Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment

Hunt Allcott, Richard L. Sweeney

To cite this article:
Hunt Allcott, Richard L. Sweeney (2017) The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment. Management Science 63(1):21-39. http://dx.doi.org/10.1287/mnsc.2015.2327

http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 63, No. 1, January 2017, pp. 21–39
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment

Hunt Allcott,[a] **Richard L. Sweeney**[b]

[a] Department of Economics, New York University, New York, New York 10012; [b] Economics Department, Boston College, Chestnut Hill, Massachusetts 02467

Contact: hunt.allcott@nyu.edu (HA); sweeneri@bc.edu (RLS)

Received: March 16, 2015
Accepted: July 17, 2015
Published Online in Articles in Advance: February 16, 2016

https://doi.org/10.1287/mnsc.2015.2327

Abstract. With a large nationwide retailer, we run a natural field experiment to measure the effects of energy use information disclosure, customer rebates, and sales agent incentives on demand for energy-efficient durable goods. Although a combination of large rebates plus sales incentives substantially increases market share, information and sales incentives alone each have zero statistical effect and explain at most a small fraction of the low baseline market share. Sales agents strategically comply only partially with the experiment, targeting information to more interested consumers but not discussing energy efficiency with the disinterested majority. These results suggest that seller-provided information is not a major barrier to energy-efficiency investments at current prices in this context.

History: Accepted by John List, behavioral economics.
Funding: The authors are grateful to the Sloan Foundation for funding this experiment and the related research on the economics of energy efficiency.
Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mnsc.2015.2327.

Keywords: energy efficiency • energy-using durables • information disclosure • randomized field experiments

1. Introduction

Consumers often learn about new or higher-quality products from firms, and in theory, information problems can be ameliorated as sellers of relatively high-quality products inform consumers of their beneficial attributes (see Grossman 1981, Milgrom 1981, Viscusi 1978). In many cases, however, barriers to information transmission can cause a failure of "unraveling," and imperfect information could reduce demand for high-quality products. In these situations, regulators may be able to increase welfare by mandating or otherwise inducing information disclosure, subsidizing quality, or even setting minimum quality standards. Before intervening, however, a regulator might want additional information: How successful are firms at providing information? How well informed are consumers about new products and their attributes?

Every year, Americans purchase $361 billion in energy-using durable goods such as cars and air conditioners and spend $570 billion on energy for those goods (Bureau of Labor Statistics 2011). Regulators intervene in durable goods markets by mandating energy use information disclosure and by encouraging additional marketing of energy efficiency through initiatives such as the Energy Star Retail Partner program. Imperfect information is also commonly used to justify extensive subsidies for energy-efficient goods, as well as minimum energy efficiency standards.[1] Despite the importance of imperfect information in the energy policy debate, however, there is limited evidence on how energy cost information disclosure affects durable goods demand. Dranove and Jin's (2010) definitive review of the information disclosure literature mentioned zero studies related to energy efficiency, although we discuss below how this literature has recently received more attention.

We study water heaters, which are interesting and important precisely because they are so mundane. Consumers rarely think about their water heater until it breaks unexpectedly, and at that point they want to replace it quickly, with limited time for search and information acquisition. Retailers thus play a pivotal role in guiding purchases. At average product lifetimes and usage rates, purchasing an energy-efficient Energy Star natural gas water heater instead of a standard model is an investment with 13%–18%return, and this is before the generous subsidies offered by many local utilities. Despite this, the Energy Star market share is only approximately 3% at the retailer (hereinafter the Retailer) we study. These choices are expensive: water heating is the second largest home energy use in the United States (U.S. Department of Energy 2009), consuming approximately $300 in energy annually per household, or approximately $29 billion per year nationwide.

Motivated by these issues, we carried out a natural field experiment with a large nationwide retailer that sells water heaters and many other goods. We worked at the Retailer's call center, which sells approximately 45,000 water heaters each year. More than 20,000 callers were randomly assigned between treatments in which sales agents were instructed to provide energy cost savings information and/or offer customer rebates for Energy Star models. We also offered $25 sales incentives for agents who sold Energy Star on randomly selected calls, and we crossed these incentives with the customer rebates.

A crucial feature of our experiment is that the seller's interactions with customers are intermediated by sales agents. This is not uncommon: consumers learn about life insurance, mutual funds, and many consumer goods at least partially through agents. In our setting, sales agent behavior is important for two reasons. First, it directly determines the Retailer's ability to market Energy Star products: if sales agents do not provide information on a call, callers will likely remain uninformed. Second, in equilibrium, it is indirectly informative about consumers' responsiveness to information: given that information disclosure takes time and focus away from other sales tasks, if consumers are not interested in information, agents will not provide it. To document sales agent behavior, our research team independently audited more than 2,000 phone calls, quantifying the interactions between agents and consumers. Our ability to observe agent behavior, instead of simply the equilibrium outcome of attempting to disclose information, is one feature that distinguishes this paper from previous work.

There are several reasons to expect that the treatments could substantially increase Energy Star market share. Because consumers are thought to be poorly informed about water heater features, they often accept sales agents' recommendations about what model to purchase. The $100 customer rebate increases the average consumer's return on investment in an Energy Star model to 28%–37%, and when combined with additional subsidies available from many local utilities, our $100 rebate brings the incremental purchase price of the Energy Star model close to zero. The $25 sales incentives are equal to two times the agents' fixed hourly wage and are 10 times larger than their usual sales incentive.

Against this backdrop, our results are surprising. Our audits show that agents comply with delivering the information and rebates on only an approximate one-fifth of calls. Of course, even with incomplete compliance, we can still estimate local average treatment effects on customers who do receive the treatments by analyzing the experiment as a randomized encouragement design. Information has zero statistical effect on demand, and confidence intervals rule out that demand increases by more than 4.9 percentage points on calls when the information is delivered and the consumer is considering a substitutable model. Although this bound is large relative to the baseline Energy Star market share, it suggests that the market share would still be very low even if agents informed all consumers. The $100 customer rebates do increase Energy Star purchases, however, and the combination of a $25 sales incentive and $100 customer rebate appears to have particularly strong complementary effects.

We show that agents preferentially market Energy Star to consumers with higher latent demand for Energy Star. Furthermore, on calls with the $25 sales incentive but no experimental customer rebate and no explicit direction to deliver an informational script, agents exert very little effort to sell Energy Star models. Combined with the small information effect, these results suggest that agents' noncompliance is better described as "strategic" instead of "shirking": agents do not inform consumers about Energy Star because they know that, at the retailer's base price, most consumers are not interested in the product once informed. Discussing energy efficiency (or any other nonessential issue) can reduce the probability of a sale by extending call times or potentially irritating customers.

There are two potential explanations for why the Retailer's attempts at information disclosure did not increase Energy Star demand. First, consumers may tend to be unaware of Energy Star or underestimate its benefits, but sales agents may not be able to address this because their "disclosure technology" is limited: they work in time-constrained sales interactions and may have a perceived lack of credibility when promoting a higher-priced model. Second, consumers might already be relatively well informed about Energy Star availability and benefits, and most choose not to buy because they do not think that the reduced energy use is worth the incremental upfront cost. These two explanations have very different implications for whether regulators should intervene to provide information or otherwise encourage energy efficiency.

We carried out an extensive set of customer follow-up surveys to shed light on these two explanations. On the one hand, there is evidence that consumers are confused: even when "Energy Star" is precisely defined, 52% of consumers report believing that they had bought an official Energy Star model, whereas only 2.1% of the survey sample actually had. Of the consumers who thought they had not bought Energy Star, 15% reported that this was because they were not aware that there was an Energy Star option. On the other hand, the great majority of consumers were aware of Energy Star, and their foremost reason for not purchasing was that the price was too high. Furthermore, although there is wide dispersion of beliefs, the

average consumer actually overestimates the potential energy cost savings from Energy Star. Although the survey results are not as conclusive as the experimental results, they at least suggest that lack of awareness and cost savings information are not the primary barriers to energy efficiency in this context.

The remainder of this section discusses related literature. Section 2 provides an overview of the water heater market, §3 details the experimental design and data, and §4 presents the empirical results. Section 5 presents follow-up survey data, and §6 concludes. The online appendix (available as supplemental material at https://doi.org/10.1287/mnsc.2015.2327) includes a simple theoretical model—an extension of Grossman and Shapiro's (1984) analysis of informational advertising in a Hotelling spatial model—that helps to motivate the experiment and interpret results.

1.1. Related Literature

Our study is broadly connected to the literature on information disclosure, as reviewed by Milgrom (2008) and Dranove and Jin (2010).[2] There are a number of papers studying energy information disclosure that differ from ours in only one respect. Some studies analyze the effects of energy use information disclosure on stated preferences or other proxies for actual purchases of durable goods, including Davis and Metcalf (2014), Deutsch (2010a, b), Newell and Siikamaki (2013), and Ward et al. (2011). Some studies either use observational data (Kallbekken et al. 2013) or randomly assign a very small number of units (Anderson and Claxton 1982).[3] Also related is Houde's (2014a) analysis of the Energy Star label and other studies of how various kinds of information affect total household energy use, such as Allcott (2011), Dolan and Metcalfe (2013), Jessoe and Rapson (2014), and others. Along with Allcott and Taubinsky (2015), our paper is slightly different in that it uses large-sample RCTs to study how providing information about a durable good's energy use affects actual purchases of that good. This particular question is important given the regulatory resources devoted to durable good energy use disclosure and given the costly energy-efficiency standards and subsidies that are partially predicated on the idea that consumers remain imperfectly informed when purchasing durable goods. Furthermore, our paper is substantially conceptually different from the rest of the energy information literature because of its focus on a situation where the information provision process is intermediated by sales agents.

Our experiment is also related to studies of behavior by sales agents and advisers, including field experiments by Anagol et al. (2013), Mullainathan et al. (2012), and Nagin et al. (2002), as well as theoretical analyses by Hoffman et al. (2013), Inderst and Ottaviani (2009), and Inderst and Ottaviani (2012). This literature largely focuses on information asymmetries between sales agents and consumers or alternatively on agency problems between firm managers and workers. Although these issues could be at play in our context, they are not our focus. Instead, we highlight agents' imperfect and differential compliance with management directives to provide information. This angle is comparable to findings by Duflo et al. (2006) that tax preparation professionals have different levels of success in encouraging tax filers to contribute to retirement accounts. Our experimental tests of sales force incentives connect us to the sales force management literature (see Mantrala et al. 2010 for a review and Chan et al. 2014 for a recent example), to the "insider econometrics" approach to studying employee compensation and management practices (see Ichniowski and Shaw 2003 for an early review), and to "process" field experiments in strategy research (Chatterji et al. 2016).

A key innovation relative to the literatures on information disclosure and employee incentives is that we highlight the interaction between consumer and firm behavior in equilibrium. Much of the recent empirical work on information disclosure studies consumer responses to information that is experimentally provided with certainty or disclosed by firms under mandate, which isolates consumer behavior independent of the firm. Conversely, much of the employee incentives literature focuses on worker behavior in isolation of the consumer.[4] By contrast, a central feature of our setting is that the firm's ability to motivate its sales agents to promote a product depends crucially on consumer interest in that product, which is in turn determined by the firm's pricing decisions.[5] The theoretical model in the online appendix makes clear how these interactions play out in equilibrium.

2. Market Overview

2.1. The Water Heater Market

In 2012, there were 7.69 million residential storage water heaters sold in the United States, of which 51% were fueled by natural gas and 49% used electricity (AHRI 2013). Replacement units (as opposed to units installed in new homes) historically represent 82% of total sales, although this varies with the housing market.[6] The typical water heater remains in use for 13 years.

Approximately half of all units are sold through wholesale distributors; of these, 87% are purchased and installed by plumbers. The remaining half of all units are purchased through retail channels such as our partner Retailer. In 2010, the Retailer had a 9% share of the retail market, third behind two other retailers that had 23% and 19% shares. Thirty percent of the Retailer's sales are made through the call center where our experiment takes place, whereas the

remaining 70% are made in physical retail establishments. Our sample thus includes a small but nontrivial share of all water heaters sold in the United States over the study period.

2.2. Water Heater Attributes

Water heaters are a convenient product to study because they are differentiated only on a few dimensions.[7] Key characteristics are as follows:

- Fuel type: natural gas, propane, or electric. Consumers' choices depend on what fuels are available in their houses, so choices along this dimension are effectively exogenous. We limit our study to natural gas and propane water heaters, because there are no Energy Star electric models. Less than 1% of sales in the sample were propane-fueled models.
- Storage tank size. Residential tank sizes range from 30 to 80 gallons. In our sample, 90% of sales are either 40 or 50 gallons.
- Warranty length. The Retailer offers models with warranty lengths of 3, 6, 9, and 12 years. Models with longer warranties are typically higher quality, using additional or improved anode rods to delay or fully prevent rusting.
- Tank height. Some consumers need to install water heaters in basements with low ceiling heights. In our sample, approximately eight percent of sales are "short" models, whereas the rest are standard height.
- Energy use.

2.3. Water Heater Energy Use and Energy Star Technology

Each model's energy use is tested at an independent laboratory using U.S. government test protocols. Test results are used for the Federal Trade Commission's EnergyGuide labels, or "yellow tags," which are energy use information labels that provide estimated annual energy costs based on national average usage and energy prices. The yellow tags report total energy costs across all fuels used, which for the Energy Star models includes both natural gas and electricity. By law, yellow tags must be displayed on water heaters in a showroom, and the Retailer's website also includes PDFs of the yellow tag next to each model's description. Thus, although many consumers may not see or attend to energy cost information, it is easily verifiable. Of course, each household's actual energy costs may differ from the average because of utilization rates, climate, and other factors.

The energy use test protocols are also used to calculate statistics called "energy factors," which represent the share of energy input into the water heater that is transformed into hotter water instead of otherwise dissipated. To qualify for Energy Star status, a natural gas water heater must achieve an energy factor of 0.67 or above, compared to the standard 0.59.

Table 1. Water Heater Model Overview

	40 Gallon		50 Gallon	
Warranty (years)	6	12	6	12
Price ($)				
Standard	420	620	485	665
Energy Star	645	969	700	1,020
Annual energy cost ($/year)				
Standard	309	290	315	294
Energy Star	272	261	272	261
Undiscounted payback period (years)	6.1	12.0	5.0	10.8
IRR (at 13-year average life) (%)	13	1	18	3
Market share (%)				
Standard	17.6	6.1	10.1	10.4
Energy Star	0.6	0.5	0.2	0.7

Notes. This table presents information on the four different Energy Star natural gas water heater models sold by the Retailer, as well as their closest non-Energy Star substitutes. The standard and Energy Star 6-year warranty models are essentially undifferentiated other than price and energy use, whereas the 12-year warranty Energy Star models have other premium features. IRR, internal rate of return.

During our study period, the Retailer sold four natural gas Energy Star models.[8] Two are modified versions of their standard 40- and 50-gallon models with 6-year warranties. To improve energy efficiency, the manufacturer adds another inch of insulation around the tank and uses electric ignition instead of a continuously burning pilot light. To accommodate electric ignition, the Energy Star models must be plugged into a power outlet, and they consume a small amount of electricity. They also have electronic thermostats and a more advanced flue damper that opens and closes depending on whether gas is currently being burned. These differences only affect energy use and have no material impact on unit performance, and there are no other differences between the standard and Energy Star 6-year warranty models. (Of course, consumers may attach other connotations to the Energy Star label, and as we mention below, local in-stock availability is an additional differentiator.)

The other two Energy Star models are 40- and 50-gallon premium models with 12-year warranties. The premium models have the same amount of insulation as the standard 12-year models; they achieve higher energy efficiency through electric ignition and other modifications to the combustion process. The premium Energy Star models also differ from the standard 12-year models on other dimensions that make them generally higher-quality.

Table 1 presents information on these four Energy Star models and how they compare to their closest substitutes. Standard models cost $400–$700, not including installation. According to the yellow tags, standard models use approximately $300 worth of energy each year, meaning that lifetime cost is much larger than upfront purchase price. Energy Star models save approximately $30 per year. Because the 6-year

warranty models are very close substitutes except for purchase price and energy cost, the 13% and 18% internal rates of return are reasonable approximations of the expected net benefits of Energy Star. By contrast, the 12-year Energy Star model also has other premium features, so 1% and 3% internal rates of return do not capture the premium model's full benefits.

2.4. The Sales Process

According to U.S. Department of Energy (2010), 35%–40% of replacement purchases nationwide arise suddenly due to complete unit failure, typically when water rusts through the steel tank and escapes onto the floor. Our follow-up customer surveys show that 83% of purchases in our sample were due to unexpected breaks instead of planned replacements. Because most people do not like cold showers, consumers typically want to replace their water heater within 24 hours if possible. This hurry has several implications. First, consumers have not saved money in anticipation of a large expenditure, so they may be especially price sensitive. Second, consumers tend to prefer models that are in stock locally and can thus be installed quickly. Because sales volumes are lower, the Retailer stocks Energy Star water heaters at fewer locations than its standard models. Third, consumers make little time to acquire information about different types of water heaters and their attributes.

A quote from our survey of the Retailer's sales agents nicely summarizes these issues: "Customers that were shopping ahead [i.e., not responding to an unexpected unit failure] seemed to be making more educated decisions . . . they were more inclined to use the Energy Star water heaters as item they wanted their quote for. I feel that whenever there was not such a sense of urgency . . . customers were in a position to spend more on a better water heater and also able to wait for it to be ordered."

When customers call the Retailer's water heater call center, sales agents have significant influence over their decisions. Some callers have done background Internet research and think they know what model they want before calling, whereas the majority know only that they need a new water heater. Agents work with these callers to determine which model is best for them based on the attributes discussed above, such as fuel type, ceiling height, local low-NOx regulations, and appropriate tank size.

Before the experiment started, we called the Retailer's call center a number of times, acting as "mystery shoppers." We found that the Retailer's sales agents have been successfully trained to look up Energy Star rebates offered by local utilities, discuss Energy Factors, and discuss information on yellow tags. Unless the caller asks about energy efficiency, however, agents never discussed the issue with us, because information disclosure is costly. As one agent wrote on our survey, "I would say about 90% of our customers only care about how cheaply can they get away with the purchase of a water heater." Sharing extraneous information increases call times, and many call centers evaluate agents on call times in order to keep labor costs low. Longer call times can reduce customer satisfaction and increase the probability that the customer gets distracted and does not complete the sale.

These features of the water heater sales process motivated our experiment: perhaps Energy Star sales are low because consumers are unaware of the product and its benefits, and agents' influence over consumers could be leveraged to increase awareness.

3. Experimental Design and Data

3.1. Sales Associates and the Sales Process

The Retailer's water heater division operates two call centers. There are 77 sales agents who take at least one call during our sample period. These sales agents sell only water heaters, not other goods. The agents report to team managers, who in turn report to shift managers, who report to the call center manager. Agents make between \$11 and \$14 per hour, depending on seniority, along with sales incentives that typically scale closely with purchase price and average approximately \$4. Interestingly, however, sales incentives are only slightly higher for the 6-year warranty Energy Star models compared to their closest non-Energy Star substitutes, despite purchase prices that are approximately \$200 higher. Sales incentives for the premium 12-year warranty Energy Star models are approximately \$5 larger than for the standard 12-year warranty non-Energy Star models.

The Retailer has an established set of processes that sales agents are to follow on each call. Approximately 60%–65% of calls are recorded at random, and managers monitor a subset of these calls for evaluation and quality assurance. The sales agents meet with their managers weekly to review performance and talk about sales initiatives and modifications to the sales process.

When a customer calls, he or she is routed to the first available sales agent. The call centers use caller ID, and the agents verbally confirm the caller's phone number. Using this phone number, the customer is assigned a "reference number." We define a "consumer" as a unique reference number. Individuals often call more than once as they comparison shop or gather more information. If an individual calls more than once from the same phone number or verbally gives the same number to a sales agent, then he or she is tracked as a unique customer.

Once the sales agent and consumer agree on a water heater model, the sales agent checks whether the

model is in stock in the customer's region, arrives at a price quote, records the customer address, and charges the customer's credit card. Customers can install the unit themselves, hire a third-party plumbing contractor, or pay the Retailer to do the installation.

3.2. Experimental Design

The sales agents have a standard computer interface that has the Retailer's sales program plus Internet access. To implement the experiment, the Retailer's staff redesigned the interface to open the experiment website each time a reference number is entered. Our research team designed and programmed the experiment website, which afforded us full control over the randomization and other content.

On the website's initial screen, the agent would enter the customer's needed fuel type (gas or electric) and click "GO." After the agent clicked "GO," the website would display call-handling instructions, including a script that the agent was to cover with the customer. The different treatments were implemented through these scripts. Electric customers are excluded from the experimental population, with the website displaying "No Script." Natural gas customers form the experimental population, and they are randomly assigned to one of the treatment groups. The online appendix includes screen shots from the experiment website.

Both agents and callers were randomized. Agents were randomly assigned as information treatment agents or information control agents. Callers were randomly assigned to treatment groups based on their reference number. Thus, consumers who called multiple times but kept the same reference number remained in the same treatment condition. Consumers who first spoke with an information control agent were automatically assigned to Information Control, whereas callers who first spoke with an information treatment agent were randomly assigned to either Information Treatment or Information Control.

Table 2 displays the experimental timeline and treatment groups. Phase 1 ran from November 2012 to April 2013. During this time, there was a three-by-two matrix of treatments: customers were randomly assigned to \$0, \$25, or \$100 rebate, which was crossed with Information Treatment or Control. Phase 2 of the experiment ran from early April to early June 2013. In this phase, we added a sales incentive, which the Retailer calls a "spiff." Phase 3 lasted from early June to early July. In this phase, we added two final treatments, which were interactions of the spiff with the two rebate levels. In phase 4, the Retailer ended the spiff treatments but continued the rest of the experiment for several weeks. In total, there were eight different treatment cells, plus control.

Table 2. Experiment Timeline

Phase	Dates	Info, rebates, and info x rebates	Spiff	Spiff x rebates	Consumers in sample	Sales
1	Nov 21–April 6	Yes			12,629	4,675
2	April 7–June 13	Yes	Yes		7,254	2,523
3	June 14–July 6	Yes	Yes	Yes	1,974	715
4	July 7–July 26	Yes			1,490	362

Below, we give examples of the call-handling instructions for several example treatments. In the Information Treatment condition with no rebate, the website instructed the agent to read the following script to the customer:

> *Let me take a moment to tell you about our Energy Star models. Energy Star water heaters cost about* \$220 *more than a standard model, but they save a typical household* \$40 *each year, so you would make up that price difference in about six years. Over* 12 *years, which is the normal life of a water heater, you would save* \$480 *in energy bills. Energy Star models may not be available for every home. If possible, would an Energy Star water heater be of interest to you*?

In the rebate condition with no information, the agent was instructed to say:

> *I have good news.* [*Retailer*] *has specially selected you for a* \$100 *rebate on any Energy Star water heater. Energy Star models may not be available for every home. If possible, would an Energy Star water heater be of interest to you*?

If the customer was assigned to the spiff, the call handling instructions read:

> *ENERGY STAR SPIFF CALL*
>
> *You* (*the Retail Hotline Associate*) *will receive* \$25 *on your next paycheck if this caller buys any Energy Star water heater. You can share with the caller any useful information about the benefits of Energy Star, perhaps including environmental benefits or long-run energy cost savings. The caller does not need to purchase on the initial call. If the same caller calls back later and uses the same reference number, all RHAs that spoke with that reference number earn the* \$25.

In the combined spiff-plus-rebate conditions that were added in phase 3, the sales agent was not instructed to read a specific script. Instead, the call-handling instructions told sales agents that the customer was eligible for a rebate and left it to the sales agent to decide how to phrase that information.

> *ENERGY STAR SPIFF CALL* + \$25 *CUSTOMER REBATE*
>
> *You* (*the Retail Hotline Associate*) *will receive* \$25 *on your next paycheck if this caller buys any Energy Star water heater.*
>
> *The customer will also receive a* \$25 *rebate off of any qualifying Energy Star model.*

In the control group, the instructions read:

> *CONTROL GROUP*: *NO SCRIPT*
>
> *This customer is in the control group. Proceed with the call as you normally would. Answer any questions the customer has, but try not to use any of the language in the information treatment script.*

At the end of the call, the sales agents reported in the experiment website whether or not they delivered the script. As we shall see, these self-reports overstate compliance relative to our independent audits. The website and the team managers instructed the agents that the only reasons not to deliver the script were if the customer needed a low-NOx, short tank, or other specialty model that was not substitutable with Energy Star models. If the agent did not complete the script, the website required the agent to select the reason for noncompliance from a dropdown menu.

The experiment was closely integrated into the call center processes. At the outset, managers trained the agents on the scripts and how to use the website, and this was also part of training for newly hired agents during the experiment. We also communicated directly with the agents through several group emails and two videos that explained the importance of compliance with the experiment. Specifically, we emphasized the importance of both delivering the scripts on treatment group calls and not discussing elements of the scripts on control group calls.

Every week of the experiment, we provided the Retailer with agent-specific compliance reports based on self-reported compliance from the website. We had biweekly calls with managers to discuss these compliance reports, and managers could then discuss with agents in their weekly meetings. The Retailer's internal call monitors also audited calls for compliance with treatment assignment. Agents with low compliance with the experiment were pressured by managers to do better. To encourage competition between the two call centers, managers also reviewed average compliance for each call center, as well as trends over time. In the endline survey, sales agents reported that managers frequently emailed and talked with them about the experiment. In sum, agents did face some costs if they did not at least report compliance with the experimental protocols. However, this experiment was only one of many issues that managers and agents needed to attend to.

Individuals who call multiple times from multiple phones and do not tell the sales agents that they have previously called would have been assigned different reference numbers, and thus potentially different treatments. This could generate spillovers, for example, if a caller who purchases using an Information Control reference number had been assigned an Information Treatment reference number on a previous call. Based on our conversations with Retailer staff, we do not believe that this happens on more than a handful of calls, although we do not have a precise estimate.

Some consumers, perhaps plumbing contractors or landlords of multiple homes, order multiple water heaters during the experiment. The Retailer gives these individuals separate reference numbers on their separate purchases, and they are thus treated as separate "consumers" in the experiment. Although this also could generate spillovers, it could not have more than a negligible impact on the estimates because it affects only a very small share of the sample: there are 104 individuals, or 0.4% of the sample, who order two water heaters from the same phone number, and no phone number appears more than twice in the sales data.[9]

3.3. Data

There are several main data sources. The first is the Retailer's call database. An observation consists of the unique customer reference number, date and time of the call, and the agent receiving the call. This database includes only sales calls, not warranty service, repairs, or other types of calls.[10] Using the reference number, this is matched to the Retailer's purchase data, which include the model purchased, price paid, and other details.

The Retailer's call database is also matched by reference number to the experiment website database. This database includes the treatment assignment and the agent's self-reported compliance for each reference number where the website was opened. For the 1.3% of reference numbers that appear in the website data multiple times, we code that the script was read to the consumer if any agent reported that he or she had done so on any call. In the regressions, each consumer i must be associated with an individual agent a; we use the last agent in the website who spoke with customer i.[11] We define a variable N^s_{iat} that takes value 1 if agent a reported compliance on a treatment group call with consumer i, and 0 otherwise. Agents were not explicitly asked to read a script on spiff treatment calls or on control calls. We define N^s_{iat} as missing for spiff treatment calls and zero for control calls.

The total number of consumers (reference numbers for consumers interested in natural gas water heaters) recorded in the call database during the experiment is 38,179. Of these, 23,347 (61%) are in the website and are randomly assigned to a treatment group; these calls comprise our "experimental population." The calls that are not recorded in the website are largely conversations that did not last long enough for the sales agent to activate the website. As Table 2 shows, 35% of consumers (8,275 in total) purchased from the Retailer. Because consumers are effectively unable to substitute across some features, we define a subset of "substitutable" models that includes all Energy Star and non-Energy Star natural gas tank water heaters except for low-NOx, short tank height, mobile home, power vent, and propane models; 73% of sales were substitutable models. Of this substitutable group, only 3.5% were Energy Star.

For consumers who purchased water heaters, the Retailer recorded their name and address. Zip codes were used to match median income from the most recent American Community Survey (ACS) five-year estimates and the share of vehicles registered in the zip code that are hybrids, which could be an important correlate of environmentalism and interest in energy efficiency. Using each purchaser's name and address, a marketing data company called Acxiom provided assessed home value, college graduate indicator, age, household size, and political affiliation. Acxiom gathers data from public records, magazine subscriptions, voting records, scanner data, online purchases, and other sources, and their data are certainly measured with error. For the approximately 10% of addresses missing the college graduate indicator, age, or household size, we substitute zip code-level means, again from the ACS five-year estimates.

Using the Acxiom political affiliation data, we construct a variable called *Democrat* that takes value 1 if the purchaser is a registered Democrat, and 0 otherwise. If political affiliation is missing, we replace Democrat with the county-level ratio of Democrat to Democrat plus Republican votes in the 2004 and 2008 presidential elections, using data from the Atlas of U.S. Presidential Elections (Leip 2013). Acxiom also provided two additional levels of environmentalism. *Environmentalist* is an indicator variable for whether the consumer subscribes to environmental magazines or contributes to environmental or animal welfare charities. *Green Living* is an indicator that takes value 1 if *Environmentalist* equals 1 or if the household purchases environmentally healthy products such as eco-friendly soaps and organic foods.

Table 3 presents sample means and standard deviations for our nine demographic variables. Consumers in our sample are older and wealthier than the general population, likely related to the fact that they are almost entirely homeowners. They are also more liberal and environmentalist, as illustrated by their Democrat scores and zip code hybrid vehicle shares. In our data, Energy Star demand is positively associated with zip code median income, home value, and zip code hybrid share, reminiscent of the findings of Kahn (2007). This suggests that the Energy Star market share would be even lower in a nationally representative sample.[12]

Table 3. Representativeness

	Sample mean	Sample std. dev.	National average
Zip Median Income ($ thousands)	71	27.3	56.9
Home Value ($ thousands)	338	293	246
College Graduate	0.61	0.43	0.32
Age	57.3	13.4	37.3
Household Size	3.2	1.5	2.4
Democrat	0.62	0.34	0.53
Zip Hybrid Share (of 100)	1.3	1.1	0.94
Acxiom Green Living	0.31	0.46	—
Acxiom Environmentalist	0.14	0.35	—

Notes. This table gives the mean and standard deviation of customer demographics. These variables are matched based on addresses and are thus available only for consumers who purchase water heaters. National average college graduate share is for people older than age 25. National averages for the *Acxiom Green Living* and *Environmentalist* variables are not available.

Table A.1 in the online appendix presents balance tests for the nine demographic variables plus two measures of call start time, for each of the eight treatment cells relative to control. Only one of the 88 t-tests rejects equality with greater than 90% confidence, and all 11 F-tests easily fail to reject that the treatment groups are balanced on observables.

3.3.1. Customer Follow-Up Surveys. We hired an independent survey research firm to conduct telephone follow-up surveys of customers who had called between December 10 and June 29. We designed two separate surveys, one for consumers who had purchased from the Retailer and one for consumers who had called but not purchased. For purchasers, we asked a battery of questions covering household information, the water heater purchase process, and the Energy Star product. For nonpurchasers, we asked whether they had purchased an Energy Star water heater and why they had decided not to buy from the Retailer. For this analysis, we focus on questions related to consumers' knowledge of the Energy Star model, which we only asked of purchasers. Any other results are certainly available upon request, and the survey protocols are available as part of the replication files.

We directed the call center to complete no more than 200 surveys of nonpurchasers and as many surveys as possible of purchasers. To maximize response rates, we offered a $25 gift card from the Retailer to any respondents who initially attempted to refuse; 149 people accepted. In total, there were 1,091 completed surveys (including 891 from purchasers) from 6,342 attempts, for a response rate of 17%.

3.3.2. Independent Audits of Recorded Phone Calls. Our research assistant (RA) audited 2,122 calls from natural gas water heater consumers recorded between May 1 and July 18. These 2,122 calls are all recorded calls assigned to any of the treatment groups during that period, along with approximately five calls per day from the control group. The audits were blind, meaning that the RA did not know the treatment assignment when auditing a call.

There are two reasons that a call is not observed in our audit data. First, the Retailer's software records only a randomly selected 60%–65% of calls. Second, the database of recorded calls is not organized by reference number, so our RA needed to match recordings

to reference numbers using phone number, time and duration of call, and other information; not all calls could be matched.

We worked with the RA to develop a protocol for quantifying the content of the interaction between the sales agent and the customer.[13] We measure the information provision process using six variables:

- *Mentioned E-Star*: Did the agent mention energy efficiency, energy use, or Energy Star?
- *Rebate*: Did the agent mention the experiment's Energy Star rebate from [the Retailer]?
- *Saves Money*: Did the agent mention that an Energy Star (or energy-efficient) water heater saves money in energy costs?
- *Payback Period*: Did the agent quote a payback period?
- *Read Script*: Does the agent say one of the experiment scripts, either exactly or approximately?
- *ln*(*E-Star Seconds*): For how many seconds did the agent and the customer talk about energy efficiency, energy use, or Energy Star? We use the natural log of one plus this number.

Because there is a small share of consumers who call multiple times, the audit data set includes multiple observations of some consumers. Thus, there are 2,069 unique consumers in our experimental population for whom we have audit data. For the five binary variables above, we collapse using the maximum. In other words, consistent with our construction of agents' self-reported compliance N^s_{iat}, we measure whether a consumer was ever informed about Energy Star or a rebate. For the sixth variable, the number of seconds discussing Energy Star, we take the sum across all of a consumer's calls.

4. Experiment Results

4.1. Sales Agent Behavior

Define N_{iat} as a measure of whether agent a provides information to consumer i during phase t of the experiment. We observe sales agent behavior from two sources: self-reported compliance N^s_{iat} and the independent audits. Although the latter measure is only available for a smaller subset of calls, it is an independent assessment and also provides multiple measures of what the sales agent said. $\mathbf{T}_i$ is a vector of indicator variables for each of the eight treatment cells. We estimate how N_{iat} varies across treatments using the following equation, where$\boldsymbol{\phi}_t$ is a vector of indicators for the four phases of the experiment, $\boldsymbol{\mu}_a$ is an agent fixed effect, and υ_{iat} is the error term:

$$N_{iat} = \beta \mathbf{T}_i + \boldsymbol{\phi}_t + \boldsymbol{\mu}_a + \upsilon_{iat}. \quad (1)$$

The ϕ_t and μ_a controls are necessary because assignment probabilities varied across phases as we added treatment groups and across agents who were assigned to Information Treatment or Control. This equation is estimated as a linear probability model (LPM) in ordinary least squares (OLS) with robust standard errors. In typical cases like ours where the true probability model is not known, Angrist and Pischke (2012) advocate for using the LPM instead of an arbitrary nonlinear model such as probit or logit, and we follow their recommendation. In practice, our results are qualitatively and quantitatively very similar when using probit, logit, or the LPM.

Column (1) of Table 4 presents the results using self-reported compliance N^s_{iat} as the measure of compliance. The sample excludes the spiff treatment calls because agents were not explicitly asked to read a script on these calls. Agents report that they read the script on 46%–49% of calls, and this depends little on treatment assignment. Columns (2)–(7) use data from our independent audits, showing that self-reports substantially overstate compliance. Column (2) shows that, relative to control group calls, agents were approximately 10% more likely to mention Energy Star in Information Only treatment group calls and approximately 14% more likely to do so in Rebate Only calls. The dependent variable in column (3) is the natural log of the estimated number of seconds that the agent and customer discussed energy efficiency; this measures the intensive margin of the dependent variable in column (2). It is 40–80 log points larger in the information and rebate treatment groups.

Columns (4)–(7) of Table 4 directly measure compliance with the experiment scripts. Column (4) shows that agents mentioned the experiment's Energy Star rebate on 14%–24% of calls when the website directed them thus. Columns (5) and (6) show that agents disclosed elements of the information script (quoted a payback period and/or said that Energy Star saves money) on 9%–17% more information treatment calls relative to control. Column (7) reports whether an agent said something approximating one of the treatment scripts during the call. These results are consistent with the results in columns (4)–(6), in that agents appear to comply with the experiment protocol on approximately 15%–20% of calls. The number of observations is lower in this column because we did not begin to record this variable until after the first 326 audits were completed.

The bottom row of Table 4 gives the mean of each dependent variable in the control group. Agent-reported compliance in column (1) is 0 by definition in the control group because the website did not ask agents to report whether they complied on control group calls. Although agents do mention Energy Star on some control group calls, they almost never deliver a script erroneously to the control group: of the more than 400 control group calls that were audited, agents quoted a payback period once and mentioned

Table 4. Compliance by Treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	*Agent Reported*	*Mentioned E-Star*	ln(*E-Star Seconds*)	*Rebate*	*Saves Money*	*Payback Period*	*Read Script*
1(information only)	0.481	0.108	0.486	0.028	0.117	0.133	0.149
	(0.010)***	(0.037)***	(0.272)*	(0.013)**	(0.031)***	(0.024)***	(0.027)***
1(info and $25 rebate)	0.491	0.130	0.635	0.239	0.168	0.149	0.193
	(0.010)***	(0.039)***	(0.274)**	(0.030)***	(0.033)***	(0.025)***	(0.033)***
1(info and $100 rebate)	0.460	0.057	0.788	0.143	0.108	0.093	0.156
	(0.019)***	(0.060)	(0.441)*	(0.045)***	(0.052)**	(0.036)***	(0.057)***
1($25 rebate only)	0.491	0.147	0.413	0.182	0.019	0.000	0.214
	(0.006)***	(0.025)***	(0.234)*	(0.018)***	(0.016)	(0.005)	(0.020)***
1($100 rebate only)	0.494	0.143	0.899	0.190	0.061	0.001	0.180
	(0.010)***	(0.037)***	(0.281)***	(0.031)***	(0.027)**	(0.009)	(0.033)***
1(spiff only)		0.027	0.228	0.032	0.008	−0.002	0.011
		(0.024)	(0.261)	(0.010)***	(0.016)	(0.005)	(0.013)
1(spiff and $25 rebate)		0.051	0.705	0.092	0.073	−0.014	0.036
		(0.050)	(0.410)*	(0.035)***	(0.044)*	(0.010)	(0.029)
1(spiff and $100 rebate)		0.063	0.706	0.172	0.107	−0.033	0.155
		(0.067)	(0.572)	(0.064)***	(0.066)	(0.016)**	(0.065)**
R^2	0.39	0.17	0.15	0.22	0.14	0.20	0.25
N	20,240	2,068	463	2,068	2,068	2,067	1,742
Dep. var. control mean	0	0.13	2.95	0	0.05	0	0.01

Notes. This table reports the estimates of Equation (1). All regressions include agent and phase indicator variables. Robust standard errors are given in parentheses. Column (3) reports the intensive margin of column (2), and the sample is restricted to calls where the agent mentioned Energy Star.

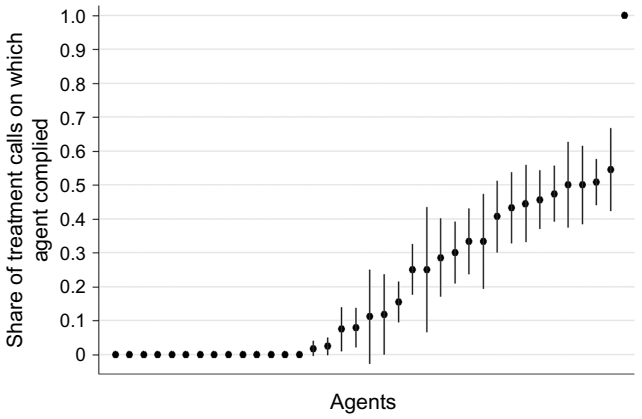
*, **, and *** indicate statistical significance with 90%, 95%, and 99% confidence, respectively.

an Energy Star rebate twice. In total, they gave information that sounded like one of the treatment scripts to the control group less than one percent of the time. Column (4) shows that agents did mention a rebate on a small but statistically significant share of nonrebate treatment group calls—both Information Only and Spiff Only calls. This may reflect some small amount of recording error in the audits or mistakes by the sales agents.

4.1.1. Measuring Sales Agent Compliance. For the next section, we use the audit data to construct a measure of compliance with the experiment scripts. This variable is intended to measure explicit compliance with the experiment scripts, not any other form of discussion of energy efficiency. For consumers whose calls were audited, we define an indicator variable N^+_{iat} that takes value 1 if an agent read a script on any treatment group call, mentioned an Energy Star rebate on a rebate call, or quoted a payback period on an information call. N^+_{iat} takes value 0 otherwise.

There is substantial variation in compliance across agents. Define $\bar{N}^+_a$ as the mean of N^+_{iat} across all of agent a's calls, excluding the Spiff Only and control calls. Figure 1 shows the cumulative distribution function (CDF) of $\bar{N}^+_a$ for all agents who were audited more than five times. Approximately one-fourth of agents never comply, the median $\bar{N}^+_a$ is 11%, and one-fourth of agents comply more than 40% of the time. This dispersion implies that we can exploit variation in compliance rates across agents to improve power in tests of the effects of information provision on demand. Using the data in this graph, we group agents into four compliance groups with $\bar{N}^+_a = 0$, $0 < \bar{N}^+_a \leq 0.2$, $0.2 < \bar{N}^+_a \leq 0.4$, and $\bar{N}^+_a \geq 0.4$. We define G_a as the mean of $\bar{N}^+_a$ across all agents within agent a's compliance group, where agent a is the agent with whom consumer i had his or her final call. The mean values of G_a for agents in

Figure 1. Audit-Based Compliance Rates by Agent



Notes. This figure plots the average compliance rate on all calls other than Spiff Only and control group calls, for all agents that were audited more than five times. Compliance is measured by an indicator variable N^+_{ia} that takes value 1 if an agent read a script, mentioned an Energy Star rebate on a rebate call, or quoted a payback period on an information call.

the four compliance groups are 0, 0.08, 0.29, and 0.53, respectively. Agents who were audited fewer than five times are automatically assigned the G_a for the second compliance group, which includes the median $\bar{N}_a^+$.

4.1.2. Spillovers of Information Provision to Noninformation Group Calls. One reason that agents might not provide much information about Energy Star is that they might not know what to say, or might not be well practiced in discussing energy efficiency. To test whether agents learn to disclose information, we exploit the fact that the experiment induced information treatment agents to repeatedly deliver the Energy Star informational script, whereas information control agents were never directly exposed. We regress the same compliance measures from Table 4 on the interaction of information treatment agent indicator variables with a vector of treatment group indicators. Defining I_a as an indicator variable for whether agent a is assigned as an information treatment agent, the regression is

$$N_{iat} = \boldsymbol{\gamma}\mathbf{T}_i I_a + \beta\mathbf{T}_i + \boldsymbol{\phi}_t + \upsilon_{iat}. \quad (2)$$

Standard errors are clustered by agent. Table 5 presents the estimated $\boldsymbol{\gamma}$ coefficients on the interactions of I_a with call treatment assignment indicators $\mathbf{T}_i$. The table parallels Table 4, with three exceptions. First, to increase power, we combine the \$25 and \$100 rebate groups into one indicator. Second, the samples exclude information treatment calls, because information control agents do not have any information treatment calls, and the objective is to compare information provision on noninformation group calls. Third, we do not present regressions for *Payback Period*, because agents only quoted payback periods on four audited noninformation group calls. (All four involved information treatment agents.)

Table 5 shows that information treatment agents are not more likely than information control agents to mention experimental rebates on calls in any treatment group. Information treatment agents are, however, more likely to mention Energy Star on control calls and to mention that Energy Star saves money on Spiff Only calls. Column (6) shows that information treatment agents talk about Energy Star for approximately 30% longer on Spiff Only and control group calls. The standard errors are too wide to determine whether there is a meaningful difference in information provision on combination Spiff plus Rebate calls.

These results have two implications. First, they suggest that one reason that the Retailer's sales agents do not frequently discuss information is that they were not well practiced at doing so. Once information treatment agents learned how to discuss Energy Star on information treatment calls, they began to do so without explicit instruction on control and spiff calls. Second, these results imply that the estimates of effects of information disclosure on sales should take account of spillovers, in the form of increased discussion of energy efficiency on control group calls.

4.2. Effects on Consumer Choice

We examine two binary outcomes Y_{iat}: whether the consumer purchases any model from the Retailer and whether the consumer purchases an Energy Star model. For each outcome, we run three specifications. The first specification is the intent-to-treat (ITT) estimator:

$$Y_{iat} = \boldsymbol{\tau}\mathbf{T}_i + \boldsymbol{\phi}_i + \boldsymbol{\mu}_a + \varepsilon_{iat}. \quad (3)$$

As above, this includes agent and phase indicator variables.[14]

The second specification is an instrumental variables (IV) estimator, substituting agent-reported compliance N^s_{iat} for $\mathbf{T}_i$ in Equation (3) and instrumenting for N^s_{iat} with $\mathbf{T}_i$. This sample is smaller because it excludes

Table 5. Information Treatment Agents vs. Information Control Agents

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*Agent Reported*	*Mentioned E-Star*	*Rebate*	*Saves Money*	*Read Script*	ln(*E-Star Seconds*)
1(rebate)	−0.072	0.063	−0.048	0.045	0.005	0.189
	(0.098)	(0.102)	(0.085)	(0.029)	(0.100)	(0.339)
1(spiff only)		0.092	0.004	0.062	0.030	0.314
		(0.059)	(0.035)	(0.031)*	(0.023)	(0.180)*
1(spiff and rebate)		−0.143	−0.136	−0.076	−0.090	−0.555
		(0.144)	(0.109)	(0.122)	(0.084)	(0.532)
1(control)		0.109	0.003	0.044	0.007	0.320
		(0.052)**	(0.007)	(0.029)	(0.010)	(0.157)**
R^2	0.01	0.04	0.09	0.02	0.11	0.04
N	8,276	1,642	1,642	1,642	1,408	1,642
Dep. var. control agent mean	0.527	0.178	0.101	0.051	0.101	0.582

Notes. This table reports the estimates of $\boldsymbol{\gamma}$ in Equation (2). All regressions include phase and treatment group indicator variables. Robust standard errors, clustered by agent, are given in parentheses.

* and ** indicate statistical significance with 90% and 95%, confidence, respectively.

the spiff treatment calls, because N^s_{iat} is undefined for these calls. The third specification is what we call the "Scaled ITT" estimate: we interact $\mathbf{T}_i$ with G_a, which reflects the probability that agent a delivered the specific information or rebate script to consumer i. Intuitively, multiplying by compliance probability scales the τ coefficient to be equivalent to a Wald estimator of the local average treatment effect (LATE).[15]

In this context, the ITT and LATE are likely to bound the average treatment effect of providing information to all consumers. Although the ATE could theoretically be larger than the LATE if sales agents targeted information at the least responsive consumers, we shall see momentarily that agents appear to target consumers who are more interested in Energy Star. If all treatment group consumers with $N_{iat} = 0$ would have had zero treatment effect, then the ITT equals the ATE. On the other hand, if the agents quasi-randomly chose to whom to disclose, then consumers with $N_{iat} = 0$ and $N_{iat} = 1$ would have the same treatment effect, and the LATE would equal the ATE.

Table 6 presents results. Columns (1)–(3) show effects on overall sales of any model from the Retailer. Columns (4)–(6) show effects on Energy Star sales. Within each set of three columns, the first is the ITT, the second is the IV using agent-reported compliance, and the third regression is the scaled ITT. Because the interaction effects between information and the two rebate level are never statistically significant, we drop these terms and report results for six major treatment groups relative to control. At the bottom of the table, we report the mean purchase probabilities in the control group: approximately 36% of consumers purchase from the Retailer, and approximately 0.9% of consumers purchase an Energy Star model.

The treatments have no effect on overall sales, except that the ITT and self-report IV suggest that a $25 rebate may reduce purchase probability. One explanation for this is that even mentioning a small rebate for a different model generates a version of choice overload, complicating the sales interaction and causing a slight decrease in purchase probability. This would be consistent with other evidence of choice overload, such as Iyengar and Lepper (2000). Another explanation is that the result is spurious and would not replicate. We find this latter explanation more plausible, partially because there is no negative effect in the scaled ITT. Intuitively, the scaled ITT differs from the ITT because it weights more heavily the treatment effects from more compliant agents. The fact that the negative association disappears in the scaled ITT implies that the ITT effect is driven by agents who aren't actually doing anything to comply with the experiment, which suggests a spurious correlation. The first two columns of Table A.2 in the online appendix replicate column (1) for agents in the bottom two versus top two compliance groups, confirming that none of the treatments affect sales in the subsample of more compliant agents.

Although there appears to be little or no effect on the Retailer's overall sales, columns (4)–(6) show that the treatments do shift the composition of sales toward

Table 6. Treatment Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	1(*Sale*)	1(*Sale*)	1(*Sale*)	1(*E-Star*)	1(*E-Star*)	1(*E-Star*)
1($100 rebate)	−0.010	−0.024	0.030	0.006	0.012	0.037
	(0.011)	(0.022)	(0.045)	(0.003)**	(0.005)**	(0.013)***
1($25 rebate)	−0.019	−0.040	−0.015	0.001	0.002	0.003
	(0.007)***	(0.015)***	(0.029)	(0.001)	(0.003)	(0.005)
1(information)	−0.005	−0.009	−0.059	0.000	0.000	0.004
	(0.009)	(0.020)	(0.036)	(0.002)	(0.004)	(0.007)
1(spiff)	0.007		0.025	−0.002		0.001
	(0.012)		(0.042)	(0.002)		(0.007)
1(spiff and $25 rebate)	0.000		0.055	−0.004		−0.007
	(0.032)		(0.111)	(0.002)**		(0.005)
1(spiff and $100 rebate)	0.041		0.200	0.040		0.219
	(0.053)		(0.193)	(0.022)*		(0.118)*
R^2	0.02	0.02	0.02	0.01	0.01	0.01
N	23,347	20,240	23,347	23,347	20,240	23,347
Dep. var. control mean	0.364	0.364	0.364	0.009	0.009	0.009
Regression type	ITT	Self-report IV	Scaled ITT	ITT	Self-report IV	Scaled ITT

Notes. This table reports the estimates of Equation (3). The dependent variable in the first three models is an indicator for whether a caller purchased any water heater, and the dependent variable in the last three models is an indicator for whether a caller bought an Energy Star water heater. All regressions contain agent and period indicator variables. Robust standard errors are given in parentheses.

*, **, and *** indicate statistical significance with 90%, 95%, and 99% confidence, respectively.

Energy Star. The $100 rebate increases Energy Star purchase probability by 0.006 to 0.037 percentage points in the three specifications.[16] This effect is large against a control group market share of 0.9%, although it is small as a share of the total potential market. The combination of a spiff and $100 rebate has a very large point estimate, increasing purchase probability by 4–22 percentage points. The standard errors are wide, because this treatment was only offered to a small share of consumers in phase 3 of the experiment.

The information treatments have a tightly estimated zero statistical effect. In column (6), the standard errors are tight enough to bound the local average treatment effect at less than 1.5 percentage points. Although this would represent a large percent increase in Energy Star purchase probability given the small base, the effect in percentage point terms is economically small.

Earlier, we documented that, although agents effectively never explicitly read a treatment script on a control call, information treatment agents are more likely to mention Energy Star on control calls. If this increases the probability of Energy Star purchase, then the estimates in Table 6 would understate the true effects of the information scripts. The audit data allow us to address such spillovers. We construct an alternative "Scaled ITT" estimator, interacting $\mathbf{T}_i$ with each compliance group's average difference in *Mentioned E-Star* between treatment and control calls. Intuitively, this scales the treatment effect to equal a LATE with *Mentioned E-Star* as the endogenous variable. The third column of Table A.2 in the online appendix presents the results. The coefficients and standard errors are inflated, as one should expect from the fact that the coefficients in Table 4 are smaller for *Mentioned E-Star* than for the variables used to construct G_{ia}. The qualitative results are similar: the Spiff and $100 Rebate combination still has a very large effect, and the standard errors bound the information effect at no more than 2.7% with 90% confidence.

As discussed in §2.3, Energy Star models are only substitutable with some standard models. Under the assumption that the treatments do not affect whether or not consumers purchase a substitutable model, we can consistently estimate treatment effects within the sample to the set of consumers who purchase substitutable models. Table A.3 in the online appendix presents results. The coefficients are larger than in Table 6, as one should expect from excluding consumers with smaller treatment effects.[17] The qualitative results are also similar: the Spiff and $100 Rebate combination has a very large effect, and the standard errors bound the information effect at no more than 4.9 percentage points with 90% confidence in the Scaled ITT. For comparison, Energy Star represents 3.4% of substitutable models. The standard errors suggest that, even if sales agents provided information to all callers, Energy Star would still not represent more than 4.9 + 3.4 = 8.3% of the market of substitutable models. Thus, the lack of seller-provided information does not explain much of the low takeup of energy-efficient water heaters.

Table A.4 in the online appendix presents alternative estimates using a probit estimator; the signs and discrete significance levels are the same or stronger.

4.3. Targeted Information Disclosure

Table 4 documents that the sales agents only partially comply with the experiment. Are agents strategic in providing information to consumers who are more interested in energy efficiency? Recall that N^s_{iat} is an indicator for whether the consumer is in a treatment group and the agent reported compliance with delivering the script, and define T_i as an indicator for whether consumer i is in any rebate or informational treatment group. We exclude the spiff treatment calls because there was no script for the agent to "comply" with on these calls. Table 7 reports estimates of the following regression:

$$Y_{iat} = \xi T_i N^s_{iat} + \kappa T_i (1 - N^s_{iat}) + \phi_i + \mu_a + \varepsilon_{iat}. \quad (4)$$

As outcome variables Y_{iat}, we use two different measures of interest in energy efficiency, both of which could be affected by the treatments. To construct the dependent variable in column (1), we exploit an open-ended question from the follow-up survey: "What were the two most important factors in your water heater purchase decision?" The dependent variable is an indicator taking value 1 for the consumers who had one of their two factors coded as "saving energy and/or environmental conservation." The dependent variable in columns (2)–(4) is an indicator for whether the consumer purchased Energy Star. These regressions use agent-reported compliance because the sample of audits is too small for sufficient power.

The coefficient κ measures the difference in Y_{iat} between treatment group consumers who did not

Table 7. Targeted Information Provision

	(1)	(2)	(3)	(4)
Dependent variable:	1(*Factor*)	1(*E-Star*)	1(*E-Star*)	1(*E-Star*)
$T \times$ *Agent-Reported Compliance*	−0.014 (0.030)	0.013 (0.002)***	0.038 (0.007)***	0.034 (0.006)***
$T \times (1 -$ *Agent-Reported Compliance*)	−0.057 (0.021)***	−0.009 (0.001)***	−0.031 (0.005)***	−0.022 (0.004)***
R^2	0.18	0.02	0.05	0.05
N	404	20,240	5,180	6,123
Dep. var. control mean	0.061	0.009	0.033	0.025

Notes. This table reports the estimates of Equation (4). All regressions contain agent and period indicator variables. Robust standard errors are given in parentheses.

***Indicates statistical significance with 99% confidence.

receive the treatment and the control group, and ξ measures the difference between treatment group consumers who did receive the treatment (according to the agents' self-reports) and the control group. ξ is a mix of selection effects and treatment effects, whereas κ is purely selection effects. If $\kappa < 0$, this implies that the consumers to whom agents are not disclosing information are less likely than average to purchase Energy Star, and thus that consumers who are provided with information are more likely than average to purchase Energy Star. Although this is not the same as targeting information disclosure at consumers who will have the largest treatment effects, it implies targeting in a different and likely related sense.

In all four columns of Table 7, the statistically negative estimates of κ show that agents are more likely to report delivering the script to consumers who are more interested in energy efficiency. Column (3) includes in the sample only consumers who purchase substitutable models. Even within this group, agents still target consumers with a higher probability of purchasing Energy Star. Column (4) includes controls for the nine address-based demographic variables. Because these covariates are missing for consumers who did not purchase from the Retailer, the sample is limited to consumers who purchased from the Retailer. The fact that κ is still negative after conditioning on observable factors shows that agents target information provision based on other unobservable factors.

How is such targeting possible? Based on conversations with sales agents and with our research assistant who carried out the audits, we believe that agents learn about the consumer's preferences as the call continues. Agents may have delayed reading the Energy Star script until later in the call, after having the chance to gauge the consumer's receptiveness.

4.4. Relationship to a Model of Information Disclosure

In the online appendix, we include a simple theoretical model that helps to motivate the experiment and to interpret results. The model is a two-firm version of Grossman and Shapiro's (1984) analysis of informational advertising in a Hotelling spatial model. We extend this framework to include two goods, a base good and a "high-quality" (energy-efficient) good, and two consumer types, "high" and "low." All consumers are initially uninformed about the high-quality good, but while the low type always prefers the base good, the high type will purchase the high-quality good in equilibrium when provided information. Nested within the firm is a set of optimizing sales agents who cannot observe consumer type but can provide quality information to an increasing share of consumers at convex cost.

The model highlights the importance of interactions between the firm's management, sales agents, and consumers. Because information provision is costly, sales agents will not inform many consumers if they know that there is little interest in the high-quality good; interest will be low if either there are few high types or a firm has set a high price relative to its competitor. If, as in the experiment, one firm lowers the high-quality good's price out of equilibrium, more high type consumers would then buy from that firm, and its sales agents exert more effort in marketing it.

This logic is formalized in three simple propositions that link closely to our empirical findings. Proposition 1 states that if information has little impact on demand, then agents will not incur the cost to inform many consumers. Results in Table 6 show that information increases Energy Star sales by at most a few percentage points when delivered. As Proposition 1 would then predict, the audit data show that agents spent very little time discussing Energy Star unless the experiment website directed them to. Agents mentioned Energy Star on only 13% of control group calls, and the median time spent discussing Energy Star within this 13% was 15 seconds.

Proposition 2 states that agents will provide more information when incentivized to do so, but they will not be very responsive to sales incentives if information has small effects on demand. Consistent with this, the audit data show that agents provide at best slightly more information about Energy Star on no-rebate spiff calls compared to control: in Table 4, the coefficient on 1(Spiff Only) is positive but not statistically significant in columns (2), (3), (5), and (7).

Proposition 3 states that sales incentives and customer rebates reinforce each other, because the rebate increases the share of high type consumers who would buy from the firm instead of the competitor when informed, which increases the firm's agents' gains from providing information in response to the sales incentive. Although the standard errors are wide, the point estimates in the sales data in Table 6 strongly support this: the combined Spiff and $100 rebate treatment substantially increases Energy Star purchase probability. The audit data provide evidence that is suggestive of the microfoundations for this result. Specifically, agents appear to be slightly more likely to market Energy Star on the spiff treatment calls as the rebates increase from $0 to $25 to $100: in Table 6, the point estimates in columns (2), (3), (4), (6), and (7) all increase as the rebates increase.

We also use the model to consider the potential role for government-provided sales incentives. There is a subset of parameter space where agents under-provide information in equilibrium relative to the social optimum. In this case, a social planner can increase welfare by directly incentivizing sales agents to sell the

high-quality good. As long as firms cannot "undo" the government incentive through a change in agents' contract structure, agents respond to the social planner's incentive by increasing information provision, which increases welfare. This idea of a government-provided sales incentive for energy-efficient goods could perhaps be a complement to the large product subsidies traditionally offered by governments and utility energy-efficiency programs. The empirical results in Table 6 suggest that reallocating part of these large product subsidies to sales agent incentives could substantially increase sales of Energy Star models, because there is a strong complementarity between price reductions and moderate sales incentives.[18]

Our empirical parameter estimates map to several of the model's theoretical parameters, including the information disclosure cost, the share of high type consumers, and the "transport cost," which maps to the Retailer's residual demand slope identified by the (unilateral) experimental price reduction. Although our exact parameter estimates surely are difficult to generalize, the more basic conceptual points about the equilibrium importance of sales agent incentives likely have "local" or "global" generalizability in the framework of Al-Ubaydli and List (2012).

5. Why Is Demand for Energy Star Low?

The theoretical model in Online Appendix A shows that a minimum energy-efficiency standard is more likely to increase welfare if the energy-efficient product generates utility gains for a larger share of the population or if the cost of providing information is high. Our empirical results show that the experimental attempts at information provision do not significantly increase Energy Star demand. Does this suggest that the information provision cost is high, and thus that minimum energy-efficiency standards might increase welfare by addressing imperfect information?

A first potential explanation for our empirical results is that although the Energy Star product would generate gains for many consumers, the Retailer is not able to credibly inform consumers of this, perhaps because of time constraints on the sales interactions or an inability to credibly convey attributes of this higher-priced product. If this were true, many consumers would remain unaware of Energy Star and unconvinced of the cost savings. A second explanation is that the Energy Star model would not in fact generate utility gains for many consumers, and consumers make an informed decision to not purchase the product. If this were true, consumers would be aware of the Energy Star product and the potential energy cost savings.

We use results from the follow-up surveys to measure the importance of these two explanations. Near the end of the survey, we asked consumers, "Some water heater models that use less energy are officially designated as Energy Star. Did you buy an Energy Star model?" The results suggest substantial confusion, consistent with the first explanation. The top panel of Table 8 shows that although only 2.1% actually purchased Energy Star, 52% of survey respondents think that they did. This should be interpreted cautiously, because there could have been experimenter demand effects: respondents may not have wanted to tell the interviewer that they had not purchased the Energy Star option. We designed the survey specifically to reduce demand effects: this was the first question we asked about energy efficiency, and we asked it after a series of other questions that signaled that the interviewer was not particularly interested in energy.

Respondents who said they had not purchased Energy Star, or who were "Not Sure" but in fact had not, were asked why they had not purchased Energy Star. The bottom panel of Table 8 shows that the primary reason was high prices. This is consistent with the second explanation of informed consumers choosing not to purchase Energy Star. Fifteen percent of these consumers (or approximately 7% of the entire surveyed population) report that they were unaware that there was an Energy Star option.

The follow-up survey also elicited beliefs over energy costs for standard models and energy cost savings from Energy Star models. Table 9 shows the 10th, 50th, and 90th percentiles of beliefs, along with the

Table 8. Survey Results: Energy Star Purchases

Self-reported Energy Star purchase *Some water heater models that use less energy are officially designated as Energy Star.* *Did you buy an Energy Star model?*	
Response	%
Yes	52
No	24
Not sure	24
True Energy Star share in survey sample	2.1
Reasons to not purchase Energy Star *Why did you decide to buy a standard, non-Energy Star model over an Energy Star model?*	
Response	%
Upfront price too high	33.3
I was not aware that there was an Energy Star option	15.4
Energy Star not in stock	8
Needed a short tank	5.9
No electrical outlet	2.8
Needed low NOx	2.3
Wanted longer warranty	1.4
Non-Energy Star heats water faster	1.2
Other	30.7

Notes. This table presents responses to two questions from our follow-up survey of consumers who purchased from the Retailer. Sample size for the first question is 891. The second question was asked only of people who thought they had not purchased an Energy Star model or were not sure but in fact had not; sample size is 423.

Table 9. Survey Results: Beliefs About Energy Star

10th	50th	90th	Mean	Yellow tag
How much money do you think the natural gas for the water heater will cost each year?				
$50	$200	$600	$305	Approx. $300
How much less money do you think the natural gas would cost each year for an Energy Star water heater compared to a similarly sized non-Energy Star water heater?				
$0	$50	$300	$129	Approx. $30
Implied percent savings from Energy Star				
5%	25%	67%	32%	Approx. 10%

Notes. This table presents responses to three questions from our follow-up survey of consumers who purchased from the Retailer. Sample size is 891.

mean and the best estimate of the true value, from the yellow tags. The first question in Table 9 shows that consumers' mean beliefs about water heater energy costs are approximately in line with the yellow tags. The second question, however, suggests that the mean and median consumers actually overestimate the average dollar value of potential cost savings. The third question takes the ratio of the second question to the first, which translates to a percent savings and can account for heterogeneity in consumers' utilization. The "true" mean energy cost savings on the yellow tags are approximately 10%, whereas the median and mean consumers report believing that Energy Star could save 25% and 32%, respectively.[19] Of course, the beliefs were elicited in a phone survey and not made incentive compatible, so they should be interpreted cautiously. Notwithstanding, they suggest that underestimating energy cost savings is not a barrier to Energy Star takeup.

6. Conclusion

Imperfect information is an often cited reason that regulators intervene in markets for energy-using durables through mandatory information disclosure, subsidies, and standards. In theory, one natural way for consumers to learn about the benefits of energy-efficient products is through retailers. Motivated by this, we partnered with a large nationwide retailer to test the effects of information provision, customer rebates, and sales incentives on the behavior of both sales agents and consumers. Results show that retailer-provided information is ineffective at increasing Energy Star demand, even after adjusting for partial compliance by sales agents. Knowing that information would be ineffective, sales agents appear to have marketed Energy Star to only the most receptive consumers, strategically failing to provide information to the majority. Follow-up surveys provide some evidence that at least some consumers who bought from the Retailer are unaware of the Energy Star product or confused about what they have bought. The majority of consumers who bought from the Retailer and completed our survey, however, are aware of Energy Star and may even overestimate its benefits. These results highlight the difficulties that retailers can face in increasing demand for energy-efficient or otherwise high-quality products. In this context, a key difficulty appears to be that many consumers still do not view energy efficiency as a privately-beneficial investment, even after the retailer's attempts to inform them.

The water heater market has some unusual features, and our sample comprises only a small share of that market. What broader lessons can be drawn from our results? First, our results fit into a broader set of results from the literature suggesting that, although social comparisons or various forms of persuasion can motivate proenvironmental behaviors, "hard" information about energy costs seems to have more limited effects. Such hard information about product availability (as well as product costs and benefits) is more relevant from a welfare perspective because a regulator's inability to address imperfect information could justify corrective policies, whereas the welfare implications of persuasion and marketing are less clear.

Second, our results provide unusually granular insight into the role of sales agents in the process of information disclosure. In many situations like ours, a firm or policymaker must consider how to incentivize agents if they want to increase information disclosure. In equilibrium, firms and policymakers can also potentially learn about the value of information from whether experienced sales agents choose to provide it.

Acknowledgments

The authors thank Judd Boomhower, Lucas Davis, Stefano DellaVigna, Carolyn Fischer, Ginger Jin, Katy Milkman, Erich Muehlegger, David Rapson, Imran Rasul, and Jim Sallee, as well as seminar participants at the 2014 Allied Social Science Associations (ASSA) Annual Meeting, University of California at Berkeley, Harvard University, and the Program on Workable Energy Regulation conference for helpful comments. Alex Cheng provided exceptional research assistance, and Larry Ehrhardt programmed the experiment website. The authors' legal agreement does not allow them to identify the Retailer or its employees, but the authors thank sales agents and management, in particular P.C. and J.W. Customer follow-up survey protocols and Stata code to replicate the analysis are available from Hunt Allcott's website.

Endnotes

[1] There are many examples. The American Council for an Energy-Efficient Economy argues that minimum efficiency standards are merited for several reasons, including "rush purchases when an existing appliance breaks down, providing no time to comparison shop" (U.S. Senate 2011). The regulatory impact analysis for the increase in the Corporate Average Fuel Economy standard for 2012–2016 argues that, even without counting the value of externality reductions, the regulation increases consumer welfare, perhaps because consumers are not correctly informed about the value of fuel

economy (NHTSA 2010, p. 2). The regulatory impact statement for Australia's ban on energy-inefficient lightbulbs argues that "information failures" help to justify that policy (DEWHA 2008, p. vii).

[2] Empirical papers on the effects of information disclosure include Choi et al. (2010), Duarte and Hastings (2012), and Duflo and Saez (2003) on financial decisions; Greenstone et al. (2006) on securities; Bhargava and Manoli (2013) on takeup of social programs; Jin and Sorensen (2006), Kling et al. (2012), and Scanlon et al. (2002) on health insurance plans; Jin and Leslie (2009) on restaurant hygiene; Pope (2009) on hospitals; Bollinger et al. (2011) on health and nutrition; Dupas (2011) on HIV risk; Figlio and Lucas (2004) and Hastings and Weinstein (2008) on school choice; and many others.

[3] There are other large-sample randomized controlled trials (RCTs) that study how peer energy use comparisons affect purchases of durable goods, including Allcott and Rogers (2014), Brandon et al. (2014), and Herberich et al. (2011), but this social information is conceptually distinct from information about the durable good itself.

[4] For example, Ashraf et al. (2014), Barankay (2012), Bandiera et al. (2005), Bandiera et al. (2007), Bandiera et al. (2009, 2010), Larkin (2014), Lazear (2000), Shearer (2004), and others focus on how employees respond to different types of incentives, but these responses do not meaningfully depend on behavior by the firm's customers.

[5] A handful of other papers have highlighted other types of equilibrium interactions between the supply and demand sides of the market for energy efficiency, including Fischer (2005), Fischer (2010), and Houde (2014b).

[6] U.S. Department of Energy (2010) provides an overview of the U.S. water heater market; this is the source of most of the statistics presented here.

[7] In addition to these key characteristics, some consumers need niche models that fit in mobile homes or that push vented gases through a horizontal (rather than a vertical) pipe. In the San Francisco Bay Area and southern California, air quality regulations require consumers to purchase water heaters with lower NOx emissions.

[8] Because of the fixed costs of developing, certifying, and manufacturing a unique model, manufacturers do not produce Energy Star versions of each standard water heater. Furthermore, because of the fixed costs of procuring and stocking each model, the Retailer also does not carry all the Energy Star models that the manufacturer offers.

[9] Other than perhaps these 104 individuals, the consumers who call the call center are the final owners of the water heater; contractors do not order through the retail hotline. Furthermore, consumers typically are not already working with a contractor before calling the Retailer, because a contractor would typically procure the water heater on behalf of a consumer.

[10] Approximately 2%–3% of reference numbers are repeated, typically as the sales agent updates information. In these cases, we use the most recent observation. We drop six reference numbers that appear to be used twice for two distinct individuals.

[11] There are other ways to code this, but it would not matter, because almost all of the 1.3% of reference numbers that appear multiple times were entered by the same agent. Only four reference numbers were entered by two separate agents in the website, and none are entered by three or more. Because treatment groups were assigned by reference number, a consumer's treatment assignment is maintained even if an agent enters the same reference number multiple times.

[12] One could imagine reasons that in-store interventions could be more or less effective. On one hand, the phone sales interaction is very focused, and the customer relies heavily on the sales agent for advice. Furthermore, the store environment has yellow tags and other information that could make the control group better informed and incremental information less effective. On the other hand, some phone customers are also looking at Internet resources that could include energy information. Consumer types could also be different between phone and in-store populations.

[13] By necessity, this was done before writing the paper, because the Retailer deletes call recordings after 30 days.

[14] Once we control for phases, additional time controls do not improve consistency or efficiency. Adding month-of-sample indicators, for example, does not change coefficients or standard errors. Furthermore, the estimates do not change when we exclude phase 1, which pre-dated the introduction of the spiff.

[15] The Wald estimator is the reduced form coefficient divided by the first-stage coefficient. If G_a is a first-stage coefficient, multiplying $\mathbf{T}_i$ by G_a means that the τ coefficient estimates equal the Wald estimates. Although G_a is a mean calculated with sampling error using the audit data, we calculate that adjusting for sampling error in this generated regressor would have only a small impact on the standard errors. N^+_{iat} is not statistically different across phases or treatment groups, so we cannot increase precision by also projecting ϕ_i or $\mathbf{T}_i$ onto G_a.

[16] In column (4), the combination of the spiff and a $25 rebate appears to reduce demand for Energy Star. We suspect that this result also would not replicate, and it is not statistically significant in the scaled ITT in column (6).

[17] Energy Star model availability and consumer preferences vary by geography, so if geography were somehow imbalanced across treatment groups, our coefficient estimates would be biased. Table A.3 in the online appendix also shows that the coefficient estimates are very similar when also including indicators for each purchaser's three-digit zip code. Of course, this is to be expected in a randomized experiment.

[18] Of course, this result is only suggestive. Consumer rebates versus sales agent incentives likely have nonlinear demand effects, and we have tested one firm's unilateral price reduction instead of a market-wide subsidy. An additional important caveat is that the model does not include the possibility that agents would provide deceptive or misleading information that could increase Energy Star market share while decreasing consumer welfare.

[19] Additional (unreported) regressions show that confusion about Energy Star purchases and beliefs about Energy Star savings do not vary across treatment groups, although estimates are imprecise.

References

AHRI (Air-Conditioning, Heating, and Refrigeration Institute) (2013) Residential Automatic Storage Water Heaters Historical Data. Accessed February 5, 2014, http://www.ahrinet.org/site/495/Resources/Statistics/Historical-Data/Residential-Storage-Water-Heaters-Historical-Data.

Allcott H (2011) Social norms and energy conservation. *J. Public Econom.* 95(9–10):1082–1095.

Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *Amer. Econom. Rev.* 104(10):3003–3037.

Allcott H, Taubinsky D (2015) Evaluating behaviorally-motivated policy: Experimental results from the lightbulb market. *Amer. Econom. Rev.* 105(8):2501–2538.

Al-Ubaydli O, List J (2012) On the generalizability of experimental results in economics. Frechette G, Schotter A, eds. *The Methods of Modern Experiments* (Oxford University Press, Oxford, UK).

Anagol S, Cole S, Sarkar S (2013) Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance market. HBS Working Paper 12-055, Harvard Business School, Boston.

Anderson D, Claxton J (1982) Barriers to consumer choice of energy efficient products. *J. Consumer Res.* 9(2):163–170.

Angrist J, Pischke J-S (2012) Probit better than LPM? Accessed January 12, 2016, http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/.

Ashraf N, Bandiera O, Jack K (2014) No margin, no mission? A field experiment on incentives for pro-social tasks. *J. Public Econom.* 120:1–17.

Bandiera O, Barankay I, Rasul I (2005) Social preferences and the response to incentives: Evidence from personnel data. *Quart. J. Econom.* 120(3):917–962.

Bandiera O, Barankay I, Rasul I (2009) Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica* 77(4):1047–1094.

Bandiera O, Barankay I, Rasul I (2010) Social incentives in the workplace. *Rev. Econom. Stud.* 77(2):417–458.

Bandiera O, Rasul I, Barankay I (2007) Incentives for managers and inequality among workers: Evidence from a firm level experiment. *Quart. J. Econom.* 122:729–775.

Barankay I (2012) Rank incentives: Evidence from a randomized workplace experiment. Working paper, University of Pennsylvania, Philadelphia.

Bhargava S, Manoli D (2013) Why are benefits left on the table? Assessing the role of information, complexity, and stigma on take-up with an IRS field experiment. Working paper, University of Texas at Austin, Austin.

Bureau of Labor Statistics (Bureau of Labor Statistics) (2011) Consumer expenditure survey: 2011 expenditure tables. Accessed January 12, 2016, http://www.bls.gov/cex/csxstnd.htm.

Bollinger B, Leslie P, Sorensen A (2011) Calorie posting in chain restaurants. *Amer. Econom. J.: Econom. Policy* 3(1):91–128.

Brandon A, List J, Metcalfe R, Price M (2014) What drives the adoption of energy efficient technologies? Theroleof social information and advertisements. Working paper, University of Chicago, Chicago.

Chan T, Li J, Pierce L (2014) Compensation and peer effects in competing sales teams. *Management Sci.* 60(8):1965–1984.

Chatterji A, Findley M, Jensen N, Meier S, Nielson D (2016) Field experiments in strategy research. *Strategic Management J.* 37(1):116–132.

Choi J, Laibson D, Madrian B (2010) Why does the law of one price fail? An experiment on index mutual funds. *Rev. Financial Stud.* 23(4):1405–1432.

Davis L, Metcalf G (2014) Does better information lead to better choices? Evidence from energy-efficiency labels. NBER Working Paper 20720, National Bureau of Economic Research, Cambridge, MA.

Deutsch M (2010a) Life cycle cost disclosure, consumer behavior, and business implications: Evidence from an online field experiment. *J. Indust. Ecology* 14(1):103–120.

Deutsch M (2010b) The effect of life-cycle cost disclosure on consumer behavior: Evidence from a field experiment with cooling appliances. *Energy Efficiency* 3(4):303–315.

DEWHA (Australian Department of the Environment, Water, Heritage, and the Arts) (2008) Regulatory impact statement: Proposal to phase-out inefficient incandescent light bulbs. Accessed July 21, 2014, http://www.energyrating.gov.au/wp-content/uploads/Energy_Rating_Documents/Library/Lighting/Incandescent_Lamps/200808-ris-phaseout.pdf.

Dolan P, Metcalfe R (2013) Neighbors, knowledge, and nuggets: Two natural field experiments on the role of incentives on energy conservation. CEP Discussion Paper No. 1222, London School of Economics and Political Science, London.

Dranove D, Jin GZ (2010) Quality disclosure and certification: Theory and practice. *J. Econom. Literature* 48(4):935–963.

Duarte F, Hastings J (2012) Fettered consumers and sophisticated firms: Evidence from Mexico's privatized social security market. NBER Working Paper 18582, National Bureau of Economic Research, Cambridge, MA.

Duflo E, Saez E (2003) The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quart. J. Econom.* 118(3):815–842.

Duflo E, Gale W, Liebman J, Orszag P, Saez E (2006) Saving incentives for low- and middle-income families: Evidence from a field experiment with H&R Block. *Quart. J. Econom.* 121(4):1311–1346.

Dupas P (2011) Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *Amer. Econom. J.: Appl. Econom.* 3(1):1–34.

Figlio D, Lucas ME (2004) What's in a grade? School report cards and the housing market. *Amer. Econom. Rev.* 94(3):591–604.

Fischer C (2005) On the importance of the supply side in demand-side management. *Energy Econom.* 27(1):165–180.

Fischer C (2010) Imperfect competition, consumer behavior, and the provision of fuel efficiency in light-duty vehicles. RFF Discussion Paper 10-60, Resource for the Future, Washington, DC.

Greenstone M, Oyer P, Vissing-Jorgensen A (2006) Mandated disclosure, stock returns, and the 1964 securities acts amendments. *Quart. J. Econom.* 121(2):399–460.

Grossman G, Shapiro C (1984) Informative advertising with differentiated products. *Rev. Econom. Stud.* 51(1):63–81.

Grossman SJ (1981) The informational role of warranties and private disclosure about product quality. *J. Law Econom.* 24(3):461–483.

Hastings J, Weinstein J (2008) Information, school choice, and academic achievement: Evidence from two experiments. *Quart. J. Econom.* 123(4):1373–1414.

Herberich D, List J, Price M (2011) How many economists does it take to change a light bulb? A natural field experiment on technology adoption. Working paper, University of Chicago, Chicago.

Hoffman F, Inderst R, Ottaviani M (2013) Hypertargeting, limited attention, and privacy: Implications for marketing and campaigning. Working paper, Bocconi University, Milan.

Houde S (2014a) How consumers respond to environmental certification and the value of energy information. Working paper, University of Maryland, College Park.

Houde S (2014b) Bunching with the stars: How firms respond to environmental certification. Working paper, University of Maryland, College Park.

Ichniowski C, Shaw K (2003) Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *J. Econom. Perspect.* 17(1):155–180.

Inderst R, Ottaviani M (2009) Misselling through agents. *Amer. Econom. Rev.* 99(3):883–908.

Inderst R, Ottaviani M (2012) Financial advice. *J. Econom. Literature* 50(2):494–512.

Iyengar S, Lepper M (2000) When choice is demotivating: Can one desire too much of a good thing? *J. Personality Soc. Psych.* 79(6):995–1006.

Jessoe K, Rapson D (2014) Knowledge is (less) power: Experimental evidence from residential energy use. *Amer. Econom. Rev.* 104(4):1417–1438.

Jin GZ, Leslie P (2009) Reputation incentives for restaurant hygiene. *Amer. Econom. J.: Microeconomics* 1(1):237–267.

Jin GZ, Sorensen A (2006) Information and consumer choice: The value of publicized health plan ratings. *J. Health Econom.* 25(2):248–275.

Kallbekken S, Saelen H, Hermansen E (2013) Bridging the energy efficiency gap: A field experiment on lifetime energy costs and household appliances. *J. Consumer Policy* 36(1):1–16.

Kahn M (2007) Do greens drive hummers? Environmental ideology as a determinant of consumer choice. *J. Environmental Econom. Management* 54(2):129–145.

Kling J, Mullainathan S, Shafir E, Vermeulen L, Wrobel M (2012) Comparison friction: Experimental evidence from medicare drug plans. *Quart. J. Econom.* 127(1):199–235.

Larkin I (2014) The cost of high-powered incentives: Employee gaming in enterprise software sales. *J. Labor Econom.* 32(2): 199–227.

Lazear E (2000) Performance pay and productivity. *Amer. Econom. Rev.* 90(5):1346–1361.

Leip D (2013) Dave Leip's atlas of U.S. presidential elections. Accessed January 12, 2016, http://uselectionatlas.org/.

Mantrala M, Albers S, Caldieraro F, Jensen O, Joseph K, Krafft M, Narasimhan C, Gopalakrishna S, Zoltners A, Lal R, Lodish L (2010) Sales force modeling: State of the field and research agenda. *Marketing Lett.* 21(3):255–272.

Milgrom P (1981) Good news and bad news: Representation theorems and applications. *Bell J. Econom.* 12(2):380–391.

Milgrom P (2008) What the seller won't tell you: Persuasion and disclosure in markets. *J. Econom. Perspect.* 22(2):115–132.

Mullainathan S, Noeth M, Schoar A (2012) The market for financial advice: An audit study. NBER Working Paper 17929, National Bureau of Economic Research, Cambridge, MA.

Nagin D, Rebitzer J, Sanders S, Taylor L (2002) Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *Amer. Econom. Rev.* 92(4):850–873.

NHTSA (National Highway Traffic Safety Administration) (2010) Final regulatory impact analysis: Corporate average fuel economy for MY 2012-MY 2016 passenger cars and light trucks. Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis, National Highway Traffic Safety Administration, Washington, DC.

Newell R, Siikamaki J (2013) Nudging energy efficiency behavior: The role of information labels. Discussion Paper 13-17, Resources for the Future, Washington, DC.

Pope D (2009) Reacting to rankings: Evidence from "America's best hospitals." *J. Health Econom.* 28(6):1154–1165.

Scanlon D, Chernew M, McLaughlin C, Solon G (2002) The impact of health plan report cards on managed care enrollment. *J. Health Econom.* 21(1):19–41.

Shearer B (2004) Piece rates, fixed wages, and incentives: Evidence from a field experiment. *Rev. Econom. Stud.* 71(2):513–534.

U.S. Department of Energy (2009) Residential Energy Consumption Survey. Accessed 12, 2016, http://www.eia.gov/consumption/residential/data/2009/.

U.S. Department of Energy (2010) Energy Star water heater market profile. Report, U.S. Department of Energy, Washington, DC. https://www.energystar.gov/ia/partners/prod_development/new_specs/downloads/water_heaters/Water_Heater_Market_Profile_2010.pdf.

U.S. Senate Committee on Energy and Natural Resources (2011) Hearing on Appliance Standards Regulation: Testimony of Steven Nadel, Executive Director, American Council for an Energy-Efficient Economy (ACEEE). Senate Hearing 112-10, 112th Cong., 1st sess., March 10, 2011.

Viscusi WK (1978) A note on "lemons" markets with quality certification. *Bell J. Econom.* 9(1):277–279.

Ward D, Clark CD, Jensen KL, Yen ST, Russell CS (2011) Factors influencing willingness-to-pay for the ENERGY STAR label. *Energy Policy* 39(3):1450–1458.