Comparative feedbacks under incomplete information

42nd IAEE International Conference - Montreal

Nicolas Astier
(nicolas.astier@stanford.edu)

May 2019

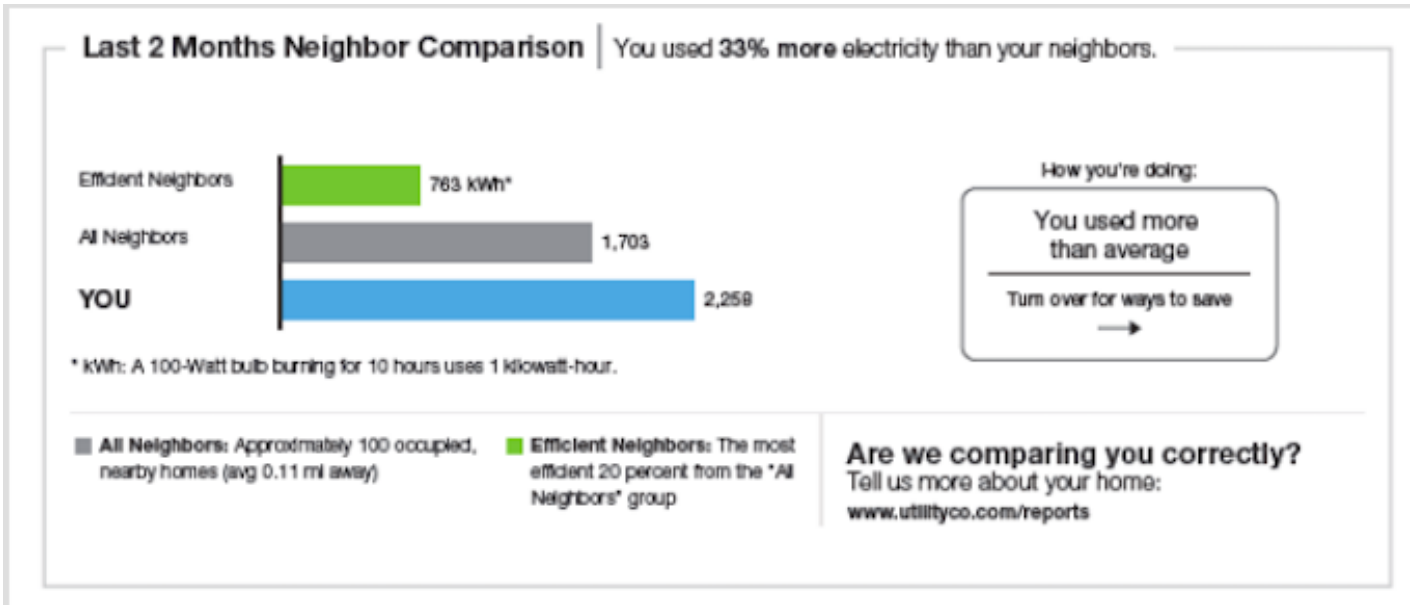
Motivation

Motivation

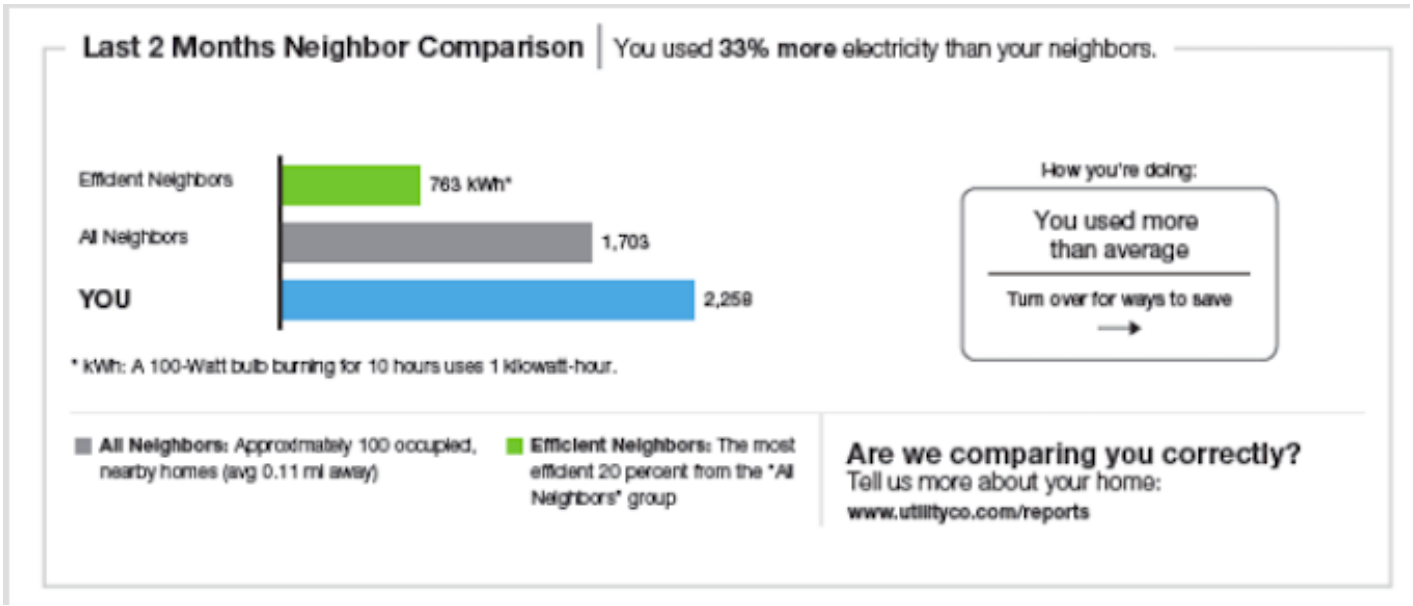
"*Social influences come in two basic categories.* ***The first involves information****. If many people do something or think something, their actions and their thoughts convey information about what might be best for you to do or think.* ***The second involves peer pressure****. If you care about what other people think about you [...], then you might go along with the crowd to avoid their wrath or curry their favor.*"

(Thaler and Sunstein, 2008)

Comparative feedbacks for residential energy consumers



Comparative feedbacks for residential energy consumers



Millions of residential consumers are receiving comparative feedbacks, a popular **nudge** aimed at decreasing energy consumption.

Two main reasons to change your behavior

When choices are made under incomplete information, peer comparisons may induce people to change their behavior for at least **two reasons**:

- **An informative channel:** agents will update their beliefs about *the way their choices map into outcomes* (monthly bill, daily comfort, etc.), that is the net utility they would derive privately from choosing a particular action.
- **A normative channel:** agents will update their beliefs about *the way their choices map into self or social esteem*, that is how an external observer would assess their social status from their actions.

Two main reasons to change your behavior

When choices are made under incomplete information, peer comparisons may induce people to change their behavior for at least **two reasons**:

- **An informative channel:** agents will update their beliefs about *the way their choices map into outcomes* (monthly bill, daily comfort, etc.), that is the net utility they would derive privately from choosing a particular action.
- **A normative channel:** agents will update their beliefs about *the way their choices map into self or social esteem*, that is how an external observer would assess their social status from their actions.

$\Rightarrow$ **This paper uses an online experiment to study the informative channel of comparative feedbacks**, in the absence of any measurable peer pressure.

Link with residential energy consumption

- **What is your monthly electricity bill?**

Link with residential energy consumption

- **What is your monthly electricity bill?**
  44% of 1721 Dutch households “had no idea” (Brounen et al. (2013)).
- **How much electricity does your coffee maker consume compared to your washing machine?**

Link with residential energy consumption

- **What is your monthly electricity bill?**
  44% of 1721 Dutch households “had no idea” (Brounen et al. (2013)).
- **How much electricity does your coffee maker consume compared to your washing machine?**
  Consumers tend to hold biased beliefs about the electricity consumption of individual appliances (Wood and Newborough (2003), Attari et al. (2010)).

Link with residential energy consumption

- **What is your monthly electricity bill?**
  44% of 1721 Dutch households "had no idea" (Brounen et al. (2013)).
- **How much electricity does your coffee maker consume compared to your washing machine?**
  Consumers tend to hold biased beliefs about the electricity consumption of individual appliances (Wood and Newborough (2003), Attari et al. (2010)).
- Energy audits (Armel et al., 2013) and people's response to energy crisis (Leighty and Meier, 2011) suggest that substantial savings come from unplugging "forgotten" or "spare" devices inadvertedly left on.

Link with residential energy consumption

- **What is your monthly electricity bill?**
  44% of 1721 Dutch households "had no idea" (Brounen et al. (2013)).
- **How much electricity does your coffee maker consume compared to your washing machine?**
  Consumers tend to hold biased beliefs about the electricity consumption of individual appliances (Wood and Newborough (2003), Attari et al. (2010)).
- Energy audits (Armel et al., 2013) and people's response to energy crisis (Leighty and Meier, 2011) suggest that substantial savings come from unplugging "forgotten" or "spare" devices inadvertedly left on.

$\Rightarrow$ **Incomplete information** is a pervasive feature of residential energy consumption: "*Consider groceries in a hypothetical store totally without price markings, billed via a monthly statement like US$527 for 2362 food units in April*" (Kempton and Layne, 1994)

Motivation

- Comparative feedback programs effectiveness has mainly been assessed in terms of **easily measurable metrics**, typically the decrease in energy consumption.

Motivation

- Comparative feedback programs effectiveness has mainly been assessed in terms of **easily measurable metrics**, typically the decrease in energy consumption.
- However **welfare consequences can be very different** depending on the mechanism underlying consumers' response.

Motivation

- Comparative feedback programs effectiveness has mainly been assessed in terms of **easily measurable metrics**, typically the decrease in energy consumption.
- However **welfare consequences can be very different** depending on the mechanism underlying consumers' response.

$\Rightarrow$ This paper contributes to improving our understanding of this underlying mechanism. It implements an **online experiment** that allows us to **focus on the purely informative aspect** of comparative feedbacks, in a setting that reproduces the most salient features of residential energy consumption.

Main results

- Comparative feedbacks are found to trigger **comparable or even greater changes in behaviors** than other kinds of informative and more accurate feedbacks, **despite the absence of normative pressure**.
- **Learning about transaction costs** may represent a consistent explanation: comparative feedbacks may convey the idea that changing one's behavior should not be too difficult.
- Learning about unknown parameters of the choice environment is indeed found to be very difficult when customers only receive infrequent bills based on aggregate consumption.

Overview of the experimental design

We create a controlled environment that reproduces important stylized features of residential energy consumption:

1. Participants do not know the price of some services they may consume.
2. By default participants only receive a "bill" aggregating consumption over time and services.

$\Rightarrow$ Importantly, we take advantage of implementing a lab/online experiment to use a **randomized control trial design**. A key advantage of our experimental set up compared to field experiments is to enable us to **perfectly observe the payoff function of participants**.

Overview of the experimental design

We create a controlled environment that reproduces important stylized features of residential energy consumption:

1. Participants do not know the price of some services they may consume.
2. By default participants only receive a "bill" aggregating consumption over time and services.

$\Rightarrow$ Importantly, we take advantage of implementing a lab/online experiment to use a **randomized control trial design**. A key advantage of our experimental set up compared to field experiments is to enable us to **perfectly observe the payoff function of participants**.

Participants are randomized between:

1. A **complete information environment** where the parameters needed for optimization are known.
2. An **incomplete information environment** where some parameters are unknown.

Framing

- Online game where participant have to feed a virtual pet for two virtual weeks:

- Given weekly budget, constant prices.
- Food choices between a **numeraire good**, and a **good (potentially) earning inframarginal utility**. Leftover money spent on the numeraire good.
- **Final score = Total virtual utility of the pet**
- Incentivized via a monetary reward based on the final score.
- **Same parameters** (prices, utility) for all participants:
  $\Rightarrow$ no horizontal differentiation in tastes.

Three treatments

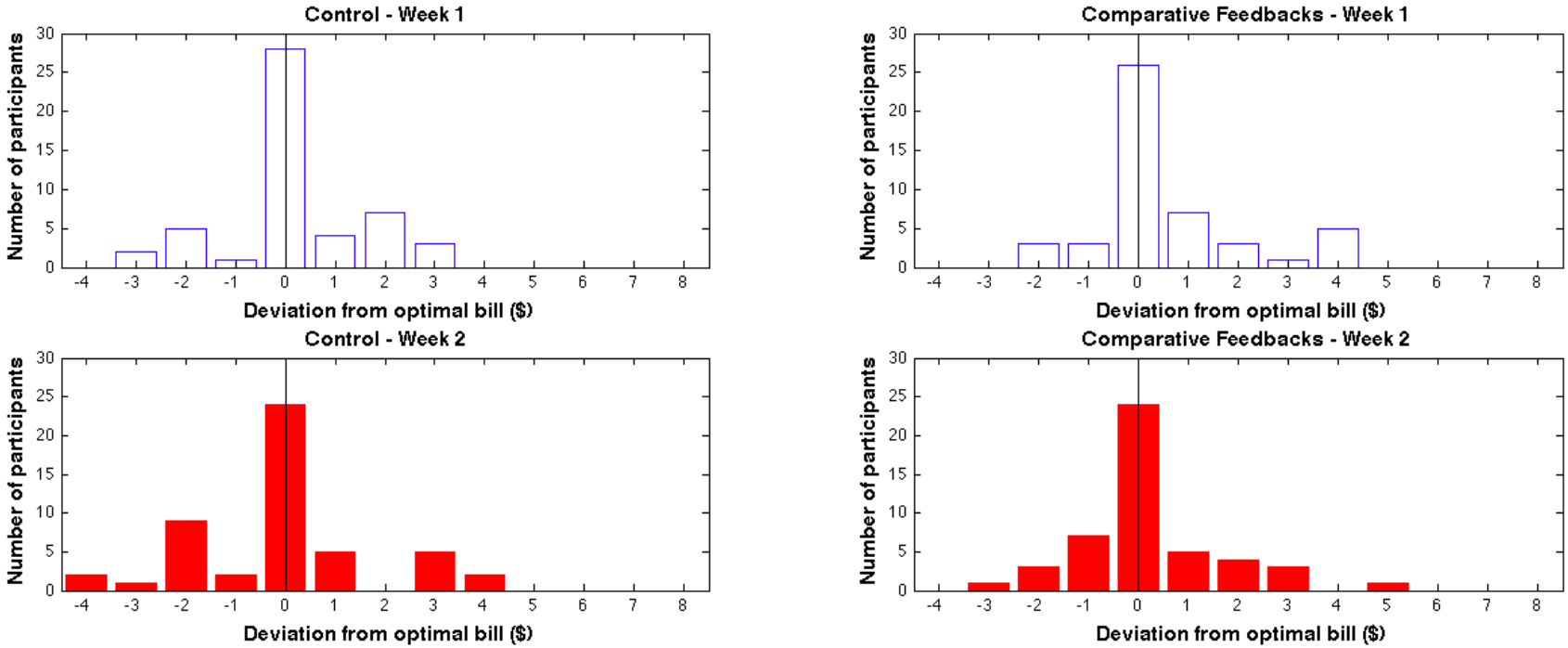
At the end of week 1, three randomly assigned treatments:

- Comparative feedback on week 1 bill;
- Information about the score-maximizing bill;
- Warning to outliers.

The experiment was implemented on Amazon Mechanical Turk.

	Control	Comparative feedback	Optimal bill	Warning
CI	50	48		
II	50	51	50	52

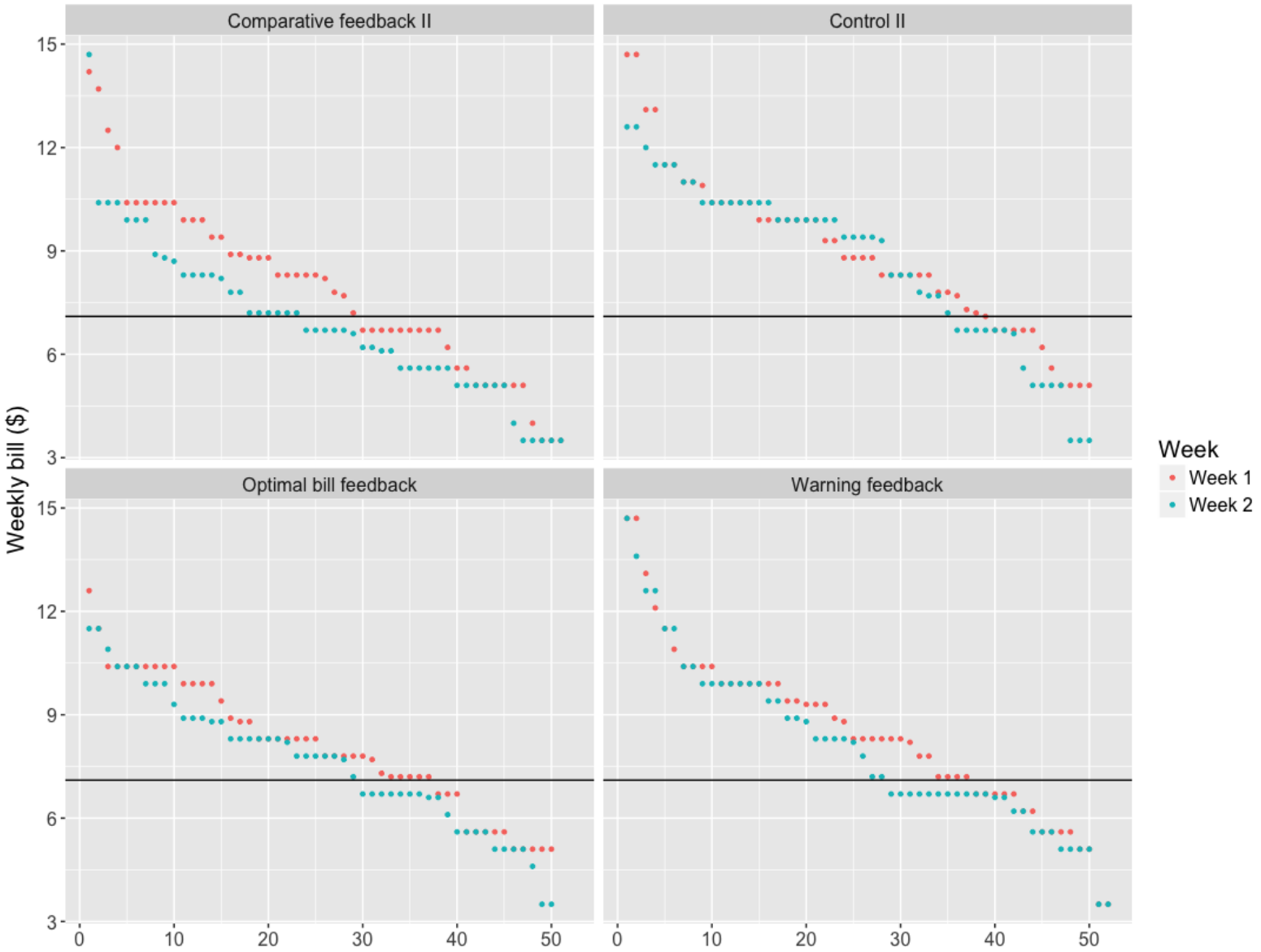
Absence of normative pressure under CI



Second-week bills in red, treatment on the right panels.

⇒ **Our experiment thus credibly focuses on the role played by the informative channel.**

Graphical results under incomplete information (all)



Summary of statistical results

Week 1		Week 2		Bill $w2_T$ - Bill $w2_C$ (controlling for w1 bill)
Information	Average bill	Information	Average bill	
Prices	7.27	Prices	7.11	
Prices	7.70	Prices and a comp. feedback	7.36	0.01 (0.28)
	8.94		8.63	
	7.90	Comp. feedback	6.91	$-1.36\ (0.48)^{***}$
	8.09	Optimal bill	7.62	$-0.63\ (0.41)$
	8.47	Warning	8.06	$-0.45\ (0.49)$

Table: Summary of obtained results ($^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$)

- Comparative feedbacks have triggered greater responses than feedbacks framed without referring to other participants.
- Quite surprising result since the latter feedbacks conveyed accurate information while comparative feedbacks only conveyed information about the realized bills of other participants.

Cognitive costs and incomplete information

- Many evidence suggest a **widespread reliance on heuristics**.
- Learning was indeed difficult: no accurate guess of unknown parameters out of 200+ guesses!

"What do you think are the chances (in %) that you could have increased the number of smiley units have collected so far by making different food choices?"

Treat.	CF	Opt. Bill	Warn. (all)	Warn. (treated)
Const.	43.71 (10.10)***	40.02 (10.96)***	41.22 (9.79)**	95.00 (22.20)***
Treat.	13.40 (4.06)***	8.11 (4.49)*	4.18 (4.47)	11.94 (6.70)*
W1 bill	2.31 (1.04)**	2.72 (1.15)**	2.59 (1.01)**	−2.43 (2.02)
R^2	0.132	0.081	0.073	0.121

Effectiveness in raising awareness
(*** : $p < 0.01$, ** : $p < 0.05$, * : $p < 0.1$)

$\Rightarrow$ comparative feedbacks were more effective in **convincing participants they could improve on their first-week score**.

Impact on participants' score

Finally, we can investigate the impact of the different feedbacks on the average score reached by participants during the second week.

Treat.	CF	Opt. Bill	Warn. (all)	Warn. (treated)
Const.	$1.58\,(0.45)^{***}$	$1.62\,(0.32)^{***}$	$1.92\,(0.39)^{***}$	$2.40\,(0.83)^{***}$
Treat.	$-0.21\,(0.20)$	$0.01\,(0.18)$	$-0.14\,(0.16)$	$-0.45\,(0.24)$
W1 bill	$0.11\,(0.05)^{**}$	$0.11\,(0.03)^{***}$	$0.08\,(0.04)$	$0.04\,(0.07)$
R^2	0.108	0.072	0.056	0.096

Impact of treatments on week 2 scores
($^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$)

$\Rightarrow$ while obviously lacking external validity, this result illustrates that **bills may be a poor proxy for welfare**.

Main results

- We design an online experiment that makes it possible to credibly focus on the role played by the **purely informational content** of comparative feedbacks. Within our setting, incomplete information is indeed shown to be a necessary condition to get a significant treatment effect.
- Despite an absence of normative pressure, comparative feedbacks are found to trigger **comparable or even greater changes in behaviors** than other kinds of informative and more accurate feedbacks.
- We suggest a possible explanation for the higher effectiveness of comparative feedbacks: **learning about cognitive costs**. Comparative feedbacks may also convey the idea that changing one's behavior should not be too difficult.
- Despite being effective in decreasing "consumption", feedbacks did not make participants better off within our experiment.

Thank you for your attention!

Distracting task used

First week, day 1

Try to guess what Noney feels like doing today!
HINT:

SL--P--- -- TH- B--

LETTER(S) GUESS:
(Tip1: you can type several letters in a row)
(Tip2: to skip, just type the whole alphabet)

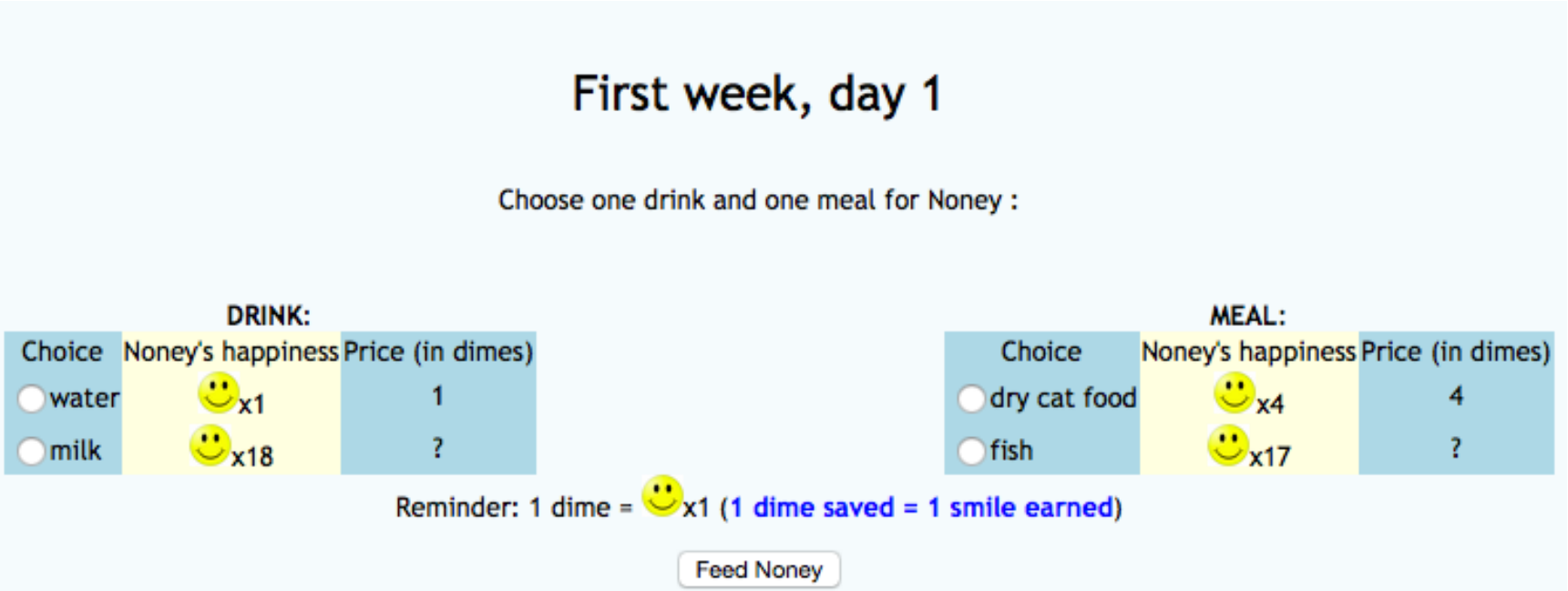
Submit letter guess

Two information environments

Complete information (CI):

Two information environments

Incomplete information (II):



Treatment 1: comparative feedbacks (both under CI and II)

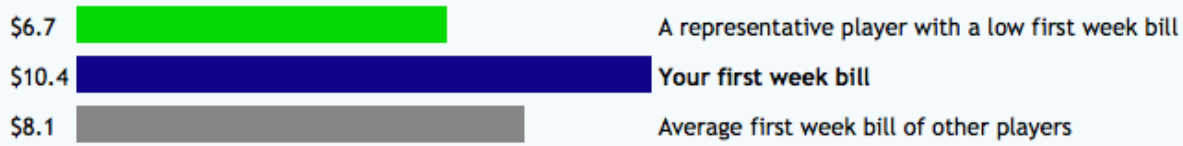
End of First Week

YOUR WEEKLY FOOD BILL IS: $10.4

You can thus buy Noney a $4.6 gift, which makes her happy by x46 units

For your information, the following graph summarizes how your bill compares to a representative sample of other players:

First week bill:



Note: the representative low-bill player is chosen such as his first week bill is lower than 80% of players, and higher than the remaining 20%.

Reminder: Milk is more expensive than water, and fish is more expensive than dry cat food.

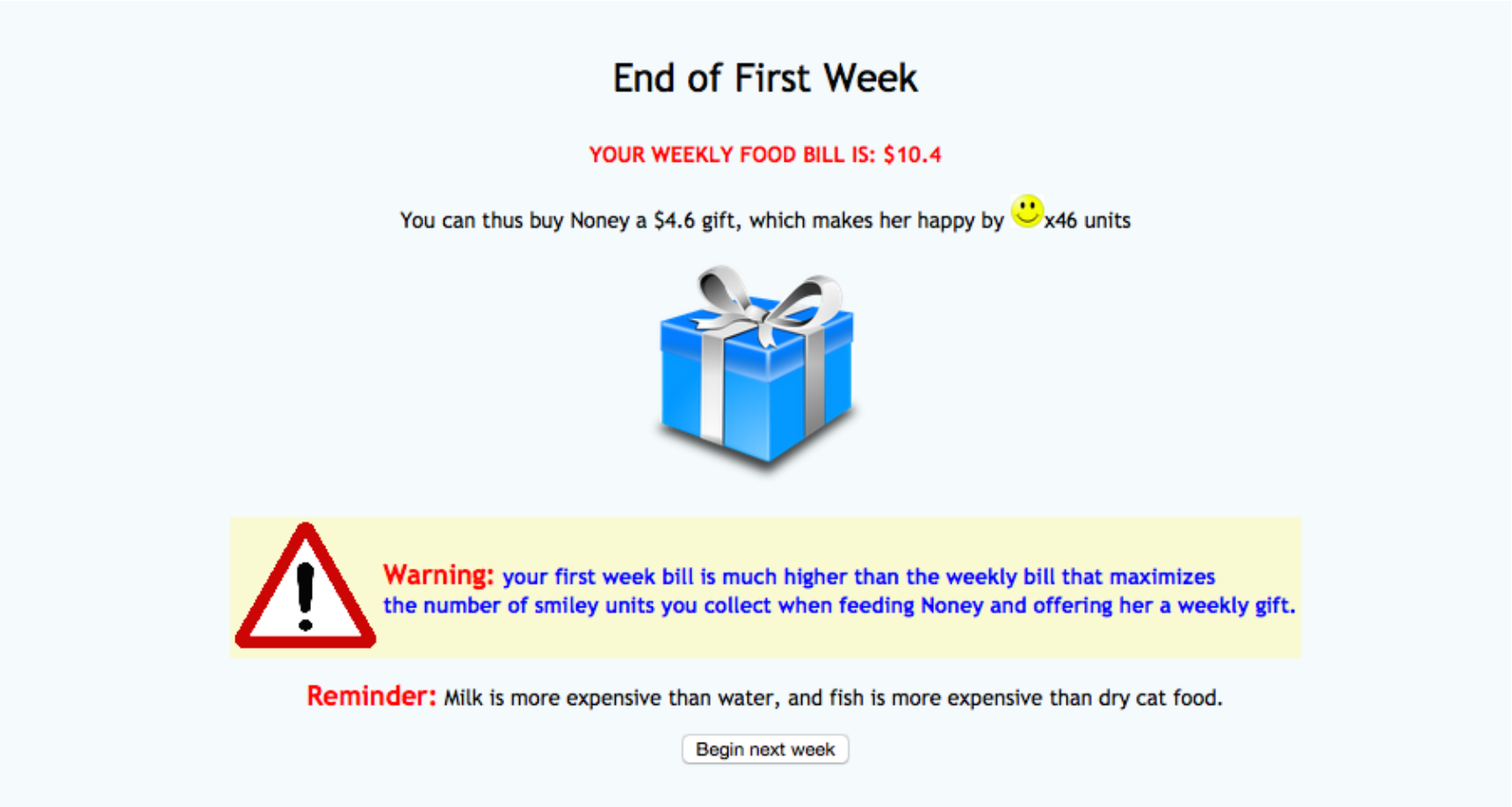
Begin next week

Treatment 2: optimal bill feedback (under II)

(warning sign was flashing)

Treatment 3: optimal bill feedback (under II)

Treatment 3: warning feedback (under II) if bill lies in the top two deciles of control groups (greater than $9.9)



(warning sign was flashing)

Validity of randomization under complete information

Summary statistics suggest the randomization procedure has worked reasonably well under complete information.

Treatment	Control	Comp. feedback
First-week bill (in $)	7.27 (1.34)	7.70 (1.60)
Time spent on tutorial (in min)	7.08 (2.84)	8.15 (4.00)
Time spent on week 1 (in min)	4.31 (1.49)	4.86 (1.36)
Tutorial passed at 1st attempt (%)	82	93.75
Sample Size	50	48

Table: Summary statistics for week 1 (CI environment)

Validity of randomization under incomplete information

Not surprisingly, the distribution of first-week bills is more dispersed under incomplete information than it was under complete information. Because of this higher dispersion in first-week outcomes, **our preferred empirical strategy will include a control for the first-week bill**.

Treatment	Control	CF	Opt. bill	Warning
1st bill (in $)	8.94 (2.31)	7.90 (2.52)	8.09 (1.91)	8.47 (2.42)
Tutorial (min)	8.06 (4.09)	7.29 (3.86)	7.27 (3.45)	7.46 (4.89)
Week 1 (min)	4.90 (3.77)	4.26 (1.95)	4.51 (1.55)	3.93 (1.86)
1st attempt (%)	90	84.3	88	76.9
Sample Size	50	51	50	52

Table: Summary statistics for week 1 (II environment)

Statistical results

Empirical strategy and main specification:

$$Y_i = \alpha + \tau R_i + \beta X_i + \epsilon_i$$

where Y_i=2nd week bill, R_i=treatment dummy, X_i=1st week bill.

Treat.	Comp. Feed.	Opt. Bill	Warn. (all)	Warn. (treated)
Const.	$5.56\,(0.98)^{***}$	$4.61\,(0.83)^{***}$	$6.32\,(0.99)^{***}$	$7.90\,(2.74)^{**}$
Treat.	$-1.36\,(0.48)^{***}$	$-0.63\,(0.41)$	$-0.45\,(0.48)$	$-2.30\,(0.66)^{***}$
W1 bill	$0.34\,(0.10)^{***}$	$0.45\,(0.08)^{***}$	$0.26\,(0.10)^{**}$	$0.14\,(0.23)$
R^2	0.236	0.229	0.076	0.265

$(^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1)$