Direct and cross-resource spillover effects of social information about water usage

Pre-Analysis Plan

Jacopo Bonan[1,2], Giovanna d'Adda[3,2], Arianna Galliera[4], and Massimo Tavoni[1,2]

[1]Politecnico di Milano
[2]RFF-CMCC European Institute on Economics and the Environment (EIEE)
[3]University of Milan
[4]Università Cattolica del Sacro Cuore

Fieldwork locations:	Italy
Fieldwork dates:	October 2019 to January 2021
Date of Pre-Analysis Plan:	September 29, 2020

1 Introduction

This document outlines our pre-analysis plan for a field experiment on the impact of social information on water consumption, conducted with customers of a large Italian electric utility. The document summarizes (i) our experiment and resulting data, (ii) our research questions and the plan of regressions, and (iii) power calculations.

At the time of writing this plan, we designed and launched the RCT. We accessed pre-experimental administrative data which we used to identify the study sample and randomize treatment assignment.
We intend to submit this Pre-Analysis Plan to the AEA RCT Registry.

2 Literature review

Social information programs are widely used by policy makers to nudge behavioural change. Their popularity is attributable, at least in part, to existing evidence on their ability to influence behaviour in a variety of settings, from energy and water consumption (Allcott et al., 2011; Allcott and Rogers, 2014; Ayres et al., 2013; Brent et al., 2015a; Ferraro and Price, 2013; Ferraro et al., 2011; Ferraro and Miranda, 2013a; Jaime Torres and Carlsson, 2018), contributions to charitable causes (Frey and Meier, 2004; Shang and Croson, 2009a), voting (Gerber and Rogers, 2009) and financial decisions (Beshears et al., 2015).[1]

Our first research question looks at the direct impact of a customized household report about water usage on water consumption. Different works have tackled this question in both developed and developing contexts. In developed contexts, the majority of works are carried out in the US. It is documented that the social norm content of the home water reports have short-term water conservation effects between 2.7 and 5%, which tend to persist significantly over longer time horizons (50% decrease in impact after one year, but still significant and which remain persistent four years later). It is also shown that these effects are explained not only by short-lived behavioral adjustments, but also by longer-lived adjustments to habits or physical capital (Ferraro et al., 2011; Ferraro and Price, 2013; Ferraro and Miranda, 2013b; Bernedo et al., 2014 in Georgia and Brent et al., 2015b in Texas). More recently, Jessoe et al. (2019) use high frequency water use to assess a home water program in California in a drought period. They find 4 to 5% reduction in water usage. They also find no differential effect of different components and contents of the report and that the effects are relatively short-lived, documenting a reversion 4 to 5 months post-treatment. In Singapore, Goette et al. (2019) find average treatment effect of a feedback on water consumption in the order of 4 Litres of water per person per day. In developing contexts, Miranda et al. (2020) find 3 to 5% effects in Costa Rica, while Jaime Torres and Carlsson (2018) find 6.8% water reduction on customers targeted by home water report and 5.5% decrease on untargeted customers living close-by (cross-individual spillover) in Colombia. We contribute to this research line by extending the geographical scope of water conservation programs to a European country, Italy[2] Hence, we complement the existent works by testing the external validity and transferability of these programs to other contexts. We also contribute to the growing literature on the context-dependency of causal effects measured in a particular policy population (e.g. Allcott, 2015; Vivalt, 2015).

Our second research question relates to the spillover effects of the water report on the individual consumption of other resources, such as gas and electricity. Environmental decisions may be subject to behavioural spillover effects whereby the adoption of one pro-environmental action reduces (negative spillover) or

[1] See Gillingham et al. (2018); Abrahamse (2019); Gerarden et al. (2017) for a broader discussion of the energy efficiency gap and the assessment of energy efficiency policies.

[2] To the best of our knowledge no evaluation of water conservation programs using social comparison has been conducted in Europe, while in the electricity sector work has been done by Andor et al. (2017) and Bonan et al. (2019a).

increases (positive spillover) the probability of another pro-environmental action. Negative spillover is typically attributed to moral licensing. Positive spillover realizes when people seek consistency across behaviors or environmental concerns are primed. The evidence on behavioural spillover effects is mixed. Such variability in the results can be partially explained by two facts. First, there is a high variability in the methods used to quantify impacts (Galizzi and Whitmarsh, 2019). Second, there is a high variability in the way behavioural outcomes are measured (behavioural intentions, policy support, self-reported behaviours, actual behaviours) (Maki et al., 2019)[3]. Investigating spillover is relevant, as it allows to shed light on hidden effects which may further improve (worsen) the cost-effectiveness of social information nudges aimed at stimulating pro-environmental behaviours. However, relatively few papers rigorously address this topic. Jessoe et al. (2017) examine "cross-sectoral" spillover using data on real world actual behaviours, namely water and electricity usage. They find that home water reports induce a 1.3 to 2.2% reduction in summertime electricity use. They also provide evidence suggesting that impacts are not solely imputable to mechanic resource complementarities. Carlsson et al. (2020) find that a social information campaign on water use decreased water use for two groups of households: those with inefficient use of water and those with efficient use of water before the information campaign. However, it is only for the households with efficient use of water that a positive spillover effect on electricity is observed. The effect is sizeable; this group has around 9 percent lower use of electricity compared with the control group 11 months into the information campaign. Brent et al. (2015b) examine the interaction between the social comparison treatment and existing utility conservation programs such as free home water audits and rebates for efficient toilets or irrigation controllers. They find that receiving the home water report increases program participation by six percentage points. Far smaller effects, in the order of 0.4 percentage points increase, are found by Allcott and Rogers (2014). In a lab experiment, d'Adda et al. (2017) find little evidence of behavioural spillover from fairness to cooperation, while in another setting with privately provided public goods, Shang and Croson (2009b) find that providing information on the charitable donations of others increases donations in both the current campaign as well as future campaigns. We contribute to this strand of literature by focusing on the impacts of the home water report on several sectors besides water usage, namely electricity and gas. We also assess the program impact on customers' engagement with the utility (relevant in a business management/marketing perspective).

The third research question, perhaps the most original contribution, is to assess the effect of receiving different nudges. In our case, people receive different social information on different resources. The net effect is an empirical question. One may expect crowd-out if an information overload occurs and additional nudges tire the individuals who stop paying attention Sunstein (2016). Conversely, crowd-in occurs if more information leads to higher awareness for example about behavioural synergies. As

[3] General reviews of behavioural spillovers can be found in Truelove et al. (2014), Dolan and Galizzi (2015), Nilsson et al. (2017).

nudges become more popular, individuals are likely to be exposed to multiple behavioural policies, hence the investigation of crowding-in or crowding-out effects is of extreme importance. Only few works have tackled this issue. Brandon et al. (2019) considers two social nudges, the first targeting conservation during peak load events, the second promoting aggregate conservation. They find no evidence of crowding-out effects. Bonan et al. (2019a) combine different nudges, by augmenting the standard social information message on electricity usage with an environmental self-identity prime. They find that social information is more effective when environmental self-identity is made more salient, but only among individuals who acted pro-environmentally in the past. Ultimately, this is relevant for the assessment of the impact of nudges on welfare (Allcott and Kessler, 2015). We contribute to this nascent literature by providing evidence on the marginal impact of receiving multiple reports. Differently from previous works, we look at the impact of reports providing social comparisons with a similar structure, but on different resources.
Finally, we explore the heterogeneity of the effects along resource pre-treatment usage and engagement with the utility. Important sources of treatment heterogeneities found in the literature are pre-treatment usage (Byrne et al., 2018; List et al., 2017; Allcott, 2011; Ferraro and Price, 2013; Bhanot, 2017; Ferraro and Miranda, 2013a)[4], political and environmental preferences (Costa and Kahn, 2011; Bonan et al., 2019a), beliefs (Byrne et al., 2018; Jachimowicz et al., 2018), strength and consistency of norms (Bonan et al., 2019b).

3 Treatment, sample and randomization

The project is centered around a customer water efficiency program conducted with utility's customers. The study is realized on a sample of customers from an Italian multi-utility who have a contract for water provision and a valid e-mail address. The multi-utility also provides, besides water, electricity and gas. Since 2018, the utility engaged in customers' information campaigns through the design and launch of "Opower-style" home energy reports for both gas and electricity consumption. A new water consumption report kicked-off in October 2019.
The report includes the following elements:

- *Static neighbor comparison*: comparison of one's own average water consumption in the reporting period (about two months) with that of two groups: "virtuous" and "average" customers. For the former the utility computes the average consumption of the top 75% most water efficient customers with the same household size in the customer base. For the latter, the average water consumption of customers with the same household size is used.

[4] In particular, several works find that social comparisons are more effective for high water users (Ferraro and Miranda, 2013b; Ferraro and Price, 2013; Brent et al., 2015b).

- *Feedback on water consumption*: based on relative performance customer receives emoticon and % difference with respect to the average customers

- Dynamic neighbor comparison: comparison of own consumption over the months in the reporting period and same months of previous year; comparison of the rolling yearly water consumption with that in the previous year

- Water saving tips: tips on how to save water, divided into categories (behavior change, small investment, large investment). Tips are seasonal-specific.

The report is sent to customers by email bimonthly soon after water bill delivery, following the rolling billing cycle. The reference period of the report is the same of that reported in the bill. Figure 1 depicts the structure and the contents of the report.

To be eligible for the study sample, customers need to have one single water contract in their main address of residence, meaning that customers with multiple contracts and/or multiple houses are excluded. Then, we exclude customers with the number of people living in the house above 20, possibly indicating condominiums, and with excessively high water usage in the past year. In the latter, we exclude customers with consumption above ten times the sample median consumption. The experimental sample is split in two waves. The first wave starts receiving the report in October 2019, while the second in January 2020. The first and second wave includes 70,210 and 38,830 customers, respectively. The total study sample is 109,040 customers.

The experimental design relies on the random assignment of eligible customers to a treatment group which receives the report and a control group which does not, for each wave. We follow a stratified individual level randomization procedure, to maximize ex-ante balance across treatment and control group along a battery of important observable characteristics (Bruhn and McKenzie, 2009). Strata are obtained from the combination of the following variables:

- Having an electricity and/or gas contract

- Receiving the report about electricity and/or gas consumption

- Having an electricity and/or gas contract in the free vs "protected" market[5]

- Having performed a water self-reading in the last twelve months

[5] Once a monopoly, the Italian energy retail market, including both electricity and gas, was liberalized in 1999. The liberalization process has been slow, with over 60 per cent of domestic customers still buying their energy at the conditions set by the public authority for energy as of 2017 (ARERA, 2018). The market is expected to be completely liberalized in 2020, when everyone has to purchase electricity from the free market. Conversely, the market for water supply has been liberalized, although it remains highly regulated. Only customers with free market contracts can be contacted to receive the reports.

We exclude strata with less than ten observations. We end up with 32 strata. Within each stratum, we sort customers by water consumption in 2018 and assign adjacent customers to treatment and control group. In particular, in the first wave we assigned to the different groups every other customer (50% treatment and 50% control). In the second wave, every eleventh customers was assigned to the control group (91% treatment and 9% control). Eventually, we end up with 70,418 treated and 38,622 control customers, corresponding to 65 and 35% of the sample, respectively.

4 Research questions and analysis

4.1 Experimental analysis

The study addresses the following research questions. For each of them the specification, the test of hypothesis and the sample of analysis are indicated.

Research Question 1 *What is the impact of receiving the water report on water consumption?*

We estimate the intention to treat effect (ITT) of receiving the report on water consumption:

$$y_{it} = \beta_1 Post_{it} + \beta_2 Prog_i * Post_{it} + h_t + g_i + \varepsilon_{it} \tag{1}$$

where y_{it} is the normalized average daily water consumption over the billing period in month t. $Prog$ is a treatment indicator, $Post$ is a dummy variable which becomes one when customers receive the first report. The regression also includes month-by-year fixed effects, h_t, and household fixed effects g_i. Standard errors are clustered at the level of household, to allow for the presence of within customer correlation over time in the error term (Bertrand et al., 2004). The analysis is conducted on the whole study sample over a period ranging from September 2018 to January 2021.

Research Question 2 *What is the impact of receiving the water report on electricity and gas consumption?*

In order to assess the broader impact of social information intervention on resource consumption, we run the following pooled model:

$$y_{irt} = \beta_1 Post_{it} + \beta_2 Prog_i * Post_{it} + h_t + g_i + f_r + \varepsilon_{irt} \tag{2}$$

where y_{irt} is customer's i normalized consumption of resource r (electricity or gas) in month t, while f_r are resource fixed effects. The sample of analysis includes customers with at least one contract (electricity, gas or both), besides the water's one.

Then, we estimate the resource-specific impacts, by estimating models similar to 1 on the different subsample. The dependent variable y_{it} is the average normalized daily electricity (gas) consumption in month t. For the estimation of the spillover effect on electricity, the sample of analysis includes customers with an electricity contract, besides the water one. In the case of gas, the sample of analysis includes customers with a gas contract, besides the water one.

Research Question 3 *What is the impact of receiving multiple nudges on resource consumption?*

We tackle this question by estimating the heterogeneous effects of receiving the water report across customers already receiving other reports (electricity, gas or both) at the baseline. We separately estimate the direct impact on water, gas and electricity usage, y_{it}, in the following model:

$$\begin{aligned} y_{it} = \beta_1 Post_{it} + \beta_2 Prog_i * Post_{it} + \beta_3 Post_{it} * OtherRep_i + \\ + \ \beta_4 Prog_i * Post_{it} * OtherRep_i + h_t + g_i + \varepsilon_{it} \end{aligned} \tag{3}$$

$OtherRep_i$ is a dummy on whether customer i receives any other report at the baseline.

Research Question 4 *How do treatment effects vary with respect to baseline household characteristics?*

We explore the heterogeneities of the effects estimated in the previous research questions along the following variables:

- pre-treatment usage of the resource considered as outcome
- consumption awareness/ engagement with the utility (self-reading, activities on the web portal, activities on the app, contacts with customer service). Variables are aggregated in an index.

Research Question 5 *What is the impact of receiving multiple nudges on customers' engagement with the utility?*

To tackle the question we estimate models similar to 1 and 3 with different dependent variables. In particular, we focus on the following indicators of customers' engagement with the utility.

- Monthly frequency of water self-readings
- Digital activities: monthly access to online portal, access to app
- Churn in electricity and gas contracts
- Customer satisfaction: the number of contacts and complaints with the utility through the different communication channels

5 Data

All data used in the analysis are provided by our partner utility, after being anonymized. Data on water and gas consumption are based on meter reading occurring periodically by the distributor for all customers and on self-readings provided by customers on a voluntary basis. Data on electricity usage rely on actual consumption measured through smart meters on a monthly basis. Usage data are expressed as daily level in each month normalized with respect to the control group consumption in the intervention period.
Data on customer engagement with the utility are measured monthly as follows:

- The number of water self-readings
- The number of days with at least one portal accesses
- The number of days with at least one access to the app
- A dummy for churn in period t, i.e. the individual ceases to be a customer
- The number of customers' care contacts through the different communication channels
- The number of complaints through different communication channels

The following variables are used for the construction of strata which are used for randomization.

- A dummy for whether the customer has an electricity contract
- A dummy for whether the customer has a gas contract
- A dummy for whether the customer receives the report about electricity consumption
- A dummy for whether the customer receives the report about gas consumption
- A binary variable for having an electricity contract in the free vs "protected" market
- A binary variable for having a gas contract in the free vs "protected" market
- A dummy for whether the customer performed a water self-reading in the last twelve months
- Average water consumption in 2018.

Treatment heterogeneity is run using the average value at baseline, i.e. in 2018, or related variables capturing properties of the distribution (e.g. median and quartiles) for the variables described above (pre-treatment resource usage and engagement).

5.1 Multiple Hypothesis testing

We control for multiple hypothesis testing in several ways. First, when looking at the main treatment effect across outcomes, i.e. water, gas, electricity usage and engagement (summarized in an index), we compute the sharpened two-stage q-values (FDR-adjusted q-values)**?**.
Second, within the same outcomes families, when we test treatment heterogeneities (i.e. by presence of other reports, pre-treatment usage and engagement index), we compute the sharpened two-stage q-values (FDR-adjusted q-values)**?**.

5.2 Sample Balance at Baseline

For each variable available at the time of treatment assignment, we conduct balance tests across treatment groups. We denote these variables as y_{i0} and for each of them we estimate the following equation:

$$y_{i0} = \beta_0 + \beta_1 Program_i + \beta_2 Wave_i + \varepsilon_{i0} \tag{4}$$

where β_1 provides the difference in variable y between customers assigned to receive the water report and the ones assigned to the control group. $Wave_i$ is a dummy variable which takes the value of one if customer is part of the first (main) wave of reports' delivery and zero for the second one. This represents an important control, as the two samples are systematically different and have different proportions of treatment and control customers.
Table 1 reports, for the variables employed for the construction of strata and the dimensions of heterogeneity available at the time of this writing, the mean and standard error in the control group as well as the coefficient β_1 which indicates the average difference between treatment and control group. As expected, we do not detect any significant difference in observable characteristics across the two samples.

5.3 Power calculations

We compute the minimum detectable effects allowed by the design through simple power calculation. In particular, we consider a simple non-repeated treatment/control design, with unequal sample size. In Table 2 we compute power for all the hypothesis under consideration for each research question. In particular, we use the baseline values of the main outcomes proposed in the analysis, namely for water, gas and electricity usage[6]. We obtain minimum detectable effects ranging from 1.2 to 1.4% for the main effects (RQ1 and RQ2). As for the heterogeneous effects (RQ3 and RQ4), MDEs are between 1.4 and 2.2%. These are commonly considered as very small effects and are aligned with the magnitude of effects of

[6] Data on engagement are not available at the current stage of the project.

social information programs. Moreover, these estimates are conservatives as they do not incorporate the fact that we will use repeated measures of the outcomes and individual fixed effects in our estimation strategy. These are likely to decrease the unexplained variation in the outcome and further decrease the minimum detectable effects.

Tables and figures

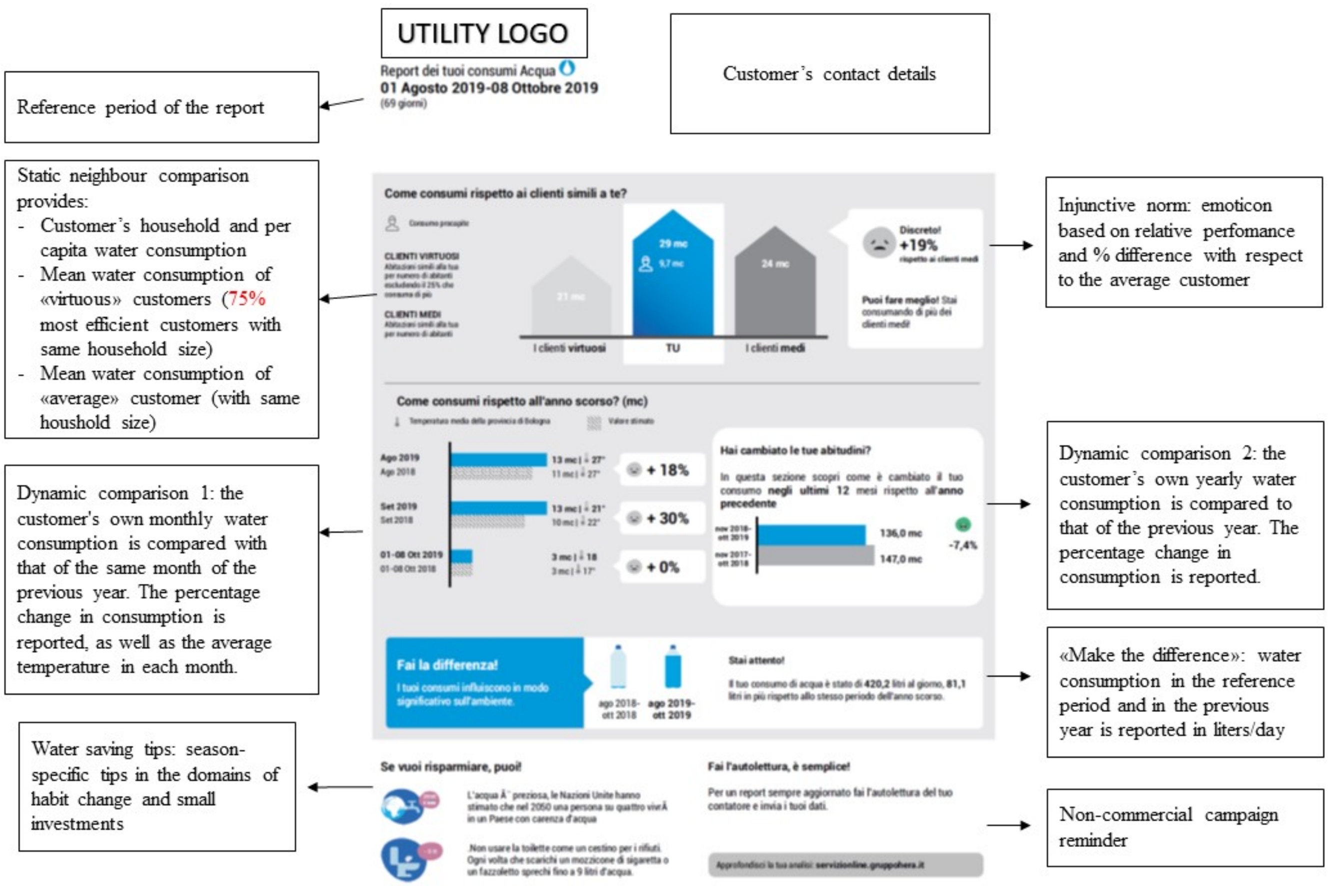

Figure 1: Water consumption report

Table 1: Summary statistics and balance

	N	Control Mean	Control SE	ITT	p-value
Gas contract	109040	0.775	0.002	0.000	0.950
Electricity contract	109040	0.655	0.002	0.000	0.977
Gas contract in the free market	109040	0.694	0.002	0.000	0.946
Electricity contract in the free market	109040	0.614	0.002	0.000	0.945
Gas report	109040	0.236	0.002	0.000	0.979
Electricity report	109040	0.198	0.002	0.000	0.966
Self-reading in the past 12 months	109040	0.367	0.002	0.000	0.942
Gas usage in 2018	89527	1041.6	3.4	4.987	0.341
Electricity usage in 2018	76382	2312.6	6.2	7.873	0.405
Water usage in 2018	109040	115.2	0.3	-0.143	0.764
N. of occupants in the house	109040	2.417	0.005	-0.009	0.290

Note: This table reports customer level summary statistics (n. of observation, mean and standard error in the control group) and assesses balance across treatment and control groups using a regressions. the dependent is the observable in the first column and is regressed on a treatment dummy and a binary variable for the wave.

Table 2: Power calculations for main outcomes

Outcomes	Hypothesis	N. control	N. treated	Mean	SD	Effect size	% change
RQ1							
Water usage	treat vs control	38622	70410	112.703	77.015	1.366	0.012
RQ2							
Gas usage	treat vs control	34988	54539	1036.125	769.781	14.772	0.014
Electricity usage	treat vs control	30248	46134	2268.783	1298.068	26.906	0.012
RQ3							
Water usage	treat & other report vs control & no report	21606	19976	112.703	77.015	2.118	0.019
	treat & other report vs treat & no other report	17016	19976	112.703	77.015	2.251	0.02
Gas usage	treat & other report vs control & no report	18929	18747	1036.126	769.7808	22.22205	0.021
	treat & other report vs treat & no other report	16059	18747	1036.126	769.7808	23.18912	0.022
EE usage	treat & other report vs control & no report	15651	16895	2268.783	1298.068	40.347	0.018
	treat & other report vs treat & no other report	14597	16895	2268.783	1298.068	41.096	0.018
RQ4							
Water usage	treat & high usage vs control & low usage	19311	35209	112.703	77.015	1.932	0.017
	treat & high usage vs treat & low usage	35209	35209	112.703	77.015	1.626	0.014
	treat & high engagement vs control & low engagement	19311	35209	112.703	77.015	1.932	0.017
	treat & high engagement vs treat & low engagement	35209	35209	112.703	77.015	1.626	0.014

Note: Effect size is calculated with a power of 0.8 and level of significance level of 0.05. Mean and SD of outcomes are computed for 2018.

References

Abrahamse, W. (2019). How can people save the planet? *Nature Sustainability*, 2(4):264–264.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9-10):1082–1095.

Allcott, H. (2015). Site Selection Bias in Program Evaluation *. *The Quarterly Journal of Economics*, 130(3):1117–1165.

Allcott, H. and Kessler, J. B. (2015). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. Technical report, National Bureau of Economic Research.

Allcott, H., Mullainathan, S., and Taubinsky, D. (2011). Externalizing the Internality. *New York University Working Paper*.

Allcott, H. and Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, 104(10):3003–3037.

Andor, M. A., Gerster, A., Peters, J., and Schmidt, C. M. (2017). Social norms and energy conservation beyond the us. Ruhr economic papers 714.

ARERA (2018). Monitoraggio dei mercati di vendita al dettaglio dell'energia elettrica e del gas. Technical report, Autortità di Regolazione per Energia Reti e Ambiente.

Ayres, I., Raseman, S., and Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *The Journal of Law, Economics, and Organization*, 29(5):992–1022.

Bernedo, M., Ferraro, P. J., and Price, M. (2014). The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation. *Journal of Consumer Policy*, 37(3):437–452.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Milkman, K. L. (2015). The Effect of Providing Peer Information on Retirement Savings Decisions. *The Journal of Finance*, 70(3):1161–1201.

Bhanot, S. P. (2017). Rank and response: A field experiment on peer information and water use behavior. *Journal of Economic Psychology*, 62:155 – 172.

Bonan, J., Cattaneo, C., d'Adda, G., and Tavoni, M. (2019a). Can we make social information programs more effective? the role of identity and values. RFF Working Paper 19-21.

Bonan, J., Cattaneo, C., d'Adda, G., and Tavoni, M. (2019b). Civic Capital and the Effect of Social Information. Development Working Papers 431, Politecnico di Milano.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences*, 116(12):5293–5298.

Brent, D. A., Cook, J. H., and Olsen, S. (2015a). Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627.

Brent, D. A., Cook, J. H., and Olsen, S. (2015b). Social Comparisons, Household Water Use, and Participation in Utility Conservation Programs: Evidence from Three Randomized Trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Byrne, D. P., Nauze, A. L., and Martin, L. A. (2018). Tell me something i don't already know: Informedness and the impact of information programs. *The Review of Economics and Statistics*, 100(3):510–527.

Carlsson, F., Jaime, M., and Villegas, C. (2020). Behavioral spillover effects from a social information campaign. *Journal of Environmental Economics and Management*, page 102325.

Costa, D. L. and Kahn, M. E. (2011). Electricity Consumption and Durable Housing: Understanding Cohort Effects. *American Economic Review*, 101(3):88–92.

Dolan, P. and Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, 47:1–16.

d'Adda, G., Capraro, V., and Tavoni, M. (2017). Push, don't nudge: Behavioral spillovers and policy instruments. *Economics Letters*, 154:92–95.

Ferraro, P. J. and Miranda, J. J. (2013a). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics*, 35(3):356 – 379.

Ferraro, P. J. and Miranda, J. J. (2013b). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics*, 35(3):356–379.

Ferraro, P. J., Miranda, J. J., and Price, M. K. (2011). The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment. *American Economic Review*, 101(3):318–22.

Ferraro, P. J. and Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1):64–73.

Frey, B. S. and Meier, S. (2004). Social Comparisons and Pro-social Behavior: Testing Conditional Cooperation in a Field Experiment. *American Economic Review*, 94(5):1717–1722.

Galizzi, M. M. and Whitmarsh, L. (2019). How to Measure Behavioral Spillovers: A Methodological Review and Checklist. *Frontiers in Psychology*, 10.

Gerarden, T. D., Newell, R. G., and Stavins, R. N. (2017). Assessing the Energy-Efficiency Gap. *Journal of Economic Literature*, 55(4):1486–1525.

Gerber, A. S. and Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, 71(1):178–191.

Gillingham, K., Keyes, A., and Palmer, K. (2018). Advances in evaluating energy efficiency policies and programs. *Annual Review of Resource Economics*, 10(1):null.

Goette, L., Leong, C., and Qian, N. (2019). Motivating household water conservation: A field experiment in Singapore. *PLOS ONE*, 14(3):e0211891.

Jachimowicz, J. M., Hauser, O. P., O'Brien, J. D., Sherman, E., and Galinsky, A. D. (2018). The critical role of second-order normative beliefs in predicting energy conservation. *Nature Human Behaviour*, 2(10):757–764. Number: 10 Publisher: Nature Publishing Group.

Jaime Torres, M. M. and Carlsson, F. (2018). Direct and spillover effects of a social information campaign on residential water-savings. *Journal of Environmental Economics and Management*, 92:222–243.

Jessoe, K., Lade, G., Loge, F., and Spang, E. (2017). Spillovers from Behavioral Interventions: Experimental Evidence from Water and Energy Use. *Economics Working Papers*.

Jessoe, K., Lade, G. E., Loge, F., and Spangi, E. (2019). Residential water conservation during drought: Experimental evidence from three behavioral interventions. mimeo.

List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2017). Harnessing Policy Complementarities to Conserve Energy: Evidence from a Natural Field Experiment. Working Paper 23355, National Bureau of Economic Research.

Maki, A., Carrico, A. R., Raimi, K. T., Truelove, H. B., Araujo, B., and Yeung, K. L. (2019). Meta-analysis of pro-environmental behaviour spillover. *Nature Sustainability*, 2(4):307–315.

Miranda, J. J., Datta, S., and Zoratto, L. (2020). Saving Water with a Nudge (or Two): Evidence from Costa Rica on the Effectiveness and Limits of Low-Cost Behavioral Interventions on Water Use. *The World Bank Economic Review*.

Nilsson, A., Bergquist, M., and Schultz, W. P. (2017). Spillover effects in environmental behaviors, across time and context: a review and research agenda. *Environmental Education Research*, 23(4):573–589.

Shang, J. and Croson, R. (2009a). A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *The Economic Journal*, 119(540):1422–1439.

Shang, J. and Croson, R. (2009b). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal*, 119(540):1422–1439.

Sunstein, C. R. (2016). Nudges That Fail. SSRN Scholarly Paper ID 2809658, Social Science Research Network, Rochester, NY.

Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., and Vandenbergh, M. P. (2014). Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change*, 29:127–138.

Vivalt, E. (2015). Heterogeneous treatment effects in impact evaluation. *American Economic Review*, 105(5):467–70.