resource use and only perceive a fraction of their consumption cost; and (2) Stigler and Becker (1977)'s classic model of habit formation, which entails a consumption-based mechanism for persistence whereby the marginal utility of current consumption depends on past consumption. Combining these models yields a baseline framework for studying the dynamic consumption effects of salience-enhancing feedback. We can apply this model to our experimental data, and in doing so, interpret feedback-induced consumption dynamics as arising from state dependence in consumption.

We then consider an alternative mechanism where state-dependence arises not from consumption but from attention. We mirror Stigler and Becker (1977)'s formulation of state-dependence by allowing current attention to costs to depend, in a parallel way, on past levels of attention to such costs.

Our analysis proceeds in seven parts. We describe the model set-up in Section 4.1 and characterize optimal consumption decisions and the model's steady states in Section 4.2. Section 4.3 describes how we empirically implement the model leveraging our experimental variation in salience-enhancing feedback. Here, we consider estimation of two restricted versions of the model, one that allows for consumption-based state dependence, and the other that allows for attention-based state dependence. Given our empirical specifications, Section 4.4 describes model identification and estimation. We then compare the performance of these respective empirical specifications in terms of within-sample fit in Section 4.5 and through an out-of-sample cross-validation analysis in Section 4.6. Finding strong empirical support for an attention-based mechanism, in Section 4.7 we discuss and rule out other potential persistence mechanisms for feedback in our setting. These include automatic control models of decision-making and experimentation and learning.

4.1 Model set-up

In period t an individual realizes utility U_t, which depends on their current consumption level c_t, an attention parameter θ, and exogenously–given per-unit price p:

$$U_t = u(c_t) - \theta p c_t. \tag{4}$$

In our setting, we interpret price as including an individual's private water usage cost *and* their moral cost from using the natural resource, following Allcott and Kessler (2019) and Levitt and List (2007).

Limited attention

The parameter $\theta \in [0,1]$ is the individual's level of attention to their resource use, and, hence its associated cost. In our context, this entails the salience of incurring private and moral costs from water usage, and how these costs rise with higher levels of usage. As in Chetty et al. (2009) or DellaVigna (2009), we assume that individuals only give weight θ to their consumption, and hence also their consumption expenditure pc_t, due to limited attention. The interpretation in our research context is that while an individual immediately and correctly feels a shower's pleasant sensation, the associated resource use is difficult to perceive. As $\theta \to 1$, the quantity and cost of consumption is correctly perceived, and any associated salience bias in consumption (relative to the fully attentive case) goes to 0. In the context of our experiment, we will interpret real-time feedback from the smart shower meter (i.e., when the binary variable $ON_t = 1$) as causing $\theta = 1$.[16]

This specification for attention in a demand model is reduced-form as it does not specify a deeper micro foundation. 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 $E(x|z) = \underbrace{\theta x + \theta u}_{\equiv \theta z} + (1 - \theta_t)x^D$. Thus, the attention parameter can be thought of as the signal-to-noise ratio $\theta = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_{u,t}^2}$, arising from this signal-extraction problem under limited attention. It implies that an individual perceives a one-liter increase in actual water consumption as only a $\theta \leq 1$ liter increase.

Consumption-based habit formation

To study persistence effects of salience-enhancing nudges, we combine the standard model of salience bias above with Stigler and Becker (1977)'s canonical model of consumption-based habit formation. Combining the models involves updating the utility function from (4) to

$$U_t = u(c_t, h_t) - \theta p c_t, \tag{5}$$

where h_t is the *consumption-based habit stock*. We further keep with common applications of Stigler and Becker (1977) and assume a quadratic utility function:[17]

$$u(c_t, h_t) = (a + \gamma h_t)c_t - \frac{1}{2} b c_t^2, \tag{6}$$

where a, b, and γ are parameters. The consumption-based habit stock h_t summarizes past consumption, which, like capital stock, can build up or decay over time

$$h_t = \delta c_{t-1} + (1 - \delta) h_{t-1}, \tag{7}$$

where δ is a parameter that governs the rate of change in the consumption-based habit stock h_t and hence the degree of consumption-based habit persistence. The stock h_t changes more quickly as $\delta \to 1$. The parameter γ in equation (6) governs how much a one-unit increase in h_t changes the marginal utility of consumption.

[16] We discuss parameter identification in Section 4.3 when describing how we empirically implement the model.

[17] Assuming quadratic utility for studying consumption-based habit formation dates back to e.g., Becker and Murphy (1988) or Gruber and Kőszegi (2001), and has more recently been used by Allcott et al. (2022). This utility specification is popular because of its tractibility. It may also be plausible that marginal utility is convex, i.e. that marginal utility of water consumption drops much faster initially than for later units, which is often captured by a constant absolute risk aversion (CARA) utility function. As we show in Appendix D, CARA utility lends itself to a formulation of a consumption-based habit model that exhibits similar predictions regarding consumption responses to salience-enhancing feeback.

In sum, this model views habits as arising from long-term consumption complementarities, which can in turn induce persistence in the effects of real-time feedback causing $\theta = 1$ (i.e., when $ON_t = 1$). Consumption complementarities in water usage can arise, for example, from self-signaling behavior as in Dal Bó and Terviö (2013). In particular, an individual can develop an introspective moral capital (or reputation) over time for "doing good" (e.g., taking shorter showers, which helps the environment), which creates a predisposition to doing good in the future (e.g., continuing to take short showers). Likewise, past wrongdoing (e.g., taking longer showers) destroys moral capital, leading to further wrongdoing.

Attention-based habit formation

Finally, we further allow for the possibility that feedback affects an individual's habitual attentiveness to consumption expenditures. We are motivated by research in neuropsychology, reviewed by Anderson (2016), that documents persistence in attention even after salience-enhancing stimulus is withdrawn. They also show that longer exposure leads to stronger post-exposure persistence in attention. Mirroring Stigler and Becker (1977)'s formulation, we allow current attention to expenditures to depend on past levels of attention. In other words, we consider the case where attention, and not consumption, is the source of state-dependence.

Incorporating state-dependent attention involves two steps. First, we add a time subscript to the attention parameter in (5), updating it to θ_t. Second, we specify θ_t in period t as follows :

$$\theta_t = \begin{cases} 1 & \text{if } ON_t = 1 \\ \omega_t & \text{if } ON_t = 0, \end{cases} \tag{8}$$

When real-time feedback is turned on, i.e. when binary $ON_t = 1$, attention is full: $\theta_t = 1$. When real-time feedback is off, i.e. when $ON_t == 0$, attention is equal to ω_t, the *attention stock*. We formulate an attention stock process which, like the consumption-based habit stock, builds up and decays,

$$\omega_t = \begin{cases} 1 \cdot \alpha_{ON} + (1-\alpha_{ON})\omega_{t-1} & \text{if } ON_t = 1 \\ \theta \cdot \alpha_{OFF} + (1-\alpha_{OFF})\omega_{t-1} & \text{if } ON_t = 0, \end{cases} \tag{9}$$

where $\alpha_{ON}, \alpha_{OFF} \in (0,1)$ govern attention stock build up and decay rates. The stock trends towards 1 when feedback is on, and decays towards a lower bound $\theta \in (0,1)$ when feedback is off.

In summary, our model nests three behavioral mechanisms that can shape how salient feedback affect behavior. If the utility function does not depend on h_t and has $\theta = \theta_t$ for all t as in (4) then the model corresponds to the standard non-dynamic salience framework of Chetty et al. (2009). Allowing h_t to affect utility as in (5) combines the Chetty et al. (2009) and Stigler and Becker (1977) models to allow dynamic consumption responses to salience through state-dependent consumption complementarities per (7). If θ_t is time-varying and evolves per (8) and (9), then state-dependent attention can also shape consumption responses to feedback.

4.2 Optimal consumption and steady state

We assume individuals maximize utility ignoring the impact of current consumption on future habit and attention stocks. This amounts to assuming full projection bias as in Loewenstein et al. (2003).[18] The first order condition that determines the optimal level of consumption in period t is

$$\frac{\partial U}{\partial c_t} = a + \gamma h_t - bc_t - \theta_t p = 0 \Rightarrow c_t = \frac{a + \gamma h_t - \theta_t p}{b}. \tag{10}$$

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

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

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

The model has two steady states of interest. First, consider the *feedback-off* (**OFF**) steady-state, which corresponds to a setting where the individual makes consumption decisions in the absence of feedback for a long time. In our research context, we envisage individuals being in this steady-state at the start of our field experiment, before having a smart shower meter. At this steady-state, salience bias is at its long-run level with $\theta_t = \omega^* = \theta$.

From the first-order condition and the habit and attention stock processes, the steady-state values for consumption, the consumption-based habit stock, and attention stock are given by

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

The second *feedback-on* (**ON**) steady state corresponds to the opposite scenario, namely when the individual makes consumption decisions in the presence of feedback for a long-time. In our field experiment, we anticipate that individuals reach this steady-state after showering with a smart shower meter for months. Here, the individual's consumption costs are fully salient such that $\theta_t = 1$. In this steady-state, consumption, the consumption-based habit stock, and attention stock are given by

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

4.3 Empirical model specifications

In this section we describe how we empirically implement the model using our experimental data. Our goal is to compare the quantitative relevance of the consumption-based and attention-based habit mechanisms in our setting. To this end, we consider two restricted versions of the

[18] A common assumption in consumption-based habit models is for individuals to be forward looking and correctly anticipating how their current choices affect future habit stocks and, hence, behavior (Becker and Murphy, 1988; Becker et al., 1994; Gruber and Kőszegi, 2001; Allcott et al., 2022). By contrast, projection bias implies that individuals underestimate the extent to which states change in the future. We consider the limiting case of full projection bias, which amount to the individual maximizing the static utility function (6). We view this assumption as reasonable in our context, as individuals did not know the feedback schedule they were facing, in contrast to other studies (e.g., Hussam et al., 2022).

model that only have one of these two respective mechanisms driving consumption responses to turning feedback on and off. Respectively, we will call these restricted models the *consumption-habit model* and *attention-based habit model*. In what follows, we illustrate how the key parameters underlying these models can be estimated.

Consumption-based habit model

Under consumption-based habit formation we set $\theta_{it} = \theta$ when feedback is turned off for some $\theta \in [0,1]$ and $\theta_{it} = 1$ when feedback is turned on (i.e., when $ON_{it} = 1$).[19] The behavioral equations are the first-order condition that governs individual i's consumption in period t

$$c_{it} = \frac{a_i - \theta p + \gamma h_{it}}{b},$$

and the consumption-based habit stock accumulation and decay process

$$h_{it} = \delta c_{it} + (1-\delta) h_{it-1}.$$

In Appendix C we derive consumption paths for this model for arbitrary histories of feedback. We show that while feedback is on, the difference between consumption in period t, c_t, and the **OFF** steady-state $c^*_{i,\text{OFF}}$ for individual i is

$$c_{it} - c^*_{i,\text{OFF}} = \frac{\theta - 1}{b} p + \frac{\gamma}{b}(h_{it} - h^*_{i,\text{OFF}}), \tag{11}$$

where

$$h_{it} - h^*_{i,\text{OFF}} = \delta(c_{it-1} - c^*_{i,\text{OFF}}) + (1-\delta)(h_{it-1} - h^*_{i,\text{OFF}}). \tag{12}$$

Intuitively, the first part of $c_{it} - c^*_{i,\text{OFF}}$ corresponds to the initial change in consumption relative to its **OFF** steady-state level when real-time feedback is turned on, $\theta_1 = 1$, and salience bias goes away. The second part is the subsequent adjustment in consumption due to the impulse response of the consumption-based habit stock.

Equation (11) highlights the consumption-based habit model's key implication for feedback persistence. An initial push to consumption from real-time feedback reduces consumption by $\frac{\theta-1}{b}p$. This leads to a subsequent drop in the consumption-based habit stock, now affecting the change in consumption in the following period through the second term in (11). In the period after that, consumption drops again, because of this secondary change in the consumption-based habit stock, and so on. Thus, the consumption-based habit model predicts an initial jump, and then a gradual reinforcement that converges to the new steady state $c^*_{i,\text{ON}}$.

By contrast, turning off the feedback leads to the *exact* reverse process: an initial jump to a higher consumption level, and then, through the accumulation effects in the habit stock, a gradual convergence back to $c^*_{i,\text{OFF}}$. With the common choice of a quadratic utility function, the adjustment

[19] We reintroduce an i subscript in this section as we develop the empirical model specifications.

process is perfectly symmetric with reversed signs.[20]

We further show in Appendix C that these feedback-induced dynamics in $c_{it} - c^*_{i,\text{OFF}}$ lends themselves to a recursive formulation which enables us to bring the model to the data. Depending on whether feedback is on ($ON_{it} = 1$) or off ($ON_{it} = 0$), the recursion can be written as

$$c_{it} - c^*_{i,\text{OFF}} = \begin{cases} \left(1+\frac{\gamma}{b}\phi_t\right)\frac{\theta-1}{b}p & \text{where } \phi_t = \delta\left(1+\frac{\gamma}{b}\phi_{t-1}\right)+(1-\delta)\phi_{t-1} \quad \text{if } ON_{it} = 1 \\ \frac{\gamma}{b}\phi_t\frac{\theta-1}{b}p & \text{where } \phi_t = \delta\frac{\gamma}{b}\phi_{t-1}+(1-\delta)\phi_{t-1} \quad \text{if } ON_{it} = 0. \end{cases} \tag{13}$$

We build a (non-linear) estimating equation for the consumption-based habit model from this recursion:

$$y_{is} = \eta_i + \left(ON_{is} + \frac{\gamma}{b}\phi_{is-1}\left(\frac{\gamma}{b}, \delta, \mathbf{ON}_{is}\right)\right) \times \varphi + \nu_s + \varepsilon_{is}, \tag{14}$$

where we update the time index from t to shower s,[21] $\varphi = \frac{\theta-1}{b}p$, and we make the dependence of ϕ_s on $\frac{\gamma}{b}$, δ, and $\mathbf{ON}_{is}$ from (13) explicit. The individual fixed effect η_i has a structural interpretation as it corresponds to household i's baseline **OFF** level of consumption, $c^*_{i,\text{OFF}}$. The other two shocks, ν_s and ε_{is}, are not structural and enter the equation additively. Previewing our discussion of identification below, we assume that these other shocks that might persist through habit stock are orthogonal to feedback-driven persistence from ϕ_s arising from experimentally-manipulated (and hence exogenous) variation in ON_{is}.

Note that we cannot separately identify the underlying γ and b parameters as they enter as a ratio everywhere in (13) and (14). However, δ and $\frac{\gamma}{b}$ are estimable and fully characterize persistence effects under the consumption-based habit mechanism. This result allows us to evaluate the consumption-bsaed habit model in predicting consumption responses to feedback without estimating all the consumption-based habit model parameters.

Asymmetric consumption-based habit model

To our knowledge, no form of asymmetry in consumption-based habit formation has been previously examined. That said, given our evidence in Figure 4 and Table 4 against symmetry per **Hypothesis 1**, we want to give the consumption-based habit model as favorable a treatment as possible relative to the attention-based model. We therefore consider a variation on Stigler and Becker (1977)'s model where we allow for asymmetric speed in the build-up and decay of the habit stock, depending on whether feedback is on or off. While this model adds another parameter for predicting consumption dynamics in response to exogenous feedback, it provides an important comparison in our assessment of behavioral mechanisms below.

[20] In Appendix D, under CARA utility we obtain a nearly identical recursion characterizing how consumption adjusts to feedback over time that is symmetric and linear in the habit stock. Thus, linearity of marginal utility is not necessary to find the result of symmetric feedback effect accumulation and decay when feedback is turned on and off, starting from steady state.

[21] Recall from Section 2.1 that our smart shower meters record shower count and not date because they do not have an internal clock.

In particular, we modify the habit stock accumulation equation to be

$$h_t = \begin{cases} \delta_{ON} c_{t-1} + (1-\delta_{ON}) h_{t-1} & \text{if } ON_t = 1 \\ \delta_{OFF} c_{t-1} + (1-\delta_{OFF}) h_{t-1} & \text{if } ON_t = 0, \end{cases} \tag{15}$$

where $\delta_{ON}, \delta_{OFF} \in (0,1)$ govern the habit stock build-up and decay rates.[22] Estimation of this model is analogous to how we estimate the (symmetric) consumption-based habit model in equations (13) and (14). The only differences are that we respectively replace δ with δ_{ON} and δ_{OFF} when $ON_t = 1$ and $ON_t = 0$ in (13), and ϕ_{is-1} becomes dependent on δ_{ON} and δ_{OFF} in (14).

Attention-based habit model

Shutting down the influence of consumption-based habit formation ($\gamma = 0$), the behavioral equations for the attention-habit model are the first-order condition

$$c_{it} = \frac{a_i - \theta_t p}{b}$$

and the attention stock accumulation and decay process

$$\theta_t = \begin{cases} 1 \\ \omega_t \end{cases} \text{and } \omega_t = \begin{cases} 1 \cdot \alpha_{ON} + (1-\alpha_{ON}) \omega_{t-1} & \text{if } ON_t = 1 \\ \theta \cdot \alpha_{OFF} + (1-\alpha_{OFF}) \omega_{t-1} & \text{if } ON_t = 0, \end{cases}$$

where recall the attention stock can build and decay at different rates depending on whether real-time feedback is on or off.[23]

In Appendix C, we also derive consumption paths for the model when feedback is turned on and off for an arbitrary history of feedback. Importantly, we show that while feedback is on, the difference between consumption in period t c_t and its **OFF** steady-state level c_i^* is

$$c_{it} - c^*_{i,\text{OFF}} = \begin{cases} \frac{\theta-1}{p} & \text{if } ON_t = 1 \\ \lambda_t(\alpha_{ON}, \alpha_{OFF}) \frac{\theta-1}{p} & \text{if } ON_t = 0 \end{cases} \tag{16a}$$

$$\lambda_t(\alpha_{ON}, \alpha_{OFF}) = \begin{cases} \alpha_{ON} + (1-\alpha_{ON}) \lambda_{t-1} & \text{if } ON_t = 1 \\ (1-\alpha_{OFF}) \lambda_{t-1} & \text{if } ON_t = 0. \end{cases} \tag{16b}$$

Equation (16a) highlights the key dynamics for persistence: the feedback intervention leads to a stable response while the feedback is on (because no attention is required), and the attention weight, $\lambda_t(\alpha_{ON}, \alpha_{OFF})$, builds according to (16b), starting from $\lambda_0 = 0$. When feedback is turned off, the attention stock becomes relevant for consumption, and gradually declines over time.[24]

[22] Keeping with the Dal Bó and Terviö (2013) self-signaling interpretation of consumption complementarities from above, if, for example $\delta_{ON} > \delta_{OFF}$, then households' moral capital grows quicker from taking shorter showers when feedback is on than the destruction of moral capital from taking longer showers when feedback is off.

[23] Note that we do not need different attention stock build-up and decay parameters to predict asymmetric consumption responses to feedback being turned on and off. Allowing such differential build-up and decay puts the consumption and attention-based habit models on equal footing in terms of their number of parameters. As a robustness check below, we report attention-based habit model estimates under the constraint $\alpha_{ON} = \alpha_{OFF}$.

[24] These recursions hold under our assumption of full projection bias, i.e. the individual fully ignoring the impact of her choices on future habit stocks. Similar dynamics and symmetry arise for the forward-looking version of the model, as our Monte-Carlo simulations for the forward-looking model show (not reported). Thus, the qualitative feature of an initial jump followed by gradual adjustment, which is crucial to distinguishing between attention and

From this recursion, we can specify a non-linear estimating equation for this model specification. Translating the model's time index t to our data's shower index s and consumption c_{it} to our shower water usage variable for household i y_{is}, the estimating equation for the attention-based habit model from the recursion in (16a) and (16b) is

$$y_{is} = \eta_i + (ON_{is} + OFF_{is} \times \lambda_{is-1}(\alpha_{ON}, \alpha_{OFF}, \mathbf{ON}_{is}))\,\varphi + \nu_s + \varepsilon_{is}, \tag{17}$$

where $\varphi = \frac{\theta-1}{p}$. Notice we make the dependence of λ_s on α_{ON}, α_{OFF}, and $\mathbf{ON}_{is}$ from (16b) explicit, where the vector $\mathbf{ON}_{is}$ contains the sequence of ON_{ik} values for household i for $k \leq s$.

4.4 Identification and estimation

How does randomization of feedback on and off cycles, and hence ON_{is}, help to identify the consumption-based and attention-based habit models? Regarding the consumption-based habit model, as mentioned above, the η_i fixed effect in its estimation equation (14) corresponds to individual i's baseline **OFF** steady state consumption $c^*_{i,\text{OFF}}$. We then treat ν_s and ε_{is} as non-structural shocks orthogonal to the process by which real-time feedback affects consumption through attention stock build-up and decay. Given this orthogonality assumption, exogenous variation in ON_{is} from our experiment allows us to identify φ, γ, and δ. In particular, φ is identified by the immediate shifts in consumption when feedback is turned on (off). The parameter γ is identified by the extent to which the new long-run consumption levels differ from the immediate shifts. The parameter δ_{ON} is identified by the speed with which the change in consumption nears the predicted long-run treatment effect when feedback initially turned on (i.e., conditional on φ). Similarly, the δ_{OFF} parameter is identified by the rate of decay back towards the baseline level of water use after feedback had been turned off.

Given this orthogonality assumption, exogenous variation in ON_{is} from our experiment allows us to identify φ, γ, and δ. In particular, φ is identified by the immediate shift in consumption when feedback is turned on. The parameter γ is identified by the extent to which the new long-run consumption level with feedback differs from the immediate shift. Similarly, the δ_{OFF} parameter is identified by the rate of decay back towards the baseline level of water use after feedback had been turned off.

Identification of the attention-based habit model parameters φ, α_{ON}, α_{OFF} in equation (17) follows a similar argument. In particular, φ is identified by the immediate shift in consumption when feedback is turned on across conditions T1–T7. The α_{ON} parameter, which recall governs how fast the attention stock grows, is identified by the extent to which feedback effects are more persistent in experimental conditions with longer feedback on cycles (e.g., as we illustrated in Figure 5). Exogenous variation in feedback-on duration *across* conditions T3–T7 is thus crucial

consumption-based habit persistence mechanisms, is preserved with forward-looking households.

for identifying α_{ON}. The α_{OFF} parameter is identified by the rate at which consumption converges back to baseline levels after feedback is turned off, conditional on the length of a given condition's previous feedback-on cycle.

Estimation and inference

We estimate the parameters in equations (14) and (17) by two-step non-linear least squares. Consider first the attention-based habit model. With the consumption-based habit model, conditional on the habit stock parameters $\frac{\gamma}{b}$ and δ (or δ_{ON} and δ_{OFF} in the asymmetric model), the remaining parameters in (14) can be estimated by OLS after we forward-simulate $\phi_{i,s-1}$ for each individual given $\frac{\gamma}{b}$ and δ (or δ_{ON} and δ_{OFF}). We then obtain estimates $\widehat{\frac{\gamma}{b}}$ and $\hat{\delta}$ (or $\hat{\delta}_{ON}$ and $\hat{\delta}_{OFF}$) by finding the values that minimize the sum of squared residuals with OLS finding the remaining parameters in (14) given $\widehat{\frac{\gamma}{b}}$ and $\hat{\delta}$ (or $\hat{\delta}_{ON}$ and $\hat{\delta}_{OFF}$). Likewise with the attention-based habit model, given a candidate $(\alpha_{ON}, \alpha_{OFF})$ pair, we forward-simulate λ_{is-1} and estimate the remaining parameters in (17) by OLS. We then obtain $\hat{\alpha}_{ON}$ and $\hat{\alpha}_{OFF}$ estimates by searching for the $(\alpha_{ON}, \alpha_{OFF})$ pair that minimizes the sum of squared residuals.

The two-step non-linear optimization routines for both models are stable and rapidly converge to global optima. For inference, we use cluster bootstrap standard errors and hypothesis tests from Cameron et al. (2008), sampling at the household level (our level of randomization) with 1000 bootstrap samples. Doing so accounts for our two-step non-linear least-squares procedure as well as household-level persistence in ε_{is} in equations (14) and (17).

4.5 Parameter estimates and within-sample fit

Table 5 presents the parameter estimates for the consumption-based and attention-based habit models. A comparison based on the RMSE of columns (1) and (2) rejects the symmetric consumption-based habit model in favor of an asymmetric model. We obtain a noisy estimate of $\hat{\delta} = 0.087$ in column (1). In contrast, in column (2), the estimator drives the point estimated of $\hat{\delta}_{ON}$ to its boundary 1.[25] We therefore constrain the parameter $\delta_{ON} = 1$ for our preferred specification. Doing so, we obtain a precisely estimated $\hat{\delta}_{OFF} = 0.058$, and also substantially lower standard errors for the other parameters.

For the attention-based habit model, columns (3) and (4) also reveal asymmetry in attention stock build-up and decay. In column (3) we reject $H_0 : \alpha_{ON} = \alpha_{OFF}$ in favor of $H_1 : \alpha_{ON} \neq \alpha_{OFF}$ with $p < 0.01$. Therefore, our preferred specification for the attention-based habit model has asymmetric attention stock build-up and decay.

We then compare the two asymmetric models. We compute the within-sample Root Mean

[25] Figure E.1 in Appendix E confirms that $\hat{\delta}_{ON} = 1$ indeed yields a global minimum for the RMSE.

Table 5: Consumption-Based and Attention-Based Habit Model Parameter Estimates

	Consumption-Based Habit Model		Attention-Based Habit Model	
	(1)	(2)	(3)	(4)
φ	-3.491	-3.190	-5.623	-7.424
	(0.349)	(0.349)	(0.448)	(0.656)
		[0.32]		
$\delta_{ON} = \delta_{OFF}$	0.087			
	(0.126)			
δ_{ON}		1.000		
		(0.394)		
δ_{OFF}		0.058		
		(0.383)		
		[0.023]		
γ/b	0.477	0.515		
	(0.099)	(0.397)		
		[0.058]		
$\alpha_{ON} = \alpha_{OFF}$			0.052	
			(0.038)	
α_{ON}				0.081
				(0.025)
α_{OFF}				0.021
				(0.010)
Within-Sample RMSE	3.745	2.824	3.270	2.284

Notes: $N = 86,376$ (household, shower) observations in each sample consisting of 1078 individuals and 555 households. 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. Boostrap standard errors clustered at the household level reported. Standard errors for unconstrained models are in parentheses. The column (2) estimates constrain $\delta_{ON} = 1$ in estimation, which thus does not have a standard error. Constrained standard errors for the other parameters are in brackets. See the text for the calculation of within-sample RMSE.

Squared Error (RMSE) based on our time-varying treatment effects as

$$RMSE = \sqrt{\frac{\sum_{b=4}^{39} (\tilde{\beta}_{b,g} - \hat{\beta}_{b,g})^2}{36}}, \tag{18}$$

where $\hat{\beta}_{b,g}$ is our time-varying treatment effect estimate for 3-shower block b for experimental group g from equation (1) above, and $\tilde{\beta}_{b,g}$ is the model's corresponding prediction for this time-varying treatment effect. This RMSE thus quantifies the ability of a given model specification to predict the path of time-varying treatment effects in panels (a)-(f) of Figure 4.[26]

[26]Importantly, this RMSE abstracts from the household and shower fixed effects in (17). The fixed effects are important for identifying the time-varying treatment effects and improving their precision, but they are not per-se important for comparing the relative performance of the consumption and attention-based habit models in fitting the symmetry and persistence of feedback effects from turning feedback on and off. The RMSE measure in (18) focuses precisely on this key aspect of the models' predictive ability.

The bottom panel of Table 5 shows that the attention-based habit model in column (4) is superior in terms of within-sample fit. In particular, the attention-based habit model in column (4) has a 39% lower RMSE than the symmetric consumption-based habit model in column (1) and a 19% lower RMSE than the asymmetric consumption-based habit model in column (2). We emphasize that the attention-based habit model has better within-sample fit despite having *one less parameter* than the consumption-based habit models.

As a robustness check, Table E.1 in Appendix E includes OFF_{is} variable as a control to all model specifications. Intuitively, this tests whether the model "leaves any persistence on the table" that could be captured by the control. In all specifications of the consumption habit model, the coefficient of OFF is statistically significant, as well as in the symmetric attention habit model. However, for our preferred specification in column (4) of Table 5), the point estimate of the OFF coefficient is small and not distinguishable from zero.

4.6 Predicting treatment effects out-of-sample

Our experimental data allows us to validate the attention-based and consumption-based habit models out-of-sample by: (1) estimating each of the three models on six out of our seven experimental conditions; and (2) comparing each model's predicted time-varying treatment effects to the realized treatment effects for the left out condition. We undertake this validation for each of the three models six times, for each of conditions T2–T7. This validation exercise illustrates each model's sensitivity to the estimation sample in predicting treatment effects out-of-sample. A model that contains structural parameters governing behavior should robustly make accurate out-of-sample predictions. Non-robust or inaccurate predictions indicate model misspecification.[27]

Figure 6 visualizes our results by reproducing the estimated time-varying treatment effects from equation (2) for the left-out experimental condition (solid grey line) and each of the three models' out-of-sample predictions for these time-varying treatment effects (darker solid and dashed lines in black, red, and green). Table 6 reports the out-of-sample RMSE and parameter estimates.[28]

Our overall out-of-sample RMSE values in column (1) of Table 6 re-affirm our within-sample RMSE results in Section 4.5. The attention-based habit model's RMSE of 2.405 is 15% smaller than the asymmetric consumption-based habit model's RMSE of 2.837. Both models vastly out-

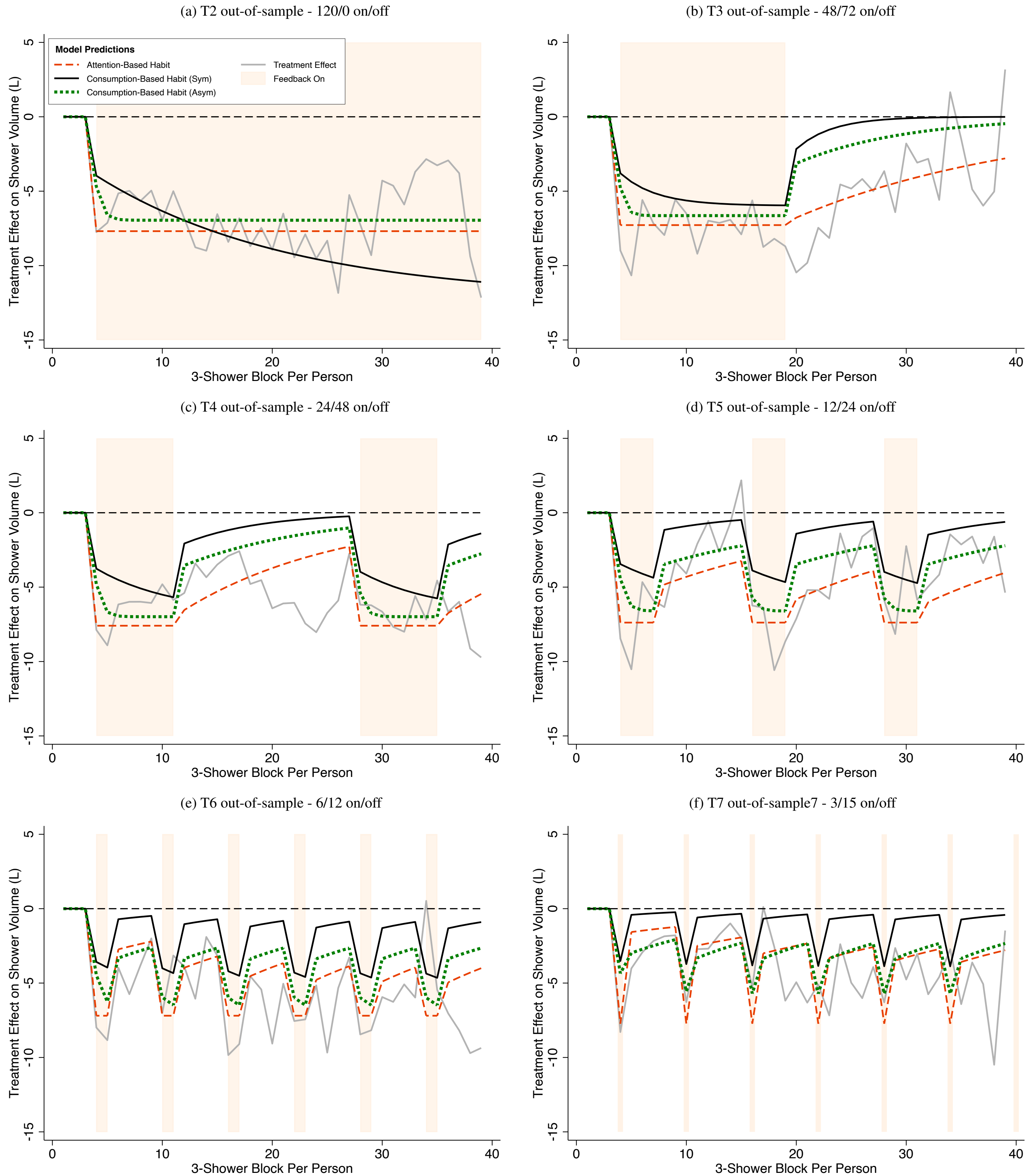
[27]Our validation approach follows Todd and Wolpin (2006) and Ferrall (2013) who use experimental variation from large-scale randomized control trials in labor markets to validate structural models out-of-sample to establish internally validity in identifying policy-invariant (i.e., structural) parameters.

[28]The overall RMSE in the table averages the out-of-sample RMSE across columns (2)-(7) in Table 6:

$$RMSE = \sqrt{\frac{\sum_{g=2}^{7} \sum_{b=4}^{39} (\tilde{\beta}_{b,g} - \hat{\beta}_{b,g})^2}{6 \times 36}},$$

where $\tilde{\beta}_{b,g}$ is the model's predicted consumption in shower block b based on an estimation sample that includes all experimental conditions *except* group Tg (i.e., the left-out condition for validating the model).

Figure 6: Predicted and Actual Treatment Effects by Experimental Condition for the Consumption-Based and Attention-Based Habit Models



Notes: See Figure 3 for details on the experimental design and equation (2) and associated discussion in the text for the regression equation used to generate these plots. For brevity, confidence intervals are not displayed.

Table 6: Out-of-Sample Validation for the Consumption-Based and Attention-Based Habit Models

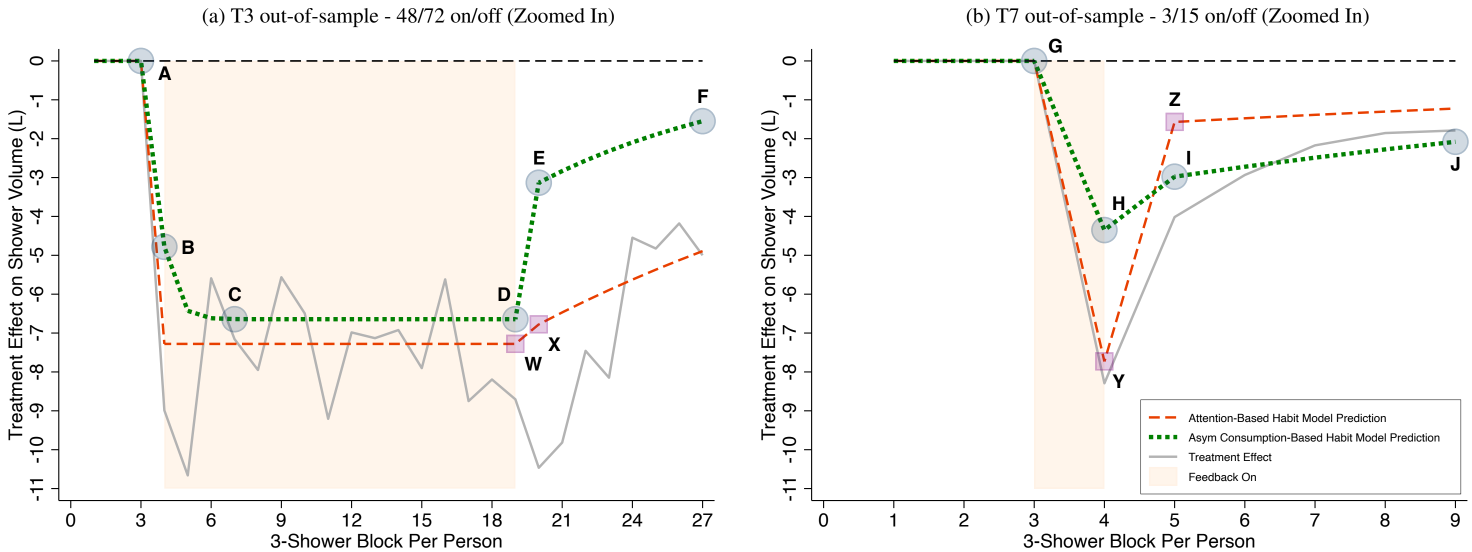
		Leave-Out Experimental Condition for Validation					
		T2 0/120 on/off	T3 48/72 on/off	T4 24/48 on/off	T5 12/24 on/off	T6 6/12 on/off	T7 3/15 on/off
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attention-Based Habit Model							
ϕ		–7.686	–7.279	–7.593	–7.382	–7.194	–7.731
α_{ON}		0.080	0.065	0.092	0.090	0.081	0.076
α_{OFF}		0.019	0.015	0.023	0.019	0.023	0.021
Out-of-sample RMSE		2.474	2.144	2.297	2.350	2.689	2.440
Overall RMSE	2.405						
Consumption-Based Habit Model (Symmetric)							
ϕ		–3.807	–3.573	–3.615	–3.332	–3.444	–3.366
δ		0.060	0.162	0.088	0.078	0.081	0.090
$\frac{\gamma}{b}$		0.689	0.402	0.468	0.479	0.490	0.498
Out-of-sample RMSE		3.573	4.156	4.025	3.198	4.765	3.791
Overall RMSE	3.948						
Consumption-Based Habit Model (Asymmetric)							
ϕ		–3.248	–3.405	–3.351	–3.071	–3.090	–2.951
δ_{ON}		1.000	1.000	1.000	1.000	1.000	1.000
δ_{OFF}		0.052	0.064	0.057	0.045	0.056	0.066
$\frac{\gamma}{b}$		0.533	0.487	0.520	0.535	0.530	0.557
Out-of-sample RMSE		2.387	3.347	3.143	2.183	3.313	2.405
Overall RMSE	2.837						
Reduced-Form Regression							
ON		–7.697	–7.331	–7.481	–7.350	–7.369	–7.570
postON		0.027	0.009	0.009	0.012	0.009	0.011
OFF		–4.941	–4.366	–4.876	–5.004	–4.970	–5.297
postOFF		0.083	0.072	0.092	0.079	0.082	0.090
Out-of-sample RMSE		2.541	2.773	2.842	2.232	2.645	2.445
Overall RMSE	2.588						
Households within-sample		914	920	900	934	942	928
Households out-of-sample		108	102	103	117	94	84
Total Households		1022	1022	1003	1051	1036	1012

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. For the sake of brevity in exposition, bootstrap standard errors are not reported. See the text for the calculation of out-of-sample and overall RMSE.

perform the symmetric consumption-based habit model which has an RMSE of 3.948.

Figure 6 allows us to investigate what drives the differences in the RMSEs across the three models. As expected, the symmetric consumption-based model fails to predict asymmetric consumption responses to feedback being turned on and off. This finding is evident in Figure 6(b) for T3: consumption does not fall fast enough with feedback on and rises too fast with feedback off.

Figure 7: Predicted and Actual Treatment Effects for the First Feedback Cycle for Experimental Conditions T3 and T7 for the Asymmetric Consumption-Based and Attention-Based Habit Models



We see similar patterns in the other conditions with shorter feedback on durations. In addition, the model predicts too high an overall level of consumption. This overprediction occurs because the model's symmetry forces its prediction error-minimizing path to tradeoff the extent to which it predicts the drop in consumption in feedback-on periods and the return in consumption to baseline in feedback-off periods.[29]

The asymmetric consumption-based habit model yields a large improvement over to the symmetric consumption-based habit model in Table 6 and Figure 6. Driving $\hat{\delta}_{ON} = 1$ enables the model to match sharp reductions in consumption when feedback is turned on. The $\hat{\delta}_{OFF} \approx 0.058$ estimate allows the model to better capture the sluggish return to baseline in the periods after feedback is initially turned off, especially in conditions T5–T7. Moreover, panels (e) and (f) of Figure 6 show that the model better predicts the overall price level after feedback starts cycling on and off, as well as the reductions and rebounds in consumption in conditions T6 and T7 compared to the symmetric consumption-based habit model.

However, the asymmetric consumption-based habit model has two key shortcomings. First, Figure 6(b) reveals that the model predicts a large jump in consumption once feedback is turned off after a long cycle of feedback-on, whereas the data exhibits a slow rise in consumption at

[29]More specifically, suppose that the model successfully predicts the large and sharp drops in consumption when feedback is turned on. In this case, the model would predict a symmetric large and sharp jump in consumption when feedback is turned off, creating a large squared prediction error as consumption gradually returns to baseline when feedback is off in the data. In minimizing the squared prediction error, we thus obtain an estimated symmetric consumption-habit model that introduces some, but not extreme error in predicting symmetric drops in consumption with feedback on and jumps in consumption with feedback off. Because these predictions imply, relative to what is observed, a drop in consumption with feedback on that is too small and return to baseline in consumption with feedback off that is too fast, we end up with predicted overall consumption levels that are too high.

those times. Figure 7(a) zooms in on the feedback cycle in condition T3, which applies a long 48-shower duration of feedback-on. Under this condition, the model reaches the **ON** steady state, represented by the flat line between **CD**. The asymmetry created by the different $\hat{\delta}_{ON}$ and $\hat{\delta}_{OFF}$ values can be seen by the differences in how consumption adjusts between **BC** and **EF**. But the model also predicts instantaneous symmetric jumps between **AB** (feedback-on) and **DE** (feedback-off). In steady state, lagged differences in consumption and the consumption-habit stock are zero, so consumption has to change by the full salience effect when feedback is turned on or off.

Figure 7(b) highlights the second empirical challenge for the asymmetric consumption-habit model. The figure zooms in on the first cycle of the shortest 3-shower feedback-on treatment from Figure 6(f), expanding the x-axis to highlight the dynamics. The figure shows that if individuals receive a short sequence of feedback then the consumption-based habit stock does not have enough time to reach the **ON** steady state. In this case the post-feedback jump up in **HI** does not equal the instantaneous salience effect of feedback in **GH**.[30] The predicted magnitude of the jump in **HI** when the habit stock has not yet reached its **ON** steady state depends the change in the habit stock h_t between **GH**. Intuitively, the full salience effect from removing feedback at **H** is tempered, with a lag, by the state-dependence in consumption associated with **GH**. Therefore, **HI** is much smaller than **GH**. Comparing across our experimental conditions, we see that the asymmetric consumption-habit model predicts smaller post-feedback jumps the shorter the duration of feedback-on, which goes against **Hypothesis 2** and what we observe empirically.

In contrast, the attention-based habit model robustly predicts time-varying treatment effects out-of-sample. In feedback-on periods, it successfully predicts immediate and stable changes in consumption in response to salient feedback. And in feedback-off periods, asymmetric attention stock build-up (through $\hat{\alpha}_{ON}$) and decay (through $\hat{\alpha}_{OFF}$) successfuly predicts two key features of post-feedback consumption: (1) *differential jumps* in consumption across conditions when feedback is turned off as a function of prior feedback-on duration; and (2) *gradual trends* in consumption back to baseline if feedback is off for sufficiently long.

To see why the $\hat{\alpha}_{ON}$ parameter estimate enables the model to predict *differential jumps*, compare the first feedback cycles in conditions T3 and T7 from Figure 7. In panel (a), after 48 showers of feedback, the individual's attention stock ω_t is near the steady-state level of $\omega^* = 1$. Therefore, when feedback is turned off at **W** and θ_t switches from one to the value of ω_t the attention-based habit model predicts a minimal change in consumption at **WX**. At the other extreme in panel (b) there are only three showers of feedback-on, and hence the attention stock ω_t is closer to its initial value $\theta << 1$ when feedback is turned off at **Y**. Therefore, when feedback is turned off and θ_t

[30]The small differences in the instantaneous salience effects predicted by the asymmetric consumption-habit model between **AB** and **GH** in panels (a) and (b) of Figure 7 is due to the model being estimating on different sub-samples across the figures, per our out-of-sample validation approach. Panels (a) and (b) respectively exclude conditions T3 and T7 from the estimation sample.

switches from one to ω_t, there is a large change in the individual's attention to consumption costs, and hence a large change in consumption at **YZ**. In sum, the differential predicted jumps at **WX** and **YZ** illustrate how the duration of feedback-on and attention stock growth rate, which $\hat{\alpha}_{ON}$ governs, enables the attention-habit model to reconcile the feedback-persistence gradient in the data per **Hypothesis 2**.

Regarding *gradual trends*, conditional on $\hat{\alpha}_{ON}$ and associated post-feedback jumps in consumption, our estimator finds the value of $\hat{\alpha}_{OFF}$ that gradually reduces the attention stock when feedback is off to match the decay in post-feedback treatment effects in the data. Such gradual decay, combined with the model's prediction of immediate and stable changes in consumption with feedback-on goes against **Hypothesis 1** and aligns with the asymmetry observed in the data.[31]

Summary

In summary, a model that combines standard behavioral models of limited attention and consumption-based habit formation is unable to predict the micro-dynamics of how consumption responds to feedback from our field experiment. However, by switching the model's source of state dependence to attention in an intuitive way that is inspired by Stigler and Becker (1977)'s original habit-formation mechanism and in-line with findings from neuropsychology, we obtain an attention-based model of habit formation that accurately and robustly predicts consumption's dynamic response to feedback. These out-of-sample validation results indicate that the attention-habit model captures structural econometric primitives that drive individual behavior.

Reduced-form model comparison

As a final validation exercise, we explore how the (non-linear) attention-based habit model compares to the (linear) treatment effect regression in equation (2) in terms of out-of-sample fit. The bottom panel of Table 6 produces an analogous set of out-of-sample RMSE results based on our equation (2) estimates. In terms of overall RMSE, the attention-based habit model's is 7% lower than the regression's. However, we find a large, 23% difference in column (2) when the models predict feedback effects with long feedback on and off durations. These findings highlight the value of the attention-based habit model in both predicting and interpretting non-linear decay of feedback effects over longer horizons after a long duration of feedback being on.

4.7 Other persistence mechanisms

Overall, our reduced-form and structural analyses together establish the quantitative relevance of time-varying attention as a key behavioral mechanism for persistent feedback effects. Before

[31]Notably, the attention-based habit model also captures treatment effect 'action and backsliding' that diminishes with repeated feedback nudges, similar to that found by Allcott and Rogers (2014) in their study of repeated home energy reports. These patterns are particularly visible in panels (e) and (f) of Figure 6, and our framework provides an explanation for these patterns. In particular, with attention stocks that asymmetrically grow quickly and decay sluggishly, there exist spillovers in individuals' attention stocks across feedback cycles.

examing associated implications for policy, we consider and rule out two additional potential mechanisms for feedback persistence in our setting.

Automatic control

Camerer et al. (2020) introduce automatic control models from neuropsychology as a behavioral mechanism for habit formation. In their model, individuals use a fixed decision rule that generates predicted and realized levels of consumption utility at any point in time. As long as prediction errors are not too large, an individual will continue to use the rule. If, however, predicted utility starts departing from the realized utility, evidence begins to accumulate against the behavioral rule. Should evidence against the current rule becomes sufficiently large, individuals will, at a utility cost, re-optimize and update their fixed decision rule, resulting in a discrete shift in behavior that persists thereafter.

In our experiment, the introduction and removal of feedback could create differences between predicted and realized utility from showers of different lengths, inducing re-optimization with discrete changes in shower usage. However, in Appendix F.1 we examine and fail to find discrete shifts in consumption at the individual level after feedback is turned off. Moreover, allowing for such jumps in consumption in our econometric models leaves our persistence effects estimates in Section 3.2 unaffected. In sum, Camerer et al. (2020)'s automatic control model does not characterize persistence observed in our setting.

Experimentation and learning

We also consider the potential for feedback-induced experimentation and learning. Here, we have in mind persistence in behavioral change like that studied by Larcom et al. (2017) following a 2014 London Tube strike. The strike temporarily forced individuals to experiment with and learn about other forms of public transit. Some individuals stayed with new forms of transit after the strike ended, underlining experimentation and learning as a potential mechanism for persistence.

In Appendix F.2, we conduct another series of auxiliary analyses using the feedback on/off cycles from conditions T4–T7. We test for a permanent long-run shift in shower consumption after individuals experience their first feedback cycle. Intuitively, the consumption response to the first cycle potentially entails both salience and experimentation and learning effects, whereas subsequent cycles primarily have salience effects. A learning mechanism implies that we should find a permanent level shift in consumption after the first cycle. However, we find no evidence of such permanent shifts in our data.

5 Attention stocks, behavioral change, and intervention design

In light of our results in Section 4, we move forward with the attention-based model of habit persistence to provide model-grounded estimates of habit formation and decay rates, and examine implications for the design of feedback interventions.

5.1 Asymmetric habit formation and decay

We examine how attention stock build-up and decay affects consumption over time in panels (a) and (b) of Figure 8. In particular, panel (a) plots the point estimate and 95% bootstrap confidence interval for the persistence effect from feedback starting from the (baseline) **ON** steady-state and turning feedback on for 120 consecutive showers.[32] We estimate that it takes 18 showers to reach the **ON** steady state. This duration is identified where the 95% confidence interval for the persistence effect of feedback contains the **ON** steady-state persistence effect of -7.35 L/shower.[33] The half-life for feedback persistence effect build-up relative to its **ON** steady-state level is just nine showers [95% CI 6, 15].

Panel (b) of Figure 8 examines attention stock decay and the common question of "how long do these habits last?". Here, we plot the point estimate and 95% confidence interval for the persistence effect of feedback starting from the **ON** steady-state and keeping feedback off for 240 consecutive showers. Panel (b) shows that we are not able to reject persistence effects greater than -0.037 L/shower (e.g., 5% of the **ON** steady-state persistence effect of -7.35 L/shower) at the 5% level of significance after 59 showers.[34] In other words, if an individual takes one shower per day, starting from the **ON** steady state, we estimate that shower water-conserving attention stock effects persist for approximately two months.

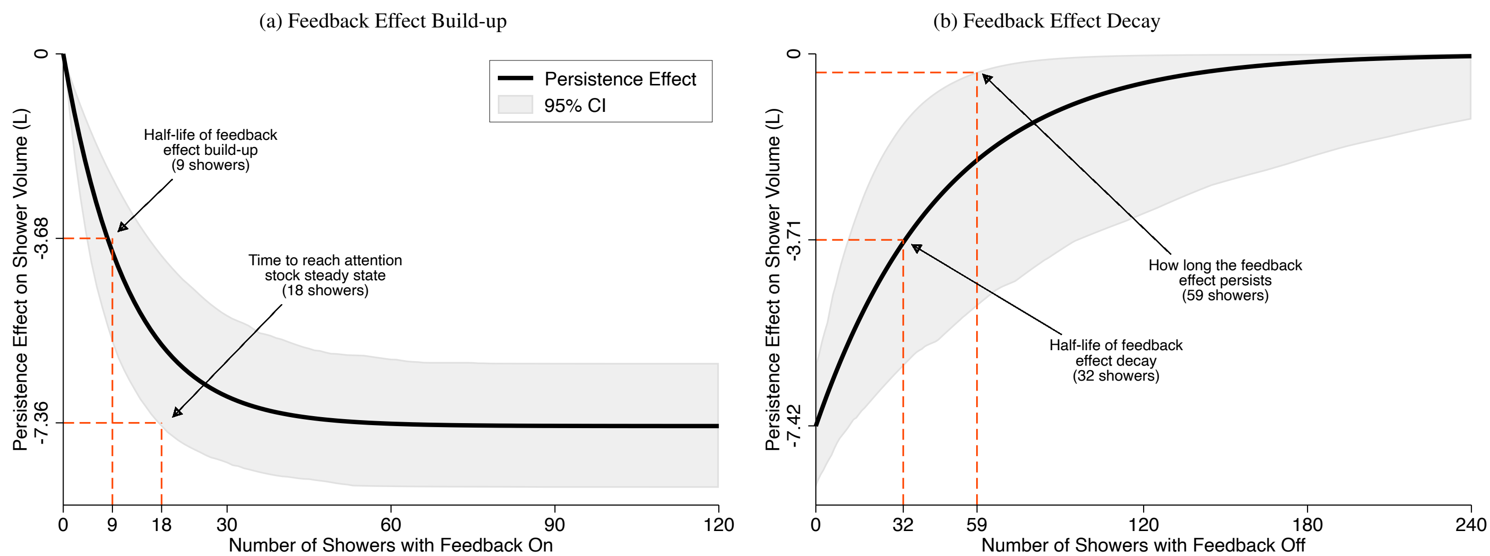
Panels (a) and (b) of Figure 8 together highlight asymmetry in attention stock and persistence effect build-up and decay. Whereas it takes 19 showers for the persistence effect converge to its **ON** steady state level while feedback is on, it takes 59 showers without feedback for the persistence effect to dissipate. Likewise, while the half-life for persistence effect build-up is nine showers [95% CI 6, 15], the half-life for persistence effects in panel (b) is 33 showers [95% CI 15, 95]. In sum, attention stocks and associated consumption effects build quickly and decay sluggishly.

[32] Formally, the persistence effect of feedback from equation (17) is $\lambda_{is-1}(\alpha_{ON}, \alpha_{OFF}, ON_{is})\varphi$. This effect quantifies an individual's change in consumption relative to their baseline level immediately after feedback is turned off. For example, starting from the **OFF** steady state, after nine showers of feedback, an individual exhibits 3.68 L/shower water savings relative to their **OFF** steady-state level if feedback is suddenly turned off. Panel (a) of Figure 8 illustrates this persistence effect from nine showers of feedback.

[33] We identify the steady-state persistence effect at the point where $\omega_t - \omega_{t-1} < 0.01$, that is, when the predicted change in the attention stock based on our point estimates is less than 1% of the maximal attention stock value of one. Using confidence intervals to identify habit persistence effect build-up and decay in panels (a) and (b) of Figure 8 aligns with reduced-form approaches for identifying the duration of habit persistence. See, for example, Charness and Gneezy (2009) or Royer et al. (2015).

[34] We require a threshold rule for classifying "how long habits last" because the attention-based habit model's persistence effect never reaches exactly 0.

Figure 8: Rates of Feedback Effect Build Up and Decay



5.2 Feedback duration and persistence: simulation analysis

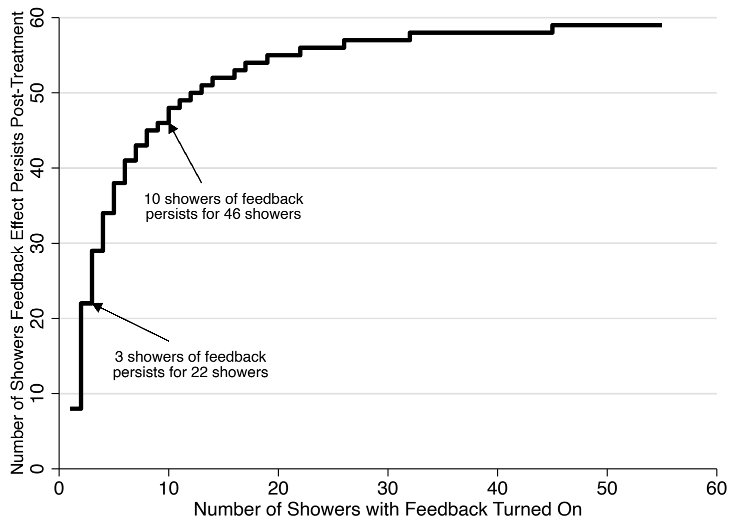
Our structural model allows us further to explore the relationship between feedback duration and consumption persistence. Following on our approach to quantifying "how long habits last" from panel (b) of Figure 8, we conduct a series of simulations, starting from the **OFF** steady-state, where we: (1) provide feedback for k periods; (2) permanently turn feedback off; and (3) determine how many periods it takes until we no longer statistically reject a persistence effect of -0.037 L/shower at the 5% level of significance. Figure 9 plots the results from these simulations, with $k \in \{1, \dots, 55\}$ on the horizontal axis and the number of showers that a habit persists on the vertical axis. In line with the short half-life for persistence effect build-up, the figure shows attention-based persistence effects form rapidly. For example, our estimates imply that feedback effects persist for up to 22 showers (e.g., $\approx$ three weeks) after feedback is turned off in our least intensive feedback condition T7 with only three periods of feedback (e.g., $\approx$ three days).

Figure 9 also highlights a diminishing impact of feedback duration on persistence. The gradient is relatively steep up to $k = 15$ showers of feedback and flattens after that. In other words, feedback is valuable for inducing persistent behavioral change up to some point, beyond which the marginal gains diminish substantially. What does diminishing marginal impacts of feedback on behavior change imply for the design of feedback interventions? We now turn to addressing this question.

5.3 Feedback intervention design and behavioral change

In this final part of our structural analysis, we explore how feedback intervention design affects total behavioral change and design implications of attention stock build-up and decay. The emer-

Figure 9: Feedback Duration and Post-Feedback Persistence – Simulation Results



gence of repeating, salience-enhancing feedback enabled by data and digital technology motivates this analysis.

Specifically, we use our model to inform the following question: suppose a policymaker designs a T period feedback intervention facing a $K < T$ period feedback budget; how should they allocate the feedback over time to maximize behavioral change from the intervention?[35] A feedback budget can emerge if it is technologically infeasible to provide individuals with real-time feedback at all times or if implementers worry about consumer backlash from constant feedback. While feedback in our particular context is not overtly costly, our attention-based persistence mechanism may be relevant elsewhere. Our structural model allows us to investigate how different optimal policies can look when our attention-based mechanism shapes behavior in the presence of costly feedback.[36]

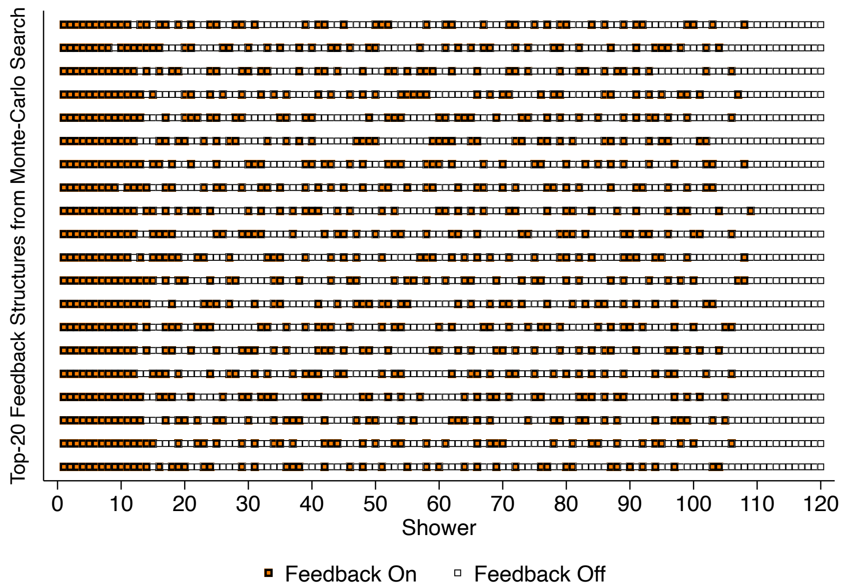
Consider designing a $T = 120$ intervention under a $K = 48$ period feedback budget. This set-up corresponds to experimental conditions T3-T5, which have 48 periods in total with feedback is turned on.[37] For non-negligible K and T values, finding the feedback sequence that maximizes behavioral change faces a Curse of Dimensionality. For instance, in our set-up, there are more

[35]This policy objective is, of course, different from welfare maximization, particularly if individuals face utility costs from feedback (Allcott and Kessler, 2019). We do not examine welfare maximization because we do not identify all of the utility function parameters in equations (4) and (6). In practice, however, maximizing behavioral change given a feedback budget is often the goal of feedback interventions, as was the case with our research partner.

[36]Separately, we also do not consider "attention budgets" in our analysis, whereby attention drawn toward the costs of one behavior (shower water usage) reduces the amount of attention an individual has to expend on other decisions during the day. See Shenhav et al. (2017) or Bronchetti et al. (2022) for overviews of research from neuropsychology on attention spillovers across behaviors which, to our knowledge, has not been formalized and empirically examined within an economic framework. Our counterfactual simulations complement this emerging research area by examining the implications of attentional spillovers within a behavior (showering) across time and not across behaviors.

[37]Working with $K = 48$ and $T = 120$ ensures our counterfactuals stem from parameter estimates identified from experimental variation in feedback and consumption from a similar intervention setting.

Figure 10: Top-20 Treatment-Effect Maximizing Feedback Cycles
(from 1 trillion randomly-generated feedback cycles)



than 1×10^{95} permutations of feedback sequences to search over in finding the behavior change-maximizing feedback structure.

We therefore undertake a Monte-Carlo study to guide our analysis. Specifically, we compute the average per-period feedback effect predicted by our estimated attention-based habit model for one trillion randomly-generated feedback on/off sequences. Figure 10 presents the top-20 treatment effect-maximizing feedback structures from these simulations. The figure reveals a particular class of feedback structures involving an initial continuous sequence of feedback followed by regular and intermittent feedback on impulses. Interestingly, these feedback structures can be implemented by an (S,s)-type rule (Scarf, 1959), but which also involves an initial attention stock build-up phase. We label our modified rule (I,S,s), with I defining the number of consecutive periods feedback is initially provided to build-up an individual's attention stock. Formally, the (I,S,s) rule is defined as

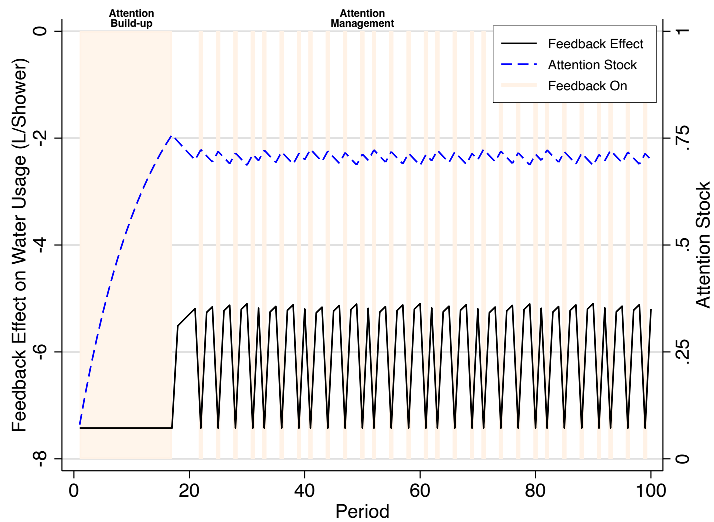
$$ON_t = \begin{cases} 1 & \text{if } t \leq I & \text{(attention build-up)} \\ 1 & \text{if } t > I \text{ and } [(\omega_t < s) \text{ or } (ON_{t-1} == 1 \text{ and } \omega_t < S)] & \text{(attention management)} \\ 0 & \text{otherwise.} \end{cases}$$

Constraining our search to (I,S,s)-based feedback structures, we can solve for the feedback-effect maximizing rule using a grid search. Based on our model point estimates, the optimal rule is not unique with 50 slightly different $(I,S,s) \in \{\{15\} \times \{0.701, 0.702, \ldots, 0.719\} \times \{0.700, 0.701, 0.702\}\}$ generating a maximal per-period feedback effect of -5.97 L / shower. However, each of these rules implement a unique optimal feedback structure, which the shaded areas in Figure 11 illustrate. The figure also highlights the corresponding optimal attention build-up and management phases.

Figure 11: Optimal Consumption and Attention Stocks

$T = 120$ Intevention Period, $K = 48$ Feedback Budget

Optimal Rule $(I,S,s) = (17, 0.74, 0.7)$

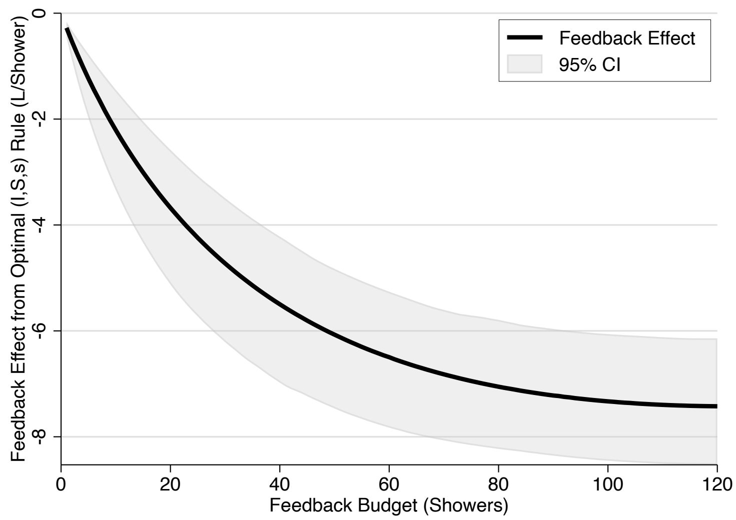


To visualize how the (I,S,s) rule works, Figure 11 plots the unique optimal consumption and attention stock paths implemented by one of the optimal (I,S,s) rules. The figure shows the attention stock ranges between 0.700 and 0.719. In other words, it is optimal to keep a consumer aware of $\approx 70\%$ of their consumption costs when feedback is turned off on average. In doing so, the optimal (I,S,s) rule generates an average feedback effect of -5.97 L/shower [95% CI -4.73, -7.38].

A relevant benchmark for these optimal feedback effects is the implied average feedback effect under the optimal feedback structure under Stigler and Becker (1977)'s commonly-used (symmetric) consumption-based habit model with quadratic utility. Their model's symmetry of habit formation and decay implies that a policymaker should immediately exhaust the K period feedback budget to maximize the average feedback effect. Under this rule, if we immediately exhaust the $K = 48$ feedback budget, we obtain an average feedback effect of -5.18 L/shower from our attention-based habit formatiom model, 15% lower than the average feedback effect under the optimal (I,S,s) rule. This contrast in optimal policy rules highlight yet another important difference between the behavioral mechanisms for feedback persistence.

Lastly, how does the treatment effect from the optimal (I,S,s) rule vary with the feedback budget K? Figure 12 answers the question by plotting the average treatment effect across $T = 120$ showers from the optimal (I,S,s) rule and unique feedback structure as a function of the feedback budget. There is a convex relationship implying as K becomes large. For instance, the point estimates imply that increasing a feedback budget from 20 to 40 periods increases the optimal feedback effect by 50%, from -3.67 L/shower to -5.50 L/shower. An equivalent feedback budget increase from 40 to 60 showers increases the optimal feedback effect by 18% to -6.50 L/shower.

Figure 12: How Feedback Effect from the (I,S,s) Rule Varies with the Feedback Budget in a 120 Period Intervention



The figure also reveals similar average feedback effects under 80 and 120-period feedback budgets.

In sum, Figure 12 highlights the importance of thinking in terms of feeback budgets whenever there is a marginal cost of providing or receiving feeback. In a setting where feeback is applied strategically and have diminishing effects in generating behavioral change, even under constant marginal cost, it is *not* optimal to provide feedback all the time.

6 Conclusion

By studying the micro-dynamics of behavioral responses to feedback, this paper has identified a new mechanism for habit formation through state-dependent attention. We designed and implemented a field experiment that uncovered *asymmetry* in behavioral responses to introducing and removing repeated feedback, and a gradient between treatment duration and post-duration *persistence*. Motivated by these reduced-form results, we developed and estimated a structural model of habit formation that nests consumption- and attention-based habit mechanisms, and adds more flexibility to the model specifications. Using the models, we validated the attention-based mechanism and uncovered a new (I,S,s) rule for optimizing feedback-based interventions to create sustained behavioral change.

From a methodological perspective, integrating experimental and structural econometric analyses of habit formation, as we have done, motivates and informs future research. While we have examined a context that abstracts from metabolically-driven persistence, there are feedback interventions that target areas where metabolically-driven persistence may play a role, like programs to reduce smoking or encourage physical activity. Similarly, while we designed our experiment to avoid forward-looking confounds, one could extend our model to interventions with a time profile known to participants, as was done in Hussam et al. (2022). A structural approach, where experimental variation identifies the model, can quantify the relative strengths of competing channels for persistence in behavior change. Such quantitative results, in turn, directly inform the design of nudge-based interventions to help individuals overcome "bad" habits and create "good" ones.

References

Aguirregabiria, Victor, "The Dynamics of Markups and Inventories in Retailing Firms," *Review of Economic Studies*, 1999, *66* (2), 275–308.

Alessandria, George, Joseph P. Kaboski, and Virgiliu Midrigan, "Inventories, Lumpy Trade, and Large Devaluations," *American Economic Review*, 2010, *100* (5), 2304–2339.

Allcott, Hunt and Judd Kessler, "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons," *American Economic Review*, 2019, *11*, 236–276.

_ and Todd Rogers, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, 2014, *104*, 3003–3037.

_ , Matthew Gentzkow, and Lena Song, "Digital Addiction," *American Economic Review*, 2022, *112* (7), 2424–2463.

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.

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

Attanasio, Orazio P., "Consumer Durables and Inertial Behaviour: Estimation and Aggregation of (S,s) Rules for Automobile Purchases," *Review of Economic Studies*, 2000, *67* (4), 667–696.

Baltagi, Badi H. and James M. Griffin, "Rational Addiction to Alcohol: Panel Data Analysis of Liquor Consumption," *Health Economics*, 2002, *11* (6), 485–491.

Becker, Gary S. and Kevin M. Murphy, "A theory of rational addiction," *The Journal of Political Economy*, 1988, pp. 675–700.

_ , Michael Grossman, and Kevin M. Murphy, "Rational addiction and the effect of price on consumption," *American Economic Review*, 1991, *81* (2), 237–241.

_ , _ , and _ , "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 1994, *84* (3), 396–418.

Bó, Ernesto Dal and Marko Terviö, "Self-esteem, moral capital, and wrongdoing," *Journal of the European Economic Association*, 2013, *11* (3), 599–633.

Bronchetti, Erin, Judd Kessler, Ellen Magenheim, and Eric Zwick, "Is Attention Produced Optimally? Theory and Evidence from Experiments with Bandwidth Enhancements," *Econometrica*, 2022. forthcoming.

Bruhin, Adrian, Lorenz Goette, Simon Haenni, and Lingqing Jiang, "Oops! I Did It Again: Understanding Mechanisms of Persistence in Prosocial Behavior," Working Paper, National University of Singapore 2020.

Caballero, Ricardo J. and Eduardo M.R.A. Engel, "Price Stickiness in Ss Models: New Interpretations of Old Results," *Journal of Monetary Economics*, 2007, *54* (1), 100–122.

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.

Cameron, Samuel, "Nicotine Addiction and Cigarette Consumption: A Psycho-Economic Model," *Journal of Economic Behavior & Organization*, 2000, *41* (3), 211–219.

Chaloupka, Frank, "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 1991, *99* (4), 722–742.

Charness, Gary and Uri Gneezy, "Incentives to Exercise," *Econometrica*, 2009, *77*, 909–931.

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

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

Eberly, Janice C., "Adjustment of Consumers' Durable Stocks: Evidence from Automobile Purchases," *Journal of Political Economy*, 1994, *102* (3), 403–436.

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

Ferrall, Christopher, "Explaining and Forecasting Results of the Self-sufficiency Project," *Review of Economic Studies*, 2013, *79* (4), 1495–1526.

Ferraro, Paul J and Michael K Price, "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment," *Review of Economics and Statistics*, 2013, *95* (1), 64–73.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl, "Habit Formation in Voting: Evidence from Rainy Elections," *American Economic Journal: Applied Economics*, 2016, *8* (4), 160–188.

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

Gilbert, Ben and Joshua Graff Zivin, "Dynamic Salience with Intermittent Billing: Evidence from Smart Electricity Meters," *Journal of Economic Behavior & Organization*, 2014, *107* (1), 176–190.

Gruber, Jonathan and Botond Kőszegi, "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

Hussam, Reshmaan, Atonus Rabbani, Giovanni Reggiani, and Natalia Rigol, "Rational Habit Formation: Experimental Evidence from Handwashing in India," *American Economic Journal: Applied Economics*, 2022, *14* (1), 1–41.

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.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems, "The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network," *Quarterly Journal of Eco-*

nomics, 2017, *132* (4), 2019–2055.

Levitt, Steven D and John A List, "What do laboratory experiments measuring social preferences reveal about the real world?," *Journal of Economic perspectives*, 2007, *21* (2), 153–174.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection bias in predicting future utility," *the Quarterly Journal of economics*, 2003, *118* (4), 1209–1248.

Olekalns, Nilss and Peter Bardsley, "Is Addiction "Rational"? Theory and Evidence," *Journal of Political Economy*, 1996, *104* (5), 1100–1104.

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experimental with Workers at a Fortune-500 Company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Scarf, Herbert E., "The Optimality of (S,s) Policies in the Dynamic Inventory Problem," in Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes, eds., *Mathematical Methods in Social Sciences*, Stanford: Stanford University Press, 1959, pp. 196–202.

Shenhav, Amitai, Sebastian Musslick, Falk Lieder, Wouter Kool, Thomas L Griffiths, Jonathan D. Cohen, and Matthew M. Botvinick, "Toward a Rational and Mechanistic Account of Mental Effort," *Annual Review of Neuroscience*, 2017, *40*, 99–124.

Sheshinkski, Eytan and Yoram Weiss, "Inflation and Costs of Price Adjustment," *Review of Economic Studies*, 1977, *44* (2), 287–303.

Steiny Wellsjo, Alexandra, "Simple Actions, Complex Habits: Lessons from Hospital Hand Hygiene," Working Paper, UC Berkeley 2022.

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

Thaler, Richard H and Cass R Sunstein, *Nudge: Improving decisions about health, wealth, and happiness*, Penguin, 2009.

Todd, Petra E. and Kenneth I. Wolpin, "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 2006, *96* (5), 1384–1710.

Wood, Wendy and David T Neal, "Healthy through habit: Interventions for initiating & maintaining health behavior change," *Behavioral Science & Policy*, 2016, *2* (1), 71–83.

World Bank, *World development report: Mind, society, and behavior*, The World Bank, 2014.

Online Appendix

A Supplemental analyses and material

A.1 Sample recruitment, instructions, baseline, and endline surveys

As discussed in the paper, we recruited 700 customers for the trial. The following five steps describe how we recruited these customers from South East Water's customer base.

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 before the experiment, we emailed an online survey to 45,685 randomly selected households from the 140,407. The survey asked about household characteristics, water usage, and shower type. Figure A.1 presents the baseline survey questions.

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

4. We sent a recruitment email and Plain Language Statement (PLS) 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. These households represent our eligible sample. Figure A.2 presents the PLS.

5. From the 1,200 eligible households, we randomly selected 700 for the experiment and randomly allocated 100 households to each experimental group T1–T7. We stratified allocation by household size to prioritize all single-person households. The Amphiro B1s were then mailed to households with the installation instructions in Figure A.3.

6. At the end of the trial, we emailed households an endline survey asking questions related to how households felt the Amphiro B1 affected their behavior, focusing on salience effects and habit formation. Figure A.5 presents the endline survey questions.

All 700 randomly selected households for the trial answered the baseline survey. From Table 2 in the paper, 653 (93%) households had their Amphiro B1's accounted for at the end of the trial, and 555 (79%) returned their devices used with data for extraction. 428 (61%) of households answered the endline survey after the trial.

Figure A.1: Baseline Survey

Baseline 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, knowing a little more about your household's water usage and lifestyle would be useful.

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 these 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.

Thank you for taking the time to complete our survey. Your answers will help shape develop tools to support customers.

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?
   [3 or less, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more]
9. What is your best guess 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 noticing and 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?
    [Home is brand new, 2-5 years, 5-10 years, 10-15 years, 15+ years or never remodeled, Don't know]
18. Which of the following do you have? [You can tick more than one answer.]
    [Balcony garden, Lawn grass, Vegetable garden, 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 commercial car wash, At home with a hose, At home with 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]

End of Survey

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.

Figure A.2: Plain Language Instructions Sent to Households During Recruitment

THE UNIVERSITY OF MELBOURNE

Plain Language Statement

Project Title: Real-time Feedback and Shower Water Usage

You are invited to participate in an anonymous four-month trial study. We are interested in understanding how customers respond to real-time feedback on shower water temperature and water use from Amphiro B1 personal shower displays. Dr David Byrne and Dr Leslie Martin from the Department of Economics at the University of Melbourne are the responsible researchers for this project. They are working with South East Water in conducting this trial study.

The Faculty of Business and Economics Human Ethics Advisory Committee has approved this project. The Energy Markets Program in the Centre of Market Design is providing funding along with a Business and Economics Faculty Research Grant from The University of Melbourne, and South East Water.

What you will need to do

1. [NOW] INSTALL THE AMPHIRO B1. You will need to install your Amphiro B1 in your primary shower in your home. This must be a detachable hose shower. We have included paper instructions that show you how to install the B1. This video also shows you how to install the B1: http://go.unimelb.edu.au/7g66.

When you take your first shower after installing the B1 the display should activate automatically. It will display the current water temperature. As long as you see the temperature, you will know that your B1 is working.

The B1 will always show you water temperature. From time-to-time during the four-month trial period, the B1 may also show you water use, both in litres and via a polar bear animation.

2. [SEPTEMBER] RETURN THE AMPHIRO B1. At the end of the four-month trial period you will need to uninstall the Amphiro B1 and return it to The University of Melbourne so that the researchers can extract the B1 data on shower water usage (in litres). No other data will be collected from the B1. You can uninstall the B1 by following the included paper instructions with the device. To return the device to The University of Melbourne, you simply place it in the self-addressed stamped envelope that we have included with the B1, and put it in your nearest mailbox (Please keep the self-addressed stamped envelope somewhere safe for the duration of the trial) If you wish to keep the Amphiro B1 for free, we will mail it back to you immediately after the shower usage data has been extracted.

3. [SEPTEMBER] FOLLOW-UP SURVEY. After the B1 trial finishes, you will be invited to participate in a follow-up survey to provide feedback on how you used your B1.

How will my confidentiality be protected?

Your name, customer account, and location will be completely de-identified in all future publications that arise from this study. Customers involved in the trial will be labeled anonymously as “Customer 1”, “Customer 2”, and so on. All raw shower usage data and follow-up survey data will be stored securely at The University of Melbourne for a minimum of 5 years. South East Water will also securely store the follow-up survey data. Your participation is completely voluntary and you may withdraw consent to participate at any time.

If, prior to processing the data, you do not wish for your data to be included in the project, you may request to have your data withdrawn from the project’s dataset. Once the research project is completed, the results from the anonymised data may be presented at academic conferences and published in academic journals.

Where can I get further information?

Should you require any further information, or have any concerns, please contact Dr David Byrne (byrned@unimelb.edu.au, 03 8344 3880) or Dr Leslie Martin (leslie.martin@unimelb.edu.au, 03 8344 5312). If you have any concerns about the conduct of the project, you are welcome to contact the Executive office, Human Research Ethics at The University of Melbourne (03 8344 2073 or 03 9347 6739 (fax)).

FHEAC No.: 1544989.1
Date: 30/04/2017

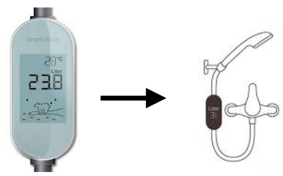
Department of Economics
The University of Melbourne
VIC 3010, Australia
Tel: +61 03 8344 5289
Fax: +61 03 8344 6899
http://www.economics.unimelb.edu.au

Figure A.3: Amphiro B1 Trial and Installation Instructions Mailed to Households

(a) Trial Description

Thank-you for participating in the Amphiro B1 trial. Please find enclosed the Amphiro B1 personal shower display, installation instructions, and details of the trial. Here is a summary of the four-month trial timeline:

Step 1 (Now)
Install the Amphiro B1 in your home's main handheld shower



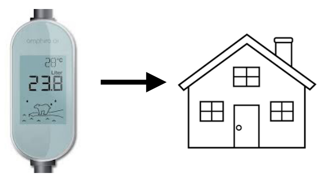
Notes:
- Please keep the original Amphiro B1 box and self-addressed pre-paid envelope somewhere safe for the duration of the trial.
- Installation video:
http://go.unimelb.edu.au/7g66

Step 2 (Sept. 2017)
Mail the Amphiro B1 back to The University of Melbourne

THE UNIVERSITY OF MELBOURNE

Notes:
- We will email you when the trial finishes.
- Please mail back the Amphiro B1 in its original box using the self-addressed pre-paid envelope.

Step 3 (Oct. 2017)
The University of Melbourne mails the Amphiro B1 back to you to keep for free, if you wish



Notes:
- We will record water usage data, reset the Amphiro B1 so it is no longer in 'trial mode', and send it back to you with extra features.

(b) Amphiro B1 Modes During the Trial

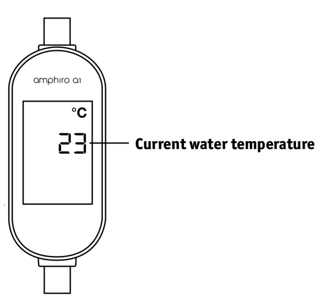
No batteries – powered by water

Amphiro b1 does not use batteries. The device is powered only by water flow. Once the water flow stops, the device will turn off automatically after a few minutes. When on, it will run in one of two modes:

Act consciously and conserve precious resources. Every drop counts!

Mode 1: water temperature

On some days, your Amphiro b1 will display only water temperature. Water temperature has a large impact on your energy usage whilst showering.



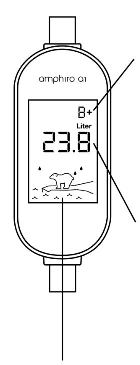
We ask for your patience during this important study phase. When the study ends, you will have the opportunity to keep your Amphiro b1 with extra-added features.

Mode 2: water and energy use

On other days, your Amphiro b1 will display water use and energy use.

Energy efficiency class

A	0 Wh	-	700 Wh
B	700 Wh	-	1,225 kWh
C	1,225 kWh	-	1,750 kWh
D	1,750 kWh	-	2,275 kWh
E	2,275 kWh	-	2,800 kWh
F	2,800 kWh	-	3,325 kWh
G	more than 3,325 kWh		



Water and energy consumption

- During the shower: Consumption in litres and current temperature in Celsius.
- After the shower: Alternating display of a) water consumption in litres and b) energy consumption in Watt-hours (Wh) / kilowatt-hours (kWh)

Climate animation

Iceberg melts with increasing energy consumption.

Amphiro b1 records water usage every time you take a shower. During this study, these data are recorded for research purposes and are only stored locally within the device.

If problems occur or if you have any questions, suggestions, or comments, please contact us at customerprograms@sew.com.au or phone 03 9552 3681.

Figure A.4: Amphiro B1 Trial and Installation Instructions Mailed to Households (continued)

(c) Physical Installation Instructions

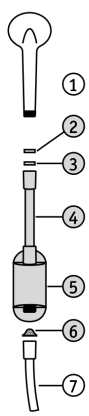
Taking care of your shower display

- Do not submerge the device in water. Do not let it float in the bathtub.
- Do not apply any decalcifiers or abrasive cleaning agents (scouring powder/cream), as they may damage the display.
- This device is not a toy and not suitable for small children. It contains small parts which may be swallowed.
- Do not open the device. Doing so will irreparably damage it, as the compartment sealing will be compromised.
- The device contains a strong permanent magnet as found for example in earphones or name tags. Individuals with a pacemaker should maintain an appropriate safety distance.
- The displayed energy consumption does not account for the boiler/furnace efficiency or transportation losses. The values are therefore not suitable for billing purposes.

This package contains

The blue/shaded elements are part of this package:

1. Shower head (part of your shower)
2. Protective cap (blue, to be removed before installation)
3. Sealing ring (underneath the blue cap at one end of the short hose)
4. Short hose
5. Amphiro b1 personal shower display
6. Sieve with sealing gasket
7. Shower hose (part of your shower)



Technical information

All terminals	½" standard screw
Application	water flow range: 5 to 20 litres per minute Temperature: 5 °C to 65 °C Max. pressure: 10 bar
Protection against moisture	waterproof in accordance with IP65. Do not submerge device in water.
Accuracy	± 10% at 12 litres per minute. Greater deviations may occur at different flow rates.

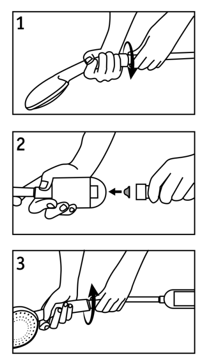
Manufacturer: Amphiro AG, Limmatstrasse 183, CH-8005 Zurich, Switzerland

Your personal shower display

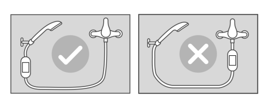
Save water and energy

You decide how much water and energy you use. With Amphiro b1, you can conserve water and energy in the shower, or simply get a sense of your personal resource use.

Installation in three easy steps



1. Unscrew the shower head from the hose.
2. Remove the red adhesive tape at the bottom end of Amphiro b1 and ensure that the sieve stays in place. As shown in the picture, make sure the head of the sieve goes into the device. Screw the shower hose onto this connector with the sieve.
3. Screw the shower head into the short hose attached to the Amphiro b1. Be sure to remove the white safety cap first and ensure that the sealing ring stays in place inside the hose.



For a video demonstrating how to install the Amphiro, please see http://go.unimelb.edu.au/7g66

Figure A.5: Endline Survey

Follow-up Survey Invitation Email

Subject line: Amphiro B1 End-of-Trial Survey

Thank you for participating in the Amphiro B1 trial! You have now reached the final stage. Please share your experience of the personal shower display with us via the following online survey.

["Complete Survey" button here]

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

If you have any additional feedback to share regarding your experience with the Amphiro B1 that is not addressed in our survey, plese contact us at [insert email].

Questions

1. How many household members regularly used the Amphiro B1 each day?
   [1, 2, 3, 4, 5+]

2. Which members of the household regularly used the Amphiro B1 each day? [Please select all that apply]
   [Only Myself, Other Adults, Teenagers (12-18), Children (under 12)]

3. How easy was it for you to read the display once it was installed in your shower?
   [Very easy to read, Somewhat easy to read, Somewhat difficult to read, Difficult to read]

4. Did you pay attention to the Amphiro display day-to-day?
   [Yes regardless of display mode,
   Yes but only when water usage/polar bear were displayed,
   Yes but only when temperature alone was displayed,
   Rarely,
   Never]

5. Did your attentiveness to the device wane over time?
   [Yes, No, I don't know or can't remember]

6. If you share the shower with a family member, did you compare shower times/energy efficiency ratings with each other?
   [Yes always, Yes often, Rarely, Never]

7. When first using the device in the mode that displayed water consumption and energy efficiency, did you consciously set a goal to reduce your usage? If so, please provide details.
   [Yes, No] [Textbox provided for details if Yes]

8. If you set a goal, did you adjust that goal over the course of the trial?
   [No, Yes, adjusted down (shorter showers), Yes, adjusted up (longer showers)]

9. Some customers mentioned that the Amphiro device was only helpful for family members with short hair, i.e. who find it easier to take shorter showers. How many of the regular users of the Amphiro had long hair at the time of the experiment?
   [1, 2, 3, 4, 5+]

10. Did sense that you took shorter showers when the device was in the mode that displayed the polar bear?
    [Yes definitely, I think so, I really could not tell, I don't remember]

11. Did your experience showering with the polar bear lead you to take shorter showers when the display was in temperature-only mode?
    [Yes definitely, I think so, I really could not tell, I don't remember]

12. How did your showering behaviour change after you removed the device and sent it to us?
    [No difference, I took longer showers without the display, I don't know, I can't remember]

13. What best describes how you or your family's attitude towards shower times has changed as a result of your participation in the Amphiro study?
    [Has not changed,
    We pay more attention to how much time we spend in the shower,
    We pay less attention to how much time we spend in the shower]

14. Today, how many minutes long is a typical shower in your home?
    [3 or less, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more]

15. What is your best guess of how much water is used (in litres) during a typical shower in your home today?
    [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]

16. Did you remove the device before the end of the trial? If so, please provide details.
    [Yes I removed the device early, No I only removed it at the end of the trial]
    [Textbox provided for details if Yes]

17. Did you upgrade your showerhead at any point before the end of the trial?
    [No
    Yes I installed a more water-efficient showerhead
    Yes I installed a more powerful (less water-efficient) showerhead]

18. Have you reinstalled the device since receiving it back?
    [Yes, No]

A.2 Robustness to multi-person and multi-shower households

Our structural model focuses on a single individual using a single shower in the home (the "main" shower, per our trial instructions). Yet our experimental data contains households with multiple people and multiple showers. This appendix shows that our main reduced-form estimates of habit build-up and decay are unchanged if we focus on subsamples containing households with one person or one shower. We show this in Tables A.1 and A.2. These tables essentially replicate the main results from Table 4 in the paper: (1) feedback effects emerge immediately and do not evolve while feedback is on, and (2) feedback effects decay sluggishly at a similar rate to our main estimates when feedback is turned off.$^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$ indicates statistical significance in all Appendix tables. Relative to our main results, we lose precision with our standard errors in Tables A.1 and A.2, reflecting that we use smaller subsamples.

Table A.1: Treatment Effects by Experimental Condition

Subsample:
1 Person Households

	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	-6.77***	-8.96***	-8.57***	-8.70***	-9.24***	-6.36***	-7.39***
	(1.18)	(2.62)	(1.97)	(2.06)	(1.72)	(1.66)	(1.97)
OFF	-3.42***		-9.17***	-4.19*	-5.43***	-2.65	-3.73*
	(1.30)		(2.99)	(2.38)	(1.98)	(1.80)	(2.13)
R-Squared	0.59	0.54	0.52	0.56	0.56	0.57	0.57
Observations	13853	4211	5879	5988	6065	6709	6056
ON	-6.44***	-6.31**	-6.37***	-7.91***	-8.44***	-5.22***	-6.26***
	(1.16)	(2.74)	(1.92)	(2.34)	(1.72)	(1.55)	(1.88)
PostON	-0.02	-0.07*	-0.07*	-0.03	-0.04	-0.07**	-0.06*
	(0.02)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
OFF	-4.74***		-13.47***	-6.11**	-5.47***	-3.39*	-5.67**
	(1.42)		(3.17)	(2.90)	(2.06)	(1.85)	(2.36)
PostOFF	0.08		0.13	0.07	0.01	0.09	0.16*
	(0.05)		(0.09)	(0.11)	(0.09)	(0.12)	(0.09)
R-Squared	0.59	0.54	0.52	0.56	0.56	0.57	0.57
Observations	13853	4211	5879	5988	6065	6709	6056

A.3 Cumulative shower counts by experimental conditions

This appendix illustrates that the number of showers recorded by the Amphiro B1 during the trial does not vary across experimental conditions. These findings provide evidence against households responding to the feedback cycles on the extensive margin (e.g., the number of showers taken

Table A.2: Treatment Effects by Experimental Condition

Subsample:
1 Shower Households

	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	-7.82***	-6.13***	-7.99***	-7.33***	-6.48***	-6.71***	-6.47***
	(1.25)	(2.23)	(1.84)	(1.73)	(1.94)	(1.84)	(1.69)
OFF	-4.00***		-4.77**	-5.38**	-3.37	-2.29	-2.24
	(1.27)		(2.28)	(2.07)	(2.04)	(2.02)	(1.63)
R-Squared	0.42	0.44	0.41	0.44	0.43	0.43	0.43
Observations	28298	8989	13132	13313	12032	12061	13716
ON	-8.02***	-6.31**	-8.12***	-8.23***	-7.16***	-7.18***	-6.56***
	(1.25)	(2.38)	(1.95)	(1.79)	(2.03)	(1.86)	(1.75)
$PostON$	0.03	0.01	0.01	0.04*	0.03	0.02	0.00
	(0.02)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
OFF	-5.31***		-8.98***	-6.36***	-6.31***	-3.33	-2.28
	(1.31)		(2.42)	(2.02)	(2.26)	(2.23)	(1.87)
$PostOFF$	0.12***		0.19***	0.06	0.24**	0.13	0.00
	(0.04)		(0.05)	(0.08)	(0.11)	(0.16)	(0.13)
R-Squared	0.42	0.44	0.41	0.44	0.43	0.43	0.43
Observations	28298	8989	13132	13313	12032	12061	13716

in the "main" shower of the home that we instruct trial participants to install their Amphiro B1 in). We examine this among the 555 of 700 households who successfully installed and returned the device with shower meter data. To illustrate this, we estimate the following regression:

$$n_i = \beta_0 + \sum_{j=1}^{6} \beta_j 1\{\mathrm{T}_j \times 1\{i \in T_j\}\} + \varepsilon_i, \tag{A.1}$$

where n_i is the cumulative total number of showers recorded by household i's Amphiro B1 during the trial, and $1\{\mathrm{T}_j \times 1\{i \in T_j\}\}$ is a dummy equaling one if household i is randomly assigned to experimental condition Tj. Table A.3 presents the estimation results. Columns (1)-(3) reveal no significant differences in cumulative shower counts across the experimental conditions in the entire sample, among single-person or multi-person households. Respectively, based on the estimation results in columns (1)-(3), joint tests of the null of an equal number of showers taken during the trial across the trial groups fail to reject with $F(5,548) = 0.93, p = 0.46$, $F(5,143) = 0.44, p = 0.82$, and $F(5,398) = 0.70, p = 0.63$.

Table A.3: Number of Household Showers Taken Within Trial Period by Experimental Condition

	All Households (1)	Single-Person Households (2)	Multi-Person Households (3)
T1	-7.76	0.97	-7.37
	(11.77)	(11.12)	(13.37)
T2	1.90	-8.33	8.18
	(11.44)	(10.09)	(12.16)
T3	11.64	-5.12	13.75
	(11.46)	(10.73)	(11.94)
T4	3.01	0.47	0.17
	(10.98)	(10.01)	(11.91)
T5	0.26	-11.13	6.32
	(11.56)	(9.76)	(12.22)
T6	-10.37	-7.61	3.65
	(11.91)	(9.49)	(13.50)
Constant	164.91***	100.89***	186.62***
	(8.24)	(6.97)	(9.11)
R-Squared	0.01	0.02	0.01
Observations	555	150	405

A.4 Selection into returning devices at the end of the trial

This Appendix examines household selection into returning their device at the end of the trial with data based on pre-trial shower water usage. Recall from Table 2 that 555 of the 700 Amphiro B1's we mailed out were eventually returned used for data extraction. In our baseline survey in Appendix A.1, we ask households *How many minutes long is a typical shower in your home?* All 700 households who were mailed an Amphiro B1 answered the survey as part of opting into the trial. We can explore selection on our dependent variable – shower usage – using these self-reported shower lengths. To this end, we estimate the following Linear Probability Model:

$$1\{\text{Returned with Data}_i\} = \beta_0 + \beta_1 \text{Pre-Trial-ShowerTime}_i + \varepsilon_i \qquad \text{(A.2)}$$

where $1\{\text{Returned with Data}_i\}$ is a dummy variable equaling one if household i returns their device at the end of the trial with data, and ShowerTime_i is household i's self-reported shower time.

Table A.4 presents the results. Column (1) shows our β_1 estimate from (A.2) is a statistically insignificant from 0. We find no evidence of selection into returning a device with data based on stated pre-trial shower usage. In columns (2)-(8) of Table A.4, we estimate β_1 based on seven difference subsamples, wherein each we consider one of our seven experimental conditions. Here, we find mixed coefficient estimates in terms of magnitudes (all small) and statistical significance. These auxiliary results do not suggest any selection into device return based on the duration or number of feedback cycles across the experimental conditions.

Table A.4: Selection into Returning a Device Used with Data Based on Pre-Trial Water Usage

		Experimental Condition Subsample						
	All	T1	T2	T3	T4	T5	T6	T7
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-Trial Shower Time	-0.001	0.003	-0.003	-0.019	0.026***	-0.028**	-0.019	0.015
	(0.005)	(0.014)	(0.013)	(0.015)	(0.007)	(0.013)	(0.014)	(0.011)
Constant	0.803***	0.749***	0.864***	0.918***	0.676***	0.971***	0.891***	0.637***
	(0.036)	(0.103)	(0.097)	(0.100)	(0.082)	(0.094)	(0.098)	(0.098)
R-Squared	0.000	0.001	0.001	0.019	0.070	0.048	0.018	0.017
Observations	700	100	100	100	100	100	100	100

A.5 Distribution of baseline shower usage by household size

Here, we show baseline shower water usage does not differ across households of different sizes. Figure A.6 illustrates this by plotting the distributions of baseline shower water usage by household size. We test for differences in mean water usage by household size using a regression:

$$y_{is} = \beta_1 1\{\text{HHsize}_i = 1\} + \beta_2 1\{\text{HHsize}_i = 2\} + \beta_3 1\{\text{HHsize}_i = 3\} + \beta_4 1\{\text{HHsize}_i \geq 4\} + \varepsilon_{is}, \tag{A.3}$$

where y_{is} is water usage in shower s by household i, $1\{\text{HHsize} = 1\}$ is a dummy equaling one if household i has 1 person and 0 otherwise, and likewise for the other dummy variables. Noting we omit the constant from the regression, the coefficients can be interpreted as mean shower water usage levels by household size in a given sample. We estimate the model using baseline shower data and present the results in column (1) of Table A.5. We find minimal differences in mean usage by household size. A joint test that $\beta_1 = \beta_2 = \beta_3 = \beta_4$ fails to reject the null ($F(3,554) = 0.91$, $p = 0.436$). Columns (2)-(4) of Table A.5 further confirm that pairwise tests of $\beta_i = \beta_j$ for $i, j \in \{1,2,3,4\}$ and $i \neq j$ likewise all fail to reject the null of equal baseline mean water usage between different sized households.

A.6 Water flow rates do not change during the trial

The appendix shows that households' water flow rates do not change during the trial period in our experimental conditions. The results provide evidence against the notion that households respond to feedback cycles by making technological investments, particularly low-flow shower heads. Such investments are of first-order importance regarding potential investments households can make to reduce their shower usage in response to personalized feedback permanently.

To investigate how water flow changes over the trial, we estimate the following regression:

$$f_{is} = \eta_i + \sum_{j=1}^{6} \beta_j 1\{\text{T}_{js} \times 1\{i \in T_{js}\}\} + \tau_k + \varepsilon_{is}, \tag{A.4}$$

where f_{is} is the shower flow rate in terms of liters per minute in shower s for household i recorded

Figure A.6: Distribution of Baseline Shower Water Usage by Household Size

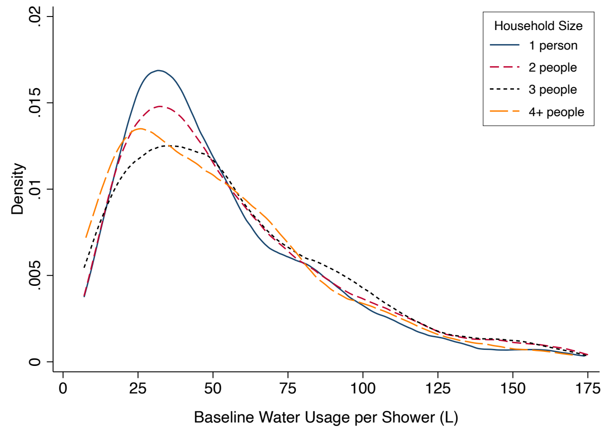


Table A.5: Mean Baseline Shower Water Usage by Household Size

		Pairwise Tests of Differences in Means					
	(1)		(2)		(3)		(4)
$\hat{\beta}_1$	52.284	$\hat{\beta}_1 - \hat{\beta}_2$	-3.606	$\hat{\beta}_2 - \hat{\beta}_3$	-.648	$\hat{\beta}_3 - \hat{\beta}_4$	4.408
	(2.199)		(2.595)		(3.3)		(4.926)
$\hat{\beta}_2$	55.89	$\hat{\beta}_1 - \hat{\beta}_3$	-4.254	$\hat{\beta}_2 - \hat{\beta}_4$	3.76		
	(1.377)		(3.719)		(4.144)		
$\hat{\beta}_2$	56.538	$\hat{\beta}_1 - \hat{\beta}_4$	.154				
	(2.999)		(4.485)				
$\hat{\beta}_2$	52.13						
	(3.909)						
R-Squared	.715						
Observations	9383						

by the Amphiro B1. The dummy variable $1\{T_j \times 1\{i \in T_j\}\}$ equals one if household i is in experimental condition Tj and if shower s occurs after the baseline phase. η_i and τ_k are household and shower count fixed effects. Given our within-subject experimental design, we can test whether the flow rate changes within the feedback period in each experimental condition by testing whether β_j equals 0. A statistically significant negative β_j would suggest a technology investment like installing a low-flow shower head in response to feedback.

Table A.6 reports our results. Most coefficient estimates are statistically insignificant with mixed signs. In terms of magnitude, they are all small relative to the baseline shower flow rate of 8.38 L/minute. Overall, there is no evidence of changing shower flow rates during our trial.

Table A.6: Shower Flowrate (L/min) by Experimental Condition

	All Households (1)	Single-Person Households (2)	Multi-Person Households (3)
T1	0.26**	0.22	0.19
	(0.10)	(0.21)	(0.12)
T2	-0.01	0.11	-0.11
	(0.11)	(0.20)	(0.13)
T3	0.07	-0.10	0.04
	(0.11)	(0.21)	(0.14)
T4	-0.09	-0.05	-0.17
	(0.11)	(0.18)	(0.14)
T5	-0.07	-0.19	-0.11
	(0.10)	(0.15)	(0.13)
T6	0.02	-0.13	-0.03
	(0.09)	(0.16)	(0.12)
T7	0.00	-0.10	-0.05
	(0.10)	(0.20)	(0.12)
Constant	8.37***	7.88***	8.45***
	(0.06)	(0.12)	(0.07)
R-Squared	0.72	0.79	0.70
Observations	91496	14434	77062

B Steady state transitions

B.1 Steady state transitions predicted by the attention-based habit model

In the attention-based habit model, we shut down the influence of consumption-based habit by setting $\gamma = 0$. Thus, by the first order condition in equation (10),

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

Jump in c_t when feedback is first turned on ($ON_t = 1$) and we depart OFF steady state

Suppose at $t = 0$ we are at **OFF** steady state, i.e., $\theta_0 = w_0 = \theta$. Further suppose that at $t = 1$ we switch $ON_t = 0$ to $ON_t = 1$. Then θ_1 changes from θ to 1. By (B.1),

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

and

$$\Delta c_1 + c^*_{OFF} = \frac{a - \theta p}{b} + \frac{\theta - 1}{b}p = \frac{a-p}{b} = c^*_{ON},$$

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

No jump in c_t when feedback is first turned off ($ON_t = 0$) and we leave the ON steady state

Suppose at $t = 0$ we are at **ON** steady state, i.e., $\theta_0 = w_0 = 1$ and that at $t = 1$ we switch $ON_t = 1$ to $ON_t = 0$. Then θ_1 changes from 1 to w_1. When $ON_t = 0$, w_t is defined recursively as

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

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

Path of c_t when feedback is left off ($ON_t = 0$) and we converge to the OFF steady state

Suppose we leave $ON_t = 0$ from $t = 2$ to $t = T$, at which point the **OFF** steady state is reached. Recall that $\theta_t = w_t$ so long as $ON_t = 0$. At:

- $t = 1$: $\theta_1 = \theta\alpha_{ON} + (1 - \alpha_{ON})$
- $t = 2$: $\theta_2 = \theta\alpha_{ON} + (1 - \alpha_{ON})[\theta\alpha_{ON} + (1 - \alpha_{ON})] = \theta + (1 - \theta)(1 - \alpha_{ON})^2$

We can show by induction for $1 \leq t \leq T$,

$$\theta_t = \theta + (1 - \theta)(1 - \alpha_{ON})^t. \tag{B.2}$$

As shown above, $\theta_1 = \theta\alpha_{ON} + (1 - \alpha_{ON})$, satisfying (B.2). Now suppose

$$\theta_k = \theta + (1 - \theta)(1 - \alpha_{ON})^k,$$

$1 \leq k < T$. Then

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

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

$$\Delta\theta_t = \theta + (1 - \theta)(1 - \alpha_{ON})^t - \theta - (1 - \theta)(1 - \alpha_{ON})^{t-1} = -a(1 - \theta)(1 - \alpha_{ON})^{t-1}$$

for $t \geq 1$. By (B.5), the change in consumption along the transition path at t is $\Delta c_t = -\frac{p}{b}\Delta\theta_t$.

Consumption transition paths between steady states

Infinite sum of Δc_t toward **OFF** steady state. The transition path of Δc_t from the **ON** to **OFF** steady state is described by a geometric sequence that takes the form xr^k, where $x = \frac{ap}{b}(1 - \theta), r = 1 - \alpha_{ON}$, 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:

$$\sum_{t=1}^{\infty} \Delta c_t = \frac{\frac{ap}{b}(1 - \theta)}{1 - 1 + \alpha_{ON}} = \frac{p}{b}(1 - \theta), \tag{B.3}$$

which, when added to $c^*_{ON} = \frac{a-p}{b}$, yields $c^*_{OFF} = \frac{a-\theta p}{b}$.

Finite sum of Δc_t toward the **OFF** steady state. 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 (B.3):

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

B.2 Steady state transitions predicted by the consumption-based habit model

To compute these transitions, we assume away transitions in salience levels and fix $\theta_t = \theta$ if $ON_t = 1$, and $\theta_t = 1$ if $ON_t = 0$. This is Stigler and Becker (1977)'s consumption-based habit model, combined with the (non-dynamic) price salience model of Chetty et al. (2009).

Consumption consumption-based habit h_{t+1} is given by

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

so Δh_{t+1} is given by

$$\Delta h_{t+1} = (1-\delta)\Delta c_t + \delta \Delta h_t. \tag{B.4}$$

The general form for optimal consumption choice c_t from the first order condition is

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

so the general form for Δc_{t+1} is

$$\Delta c_{t+1} = \frac{\gamma}{b}\Delta h_{t+1} - \frac{p}{b}\Delta\theta_{t+1}. \tag{B.5}$$

Jump in c_t when feedback is first turned on and we initially depart from the OFF steady state

Suppose at $t = 0$ we are at the **OFF** steady state, i.e., $\theta_0 = w_0 = \theta$ and $h_0 = c_0$. Suppose then that at $t = 1$, we turn $ON_t = 1$. Then by (B.4),

$$\Delta h_1 = (1-\delta)0 + \delta 0 = 0, \tag{B.6}$$

and θ_1 changes from θ to 1. By (B.5), the jump in c_t is computed as

$$\Delta c_1 = 0 - \frac{p}{b}(1-\theta) = \frac{\theta - 1}{b}p. \tag{B.7}$$

Path of c_t when feedback is left on ($ON_t = 1$) and we converge to the ON steady state

Suppose we leave $ON_t = 1$ from $t = 2,\ldots,T$ at which point the **ON** steady state is reached. With $ON_1 = 1$ such that $\theta_t = 1$ in period 1, so long as $ON_t = 1$, $\Delta\theta_t = 0$ for $t = 2,\ldots,T$. So by (B.5), Δc_t is driven only by Δh_t. The transitional dynamics of Δh_t can be computed as:

- $t = 2$: $\Delta h_2 = (1-\delta)\Delta c_1$ by (B.4) and (B.6), and $\Delta c_2 = \frac{\gamma}{b}\Delta h_2$ by (B.5).

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

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

$$\Delta h_t = (1-\delta)\left[\delta + \frac{\gamma}{b}(1-\delta)\right]^{t-2}\Delta c_1. \tag{B.8}$$

By (B.4) and (B.6), $\Delta h_2 = (1-\delta)\Delta c_1$, which satisfies (B.8). 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 (B.5), $\Delta c_k = \frac{\gamma}{b}\Delta h_k$. Then

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

as desired. By (B.5), the period t change in consumption along the transition is $\Delta c_t = \frac{\gamma}{b}\Delta h_t$.

Jump in c_t when feedback is first turned off ($ON_t = 0$) and we leave the ON steady state

Suppose instead that at $t = 0$ we are at **ON** steady state, i.e. $\theta_0 = w_0 = 1$ and $h_0 = c_0$. Suppose then that at $t = 1$, we turn feedback off ($ON_t = 0$). Then by (B.4),

$$\Delta h_1 = (1-\delta)0 + \delta 0 = 0. \tag{B.9}$$

Further, $w_1 = \theta\alpha_{ON} + (1-\alpha_{ON})\theta = \theta$, implying that θ_1 changes from 1 to θ. By (B.5),

$$\Delta c_1 = 0 - \frac{p}{b}(\theta - 1) = \frac{1-\theta}{b}p. \tag{B.10}$$

which shows there is an initial jump when feedback is initially turned off ($ON_t = 0$) and we depart the **ON** steady state. Notice the size of the jump does not depend on prior feedback duration.

Path of c_t when feedback is left off ($ON_t = 0$) and we converge back to the OFF steady state

Suppose we leave $ON_t = 0$ from $t = 2, \ldots, T$, at which point the **OFF** steady state is reached. With $\theta_t = \theta$ if $ON_t = 0$, $\Delta\theta_t = 0$ for all $t = 2, \ldots, T$. By (B.4) and (B.9), $\Delta h_2 = (1-\delta)\Delta c_1$, which satisifies (B.8). Through an identical induction step, we find that Δh_t follows the same transition path described in (B.8). By (B.5), the change in consumption at t along the path is $\Delta c_t = \frac{\gamma}{b}\Delta h_t$.

Symmetry in jumps and transition paths

By (B.7) the magnitude of the first jump is $\frac{\theta-1}{b}p$. By (B.10), the magnitude of the second jump is $\frac{1-\theta}{b}p$. Hence the two jumps have equal magnitudes in opposite directions.

We further find that the transition path of consumption-based habit is described by (B.8):

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

for both the **OFF** $\to$ **ON** transition and the **ON** $\to$ **OFF** transition.

Consumption transition paths between steady states

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:

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

We can then add Δc_1 to obtain the infinite sum from $t=1$ onwards. Suppose we transition from the **OFF** to the **ON** steady state, i.e., Δc_1 is given by (B.7):

$$\sum_{t=1}^{\infty} \Delta c_t = \frac{\gamma}{b-\gamma}\Delta c_1 + \Delta c_1 = \left(1+\frac{\gamma}{b-\gamma}\right)\frac{\theta-1}{b}p = \frac{\theta p - p}{b-\gamma}, \tag{B.11}$$

which when added to $c^*_{OFF} = \frac{a-\theta p}{b-\gamma}$ yields $c^*_{ON} = \frac{a-p}{b-\gamma}$. Similarly, adding the sum of Δc_t in the opposite direction to the **ON** steady state yields the **OFF** steady state.

Finite sum of Δc_t. The 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 $t = 1 \dots T$ from the infinite sum in (B.11):

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

B.3 General analytic expressions for the accumulation of w_t

Here we compute general expressions of the accumulation of w_t from any starting point (not necessarily a steady state) when feedback is turned on ($ON_t = 1$) or off ($ON_t = 1$).

When Feedback is Turned On

When feedback is on ($ON_t = 1$), $w_t = \alpha_{ON} + (1-\alpha_{ON})w_{t-1}$. Suppose we turn $ON_t = 1$ at time $t = 0$, at which point w_t 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

$$w_t = \alpha_{ON} + (1-\alpha_{ON})w_{t-1} \implies u_t + k = \alpha_{ON} + (1-\alpha_{ON})u_{t-1} + (1-\alpha_{ON})k.$$

Set $k = \alpha_{ON} + (1-\alpha_{ON})k$, such that $k = 1$. We can then write

$$u_t + 1 = \alpha_{ON} + (1-\alpha_{ON})u_{t-1} + 1 - \alpha_{ON} \implies u_t = (1-\alpha_{ON})u_{t-1} \text{ and } u_t = (1-\alpha_{ON})^t u_0.$$

It follows that $w_t = u_t + k = (1-\alpha_{ON})^t u_0 + 1$. Since, by definition, $u_0 = w_0 - 1$, we can rewrite

$$w_t = (1-\alpha_{ON})^t (w_0 - 1) + 1,$$

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

When Feedback is Turned Off

When feedback is off ($ON_t = 0$), $w_t = \theta\alpha_{ON} + (1-\alpha_{ON})w_{t-1}$. The proof proceeds similarly. Suppose $ON_t = 0$ at $t = 0$, at which point w_t takes on an initial value of w_0. 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

$$w_t = \theta\alpha_{ON} + (1-\alpha_{ON})w_{t-1} \implies u_t + k = \theta\alpha_{ON} + (1-\alpha_{ON})u_{t-1} + (1-\alpha_{ON})k.$$

Set $k = \theta\alpha_{ON} + (1-\alpha_{ON})k$, such that $k = \theta$. We can then write

$$u_t + \theta = \alpha_{ON} + (1-\alpha_{ON})u_{t-1} + \theta - \alpha_{ON} \implies u_t = (1-\alpha_{ON})^t u_0 \text{ and } w_t = (1-\alpha_{ON})^t(w_0 - \theta) + \theta,$$

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

Feedback On and Off Cycles

These non-recursive expressions simplify the process of solving for w_t after alternating periods of feedback being on ($ON_t = 1$) and off ($ON_t = 0$). Let w_t be in the **OFF** 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 $t = j$,

$$w_j = (1-\alpha_{ON})^j(\theta - 1) + 1.$$

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

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

C Recursive implementations of the models

C.1 Attention-based habit formation model

We can re-write the recursive formula as follows:

$$\omega_t - \theta = \begin{cases} \alpha_{ON}\cdot(1-\theta) + (1-\alpha_{ON})\cdot(\omega_{t-1}-\theta) & \text{if } ON_t = 1 \\ (1-\alpha_{ON})\cdot(\omega_{t-1}-\theta) & \text{if } ON_t = 0 \end{cases} \tag{C.1}$$

Consider t = 1, 2, ..., k during which feedback is on ($ON_t = 1$):

At $t = 1$: $\phi_1(\alpha_{ON})\cdot(1-\theta) := \omega_1 - \theta = \alpha_{ON}(1-\theta) \implies \phi_1(\alpha_{ON}) = \alpha_{ON}$

At $t = 2$: $\phi_2(\alpha_{ON})\cdot(1-\theta) := \omega_2 - \theta = \alpha_{ON}(1-\theta) + \alpha_{ON}(1-\alpha_{ON})\cdot(1-\theta)$

$$\implies \phi_2(\alpha_{ON}) = \alpha_{ON} + (1-\alpha_{ON})\cdot\phi_1(\alpha_{ON})$$

We iterate the process to obtain

$$\phi_k(\alpha_{ON}) = \alpha_{ON} + (1-\alpha_{ON})\cdot\phi_{k-1}(\alpha_{ON})$$

Consider t = k+1, k+2, ..., k+M during which feedback is off ($ON_t = 0$):

At $t = k+1$:

$$\phi_{k+1}(\alpha_{ON}) \cdot (1-\theta) := \omega_{k+1} - \theta = (1-\alpha_{ON})(\omega_k - \theta) = (1-\theta) \cdot (1-\alpha_{ON}) \cdot \phi_k(\alpha_{ON})$$
$$\implies \phi_{k+1}(\alpha_{ON}) = (1-\alpha_{ON}) \cdot \phi_k(\alpha_{ON})$$

At $t = k+2$:

$$\phi_{k+2}(\alpha_{ON}) \cdot (1-\theta) := \omega_{k+2} - \theta = (1-\alpha_{ON})(\omega_{k+1} - \theta) = (1-\theta) \cdot (1-\alpha_{ON}) \cdot \phi_{k+1}(\alpha_{ON})$$
$$\implies \phi_{k+2}(\alpha_{ON}) = (1-\alpha_{ON}) \cdot \phi_{k+1}(\alpha_{ON})$$

We iterate the process to obtain

$$\phi_{k+M}(\alpha_{ON}) = (1-\alpha_{ON}) \cdot \phi_{k+M-1}(\alpha_{ON})$$

Consider t = k+M+1, k+M+2, ... during which feedback is on ($ON_t = 1$):

At $t = k+M+1$:

$$\phi_{K+M+1}(\alpha_{ON}) \cdot (1-\theta) := \omega_{K+M+1} - \theta = [\alpha_{ON} + (1-\alpha_{ON}) \cdot \phi_{K+M}(\alpha_{ON})](1-\theta)$$
$$\implies \phi_{K+M+1}(\alpha_{ON}) = \alpha_{ON} + (1-\alpha_{ON}) \cdot \phi_{K+M}(\alpha_{ON})$$

At $t = k+M+2$:

$$\phi_{K+M+2}(\alpha_{ON}) \cdot (1-\theta) := \omega_{K+M+2} - \theta = [\alpha_{ON} + (1-\alpha_{ON}) \cdot \phi_{K+M+1}(\alpha_{ON})](1-\theta)$$
$$\implies \phi_{K+M+2}(\alpha_{ON}) = \alpha_{ON} + (1-\alpha_{ON}) \cdot \phi_{K+M+1}(\alpha_{ON})$$

By induction, we can verify that

$$\phi_t(\alpha_{ON}) = \begin{cases} 0 & \text{if t = 0} \\ \alpha_{ON} + (1-\alpha_{ON})\phi_{t-1} & \text{if } ON_t \ = \ 1 \\ (1-\alpha_{ON})\phi_{t-1} & \text{if } ON_t \ = \ 0 \end{cases} \tag{C.2}$$

We wish to formulate the attention-based habit model in the form

$$y_{it} = \eta_i + (ON_{it} + OFF_{it}\phi_{it}(\alpha_{ON}))\varphi + \delta_t + \varepsilon_{it} \tag{C.3}$$

That is, we express the persistence effect as a fraction $\phi_{it}(\alpha_{ON})$ of the feedback effect φ, depending on the parameter α_{ON}, as well as the feedback on and off-period history of individual i up to period t. The rest of the model is linear so that it can be estimated by OLS conditional on α_{ON} and the recursion for $\phi_{it}(\alpha_{ON})$.

It is straightforward to extend the model to allow for different rates of buildup (α_{ON}) and decay

(α_{OFF}) of attention stock. The function ϕ_t then becomes

$$\phi_t(\alpha_{ON}, \alpha_{OFF}) = \begin{cases} 0 & \text{if t = 0} \\ \alpha_{ON} + OFF\phi_{t-1} & \text{if } ON_t \ = \ 1 \\ OFF\phi_{t-1} & \text{if } ON_t \ = \ 0 \end{cases} \tag{C.4}$$

C.2 Consumption-habit habit formation model

The level of consumption-based habit is:

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

and the optimal level of consumption is given by:

$$\frac{\partial U}{\partial c_t} = 0 \implies c_t = \frac{a + \gamma h_t - \theta_t p}{b}$$

And the model's **ON** and **OFF** steady states are

$$c^*_{OFF} = \frac{a + \gamma h^* - \theta p}{b} \implies c^*_{OFF} = \frac{a - \theta p}{b - \gamma}$$

$$c^*_{ON} = \frac{a + \gamma h^* - p}{b} \implies c^*_{ON} = \frac{a - p}{b - \gamma}$$

It follows that the change in consumption from the **OFF** steady state (i.e., at baseline) is

$$c_t - c^*_{OFF} = \frac{a + \gamma h_t - \theta_t p}{b} - \frac{a + \gamma h^* - \theta p}{b} = \frac{\gamma}{b}(h_t - h^*) + \frac{\theta - \theta_t}{b}p.$$

The corresponding change in the consumption-based habit stock is

$$h_t - h^* = (1-\delta)h_{t-1} + \delta c_{t-1} - h^* = (1-\delta)(h_{t-1} - h^*) + \delta(c_{t-1} - h^*)$$

Consider t = 1, 2, ..., k during which $ON_t = 1$:

At $t = 1$:

$$h_1 - h^* = 0; \qquad c_1 - c^* = \frac{\gamma}{b}(0) + \underbrace{\frac{\theta - 1}{b}p}_{\varphi}$$

At $t = 2$:

$$h_2 - h^* = (1-\delta)(h_1 - h^*) + \delta(c_1 - h^*) = \underbrace{\delta}_{\phi_2}\varphi$$

$$c_2 - c^* = \frac{\gamma}{b}(h_2 - h^*) + \varphi = \frac{\gamma}{b}\phi_2\varphi + \varphi$$

By iterating the above process, at $t = k$:

$$h_k - h^* = \phi_k\varphi; \qquad c_k - c^* = \frac{\gamma}{b}\phi_k\varphi + \varphi$$

Consider t = k+1, k+2, ..., k+M during which $ON_t = 0$:

At $t = k+1$:

$$h_{k+1} - h^* = \phi_{k+1}\varphi$$
$$c_{k+1} - c^* = \frac{\gamma}{b}(h_{k+1} - h^*) + \frac{\theta - \theta}{b}p = \frac{\gamma}{b}\phi_{k+1}\varphi$$

At $t = k+2$:

$$h_{k+2} - h^* = (1-\delta)(h_{k+1} - h^*) + \delta(c_{k+1} - h^*) = \underbrace{\left[(1-\delta)\phi_{k+1} + \delta(\frac{\gamma}{b}\phi_{k+1})\right]}_{\phi_{k+2}}\varphi$$

$$c_{k+2} - c^* = \frac{\gamma}{b}(h_{k+2} - h^*) + \frac{\theta - \theta}{b}p = \frac{\gamma}{b}\phi_{k+2}\varphi$$

By iterating the above process, at $t = k+M$:

$$h_{k+M} - h^* = (1-\delta)(h_{k+M-1} - h^*) + \delta(c_{k+M-1} - h^*) = \phi_{K+M}\varphi$$
$$c_{k+M} - c^* = \frac{\gamma}{b}(h_{k+M} - h^*) + \frac{\theta - \theta}{b}p = \frac{\gamma}{b}\phi_{k+M}\varphi$$

Consider t = k+M+1, k+M+2, . . . during which $ON_t = 1$:

At $t = k+M+1$:

$$h_{k+M+1} - h^* = (1-\delta)(h_{k+M} - h^*) + \delta(c_{k+M} - h^*) = \phi_{K+M+1}\varphi$$
$$c_{k+M+1} - c^* = \frac{\gamma}{b}(h_{k+M+1} - h^*) + \frac{\theta - 1}{b}p = \frac{\gamma}{b}\phi_{k+M+1}\varphi + \varphi$$

At $t = k+M+2$:

$$\begin{aligned} h_{k+M+2} - h^* &= (1-\delta)(h_{k+M+1} - h^*) + \delta(c_{k+M+1} - h^*) \\ &= \underbrace{\left[(1-\delta)\phi_{k+M+1} + \delta(\frac{\gamma}{b}\phi_{k+M+1} + 1)\right]}_{\phi_{k+M+2}}\varphi \end{aligned}$$

$$c_{k+M+2} - c^* = \frac{\gamma}{b}(h_{k+M+2} - h^*) + \frac{\theta - 1}{b}p = \frac{\gamma}{b}\phi_{k+M+2}\varphi + \varphi$$

By induction, we can verify that

$$\phi_t = \begin{cases} 0 & \text{if } t = 1 \\ \delta & \text{if } t = 2 \\ (1-\delta)\phi_{t-1} + \delta(\frac{\gamma}{b}\phi_{t-1} + 1) & \text{if } ON_{t-1} \;=\; 1,\ t \neq 2 \\ (1-\delta)\phi_{t-1} + \delta(\frac{\gamma}{b}\phi_{t-1}) & \text{if } ON_{t-1} \;=\; 1 \end{cases} \tag{C.5}$$

$$h_t - h^* = \begin{cases} 0 & \text{if } t = 1 \\ \phi_t\varphi & \text{if } t > 1 \end{cases} \tag{C.6}$$

$$c_t - c^*_{OFF} = \begin{cases} \frac{\gamma}{b}\phi_t\varphi + \varphi & \text{if } ON_t \; = \; 1 \\ \frac{\gamma}{b}\phi_t\varphi & \text{if } ON_t \; = \; 0 \end{cases} \tag{C.7}$$

We want to formulate the consumption-based habit model in the form

$$y_{it} = \eta_i + (ON_{it} + \frac{\gamma}{b}\phi_{it-1}(\delta, \frac{\gamma}{b}))\varphi + \delta_t + \varepsilon_{it} \tag{C.8}$$

That is, we express the treatment effect in terms of φ, the immediate effect of feedback, depending on the parameters δ and $\frac{\gamma}{b}$, as well as the feedback on-period history of individual i up to period t. The rest of the model is linear so that it can be estimated by OLS conditional on δ and $\frac{\gamma}{b}$, and the recursion for $\phi_{it}(\delta, \frac{\gamma}{b})$.

D Habit formation with CARA utility

Consider the constant absolute risk aversion (CARA) utility function

$$u(c_t, h_t) = 1 - \frac{e^{-ac_t + \gamma h_t}}{a}$$

where $a > 0$, and c_t, h_t, and γ are defined as in Section 4 in the paper. The marginal utility of consumption is

$$\frac{\partial u}{\partial c} = e^{-ac + \gamma h} \tag{D.1}$$

The utility function is concave as the second derivative is negative; marginal utility is convex, as the third derivative is positive. Notice also that consumption is habit forming, as

$$\frac{\partial^2 u}{\partial c \partial h} = \gamma \cdot e^{-ac + \gamma h} > 0$$

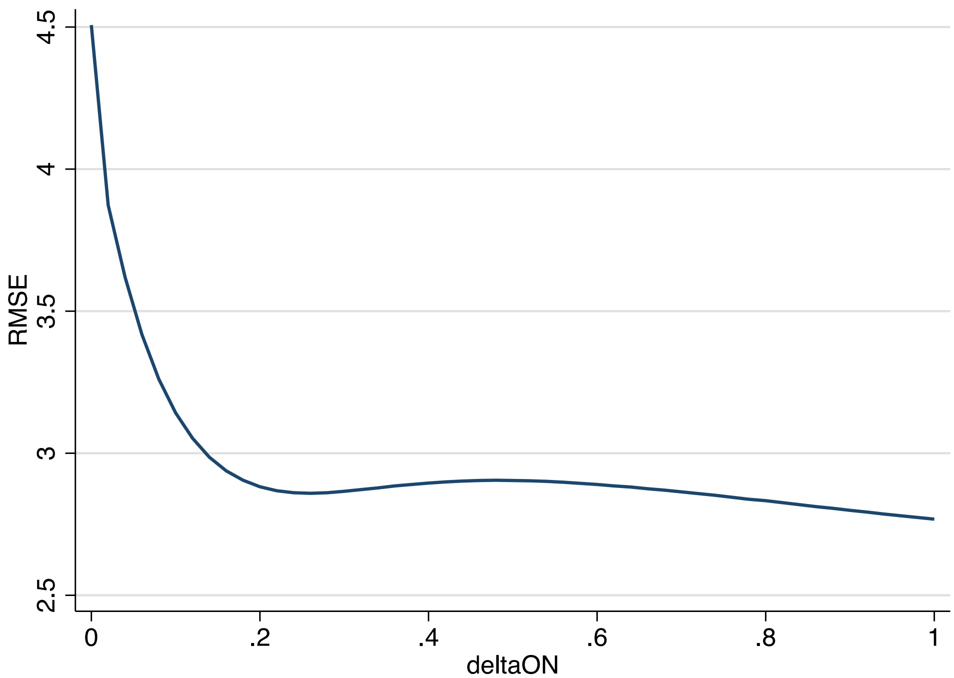
Consumption is given by

$$e^{-ac + \gamma h} = \theta_t p \Rightarrow c^*_t = \frac{\gamma}{a}h_t - \frac{1}{a}\left[\log(\theta_t) + \log(p)\right] \tag{D.2}$$

Thus, equation (D.2), just like equation (10) in the paper, is linear in in the consumption-based habit. Thus, this example shows that constant marginal utility of consumption is not necessary to display symmetric and, in fact, linear, consumption-based habit effects on consumption.

E Auxiliary Structural Model Results

Figure E.1: Asymmetric Consumption-Based Habit Model RMSE for all Possible δ_{ON} Values



Notes: The figure displays the estimated RMSE of the asymmetric consumption habit model from column (2) of Table 5. We plot the RMSE for each possible value of δ_{ON} and estimate all other parameters conditionally on that value. The figure illustrates that the model's RMSE is minimized at $\hat{\delta}_{ON} = 1$.

Table E.1: Consumption-Based and Attention-Based Habit Model Parameter Estimates Including the OFF Variable

	Consumption-Based Habit Model		Attention-Based Habit Model	
	(1)	(2)	(3)	(4)
φ	–6.531 (0.683)	–5.322 (0.769) [0.746]	–7.701 (0.653)	–7.623 (0.663)
$\delta_{ON} = \delta_{OFF}$	0.077 (0.260)			
δ_{ON}		1.000 (0.424)		
δ_{OFF}		0.080 (0.073) [0.069]		
γ/b	0.279 (0.089)	0.326 (0.100) [0.076]		
$\alpha_{ON} = \alpha_{OFF}$			0.041 (0.368)	
α_{ON}				0.062 (0.096)
α_{OFF}				0.026 (0.039)
OFF	–3.460 (0.691)	–2.649 (0.735) [0.771]	–2.082 (1.115)	–0.871 (1.077)
Within-Sample RSME	2.472	2.407	2.290	2.275

Notes: $N = 86,376$ (household, shower) observations in each sample consisting of 1078 individuals and 555 households. 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. Boostrap standard errors clustered at the household level reported. Standard errors for unconstrained models are in parentheses. The column (2) estimates constrain $\delta_{ON} = 1$ in estimation, which thus does not have a standard error. Constrained standard errors for the other parameters are in brackets. See the text for the calculation of within-sample RMSE.

F Other persistence mechanisms

F.1 Automatic control

Camerer, Landry and Webb (2020) develop a dual-systems model of habit involving default decision-making when one's decision-making environment is stable, and deliberation and infrequent updating of decision rules occurs when a decision-making environment changes sufficiently. In our setting, the introduction and removal of feedback could represent a substantial change in an individual's decision-making environment such that feedback-induced changes in consumption over time reflect discrete changes in default decision rules at the individual level.

Exploiting the richness 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 (2) as follows:

$$y_{is} = \eta_i + \beta_1 ON_{is} + \beta_2 PostON_{is} + \beta_3 OFF_{is} + \beta_4 PostOFF_{is} + \sum_{i}^{N} \lambda_i PostOFF_{is_i^*} + \tau_k + \varepsilon_{is}, \quad \text{(F.1)}$$

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

1. Initialize (F.1) for household i by setting $PostOFF_{is_i^*} = PostOFF_{is} \times \eta_i$, where η_i is a household i fixed effect. We construct these initial household–specific $PostOFF_{is_i^*}$ variables for all N households in conditions T3–T7 for whom we observe post–feedback showers. In effect, this initializes $s_i^* = 1$ (1 shower since feedback has been turned off) for all $i = 1, \ldots, N$. Starting from these initial values, we search household-by-household for the set of s_i^* values that best rationalize our data.

2. With the initialized $PostOFF_{is_i^*}$ variables, run the regression in (F.1). This yields a distribution of household–specific post–feedback treatment effects $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_N$ above and beyond the common post–feedback treatment effect β_4. Let $i = 1$ as the household with the smallest device identifier in T3, and $i = N$ as the household with the largest device identifier in T7. To avoid perfect collinearity with $PostOFF_{is}$, we drop one of the households in conditions T3–T7 in estimating the coefficients in (F.1).

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:

$$H_0 : \lambda_1 = 0 \;\text{ vs. }\; H_1 : \lambda_1 \neq 0 \quad \text{(F.2)}$$

Denote the F-statistic from this test by $F_{1,1}$, where the first "1" in the subscript corresponds to household $i = 1$. 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 or equal to $\tau = 2$ showers since feedback was turned off for household 1.

5. Run regression (F.1) and test the hypothesis in (F.2) again. Denote the corresponding F-statistic for the test for $i = 1$ and $\tau = 2$ by $F_{1,2}$.

6. Iterate between steps 4 and 5 for household 1, each time incrementing τ by 1 and re-defining $PostOFF_{1s^*}$ such that it equals 0 if it has been less than or equal to τ showers since feedback was turned off for household 1. Denote the F-statistic for the hypothesis test in (F.2) for household $i = 1$ at iteration $\tau = j$ by $F_{1,j}$.

7. Find the value of j that corresponds to the maximum F-statistic from $F_{1,1}, F_{1,2}, \ldots, F_{1,J_1}$, where J_1 is the maximum number of consecutive post–feedback showers for household 1. Define s_1^* as the shower corresponding to this maximum F-statistic for household $i = 1$. Shower s_1^* is our initial estimate of the timing of the post-feedback jump for household 1.

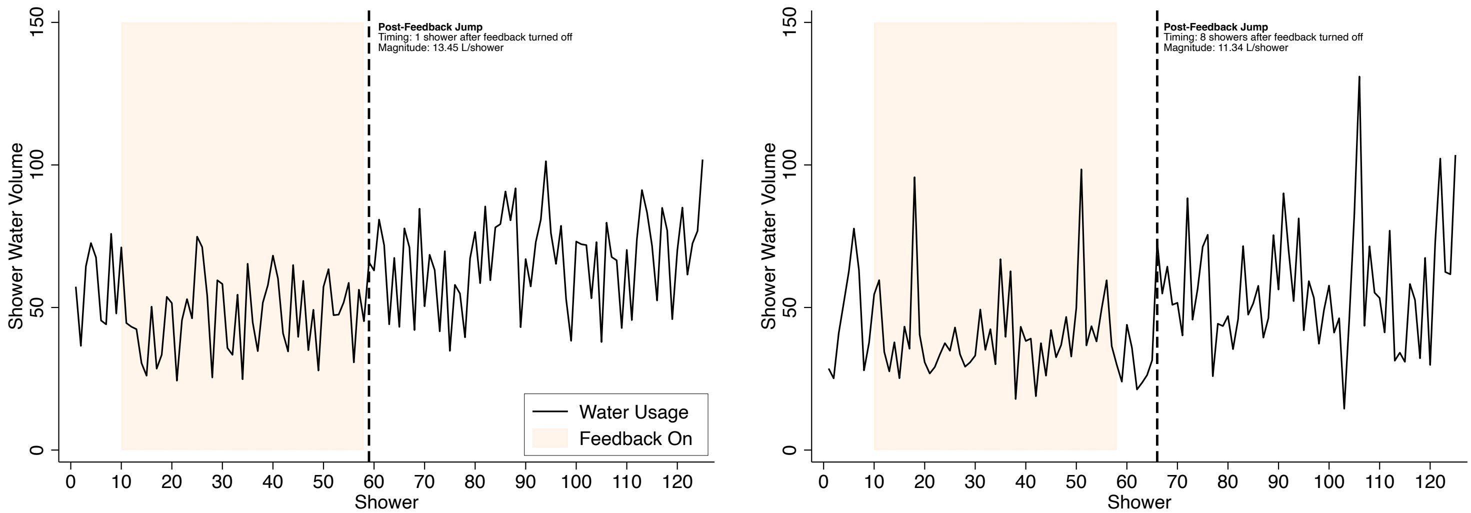
8. Move to household $i = 2$ and repeat steps 2–7, holding fixed s_i^* at their current values for all other households $i \neq 2$, including s_1^* at the value previously found from steps 2–7 for household 1. [...] Repeat steps 2–8 for households $i = 3, \ldots, N$. and find s_k^* for household k at iteration k holding fixed s_i^* at their current values for all other households $i \neq k$.

Once we have looped through all N households, we obtain estimates of the *timing* of the post–feedback jumps $s_1^*, s_2^*, \ldots, s_N^*$ and their *magnitudes* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_N$. We obtain the latter coefficients by running the regression in (F.1) where each $PostOFF_{is_i^*}$ regression corresponds to the s_i^* value from steps 1–8.

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

Figure F.1: Examples: Identifying Structural Breaks with Individual Households

(a) Example Household 1 (b) Example Household 2



Implementation

Various considerations exist in implementing this search for household-specific post-feedback jumps. Households must have sufficiently many observations without feedback. We restrict households to having a minimum of 12 showers without feedback to implement our test. This leaves 328 of 394 households from conditions T3-T7 in our sample. We also use the 161 households from conditions T1 and T2 in robustness checks.

In line with Andrews (1993), we check for s_i^* values for a given household up until the last 20% of observations during the post-feedback phase. That is, we search for s_i^* from showers $\tau = 1, \ldots, 0.8 \times J_i$ after feedback is turned off. This restriction helps ensure sufficient data after a candidate s_i^* to implement the F-test for testing for a structural break at each τ value.

After finding $s_1^*, \ldots, s_N^*$, we could iterate on steps 2–8, starting from the breakpoints found in the first iteration. From there, we could continue iterating on steps 2–8 until the $s_1^*, \ldots, s_N^*$ converge according to some metric. In practice, however, when we iterate on steps 2–8 a second time, we find virtually no difference in our coefficients of interest β_4 below nor in the timing and magnitude of jumps across households. Therefore, our results reflect just one iteration of steps 1–8.

Finally, as in the paper, we cluster standard errors at the household level.

Figure F.1 shows how our routine identifies household-specific post-feedback jumps in water usage for two example households in condition T3. Visually, we see that water usage for Household 1 in panel (a) sharply drops and rises when feedback is turned on and off. Our routine identifies the immediate 13.45 L/shower jump in consumption one shower after feedback is turned off for this household. Likewise, for Household 2 in panel (b), our routine successfully identifies a delayed 11.34 L/shower jump in shower water usage eight showers after feedback is turned off.

Results

Table F.1 presents our main findings. For reference, panel (a) of the table reproduces the bottom panel of Table 4 from the paper. Panel (b) reports analogous estimates to panel (a) based on the 489 of 555 households for which we test for post-feedback jumps in consumption. We obtain similar results across all columns by comparing panels (a) and (b). Thus, there is no evidence of sample selection bias in the coefficient estimates from our conditioning on households with sufficiently many post-feedback periods in testing for heterogeneous post-feedback jumps.

Panel (c) of Table F.1 presents our estimation results from equation (F.1), where we allow for household-specific post-feedback jumps. Comparing Panels (b) and (c), we obtain similar magnitude estimates on the $PostOFF_{it}$ coefficients. In other words, we estimate a similar reduce-form decay rate 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 persistence and asymmetric results in Table F.1 support our attention-based theory of habit in favor of a habit–as–automatic–control model in our particular research context.

F.2 Experimentation and learning

This section studies an experimentation and learning mechanism for consumption responses to real-time feedback. Our motivation comes from Larcom et al. (2017), who examine persistent changes in individuals' commuting behavior after the 2014 London Tube Strike, which temporarily force experimentation and learning about other forms of commuting. The analogue in our setting is feedback-induced experimentation with shorter shower lengths by households who learn about the costs and benefits of doing so, leading to permanent changes in shower length and water consumption in the long run.

Converging to a new long-run consumption level

The time-varying persistence effects from experimental condition T3 provide a natural way to test for permanent long-run changes in shower water usage from experimentation and learning. We can test for this directly using the column (3) estimates from Table 4, which recall estimates our treatment effects from the regression in equation (2) using experimental conditions T1–T3. The following hypothesis test establishes whether there is a long-run shift in consumption from feedback:

$$H_0 : \beta_3 + 72 \times \beta_4 = 0 \text{ vs. } H_1 : \beta_3 + 72 \times \beta_4 \neq 0$$

where recall β_3 and β_4 are the coefficients on OFF_{is} and $PostOFFis$ in (2). This test determines whether there is a persistent effect from the feedback-on phase in the first 48 showers of condition T3 after 72 showers at the end of the feedback-off phase in T3. Using our OLS estimates $\beta_3 = -7.31$ and $\beta_4 = 0.11$, we obtain a minimal persistent effect of -0.22 L / shower after 72 showers,

Table F.1: Treatment Effects by Conditions T1–T7 Accounting for Post-Feedback Jumps

	T1-T7 (1)	T1,T2 (2)	T1,T2,T3 (3)	T1,T2,T4 (4)	T1,T2,T5 (5)	T1,T2,T6 (6)	T1,T2,T7 (7)
			Panel (a): Full Sample (Main Results from Paper)				
ON	-7.39***	-6.65***	-7.31***	-7.31***	-7.13***	-7.53***	-6.96***
	(0.70)	(1.39)	(1.10)	(1.06)	(1.04)	(1.09)	(1.05)
PostON	0.01	-0.01	-0.01	0.00	-0.01	0.00	-0.00
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
OFF	-4.85***		-7.70***	-5.44***	-5.17***	-4.68***	-3.60***
	(0.73)		(1.44)	(1.28)	(1.24)	(1.22)	(1.19)
PostOFF	0.08***		0.11***	0.07*	0.19***	0.08	0.02
	(0.02)		(0.04)	(0.04)	(0.06)	(0.09)	(0.07)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.45	0.44
Observations	86376	24648	37798	38286	36885	35862	36137
			Panel (b): Jumps Sample				
ON	-7.20***	-6.65***	-8.13***	-7.33***	-7.07***	-6.75***	-6.68***
	(0.73)	(1.39)	(1.13)	(1.06)	(1.06)	(1.08)	(1.10)
PostON	0.01	-0.01	-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	-4.95***		-8.17***	-5.56***	-5.05***	-5.18***	-3.60***
	(0.78)		(1.45)	(1.28)	(1.26)	(1.23)	(1.34)
PostOFF	0.08***		0.10***	0.07**	0.18***	0.13	0.00
	(0.02)		(0.04)	(0.04)	(0.06)	(0.09)	(0.07)
R-Squared	0.43	0.44	0.42	0.44	0.43	0.46	0.45
Observations	79320	24648	35856	37849	36705	34435	33067
			Panel (c): Jumps Sample (Controlling for Post–Feedback Jumps)				
ON	-7.28***	-6.65***	-8.05***	-7.19***	-6.96***	-7.02***	-6.71***
	(0.73)	(1.39)	(1.12)	(1.06)	(1.07)	(1.08)	(1.09)
PostON	0.00	-0.01	-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	-5.04***		-7.46***	-4.63***	-5.99***	-4.90***	-3.66***
	(0.80)		(1.29)	(1.17)	(1.18)	(1.22)	(1.34)
PostOFF	0.10***		0.08***	0.15***	0.05	0.23***	0.16**
	(0.02)		(0.03)	(0.04)	(0.05)	(0.09)	(0.07)
R-Squared	0.44	0.44	0.42	0.45	0.43	0.46	0.45
Observations	79320	24648	35856	37849	36705	34435	33067

which is just 0.4% of mean baseline shower usage of 57 L / shower. The hypothesis test implies a statistically insignificant with $F(1,239) = 0.02$ and $p = 0.90$. There is no evidence of a long-run level shift in consumption after feedback is turned off that an experimentation and learning mechanism would imply.

Larger consumption responses in the first feedback cycle

We can also use experimental conditions T4-T7 and their feedback on/off cycles to test for experimentation and learning effects. Suppose that an individual "learns" about a new optimal

Table F.2: Treatment Effects by Condition T3–T7 for Experimentation and Learning

	Panel (a): Main Results from Paper			
	T1,T2,T4 24/48 on/off (1)	T1,T2,T5 12/24 on/off (2)	T1,T2,T6 6/12 on/off (3)	T1,T2,T7 3/15 on/off (4)
ON	-7.31***	-7.13***	-7.53***	-6.96***
	(1.06)	(1.04)	(1.09)	(1.05)
PostON	0.00	-0.01	0.00	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)
OFF	-5.44***	-5.17***	-4.68***	-3.60***
	(1.28)	(1.24)	(1.22)	(1.19)
PostOFF	0.07*	0.19***	0.08	0.02
	(0.04)	(0.06)	(0.09)	(0.07)
R-Squared	0.44	0.43	0.45	0.44
Observations	38286	36885	35862	36137
	Panel (b): Results for Cycling Feedback in T4–T7 Not Controlling for T2			
ON	-8.35***	-7.98***	-8.15***	-6.62***
	(1.34)	(1.59)	(1.77)	(1.87)
PostON	0.08*	0.06	-0.13	-0.35
	(0.05)	(0.12)	(0.22)	(0.68)
OFF	-5.62***	-4.95***	-4.91***	-3.61***
	(1.52)	(1.46)	(1.51)	(1.33)
PostOFF	0.06*	0.15**	0.08	0.02
	(0.04)	(0.07)	(0.09)	(0.07)
R-Squared	0.46	0.45	0.48	0.46
Observations	24958	23557	22534	22809
	Panel (c): Results for Cycling Feedback in T4–T7 Not Controlling for T2, Allowing the First Feedback Cycle to Have a Differential ON Effect			
ON	-8.53***	-7.60***	-7.89***	-5.96***
	(1.82)	(2.04)	(1.89)	(2.06)
PostON	0.08	0.08	-0.12	-0.28
	(0.05)	(0.12)	(0.21)	(0.68)
OFF	-5.66***	-4.86***	-4.82***	-3.43**
	(1.58)	(1.53)	(1.54)	(1.37)
PostOFF	0.06*	0.15**	0.08	0.02
	(0.04)	(0.07)	(0.09)	(0.07)
FirstON	0.26	-0.90	-0.99	-2.89*
	(1.52)	(1.72)	(1.45)	(1.65)
R-Squared	0.46	0.45	0.48	0.46
Observations	24958	23557	22534	22809

level of shower water usage when provided feedback. Then, the first feedback-on phase in a given experimental condition will have both a learning and salience effect. After experimentating and learning about a new optimal consumption level during the first feedback-on phase, subsequent feedback-on phases will only entail salience effects. Therefore, to test for experimentation and learning effects, we adapt our treatment effects regression from equation (2) as follows:

$$y_{is} = \eta_i + \beta_1 ON_{is} + \beta_2 PostON_{is} + \beta_3 OFF_{is} + \beta_4 PostOFF_{is} + \beta_5 FirstON_{is} + \tau_k + \varepsilon_{is}, \qquad \text{(F.3)}$$

where $FirstON_{is}$ equals 1 if feedback is on and it is the first feedback-on cycle for household i, where we consider households in experimental conditions T4-T7 with multiple feedback on/off phases.The following test allows us to test for an experimentation and learning mechanism:

$$H_0 : \beta_5 = 0 \text{ vs. } H_1 : \beta_5 \neq 0$$

Panels (b) and (c) of Table F.2 present our test results. Panel (b) produces analogous estimates to our primary treatment effects from the paper in panel (a), except we exclude including T2 in the estimation samples in columns (4)-(7) of the table. Doing so allows us to focus on the differences in feedback effects across the feedback-on phases *within* conditions T4-T7 at the cost of not controlling for secular trends in feedback effects by including T2 households in the samples.[38] Panel (c) then adds the $FirstON_{is}$ regressor. Its coefficient estimate is statistically insignificant across all columns. Moreover, comparing the *ON* coefficients in panels (b) and (c), we find little impact on the estimates from including $FirstON_{is}$ in the regression. This finding implies that salience effects identified by later feedback-on phases are robust to controlling for any experimentation and learning during the first phase. In sum, as with the time-varying treatment effects from experimental condition T3, our results in panels (b) and (c) of Table F.2 suggest minimal impacts of experimentation and learning in our setting.

[38]Including T2 households in the samples has no impact on our results, but makes it harder to interpret the results from the hypothesis tests regarding β_5.