Nudging Energy Efficiency Audits: Evidence from a Field Experiment*

Kenneth Gillingham[†] Tsvetan Tsvetanov[‡]

June 22, 2018

Abstract

This paper uses a randomized field experiment to test how information provision leveraging social norms, salience, and a personal touch can serve as a nudge to influence the uptake of residential energy audits. Our results show that a low-cost carefully-crafted notecard can increase the probability of a household to follow through with an already scheduled audit by 1.1 percentage points on a given day. The effect is very similar across individuals with different political views, but households in rural areas display a substantially greater effect than those in urban areas. Our findings have important managerial and policy implications, as they suggest a cost-effective nudge for increasing energy audit uptake and voluntary energy efficiency adoption.

Keywords: residential energy efficiency; home energy audits; non-price interventions; information provision; social norms; field experiment
JEL: D03, Q41, Q48

*The authors gratefully acknowledge the financial support of the Center for Business and the Environment at Yale (CBEY) for this project through the F.K. Weyerhaeuser Memorial Fund. We would like to thank "Next Step Living" for providing the data used in the analysis and Will Fogel, Peter Weeks, and Brian Sewell for their input in coordinating this project. We are also thankful to seminar audiences at the AERE and SEA Annual Meetings for helpful discussions.

[†]Yale University, School of Forestry and Environmental Studies, 195 Prospect Street, New Haven, CT 06511, phone: 203-436-5465, e-mail: kenneth.gillingham@yale.edu.

[‡]University of Kansas, Department of Economics, 1460 Jayhawk Boulevard, Lawrence, KS 66045, phone: 785-864-1881, e-mail: tsvetanov@ku.edu.

1 Introduction

Investments in residential energy efficiency upgrades are a widely used strategy by policymakers for reducing energy consumption and emissions from the residential sector, which currently amounts to 22% of the total annual energy use in the United States (U.S. EIA 2016). Although the role of energy efficiency in curbing greenhouse gas emissions and addressing energy security concerns is well-recognized, the use of price-based policies for encouraging optimal energy use and adoption of energy-efficient technologies continues to face political opposition, which is likely to persist in the near future in the United States. Moreover, there is a continued discussion in the literature over why there might be an "energy efficiency gap" slowing the diffusion of seemingly cost-effective energy efficiency investments, with plausible explanations including lack of information or behavioral anomalies in processing information (Gillingham and Palmer 2014). Increasing attention has therefore been given to non-price interventions that serve to inform households about the cost savings possible under energy conservation activities.

One such non-price intervention is a "residential energy audit," which is often targeted to lower-income households. An energy efficiency audit is a professional home assessment that identifies the energy efficiency investments through which a household can reduce its energy consumption and provides engineering estimates of the monthly savings associated with these investments. The reduction in residential electricity usage among audit participants can be in excess of 5 percent, as shown by previous studies (Delmas, Fischlein, and Asensio 2013; Alberini and Towe 2015). A growing body of literature, most recently summarized by Gillingham and Palmer (2014) and Allcott (2016), has established that households face uncertainty about the payoffs from investing in energy-using durable products due to imperfect information about energy costs or product-specific attributes. In this regard, a household's decision to participate in an audit resembles information search behavior that is associated with certain time and monetary costs (Holladay et al. 2016). Audit uptake would therefore depend on how these search costs compare to the perceived individual gains from information acquisition.

According to recent evidence, audit participation rates remain low, with only about 4% of the U.S. homeowners having completed an audit (Palmer, Walls, and O'Keeffe 2015). In a field experiment, Fowlie, Greenstone, and Wolfram (2015) find very low audit uptake (approximately 5% of their total sample) even in the absence of direct monetary costs, possibly due to substantial time costs associated with the application process and the audit itself. In line with this, the findings from a number of other studies imply that households may be responding rationally to the perceived benefits and costs of the audit. For instance, Holladay et al. (2016) and Allcott and Greenstone (2017) find that in instances where audits are not offered free of charge, lowering audit prices leads to an increase in completion rates. In addition, survey evidence by Palmer and Walls (2015) suggests that households may be inattentive to energy issues, which would substantially lower their perceived expected benefit from an audit and lead to low uptake. This is consistent with the findings of Holladay et al. (2016), whose results suggest that informing households about their monthly electricity usage relative to their neighbors (both in kWh and dollars) increases the probability of completing an audit. Allcott and Greenstone (2017) also find a positive effect of providing residents with information about the private and social gains from audits.

From a policy perspective, devising a sufficiently low-cost yet effective method to increase audit uptake and capture the private and social benefits from reduced uncertainty in the payoffs to energy efficiency investments promises to be an appealing solution. Based on the above evidence, incomplete (or misperceived) knowledge of the expected benefits from an audit is likely to be a deterrent to uptake. This suggests information provision as one possible pathway to increasing audit participation. Specifically, information provision can serve a dual role, by not only providing important details related to the audit, but also acting as a "nudge" to encourage follow-through on the audit.

This paper examines the effectiveness of information provision in the particular stage of the audit uptake process where a decision is faced on whether or not to complete an already scheduled audit visit. To the best of our knowledge, we are the first to examine an intervention at this stage in the process. This is a particularly crucial phase, for the issue of households reversing their initial decision to schedule an assessment visit appears to be a substantial contributing factor behind the low uptake of energy audits. This was seen in previous work (e.g., Fowlie, Greenstone, and Wolfram 2015) and is clearly seen in our data, which are based on a series of energy efficiency outreach efforts by "Next Step Living" (NSL), a major residential energy efficiency provider in New England. NSL campaigns in Connecticut (CT) during 2014 faced customer attrition of more than 60 percent between the scheduling and completion of a home audit.

In collaboration with NSL, we design and conduct a randomized field experiment in order to examine the effect of information provision on the completion of scheduled energy assessment visits. Specifically, we test the impact of a carefully-crafted written message that combines the effects of social norms (information about the energy audits from other residents in the community), salience (a reminder with information about the visit date and time), and personal touch (a personalized and signed note). Our estimates imply a boost in the audit uptake rate for the sample by 1.1 percentage points on a given day as a result of the treatment. Finally, we combine data from the experiment with local demographic, socioeconomic, and voting information in order to explore possible sources of heterogeneity in our treatment impacts and better inform similar future efforts. Importantly, we find two sets of results indicative of greater effectiveness in areas with plausibly stronger social ties relating to clean energy. First, we find that our treatment is much more effective in rural areas than larger population centers, consistent with the hypothesis of stronger social networks in smaller communities. Second, we find evidence that our intervention is more successful in communities with recent energy-related social campaigns.

Our work falls within a body of recent literature that uses field experiments to explore the impact of informational treatments on audit uptake. The treatments in Fowlie, Greenstone, and Wolfram (2015) and Allcott and Greenstone (2017) involve provision of direct information about the program (and its benefits) to potential participants.[1] In terms of employing information provision as a nudge by invoking social comparisons, our study is the closest to Holladay et al. (2016), who conduct a field experiment in a medium-sized metropolitan statistical area in the southeastern United States and find that social comparisons on electricity consumption tend to be associated with higher participation in an energy audit program. Our findings further support the use of social-norm based messages to encourage audit completion.

A primary contribution of our paper arises from the stage in the audit process that we examine, which is an important stage and different than the ones previously explored in the literature. We provide evidence on the cost-effectiveness of nudges targeting individuals who have already initiated this process by signing up for an audit. In contrast, previous work on audits (e.g., Fowlie, Greenstone, and Wolfram 2015; Holladay et al. 2016) has found relatively high costs of unsolicited nudges. This suggests that the timing of behavioral interventions can be an important factor to consider in policy-making decisions. It further suggests that such interventions may be more successful by focusing on those who are already "primed" to respond (e.g., have already scheduled an audit) than those who are not.

Another important contribution above the previous literature stems from our focus on a large study area with a correspondingly heterogeneous sample of households. In particular, because our experiment draws from nearly the entire state of Connecticut, we are able to exploit the heterogeneity in location and individual characteristics in order to determine how different households respond to the information provided. This allows us to test several hypotheses about heterogeneous treatment effects. The resultant analysis carries substantial managerial and policy implications, as it offers more precise guidance on targeting the treatment in a manner than reduces costs and maximizes effectiveness.

The rest of this paper is structured as follows. Section 2 provides background on nudges and social norms, and their influence on consumer decisions. Section 3 describes the field

[1] In addition, Allcott and Greenstone (2017) also test the effect of behavioral and financial treatments.

experiment, the data used in our analysis, and the empirical specification. Section 4 presents the results and discusses their implications. Finally, Section 5 concludes.

2 Background: Nudges and Social Norms

Social norms represent individuals' beliefs about what others within the social group approve of and do. Psychology recognizes two types of social norms: injunctive and descriptive. While injunctive norms refer to the beliefs of what others typically approve of (e.g., "you should donate"), descriptive norms reflect the beliefs of what others typically do (e.g., "many people donate"). From classic early social psychological studies (Sherif 1937) to more recent works by Robert Cialdini and colleagues, research has shown that descriptive norms can be an especially powerful persuasive tool by influencing an individual to act in a way that is consistent with others in her social group.

Economic studies have tested the role of descriptive social norms in a variety of contexts, including voting (Gerber and Rogers 2009), retirement savings (Beshears et al. 2015), charitable giving (Frey and Meier 2004; Shang and Croson 2009), and employee effort (Bandiera, Barankay, and Rasul 2006). Recently, a rather substantial environmental economics literature has emerged with a focus on the use of social norms in the context of resource conservation. A series of field experiments have demonstrated that well-crafted messages offering peer comparisons–often referred to as "nudges" (Thaler and Sunstein 2008)–can reduce household consumption of electricity (Schultz et al. 2007; Nolan et al. 2008; Allcott 2011; Allcott and Rogers 2014) and water (Ferraro, Miranda, and Price 2011; Ferraro and Miranda 2013; Ferraro and Price 2013; Brent, Cook, and Olsen 2015). The above findings suggest that information campaigns built around social norms can be a potentially effective low-cost policy alternative to price-based policies.

Social norms have also been examined in a context, similar to ours, where an individual makes a decision about whether to honor a previous commitment. For instance, evidence suggests that peer pressure can serve as a commitment device in reaching a shared but individual goal (e.g., Kast, Meier, and Pomeranz 2014). Clearly, a similar situation would arise in the case of energy audits, if the goal is framed in terms of helping the environment. More broadly, a meta-analysis by Lokhorst et al. (2013) recognizes two types of social norm-based pressures to commit: external, due to fear of negative social reaction if one reneges on their commitment, and internal, due to self-directed moral obligation to the social group. Other channels through which social factors have been shown in the literature to influence commitment in sequential decisions include invoking one's sense of self-presentation (e.g., Sleesman et al. 2012), combining a reminder of one's commitment at an earlier decision stage with information about the fact that many others have followed through in a similar situation (e.g., Burger 1999), and demonstrating through social comparisons how much others have gained after committing (e.g., Hoelzl and Loewenstein 2005).

In view of this evidence, we design an informational treatment that incorporates descriptive social norms. More specifically, we provide information to households with upcoming audit visits about the total monetary savings and environmental benefits that have accrued to other residents in the community who already participated in home audits. Based on the above findings, we expect that a combination of behavioral factors–from social pressure to sense of self-presentation in the community–should result in a positive effect of this nudge on the probability of treated households completing their scheduled audits.

3 Experiment and Data

3.1 Experimental Design

The randomized field experiment employed in our analysis is coordinated with NSL. In 2013, NSL began a community outreach campaign in CT, which was implemented through the Energize CT Home Energy Solutions (HES) program, funded by ratepayers and administered by the local electric utility. Applicants to HES receive a home audit at a subsidized rate of

$75 for electric- or gas-heated homes and $99 for oil-heated homes, with generous rebates available for recommended efficiency upgrades. NSL outreach targets customers through a variety of strategies, including website, email/phone, events, door-to-door canvassing, and neighbor/peer referral. Following initial contact, the next stage of NSL's customer acquisition process involves a home occupant scheduling a home audit for a particular date, usually about three weeks from the date of scheduling. Once a customer schedules an audit, NSL refers to them as a "lead." The focus of our study is on these leads.

The experiment takes place between July 13, 2014 and September 21, 2014. During this time, NSL was hosting informational events and actively canvassing for leads. At the beginning of each week during this 71-day period, NSL compiled a list of leads with scheduled audits between 16 and 22 days from that date. Each of these weekly lists can be thought of as a "cohort" of leads. There are a total of 10 cohorts in our experiment. For each cohort, the leads were randomly assigned into a control group or treatment group.[2] Thus, the experimental unit in the study is a lead.

The goal of the intervention is to convert the leads into HES audits. The treatment is a personalized notecard, mailed to the individual's address 14 days prior to the scheduled assessment visit.[3] Below is the language included on the card:

"*Dear* [customer name]*,*

Thank you for taking the time to chat with me at the [event] *on* [date]*. After a cold winter and January's rate hike, it seems like interest in the Home Energy Solutions program has really picked up.*

I checked after the event, and over [number of residents in lead's town] *residents have participated so far this year, reducing the city's carbon footprint by* [...] *tons of carbon and saving families over $*[...] *annually.*

[2] The random assignment is performed using the random number generator in Excel. For a small number of leads in 9 ZIP codes, we identified a clerical error in the randomization, and thus we drop these from our analysis. Our results are robust to the inclusion of these ZIP codes.

[3] NSL found that many cancellations occurred within 10 days of the scheduled visit, so the intention was for the notecard to arrive just before this "high cancellation period."

Thanks for being one of these energy savers! I'm excited to hear from my colleagues how your visit goes on [date] *at* [time].

Sincerely,

[NSL staff member who recruited the lead]"

Three key elements are present in the above language. First, a reference is made to the number of residents from the lead's community who participated in the program from the beginning of 2014, and the total monetary and environmental benefits for the entire community associated with those residents' participation in the program. This is the part of the notecard that invokes social norms and information about actual savings. Second, a reminder of the exact date and time of the visit is included, thus making the upcoming audit more salient. While all households in the experiment received an email reminder and a phone call prior to the audit, it is possible that the reminder at the end of the note further increased the salience of the upcoming audit. In addition, the language "excited to hear [...] how your visit goes" also contributes to the salience by prompting the customer to create a more concrete expectation or even mental visualization of the audit process.[4] Lastly, the notecard addresses the lead by name and is signed by the same employee who established the initial contact with the lead. Behavioral studies (e.g., Garner 2005) have shown that a personal touch, which serves as evidence that someone has invested their time to personalize the request, is likely to elicit a stronger reciprocal feeling of obligation to comply with the request. In our experiment, we expect this to increase the likelihood that the customer would read the note and follow through with the audit. Through the notecards, we therefore test the combined effect of three elements: social norms, salience, and personal touch.[5]

This leads to a number of hypotheses, which are tested in Section 4.

Hypothesis 1. *The treatment increases the probability of audit completion.*

Based on the discussion in Section 2, the social comparison would be expected to impact

[4]We thank an anonymous referee for making this point.

[5]We are unable to disentangle these three effects, which would require additional experimental arms and would substantially reduce statistical power.

the probability of audit uptake positively. Our intuition is that the reminder at the end of the note and the personal touch are also likely to have a positive effect on the probability of completing the visit, implying that all three effects of our treatment work in the same direction.

In addition, we expect the impact of the notecard to vary across households and communities, with some more receptive than others. We therefore examine the following three hypotheses about the heterogeneity of the treatment effect.

Hypothesis 2a. *Households with liberal political views are more likely to respond to the treatment.*

This hypothesis stems from work by Costa and Kahn (2013), who show that social norm-based energy conservation nudges are more effective among voters registered as Democrats. We would expect a similar finding in our setting.

Hypothesis 2b. *The treatment is less effective for low-income households.*

Since audits are not administered free of charge, one might expect that lower-income households face liquidity constraints that make them less receptive to the treatment. This is in line with our earlier discussion of audit costs acting as a barrier to uptake.

Hypothesis 2c. *The treatment is more effective in communities that experienced previous campaigns designed to build social ties relating to clean energy.*

This hypothesis stems from evidence that members of close-knit groups that are characterized by more intense within-group interactions and stronger social ties are more likely to exhibit pro-social behavior (e.g., Goette, Huffman, and Meier 2012) and are thus more prone to respond to a social norm-based message.

3.2 Data

The data used in our analysis include all households with scheduled audit visits (i.e., leads) between July 13, 2014 and September 21, 2014. As shown in the ZIP code map in Figure 1, our sample is quite heterogeneous with regards to location. It spans a number of different areas across CT, with the southern and central part of the state being represented most heavily. For each of the 323 households in the sample, NSL provides us with information on the street address, type of home heating, name of staff member who recruited the lead, date when the notecard was mailed, scheduled visit date, visit outcome, and subsequent visit date(s) and outcome(s) for households who rescheduled the audit.

We draw individual-level political party affiliation data from the Office of CT's Secretary of the State (`http://www.ct.gov/sots`). We then match these data to our main dataset using the individual addresses of the leads. We are able to match 298 out of the 323 addresses in our sample and thus obtain the proportion of Democratic and Republic voters within each of these households. For the remaining addresses in the sample, we use ZIP-level voter registration averages. We use this information to construct a household-level binary indicator variable for each party that takes a value of one if at least half of the members of the household support the respective party.

Finally, we also obtain detailed demographic and socioeconomic data from Acxiom. This includes household size and income, home value, as well as information about the age, race, and education of the head of household. Detailed data are available at the household level for 47 addresses in the sample and at the ZIP code level for the remaining addresses.

Table 1 presents descriptive statistics for the households in our sample stratified by whether the household is treated or not. Quick inspection of these numbers suggests that households that receive the notecard are more likely to complete a visit and take less time to do so–our first suggestive evidence of an effect of the treatment.[6] Looking across the other available observables, it appears that the two groups are roughly similar. There are minor differences: untreated households are more likely to occupy oil-heated homes (associated with a higher audit price), treated homes tend to have higher percentage of Democratic voters and lower percentage of Republican voters in the household, and untreated individuals are

[6]The relevant outcome in our experiment is the audit uptake rate. In Online Appendix A, we provide more detail about the way that an audit is completed by different households in our sample.

more likely to be of white/Caucasian origin, are slightly better educated, earn higher income, and occupy more expensive residences. If any of the above minor differences in household characteristics between the treated and control groups turn out to be systematic, this could cast doubt on the proper random assignment of the treatment. To confirm that this is not an issue, Table 2 examines the balance of covariates across the treated and untreated households in our estimation sample. The statistics suggest no statistically significant differences between the two groups. Similarly, the normalized differences do not exceed 0.13 in absolute value. This reassures us that the randomization was successfully performed.

It is also informative to compare the household characteristics in our data to the average energy user in Connecticut. Note that we draw from a sub-population of households that have already made an initial decision to schedule energy audits. We now utilize the demographic, housing, socioeconomic, and voting data from our sample to determine the extent to which this sub-population matches the average household in the state. As shown in Table 3, our sample appears to draw more heavily from Democratic voters and from older and higher-income populations relative to the state average. Oil heating and lower home values are also more prevalent in our data. On the other hand, our sample is quite close to Connecticut's average household profile in terms of fraction of Republican voters, ethnic composition, household size, and education.

The above discussion about our data is focused on the randomization, but it misses a key feature of our empirical setting: the timing at which different households enter and exit. The relevant measure of the success of our treatment is the probability of completing an audit, which in turn depends on the size of the treatment and control groups. Recall that cohorts of households become treated in a staggered fashion, and once a household completes an audit, it is no longer appropriate to include them in either of the two groups. Thus, the treatment and control groups evolve over time, making the usual simple comparison of means between the two groups over the entire time period an unsuitable approach.

Instead, a more precise measure of the treatment effect compares audit completion rates across the two groups at different points in time. In this approach, timing becomes an important factor. For example, we would not want to compare control households that still have the possibility of completing an audit to treated households that have already completed theirs, or vice versa. The former would understate, and the latter overstate, the uptake rate in the treatment group. Similarly, we would not want to compare control and treatment households before the treatment households receive the notecard.

Due to the evolution of the treatment and control groups over time, we will first present some simple descriptive results directly from our data that begin to build evidence, and then construct a panel dataset for estimating causal effects. The panel will have the unit of observation as the household-day. Adding this time dimension allows us to nonparametrically account for the evolution of the treatment and control groups over time, which can be thought of as a sample selection problem.

Finally, although we know the exact date when a notecard is mailed by NSL, for our panel dataset we need to make an assumption about the date it was received and read by the household. In our baseline analysis, we assume that it takes 3 days for the notecard to reach the respective customer.[7]

3.3 Descriptive Evidence

A quick look at the data already provides descriptive evidence of a treatment effect. Recall that households become leads and enter the sample when they first schedule an audit and exit the sample when they have completed their audit. Table 4 divides up our sample by week to both show the evolution of the treatment and control group and to show the audit uptake in the raw data. A clear finding is that the percentage uptake for treated households is substantially higher than the uptake for the untreated households. For some weeks, such as the weeks of August 11-August 17 and August 18-August 24, the differences in the audit

[7]This assumption is based on the estimated delivery window for first-class mail by the U.S. postal service, which typically ranges between 1 and 3 days for local mail. See `http://www.usps.com/ship/mail-shipping-services.htm`. We also run robustness checks varying this time window.

uptake are dramatic (e.g., 15% compared to 2%).

Figure 2 graphically illustrates the daily difference in uptake over time in the raw data. It displays the percentage of customers in each group that complete an audit on a given day over the study period. With the exception of a few brief periods in July and at the beginning and end of August, the "notecard" group clearly dominates the "no notecard" group in terms of its audit uptake rate. Hence, the raw data suggest, both on average and within any given week of the study period, a positive effect of the treatment on the probability of completing an audit.

We perform a further set of descriptive analyses that do not address the timing issue we discussed above. First, we simply examine the difference in the mean uptake rate between the treatment and control households over the entire time period (recall Table 1). The uptake rate for the treatment is 31% and for the control is 29%. The difference is in the expected direction, but is not statistically significant (p-value for a one-sided test is 0.34). Performing a similar cross-sectional analysis of uptake, in which we include additional control variables to address possible confounding effects and improve precision of the results, yields an estimated treatment effect of 7.3 percentage points (p-value for a one-sided test of 0.09).[8] We interpret these results with caution due to the timing issues leading to sample selection concerns.

We can go one step further by stratifying the sample by cohort, focusing only on the treatment time period (i.e., the period after receiving the notecard), and examining the uptake rate within each resultant sub-sample. This approach avoids comparing treated households to controls that have completed their visits prior to the cohort's treatment date. Yet, the approach is still problematic in that it does not account for the exit of households upon completion of an audit, thus implying that we would still be comparing leads within the time window considered that have no possibility of receiving an audit to ones that do. Furthermore, by stratifying by cohort, the sample becomes much smaller, leading to

[8] The control variables we include are all household characteristics from Table 2, along with a fixed effect for the NSL employee who recruited the lead.

considerable loss of statistical power. What we find in these cohort-level comparisons of treatment and control is that all but one show a positive treatment effect, and four of the positive effects are statistically significant. The values range from a negative (and not statistically significant) 1.2 percentage point effect of treatment on cohort uptake to a positive effect of 22 percentage points (significant at the 5% level).[9]

These simple analyses set the stage for our causal empirical analysis presented in the next section.

3.4 Empirical Specification

We specify a discrete-time hazard model, in which the underlying hazard rate is a flexible function of duration. Hazard models are a class of survival models commonly used in econometric studies when the dependent variable is the probability that an absorbing state occurs (e.g., an audit occurs, leading a household to exit the sample).[10] We specify a model with time-varying covariates and fixed effects to control for unobserved heterogeneity in the timing of the scheduled audit. In our experiment, households become treated at different times. To accurately identify the causal treatment effect in such a setting, we employ a staggered difference-in-differences approach, in which the pre-treatment and treatment periods are captured by time dummies (e.g., Stevenson and Wolfers 2006).

The outcome variable is the probability of completing an audit P_{izst}, which we specify as the following linear function:[11]

$$P_{izst} = \mu(t) + \alpha T_i + \beta PT_{it} + h_z + v_s + w_t + d_t + \epsilon_{izst}, \tag{1}$$

[9]Detailed results are presented in Appendix Table A1. They are again to be interpreted with caution due to the possible sample selection problem discussed above.

[10]For examples of environmental economics studies that employ hazard models, see Kerr and Newell (2003), Snyder, Miller, and Stavins (2003), and Lovely and Popp (2011).

[11]We specify our model as a linear probability model (LPM) rather than logit or probit due to the computational advantage of the LPM, which is a key factor given the large number of fixed effects in our analysis. For other similar studies that implement LPM to estimate a discrete-time hazard model, see Currie and Neidell (2005), Currie, Neidell, and Schmieder (2009), and Knittel, Miller, and Sanders (2016).

where i indexes the customer, z indexes the ZIP code, s indexes the staff member who recruited this customer, and t indexes the date. As discussed in the previous section, households enter the sample after they become leads (or after the start date of the experiment, if they were leads prior to that) and leave the sample if they complete an audit. Following the standard modeling approach for discrete-time hazard models (e.g., Currie and Neidell 2005; Currie, Neidell, and Schmieder 2009; Knittel, Miller, and Sanders 2016), we include a function $\mu(t)$ in (1) to approximate the baseline hazard rate as a measure of duration dependence. In our analysis, we specify $\mu(t)$ as a polynomial in the number of days since the date the household joined the sample. In addition, the dummy variable T_i indicates whether the household belongs to the treatment group. Our key covariate is PT_{it}, which takes a value of one for the period after the treated household receives the notecard. The coefficient of interest β measures the effect of treatment on the probability of completing an audit on a given date and is identified through a difference-in-differences approach. As we cannot verify whether a treated household has opened and read the notecard, we interpret the estimated effect as an intent-to-treat estimate. Our empirical specification also controls for a wide range of potential unobservables by including ZIP code, recruiting staff member, calendar week, and day-of-week fixed effects, denoted by h_z, v_s, w_t, and d_t, respectively. Finally, ϵ_{izst} is the idiosyncratic error term.

4 Results

4.1 Baseline Results

The model, specified in equation (1), is estimated using ordinary least squares, with the dependent variable set equal to 1 on the day the customer completes an audit and 0 otherwise. We model the time trend as a quadratic polynomial and later perform robustness checks with other functional forms. Table 5 presents the results from a number of alternative specifications of the empirical model. The estimated treatment coefficient β is statistically significant in all cases, which is consistent with Hypothesis 1. The estimates are also very similar across the specifications, as expected in a properly randomized field experiment.[12]

Columns (1) and (2) in Table 5 show the results without controlling for time-invariant unobservable factors. In principle, given that the randomization process in this experiment is cross-sectional and at the individual household level, the estimated treatment effect should not be biased by time-invariant heterogeneity. Nonetheless, it is possible that factors, such as NSL employee skill and effort, which we do not randomize across, confound our results. Hence, we include staff member fixed effects in our regression. We also add ZIP code fixed effects to flexibly control for any potential differences in the composition of households across regions. As shown in column (3), introducing these two sets of fixed effects leads to a slightly lower estimate of the treatment effect. In addition, given that audits are priced differently depending on heating type, we estimate another specification, in which we add an indicator for oil-heated home. The result from this specification, shown in column (4), is identical to our estimate in column (3), likely due to the randomization and the lack of substantial variation in heating type within ZIP codes. Lastly, we also examine two additional model specifications, in which we control more flexibly for time-varying unobserved effects. In the first specification, we replace calendar week with date fixed effects to account for potential day-to-day unobserved factors. As shown in column (5), this does little to change the estimated treatment effect, indicating that calendar week fixed effects are quite effective at capturing all location-invariant unobserved heterogeneity. Alternatively, we can also control for unobserved factors that vary both across time and region. We do so by replacing the ZIP code and calendar week fixed effects with ZIP-month fixed effects.[13] As shown in column (6), our estimates are still very similar, which is not surprising: given the relatively short time horizon and conditional on controlling for week-specific effects, it is unlikely that the temporal variation in region-specific unobservables would be significant.

[12]In addition, as shown in Appendix Table B1, the treatment coefficient estimate is very robust to excluding the duration dependence polynomial function.

[13]Because of the large number of ZIP code-calendar week interaction terms (more than 1,000), we are not able to use ZIP-week fixed effects.

Our preferred model specification is the one presented in column (4) of Table 5. The estimate in this specification implies that, conditional on a household not having completed its audit until the current day, the probability of completing it on this day is 1.1 percentage points higher if the household has received a notecard. We now utilize our data and the regression output to calculate the effect of the treatment on the overall uptake during the study period. For every household i in the data, we first compute the predicted daily probabilities from the regression by setting $T_i = 0$ and $PT_{it} = 0$ for each day t the household remains in the sample. We then use these daily probabilities to obtain the probability of completing an audit at some point during the study period.[14] We repeat the same procedure, this time setting $T_i = 1$, $PT_{it} = 0$ for t up to 11 days prior to its first scheduled visit (i.e., the date the notecard would have been received if treated), and $PT_{it} = 1$ for all remaining t. For each household, we thus have the predicted uptake for both states of the world: one in which they are part of the control and one in which they are in the treatment group. Subtracting the former figure from the latter provides a measure of the treatment effect. For the average household in our sample, this difference is 0.065, implying a 6.5 percentage points higher overall uptake for an average treated lead in this experiment.

4.2 Robustness Checks

We conduct a number of robustness checks, shown in Table 6. Our baseline estimate is 1.1 percentage points, which was obtained in column (4) of Table 5. First, we test the sensitivity of our treatment effect estimate to the functional form of the duration dependence function. As seen from the results in specifications I and II, our estimate is very robust to the use of a higher-degree polynomial.[15] In addition, as discussed earlier, due to the lack of information about the precise date when the notecard is received and read by its recipient, our baseline

[14]The probability of completing the audit during any T-day period is $1 - \pi(T)$, where π denotes the probability of *not* completing an audit on any of these days. Let p_t denote the probability of completing an audit on a given day t, conditional on not having completed the audit until then. Then, $\pi = \prod_{t=1}^{T}(1 - p_t)$.

[15]In Online Appendix B, we also show that modeling the duration dependence function as a series of splines yields very similar results.

specification assumes a 3-day time window between mailing and receipt of the notecard. As a robustness check, we extend this window to 6 days and find a treatment effect of 1.4 percentage points, which is close to our main result. As a further test, we make the assumption that the cards are received and read within a day after being mailed. Even in this case, the estimated treatment effect of 0.9 percentage points is still very similar to our baseline estimate.

We also perform two robustness checks in which we include additional control variables in our regression. In specification V, we add all voting, demographic, and socioeconomic variables from Table 1 as controls. While the random assignment of the treatment should eliminate any bias resulting from individual-specific heterogeneity (which is even further alleviated by the inclusion of ZIP-specific fixed effects), it is comforting to note that, even after explicitly controlling for those household characteristics in specification V, our treatment effect estimate remains almost unchanged. This is also consistent with our findings in Table 2, which indicated a very good balance across the control and treatment group.

Lastly, we estimate a specification in which we include a "lead-to-visit" control variable to measure the number of days between the date a lead is recruited (i.e., enters the sample) and their initial scheduled visit. This provides an additional control over the timing of the treatment, which would depend on the length of the "lead-to-visit" window. As shown in specification VI, even after conditioning on the "lead-to-visit" measure, our estimated treatment effect is very similar to the baseline result.

4.3 Cost-effectiveness of the Treatment

We combine the estimated treatment effect with cost data from NSL in order to gain insight into the cost-effectiveness of our intervention. NSL staff members were paid a total of $288 for their time over the course of the experiment. With 161 notecards mailed, this translates into approximately $1.79 per card. After adding in the costs of the notecard ($0.12) and postage ($0.49), total NSL expenses per treated household amount to $2.40. Hence, our estimate of a 6.5 percentage points increase in uptake due to the treatment implies a cost of $36.90 for securing an additional audit through this intervention.

Through back-of-the-envelope calculations based on estimates from the literature and average figures for Connecticut, we can obtain a rough value of the environmental benefits from an audit. A meta-analysis of experimental studies by Delmas, Fischlein, and Asensio (2013) derives a central estimate for the household's post-audit energy savings of approximately 5 percent. By combining residential energy use data from the State Energy Data System (`https://www.eia.gov/state/seds`) with demographic data from the 2010-2014 wave of the American Community Survey, we obtain a value for the annual energy consumption of an average CT household of 53,810 kilowatt-hours (kWh). This implies average annual energy savings of 2,690 kWh from a completed audit. Converting these into lifetime savings requires assumptions about the household's long-term energy use patterns, appliance lifespan and replacement decisions, and housing unit maintenance and length of occupancy. We use a conservative estimate of 5 years of sustained energy savings, based on information about wear and tear of homes and average life expectancy of certain energy-using durables by InterNACHI (`https://www.nachi.org/life-expectancy.htm`).[16] Finally, employing the estimate of average emissions per kWh in the Northeast by Graff Zivin, Kotchen, and Mansur (2014), we find that the lifetime energy savings from a completed audit result in avoiding a total of 3.47 tons of carbon dioxide emissions.

The above estimates imply that the cost-effectiveness of our treatment in terms of carbon dioxide emissions reduction benefits is $10.63/tCO2.[17] Note that that this estimate is lower than the IAWG (2013) central value of the social cost of carbon of $42/tCO2 (in 2014 dollars) and also much lower than cost estimates for alternative carbon abatement programs (Johnson 2014) or renewable energy subsidies (Gillingham and Tsvetanov 2018) in the region. Thus, our findings are qualitatively consistent with the notion that this particular nudge is a low-

[16] Alternatively, this could be viewed as equivalent to more prolonged savings that are decreasing over time.

[17] Note this does not include environmental benefits from reduced criteria air pollutants.

cost tool to reduce emissions, although the exact cost per ton value should be interpreted with caution due to the number of assumptions involved in our calculations.

4.4 Heterogeneous Treatment Effects

We proceed to test Hypotheses 2a-c by estimating a number of additional specifications.[18] First, we examine the impact of political views on the effectiveness of our treatment. We test Hypothesis 2a with our data by augmenting the empirical model with indicators for support of the Democratic and Republican parties and interactions between these indicators and the treatment variable. The omitted group category is "independent." Recall that the Democrat and Republican indicator variables are defined as follows: each indicator variable equals 1 for a given household if at least half of the household members are registered as supporters of the respective party.[19] Due to concerns that political affiliation may be correlated with household characteristics, such as race and income, we include our complete suite of demographic and socioeconomic controls in this regression.

As shown in column (1) in Table 7, support for either of the two major parties does not have significant influence on the effectiveness of the treatment. The estimates of the coefficients on the interactions of treatment with party affiliation are not statistically significant. The average treatment effects across supporters of the two parties are very close: 1 percentage point for Democrat supporters vs. 1.1 percentage points for Republican supporters. Furthermore, a joint test fails to reject the equality of the treatment effects across the two camps, with a p-value of 0.56. Hence, in contrast to the findings of Costa and Kahn (2013), we find no evidence of heterogeneity in the treatment effects based on political ideology and thus reject Hypothesis 2a.[20]

[18]Our findings hold even after dropping some or all of the time-invariant fixed effects (Appendix C).

[19]Under this definition, 14 households in the sample are categorized as both Democrat and Republican. In Appendix C, we explore alternative definitions, such as one where the indicator variable equals 1 only if *all* household members are affiliated with the respective party. We obtain very similar results.

[20]In addition, as shown in Appendix Table C1, political views do not appear to be an important driver of audit uptake for the sample as a whole, with neither of the two coefficients on the standalone party affiliation terms in the regression being statistically significant. The two coefficients are also not significantly different

Heterogeneous treatment effects could also arise due to individual-specific factors, as suggested by Hypothesis 2b. To examine how individual heterogeneity may affect the treatment outcome, we augment our model to allow the treatment effect to vary by demographic and socioeconomic household characteristics in our data. As in the earlier analysis, we also include these variables separately as controls. In addition, because certain household characteristics may be correlated with political views, we control for political party affiliation by including the two party indicators, which are defined as before.[21] The results are presented in column (2) of Table 7. In line with Hypothesis 2b, we find evidence suggesting heterogeneity in the treatment effect across households with different income. In particular, the treatment appears to be more effective among higher-income households. A 1% increase in income for the average household in the sample results in a 0.02 percentage point boost in the treatment effect. Furthermore, we find that the treatment effect also varies by household ethnicity, with an estimated gap of 2.4 percentage points between the effect on Caucasian and non-Caucasian households.

As stated in Hypothesis 2c, we anticipate that the effectiveness of our treatment also depends on community-specific characteristics that influence the impact of the social norm-based message in the notecard. For instance, the number of households with already completed energy audits that is referenced in the notecard would likely be higher in towns with previous energy-related social campaigns. Notable examples of such campaigns include the "Solarize CT" marketing program and a similar (firm-created) "CT Solar Challenge" marketing program (Gillingham and Bollinger 2017; Gillingham and Tsvetanov 2018). These are limited-time programs with group discount pricing for residential solar photovoltaic installations in participating towns, along with extensive promotion of word-of-mouth to disseminate information about solar panels by members of the community. Hence, the presence of such community-based programs points to higher awareness of energy and environmental issues and potentially stronger social ties within the community. All of these are factors that could

from each other, with a p-value of 0.74.

[21] As shown in Appendix Table C2, dropping the political party controls leaves our results unchanged.

influence the effectiveness of our intervention. To test this, we interact the treatment variable with an indicator for past Solarize or Solar Challenge programs in the respective town. As shown in column (3) of Table 7, there is evidence of NSL leads in these towns responding more positively to the treatment, with an average boost in "solar" towns of approximately 0.9 percentage points (statistically significant at the 10% level).[22] This result is consistent with our Hypothesis 2c, suggesting that towns which have experienced campaigns that built social ties relating to clean energy are more responsive to the treatment.

Lastly, we can also test Hypothesis 2c by recognizing that the effectiveness of our social norm-based intervention is likely to hinge on the perceived connection of the household with their respective community. Recall that one of the key elements in the notecard language– "Thanks for being one of these energy savers!"–targets individuals identifying with their local community. This is more likely have an impact in smaller, closely-knit communities than larger population centers, suggesting possible disparity in the treatment effects between urban and rural communities. While there are a total of 19 municipalities in CT that are historically identified as cities, some of them have current populations below 10,000 and are less likely to represent truly urbanized areas. In order to ensure that we have a more accurate urban proxy, we designate a household as being in an urban area if it is in one of the 8 largest cities, each of which have population exceeding 70,000.[23] We then re-estimate our regression allowing for the treatment effect to vary based on this definition of urban centers. As shown in column (4) of Table 7, there appears to be a considerable gap between the effect of our treatment in urban versus rural areas.[24] Specifically, the notecard effectiveness is 1.2 percentage points higher in rural communities. As demonstrated

[22]Due to concerns about a potential correlation between town and resident characteristics, we also include household demographic, socioeconomic, and voting controls in this regression. As shown in the Appendix, dropping these controls does not affect the magnitude or significance of our estimates.

[23]These eight cities are Bridgeport, Danbury, Hartford, New Britain, New Haven, Norwalk, Stamford, and Waterbury. These cities cover just over 80% of the sample. Population figures are obtained from the 2010-2014 wave of the American Community Survey. For a complete list of CT towns and cities, see `http://portal.ct.gov`.

[24]We also examine several other possible thresholds for designating an urban area, and find that our results are quite robust to the exact threshold.

in Appendix Table C3, this result is robust to excluding demographic, socioeconomic, and voting controls from the regression. The estimates translate into an average treatment effect of only 0.1 percentage points in urban centers versus 1.3 percentage points in more rural areas. This implies potentially significant cost-savings and efficiency gains from refocusing audit recruitment and retention efforts using this nudge towards smaller communities.

5 Conclusion

Political barriers to the widespread use of price-based policy instruments for encouraging energy conservation and energy efficiency adoption have led to an increased interest in non-price interventions as an alternative policy tool. This study focuses on residential energy audits, which in theory are a promising example of such a non-price intervention, but in reality appear to face major challenges in reaching and retaining customers. We consider a particular stage of the customer acquisition process in a typical audit, namely the household choice of whether to complete an already scheduled audit visit. This stage is particularly important in that it exhibits considerable customer attrition and has not yet been examined in previous studies of energy audits.

Using a natural field experiment, we explore the role of information provision as a nudge targeted at increasing the uptake of initially scheduled audits. Specifically, we test the joint effect of social norms, increased salience, and personal touch. Our results suggest that a carefully crafted notecard which incorporates the above three effects can increase the uptake probability for an average customer with a scheduled but not yet completed audit by 1.1 percentage points on a given day. Back-of-the-envelope calculations suggest that this intervention could present a more cost-effective carbon reduction approach than price-based energy policies.

Interestingly, we find the treatment effect to be very similar across households with different political views. This suggests that, conditional on the decision to schedule an energy assessment visit, political ideology does not play a major role in the effectiveness of the treatment. On the other hand, we find heterogeneity in the treatment effect based on household- and community-specific characteristics. An important finding is that a nudge which exploits the individual's propensity to act in a way that is consistent with the social group can be much more effective in rural areas than larger population centers. This implies that targeting smaller communities in similar future interventions could substantially improve the cost-effectiveness of the treatment. We find similar results for communities that already built social ties relating to clean energy from previous campaigns.

Overall, our study finds considerable potential for the use of personalized social norm-based reminders to households as a means of encouraging follow-through on the commitment to a previously scheduled home assessment visit. This finding offers a viable marketing strategy for residential energy efficiency providers to help improve their customer acquisition and retention process. It also carries important policy implications, as it suggests a relatively low-cost tool for enhancing audit effectiveness as part of the effort to improve energy conservation to reduce emissions. Note that, from a policy perspective, the success of such intervention is assessed not only based on its impact on audit uptake, but also with regards to the ultimate outcome of the audit. For instance, Palmer, Walls, and O'Keeffe (2015) and Holladay et al. (2016) find that follow-up on audit recommendations can vary widely across participants. Hence, a promising next step would be to examine the impacts of such social norm-based informative nudges on audit follow-up and the decision to invest in energy efficiency upgrades.

References

Alberini, Anna and Charles Towe. 2015. "Information v. Energy Efficiency Incentives: Evidence from Residential Electricity Consumption in Maryland." *Energy Economics* 52 (S1):S30–S40.

Allcott, Hunt. 2011. "Social Norms and Energy Conservation." *Journal of Public Economics* 95 (9-10):1082–1095.

———. 2016. "Paternalism and Energy Efficiency: An Overview." *Annual Review of Economics* 8:145–176.

Allcott, Hunt and Michael Greenstone. 2017. "Measuring the Welfare Effects of Residential Energy Efficiency Programs." Working paper no. 2017-05, Becker Friedman Institute.

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

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2006. "The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment." *The B.E. Journal of Economic Analysis and Policy* 5 (2):1–28.

Beshears, John, James Choi, David Laibson, Brigitte Madrian, and Katherine Milkman. 2015. "The Effect of Providing Peer Information on Retirement Savings Decisions." *Journal of Finance* 70:1161–1201.

Brent, Daniel, Joseph H. Cook, and Skylar Olsen. 2015. "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.

Burger, Jerry. 1999. "The Foot-in-the-Door Compliance Procedure: A Multiple-Process Analysis and Review." *Personality and Social Psychology Review* 3 (4):303–325.

Costa, Dora and Matthew Kahn. 2013. "Energy Conservation 'Nudges' and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment." *Journal of the European Economic Association* 11 (3):680–702.

Currie, Janet and Matthew Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics* 120 (3):1003–1030.

Currie, Janet, Matthew Neidell, and Johannes Schmieder. 2009. "Air Pollution and Infant Health: Lessons from New Jersey." *Journal of Health Economics* 28 (3):688–703.

Delmas, Magali A., Miriam Fischlein, and Omar I. Asensio. 2013. "Information Strategies and Energy Conservation Behavior: A Meta-analysis of Experimental Studies from 1975 to 2012." *Energy Policy* 61:729–739.

Ferraro, Paul and Juan José Miranda. 2013. "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, Paul, Juan José Miranda, and Michael Price. 2011. "The Persistence of Treatment Effects with Norm-Based Policy Instruments : Evidence from a Randomized Environmental Policy Experiment." *American Economic Review* 101 (3):318–322.

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

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2015. "Are the Non-Monetary Costs of Energy Efficiency Investments Large? Understanding Low Take-up of a Free Energy Efficiency Program." *American Economic Review: Papers and Proceedings* 105 (5):201–204.

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

Garner, Randy. 2005. "Post-It® Note Persuasion: A Sticky Influence." *Journal of Consumer Psychology* 15 (3):230–237.

Gerber, Alan and Todd Rogers. 2009. "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You." *Journal of Politics* 71 (1):178–191.

Gillingham, Kenneth and Bryan Bollinger. 2017. "Social Learning and Solar Photovolatic Adoption: Evidence from a Field Experiment." Working paper, Yale University.

Gillingham, Kenneth and Karen Palmer. 2014. "Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Evidence." *Review of Environmental Economics and Policy* 8 (1):18–38.

Gillingham, Kenneth and Tsvetan Tsvetanov. 2018. "Hurdles and Steps: Estimating Demand for Solar Photovoltaics." USAEE working paper no. 18-329.

Goette, Lorenz, David Huffman, and Stephan Meier. 2012. "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4 (1):101–115.

Graff Zivin, Joshua, Matthew J. Kotchen, and Erin T. Mansur. 2014. "Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies." *Journal of Economic Behavior and Organization* 107A:248–268.

Hoelzl, Erik and George Loewenstein. 2005. "Wearing out Your Shoes to Prevent Someone Else from Stepping into Them: Anticipated Regret and Social Takeover in Sequential Decisions." *Organizational Behavior and Human Decision Processes* 98 (1):15–27.

Holladay, J. Scott, Jacob LaRiviere, David Novgorodsky, and Michael Price. 2016. "Asymmetric Effects of Non-Pecuniary Signals on Search and Purchase Behavior for Energy-Efficient Durable Goods." NBER working paper no. 22939.

IAWG. 2013. *Inter-Agency Working Group on Social Cost of Carbon, Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866*. Washington, DC.

Johnson, Erik. 2014. "The Cost of Carbon Dioxide Abatement from State Renewable Portfolio Standards." *Resource and Energy Economics* 36 (2):332–350.

Kast, Felipe, Stephan Meier, and Dina Pomeranz. 2014. "Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device." Working paper no. 12-060, Harvard University.

Kerr, Suzi and Richard G. Newell. 2003. "Policy-Induced Technology Adoption: Evidence from the US Lead Phasedown." *The Journal of Industrial Economics* 51 (3):317–343.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. "Caution Drivers! Children Present: Traffic, Pollution and Infant Health." *Review of Economics and Statistics* 98 (2):350–366.

Lokhorst, Anne Marike, Carol Werner, Henk Staats, Eric van Dijk, and Jeff L. Gale. 2013. "Commitment and Behavior Change: A Meta-Analysis and Critical Review of Commitment-Making Strategies in Environmental Research." *Environment and Behavior* 45 (1):3–34.

Lovely, Mary and David Popp. 2011. "Trade, Technology, and the Environment: Does Access to Technology Promote Environmental Regulation?" *Journal of Environmental Economics and Management* 61 (1):16–35.

Nolan, Jessica, Wesley Schultz, Robert Cialdini, Noah Goldstein, and Vladas Griskevicius. 2008. "Normative Social Influence is Underdetected." *Personality and Social Psychology Bulletin* 34 (7):919–923.

Palmer, Karen and Margaret Walls. 2015. "Limited Attention and the Residential Energy Efficiency Gap." *American Economic Review: Papers and Proceedings* 105 (5):192–195.

Palmer, Karen, Margaret Walls, and Lucy O'Keeffe. 2015. "Putting Information into Action: What Explains Follow-up on Home Energy Audits?" Discussion paper 15-34, Resources for the Future.

Schultz, Wesley, Jessica Nolan, Robert Cialdini, Noah Goldstein, and Vladas Griskevicius. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science* 18 (5):429–434.

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

Sherif, Muzafer. 1937. "An Experimental Approach to the Study of Attitudes." *Sociometry* 1:90–98.

Sleesman, Dustin J., Donald E. Conlon, Gerry McNamara, and Jonathan E. Miles. 2012. "Cleaning up the Big Muddy: A Meta-Analysis Review of the Determinants of Escalation of Commitment." *Academy of Management Journal* 55 (3):541–562.

Snyder, Lori D., Nolan H. Miller, and Robert N. Stavins. 2003. "The Effects of Environmental Regulation on Technology diffusion: The case of Chlorine Manufacturing." *American Economic Review* 93 (2):431–435.

Stevenson, Betsey and Justin Wolfers. 2006. "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress." *The Quarterly Journal of Economics* 121 (1):267–288.

Thaler, Richard and Cass Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness.* Yale University Press.

U.S. EIA. 2016. *Monthly Energy Review: November 2016.* U.S. Department of Energy: Washington, DC.

Figure 1: Distribution of Sample Across CT

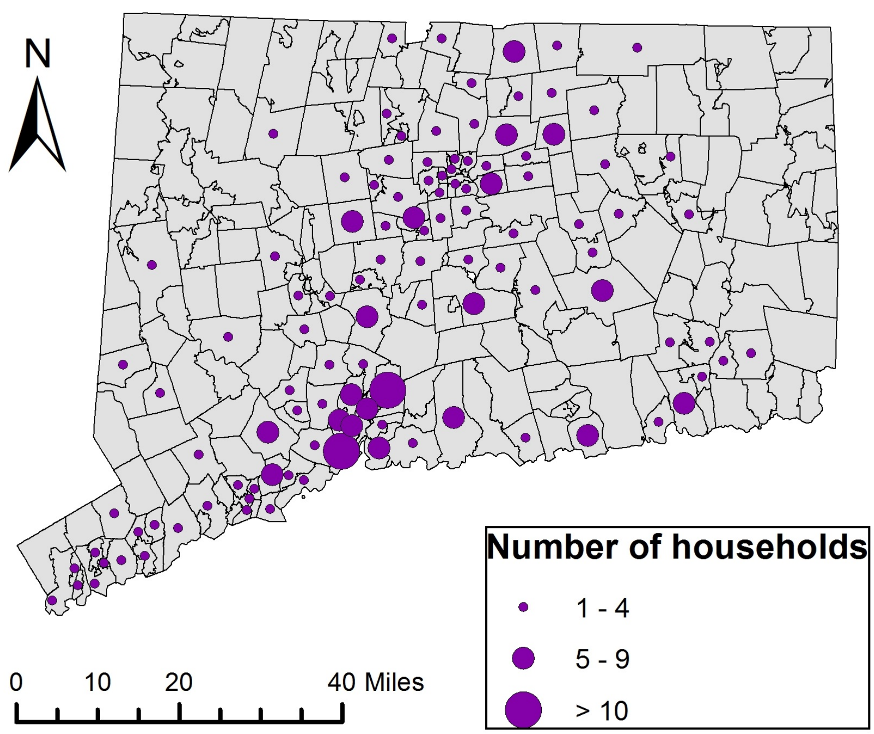


Figure 2: Total Daily Audits as Percentage of Customers in Each Group

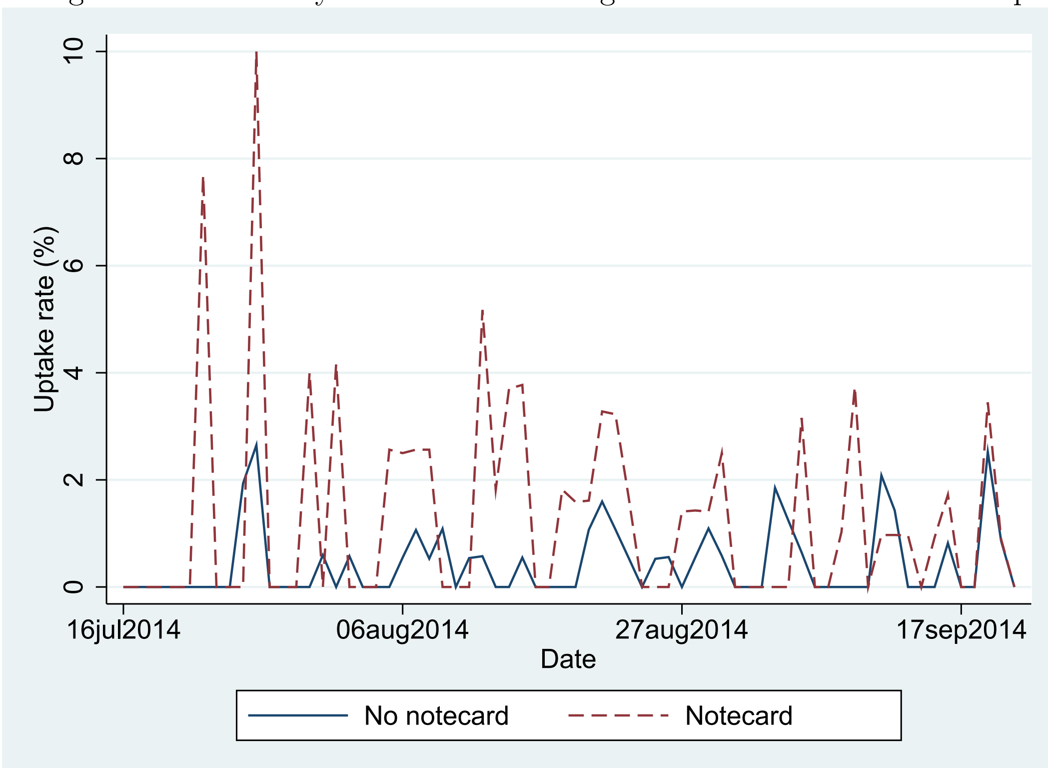


Table 1: Descriptive Statistics

	Treated households					Untreated households				
Variable	Obs	Mean	Std. dev.	Min	Max	Obs	Mean	Std. dev.	Min	Max
Completed audit	161	0.311	0.464	0	1	162	0.29	0.455	0	1
Days to complete audit	161	46.594	18.668	10	71	162	48.239	19.191	13	71
Oil-heated home	161	0.571	0.496	0	1	162	0.617	0.488	0	1
% Democrat voters	161	41.846	40.922	0	100	162	39.581	39.328	0	100
% Republican voters	161	16.29	30.408	0	100	162	20.786	33.141	0	100
Householder age	161	46.832	9.554	20	80	162	46.785	9.264	26	90
White/Caucasian	161	0.797	0.194	0	1	162	0.82	0.182	0	1
Household size	161	2.392	0.771	1	8	162	2.395	0.69	1	7
College degree or higher	161	0.34	0.218	0	1	162	0.379	0.225	0	1
Home value ($100,000's)	161	3.022	2.706	0.521	23.435	162	3.313	3.491	0.658	27.529
Income ($1,000,000's)	161	0.092	0.056	0.008	0.2	162	0.099	0.051	0.008	0.2

Note: All statistics are presented at the household level. "Completed audit," "oil-heated home," "college degree or higher," and "white/Caucasian" are dummy indicator variables. The "days to complete audit" variable is top-censored for all households who do not complete their visit by the end of the 71-day study period.

Table 2: Balance of Covatiates

Variable	Treated	Control	Difference	p-value	Norm. difference
Oil-heated home	0.571	0.617	-0.046	0.403	-0.066
% Democrat voters	0.419	0.396	0.023	0.612	0.04
% Republican voters	0.163	0.208	-0.045	0.205	-0.1
Householder age	46.832	46.785	0.047	0.964	0.004
White/Caucasian	0.797	0.82	-0.023	0.267	0.087
Household size	2.392	2.395	-0.003	0.976	-0.002
College degree or higher	0.34	0.379	-0.04	0.107	-0.127
Home value (\$100,000's)	3.022	3.313	-0.291	0.403	-0.066
Income (\$1,000,000's)	0.092	0.099	-0.007	0.264	-0.088

Note: All statistics are presented at the household level. There are a total of 161 treated households and 162 untreated households in the sample. The columns "Treated" and "Control" display the sample means for the two groups. The columns "Difference" and "p-value" show the differences in group means across treated and untreated households and the corresponding p-values from a means t-test of these differences, respectively. The column "Norm. difference" shows the normalized differences, calculated as the difference in means, normalized by the standard deviation of the covariates, i.e., $\frac{\bar{x}_T - \bar{x}_U}{\sqrt{s^2_{x,T} + s^2_{x,U}}}$.

Table 3: Sample Characteristics vs. State Averages

Variable	Sample	Connecticut
% oil-heated homes	59.4	45.2
% Democrat voters	40.7	36.4
% Republican voters	18.6	20.1
Median age	46.3	40.3
% white/Caucasian	80.8	79.9
Household size	2.4	2.6
% college degree or higher	36	37
Median home value ($)	232,368	274,500
Median income ($)	86,108	69,899

Note: State voting registration data are obtained from the Office of CT's Secretary of State. All remaining data for CT are from the 2010-2014 wave of the American Community Survey.

Table 4: Treated and Untreated Households During the Study Period

Calendar week	Treated households			Untreated households		
	Households	Audits	%	Households	Audits	%
Jul 14–Jul 20	5.1	0	0	122.2	0	0
Jul 21–Jul 27	15.1	3	20	149.6	7	5
Jul 28–Aug 3	25.6	2	8	168.6	2	1
Aug 4–Aug 10	40	4	10	182.9	6	3
Aug 11–Aug 17	53.3	8	15	180.2	3	2
Aug 18–Aug 24	60.3	8	13	186.9	8	4
Aug 25–Aug 31	71.6	5	7	180	6	3
Sep 1–Sep 7	91	3	3	160.7	6	4
Sep 8–Sep 14	103.4	8	8	142.3	5	4
Sep 15–Sep 21	113.3	8	7	120.5	5	4

Note: The number of treated and untreated households shown in the second and fifth columns is an average for the respective calendar week. Because of households being added to each group over time, as well as customers exiting the sample after completing a visit (see Section 3), the average number for a given week is usually not an integer. Completed visits refers to the total number of completed audits over the course of the respective week.

Table 5: Regression Results

Variable	Specification					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.012***	0.012***	0.011***	0.011***	0.010***	0.010***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Oil heating	no	no	no	yes	yes	yes
ZIP code fixed effects	no	no	yes	yes	yes	no
Staff member fixed effects	no	no	yes	yes	yes	yes
Duration dependence function	yes	yes	yes	yes	yes	yes
Calendar week fixed effects	no	yes	yes	yes	no	no
Day-of-week fixed effects	no	yes	yes	yes	yes	yes
Date fixed effects	no	no	no	no	yes	no
ZIP-month fixed effects	no	no	no	no	no	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. Duration dependence function is a quadratic in number of days since entering the sample. Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table 6: Robustness Checks

Variable	Specification						
	Baseline	I	II	III	IV	V	VI
Treatment	0.011***	0.011***	0.011***	0.014***	0.009***	0.011***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Oil heating	yes	yes	yes	yes	yes	yes	yes
ZIP code fixed effects	yes	yes	yes	yes	yes	yes	yes
Staff member fixed effects	yes	yes	yes	yes	yes	yes	yes
Calendar week fixed effects	yes	yes	yes	yes	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes	yes	yes	yes	yes
Duration dep. polynomial degree	2	3	4	2	2	2	2
Days until notecard received	3	3	3	6	1	3	3
Household characteristics	no	no	no	no	no	yes	no
Lead-to-visit	no	no	no	no	no	no	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table 7: Heterogeneous Treatment Effects

Variable	Specification			
	(1)	(2)	(3)	(4)
Treatment	0.011**	0.005	0.007*	0.013***
	(0.005)	(0.018)	(0.004)	(0.003)
Treatment×Democrat	-0.001	-	-	-
	(0.006)			
Treatment×Republican	0.002	-	-	-
	(0.006)			
Treatment×age	-	0.0003	-	-
		(0.0003)		
Treatment×white	-	-0.024**	-	-
		(0.011)		
Treatment×household size	-	0.0002	-	-
		(0.0035)		
Treatment×college or more	-	-0.012	-	-
		(0.017)		
Treatment×home value	-	-0.0004	-	-
		(0.0012)		
Treatment×income	-	0.187**	-	-
		(0.086)		
Treatment×solar campaign	-	-	0.009*	-
			(0.005)	
Treatment×urban	-	-	-	-0.012***
				(0.004)
Oil heating	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes
Socioeconomic controls	yes	yes	yes	yes
Political view controls	yes	yes	yes	yes
ZIP code fixed effects	yes	yes	yes	yes
Staff member fixed effects	yes	yes	yes	yes
Duration dependence function	yes	yes	yes	yes
Calendar week fixed effects	yes	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. The main effects are included in the regression, but omitted above. Total sample contains 15,318 observations. Duration dependence function is a quadratic in number of days since entering the sample. "Democrat" and "Republican" are binary indicator variables equal to one for a given household if at least half of the household members are registered as supporters of the respective party. Omitted party affiliation group is "independent." Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

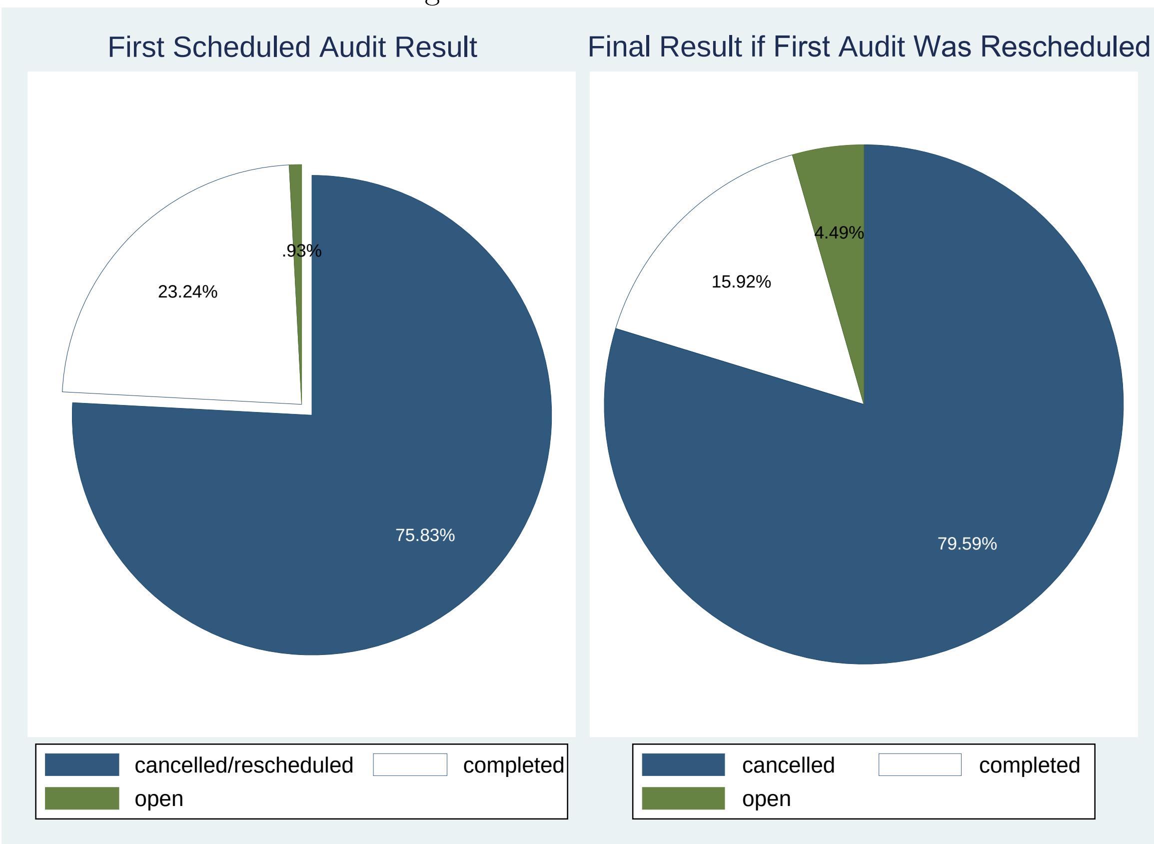
Online Appendix A: Visit Outcomes in the Sample

The focus of our analysis is on the uptake of audits for those who previously signed up for them. In this appendix section, we provide more background information about the process through which households in our sample reach the audit outcome and present some descriptive evidence of the treatment effect on uptake.

Once a household has scheduled an audit visit, there are four different possible outcomes: (i) complete the visit on the scheduled date: (ii) cancel the visit; (iii) reschedule for a later date within our study period; or (iv) reschedule for a later date outside of our study period. As shown in the left panel of Figure A1, about 23 percent of the households in our sample complete the audit on the initially scheduled date. For the cases where the audit is not completed on that date, the customer either cancelled or rescheduled. The outcome is listed as "open" if the household rescheduled for a later date that lies outside of our study period. Finally, we combine the outcomes in which the first scheduled audit is cancelled or rescheduled for a date within our study period due to multiple instances in which customers who cancelled the first scheduled audit contacted NSL to reschedule (11 percent of the cases in our data). Almost 76 percent of the outcomes of the initial visit fall within this broad category, which we then track further. As shown in the right panel of Figure A1, roughly 16 percent of these leads complete their audit on a later date within our study period.

Table A1 presents detailed results from our cohort-level comparisons of audit uptake rates. There are 10 cohorts in the sample. The time horizon for each cohort begins in the week the notecard was received by the cohort's treatment group and ends once all treated households have completed an audit or the end of the study period has been reached, depending on which date comes first. The second and third columns in Table A1 show the difference between the uptake in the treatment and control groups in each cohort (i.e., the implied average treatment effect) and the p-value from a t-test of the means across the two groups. We find statistically significant positive difference, indicating a positive treatment effect, for cohorts 4, 5, 8, and 9.

Figure A1: Audit Outcomes



<u>Note:</u> Results based on a sample of 323 households.

Table A1: Cohort-level Comparisons

Cohort	Difference	p-value
#1	0.138	0.435
#2	0.21	0.208
#3	0.031	0.812
#4	0.22	0.043
#5	0.181	0.069
#6	0.168	0.127
#7	-0.012	0.91
#8	0.168	0.029
#9	0.155	0.044
#10	0.014	0.789

Note: Each household within a weekly cohort is represented by a single binary outcome measure, which equals one if the household completed an audit during the time period considered and zero otherwise. "Difference" is calculated by subtracting the average outcome in the control group from the average outcome in the treatment group within each cohort. "p-value" displays the p-value from a test comparing the means of the cohort's treatment and control group.

Online Appendix B: The Duration Dependence Function

In this appendix section, we demonstrate that our results are robust to the way we model the duration dependence function.

Columns (1) and (2) in Table B1 display the treatment coefficient estimates without controlling for duration dependence. Column (1) presents a parsimonious specification of our model in which we do not control for any observable or unobservable factors other than the assignment of the treatment. This is equivalent to removing the duration dependence function from specification (1) in Table 5. Then, column (2) shows the treatment estimate under our baseline specification, once again excluding the duration dependence function. In both cases we obtain estimates very similar to the ones in columns (1) and (4) of Table 5, indicating that our results are robust to excluding the duration dependence function.

Furthermore, we also show that, once the duration dependence function is included in the model, our estimates are not impacted by the choice of functional form. In our baseline specification we use a polynomial time trend, and, as shown in the robustness checks in Section 4, the order of that polynomial does not affect the magnitude or significance of the estimated treatment effect. Now, in column (3) of Table B1 we instead model the duration dependence function as a linear spline function. We choose weekly break points for the series of splines, i.e., break points after days 7, 14, 21, and so on. Once again, our estimate is quite close to the baseline result of 1.1 percentage points.

Table B1: Duration Dependence

Variable	Specification		
	(1)	(2)	(3)
Treatment	0.012***	0.013***	0.010***
	(0.002)	(0.003)	(0.003)
Oil heating	no	yes	yes
ZIP code fixed effects	no	yes	yes
Staff member fixed effects	no	yes	yes
Duration dependence function	no	no	yes
Calendar week fixed effects	no	yes	yes
Day-of-week fixed effects	no	yes	yes
Duration function form	N/A	N/A	spline

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Online Appendix C: Heterogeneity of the Treatment Effect

This appendix section provides a more complete set of regression results for the heterogeneous treatment effect estimates. Further, it provides a robustness check by using a different definition of party affiliation and dropping some or all of the demographic, socioeconomic, and political controls.

In Tables C1-C3, we use two alternative definitions of the Democrat and Republican indicator variables. In column (1) of each table, a party indicator variable equals 1 for a given household, if at least half of the household members are registered as supporters of the respective party. In column (2), the indicator variable equals 1 only if *all* household members are affiliated with the respective party. As shown in Tables C1-C3, both definitions yield very similar results. We therefore proceed to use the definition from column (1) throughout our main analysis.

In particular, the output in Table C1 shows that the individual coefficients on the interactions of treatment with party affiliation are not statistically significant under both definitions. Furthermore, a joint test fails to reject the equality of the treatment effects across the two political camps, with p-value of 0.56 in column (1) and 0.31 in column (2). The coefficients on the standalone Democrat and Republican dummies are also not statistically significant under either definition. Lastly, a joint test of the equality of these two dummy variable coefficients yields p-values of 0.74 in column (1) and 0.96 in column (2).

Table C2 displays results with and without controlling for party affiliation. The first two columns use the same definitions of party affiliation as the respective columns in Table C1. Column (3) shows the estimates from a specification that does not include political view controls. The results in all three columns are similar both in magnitude and significance.

Table C3 presents the heterogeneity in treatment effect estimates based on community characteristics, such as past solar campaigns and population. The estimates for each of these two characteristics are obtained under three specifications. The first two specifications use the above definitions of the political view indicator variables. The third specification drops all demographic, socioeconomic, and political controls. Again, the results are very robust across these specifications.

Finally, Tables C4 and C5 replicate the regressions from Table 7, but drop some or all of the time-invariant fixed effects. A potential concern is that, since the variation in household- and community-specific characteristics is entirely cross-sectional, much of it may be absorbed by the inclusion of too many controls and time-invariant fixed effects. Such overfitting could potentially lead to Type I and Type II errors. The robustness checks shown in Tables C1-C3 already demonstrated that our estimates are strongly robust to the exclusion of demographic, socioeconomic, and political controls. We now drop ZIP code fixed effects in Table C4, as well as both ZIP code and staff member fixed effects in Table C5. As shown in Tables C4 and C5, all of our findings with regards to Hypotheses 2a-c still hold even after excluding these time-invariant fixed effects.

Table C1: Heterogeneity by Party Affiliation

Variable	(1)	(2)
Treatment	0.011**	0.011***
	(0.005)	(0.003)
Treatment×Democrat	-0.001	-0.004
	(0.006)	(0.006)
Treatment×Republican	0.002	0.009
	(0.006)	(0.011)
Democrat	0.0006	0.0028
	(0.0022)	(0.0017)
Republican	-0.0003	0.0025
	(0.0030)	(0.0039)
Oil heating	yes	yes
Demographic controls	yes	yes
Socioeconomic controls	yes	yes
ZIP code fixed effects	yes	yes
Staff member fixed effects	yes	yes
Duration dependence function	yes	yes
Calendar week fixed effects	yes	yes
Day-of-week fixed effects	yes	yes
Average Treatment Effect by Party Affiliation		
Democrat	0.010***	0.010***
	(0.003)	(0.003)
Republican	0.011***	0.012***
	(0.004)	(0.004)

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. "Democrat" and "Republican" are binary indicator variables. Omitted group is "independent." Duration dependence function is a quadratic in number of days since entering the sample. Standard errors are clustered at the ZIP code level. Standard errors for the average treatment effects are computed using the Delta method. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table C2: Heterogeneity by Household Characteristics

Variable	(1)	(2)	(3)
Treatment	0.005	0.004	0.005
	(0.018)	(0.018)	(0.018)
Treatment×age	0.0003	0.0003	0.0003
	(0.0003)	(0.0003)	(0.0003)
Treatment×white	-0.024**	-0.024**	-0.024**
	(0.011)	(0.011)	(0.011)
Treatment×household size	0.0002	0.0004	0.0002
	(0.0035)	(0.0035)	(0.0035)
Treatment×college or more	-0.012	-0.011	-0.012
	(0.017)	(0.016)	(0.016)
Treatment×home value	-0.0004	-0.0004	-0.0004
	(0.0012)	(0.0012)	(0.0012)
Treatment×income	0.187**	0.184**	0.187**
	(0.086)	(0.086)	(0.087)
Oil heating	yes	yes	yes
Demographic controls	yes	yes	yes
Socioeconomic controls	yes	yes	yes
Political view controls	yes	yes	no
ZIP code fixed effects	yes	yes	yes
Staff member fixed effects	yes	yes	yes
Duration dependence function	yes	yes	yes
Calendar week fixed effects	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. Party affiliation variables in specifications (1) and (2) defined as in respective columns in Table C1. Duration dependence function is a quadratic in number of days since entering the sample. Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table C3: Heterogeneity by Community Characteristics

Variable	Solar campaigns			Urban vs. rural		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	0.007*	0.007*	0.007*	0.013***	0.013***	0.013***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Treatment×solar campaign	0.009*	0.009*	0.009*	-	-	-
	(0.005)	(0.005)	(0.005)			
Treatment×urban	-	-	-	-0.012***	-0.012***	-0.012***
				(0.004)	(0.004)	(0.004)
Oil heating	yes	yes	yes	yes	yes	yes
Demographic controls	yes	yes	no	yes	yes	no
Socioeconomic controls	yes	yes	no	yes	yes	no
Political view controls	yes	yes	no	yes	yes	no
ZIP code fixed effects	yes	yes	yes	yes	yes	yes
Staff member fixed effects	yes	yes	yes	yes	yes	yes
Duration dependence function	yes	yes	yes	yes	yes	yes
Calendar week fixed effects	yes	yes	yes	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes	yes	yes	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. Total sample contains 15,318 observations. Party affiliation variables in specifications (1) and (2) defined as in respective columns in Table C1. Duration dependence function is a quadratic in number of days since entering the sample. Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table C4: Treatment Effects without ZIP Code Fixed Effects

Variable	Specification			
	(1)	(2)	(3)	(4)
Treatment	0.012***	0.006	0.009***	0.013***
	(0.004)	(0.016)	(0.003)	(0.003)
Treatment×Democrat	-0.002	-	-	-
	(0.005)			
Treatment×Republican	0.004	-	-	-
	(0.006)			
Treatment×age	-	0.0002	-	-
		(0.0003)		
Treatment×white	-	-0.02**	-	-
		(0.01)		
Treatment×household size	-	0.0011	-	-
		(0.0032)		
Treatment×college or more	-	-0.012	-	-
		(0.013)		
Treatment×home value	-	-0.0001	-	-
		(0.0011)		
Treatment×income	-	0.143*	-	-
		(0.078)		
Treatment×solar campaign	-	-	0.008**	-
			(0.004)	
Treatment×urban	-	-	-	-0.007*
				(0.004)
Oil heating	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes
Socioeconomic controls	yes	yes	yes	yes
Political view controls	yes	yes	yes	yes
ZIP code fixed effects	no	no	no	no
Staff member fixed effects	yes	yes	yes	yes
Duration dependence function	yes	yes	yes	yes
Calendar week fixed effects	yes	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. The main effects are included in the regression, but omitted above. Total sample contains 15,318 observations. Duration dependence function is a quadratic in number of days since entering the sample. “Democrat” and “Republican” are binary indicator variables equal to one for a given household if at least half of the household members are registered as supporters of the respective party. Omitted party affiliation group is “independent.” Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table C5: Treatment Effects without ZIP Code and Staff Fixed Effects

Variable	Specification			
	(1)	(2)	(3)	(4)
Treatment	0.014***	0.005	0.009***	0.014***
	(0.004)	(0.016)	(0.003)	(0.001)
Treatment×Democrat	-0.004	-	-	-
	(0.004)			
Treatment×Republican	0.001	-	-	-
	(0.005)			
Treatment×age	-	0.0003	-	-
		(0.0003)		
Treatment×white	-	-0.021**	-	-
		(0.01)		
Treatment×household size	-	0.0015	-	-
		(0.0031)		
Treatment×college or more	-	-0.01	-	-
		(0.013)		
Treatment×home value	-	-0.0007	-	-
		(0.001)		
Treatment×income	-	0.145*	-	-
		(0.079)		
Treatment×solar campaign	-	-	0.007*	-
			(0.004)	
Treatment×urban	-	-	-	-0.008**
				(0.004)
Oil heating	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes
Socioeconomic controls	yes	yes	yes	yes
Political view controls	yes	yes	yes	yes
ZIP code fixed effects	no	no	no	no
Staff member fixed effects	no	no	no	no
Duration dependence function	yes	yes	yes	yes
Calendar week fixed effects	yes	yes	yes	yes
Day-of-week fixed effects	yes	yes	yes	yes

Note: Dependent variable is an indicator for completed audit. Unit of observation is household-day. The main effects are included in the regression, but omitted above. Total sample contains 15,318 observations. Duration dependence function is a quadratic in number of days since entering the sample. "Democrat" and "Republican" are binary indicator variables equal to one for a given household if at least half of the household members are registered as supporters of the respective party. Omitted party affiliation group is "independent." Standard errors are clustered at the ZIP code level. $p < 0.1$ (*), $p < 0.05$ (**), $p < 0.01$ (***).