Online Appendix: Not for Publication

The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons

Hunt Allcott and Judd B. Kessler

A Phone Survey Questionnaire

Below is the phone survey questionnaire. Programming notes and comments are in italics. Bolded headers are for organizational purposes and were not read.

Introduction

Hi. I am calling on behalf of Central Hudson Gas and Electric, your local utility. Central Hudson has been sending you Home Energy Reports since last fall, and we want to know what you think about them. Do you have about two minutes to answer some questions? If yes, Central Hudson will send you a check for up to $10.

If asked, "What is a Home Energy Report?", say: "Home Energy Reports are one-page letters that compare your natural gas use to your neighbors' use and provide energy conservation tips. Central Hudson sent up to four of these reports to the address on the account associated with this phone number between late fall 2014 and early spring 2015. Do you recall receiving any Home Energy Reports in the past nine months?"

- *If "Yes", continue to Question 1.*
- *If "No", or if the customer otherwise says "I don't remember receiving any Home Energy Reports," say:* "Is there someone else in the household who may have seen these reports come in the mail? If so, may I speak to him or her?" *If there is no one else who might have seen the reports, terminate call and code response as "Does not remember Home Energy Reports." If there is someone else but not available, record that person's name and attempt to call him/her later.*

If the caller indicates that he/she has already answered these questions in a mail survey, then skip questions 1 and 2 and say: "Thank you for responding to our mail survey. We have a couple of follow-up questions that are better to ask by phone." *Then continue to Question 3.*

Question 1

To start, I'm going to ask three questions where you'll choose between some combination of continuing Home Energy Reports and receiving checks for different amounts of money. These are unusual questions, but they're designed to tell us how much you value the Reports. These are real questions: Central Hudson will use a lottery to pick one question and will actually mail you what you chose, so please answer carefully.

Survey Version B only: "Remember that Home Energy Reports compare your energy use to your neighbors' use.

Survey Version C only: "Remember that Home Energy Reports help you to reduce your environmental impact."

a. Which would you prefer: 4 more Home Energy Reports PLUS a $10 check, OR a $1 check?

b. Which would you prefer: 4 more Home Energy Reports PLUS a $10 check, OR a $5 check?
c. Which would you prefer: 4 more Home Energy Reports PLUS a $10 check, OR a $9 check?
d. Which would you prefer: 4 more Home Energy Reports PLUS a $10 check, OR a $10 check?
e. Which would you prefer: 4 more Home Energy Reports PLUS a $9 check, OR a $10 check?
f. Which would you prefer: 4 more Home Energy Reports PLUS a $5 check, OR a $10 check?
g. Which would you prefer: 4 more Home Energy Reports PLUS a $1 check, OR a $10 check?

If consumers have consistent preferences, we would not need to ask all seven MPL questions because answers to some imply answers to others. Questions 1a-1g were asked in the following order:

Ask 1d first
If 1d="HER+$10", then 1f
If 1f="HER+$5", then 1g
If 1f="$10", then 1e
If 1d="$10", then 1b
If 1b="HER+$10", then 1c
If 1b="$5", then 1a

Question 2

Think back to when you received your first Home Energy Report. Did the Report say that you were using more or less energy than you thought?

a. Much less than I thought
b. Somewhat less than I thought
c. About what I thought
d. Somewhat more than I thought
e. Much more than I thought

Question 3

Do you think that receiving four more Home Energy Reports this fall and winter would help you reduce your natural gas use by even a small amount?

a. Yes
b. No

If Yes: How much money do you think you would save on your natural gas bills if you receive four more Reports compared to if you do not receive them?

If necessary: "We just want to know your best guess."

Note to enumerators: Prompt for a dollar value, not a percentage. If necessary: "I'm supposed to ask for your best guess of how many dollars you'd save in total."

Question 4

Since last fall, Central Hudson sent up to four Home Energy Reports to many households like yours. For the average household, how much money do you think these Reports have helped them save on their natural gas bills?

If necessary: "We just want to know your best guess."

Note to enumerators: Prompt for a dollar value, not a percentage. If necessary: "I'm supposed to ask for your best guess of total dollar savings since last fall."

Question 5

How would you like the Reports if they did not have the bar graph comparing your energy use to your neighbors' use?

a. Much less
b. Somewhat less
c. About the same
d. Somewhat more
e. Much more

Question 6

Some people feel either inspired or pressured when they see their Home Energy Reports. Did you feel inspired, pressured, neither, or both?

a. Inspired
b. Pressured
c. Neither
d. Both

Question 7

Some people feel either proud or guilty when they see their Home Energy Reports. Did you feel proud, guilty, neither, or both?

a. Proud
b. Guilty
c. Neither
d. Both

Question 8

To what extent do you agree or disagree with the following statement: "The Home Energy Reports gave useful information that helped me conserve energy."

a. Strongly agree
b. Agree
c. Neither
d. Disagree
e. Strongly disagree

Question 9

Do you have any other comments about the Home Energy Reports that you'd like to share?

Open response, please write down as much as possible.

B Data Appendix

Table A1: **Balance Tests (Page 1)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	Baseline use (ccf/day)	ln(Income)	ln(Net worth)	ln(House value)	Education (years)	Male	ln(Age)	Retired	Married	Rent
				Panel A: Home Energy Report Recipient/Control						
Recipient	0.013	-0.011	0.0058	-0.029	-0.032	-0.0026	-0.0044	0.0028	-0.0064	0.0039
	(0.024)	(0.012)	(0.024)	(0.029)	(0.035)	(0.0077)	(0.0051)	(0.0030)	(0.0079)	(0.0069)
Observations	19,921	19,927	15,557	16,741	19,475	16,811	17,282	16,728	15,406	17,561
				Panel B: Survey Group						
Mail follow-up	-0.027	0.0073	-0.054	-0.043	-0.0076	0.0093	-0.0092	-0.012	0.0068	-0.012
	(0.037)	(0.018)	(0.036)	(0.044)	(0.054)	(0.012)	(0.0078)	(0.0050)**	(0.012)	(0.011)
Comparison cue	-0.039	0.00063	-0.043	-0.062	-0.025	-0.014	0.0015	0.0057	-0.00061	-0.011
	(0.043)	(0.021)	(0.042)	(0.051)	(0.063)	(0.014)	(0.0090)	(0.0054)	(0.014)	(0.012)
Environmental cue	0.011	0.0056	0.012	-0.049	0.0016	-0.018	0.015	0.011	0.0058	-0.0100
	(0.043)	(0.021)	(0.042)	(0.051)	(0.063)	(0.014)	(0.0090)*	(0.0055)**	(0.014)	(0.012)
Observations	9436	9439	7466	7965	9226	8036	8251	8004	7255	8371
F-test p-value	0.54	0.97	0.26	0.45	0.97	0.46	0.18	0.023	0.91	0.53

Notes: This table presents tests of balance on observables between randomly-assigned groups. Samples in Panel A include the full HER recipient and control groups, while samples in Panel B are limited to the households that were sent Home Energy Reports and were thus eligible for our surveys. Observation counts differ between columns because regressions include only non-missing observations of the dependent variable. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A2: **Balance Tests (Page 2)**

	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Dependent variable:	Single family home	ln(House age)	Democrat	Hybrid auto share	Green consumer	Wildlife donor	Profit score	Buyer score	Mail responder	Home improvement interest
				Panel A: Home Energy Report Recipient/Control						
Recipient	0.0018	-0.032	0.0018	0.021	0.0042	0.0033	-0.0033	-0.022	-0.012	0.00024
	(0.0070)	(0.016)*	(0.0082)	(0.047)	(0.0052)	(0.0036)	(0.014)	(0.016)	(0.0069)*	(0.0051)
Observations	17,734	14,885	18,080	19,728	18,883	16,728	19,784	14,967	17,734	16,728
				Panel B: Survey Group						
Mail follow-up	-0.0097	-0.021	0.0037	0.0013	-0.0053	-0.0094	0.018	-0.017	0.0017	-0.025
	(0.011)	(0.025)	(0.013)	(0.053)	(0.0081)	(0.0058)	(0.022)	(0.025)	(0.011)	(0.0081)***
Comparison cue	0.0048	-0.047	-0.00065	-0.10	0.0058	-0.0035	-0.022	0.0036	-0.012	-0.0037
	(0.012)	(0.029)	(0.015)	(0.071)	(0.0094)	(0.0066)	(0.025)	(0.029)	(0.012)	(0.0092)
Environmental cue	0.011	-0.048	0.0038	-0.18	-0.0056	-0.0045	0.010	-0.014	-0.0079	-0.0015
	(0.012)	(0.029)*	(0.015)	(0.13)	(0.0092)	(0.0066)	(0.026)	(0.029)	(0.012)	(0.0092)
Observations	8464	7109	8617	9340	8977	8004	9377	7143	8464	8004
F-test p-value	0.67	0.23	0.98	0.35	0.59	0.36	0.49	0.84	0.80	0.020

Notes: This table presents tests of balance on observables between randomly-assigned groups. Samples in Panel A include the full HER recipient and control groups, while samples in Panel B are limited to the households that were sent Home Energy Reports and were thus eligible for our surveys. Observation counts differ between columns because regressions include only non-missing observations of the dependent variable. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A3: **Survey Response Counts by Attempt**

	(1)	(2)
Attempt	Mail	Phone
1	402	523
2	497	358
3		229
4		172
5		163
6		83
7		80
8		80
Overall	899	1690

Notes: For the mail survey, attempt 1 refers to the survey included in the final Home Energy Report, and attempt 2 refers to the follow-up survey sent to 2/3 of households. For the phone survey, attempt refers to the number of times that the phone number was called before completing the survey.

Table A4: **Correlations of Willingness-to-Pay with Qualitative Survey Responses**

	(1)	(2)	(3)	(4)	(5)	(6)
Expected savings	0.11					
	(0.0089)***					
Like without comparisons		-1.05				
		(0.16)***				
Useful info			2.26			
			(0.18)***			
Inspired				3.36		
				(0.38)***		
Pressured				-1.02		
				(0.50)**		
Proud					1.18	
					(0.41)***	
Guilty					1.39	
					(0.49)***	
Positive comment						4.34
						(0.44)***
Observations	1365	1581	1570	1571	1571	2137
R^2	0.094	0.026	0.093	0.047	0.011	0.042

Notes: Data are the unweighted sample of phone survey responses. Dependent variable is willingness-to-pay. The independent variables in columns 1-6 are from questions 3, 5, 8, 6, 7, and 9, respectively. Expected savings is winsorized at $50. Columns 2 and 3 consider the five-point Likert scale responses to questions 5 and 8, which we code as integers $\{-2, -1, 0, 1, 2\}$. The sample in column 6 includes both mail and phone survey respondents: the phone survey enumerators transcribed responses to question 9, and we also transcribed the 30 unsolicited comments written on the mail survey. The variable "Positive comment" takes value 1 for positive comments about HERs, -1 for negative comments, and 0 for neutral or no comments. Sample sizes vary due to item non-response. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A5: **Correlations of Negative Willingness-to-Pay with Qualitative Survey Responses**

	(1)	(2)	(3)	(4)	(5)	(6)
Expected savings	-0.0059					
	(0.00057)***					
Like without comparisons		0.059				
		(0.011)***				
Useful info			-0.14			
			(0.011)***			
Inspired				-0.18		
				(0.024)***		
Pressured				0.10		
				(0.034)***		
Proud					-0.11	
					(0.027)***	
Guilty					-0.026	
					(0.033)	
Positive comment						-0.22
						(0.026)***
Observations	1365	1581	1570	1571	1571	2137
R^2	0.070	0.019	0.089	0.037	0.011	0.025

Notes: Data are the unweighted sample of phone survey responses. Dependent variable is an indicator for negative willingness-to-pay. The independent variables in columns 1-6 are from questions 3, 5, 8, 6, 7, and 9, respectively. Expected savings is winsorized at $50. Columns 2 and 3 consider the five-point Likert scale responses to questions 5 and 8, which we code as integers $\{-2, -1, 0, 1, 2\}$. The sample in column 6 includes both mail and phone survey respondents: the phone survey enumerators transcribed responses to question 9, and we also transcribed the 30 unsolicited comments written on the mail survey. The variable "Positive comment" takes value 1 for positive comments about HERs, -1 for negative comments, and 0 for neutral or no comments. Sample sizes vary due to item non-response. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A6: **Within-Household Correlations of Survey Responses**

	(1)	(2)	(3)	(4)
Dependent variable	WTP from first mail survey	WTP from phone survey	1(WTP from phone survey<0)	Belief update from phone survey
WTP from second mail survey	0.819			
	(0.080)***			
WTP from mail survey		0.440		
		(0.072)***		
1(WTP from mail survey<0)			0.362	
			(0.071)***	
Belief update from mail survey				0.500
				(0.064)***
Observations	87	224	224	259
R^2	0.584	0.206	0.132	0.217

Notes: The sample for column 1 is households that returned more than one mail survey with valid WTP. The sample for columns 2-4 is households that responded to both mail and phone surveys. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

C Appendix: Effects of HERs on Energy Use, Retail Expenditures, and Retailer Net Revenues

To estimate the effect of Home Energy Reports on energy use, we limit the sample to post-treatment data and control for pre-treatment usage. The first HERs were generated on October 13th, 2014, and first HERs had been generated for 61 percent of households by November 3rd, and 98 percent by December 8th. Post-treatment is defined as any meter read after November 1, 2014.

Y_{it} is household i's average natural gas use (in hundred cubic feet (ccf) per day) over the billing period ending on date t, and T_i is an indicator for whether household i was randomly assigned to the initial HER recipient group. We define S_{st} as indicators for whether date t falls within a group of months s, which we will implement in two different ways below, and we allow treatment effect τ_s to vary by s. We define the baseline period as the earliest 365 days in the data: July 1, 2013 through June 30, 2014. $\tilde{Y}_{it}$ is "baseline usage" — more specifically, the average daily usage from the meter read in the baseline period that most closely corresponds to billing date t. For example, if t is October 14, 2015, $\tilde{Y}_{it}$ is the average daily usage from the meter read date closest to October 14, 2013. Because meters are read on a very regular bi-monthly basis, we have fairly precise matches that help account for seasonality.[27] ν_m allows separate coefficients on $\tilde{Y}_{it}$ by the month of sample that contains date t, and ω_m is a vector of month of sample indicators. The estimating equation is:

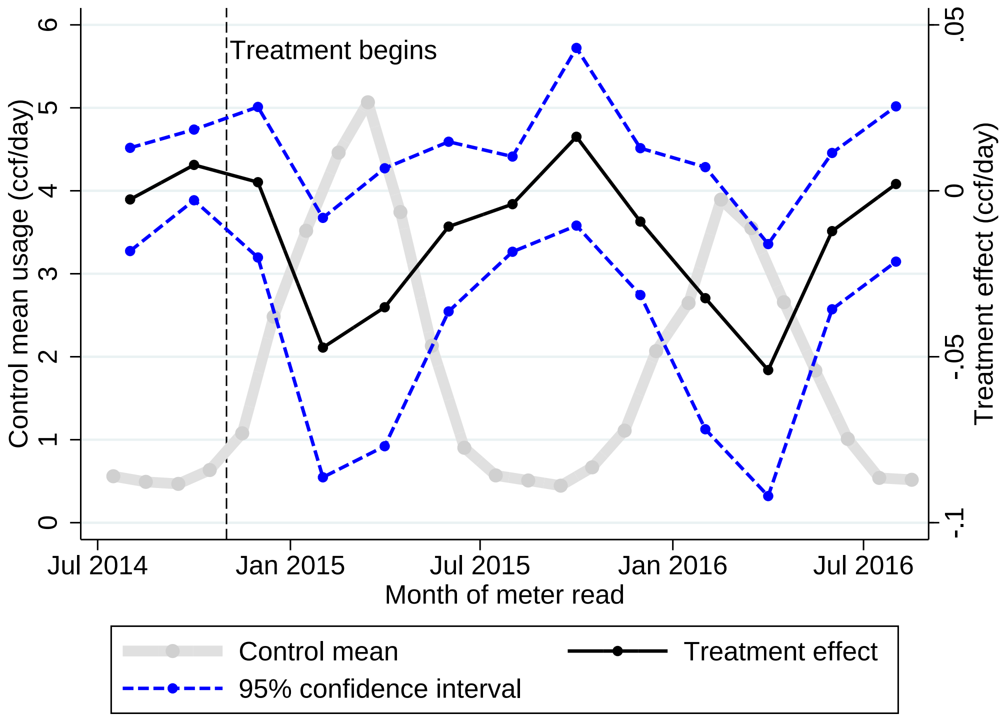
$$Y_{it} = \sum_s \tau_s S_{st} T_i + \nu_m \tilde{Y}_{it} + \omega_m + \varepsilon_{it}. \qquad (11)$$

Standard errors are clustered by household to allow for arbitrary serial correlation.

Figure A1 graphically illustrates the basic results. The thick grey line plots control group mean usage in each month of the sample, illustrating considerable seasonality. Usage is lowest during the bimonthly billing periods ending in July through October, and usage is about five times higher during the bimonthly billing periods ending in November through June. The thin black line and confidence intervals are estimates of treatment effects τ_s, where s here indexes each pair of months after the baseline period ends on June 30, 2014. The several months of pre-treatment observations allow us to test for spurious pre-treatment effects, and there are indeed zero statistical effects for meters read in July through October 2014. There are also zero statistical effects for meters read in November and December. We then see strong seasonality in the treatment effects: as much as a 0.05 ccf per day reduction in the winter periods, and zero statistical effects in any of the summer billing periods ending in July through October. This seasonality is standard in natural gas energy conservation programs: households cannot conserve much natural gas when they are not using much in the first place.

[27] Since natural gas is primarily used for heating, usage is highly seasonal, as illustrated in Figure A1. Thus, controlling for seasonal fluctuations is crucial for improving statistical efficiency. Note that estimating in logs and transforming the percent savings back into levels is not a consistent estimator of the level of average savings due to Jensen's Inequality. For this reason, Allcott (2011, 2015) and Allcott and Rogers (2014) estimate effects in levels.

Figure A1: **Effects of Home Energy Reports on Natural Gas Use**



Notes: This figure presents control group average natural gas use and the treatment effects of Home Energy Reports. Dependent variable is natural gas use in ccf/day, where "ccf" means hundred cubic feet. For context, the average marginal retail price is \$0.99/ccf during the program's first winter and \$0.80/ccf during the program's second winter. Observations weighted by billing period duration. Confidence intervals are based on robust standard errors, clustered by household.

Appendix Table A7 presents estimates of Equation (11). In all columns, we weight each observation of daily usage by the duration of the billing period (i.e., the number of days between natural gas meter reads, which is typically about two months), which gives average treatment effects in ccf/day. In columns 3 and 4, we multiply this duration weight by additional household weights for extrapolation, as discussed below. As suggested by the graphical results in Figure A1, we estimate separate treatment effects for four month groups $s = \{1, 2, 3, 4\}$: winter (November-June) of 2014-2015, summer (July-October) of 2015, winter of 2015-2016, and summer of 2016, respectively. Column 1 presents intent-to-treat effects: the average effect over time for households assigned to the treatment group.

Table A7: **Effects of Home Energy Reports on Natural Gas Use**

	(1)	(2)	(3)	(4)
Specification:	OLS	IV	IV	IV
Assigned to treatment ×	-0.0228	-0.0228	-0.0230	-0.0249
winter 2014-2015	(0.0117)*	(0.0117)*	(0.0118)*	(0.0128)*
Assigned to treatment×	0.00611	0.00611	0.00618	0.00836
summer 2015	(0.00807)	(0.00807)	(0.00809)	(0.00938)
Assigned to treatment ×	-0.0264			
winter 2015-2016	(0.0115)**			
Assigned to treatment×	0.00581			
summer 2016	(0.0112)			
2nd-year recipient ×		-0.0269	-0.0271	-0.0315
winter 2015-2016		(0.0117)**	(0.0117)**	(0.0121)***
2nd-year recipient ×		0.00592	0.00584	0.00399
summer 2016		(0.0114)	(0.0114)	(0.0110)
Observations	200,540	200,540	200,540	200,540
R^2	0.853	0.853	0.853	0.859
Weights	Duration	Duration	Duration ×	Duration ×
			IPW for $\mathcal{P}_n$	IPW for $\mathcal{P}_s$

Notes: This table presents estimates of Equation (11), using post-treatment data only. Dependent variable is natural gas use in hundred cubic feet (ccf) per day. For context, control group sample mean usage is 2.07 ccf/day, and the average marginal retail price is \$0.99/ccf during the program's first winter and \$0.80/ccf during the program's second winter. Columns 2-4 are IV regressions, where we instrument for 2nd-year recipient × winter 2015-2016 and 2nd-year recipient × summer 2016 with Assigned to treatment × winter 2015-2016 and Assigned to treatment × summer 2016. Columns 1 and 2 weight by billing period duration. Column 3 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_n$ of treatment group households that did not opt out before the second year. Column 4 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_s$ of treatment group households that did not opt out and returned a survey with valid willingness-to-pay. Robust standard errors, clustered by household, in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Define R_i as an indicator for whether household i was sent HERs in the program's second year. $R_i = T_i$, except that $R_i = 0$ for the 162 households that opted out or were dropped due to survey responses. Column 2 presents results of an instrumental variables (IV) regression where instead of the second year winter and summer treatment assignment indicators S_3T_i and S_4T_i in Equation (11), we substitute S_3R_i and S_4R_i and then instrument with S_3T_i and S_4T_i. Under the "no persistence" assumption, i.e. that HERs sent during a given winter only affect energy use in that same winter, this IV regression delivers the local average treatment effects of the second year of HERs. Because there are no always-takers (i.e., no households receive HERs in the absence of the program), $R_i = 1$ is an indicator for being a complier.

In column 3, we re-weight compliers have the same observable characteristics as target population

$\mathcal{P}_n$, the households that would normally receive reports in the program's second year.[28] In column 4, we re-weight compliers to match $\mathcal{P}_s$, the subset of households that responded to the survey and did not opt out.[29] The estimates are almost exactly the same in the first three columns: zero statistical effects in the summers, and reductions of 0.026 to 0.027 ccf per day in winter 2015-2016. The re-weighting and IV estimation hardly change the estimates because 98.4 percent of households are compliers, and only 0.08 percent of households opted out before our study. The estimated energy savings are slightly — although not statistically significantly — larger in column 4, which suggests that survey respondents have somewhat larger energy savings, perhaps because they are more engaged with the HERs. Control group usage averages 2.35 ccf/day in winter 2015-2016, so the treatment effects in columns 1-3 amount to about 1.1 percent of counterfactual usage.[30]

Invoking the "no persistence" assumption, we use $\hat{\tau}_3$, the coefficients for winter 2015-2016, to construct our estimates of $\Delta\tilde{e}$ for the welfare analysis. There were 243 days between November 1, 2015 and June 30, 2016, so these estimates imply that HERs changed energy use by $\Delta\tilde{e} \approx -0.0271 \times 243 \approx -6.59$ ccf for $\mathcal{P}_n$, the target population of treatment group households that did not opt out before the program's second year, and by $\Delta\tilde{e} \approx -0.0315 \times 243 \approx -7.65$ ccf for $\mathcal{P}_s$, the target population of MPL respondents with valid WTP that did not opt out. To economize on notation, we denote the winter 2015-2016 treatment effect τ_3 simply as τ in the body of the paper.

Because Central Hudson uses decreasing block pricing, the effects of HERs on retail expenditures (or retailer net revenues) are not simply the energy use effects multiplied by a constant retail price (or markup). Instead, we must separately estimate regressions analogous to those in Appendix Table A7, except with different dependent variables: daily average retail gas expenditures or daily average contribution to retailer net revenues (i.e., the difference between retail expenditures and wholesale acquisition costs), respectively. Appendix Tables A8 and A9 present the full results.

For winter 2015-2016, we estimate that the average treatment effects on retail expenditures and retailer net revenues for target population $\mathcal{P}_n$ (in column 3) are \$-0.0202 per day and \$-0.0104 per day, respectively. This implies that HERs reduced retail expenditures by \$0.0202 $\times$ 243 $\approx$ \$4.91 and reduced Central Hudson's net revenues by \$0.0104 $\times$ 243 $\approx$ \$2.53 for the average household in the program's second year. For population $\mathcal{P}_s$, the estimated average retail expenditure reduction is \$5.61, and the estimated average net revenue reduction is \$2.84. These net revenue effects are our estimates of $\Delta\Pi$ for welfare analysis.

[28] Specifically, we weight observations by billing period duration times a household weight, where the household weight is the inverse predicted probability of being a complier for households in the population $\mathcal{P}_n$, $\left[\hat{\Pr}(R_i = 1|\mathbf{X}_i; \mathcal{P}_n)\right]^{-1}$. This uses probit estimates from column 9 of Appendix Table A10.

[29] Specifically, we weight households by the ratio of the predicted probability of responding to the survey with valid WTP to the predicted probability of being a complier, $\frac{\hat{\Pr}(H_i=1|\mathbf{X}_i;\mathcal{P}_n)}{\hat{\Pr}(R_i=1|\mathbf{X}_i;\mathcal{P}_n)}$, where H_i is an indicator for whether the household responded to the survey and has valid WTP. The numerator of this weight is predicted from estimates in column 7 of Appendix Table A10, while the denominator is from column 9.

[30] In percent terms, this is somewhat less than the typical effect of HERs on electricity use (Allcott 2015), but Opower's natural gas-focused programs typically have smaller percent effects than their electricity-focused programs.

Table A8: **Effects of Home Energy Reports on Retail Natural Gas Expenditures**

	(1)	(2)	(3)	(4)
Specification:	OLS	IV	IV	IV
Assigned to treatment × winter 2014-2015	-0.0197	-0.0197	-0.0199	-0.0206
	(0.0120)	(0.0120)	(0.0120)*	(0.0129)
Assigned to treatment× summer 2015	0.00432	0.00432	0.00436	0.00566
	(0.00686)	(0.00685)	(0.00687)	(0.00770)
Assigned to treatment × winter 2015-2016	-0.0197			
	(0.00982)**			
Assigned to treatment× summer 2016	0.00440			
	(0.0101)			
2nd-year recipient × winter 2015-2016		-0.0200	-0.0202	-0.0231
		(0.01000)**	(0.0100)**	(0.0102)**
2nd-year recipient × summer 2016		0.00449	0.00438	0.00309
		(0.0103)	(0.0103)	(0.0101)
Observations	200,540	200,540	200,540	200,540
R^2	0.871	0.871	0.871	0.878
Weights	Duration	Duration	Duration × IPW for $\mathcal{P}_n$	Duration × IPW for $\mathcal{P}_s$

Notes: This table presents estimates of Equation (11), using post-treatment data only. Dependent variable is retail natural gas expenditures in dollars per day. For context, control group sample mean expenditure is \$2.59/day. Columns 2-4 are IV regressions, where we instrument for 2nd-year recipient × winter 2015-2016 and 2nd-year recipient × summer 2016 with Assigned to treatment × winter 2015-2016 and Assigned to treatment × summer 2016. Columns 1 and 2 weight by billing period duration. Column 3 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_n$ of treatment group households that did not opt out before the second year. Column 4 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_s$ of treatment group households that did not opt out and returned a survey with valid willingness-to-pay. Robust standard errors, clustered by household, in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A9: **Effects of Home Energy Reports on Contribution to Retailer Net Revenue**

	(1)	(2)	(3)	(4)
Specification:	OLS	IV	IV	IV
Assigned to treatment × winter 2014-2015	-0.00820	-0.00820	-0.00825	-0.00806
	(0.00557)	(0.00557)	(0.00557)	(0.00580)
Assigned to treatment× summer 2015	0.00263	0.00263	0.00265	0.00324
	(0.00420)	(0.00420)	(0.00421)	(0.00467)
Assigned to treatment × winter 2015-2016	-0.0102			
	(0.00582)*			
Assigned to treatment× summer 2016	0.00257			
	(0.00635)			
2nd-year recipient × winter 2015-2016		-0.0104	-0.0104	-0.0117
		(0.00593)*	(0.00593)*	(0.00597)*
2nd-year recipient × summer 2016		0.00262	0.00254	0.00196
		(0.00648)	(0.00648)	(0.00645)
Observations	200,540	200,540	200,540	200,540
R^2	0.831	0.831	0.831	0.839
Weights	Duration	Duration	Duration × IPW for $\mathcal{P}_n$	Duration × IPW for $\mathcal{P}_s$

Notes: This table presents estimates of Equation (11), using post-treatment data only. Dependent variable is contribution to retailer net revenue (i.e., the difference between retail expenditures and wholesale acquisition costs for household i's gas bill on date t) in dollars per day. For context, control group sample mean contribution is \$1.59/day. Columns 2-4 are IV regressions, where we instrument for 2nd-year recipient × winter 2015-2016 and 2nd-year recipient × summer 2016 with Assigned to treatment × winter 2015-2016 and Assigned to treatment × summer 2016. Columns 1 and 2 weight by billing period duration. Column 3 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_n$ of treatment group households that did not opt out before the second year. Column 4 weights by duration times a household weight that matches the compliers to the target population $\mathcal{P}_s$ of treatment group households that did not opt out and returned a survey with valid willingness-to-pay. Robust standard errors, clustered by household, in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

D Appendix to Empirical Estimates

Table A10: **Inverse Probability Weights (Page 1)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent var:	Have WTP from Paper	Have WTP, Assigned to Base	Have WTP, Assigned to Follow-up	Have WTP from Base Mail	Have WTP from Follow-up Mail	Have WTP from Phone	Have WTP	Have WTP; Base Mail Excluded	Received Second Year
Baseline use	-0.358	-0.0985	-0.258	-0.308	-0.0556	0.765	0.541	0.795	0.0182
	(0.161)**	(0.0592)*	(0.149)*	(0.110)***	(0.117)	(0.225)***	(0.255)**	(0.242)***	(0.0706)
ln(Income)	-0.306	-0.171	-0.115	-0.0962	-0.210	-0.624	-0.398	-0.137	-0.0492
	(0.457)	(0.175)	(0.419)	(0.309)	(0.336)	(0.661)	(0.744)	(0.716)	(0.216)
ln(Net worth)	0.0841	0.0691	0.0191	0.212	-0.116	-0.0450	-0.0159	-0.285	-0.202
	(0.289)	(0.110)	(0.267)	(0.194)	(0.214)	(0.395)	(0.445)	(0.427)	(0.135)
ln(House value)	-0.180	0.0255	-0.211	-0.0824	-0.0950	-0.117	-0.140	-0.0774	0.0807
	(0.158)	(0.0595)	(0.145)	(0.106)	(0.116)	(0.230)	(0.255)	(0.246)	(0.0770)
Education	0.554	0.133	0.410	0.295	0.240	0.416	0.776	0.533	-0.00590
	(0.110)***	(0.0388)***	(0.102)***	(0.0745)***	(0.0802)***	(0.166)**	(0.185)***	(0.179)***	(0.0474)
Male	-0.240	0.0661	-0.317	-0.250	0.00731	0.290	-0.126	-0.0476	-0.0277
	(0.550)	(0.221)	(0.500)	(0.372)	(0.403)	(0.812)	(0.906)	(0.875)	(0.248)
ln(Age)	1.598	0.328	1.257	0.799	0.747	1.069	2.187	1.438	-0.737
	(1.012)	(0.389)	(0.928)	(0.690)	(0.739)	(1.452)	(1.621)	(1.558)	(0.458)
Retired	0.593	0.207	0.323	-0.00655	0.544	1.049	1.696	2.163	-1.108
	(1.309)	(0.487)	(1.199)	(0.852)	(0.969)	(2.103)	(2.339)	(2.297)	(0.520)**
Married	-0.123	0.160	-0.265	0.0575	-0.153	-0.975	-1.384	-1.443	0.933
	(0.681)	(0.247)	(0.631)	(0.444)	(0.518)	(0.989)	(1.102)	(1.061)	(0.305)***
Rent	0.210	0.138	0.0469	0.149	0.0390	-2.209	-2.097	-2.317	-0.409
	(0.799)	(0.311)	(0.735)	(0.541)	(0.595)	(1.112)**	(1.253)*	(1.202)*	(0.308)

(table continues on next page)

Notes: This table presents probit estimates used to construct inverse probability weights. We report marginal effects, with coefficients multiplied by 100 for readability. In all columns other than column 8, the sample is all households assigned to the initial HER recipient group that did not opt out before the 2nd year. In column 8, the sample is the same except excluding households that returned the first mail survey. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A11: **Inverse Probability Weights (Page 2)**

Dependent var:	(1) Have WTP from Paper	(2) Have WTP, Assigned to Base	(3) Have WTP, Assigned to Follow-up	(4) Have WTP from Base Mail	(5) Have WTP from Follow-up Mail	(6) Have WTP from Phone	(7) Have WTP	(8) Have WTP; Base Mail Excluded	(9) Received Second Year
Single family	0.880	0.294	0.584	0.310	0.566	-0.374	-0.304	-0.488	0.0787
	(0.708)	(0.273)	(0.649)	(0.462)	(0.539)	(1.011)	(1.129)	(1.086)	(0.290)
ln(House age)	-0.803	-0.0767	-0.717	-0.329	-0.460	-0.997	-1.513	-1.191	-0.0472
	(0.305)***	(0.124)	(0.277)***	(0.204)	(0.224)**	(0.453)**	(0.507)***	(0.492)**	(0.153)
Democrat	0.447	0.326	0.127	0.296	0.129	0.902	1.148	0.866	0.260
	(0.461)	(0.181)*	(0.419)	(0.308)	(0.339)	(0.720)	(0.798)	(0.778)	(0.204)
Hybrid auto share	0.116	0.0262	0.0857	0.0873	0.0143	0.445	0.479	0.404	-0.0331
	(0.0829)	(0.0302)	(0.0763)	(0.0505)*	(0.0656)	(0.123)***	(0.139)***	(0.135)***	(0.0342)
Green consumer	-0.470	-0.160	-0.302	-0.696	0.256	0.816	0.440	0.941	0.150
	(0.728)	(0.293)	(0.662)	(0.491)	(0.531)	(1.115)	(1.244)	(1.205)	(0.340)
Wildlife donor	3.425	1.056	2.216	2.549	0.537	3.212	5.825	3.342	0.543
	(1.151)***	(0.442)**	(1.050)**	(0.735)***	(0.870)	(1.823)*	(2.040)***	(2.019)*	(0.548)
Profit score	1.952	0.0559	1.854	0.696	1.209	1.171	2.433	1.896	-0.411
	(0.392)***	(0.138)	(0.363)***	(0.261)***	(0.291)***	(0.574)**	(0.642)***	(0.618)***	(0.176)**
Buyer score	0.858	0.185	0.655	0.586	0.267	-0.164	0.513	-0.0173	0.140
	(0.374)**	(0.141)	(0.345)*	(0.244)**	(0.282)	(0.539)	(0.607)	(0.584)	(0.160)
Mail responder	0.333	0.0517	0.278	-0.0196	0.363	-0.739	-0.359	-0.442	-0.164
	(0.638)	(0.244)	(0.585)	(0.422)	(0.473)	(0.957)	(1.065)	(1.030)	(0.289)
Home improvement	0.322	-0.322	0.620	-0.397	0.725	1.630	1.224	1.564	0.0255
	(0.892)	(0.382)	(0.800)	(0.598)	(0.648)	(1.332)	(1.500)	(1.451)	(0.401)
Observations	9948	9948	9948	9948	9948	9948	9948	9548	9948

Notes: This table presents probit estimates used to construct inverse probability weights. We report marginal effects, with coefficients multiplied by 100 for readability. In all columns other than column 8, the sample is all households assigned to the initial HER recipient group that did not opt out before the 2nd year. In column 8, the sample is the same except excluding households that returned the first mail survey. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A12: **Correlation of Willingness-to-Pay with Phone Survey Responsiveness**

	(1)	(2)
Completed survey attempt number	0.0229	0.0651
	(0.0885)	(0.0901)
Observations	1609	1609
Weights	Equal	IPW for $\mathcal{P}_n$

Notes: Dependent variable is willingness-to-pay, sample is all phone survey respondents. For the phone survey, each respondent was dialed up to eight times; the independent variable is the attempt number on which the survey was completed. Column 2 re-weights observations to match $\mathcal{P}_n$, the target population of treatment group households that did not opt out before the program's second year. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A13: **Fitting Moral Utility**

	(1)
Expected savings	0.0870
	(0.00930)***
Inspired	2.800
	(0.452)***
Pressured	-1.224
	(0.594)**
Proud	0.119
	(0.473)
Guilty	0.662
	(0.563)
Observations	1350
R^2	0.122

Notes: Dependent variable is willingness-to-pay. Expected savings is winsorized at $50. Sample includes only phone survey respondents with non-missing data. Observations are weighted to match the target population of treatment group households that did not opt out before the program's second year. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A14: **Effect of Survey Version on Willingness-to-Pay**

	(1)	(2)	(3)	(4)
Comparison version	-0.686	-0.693	-0.661	-0.658
	(0.386)*	(0.384)*	(0.390)*	(0.387)*
Environmental version	-0.211	-0.155	-0.190	-0.143
	(0.386)	(0.388)	(0.394)	(0.399)
Mean comparison			0.106	0.0391
			(0.215)	(0.286)
Comparison version×Mean comparison			0.111	0.184
			(0.317)	(0.319)
Environmental version×Mean comparison			0.0530	0.0403
			(0.347)	(0.354)
Observations	2137	2137	2137	2137
Include **X** covariates	No	Yes	No	Yes

Notes: Dependent variable is willingness-to-pay. "Mean comparison" is the average difference (in 1000s of ccf) between own natural gas usage and mean neighbor usage on the HERs in winter 2014-2015. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

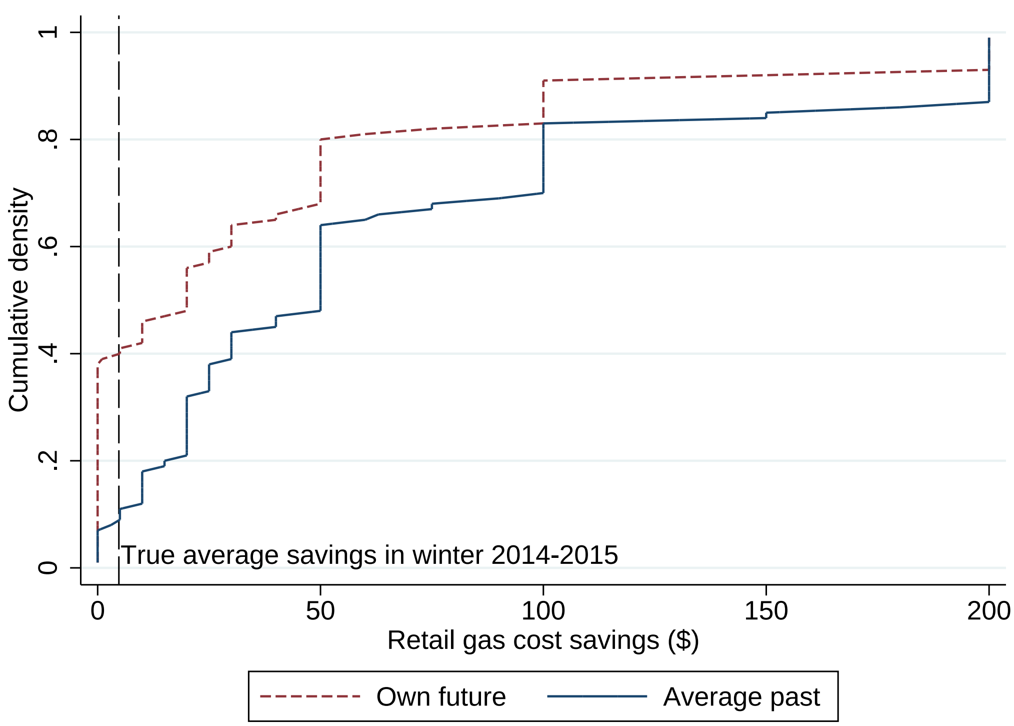
E Appendix to Welfare Estimates

E.A Testing for Biased Beliefs and Optimism Bias

For the welfare analysis, we assume that WTP equals consumer utility gain. In this context, we could imagine two reasons why this might fail: biased beliefs and optimism bias.

By biased beliefs, we mean that consumers might systematically underestimate or overestimate the energy cost savings resulting from their conservation efforts. Consumers likely know the monetary and non-monetary costs of their efforts, such as the time to adjust the thermostat or the money to install energy-saving windows, but resulting energy savings can be quite difficult to infer given that gas bills fluctuate substantially across months and years. There is empirical evidence to support this concern: Pronin, Berger, and Molouki (2007) and Nolan *et al.* (2008) find that people underestimate the motivational power of social norm messaging, and Larrick and Soll (2008), Attari *et al.* (2010), and Allcott (2013) explore various belief biases related to energy costs.

Figure A2: **Beliefs About Savings Caused by Home Energy Reports**



Notes: This figure presents the unweighted distribution of responses to the following phone survey questions: "How much money do you think you would save on your natural gas bills if you receive four more Reports?" and "For the average household, how much money do you think these Reports have helped them save on their natural gas bills?" True average savings in winter 2014-2015 was $4.77 per household.

To test this, the phone survey asked respondents how much money they thought they would save on their natural gas bills if they received four more HERs, as well as how much money they thought the average HER recipient had saved since last fall. Figure A2 shows that both the median and mean respondents overstate gas cost savings relative to the true average treatment effect. This suggests that if anything, biased beliefs could bias WTP upward instead of downward. However, we treat this result very cautiously, given that these questions were not incentive-compatible and stated beliefs are highly dispersed.

A second and more controversial reason why WTP might not equal consumer welfare gain has to do with optimism bias. Oster, Shoulson, and Dorsey (2013) show that people at high risk of Huntington disease do not get tested despite the fact that knowledge of disease status leads to very different life choices. They propose a model based on Brunnermeier and Parker (2005) in which people optimally choose beliefs while trading off the utility gain from optimistic beliefs with the utility loss from suboptimal actions. Bracha and Brown (2012) develop an alternative model in which optimism bias is constrained by the cost of holding incorrect beliefs. Evaluating information provision in these models requires the analyst to take a stand on whether to recognize optimistically biased beliefs as true utility. In these models, optimistically biased consumers may not experience a utility gain from exogenously-provided information, even though it would lead to more accurate beliefs and (in Brunnermeier and Parker's model) improved decision making. If current Home Energy Report recipients derive utility from incorrectly believing that they use less energy than their neighbors and want to remain incorrectly optimistic about their relative energy use in the future, this might reduce their WTP for HERs, and perhaps the utility loss from correcting optimism bias should not be counted as a "true" utility loss.

Even without taking a stand on this issue, we can provide suggestive tests of whether optimism bias affects WTP. On both the mail and phone surveys, we asked people whether their first HER told them they were using more or less energy than they thought. We hypothesize that people who want to be optimistic in the future are more likely to have been optimistic in the past. The initial belief update should thus be negatively correlated with WTP if optimism affects WTP. People gave meaningful responses: the belief update variable is positively correlated with baseline usage, usage relative to neighbors on the first HER, and reporting that they would like the HERs more if they did not have social comparisons. More people report underestimating their energy use than report overestimating. However, Appendix Table A15 shows that the belief update is not associated with WTP, either unconditionally or conditional on X.

Table A15: **Correlation of Willingness-to-Pay with Pre-Treatment Optimism**

	(1)	(2)
Belief update	0.0540	0.0532
	(0.134)	(0.138)
Observations	2102	2102
Include X covariates	No	Yes

Notes: This table presents regressions of WTP on the belief update using unweighted responses from both mail and phone surveys. Belief update is from question 8 on the mail survey and question 2 on the phone survey: "Think back to when you received your first Home Energy Report. Did the Report say that you were using more or less energy than you thought?" Responses are on a five-point Likert-style scale from "much less than I thought" to "much more than I thought," and we code these as integers from -2 to +2. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

E.B Program Implementation Cost

Home Energy Report programs have setup costs, per-household marginal costs, and annual fixed costs. In evaluating a program's second year, we ignore setup costs. Panel A of Table A16 presents the per-household annual marginal costs. Based on a high-volume price quote from PFL (www.PrintingForLess.com), we assume $0.4926 per HER for printing and mailing. This uses the appropriate printing and paper quality, production speed, and shipping method for HERs. HER recipients occasionally call the utility to ask questions, complain, or opt out of HERs. Opower data show that HER recipients typically call with 0.5 percent probability per year and that these calls cost the utility $5 per call to answer. We estimate $0.01 per household for server space to store data, and $0.05 to purchase household-level demographic data to enhance the HERs. Overall, we estimate that the per-household marginal cost for one year of a program involving four HERs is $2.06.

Panel B of Table A16 presents the per-utility annual costs that are fixed with respect to the number of households. Opower reported an estimated 51 hours of program design and reporting time for a client like Central Hudson. In addition, Central Hudson and Opower have in-person meetings approximately every quarter, and short phone meetings most weeks. We assume that Opower staff cost $85 per hour, on the basis of a $118,097 nationwide median annual salary for "program managers" (see http://www1.salary.com/Program-Manager-Salary.html) multiplied by a 1.5 loading factor to account for health insurance, vacation, and other benefits and divided by 2080 hours per year. Central Hudson reported to us that their fully-loaded staff time for this project costs $62.64 per hour. Total utility-level fixed costs are $16,339.

Central Hudson has four HER programs — the natural gas program we study, plus three others — with a total of about 100,000 households in treatment. Some of the per-utility fixed costs such as program design and reporting likely would increase with the number of programs, whereas others such as travel time for quarterly meetings likely would not. If the fixed cost is allocated equally to each of Central Hudson's 100,000 recipient households, this gives $0.16 per household. Given that the second year of the program we study includes 9948 households that were allocated to the treatment group and did not opt out, this would sum to $1625. Alternatively, if the fixed cost is allocated equally to each program, this is $4,085 per program. Allocating this $4,085 equally across the 9948 recipient households gives $0.41 per household.

Table A16: **Implementation Cost Estimates**

Item	Explanation	Cost ($)
Panel A: Per-Household Annual Marginal Costs		
Printing and mailing	$0.4926/HER × 4 HERS	1.97
Utility call center	0.5% call probability × $5/call	0.025
Server space	$0.01 per recipient household	0.01
Demographic data	$0.05 per recipient household	0.05
Total		*2.06*
Panel B: Per-Utility Annual Fixed Costs		
	Opower ($85/hour)	
Program design and reporting	51 hours	4,335
Quarterly meetings (time)	8 hours/quarter× 2 people	5,440
Quarterly meetings (travel)	$250/quarter × 2 people	2,000
Weekly phone meetings	20 minutes/week × 1 person	1,473
	Central Hudson ($62.64/hour)	
Quarterly meetings	2 hours/quarter× 4 people	2,004
Weekly phone meetings	20 minutes/week × 1 person	1,086
Total		*16,339*
Annual fixed cost per household	Central Hudson has ∼100,000 HER recipients	0.16
Annual fixed cost per program	Central Hudson has four HER programs	4,085

Notes: This table presents the implementation costs for an ongoing Opower Home Energy Report program. See text for details.

E.C Speculative Evaluation of a Typical Full Home Energy Report Program

In this appendix, we address two shortcomings of the welfare evaluation in Table 7. First, Table 7 evaluates only the second year of a Home Energy Report program. Second, it evaluates one specific HER program, which may or may not be typical.

Table A17 evaluates the full course of a typical Home Energy Report program. We use the energy savings from "site 2" studied by Allcott and Rogers (2014), an electricity-focused program with savings approximately equal to the average savings of other Opower programs. Using Table 8 from Allcott and Rogers (2014), four years of Home Energy Reports are projected to save 1875 kilowatt-hours (kWh) in total per household, including significant savings after the program ends. At the 2014 national average electricity price of $0.125/kWh, this amounts to $234 dollars, as shown in Panel A.[31] We assume that the long-run marginal source of electricity is a combined cycle gas plant, with cost and heat rate characteristics from the U.S. Energy Information Administration's Annual

[31]See http://www.eia.gov/electricity/monthly/pdf/epm.pdf.

Energy Outlook.[32] This gives energy acquisition cost savings of \$176 and externality reduction of \$53, using the externality damage assumptions detailed in the body of this paper. Subtracting acquisition cost savings from retail electricity cost savings gives a retailer net revenue loss of \$58.

For implementation cost, we use the price that Opower charges utilities, which is about \$8 per household per year for six HERs. We assume that this covers costs to set up and operate the HER program as well as relevant overhead costs for sales, marketing, and research and development.[33]

Panel B shows the consumer welfare and social welfare effect of the program under two assumptions. In column 1, we ignore non-energy costs, assuming that $\Delta V = -\Delta\tilde{e} \cdot p_e$. In column 2, we adjust for non-energy costs using our estimate that $\Delta V \approx -\Delta\tilde{e} \cdot p_e \times 0.57$. Failing to adjust for non-energy costs overstates social welfare gains by a factor of 2.0. We label this calculation as "speculative" because it hinges on the assumption that $\Delta V \approx -\Delta\tilde{e} \cdot p_e \times 0.57$.

These numbers grow very large when aggregated across the many households worldwide that have received home energy reports. As of January 2017, opower.com reports that Opower participant households have saved 11.6 billion kWh of energy. The total effects are thus $11.6 \times 10^9/1875 \approx 6$ million times those reported in the table. Thus, if one ignores non-energy costs, one would calculate that the total social welfare gains from HERs was \$1.22 billion, whereas adjusting for our estimate of non-energy costs gives an estimate of \$600 million. Thus, this rough calculation suggests that ignoring non-energy costs causes the total social welfare gains for home energy reports to be overstated by \$620 million.

Table A17: **Social Welfare Effects of a Full Home Energy Report Program**

	(1)	(2)
Panel A: Benefits and Costs Other than Consumer Welfare (\$/recipient)		
(+) Externality reduction	53	
(-) Retailer net revenue loss	58	
Retail electricity cost savings	234	
Electricity acquisition cost savings	176	
(-) Implementation cost	32	
(=) ΔWelfare, excluding consumer welfare	-37	
Panel B: Consumer Welfare and Social Welfare Effect (\$/recipient)		
Assumption:	$\Delta V = -\Delta\tilde{e} \cdot p_e$	$\Delta V \approx -\Delta\tilde{e} \cdot p_e \times 0.57$
Consumer welfare gain (ΔV)	234	134
ΔWelfare	197	97

Notes: See text for details.

[32] http://www.eia.gov/forecasts/aeo/electricity_generation.cfm

[33] Opower (2014) reports that the company has a 65 percent gross margin in 2013, which would suggest that the price overstates program implementation cost. On the other hand, the company operated at a net loss through that year, suggesting that the gross margin is actually not sufficient to cover sales, marketing, R&D, and other relevant overhead.

F Machine Learning Intermediate Results

Table A18: **Machine Learning: Tuning Parameters and Performance Statistics**

Panel A: Tuning Parameters

	(1)	(2)	(3)	(4)	(5)	(6)
Training Set Number	Ridge Regression λ	Elastic Net λ	Elastic Net Fraction of Full Solution	Random Forest Number of Candidate Variables at Each Split	Imai and Ratkovic (2013) λ_1	Imai and Ratkovic (2013) λ_2
1	2726	500	0.1	3	-5.19	-5.06
2	2730	500	0.15	3	-5.14	-4.82
3	1581	1	0.15	1	-5.23	-5.14
4	1480	200	0.15	1	-5.02	-5.25
5	2698	10	0.2	1	-5.22	-5.55

Panel B: Performance Statistics When Predicting WTP

	(1)	(2)
Method	Average WTP Conditional on Above-Median Predicted WTP	Root Mean-Squared Error
Elastic net	3.32	7.28
Ridge	3.29	7.28
Random forest	3.19	7.29

Panel C: Performance Statistics When Predicting Energy Savings

	(1)
Method	Average Winter 2015-2016 Treatment Effect Conditional on Above-Median Predicted Effect
Gradient forest	0.0528
Imai and Ratkovic	0.0039

Notes: See Section VI for details. In step 1 described in the text, we partitioned the sample into five test sets and predicted WTP and energy savings in each test set using training data from the other four. Panel A presents the optimal tuning parameters in each of the five training sets. For ridge regression, we used the R package ridge.cv. For elastic nets, we used the train script from the caret package. For random forests, we used tuneRF. For Imai and Ratkovic (2013), we used FindIt. For Athey, Imbens, and Tibshirani (2016), we used gradient.forest. Panel B presents performance statistics for the three algorithms used to predict WTP. We used the elastic net predictions for step 2 of the targeting procedure because it performed best on the metric in column 1. Panel C presents performance statistics for the two algorithms used to predict winter 2015-2016 energy savings. We used the gradient forest predictions for step 2 of the targeting procedure because it performed best on this metric.

References

Brandon, Alec, Paul J. Ferraro, John A. List, Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer. 2017. "Do The Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments." NBER Working Paper No. 23277 (March).