Online Appendix to "Reducing Nonpayment for Public Utilities: Experimental Evidence from South Africa"

Andrea Szabó
Department of Economics
University of Houston
aszabo2@uh.edu

Gergely Ujhelyi
Department of Economics
University of Houston
gujhelyi@uh.edu

June 1, 2015

Abstract

This document, not intended for publication, contains the additional results referred to in our paper.

1 Additional tables and figures

Table 1: Payment and consumption data in levels

Variable	Winsorized mean	Mean	Std. dev
January - March total payment	265.028	260.331	377.398
January payment	77.733	91.392	192.335
February payment	80.962	92.594	209.071
March payment	90.031	109.300	237.036
April payment	65.758	80.946	190.688
May payment	63.037	261.421	5477.822
January - March average consumption	13.134	13.983	13.097
January consumption	10.530	11.164	11.215
February consumption	15.377	17.527	27.050
March consumption	12.167	13.258	20.016
April consumption	12.563	14.260	21.470
May consumption	13.665	14.908	17.348
June consumption	10.068	14.864	50.543

Notes: : Summary statistics of payment (Rand) and consumption (kl) data. The second column presents means for the data Winsorized at 5 percent on both tails. N=966

Table 2: Treatment effects on indigent status

	(1)	(2)	(3)	(4)
Treatment	0.013	0.002	0.002	0.003
	(0.030)	(0.010)	(0.010)	(0.010)
Strata indicators	No	Yes	Yes	Yes
Baseline dep. var.	No	No	Yes	Yes
Demographic controls	No	No	No	Yes

Notes: : Each cell presents the estimated treatment effect on indigent status from a different regression. The dependent variable is an indicator for indigent status in January (the month following the treatment), mean = 0.318. Columns (1-4) correspond to different specifications. 'Demographic controls' are household size, number of employed members, education, income, and whether the household has hot water. Robust standard errors in parentheses. N = 966. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 3: Replication of Table 3 from the paper with demographic controls

Dependent variable		(1) Payment amount	(2) Payment propensity	(3) Payment frequency	(4) Consumption
Panel A: First three months after treatment					
Jan-March	Treatment effect	0.268**	0.042*	0.091**	0.005
		(0.127)	(0.022)	(0.046)	(0.030)
	Control mean	3.156	0.537	1.035	2.468
Panel B: By month					
Jan	Treatment effect	0.220	0.038		-0.050
		(0.136)	(0.026)		(0.058)
	Control mean	1.739	0.332		2.122
Feb	Treatment effect	0.227*	0.044*		0.016
		(0.136)	(0.026)		(0.017)
	Control mean	1.760	0.340		2.625
March	Treatment effect	0.129	0.024		-0.070
		(0.136)	(0.026)		(0.094)
	Control mean	1.940	0.363		1.733
Apr	Treatment effect	-0.146	-0.029		-0.052
		(0.132)	(0.025)		(0.084)
	Control mean	1.722	0.330		2.024
May	Treatment effect	-0.052	-0.019		-0.078
		(0.132)	(0.025)		(0.069)
	Control mean	1.568	0.306		2.268
June	Treatment effect				0.023
					(0.059)
	Control mean				2.097
p-values for equal treatment effects					
Feb = Jan		0.970	0.867		0.258
March = Jan		0.599	0.652		0.853
Apr = Jan		0.026	0.028		0.985
May = Jan		0.092	0.055		0.743
June = Jan					0.343

Notes: The table replicates Table 3 from the main text including demographpic control variables for which the balance test in Table 1 had a p-value below 0.5. These are: household size, number of employed members, education, income, and whether the household has hot water. Each cell presents the estimated treatment effect from a different regression. All specifications control for sampling strata indicators and the value of the dependent variable for the 3 months prior to the treatment. The p-values for equal treatment effects are from Chi2 tests on the equality of the treatment coefficients when each pair of regressions is estimated as a system. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 4: Treatment effects on payment amount and propensity: Tobit and Probit estimates

Period	Treatment effect on payment (1)	Treatment effect on payment | payment > 0 (2)	Treatment effect on payment propensity (3)
Jan-March	0.367**	0.258**	0.082**
	(0.171)	(0.120)	(0.038)
January	0.228*	0.192*	0.055*
	(0.128)	(0.108)	(0.032)
February	0.247*	0.202*	0.064**
	(0.133)	(0.109)	(0.033)
March	0.200	0.159	0.045
	(0.142)	(0.113)	(0.033)

*Notes:*Marginal effects of the treatment indicator from Tobit (columns 1 and 2) and Probit (column 3) regressions. The first column gives the period of the dependent variable. In columns (1) and (2), the dependent variable is log payment over the given period. Column (1) presents unconditional marginal effects, and column (2) marginal effects conditional on positive payments (from the same regression). In column (3) the dependent variable is an indicator equal to 1 if the household made a payment over the given period and 0 otherwise. Columns (1) and (2) control for total payment in the 3 months before the treatment, and column (3) controls for whether a payment was made during this period (marginal effects are evaluated at the means of the controls). Robust standard errors in parentheses. N = 966. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 5: Treatment effects on consumption: Tobit estimates

Period	Treatment effect on consumption (1)	Treatment effect on consumption | consumption > 0 (2)
Jan-March	0.006	0.006
	(0.030)	(0.030)
January	-0.058	-0.053
	(0.063)	(0.057)
February	0.018	0.018
	(0.018)	(0.018)
March	-0.098	-0.069
	(0.105)	(0.074)

Notes: Marginal effects of the treatment indicator from Tobit regressions. The first column gives the period of the dependent variable. The dependent variable is log consumption over the given period. Column (1) presents unconditional marginal effects, and column (2) marginal effects conditional on positive consumption (from the same regression). All regressions control for average consumption in the 3 months before the treatment (marginal effects are evaluated at the means of the controls). Robust standard errors in parentheses. N = 966. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 6: Replication of Table 5 from the paper with demographic controls

	(1) Response in kl	(2) Reads consumption from bill	(3) Consumption accurate	(4) Tariff in ballpark	(5) Tariff error	(6) Increasing tariff	(7) Quiz score
Treatment effect	0.035*	0.045	0.033	-0.017	-3.660	-0.022	0.047
	(0.020)	(0.037)	(0.023)	(0.015)	(8.892)	(0.030)	(0.065)
Control mean and std. dev.	0.085	0.379	0.095	0.054	81.796	0.711	2.456
	(0.279)	(0.486)	(0.294)	(0.225)	(232.596)	(0.454)	(0.963)
Multiple inference p-values	0.385	0.521	0.429	0.522	0.745	0.728	0.728
N	953	731	731	820	396	964	965

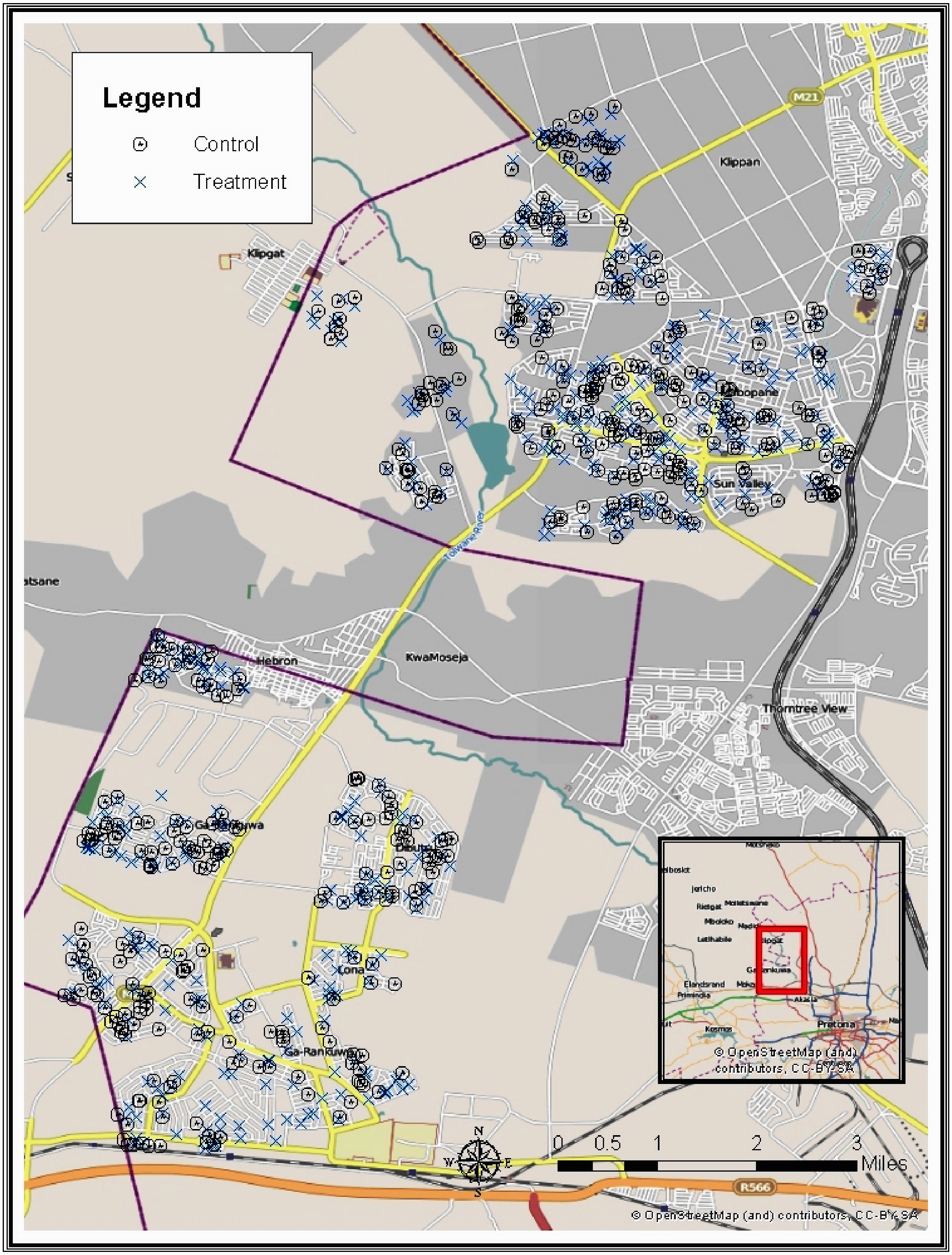
Notes: The table replicates Table 5 from the main text including demographpic control variables for which the balance test in Table 1 had a p-value below 0.5. These are: household size, number of employed members, education, income, and whether the household has hot water. Each column corresponds to a different regression. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. We also prresent p-values that control for the false discovery rate under multiple inference using the Benjamini and Hochberg (1995) method. These are computed taking into account the full set of 14 information measures we consider (see Table 14 below). Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 7: Survey effects, 3 month average

Dep. var:	March-May payment (1)	March-May payment (0/1) (2)	March-May consumption (3)
Control	0.141	0.012	0.035
	(0.131)	(0.022)	(0.050)
N	985	985	985

Notes: The table estimates survey effects by comparing the control group in our study (Control = 1) to 500 randomly selected households who did not participate in our study. Each column corresponds to a different regression. The column headings give the dependent variable (the time period is the 3 months following the follow-up survey). Every regression controls for sampling strata indicators and the pre-survey value of the dependent variable. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Figure 1: Sample area



2 Approximate cost calculation for the education campaign

Figure 2 presents a back-of-the-envelope cost calculation for the education campaign. Since the provider used its regular employees and resources for the campaign, the cost of these is necessarily a rough approximation (e.g., alternative costs of these resources are unknown). We present three cost measures. The main direct cost associated with the campaign was the cost of printing the education materials, this is shown under Total 1. Total 2 also includes the imputed wage of the education officers making the household visits. Total 3 includes an imputed wage for a supervisor as well as the cost of transportation (fuel cost, vehicle amortization). All figures are in 2012 monetary values (dollar amounts are based on exchange rates in effect at the time of the campaign).

Figure 2: Cost calculation of the education campaign

Notes	Description	Quantity	Unit cost	Total (ZAR)	Total (USD)
1	Printing	500 brochures, 5 pages each, 2500 pages total	3.06 ZAR / page	7660.16	885.56
	TOTAL 1			**7660.16** **15.32 / hh**	**885.56** **1.77 / hh**
2	Wage for education officers	500 households visited, 1 hour per visit (including travel), 500 hours total	170 ZAR / 8 hours	10625.00	1228.32
	TOTAL 2			**18285.16** **36.57 / hh**	**2113.88** **4.23 / hh**
2	Wage for supervisor	7 days, 8 hours per day, 56 hours total	455 ZAR / 8 hours	3185.00	368.21
3	Car rental, amortization	5 km per visit + 20 km, 2520 km total	2.12 ZAR / km	5342.40	617.61
3	Fuel cost	2520 km traveled, fuel economy at 6 liter per 100km, 151.2 liter total	11.23 ZAR / liter	1697.97	196.29
	TOTAL 3			**28510.13** **57.02 / hh**	**3295.99** **6.59 / hh**

Notes: 1: Based on printing cost we incurred during the pilot study using a commercial printing company (50 copies for each brochure). Since the provider used its own facilities for the remaining printing (unit cost unknown), we base our calculation on 802: Based on typical monthly salaries. Statistics South Africa reports monthly wages by education level. The median wage for a worker with completed secondary education is 3500 Rand per month. This is what we used for fieldworkers. The median earnings for a worker with tertiary education is 10,000 Rand per month, this was used as salary for the fieldworker supervisor. Both of these numbers are somewhat higher than the salary ranges for comparable jobs in the area reported on payscale.com, which publishes salaries by occupation. Not all of the education officers had completed secondary education, and this is not necessary for the campaign. 3: Fuel cost and car rental / amortization cost is based on expenditures billed by the survey company we used in our surveys.

3 Water conservation

For more direct evidence on an increased desire for water conservation, we look at various water saving activities reported by the households. Our surveys asked whether the household recently took actions to save water, listing several possibilities. Table 8 looks for treatment effects in these answers. Treated households are 10 percentage points more likely to report fixing leaks around the house and 8.5 percentage points more likely to report saving water during laundry ("Use washing machine less / use fuller loads"). Treated households also report taking more actions than control households, although the fraction of households taking no action does not differ significantly between the two groups. Subject to the usual caveats associated with self-reported behavior, this suggests that the treatment primarily increased conservation on the intensive margin, among households already taking steps to save water.[1]

4 Information

4.1 Information effects by consumption quartile

Tables 9-12 repeat the information results from the main text separately for each consumption quartile.

4.2 Other information measures

We also considered further information measures besides the ones reported in the main text. In Table 13, we present treatment effects for the full set of 14 information measures we have considered. The new measures are as follows.

Bill hard to understand: Indicator equal to 1 if the respondent reports that the bill is hard to understand, 0 otherwise.

Q1 correct: Indicator for whether the respondent answered the following question correctly: "Please take a guess: Do you think more water is used by the baths/showers your household takes during the month OR by washing your clothes during the month?"

[1] In principle, households could misreport taking conservation actions to satisfy perceived social expectations. However, it is not clear why this incentive would be different in the control and treatment groups. We also find it reassuring that there are no significant differences in reports of using rainwater or reusing household water in Table 8. Neither of these practices was mentioned in our education campaign. Households report more water saving actions in those areas that were explicitly covered in the campaign.

Table 8: Effect of treatment on water saving actions

Dep. var.	Control mean and std. dev.	Treatment effect	Multiple inference p-value
Use rainwater	0.017	-0.005	
	(0.128)	(0.007)	0.597
Reuse water	0.285	0.015	
	(0.452)	(0.029)	0.618
Repair leaks	0.362	0.104***	
	(0.481)	(0.031)	0.007
Save with laundry	0.249	0.083***	
	(0.433)	(0.029)	0.018
Save with irrigation	0.272	-0.040	
	(0.445)	(0.028)	0.263
Number of actions	1.276	0.166**	
	(1.097)	(0.070)	0.045
No action	0.241	-0.034	
	(0.428)	(0.026)	0.283

Notes: : Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. Except for the last two, these are dummies for whether the respondent reported having taken the action to save water. 'Number of actions' is the number of actions the household reported. 'No action' is a dummy equal to 1 if the household did not report taking any action. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. The last column presents p-values that control for the false discovery rate under multiple inference using the Benjamini and Hochberg (1995) method. N = 965. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 9: Effect of treatment on information, 1st consumption quartile

Dep. var	Control mean and std. dev.	N	Treatment effect
Response in kl	0.037	233	0.051
	(0.191)		(0.038)
Reads consumption from bill	0.286	176	0.146*
	(0.455)		(0.076)
Consumption accurate	0.052	176	0.105**
	(0.223)		(0.043)
Tariff in ballpark	0.052	203	0.001
	(0.222)		(0.036)
Tariff error	58.136	97	16.373
	(58.309)		(19.642)
Increasing tariff	0.773	237	-0.102
	(0.421)		(0.062)
Quiz score	2.582	237	-0.192
	(0.923)		(0.136)

Notes: Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 10: Effect of treatment on information, 2nd consumption quartile

Dep. var	Control mean and std. dev.	N	Treatment effect
Response in kl	0.075	248	0.008
	(0.265)		(0.036)
Reads consumption from bill	0.368	193	-0.046
	(0.485)		(0.078)
Consumption accurate	0.179	193	-0.039
	(0.385)		(0.056)
Tariff in ballpark	0.042	221	-0.025
	(0.202)		(0.024)
Tariff error	59.180	113	1.473
	(63.685)		(15.673)
Increasing tariff	0.695	247	-0.058
	(0.462)		(0.062)
Quiz score	2.402	248	0.036
	(1.003)		(0.143)

Notes: Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 11: Effect of treatment on information, 3rd consumption quartile

Dep. var	Control mean and std. dev.	N	Treatment effect
Response in kl	0.105	246	0.070
	(0.308)		(0.047)
Reads consumption from bill	0.395	186	0.123
	(0.492)		(0.077)
Consumption accurate	0.058	186	0.090**
	(0.235)		(0.045)
Tariff in ballpark	0.064	195	-0.001
	(0.246)		(0.034)
Tariff error	156.021	97	-23.205
	(469.983)		(34.988)
Increasing tariff	0.678	248	0.061
	(0.469)		(0.061)
Quiz score	2.357	248	0.222
	(0.984)		(0.136)

Notes: Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 12: Effect of treatment on information, 4th consumption quartile

Dep. var	Control mean and std. dev.	N	Treatment effect
Response in kl	0.118	226	0.034
	(0.324)		(0.047)
Reads consumption from bill	0.449	176	-0.107
	(0.500)		(0.082)
Consumption accurate	0.071	176	0.009
	(0.259)		(0.047)
Tariff in ballpark	0.059	201	-0.043
	(0.236)		(0.027)
Tariff error	63.011	89	-8.192
	(73.823)		(11.425)
Increasing tariff	0.702	232	0.040
	(0.459)		(0.064)
Quiz score	2.496	232	0.036
	(0.932)		(0.138)

Notes: Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Q2 correct: Indicator for whether the respondent answered the following question correctly: "Please take a guess: Do you think more water is used if you fill 2 two-liter bottles of soda with water OR if you flush the toilet once?"

Q3 correct: Indicator for whether the respondent answered the following question correctly: "Please take a guess: Do you think more water is used if you use the outside hose for 10 minutes OR if you do one load of laundry?"

Q4 correct: Indicator for whether the respondent answered the following question correctly: "Please take a guess: Do you think more water is used if you open the tap for 1 minute OR with the water a person drinks in a day?"

(*Quiz score* is the number of correct answers to the above questions.)

Can check leaks: Indicator equal to 1 if the respondent reports knowing how to check for leaks, 0 otherwise.

Can repair leaks: Indicator equal to 1 if the respondent reports knowing how to repair leaks, 0 otherwise.

Overall, our education treatment had at most a modest effect on households' knowledge on average. Out of the 14 measures in Table 13, we find three significant differences between treatment and control households, only one of which remains significant once we control for multiple inference. According to these results, more households use the word "kiloliter," have a realistic idea about the amount of water used when flushing the toilet, and more of them report knowing how to check for leaks in response to our treatment. We do not see any improvement in consumers' familiarity with their bill.

Table 13: Effect of treatment on information

Dep. var	Control mean and std. dev.	N	Treatment effect	Multiple inference p-value
Response in kl	0.085	953	0.037*	
	(0.279)		(0.020)	0.364
Bill hard to understand	0.074	952	-0.023	
	(0.262)		(0.016)	0.395
Reads consumption from bill	0.379	731	0.043	
	(0.486)		(0.037)	0.55
Consumption accurate	0.095	731	0.038	
	(0.294)		(0.023)	0.364
Tariff in ballpark	0.054	820	-0.016	
	(0.225)		(0.015)	0.55
Tariff error	81.796	396	-6.141	
	(232.596)		(9.256)	0.71
Increasing tariff	0.711	964	-0.023	
	(0.454)		(0.030)	0.676
Quiz score	2.456	965	0.059	
	(0.963)		(0.064)	0.634
Q1 correct	0.461	966	-0.005	
	(0.499)		(0.032)	0.887
Q2 correct	0.699	966	0.050*	
	(0.459)		(0.028)	0.364
Q3 correct	0.607	965	-0.005	
	(0.489)		(0.032)	0.887
Q4 correct	0.689	966	0.017	
	(0.463)		(0.030)	0.72
Can check leaks	0.659	920	0.101***	
	(0.474)		(0.030)	0.014
Can repair leaks	0.367	951	0.011	
	(0.482)		(0.031)	0.842

Notes: Each cell presents the estimated treatment effect from a different regression. The first column gives the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. The last column presents p-values that control for the false discovery rate under multiple inference using the Benjamini and Hochberg (1995) method. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

4.3 Information spillovers

Treated households could talk to their neighbors about what they have learned, or they could give them the information brochures. Even if the treatment was effective at increasing households' information, such spillovers could result in no difference in information between the treatment and control groups. Could this be responsible for the lack of large information effects we found above? Note that in most cases, we did not simply find the information of treatment and control groups to be similar, but also that they were both similarly low both before and after the treatment (post-treatment means are given in Table 13). While spillover effects from our treatment could potentially explain the first of these patterns, they are unlikely to account for the second. If the treatment had increased information and there were spillovers, we would expect to find increased knowledge in both the treatment and control groups.

To formally test whether spillovers were present in our intervention, we collected data to identify individuals who would be most likely to be exposed to information spillovers. First, our survey collected information on whether the respondent had talked to his neighbors or friends about water in the previous 6 months. If there were information spillovers, these would likely be present among the 39% who reported talking about water with others. Second, we collected each household's GPS coordinates and thus know their location relative to other households. Information spillovers could occur between neighbors, and we can capture this by creating an indicator for whether a household has other treated households nearby.

Let $Exposure$ represent one of the proxies for exposure to information spillovers. We estimate the regression

$$Y_i = \beta_0 + \beta_1 Treat_i + \beta_2 Exposure_i + \beta_3 Exposure_i \times Treat_i + \varepsilon_i,$$

where Y_i is one of our measures of respondent i's knowledge. If the treatment did have an effect on Y_i, but large spillovers caused us to find no effect, then we expect to find $\beta_1 > 0$ (treatment effect among those not exposed to spillovers) and $\beta_2 > 0$ (spillover effect in the control group). By contrast, if the treatment was indeed ineffective, we expect $\beta_1 = \beta_2 = 0$.

In Table 14, our measure of exposure is $Talks$, which takes a value of 1 if the respondent reports talking to neighbors about water in the previous 6 months. Our dependent variables are the main information measures. The table also presents an F-test and the corresponding p-value for the hypothesis that $\beta_1 = \beta_2 = 0$ (no treatment effect and no spillovers). For 6 out of 10 variables, this hypothesis is not rejected at conventional levels of significance even without adjusting the p-values for multiple inference. For the rest, in columns 1, 7, and 8 the coefficient on $Talks$ is negative: if anything, individuals who talk to others have less information. Column (9) is the only one showing evidence consistent with spillover effects: households in the control group who talk to others about water are more likely to know how to check for leaks. Based on this, the significant average treatment effect we found in Table 7 in the paper for this variable may have underestimated the true treatment effect. In no case do we find evidence that the insignificant treatment effects in Table 7 were due to information spillovers.

Table 15 presents corresponding regressions using GPS coordinates to identify a household's neighbors, and using the treatment status of a household's neighbors to capture potential exposure to information spillovers. The variable $Treated\ neighbors$ takes a value of 1 if there is one or more treated household in a 30 meter radius around the respondent. 7% of the households in our study have such a neighbor, and the number of treated neighbors ranges between 0 and 2. The results in Table 15 echo those seen with the $Talks$ measure. They strongly reject the idea that spillover effects between neighbors could explain any of the statistically insignificant information effects found above.

Table 14: Checking for spillover effects I: Talking to neighbors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Response in kl	Bill hard to under-stand	Reads consump-tion from bill	Consumption accurate	Tariff in ballpark	Tariff error	Increasing tariff	Quiz score	Can check leaks	Can repair leaks
Treatment	0.038	-0.025	0.039	0.008	-0.017	-22.946*	0.007	0.148*	0.121***	0.001
	(0.026)	(0.021)	(0.050)	(0.031)	(0.018)	(13.054)	(0.036)	(0.082)	(0.040)	(0.041)
Talks	-0.029	-0.018	-0.063	-0.030	0.005	-44.679	-0.087*	-0.074	0.168***	0.003
	(0.027)	(0.025)	(0.054)	(0.034)	(0.025)	(48.499)	(0.045)	(0.094)	(0.046)	(0.047)
Talks x Treatment	0.005	0.001	0.016	0.067	0.003	63.682	-0.064	-0.205	-0.068	0.031
	(0.040)	(0.032)	(0.077)	(0.051)	(0.032)	(53.143)	(0.062)	(0.134)	(0.060)	(0.066)
N	947	946	726	726	815	391	958	959	914	946
F test (Treatment, Talks)	2.806	0.707	1.755	0.638	0.697	1.793	2.347	2.987	7.908	0.002
p value	0.061	0.493	0.174	0.528	0.498	0.168	0.096	0.051	0.000	0.998
Multiple inference p value	0.204	0.587	0.290	0.587	0.587	0.290	0.240	0.204	0.004	0.998
Mean dep. var.	0.104	0.062	0.397	0.114	0.045	70.591	0.699	2.485	0.710	0.371

Notes: Each column lists coefficient estimates from a different regression. The column headings give the dependent variable. 'Talks' is 1 if the respondent answered Yes to 'In the previous 6 months have you talked to friends or neighbors about the way you use or save water?' All regressions control for sampling strata indicators and the baseline value of the dependent variable. Robust standard errors in parentheses. F test and corresponding p value reported for the joint hypothesis that the coefficients on Treatment and Talks are both zero. Multiple inference p values adjust these p values to control for the false discovery rate using the Benjamini and Hochberg (1995) method. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 15: Checking for spillover effects II: Treated neighbors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Response in kl	Bill hard to under-stand	Reads consump-tion from bill	Consumption accurate	Tariff in ballpark	Tariff error	Increasing tariff	Quiz score	Can check leaks	Can repair leaks
Treatment	0.033*	-0.023	0.026	0.039	-0.012	-1.838	-0.023	0.060	0.106***	-0.001
	(0.020)	(0.015)	(0.038)	(0.024)	(0.015)	(7.750)	(0.031)	(0.066)	(0.031)	(0.033)
Treated neighbors	-0.008	0.121*	-0.383***	-0.036	-0.004	154.766	-0.070	-0.082	0.112	-0.232***
	(0.057)	(0.073)	(0.070)	(0.056)	(0.050)	(141.348)	(0.092)	(0.184)	(0.087)	(0.070)
Treated neighbors x Treatment	0.055	-0.020	0.313**	-0.009	-0.053	-141.600	0.014	-0.002	-0.084	0.214**
	(0.088)	(0.093)	(0.128)	(0.093)	(0.053)	(141.282)	(0.125)	(0.282)	(0.115)	(0.107)
N	953	952	731	731	820	396	964	965	920	951
F test (Treatment, Treated neighbors)	1.441	2.896	17.619	1.712	0.300	0.633	0.494	0.602	5.943	5.829
p value	0.237	0.056	0.000	0.181	0.741	0.532	0.610	0.548	0.003	0.003
Multiple inference p value	0.395	0.140	0.001	0.362	0.741	0.678	0.678	0.678	0.010	0.010

Notes: Each column lists coefficient estimates from a different regression. The column headings give the dependent variable. 'Treated neighbors' is 1 if the repondent has a neighbor in the treatment group (in a 30 meter radius). All regressions control for sampling strata indicators and the baseline value of the dependent variable. Robust standard errors in parentheses. F test and corresponding p value reported for the joint hypothesis that the coefficients on Treatment and Treated neighbors are both zero. Multiple inference p values adjust these p values to control for the false discovery rate using the Benjamini and Hochberg (1995) method. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

4.4 Information sharing within the household

Another possible explanation for the lack of a measured information effect is that information may not be shared within the household. As described in Section 2.3 in the paper, it makes sense to consider the household as the unit of analysis since consumption and payment are measured at the household level. However, this raises the possibility that surveyed individuals within the household are different from treated individuals. To fix ideas, suppose that the education officers met with the wife, who is responsible for paying the water bill, and the treatment successfully increased her knowledge. Suppose this information channel explains the findings above. We may still measure no treatment effect on information if our surveyors in the follow-up survey talked to the husband *and* the wife failed to share her information with him.

We perform two tests to assess the possibility that information sharing within the household might be important in explaining the findings above. First, based on the respondent's age and gender, we identify households where the same respondent is likely to have answered the baseline and the follow-up survey. If the same person answered both surveys, it is more likely that (s)he was also home during the education visit. Under the information story, these households should show the biggest increase in knowledge relative to the control group. We have 28 such households in the control and 25 in the treatment group. Including this indicator and its interaction with treatment status yields a significant interaction in only 2 out of 10 cases, but one of these has the wrong sign (Table 16). In column (10), we find a significant treatment effect on knowing how to fix a leak for households with a matched respondent. For the other 9 information measures, treated households where the same person was home during both surveys do not have significantly more information than others.

Our second test is based on the idea that if information sharing within the household is a major factor, we would expect treatment effects to diminish as households get larger. This is both because information sharing within the household becomes harder in a larger household, and because a larger household makes it more likely that the education officers and the surveyors met with different members of the household. In Table 17 we include an interaction of household size with treatment status. This is never statistically significant: relative to the control group, smaller treated households do not have more information than larger households.

Table 16: Information sharing within the household I: Same respondent

	(1) Response in kl	(2) Bill hard to understand	(3) Reads consumption from bill	(4) Consumption accurate	(5) Tariff in ballpark	(6) Tariff error	(7) Increasing tariff	(8) Quiz score	(9) Can check leaks	(10) Can repair leaks
Treatment	0.043*	-0.033*	0.063	0.047*	-0.031*	-9.364	-0.028	0.087	0.122***	0.024
	(0.023)	(0.018)	(0.043)	(0.027)	(0.017)	(9.684)	(0.035)	(0.075)	(0.035)	(0.036)
Same respondent	0.005	0.068	0.262**	0.168*	-0.024	10.479	-0.143	0.146	0.121	-0.071
	(0.054)	(0.066)	(0.108)	(0.100)	(0.046)	(24.365)	(0.097)	(0.184)	(0.091)	(0.095)
Same respondent x Treatment	0.117	0.023	-0.298*	-0.125	0.049	-19.309	0.039	-0.241	-0.178	0.234*
	(0.100)	(0.104)	(0.166)	(0.142)	(0.058)	(41.208)	(0.138)	(0.289)	(0.131)	(0.140)
N	763	763	578	578	666	327	774	774	738	761
F test (Treatment $\mid$ Same respondent)	2.696	0.010	2.156	0.316	0.104	0.473	0.007	0.302	0.202	3.688
p value	0.101	0.921	0.143	0.574	0.747	0.492	0.935	0.583	0.653	0.055
Multiple inference p value	0.477	0.935	0.477	0.933	0.934	0.933	0.935	0.933	0.933	0.477

Notes: Each column lists coefficient estimates from a different regression. The column headings give the dependent variable. 'Same respondent' is 1 the followup respondent's gender and age match the baseline respondent's gender and age, age+1, or age+2. All regressions control for sampling strata indicators and the baseline value of the dependent variable. Robust standard errors in parentheses. F test and corresponding p value reported for the hypothesis of no treatment effect conditional on Same respondent = 1. Multiple inference p values adjust these p values to control for the false discovery rate using the Benjamini and Hochberg (1995) method. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 17: Information sharing within the household II: HH size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Response in kl	Bill hard to understand	Reads consumption from bill	Consumption accurate	Tariff in ballpark	Tariff error	Increasing tariff	Quiz score	Can check leaks	Can repair leaks
Treatment	0.046	0.011	0.007	-0.023	-0.067**	11.972	0.071	0.265	0.128*	0.054
	(0.042)	(0.035)	(0.091)	(0.059)	(0.034)	(44.326)	(0.071)	(0.168)	(0.077)	(0.078)
HH size	-0.006	0.015**	0.001	-0.019***	-0.003	5.837	0.002	0.061**	-0.003	0.007
	(0.005)	(0.007)	(0.013)	(0.007)	(0.005)	(9.849)	(0.010)	(0.024)	(0.012)	(0.012)
HH size x Treatment	-0.002	-0.008	0.008	0.013	0.012	-4.117	-0.022	-0.048	-0.006	-0.010
	(0.008)	(0.008)	(0.019)	(0.011)	(0.008)	(10.302)	(0.015)	(0.037)	(0.017)	(0.016)
N	950	949	730	730	817	395	961	962	917	948
Treatment at mean HH size	0.037*	-0.024	0.041	0.037	-0.015	-6.145	-0.025	0.056	0.100***	0.011
	(0.020)	(0.016)	(0.037)	(0.023)	(0.015)	(9.232)	(0.030)	(0.065)	(0.030)	(0.031)

Notes: Each column lists coefficient estimates from a different regression. The column headings give the dependent variable. 'HH size' is the number of individuals in the household. All regressions control for sampling strata indicators and the baseline value of the dependent variable. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

5 Heterogenous treatment effects

We focus on six dimensions of heterogeneity: household income, education, baseline knowledge, restricted status at baseline, indigent status at baseline, and payment behavior before the treatment. The last four of these variables were also used in our stratified sampling procedure because they are natural candidates for determinants of households' ability or willingness to respond to our treatment. We add income, education and baseline knowledge because they are obvious dimensions of heterogeneity in the context of an information campaign and payment behavior.

To maximize our sample size, we use income and education measures from the follow-up survey. We measure income using below/above median indicators, and education by whether the respondent completed high school (the share of such respondents is 58% in the control and 57% in the treatment group). For payment, we use an indicator for whether the household paid any of their water bills in the 6 months before the treatment. Tables 18 and 19 show that there are no significant differences in pre-treatment consumption or payment between the control and treatment group in any of the subgroups we use.

Heterogenous effects by income, education, and baseline knowledge are discussed in the main text. Here we look at indigent status, restricted status, and payment behavior. Table 20 studies the heterogeneity of treatment effects on payment and consumption. Each panel interacts our treatment indicator with one of the variables mentioned above. In each case, a test of heterogenous treatment effects is equivalent to asking whether the interaction term is statistically significant. To improve readability, we omit the coefficient on the grouping variables from the table. We do not find any evidence of heterogenous treatment effects for payment and consumption.

Table 21 adds the same interactions, one at a time, to the information regressions. In Panel B, we find evidence that the treatment increased the ability of indigent households to tell their consumption from their bill (columns (3) and (4)). However, as we saw in Table 20, this group was not driving the payment results.

Table 18: Average consumption across subgroups before the treatment

	Control	Treatment	Difference	p-value
Non-restricted	2.514	2.516	0.002	0.98
Restricted	2.553	2.532	-0.021	0.82
Non-indigent	2.531	2.525	-0.006	0.91
Indigent	2.511	2.509	-0.001	0.99
Low education	2.497	2.496	-0.001	0.99
High education	2.548	2.535	-0.013	0.83
Low income	2.483	2.489	0.006	0.93
High income	2.550	2.599	0.049	0.51
Tariff in ballpark: no	2.544	2.540	-0.004	0.95
Tariff in ballpark: yes	2.477	2.503	0.026	0.84
Quiz score low	2.493	2.449	-0.044	0.58
Quiz score high	2.568	2.604	0.036	0.62
Has not paid	2.383	2.490	0.107	0.23
Has paid	2.609	2.542	-0.067	0.21

Notes: The table presents log average consumption in the 3 months before the treatment in the subgroups used for the heterogenous treatment effects analysis. Reported values are the means in each group, the difference between control and treatment, and the p-value for a t-test of zero difference. For consumption and income, 'Low' and 'High' refer to below-median and above-median, respectively. For education, they refer to whether the respondent completed high school. For the quiz score, low is 0-2, high is 3 or 4. 'Has (not) paid' refers to the 6 months before the treatment.

Table 19: Average payment across subgroups before the treatment

	Control	Treatment	Difference	p-value
Non-restricted	3.680	3.468	-0.212	0.35
Restricted	2.397	1.892	-0.505	0.14
Non-indigent	3.703	3.410	-0.293	0.20
Indigent	2.302	2.061	-0.241	0.45
Low education	3.125	2.782	-0.342	0.25
High education	3.435	3.181	-0.254	0.32
Low income	3.274	2.771	-0.503	0.08
High income	3.500	3.387	-0.113	0.70
Tariff in ballpark: no	3.284	2.966	-0.318	0.17
Tariff in ballpark: yes	3.595	3.647	0.052	0.94
Quiz score low	3.194	2.982	-0.213	0.51
Quiz score high	3.448	3.066	-0.383	0.18
Has not paid	0.000	0.000	0.000	
Has paid	5.237	5.122	-0.115	0.49

Notes: The table presents log average payment in the 3 months before the treatment in the subgroups used for the heterogenous treatment effects analysis. Reported values are the means in each group, the difference between control and treatment, and the p-value for a t-test of zero difference. For consumption and income, 'Low' and 'High' refer to below-median and above-median, respectively. For education, they refer to whether the respondent completed high school. 'Has (not) paid' refers to the 6 months before the treatment.

Table 20: Heterogenous treatment effects on payment and consumption

	(1) Payment amount	(2) Payment propensity	(3) Payment frequency	(4) Consumption
Panel A: Restricted (N = 966)				
Treatment	0.180	0.028	0.069	0.026
	(0.151)	(0.026)	(0.057)	(0.035)
Interaction	0.243	0.036	0.065	-0.071
	(0.288)	(0.050)	(0.097)	(0.068)
Panel B: Indigent (N = 966)				
Treatment	0.253*	0.031	0.070	0.008
	(0.150)	(0.025)	(0.055)	(0.036)
Interaction	-0.007	0.027	0.062	-0.010
	(0.290)	(0.051)	(0.102)	(0.065)
Panel C: Past payment (N = 960)				
Treatment	0.076	0.024	0.033	-0.017
	(0.177)	(0.033)	(0.050)	(0.051)
Interaction	0.309	0.027	0.102	0.035
	(0.248)	(0.044)	(0.085)	(0.064)

Notes: Panels A-C investigate heterogenous treatment effects by different grouping variables. 'Restricted' is 1 if the consumer was restricted at baseline. 'Indigent' is 1 if the consumer was registered as indigent at baseline. 'Past payment' is 1 if the household made a payment on their bill in the 6 months preceding the treatment. The columns in each panel correspond to separate regressions. The column headings give the dependent variable. 'Payment amount' is total payment in the 3 months following the treatment in logs; 'Payment propensity' is 1 if the household made a payment during this period, and 'Payment frequency' is the number of payments made. 'Consumption' is average consumption in the 3 months following the treatment (in logs). All regressions control for sampling strata indicators and the value of the dependent variable during the 3 months prior to the treatment. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.

Table 21: Heterogenous treatment effects on information

	(1) Response in kl	(2) Bill hard to understand	(3) Reads consumption from bill	(4) Consumption accurate	(5) Tariff in ballpark	(6) Tariff error	(7) Increasing tariff	(8) Quiz score	(9) Can check leaks	(10) Can repair leaks
Panel A: Restricted										
Treatment	0.029	-0.018	0.022	0.036	-0.018	-7.030	-0.036	0.080	0.103***	0.006
	(0.024)	(0.019)	(0.044)	(0.028)	(0.015)	(11.440)	(0.036)	(0.077)	(0.036)	(0.037)
Interaction	0.027	-0.018	0.074	0.007	0.006	4.015	0.044	-0.072	-0.005	0.018
	(0.043)	(0.033)	(0.078)	(0.050)	(0.037)	(20.018)	(0.064)	(0.141)	(0.064)	(0.070)
N	953	952	731	731	820	396	964	965	920	951
Panel B: Indigent										
Treatment	0.040*	-0.036**	-0.012	0.011	-0.026	-1.933	0.007	0.079	0.095***	0.032
	(0.024)	(0.018)	(0.044)	(0.026)	(0.018)	(8.665)	(0.035)	(0.076)	(0.036)	(0.037)
Interaction	-0.010	0.045	0.191**	0.095*	0.035	-19.123	-0.105	-0.070	0.021	-0.072
	(0.043)	(0.035)	(0.079)	(0.056)	(0.031)	(32.526)	(0.065)	(0.142)	(0.065)	(0.069)
N	953	952	731	731	820	396	964	965	920	951
Panel C: Past payment										
Treatment	0.065**	-0.014	0.100*	0.077**	-0.004	-8.868	-0.079*	0.089	0.132***	-0.009
	(0.029)	(0.023)	(0.060)	(0.035)	(0.022)	(22.242)	(0.047)	(0.099)	(0.049)	(0.050)
Interaction	-0.045	-0.015	-0.088	-0.057	-0.018	1.873	0.093	-0.050	-0.044	0.038
	(0.039)	(0.031)	(0.075)	(0.047)	(0.030)	(32.739)	(0.061)	(0.131)	(0.063)	(0.064)
N	953	952	731	731	820	396	964	965	920	951

Notes: Panels A-C investigate heterogenous treatment effects by different grouping variables. 'Restricted' is 1 if the consumer was restricted at baseline. 'Indigent' is 1 if the consumer was registered as indigent at baseline. 'Past payment' is 1 if the household made a payment on their bill in the 6 months preceding the treatment. The columns in each panel correspond to separate regressions. The column headings give the dependent variable. All regressions control for sampling strata indicators and the value of the dependent variable at baseline. Robust standard errors in parentheses. ***, **, * denote significance at 1, 5, and 10 percent, respectively.