Title
Human Behavior and Intelligent Energy Metering Systems: Experimental Approaches

Author
Asensio, Omar Isaac

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Human Behavior and Intelligent Energy Metering Systems:

Experimental Approaches

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Environmental

Science and Engineering

by

Omar Isaac Asensio

2015

ABSTRACT OF THE DISSERTATION

Human Behavior and Intelligent Energy Metering Systems:

Experimental Approaches

by

Omar Isaac Asensio

Doctor of Environmental Science and Engineering

University of California, Los Angeles, 2015

Professor Magali A. Delmas, Chair

The use of field experiments and randomized controlled trials offer rich sources of inquiry to uncover causal mechanisms in the social and behavioral sciences. When these approaches are further integrated with the latest advances in engineering and information technologies, the result is an integrated research agenda that can shape new directions for innovation, science and public policy.

This dissertation combines three essays on the use of experimental methods in the study human decision making with advanced technologies. The focus of this work is on demand side innovation for energy efficiency and conservation. We engage both business and residential consumers in energy efficiency and conservation decisions, using information-based strategies, smart metering technologies, and finally grand challenges as a policy mechanism. We investigate how information changes the behavior of consumers, households and firms, advancing the literature on non-monetary incentives for behavior change and making theoretical advances on altruistic motivations for energy conservation behavior.

The dissertation of Omar Isaac Asensio is approved.

Barbara S. Lawrence

Hilary Godwin

Stephanie S. Pincetl

Magali A. Delmas, Committee Chair

University of California, Los Angeles

2015

This dissertation is dedicated to my mother and father.

CONTENTS

Abstract ii
List of Figures viii
List of Tables ix
List of Abbreviations x
Preface xi
Acknowledgments xiv
Vita xvi

I. Information strategies and energy conservation behavior 1
1.1 Introduction 2
1.2 Understanding levers for energy conservation behavior 5
1.2.1 Energy feedback 5
1.2.2 Information on problem solving strategies 6
1.2.3 Pecuniary strategies 7
1.2.4 The power of norms 8
1.3 Methods 9
1.3.1 Data collection 9
1.3.2 Overview of meta-analysis methodology 11
1.3.3 The meta-regression model 12
1.3.4 Measures 13
1.3.4.1 Dependent variable 13
1.3.4.2 Independent variables 13
1.4 Results 15
1.4.1 Descriptive statistics 15
1.4.2 Average treatment effect 17
1.4.3 Outcomes of different strategies 20
1.5 Discussion 23
1.6 Conclusion 26

References 28
Appendix 1 List of papers included in meta-analysis 33
Appendix 2 List of journals included in meta-analysis 35

II. Nonprice incentives and energy conservation 37
2.1 Introduction 38
2.2 Health externalities: a missing link in consumer choice 39
2.3 Experimental evidence 40
2.4 Measuring conservation behavior 41
2.5 Results and discussion 44
2.5.1 Heterogeneous effects on households 47
2.5.2 Appliance-level behavior 49
2.5.3 Implications for load shifting 50
2.5.4 Attitude-behavior gap 51
2.6 Policy implications 53

References 55

SI Appendix 1 Materials and Methods 59
Field site 59
Treatment messages 61
Participant recruitment 63
Empirical strategy 64
Identification 65
Dependent variable 65
Independent variables 67
Seasonality and time trends 67
Econometric model 68
Baseline characteristics 75
Robustness checks 76
Comments on external validity 78

SI References 81, 88

Appendix 2 Proof of identification 82
Appendix 3 Theoretical model 84
Appendix 4 Supplement on appliance dynamics 89
Appendix 5 Supplement on conservation messaging 91

III. Evaluating performance in societal grand challenges 104

3.1 Introduction 105
3.2 Grand challenges: the demand side 106
3.3 Impact evaluation 107
3.4 Los Angeles Better Buildings Challenge 109
3.5 Data Sources 110
3.6 Results & Discussion 111
3.7 Conclusions 123

References 125

LIST OF FIGURES

I.

Figure 1 Funnel plot of conservation effect size vs. sample size 20
Figure 2 Effect size by publication year 20

II.

Figure 1 Effects of informational messages on study households 45
Figure 2 Quantile treatment effects 48
Figure 3 Appliance-level electricity measurements 49
Figure S1 View of health based messaging feedback for consumers 60
Figure S2 Time series of hourly consumption 62
Figure S3 Peak daily consumption for the community 62
Figure S4 Effects of parameters on hypothetical model of energy saving bursts 87
Figure S5 Dynamic treatment effects at the appliance level 89
Figure S6 Health and financial savings are the top ranking messages 95
Figure S7 Rank ordering of messages on comprehension, believability and relevance 98

III.

Figure 1 Distribution of commercial building types 113-115
Figure 2 Estimated propensity scores 119
Figure 3 Bias reduction in propensity score matched samples 121

List of Tables

I.	Table 1	Descriptive statistics	17
	Table 2	Correlations	17
	Table 3	Summary of treatment effects	19
	Table 4	Meta-regression results	21
II.	Table S1	Treatment messages	61
	Table S2	Comparison of participating versus non-participating households	64
	Table S3	Distributions of dependent variables (hourly kWh measurements)	66
	Table S4	Heterogeneous treatment effects on families with children	70
	Table S5	Treatment effects by appliance	71
	Table S6	Treatment effects by time of day	72
	Table S7	Means, standard deviations and correlations	73
	Table S8	Quantile regression estimates	74
	Table S9	Comparison of baseline usage characteristics between treated and control households	75
	Table S10	ATE specifications, OLS (hourly sampling)	76
	Table S11	ATE estimates at various sampling frequencies, OLS	77
	Table S12	Per capita residential electricity consumption	78
	Table S13	Treatment messages, flash survey	94
	Table S14	Respondents' stated preference for willingness to save energy	94
	Table S15	Ordered logit regressions of respondents' willingness to save energy	100
III.	Table 1	Comparison of participating and non-participating buildings	111
	Table 2	List of balancing characteristics	116
	Table 3	Estimation of propensity score with logit	120
	Table 4	Building energy performance using various matching algorithms	122
	Table 5	Los Angeles Better Buildings Challenge energy savings	122

LIST OF ABBREVIATIONS

ASHRAE	American Society of Heating, Refrigerating, & Air-Conditioning Engineers
CCEP	Center for Corporate Environmental Performance
CCSC	California Center for Sustainable Communities
DOE	Department of Energy
IDRE	Institute for Digital Research and Education
kWh	kilowatt-hour
LABBC	Los Angeles Better Buildings Challenge
LADWP	Los Angeles Department of Water and Power
LEED	Leadership in Energy and Environmental Design
NOAA	National Oceanic and Atmospheric Administration
SABE	Society for the Advancement of Behavioral Economics

PREFACE

The use of field experiments and randomized controlled trials offer rich sources of inquiry to uncover causal mechanisms in the social and behavioral sciences. When these approaches are further integrated with the latest advances in engineering and information technologies, the result is an integrated research agenda that can shape new directions for innovation, science and public policy.

This dissertation employs experimental (and quasi-experimental) techniques to identify the causal effects of interventions aimed at solving social problems for demand side innovation policies in energy and the environment. Mitigating undesirable environmental consequences related to human activities will require two important societal changes: the first is *technological* and the second is *behavioral*. Technological change involves understanding institutions, resource use and the forces that determine the adoption and diffusion of efficient technologies; especially those that can alter the trajectory of resource consumption in major contributing sectors and industries. Behavioral change involves understanding both economic and non-economic drivers for consumers to alter the use of materials and resources. I explore these human dimensions in greater detail in three chapters.

Chapter I begins with the literature review of the dissertation. I examine 37 years of peer-reviewed behavioral field experiments in energy conservation from 1975-2012. This chapter synthesizes the experimental literature with a meta-analysis that combines statistical evidence across multiple studies. It extends previous (primarily qualitative) reviews with meta-regression techniques to quantitatively analyze the effectiveness of information-based behavioral interventions in the residential sector, while controlling for study level factors. The magnitude of the energy savings is discussed by intervention type as is the general finding that less rigorous studies tend to over-report the energy savings in the literature.

In Chapter II, I study incentives for technological and behavioral change, using high-frequency analysis and high-performance computing. I discuss results of a large scale randomized controlled trial

(RCT) with residential consumers in which we provided households with real-time feedback about their home energy consumption down to individual appliances. In this study, we demonstrate that non-monetary information strategies, which communicate the environmental and public health damages of electricity production—such as pounds of air pollutant emissions, childhood asthma and cancer—can result in economically significant changes in consumption with residential consumers.

The human health effects of fine particulate air pollution resulting from industrial activity, which include premature mortality and morbidity, exceed $62 billion in health costs for families worldwide (National Research Council, 2006). The scale and magnitude of the problem greatly affect children's health and welfare through somatic effects, those occurring in persons exposed to air pollution; along with genetic effects, those occurring in susceptible populations, particularly women in late stages of pregnancy, families with young children and urban communities more generally. In a rigorous field study, we demonstrate that tailored information disclosures about the "hidden" social costs of energy use, can incentivize household behavioral changes and even outperform traditional economic incentives for resource conservation in multifamily residential households. The behavioral savings achieved in the randomized trial were of economic significance, estimated rigorously at 8% versus control; and as predicted, were strongest in households with children with energy savings up to 19%. Remarkably, these behavioral savings enabled through recent innovations in metering and information technologies, occurred without changes to existing price structures or economic incentives. Chapter II contributes to behavioral theory on moralized consumer choice and the importance of altruism in consumer decision-making. We show that the adage of 'reducing harm to innocents' is still a powerful motivator in society. We discuss strategies for the design of effective non-price interventions, particularly where price-based policies may not be politically feasible or effective.

While the first two chapters address incentives in the residential sector, the third chapter focuses on incentives in the commercial sector. Commercial buildings readily account for half of all emissions in the United States, so its importance is central in the study of innovation policies for energy efficiency. The chapter takes a challenge-based approach at the firm level, and conducts a program evaluation of a major public-private partnership in the City of Los Angeles unified under the U.S. Department of Energy Better Buildings Challenge. I evaluate the effectiveness of energy efficiency investments and incentives for technology adoption. Using quasi-experimental techniques and matching algorithms, I assess energy savings and various market outcomes to benchmark performance for societal grand challenges. Given the tightening of public finance to bear the costs of rebuilding existing infrastructure, the blending of public information programs with private spending can drive enhanced innovative capacity and increasing cost effectiveness to achieve reductions in energy intensity.

June 2, 2015

Los Angeles, California

ACKNOWLEDGEMENTS

I thank my dissertation committee and UCLA colleagues and researchers at the Center for Corporate Environmental Performance (CCEP) and the California Center for Sustainable Communities (CCSC). Many thanks to colleagues at the Los Angeles Better Buildings Challenge (LABBC) for collaboration on program participation data. Overarching the statistical models in this dissertation, I would also like to acknowledge the Hoffman2 high performance computing facilities managed by the IDRE Research Technology group, which were invaluable throughout this time for statistical computing with large data sets.

The first chapter begins with a literature review of the dissertation. A version of this chapter was published in *Energy Policy* journal in October 2013 under the title: "Information Strategies and Energy Conservation Behavior: A Meta-Analysis of Experimental Studies from 1975-2012" Volume 61, Pages 729-739 [http://dx.doi.org/10.1016/j.enpol.2013.05.109]. This joint study with Magali Delmas uncovers original findings on the effectiveness of behavioral interventions in the peer-reviewed literature. I thank Michael Oppenheim and Rikke Ogawa for research assistance in designing the search strategy for papers in the data set. I also thank Barbara Lawrence for support and guidance in developing the manuscript. I built the data set, models and coauthored all sections. Thanks to Diana Huffaker and National Science Foundation IGERT award #0903720, which supported me during this time.

The second chapter is a version of the paper titled: "Nonprice Incentives and Energy Conservation" coauthored with Magali Delmas and published February 2015 (January 12 online) in the Proceedings of the National Academy of Sciences of the USA (PNAS), Volume 112, Issue No. 6, Pages E510-E515 [www.pnas.org/cgi/doi/10.1073/pnas.1401880112]. This paper was published in the Sustainability Sciences section and was quite unexpectedly featured on the cover of PNAS. We are grateful to the editor, William C. Clark of Harvard University and two anonymous referees, whose comments substantially improved the manuscript. I also thank Thomas Dietz, who authored a commentary on the significance of this work in the same issue of PNAS [www.pnas.org/cgi/doi/10.1073/pnas.1423686112]. None of this would be possible without William J. Kaiser or Victor L. Chen, who developed the technology used in the study. Special thanks to Hilary Godwin, Barbara Lawrence for guiding comments on the earliest versions of this manuscript; and J.R. DeShazo for valuable discussions. I also thank Arthur Winer for useful conversations with me during the design phase. I have benefitted greatly from comments received at research workshops and seminars at Columbia University; UC Berkeley; Carnegie Mellon; Caltech; UC San Diego Rady School of Management; UC Santa Barbara; and the Ivey School of Business. This work was also presented at several conferences including: the Behavior, Energy and Climate Change conference; the Society for the Advancement of Behavioral Economics; the Society for Judgment and Decision Making; the Society for Personality and Social Psychology; and the Academy of Management ONE and TIM doctoral consortia. This research was made possible by grants from the National Science Foundation Awards #0903720 and #SES-125718 and California Air Resources Board Contract #10-332.

The third chapter is work forthcoming on evaluating performance in societal grand challenges. I thank my collaborators David Hodgins and Ben Stapleton of the LABBC; and special thanks to Stephanie Pincetl who provided access to large amounts of utility data, which enabled performance comparisons. This research was supported by a grant from the University of California Center for Energy and Environmental Economics (UCE3) funded by the UC Office of the President. There are many others, who have inspired me over the years and have allowed me to get to this point.

I close by thanking my family and my personal inspiration throughout these years, my fiancée, Γιωτα Στρατου. See you in Folégandros Island.

VITA

OMAR ISAAC ASENSIO

EDUCATION:

Cert. London School of Economics –Peking University (LSE-PKU) 2009

M.S. University of Southern California 2005

B.S. University of Southern California 2002

SELECTED PUBLICATION:

Asensio, O.I. and Delmas, M.A. (2015) "Nonprice Incentives and Energy Conservation" Proceedings of the National Academy of Sciences, 112(6): E510-E515.

HONORS:

SABE Nominee for the Klaus J. Jacobs Research Prize 2015

People's Choice Award, Alliance for Research on Corporate Sustainability 2015

Best Paper Award, Ivey/ARCS Ph.D. Sustainability Academy 2014

Best Reviewer Award, Academy of Management ONE Division 2014

National Science Foundation (NSF-IGERT) Fellowship, UCLA 2010-2013

American Economic Association (NSF-AEA Fellowship) Summer Training Program 2010

PROFESSIONAL EXPERIENCE:

Sr. R&D Engineer (Corporate Strategy and R&D), Avery Dennison Corporation, Pasadena, California and Hazerswoude, Netherlands 2005-2009

Research Chemist, International Solar Electric Technology 2001-2005

I

INFORMATION STRATEGIES AND ENERGY CONSERVATION BEHAVIOR: A META-ANALYSIS OF EXPERIMENTAL STUDIES FROM 1975 TO 2012

This chapter conducts a meta-analytic review of the experimental literature in energy conservation. Over the last 30 years, there has been a growing debate on the effectiveness of information programs used to meet energy conservation targets. Some researchers argue that tailored information programs have a tremendous potential for reducing household electricity use by as much as 20%. However, other researchers have criticized these information programs for their mixed empirical record, often citing heterogeneous effects by consumers, limitations in consumer attention or information processing, or crowding out of incentives, especially when the savings or benefits are small. Our main research question we address in this chapter is: How effective are information strategies in motivating energy conservation and what can we learn from the published experimental evidence about household behavioral changes in response to these programs?

Strategies that provide information about the environmental impact of activities are increasingly seen as effective to encourage conservation behavior. In this chapter, we offer the most comprehensive meta-analysis of information based energy conservation experiments conducted to date. Based on evidence from 156 published field trials and 525,479 study subjects from 1975-2012, we quantify the reduction potential of information-based strategies for energy conservation. On average, individuals in the experiments reduced their electricity consumption by 7.4%. Our results also show that strategies providing individualized audits and consulting are comparatively more effective for conservation

behavior than strategies that provide historical, peer comparison energy feedback. Interestingly, we find that pecuniary feedback and incentives lead to a relative increase in energy usage rather than induce conservation. We also find that the effect varies with the rigor of the study, indicating potential methodological issues in the current literature.

HIGHLIGHTS

- We conduct a meta-analysis of information-based energy conservation experiments.
- We analyze 156 published trials and 524,479 study subjects from 1975 to 2012.
- On average, individuals in the experiments reduced electricity consumption by 7.4%
- Individualized feedback via audits and consulting results in the largest reductions.
- Pecuniary feedback and incentives lead to a relative increase in energy usage

***Keywords*: energy conservation, meta-analysis, feedback**

1.1 Introduction

The environmental impact of everyday activities is often invisible to consumers. Information strategies that aim at correcting this information asymmetry are increasingly common. These include ecolabels (Crespi and Marette, 2005 and Leire and Thidell, 2005), and mandatory and voluntary corporate disclosure (Khanna, 2001 and Delmas et al., 2010). Information strategies are based on the principle that more and better information about the environmental impact of activities will encourage consumers to conserve. While theory suggests that information programs may be effective, the empirical evidence seems to indicate important differences in effectiveness according to type of information provided and the context in which the information is communicated (Delmas and Grant, 2010 and Delmas et al., 2010).

Electricity conservation has been an especially active context for the deployment of information strategies. Energy use accounts for 40% of greenhouse gases across the world and effective conservation programs could contribute to significant environmental improvements. A large number of energy conservation experiments have been conducted using various information strategies to reduce energy use (Abrahamse et al., 2005, Fischer, 2008 and Vining and Ebreo, 2002). These include providing users with savings tips, historical individual usage, real time energy usage, peer usage etc. Yet despite the accumulated experimental evidence, analyses of the effectiveness of such strategies have provided mixed results. Some researchers claim that more information has little or no effect on energy use (Abrahamse et al., 2005), while others estimate that information programs could result in energy use reductions on the order of 22 to 30% over the next 5 to 8 years (Laitner et al., 2009 and Gardner and Stern, 2008). However, these claims are not backed up by rigorous empirical comparison and include very different types of information strategies.

Information strategies are varied. Pricing information has been widely used to induce individuals to save energy (Battalio et al., 1979, Katzev and Johnson, 1984, Nielsen, 1993, Reiss and White, 2008, Sexton et al., 1987 and Slavin et al., 1981). Despite the direct financial benefits of saving energy, research indicates that providing information about the cost of energy use does not necessarily affect energy use behavior among households (Lindén et al., 2006). At the same time, research on the influence of non-price strategies such as peer comparisons (Katzev and Johnson, 1983, McCalley and Midden, 2002, Schultz et al., 2007 and Stern, 1992) has highlighted approaches beyond price information that may drive conservation behavior. At this point, an authoritative comparison of price vs. non-price experiments is lacking. Comparing these different approaches may shed light on the debate of what best motivates energy conservation behavior.

In this paper, we compare the impact of different types of information strategies on energy use to strengthen our understanding of energy conservation information-based strategies. Information strategies include savings tips, energy audits, different forms of energy use feedback, and pecuniary strategies. Experiments generally use one, or at most two or three of these strategies, leaving open the question of how these strategies compare overall. We conduct a meta-analysis of existing field experiments to quantify the effect of information strategies on energy conservation. We focus on experiments trying to lower overall consumption (energy conservation) as opposed to shifting usage in time from periods of high demand to off-peak periods (load shifting). We limit our study to residential settings. We build a dataset of experimental studies within economics, psychology and related fields, incorporating all available evidence. We normalize reported effects to reflect mean changes in energy usage between control and treatment groups. We find a significant overall effect of information strategies on energy savings with a weighted average of 7.4%. Our results also show that strategies providing individualized audits and consulting are comparatively more effective for conservation behavior than strategies that provide historical, peer comparison energy feedback and pecuniary feedback. This indicates that information delivered in person might be more effective than information provided through other media such as mail or e-mail. Interestingly, we find that pecuniary feedback tends to lead to a relative increase in energy usage rather than induce conservation. We also observe that the effect differ across studies depending on the rigor of the methodology used. Indeed the savings are down to 2% for the studies of the highest quality that include a control group as well as weather and demographics controls.

Several authors have provided descriptive reviews of this research area, comparing methods used across studies (Abrahamse et al., 2005), discussing factors influencing residential energy conservation (Burgess and Nye, 2008 and Steg, 2008), classifying studies by theoretical approach (Fischer, 2008 and Vining and Ebreo, 2002), or presenting comparative case studies of residential energy efficiency programs in certain geographic areas (Faruqui et al., 2010 and Mullaly, 1998). While providing interesting insights, these qualitative reviews do not constitute a firm basis for estimating the average treatment effect of behavioral energy conservation programs. Our study is the first to quantify the conservation potential of energy conservation information-based strategies and provides insights into the relative effectiveness of different strategies, which has important policy implications for the future design of such programs.

The paper proceeds as follows. In Section 2, we develop hypotheses on the impact of different types of information strategies. In Section 3, we describe the data-collection and meta-regression methodology. In Section 4 we present the results. In Section 5 we outline steps for advancing methods and theory in this field. A conclusion follows.

1.2 Understanding levers for energy conservation behavior

The failure to engage in energy efficiency can be characterized as a market failure: individuals lack the relevant information or knowledge to engage in energy saving behaviors (DeYoung, 2000, Hungerford and Volk, 1990 and Schultz, 2002) and acquiring such information is costly. Therefore detailed and immediate feedback is a frequently proposed solution to remedy wasteful energy use patterns (Van Houwelingen and Van Raaij, 1989).

We first describe how information about individual energy usage such as historical feedback, and real time feedback, as well as information on saving approaches might facilitate conservation behavior. While these strategies aim at reducing the cost of acquiring information, they do not touch on the potential motivations that might trigger conservation. We then describe the potential effectiveness of information strategies based on social norms and pecuniary incentives.

1.2.1 Energy feedback

Feedback can be described as "the mechanism that directs attention to a specific goal" (McCalley, 2006). The most common form of feedback informs participants about their own energy usage, often drawing comparisons to the past (e.g., Nielsen, 1993 and Winett et al., 1979). Because most individuals have low awareness about their energy usage or its impacts (Attari et al, 2010; Kempton and Montgomery, 1982; Read et al., 1994), periodical energy use reminders, may render energy usage more salient and help trigger conservation activities. In addition, learning about one's own electricity use may increase the sense of relevance of taking action to conserve. If individuals perceive their own impact as negligible, they might not behave in a prosocial manner (Larrick and Soll, 2008). Consequently, making an individual more aware of his or her own energy usage may contribute to conservation. We therefore hypothesize the following:

H1. *Information on past energy use will result in reduced energy use.*

1.2.2 Information on problem solving strategies

Another set of information strategies provide participants with energy savings tips (e.g., Schultz et al., 2007 and Slavin et al., 1981) or conduct home energy audits (e.g., Nielsen, 1993 and Winett et al., 1982). Both of these information strategies involve teaching consumers about new behaviors to lower their energy consumption.

The implicit assumption behind the use of information strategies to reduce energy usage is that these strategies will result in a higher level of knowledge and therefore enable participants to change their behavior (Van Dam et al., 2010 and Ouyang and Hokao, 2009). According to norm activation theory, changes in behavior occur when a person is aware of an issue and thinks he can influence it (Fischer, 2008, Schwartz, 1977 and Vining and Ebreo, 2002). These preconditions to taking action may be enhanced if the person receives additional information on how to perform certain activities and on the outcomes of these activities. With regard to energy conservation behavior, it is conceivable that learning about the impacts of energy usage and receiving conservation tips will lower the barrier to actions. Energy savings tips and audits are likely to contribute to both awareness and perceived behavioral control. Providing such information in an easily accessible manner lowers the cost of information on conservation strategies for the consumer. We therefore formulate the following prediction about the impact of problem solving strategies on energy use:

H2. *Information on conservation strategies will result in reduced energy use.*

Conservation strategies based on energy feedback and information increase individual awareness of the problem and of the possibilities to influence the problem. Once individuals have this information, they will weigh motives versus the cost of actions. The following information strategies frame the message to motivate behavior by focusing on pecuniary incentives or social norms.

1.2.3 Pecuniary strategies

Pecuniary strategies represent another set of strategies commonly used in conservation behavior studies. Lowered energy use results in immediate financial benefits to a household, provided the household pays its own electricity bill. Individuals should be expected to take up energy conservation as long as the benefits of doing so are larger than the costs. Researchers have pointed out the importance of financial incentives and price signals for conserving energy (Hutton and McNeill, 1981).

Many energy conservation experiments inform participants about the financial expenses and/or savings potential associated with their energy usage (e.g., Bittle et al., 1979 and Wilhite and Ling, 1995). Some studies include actual price incentives. These may take the form of rewards or rebate payments (e.g., Slavin et al., 1981), where participants receive a monetary payment for achieving certain energy savings goals. Other studies change the price of electricity (e.g., Sexton et al., 1987), raising for example the price per kWh or introducing rate schedules that change with the time of day or demand levels.

Two recent meta-analysis studies found strong effects of price signals on the timing of electricity consumption (Faruqui and Sergici, 2010 and Newsham and Bowker, 2010), demonstrating that price signals affect behavior. Furthermore, several studies have shown that electricity demand responds to prices, although price-elasticity can be low in the short-term (for an overview see Branch, 1993 and Gillingham et al., 2009).

However, other studies indicate that pecuniary incentives might be counterproductive for energy conservation because they might crowd out more altruistic or prosocial motivations (Bénabou and Tirole, 2005; Bowles, 2008). Furthermore, pecuniary strategies might not be effective if the monetary incentives are negligible. Potential savings from conservation as well as price incentives used in the experiments are often small, in order to bear some relation to the actual price of electricity. For instance, a study by Hayes and Cone (1977) provided a $3 weekly rebate payment for up to a 20% reduction in energy use. In experiments using time of day pricing or critical peak pricing,[1] price differences can be more substantial (e.g., 1:9 ratio used by Aigner and Lillard (1984), as well as Sexton et al. (1987)). The literature is therefore not unanimous about the effectiveness of pecuniary strategies in the current context. Based on the above discussion, we test the following hypothesis:

H3. *Information on monetary savings will result in reduced energy use.*

1.2.4 The power of norms

Comparative feedback provides comparisons to others (e.g., Alcott, 2011, Kantola et al., 1984 and Schultz et al., 2007) and can also be called a motivational strategy, or nudge. Such strategies send non-price signals to participants that activate intrinsic and extrinsic motivation. Besides comparative feedback, motivational strategies also include the use of competitions (e.g., McMakin et al., 2002) and goal-setting (e.g., Katzev and Johnson, 1984) where participants are assigned or select non-binding goals over a defined period of time.

Recognizing the importance of social and psychological aspects, a number of studies on energy use behavior have made use of comparative feedback (Alcott, 2011 and Schultz et al., 2007). These studies illuminate other motivations for changing energy use behavior. In particular, the theory of normative conduct points to the importance of social norms in guiding conservation behavior. Norms influence behavior by giving cues as to what is appropriate and desirable. The effectiveness of social norms in

1 In time of day pricing, prices follow a daily schedule, rising during high demand times. In critical peak pricing, prices are only raised on days with high load forecasts.

bringing about conservation behavior is empirically supported by several studies. For example, Hopper and Nielsen (1991) find that recruiting neighbors to encourage and remind others in their community about recycling significantly increased recycling behavior. In an experiment presenting participants with the choice between a conventional, and a green, but inferior product, participants were more likely to choose the green product if their choices were publicly visible (Griskevicius et al., 2010). Similarly, Nolan et al. (2008) find that comparing individuals to the average energy user was more effective than other strategies at reducing energy usage. Overall, behavioral approaches predict that comparative feedback strategies making use of social norms will be effective in bringing about changes in behavior. We therefore hypothesize the following:

H4. *Information on peer consumption will result in reduced energy use.*

In our hypotheses, we focus on the most common strategies used in energy conservation experiments. In sum, we propose that providing information on past energy use, conservation strategies, financial savings and peer consumption will all contribute to increase energy savings. We now turn to a description of the methods and data collection, before testing our hypotheses and the comparative effectiveness of these strategies.

1.3 Methods

1.3.1 Data collection

We used three complementary search strategies to identify relevant field studies for our analysis. First, we consulted prior narrative review articles in energy conservation (e.g., Abrahamse et al., 2005, Darby, 2006, Fischer, 2008 and Ehrhardt-Martinez et al., 2010). Second, we did a hand search of cited papers in these reviews. Third, we searched the following online databases: (1) EconLit, (2) PsychINFO, (3) Academic Search Complete, (4) Business Source Complete, (5) JSTOR, (6) GreenFILE, (7) Environmental Sciences and Pollution Management, (8) Social Science Research Network (SSRN), (9) GeoRef, (10) Ecology Abstracts, and (11) the NBER database, covering a breadth of disciplines. We compiled a list of keywords using a Boolean search with the following logic: (i) terms relating to energy or electricity[2], e.g., ["energy usage" "energy conservation" "energy demand"], and (ii) terms relating to study type or strategy, e.g., ["behavior*," "feedback," "information," "randomized field trial," "rewards," "incentives," "smart meter," "pricing," "rebates,"] and (iii) terms relating to household or individual level as the unit of analysis, e.g., ["household," "residential," "dormitories," "building," "individual."] This resulted in a list of 6858 scholarly peer-reviewed publications, of which 3,511 were most relevant to our topic. We read all article abstracts and eliminated those not relevant to the topic. We developed a coding protocol and arrived at a short list of 365 articles, of which 59 were experimental studies used in the meta-analysis.

Studies were selected for inclusion on the basis of four criteria. First, as a measure of quality, we focused on peer-reviewed publications as well as the NBER database. Second, we selected only those studies involving behavioral experiments in electricity usage. Gas or water conservation studies, for example, were screened out. Third, only electricity feedback studies at the residential level were selected. Fourth, conditional on the above, we only included studies that reported a quantitative treatment effect, either in percentage relative to a baseline or in kilowatt hours (kWh) per unit time. Experiments focused on the timing of electricity use (e.g. dynamic pricing) were therefore included if they reported conservation effects (changes in kWh), but not if they only referenced load effects (changes in kW).

A number of relevant studies were excluded from the meta-analysis because (1) they did not report quantitative effect sizes relative to baseline levels, or (2) they did not use actual energy readings through individual metering or other verifiable measurement. Electricity use information based strictly on self-reported surveys or questionnaires were excluded from this analysis of experimental studies. Upon completion of the literature screening process, we obtained 59 unique papers, representing 156 field

2 We use the terms "energy" and "electricity" interchangeably.

experiments in 13 countries and 525,479 study subjects and covering the period from 1975 to 2012. Appendix 1 lists all included studies and information coded in the meta-analysis. Appendix 2 contains a complete listing of scholarly journals. All papers were read and coded by two researchers to assure reliable extraction of the effect size and numerical coding of behavioral strategies and of the experimental methods used.

1.3.2 Overview of meta-analysis methodology

Meta-analysis is the art of calibrating and combining statistical evidence from separate studies into a single analysis to provide a quantitative, systematic overview of an empirical effect in the literature. The goal in meta-analysis is to derive a common summary statistic for the effect size of a study and to derive corresponding confidence intervals. Meta-analysis methods have become widely used and cited in the economics and management literature (see for example Stanley and Jarrell (1989) Geyskens et al., 2009). The techniques for analysis generally result in increased statistical power—roughly equal to the sum of individual sample sizes ‐ and can result in improved parameter significance and accuracy relative to primary studies alone (see Bijmolt and Pieters, 2001).

This study uses meta-regression analysis (MRA) to estimate the effects of conservation strategies across many behavioral experiments. This advanced meta-analysis method addresses statistical issues of heterogeneity (Field, 2001, Lipsey and Wilson, 2001 and Nelson and Kennedy, 2009). Heterogeneity in this context occurs when effect sizes in primary studies do not consistently converge to a central population mean, which is certainly the case in energy conservation studies with heterogeneous treatment effects (see Alcott, 2011 and Costa and Kahn, 2010). A key advantage of meta-regression analysis is the ability to model excess heterogeneity in effect size distributions, particularly when combining empirical evidence across groups of studies.

1.3.3 The meta-regression model

For the *jth* study and *L* number of studies included in the analysis, the reported empirical estimates of average treatment effects, b_j are regressed on a vector of study-level characteristics Z_{jk} (typically dummy or indicator variables) as follows:

$$b_j = \beta_j + \sum_{k=1}^{K} \alpha_k \cdot Z_{jk} + e_j \qquad \text{where } j = (1,2,\ldots,L) \tag{1}$$

In equation (1), we adopt standard meta-analytic notation advocated by Stanley and Jarrell (1989). The meta-regression coefficients α_k provide an estimate of the biasing effect of *K* number of moderating variables, for example, the incentive type or duration of the study. Positive values of the meta-regression coefficients imply a positive bias (increased energy use relative to a control group or baseline) and negative values imply a negative bias (decreased energy use relative to a control group or baseline). β_j is the 'true value' of the treatment effect, net of the biasing effect. It is indexed by *j* because we allow for heterogeneous treatment effects by study. The residual errors are captured in e_j. When individual study standard errors are known, we normalize expression (1) by dividing each term by the respective primary study standard errors (S_b) in order to combine unequal variances and mitigate heteroskedasticity; see Stanley and Jarrell (1989) and Roberts and Stanley (2005). In reduced form, we estimate the 'true' empirical effect of moderating variables as follows:

$$t_j = \frac{b_j}{S_{bj}} = \frac{\beta_j}{S_{bj}} + \sum_{k=1}^{K} \alpha_k \frac{Z_{jk}}{S_{bj}} + \frac{u_j}{S_{bj}} \qquad \text{where } j = (1,2,\ldots,L) \tag{2}$$

In the absence of publication bias (i.e. the tendency to favor significant or positive results in published studies), observed effect sizes should vary randomly around the 'true' value and we can empirically estimate meta-regression coefficients for our moderating variables of interest directly from equation (2).

Because most of the standard errors in our data set are missing or not reported in primary studies, we take a commonly used approach in meta-regression analysis, that is, to proxy the effect size variance and hence the primary study standard errors using a monotonic transformation of the primary study sample size (see Nelson and Kennedy, 2009 and Horowitz and McConnell, 2002). We estimate equation (2) by generalized least squares (GLS) and use the square root of the sample sizes as analytical weights. We use a more conservative specification by GLS panel clustered by publication ID (as compared with standard OLS or simple weighted least squares which tend to downward bias the standard errors) to remove heteroskedasticity in the disturbances of the regression model. This offers the advantage of adequately capturing variation in the estimated effect, correlation between effect sizes within the same study and any unobserved component. Our meta-regression model mitigates known heteroskedasticity, provides analytical weights to studies with larger sample sizes, and is less sensitive to estimation bias from small sample studies. In this way, we present robust estimates that allow for multiple effect sizes, model excess heterogeneity, and differences in precision due to sample size.

1.3.4 Measures

1.3.4.1 Dependent variable

Our dependent variable is the reported "effect size" in percentage units. This is a normalized measure across all studies and is defined as the percent change in the treatment group minus the percent change in the control group. Effect sizes can take on both positive and negative values. A negative effect size estimate implies energy savings (conservation) relative to a control group or baseline, whereas a positive effect size estimate implies energy increases relative to a control group or other baseline.

1.3.4.2 Independent variables

We model the effect sizes as a function of study characteristics falling into one of three classes (i) feedback on energy usage feedback and problem solving strategies, (ii) pecuniary strategies, (iii) normative feedback, and (iv) study-level controls, such as weather or demographics. We code these behavioral strategies as dummy variables, taking the value of 1 if the strategy was applied and 0 otherwise.

Energy feedback studies employ *Usage Feedback*: this means participants receive information about their own energy use as a self-comparison to their prior energy use (within subject comparison). We also test whether more specific feedback is helpful, by including a variable called *Real-time*: participants can access energy use information updated at frequencies greater than once per hour. Conservation strategies are measured by two variables, (1) *Energy Saving Tips*: participants receive information on how to save energy (leaflets, alerts or prompts) and (2) *Audits and Consulting*: participants receive in person advice on how to conserve energy or receive visits by technical personnel for home energy audits and consulting.

Pecuniary information strategies include: *Monetary Savings Information*: participants receive information about financial impacts or potential monetary savings from actions to conserve energy. This also includes information about available incentive programs (utility rewards, rebates, tax credits, etc.) but does not involve direct financial transfers; *Monetary Incentives*: participants are involved in direct monetary incentives like rebates, cash rewards and/or tiered pricing or dynamic pricing. Participants can also receive other monetary incentives for conserving energy or achieving certain consumption targets.

Finally, the social norm strategy is presented as the variable *Comparative Feedback*: participants receive information about their own energy use in comparison with others such as their neighbor(s) or community.

Study-level controls include the following variables: *Control Group* indicates whether the study includes a control group as a measure of baseline consumption or treatment counterfactual. When a study does not contain an in-situ or blind control group, the value of this variable is set to zero. *Weather Controls* indicates if the study adjusts for the effects of weather, for example, using heating and cooling degree-days. The lack of weather controls are known to over- or under- estimate the impacts of conservation efforts, depending on the season. *Demographic Controls* specifies if the study adjusts for internal characteristics of the population such as income, education, etc. *Feedback Duration* identifies the time period for the behavioral treatment, measured in months.

1.4 Results

1.4.1 Descriptive Statistics

Table 1 presents the means, standard deviations and percentages of all observations and Table 2 presents the correlations. We see that in general, the effect sizes are not strongly correlated with treatment categories presented in Table 2. This is reasonable and to be expected, given that treatment selection is typically randomized. Among the more significant correlations presented in Table 2, we observe that monetary savings information is strongly correlated with both individual and social comparison feedback strategies. Feedback strategies in conservation studies are often combined with residential billing data, which includes cost savings information as a combined treatment. In Table 4, we quantify the separate effects of these interventions with meta-regression technique.

Our results indicate that quantitative feedback studies in energy conservation date back to at least the mid-1970s (Winett and Nietzel, 1975) with usage feedback representing 75.6% of all experimental observations and 76.9% of the papers. While direct feedback studies are far more common in this literature, the use of comparative feedback in energy conservation dates back to the early 1980s (Midden et al., 1983) but remained largely dormant as a behavioral treatment until rediscovered in the late 2000s, following an influential paper by Schultz et al. (2007)—whose insights from behavioral psychology demonstrated the potential of comparative feedback in residential electricity experiments. Since then, a number of larger studies have recently emerged that use comparative feedback or "social norms" to motivate household conservation (see for example, Alcott, 2011). In Table 1, we see that these comparative feedback studies now represent approximately 1/5 of all entries (23.7% of the observations and 20% of papers). Real-time feedback is still relatively rare and was used in only 22% of studies, for 12.2% of observations. Other incentives tested include strategies such as energy savings tips (72.4% of observations and 63.1% of papers) and audits and consulting (8.3% of observations and 6.2% of papers).

Table 1 Descriptive Statistics

Study Characteristic	Field Observations	Mean	Std. Dev	Min	Max	Percent of Observations	Percent of Papers	Weighted Average treatment effect
Dependent Variable								
Effect Size (Percent)	156	-7.441	10.02	-55.0	18.8	-	-	-7.4%
Independent Variables								
Individual Usage Feedback	156	0.7564	0.43	0	1	75.6%	76.9%	-8.5%
Energy Saving Tips	156	0.7243	0.45	0	1	72.4%	63.1%	-9.6%
Real time Feedback	156	0.1217	0.33	0	1	12.2%	22.0%	-11.0%
Audits and Consulting	156	0.0833	0.28	0	1	8.3%	6.2%	-13.5%
Monetary Savings Info	156	0.3012	0.46	0	1	30.1%	26.2%	-7.7%
Monetary Incentives	156	0.2179	0.22	0	1	21.8%	27.7%	-5.7%
Social Comparisons	156	0.2371	0.42	0	1	23.7%	20.0%	-11.5%
Study Level Controls								
Control Group	156	0.7115	0.45	0	1	71.1%	61.5%	-
Weather	156	0.3141	0.46	0	1	31.4%	24.6%	-
Demographics	156	0.1795	0.38	0	1	17.9%	15.4%	-
Treatment Duration (months)	156	7.6872	12.53	0.3	60	100%	100%	-

Table 2 Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Effect Size (Percent)	1.00											
(2) Energy Saving Tips	0.00	1.00										
(3) Audits and Consulting	-0.10	0.13	1.00									
(4) Monetary Savings Info	0.16*	0.00	-0.05	1.00								
(5) Monetary Incentives	0.12	-0.09	0.01	-0.08	1.00							
(6) Individual Usage Feedback	0.12	-0.08	-0.05	0.34*	0.08	1.00						
(7) Social Comparison Feedback	0.12	0.11	-0.17*	0.46*	-0.26*	0.32*	1.00					
(8) Real Time Feedback	-0.07	-0.38*	-0.11	-0.03	-0.01	0.21*	0.07	1.00				
(9) Control Group	-0.09	-0.20*	-0.12	0.05	-0.32*	-0.26*	0.09	0.02	1.00			
(10) Weather Control	0.12	0.05	-0.15*	0.31*	0.11	0.13	0.3*	-0.13	-0.03	1.00		
(11) Demographic Control	0.22*	0.06	-0.14	0.42*	-0.04	0.23*	0.49*	-0.02	0.15	0.55*	1.00	
(12) Treatment Duration	0.18*	-0.12	0.24*	0.28*	0.29*	0.03	-0.05	0.00	-0.18*	-0.08	0.04	1.00

N = 156 field observations
* p <0.05

Price information and incentives are also directly tested. From Table 1, we see that monetary savings information represent 30.1% of observations and 26.2% of papers; monetary incentives (rebates, cash rewards and tiered pricing) represent 21.8% of observations and 27.7% of papers. These publications generally fall into one of two categories: (1) small-scale behavioral field experiments, typically psychology, building science or engineering, or (2) utility-scale conservation pilot projects, typically economics and related fields. Our sample also distinguishes between monetary incentives and monetary savings information, as we see no significant correlation between these two strategies in Table 2.

1.4.2 Average treatment effect

Across all studies (Table 3), we find the weighted average treatment effect to be −7.44%. A typical behavioral study, on average, will produce more than 7% savings potential, although the range can span significantly from −55% to +18.5%, depending on the study. These numbers are the most comprehensive field experimental figures to date. Interestingly, the average treatment effect differs between more and less methodologically rigorous studies. A savings effect of 1.99% is found for high quality studies that include statistical controls such as weather, demographics, and – most importantly – a control group. In contrast, lower quality studies without such statistical controls find a savings effect of 9.57%. This suggests that savings effects may be overestimated in some of these studies.

We also calculated weighted average treatment effects for each type of treatment (Table 1). On average, field studies using energy audits saw the highest average energy savings, at 13.5%, followed by social comparisons, at 11.5% savings.

Table 3. Summary of Treatment Effects

Description	Field Observations	Mean (%)	St. Dev.	Min	Max
All experimental studies 1975-2012 (unweighted)	156	-7.44	10.0	-55.0	18.5
High quality studies with statistical controls (weather, demographics, and control group)	22	-1.99	1.1	-5.0	5.5
Lower quality studies without statistical controls (weather, demographics or control group)	75	-9.57	12.1	-55.0	8.18

Fig. 1 shows a funnel plot of the primary study treatment effect against a measure of sample size. While a large number of effect size observations are negative, implying energy savings, there is also a considerable number of non-negative experimental effect sizes reported. In Fig. 1, we see a high degree of symmetry (no major truncation about the vertical axis) in our group of studies, which suggests that publication bias is likely not a significant issue in this literature. More generally, we observe strong evidence for heterogeneous responses to behavioral treatments, consistent with prior literature (see Costa and Kahn, 2010 and Freedman, 2006).

In Fig. 2, we plot the reported experimental effect sizes by publication year. Fig. 2 shows little convergence across effect size over the years, suggesting that the field has not converged on optimal strategies.

Figure 1. Funnel plot of conservation effect size vs. sample size

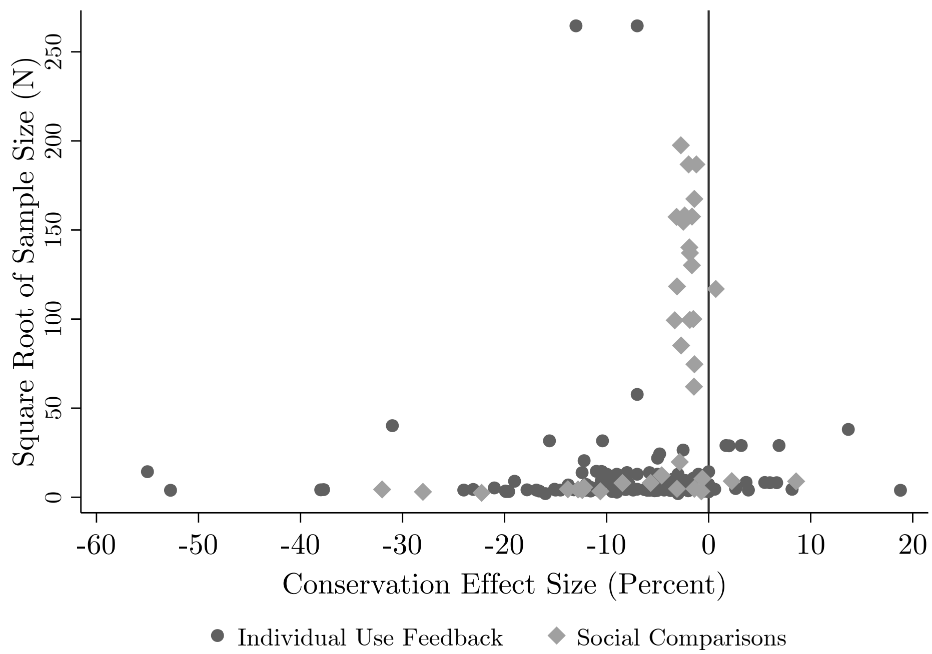
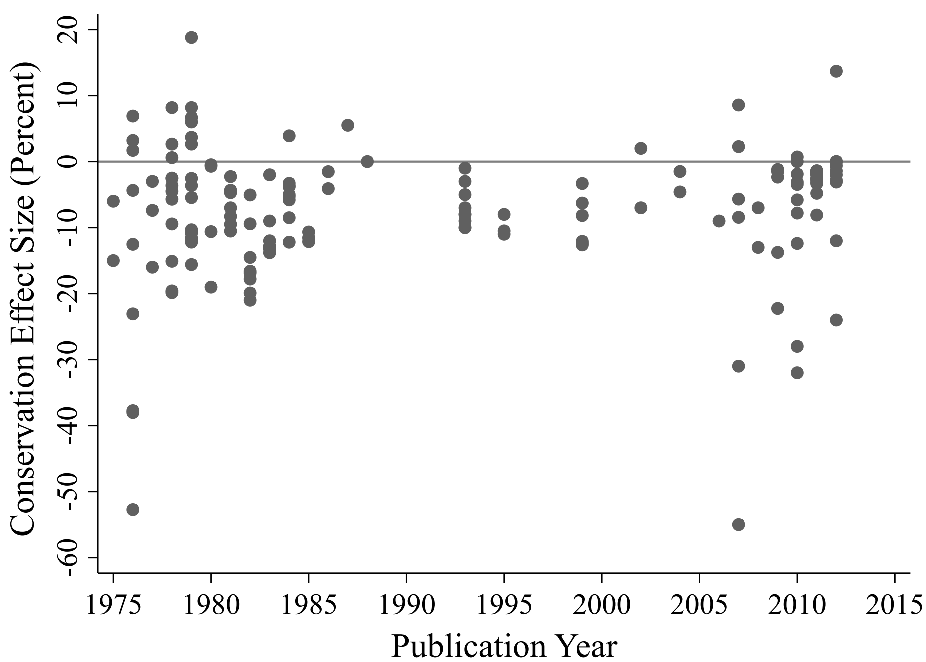


Figure 2. Effect size by publication year



1.4.3 Outcomes of different strategies

Table 4 summarizes the results of the meta-regression model for the different types of behavioral strategies.

Table 4. Meta-regression results

Study Characteristic	(1) Controls only	(2) Individual Feedback	(3) Conservation Strategies	(4) Monetary Information	(5) Comparative Feedback	(6) Full Model
Experimental Treatment						
Energy Use Feedback						
Individual Usage Feedback		1.858**			2.384***	1.346
		(0.796)			(0.787)	(1.000)
Real-time Feedback		-2.849***			-2.197***	-1.175
		(0.555)			(0.743)	(1.205)
Conservation Strategies						
Energy Saving Tips			1.695**			1.547
			(0.781)			(1.109)
Audits and Consulting			-5.124***			-5.678***
			(1.364)			(1.609)
Monetary Information						
Monetary Savings				2.225***		-0.067
				(0.521)		(0.991)
Monetary Incentives				2.189***		2.174**
				(0.435)		(1.028)
Peer Consumption Feedback					0.262	0.617
					(0.940)	(1.374)
Study-Level Controls						
Control Group	-1.950***	-2.028**	-1.369*	-1.340***	-1.227*	-0.034
	(0.735)	(0.827)	(0.792)	(0.460)	(0.735)	(1.144)
Weather Controls	-0.125	-0.756	-1.008	-0.900*	-0.882	-2.025***
	(0.668)	(0.683)	(0.700)	(0.471)	(0.653)	(0.728)
Demographic Controls	7.306***	7.294***	6.766***	6.633***	6.207***	5.833***
	(0.662)	(0.804)	(0.717)	(0.484)	(1.056)	(1.236)
Treatment Duration	0.059**	0.023	0.149***	0.011	0.016	0.120***
	(0.029)	(0.032)	(0.031)	(0.019)	(0.035)	(0.044)
Constant	-8.130***	-8.531***	-9.596***	-9.007***	-9.191***	-10.899***
	(0.718)	(1.142)	(1.115)	(0.188)	(1.024)	(2.094)
Number of Observations	156	156	156	156	156	156
Number of Publications	58	58	58	58	58	58
Wald chi-square	199.6	759.6	177.0	1446	276.5	146.5

Estimation by Generalized Least Squares (GLS) with inverse square root of the sample size as analytical weights.
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In Model 1, we include controls relating to study design. These include Control Group, Weather, Demographics, and Treatment Duration. All of these, except for weather, are significant across specifications. For studies with dedicated control groups, we find a negative bias ranging between 1 and 2% in specifications 1–5. This negative control group bias suggests that higher quality studies with control groups in the study design are more likely to report treatment effects as energy savings. We find that studies without demographic controls over-estimate energy savings between 5.8 and 7.3%. This result is consistent with the view that demographic characteristics are important statistical controls to include with randomly sampled experimental populations. In terms of treatment duration, for each additional month of treatment, there is a small, but significant increase in energy usage in both the simple and full models in Table 4.

This finding indicates that the effect of information programs may be subject to attrition over time, and the dynamic effects of repeated interventions over time merits further investigation. Close to 60% of the field studies in our sample lasted for three months or less, suggesting that studies of longer duration are needed to understand *durability* of treatment effects during experimental periods, and *persistence*, whether information effects disappear over time.

In Model 2, we test the effects of informational feedback: individual feedback about past usage and real-time feedback, controlling for various study level characteristics. Both types of feedback are significant. Interestingly, energy usage increased relative to the control group for studies employing individual feedback strategies. By contrast, conditional on providing information, real-time feedback drives significant energy savings. However, in combination with all other interventions, the effects of feedback are no longer significant in the full specification (Model 6), which combines statistical evidence with other interventions. This is an interesting finding, because it suggests that informational feedback alone (e.g., for example, via smart metering) may be a necessary but not a sufficient condition to produce conservation.

In Model 3, we test information strategies further by examining the effect of 'Energy Saving Tips' a relatively low involvement strategy and 'Audits and Consulting,' a relatively high involvement strategy. Controlling for additional study characteristics, we see that these education strategies are both significant in specification 3, but work in opposing directions. These results demonstrate the powerful role of information in motivating energy conservation and provide insight into whether experiments can stimulate learning effects to encourage conservation. As it turns out, low involvement information-based strategies, i.e. energy saving tips, are not effective at reducing energy use, while high involvement information strategies, i.e. home energy audits and consulting, do support our hypothesis that non-price, information strategies can lead to favorable energy use reductions. While most energy savings tips are provided either in billing or website data, these results suggest that simply providing energy saving tips does not sufficiently motivate subjects to conserve.

For the pecuniary strategies (Model 4), we find 'Monetary Savings Information' or in other words, providing information about potential cost savings, to be significant predictors of energy use behavior, although the effect is opposite to what is predicted by theory. Controlling for major study characteristics, monetary savings information alone did not induce conservation outcomes among study participants but in fact increased usage. However, the significant effect vanishes in the full model (Model 6). Similarly, 'Monetary Incentives' in energy feedback studies, which include rebates, tiered pricing and/or cash rewards is consistently positive and significant in Model 4 and in the full model (Model 6) which includes monetary savings in addition to other experimental treatments. There are several plausible explanations for this empirical finding. One reason for an increase in consumption in response to savings information is that people might simply ignore the potential savings, or that the actual savings might be too small to be meaningful especially given the low price elasticity of electricity use in the short term (Lijesen, 2007 and Reiss and White, 2008) and the small contribution of electricity cost to household expenses.

We also test the effect of comparative or normative feedback strategies in Model 5. Although studies using comparative feedback had the second highest average treatment effect of all strategies (Table 1), the variable is not significant in the full specification. Note, however, that our analysis places analytical weights on experimental studies with larger study sample sizes (weighted by square root of the sample size). As Fig. 1 demonstrates, there is a difference in sampling distribution between studies that use individual feedback and those that use comparative feedback. With very few exceptions, studies using social comparisons have smaller sample sizes. This suggests that further, larger scale studies using comparative feedback are needed to evaluate behavioral effects at scale.

1.5 Discussion

Our study presents the first quantitative comparison of different information strategies used in studies targeted at energy conversation. At most, individual field experiments reported in the literature compare up to three of the six different strategies evaluated in this article. Our meta-analysis allows for a more expansive comparison, because it accounts for differences in strategies across many field experiments. We test some specific predictions about the effectiveness of information with and without financial incentives finding that neither the low-level information strategies (energy saving tips), nor the two feedback strategies (individual usage feedback; comparative feedback) lead to additional energy savings. It is only when information is given in real-time (real time feedback) or includes higher involvement interventions (e.g., home energy audits) that energy conservation is triggered over the span of monitored experimental periods.

In addition, study participants actually increased their energy usage when provided information on monetary savings or monetary incentives (payments or rate changes). One potential explanation for these increases is the so-called "licensing effect" where participants may learn that their expenditures and/or potential savings are small, and they may feel entitled to benefits from energy use because they are paying for it. Overall, the strong focus of current policies on providing additional pricing information is not necessarily warranted based on our study. Rather, it indicates that non-price triggers for behavior change also merit consideration when building future conservation programs.

Although much prior research on energy conservation behavior has focused on pecuniary aspects, one limitation of this approach is that financial benefits from saving energy are often quite small (Wolak, 2011). The average monthly residential electricity bill is $110 (EIA, 2010), so saving 5% energy translates to little more than $5 saved per month. This provides little incentive to conservation behavior given the potential impact on comfort or convenience. Furthermore, a rational actor model of electricity use behaviors, where individuals are utility maximizing and are primarily motivated by self-interest, neglects the pro-social behaviors that people often engage in Penner et al. (2005), Verplanken and Holland (2002). Providing financial incentives may crowd out such prosocial motivation (Bénabou and Tirole, 2005) and this could in fact explain the observed increase energy usage in over thirty years of experimental field studies dating back to the 1970s. Bowles (2008) describes several conditions under which explicit financial incentives may be counterproductive, because self-interest and prosocial motives are not separable, but interact. According to him when incentives were framed as a transaction in terms of a market exchange, they "all but extinguished the subjects' ethical predispositions," without succeeding to "enforce the social optimum." Incentives may also evoke control aversion in individuals, who react exactly opposite to the incentives' intent (see also: reactance theory, Brehm, 1966). Overall, psychological perspectives on incentives predict that financial incentives are effective only under specific circumstances, and sometimes can be counterproductive.

Comparative feedback in the form of norms did not prove to be a significant driver of conservation behavior. As we indicated, one possibility for this finding might be the smaller size of the majority of the studies using this information strategy. Another possibility might be the delivery of the comparative feedback. For example, it is possible that comparative feedback is more effective when delivered in real time but no study in our sample includes real time comparative feedback. Variation in the type of metrics or comparison group might also be important. For example, a recent study found that privately disclosed information about a consumer own (relative) energy use was less effective than when such information was publicly disclosed (Delmas and Lessem, 2012), allowing conservation to act as a signal of "green" virtue.

Finally, the study provides insights into methodological challenges prevalent in this field. Many of the reviewed studies suffer from methodological problems. They involve small samples (e.g., Gronhoj and Thogersen, 2011 and Ueno et al., 2006), short time periods (e.g. Petersen et al. (2007)), and low level of granularity (i.e. providing overall electricity usage without appliance level information, see for example Alcott, 2011, Becker et al., 2010 and Wolak, 2011). A surprisingly large number of studies do not have control groups or do not take baseline measurements prior to reporting changes in consumption. Additionally, many studies also do not account for the impacts of weather characteristics over time or demographics, jeopardizing the reliability of estimates. The estimation methods themselves could also be improved, by adopting more rigorous statistical approaches for time series analysis that can include de-seasonalizing trends in the data or employing difference-in-difference estimation. While we controlled for these methodological factors in the meta-analysis to the best of our ability, future studies in this field should pay careful attention to these aspects to contribute to building a more solid basis of experimental evidence.

Based on our finding, the minimum recommended set of controls for experimental field studies should include:

- Dedicated control group, where subjects are monitored in-situ, but receive no treatment
- Weather controls (i.e. heating and cooling degree days or hours (see Day and Karayiannis, 1998)
- Demographic and household level controls
- Randomization (or pseudo-randomization for opt-in studies)

Randomization appears to be well understood in this literature, so the fourth item is less of an issue in the experimental literature. However, the set of all controls above is non-obvious and forms the basis of methodological issues uncovered in this review, as very few studies incorporate all of the above elements in the experimental research design.

The present study is limited in several regards. As mentioned above, the methodological shortcomings of the included case studies cast some doubt on the reliability of the reported effect sizes. A second limitation of this study is that strategies can differ in additional characteristics that are not tested in this study. For example, conservation strategies may have different levels of intensity. Participating in an energy audit requires greater involvement and time commitment than reading a tip sheet on how to conserve electricity. Studies also differ in the tailoring of the information given. Strategies can be generic (e.g., generalized energy savings tips) or tailored to the participant (e.g., appliance-specific energy-use feedback). For example, audits are custom-tailored to the particular needs of the participant. Finally, the comparison of individual information strategies suffers from confounding effects. There are very few studies that apply only one strategy per experimental group, making it difficult to identify the additional variability explained by a single strategy.

1.6 Conclusion

In this article, we provide a comparison of the quantitative evidence on behavioral strategies targeting energy usage across various literatures in behavioral psychology, economics and related fields. This study represents the most comprehensive review of experimental energy conservation studies to date. We find an overall treatment effect of 7.4% energy conservation across all experimental studies. Based on these results, we conclude that despite heterogeneous treatment effects, non-monetary, information-based strategies can be effective at reducing overall energy usage in controlled experimental studies. This is an important finding, because it suggests that information and education programs targeting conservation through behavioral change should be considered alongside with efforts to reduce energy consumption through technological improvements. As the advent of new technologies such as smart meters reduces the cost of feedback and increases the quality and reliability of information provided, policy makers would be well-served to shift some means to these high-impact, relatively low-cost information programs, which can result in real savings.

While this meta-analysis suggests that information strategies induce energy conservation, it is less clear which strategies work best, in part because many experiments simultaneously use more than one strategy leading to confounding issues and also because of the lack of methodological sophistication of some of the studies. To better identify the winning strategies, additional experiments are needed. Such experiments should learn from the previous literature by following a few guiding principles with regard to the methodological rigor. Sound experiments in energy conservation should use dedicated control groups, take sufficient baseline measurements and control for weather and demographic characteristics. It is also advisable to isolate individual strategies to assess their added value. The field could also benefit from studies of longer duration and larger sample size. With the continuing deployment of smart meters across the world, there are new and exciting opportunities to test information-based strategies for energy conservation. Providing information to encourage energy savings has enormous potential, but it is critical to carry out this research in a more methodologically rigorous manner.

References

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T., 2005. A review of intervention studies aimed at household energy conservation. Journal of Environmental Psychology 25 (3), 273–291.

Aigner, D.J., Lillard, L.L., 1984. Measuring peak load pricing response from experimental data: an exploratory analysis. Journal of Business and Economic Statistics 2 (1), 21–39.

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

Attari, S.Z.; DeKay, M.L.; Davidson, C.I. Bruine de Bruin, W., 2010. Public perceptions of energy consumption and savings. Proceedings of the National Academy of Sciences, 107(37), pp. 16054–16059.

Battalio, R.C., Kagel, J.H., Winkler, R.C., Winett, R., 1979. Residential electricity demand: an experimental study. Review of Economics and Statistics 61 (2), 180–189.

Becker, M.J., Cumming, T.D., Osborne, N.K.P., Bruining, A.M., McClean, J.I., Leland Jr., L.S., 2010. Encouraging electricity savings in a university residential hall through a combination of feedback, visual prompts, and incentives. Journal of Applied Behavior Analysis 43 (2), 327–331.

Bénabou, R. Tirole, J., 2005. Incentives and Prosocial Behavior. NBER Working Paper No 11535, August 2005. National Bureau of Economic Research: Cambridge, MA.

Bijmolt, T.H.A., Pieters, R.G.M., 2001. Meta-analysis in marketing when studies contain multiple measurements. Marketing Letters 12, 157–169.

Bittle, R.G., Valesano, R., Thaler, G., 1979. The effects of daily cost feedback on residential electricity consumption. Behavior Modification 3 (2), 187–202.

Bowles, S., 2008. Policies designed for self-interested citizens may undermine “the moral sentiments”: evidence from economic experiments. Science 320, 1605–1609.

Branch, E.R., 1993. Short run income elasticity of demand for residential electricity using consumer expenditure survey data. The Energy Journal 14, 111–122.

Brehm, J.W., 1966. A Theory of Psychological Reactance. Academic Press, New York.

Burgess, J., Nye, M., 2008. Re-materialising energy use through transparent monitoring systems. Energy Policy 36 (12), 4454–4459.

Costa, D. Kahn, M.E., 2010. Energy Conservation “Nudges” and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment. NBER Working Paper No. 15939, May 2010.

Crespi, J.M., Marette, S., 2005. Eco-labeling economics: is public involvement necessary?. In: Krarup, S., Russell, C.S. (Eds.), Environment, Information and Consumer Behavior. Edward Elgar, Northampton, MA, pp. 93–110.

Darby, S., 2006. The effectiveness of feedback on energy consumption. A Review For DEFRA Of The Literature On Metering, Billing And Direct Displays. Environ- mental Change Institute, University of Oxford, Oxford.

Day, A.R., Karayiannis, T.G., 1998. Degree-days: comparison of calculation methods. Building Services Engineering Research and Technology 19 (7), 7–13.

Delmas, M.A., Montes, M., Shimshack, J., 2010. Mandatory information disclosure policy: evidence from the electric utility industry. Economic Inquiry 48 (2), 483–492.

Delmas, M.A., Grant, L., 2010. Eco-labeling strategies and price-premium: the wine industry puzzle. Business and Society 20 (10), 1–39.

Delmas, M.A., Lessem, N., 2012. Saving Power to Conserve Your Reputation. UC Los Angeles: Institute of the Environment and Sustainability, Retrieved from: http://escholarship.org/uc/item/7d16s49d.

DeYoung, R., 2000. Expanding and evaluating motives for environmentally responsible behavior. Journal of Social Issues 56 (3), 509–526.

Ehrhardt-Martinez, K., Donnelly, K.A., Laitner, J.A., 2010. Advanced metering initiatives and residential feedback programs: a meta review for household electricity-saving opportunities. American Council for an Energy Efficient Economy. ACEEE, Washington, D.C.

EIA, 2010. Residential Average Monthly Bill by Census Division, and State 2010. Energy Information Administration. Available at: http://www.eia.gov/electri city/sales_revenue_price/pdf/table5_a.pdf

Faruqui, A., Sergici, S., 2010. Household response to dynamic pricing of electricity: a survey of 15 experiments. Journal of Regulatory Economics 38, 193–225.

Faruqui, A., Sergici, S., Sharif, A., 2010. The impact of informational feedback on energy consumption—a survey of the experimental evidence. Energy 35 (4), 1598–1608.

Field, A., 2001. Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed- and random-effects methods. Psychological Methods 6 (2), 161–180.

Fischer, C., 2008. Feedback on household electricity consumption: a tool for saving energy? Energy Efficiency 1 (1), 79–104.

Freedman, D.A., 2006. Statistical models for causation: what inferential leverage do they provide? Evaluation Review 30 (6), 691–713.

Gardner, G., Stern, P., 2008. The short list: the most effective actions U.S. households can take to curb climate change. Environment 50 (5), 12–24.

Geyskens, I., Krishnan, R., Steenkamp, J.-B.E.M., Cunha, P.V., 2009. A review and evaluation of meta-analysis practices in management research. Journal of Management 35, 393–419.

Gillingham, K., Newell, R.G., Palmer, K., 2009. Energy efficiency economics and policy. Annual Review of Resource Economics 1, 597–620.

Griskevicius, V., Tybur, J.M., Van den Bergh, B., 2010. Going green to be seen: status, reputation, and conspicuous conservation. Journal of Personality and Social Psychology 98 (3), 392–404.

Gronhoj, A., Thogersen, J., 2011. Feedback on household electricity consumption: learning and social influence processes. International Journal of Consumer Studies 35, 138–145.

Hayes, S.C., Cone, J.D., 1977. Reducing residential electrical energy use: payments, information, and feedback. Journal of Applied Behavior Analysis 10 (3), 425–435.

Hopper, J.R., Nielsen, J.M., 1991. Recycling as altruistic behavior. Normative and behavioral strategies to expand participation in a community recycling program. Environment and Behavior 23 (2), 195–220.

Horowitz, J.K., McConnell, K.E., 2002. A review of WTA/WTP studies. Journal of Environmental Economics and Management 44, 426–447.

Hungerford, H.R., Volk, T.L., 1990. Changing learner behavior through environmental education. Journal of Environmental Education 21 (3), 8–22.

Hutton, R.B., McNeill, D.L., 1981. The value of incentives in stimulating energy conservation. Journal of Consumer Research 8 (3), 291–298.

Kantola, S.J., Syme, G.J., Campbell, N.A., 1984. Cognitive dissonance and energy conservation. Journal of Applied Psychology 69 (3), 416–421.

Katzev, R.D., Johnson, T.R., 1983. A social-psychological analysis of residential electricity consumption: the impact of minimal justification techniques. Journal of Economic Psychology 3 (3-4), 267–284.

Katzev, R.D., Johnson, T.R., 1984. Comparing the effects of monetary incentives and foot-in-the-door strategies in promoting residential electricity conservation. Journal of Applied Social Psychology 14 (1), 12–27.

Kempton, W., Montgomery, L., 1982. Folk quantification of energy. Energy 7 (10), 817–827.

Khanna, M., 2001. Non-mandatory approaches to environmental protection. Journal of Economic Surveys 15 (3), 291–324.

Larrick, R.P., Soll, J.B., 2008. Economics. The MPG illusion. Science 320, 1593–1594.

Laitner J., Ehrhardt-Martinez, K. McKinney, V., 2009. Examining the scale of the energy efficiency continuum. In: Proceedings of 2009 ECEEE Summer Study. La Colle Sur Loup, Cote d'Azur, France.

Leire, C., Thidell, A., 2005. Product-related environmental information to guide consumer purchases—a review and analysis of research on perceptions, understanding and use among Nordic consumers. Journal of Cleaner Production 13, 1061–1070.

Lijesen, M.G., 2007. The real-time price elasticity of electricity. Energy Economics 29, 249–258.

Lindén, A., Carlsson-Kanyama, A., Eriksson, B., 2006. Efficient and inefficient aspects of residential energy behaviour: what are the policy instruments for change? Energy Policy 34 (14), 1918–1927.

Lipsey, M.W., Wilson, D.B., 2001. Practical Meta-Analysis. Sage, Thousand Oaks, CA.

McCalley, L., 2006. From motivation and cognition theories to everyday applications and back again: the case of product-integrated information and feedback. Energy Policy 34, 129–137.

McCalley, L.T., Midden, C.J.H., 2002. Energy conservation through product-integrated feedback: the roles of goal-setting and social orientation. Journal of Economic Psychology 23, 589–603.

McMakin, A.H., Malone, E.L., Lundgren, R.E., 2002. Motivating residents to conserve energy without financial incentives. Environment and Behavior 34 (6), 848–863.

Midden, C.J.H., Meter, J.F., Weenig, M.H., Zieverink, H.J.A., 1983. Using feedback, reinforcement and information to reduce energy consumption in households: a field-experiment. Journal of Economic Psychology 3 (1), 65–86.

Mullaly, C., 1998. Home energy use behaviour: a necessary component of successful local government home energy conservation (LGHEC) programs. Energy Policy 26 (14), 1041–1052.

Nelson, J., Kennedy, P., 2009. The use (and abuse) of meta-analysis in environ- mental and natural resource economics: an assessment. Environmental and Resource Economics 42 (3), 345–377.

Newsham, G.R., Bowker, B.G., 2010. The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: a review. Energy Policy 38 (7), 3289–3296.

Nielsen, L., 1993. How to get the birds in the bush into your hand: results from a Danish research project on electricity savings. Energy Policy 21, 1133–1144.

Nolan, J.M., Schultz, P.W., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2008. Normative social influence is underdetected. Personality and Social Psychology Bulletin 34, 913–923.

Ouyang, J., Hokao, K., 2009. Energy-saving potential by improving occupants' behavior in Urban Residential Sector in Hangzhou City, China. Energy and Buildings 41, 711–720.

Petersen, J.E., Shunturov, V., Janda, K., Platt, G., Weinberger, K., 2007. Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives. International Journal of Sustainability in Higher Education 8 (1), 16–33.

Penner, L.A., Dovidio, J.F., Piliavin, J.A., Schroeder, D.A., 2005. Prosocial behavior: multilevel perspectives. Annual Review of Psychology 56, 365–392.

Read, D., Bostrom, A., Morgan, M.G., Fischhoff, B., Smuts, T., 1994. What do people know about global climate change? Survey studies of educated laypeople. Risk Analysis 14 (6), 971–982.

Reiss, P.C., White, M.W., 2008. What changes energy consumption? Prices and public pressures. Rand Journal of Economics 39 (3), 636–663.

Roberts, C.J., Stanley, T.D., 2005. Meta-Regression Analysis: Issues of Publication Bias in Economics. Blackwell Publishing, Oxford.

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. Psychological Science 18 (5), 429–434.

Schultz, P.W., 2002. Knowledge, information, and household recycling: examining the knowledge-deficit model of behavior change. In: Dietz, T., Stern, P.C. (Eds.), New Tools for Environmental Protection: Education, Information, and Voluntary Measures. National Academy Press, Washington DC, pp. 67–82.

Schwartz, S., 1977. Normative influences on altruism. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, 10. Academic Press, New York, NY, pp. 221–279.

Sexton, R.J., Brown Johnson, N., Konakayama, A., 1987. Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. Journal of Consumer Research 14, 55–62.

Slavin, R.E., Wodarski, J.S., Blackburn, B.L., 1981. A group contingency for electricity conservation in master-metered apartments. Journal of Applied Behavior Analysis 14 (3), 357–363.

Stanley, T.D., Jarrell, S.B., 1989. Meta-regression analysis: a quantitative method of literature surveys. Journal of Economic Surveys. Wiley Blackwell 3 (2), 161–170.

Steg, L., 2008. Promoting household energy conservation. Energy policy 36 (12), 4449–4453.

Stern, P.C., 1992. What psychology knows about energy conservation. American Psychologist 47 (10), 1224.

Ueno, T., Sano, F., Saeki, O., Tsuji, K., 2006. Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. Applied Energy 83 (2), 166–183.

Van Dam, S.S., Bakker, C.A., Van Hal, J.D.M., 2010. Home energy monitors: impact over the medium-term. Building Research and Information, 38; , pp. 458–469.

Van Houwelingen, J.H., Van Raaij, W.F., 1989. The effect of goal-setting and daily electronic feedback on in-home energy use. Journal of Consumer Research 16 (1), 98–105.

Verplanken, B., Holland, R.W., 2002. Motivated decision making: effects of activation and self-centrality of values on choices and behavior. Journal of Personality and Social Psychology 82, 434–447.

Vining, J., Ebreo, A., 2002. In: Bechtel, R., Churchman, A. (Eds.), Emerging Theoretical and Methodological Perspectives on Conservation Behavior. New Handbook of Environmental Psychology, Wiley: New York, NY, pp. 541–558.

Wilhite, H., Ling, R., 1995. Measured energy savings from a more informative energy bill. Energy and Buildings 22 (2), 145–155.

Winett, R.A., Hatcher, J.W., Fort, T.R., Leckliter, I.N., Love, S.Q., Riley, A.W., Fishback, J.F., 1982. The effects of videotape modeling and daily feedback on residential electricity conservation, home temperature and humidity, perceived comfort, and clothing worn: winter and summer. Journal of Applied Behavior Analysis 15 (3), 381–402.

Winett, R.A., Neale, M.S., Grier, H.C., 1979. Effects of self-monitoring and feedback on residential electricity consumption. Journal of Applied Behavior Analysis 12 (2), 173–184.

Winett, R.A., Nietzel, M.T., 1975. Behavioral ecology: contingency management of consumer energy use. American Journal of Community Psychology 3 (2), 123–133.

Wolak, F., 2011. Do residential customers respond to hourly prices? Evidence from a dynamic pricing experiment. American Economic Review 101 (3), 83–87.

Appendix 1.

List of papers included in meta-analysis

Aigner, Lillard	1984	Journal of Business and Economic Statistics, 2(1) 21-39
Alahmad, Wheeler, et.al.	2012	IEEE Transactions on Industrial Electronics, 59(4) 2002-2013
Alcott	2011	Journal of Public Economics, 95, 1082-1095
Alcott	2011	Resource and Energy Economics 33 (1), 820–842
Ayres, Raseman, Shih	2009	NBER Working Paper No. 15386 Sep 2009
Battalio, Kagel, et al.	1979	The Review of Economics and Statistics 61(2), 180-189.
Becker	1978	Journal of Applied Psychology, 63(4), 428–433.
Becker, Cumming, et.al.	2010	Journal of Applied Behavior Analysis, 43 (2), 327-331
Bittle, Valesano, et al.	1979	Behavior Modification, 3(2), 187–202.
Bittle, Valesano, Thaler	1980	Journal of Applied Social Psychology, 10(1), 20-31
Brandon and Lewis	1999	Journal of Environmental Psychology, 19, 75–85
Costa and Kahn	2010	NBER Working Paper No. 15939 May 2010
Craig, McCann	1978	Journal of Consumer Research, 5(2), 82-88
Geelen, Keyson, Boess, Brezet	2012	Journal of Design Research, 10(1/2), 102-120
Gleerup, Larsen et.al.	2010	The Energy Journal 31 (3) 113-132
Gonzales, Aronson, Costanzo	1988	Journal of Applied Social Psychology 18(12) 1049-1066
Gronhoj, Thogersen	2011	International Journal of Consumer Studies 35, 138–145
Gustafsson, Bang	2009	Computers in Entertainment, 7(4), 2009
Hayes and Cone	1981	Journal of Applied Behavior Analysis, 14, 81–88.
Hayes and Cone	1977	Journal of Applied Behavior Analysis, 10(3) 425-435
Hutton, Mauser, et al.	1986	Journal of Consumer Research, 13, 327–336
Kantola, Syme, et al.	1984	Journal of Applied Psychology 69(3), 416–421
Katzev and Johnson	1983	Journal of Economic Psychology, 3, 267–284
Katzev and Johnson	1984	Journal of Applied Social Psychology 14, 12-27
Katzev, Cooper, Fisher	1980	Journal of Environmental Systems, 10(3), 215-227
Kua, Wong	2012	Energy Policy 47, 49-56
Lifson, Miedema	1980	Energy 6, 403-408
Matsukawa	2004	The Energy Journal, 25 (1) 1-17
McClelland and Belsten	1979	Journal of Environmental Systems, 9(1), 29-38
McMakin, Malone, et al.	2002	Environment and Behavior, 34(6), 848–863
Midden, Meter, Weenig, Zieverink	1983	Journal of Economic Psychology, 3(1) 65-86
Nielsen	1993	Energy Policy 21(11), 1133-1144
Ouyang, Hokao	2009	Energy and Buildings, 41(7) 711–720
Pallak and Cummings	1976	Personality and Social Psychology Bulletin, 2(1) 27-30
Palmer, Lloyd, Lloyd	1977	Journal of Applied Behavior Analysis, 10(4) 665-671
Peschiera, Taylor	2012	Energy and Buildings 49, 584-590
Peschiera, Taylor, Siegel	2010	Energy and Buildings, 42(8) 1329–1336
Petersen, Shunturov, et.al.	2007	Intl. Journal of Sustainability in Higher Education 8(1) 16-33
Reiss, White	2008	RAND Journal of Economics, 39(3), 636–663
Schultz, Nolan, et al.	2007	Psychological Science 18(5), 429
Seligman and Darley	1977	Journal of Applied Psychology, 62(4), 363–368
Seligman, Darley, Becker	1978	Energy and Buildings 1(3), 325-337
Sexton, Brown Johnson et al.	1987	Journal of Consumer Research, 14, 55–62
Slavin, Wodarski, Blackburn	1981	Journal of Applied Behavior Analysis, 14(3), 357–363
Staats, Harland, et al.	2004	Environment and Behavior, 36(3), 341–367
Torriti	2012	Energy 44, 576-583
Ueno, Sano et al.	2006	Applied Energy 83(2), 166-183
van Dam, Bakker, van Hal	2010	Building Research and Information, 38(5) 458-469
Vassileva, Odlare, Wallin, Dahlquist	2012	Applied Energy 93, 572-582
Wenders, Taylor	1976	Bell Journal of Economics 7(2) 531-552

Wilhite and Ling	1995	Energy and Buildings 22(2), 145-155
Winett, Hatcher, et al.	1982	Journal of Applied Behavior Analysis 15(3), 381
Winett, Kagel, et al.	1978	Journal of Applied Psychology, 63(1), 73–80
Winett, Leckliter, et al.	1985	Journal of Applied Behavior Analysis, 18, 33–44
Winett, Love, Kidd	1982	Journal of Environmental Systems, 12(1), 61-70
Winett, Neale, et al.	1979	Journal of Applied Behavior Analysis, 12(2), 173–184
Winett, Nietzel	1975	American Journal of Community Psychology, (3)2 123-133

Appendix 2

List of journals included in meta-analysis

American Journal of Community Psychology
Applied Energy
Behavior Modification
Bell Journal of Economics
Building Research and Information
Computers in Entertainment
Energy
Energy and Buildings
Energy Policy
Environment and Behavior
IEEE Transactions on Industrial Electronics
International Journal of Consumer Studies
International Journal of Sustainability in Higher Education
Journal of Applied Behavior Analysis
Journal of Applied Psychology
Journal of Applied Social Psychology
Journal of Business and Economic Statistics
Journal of Consumer Research
Journal of Design Research
Journal of Economic Psychology
Journal of Environmental Psychology
Journal of Environmental Systems
Journal of Public Economics
NBER Working Paper
Personality and Social Psychology Bulletin
Psychological Science
RAND Journal of Economics
Resource and Energy Economics
The Energy Journal
The Review of Economics and Statistics

II

NONPRICE INCENTIVES AND ENERGY CONSERVATION

In this chapter, we investigate the effectiveness of nonprice information strategies to motivate conservation behavior. We introduce environment and health-based messaging as a behavioral strategy to reduce energy use in the home and promote energy conservation. In a randomized controlled trial with real-time appliance-level energy metering, we find that environment and health-based information strategies, which communicate the environmental and public health externalities of electricity production, such as pounds of pollutants, childhood asthma, and cancer, outperform monetary savings information to drive behavioral change in the home. Environment and health-based information treatments motivated 8% energy savings versus control and were particularly effective on families with children, who achieved up to 19% energy savings. Our results are based on a panel of 3.4 million hourly appliance-level kilowatt–hour observations for 118 residences over 8 months. We discuss the relative impacts of both cost-savings information and environmental health messaging strategies with residential consumers.

Keywords: Energy conservation, decision-making, health information disclosure, environmental behavior, randomized controlled trials

2.1 Introduction

In the electricity sector, energy conservation through technological and behavioral change is estimated to have a savings potential of 123 million metric tons of carbon per year, which represents 20% of US household direct emissions (1). Although some scholars contend that improvements in energy generation technologies offer the greatest potential for carbon emission reductions (2), others argue that household-level behavioral changes can also produce significant and immediate emission reductions (1). In residential electricity markets, however, promoting conservation through behavior change is particularly challenging. Traditional economic incentives for household energy conservation are typically small and subject to problems of inattention or imperfect information, which economists often classify as information or market failures (3–7). Tailored information strategies could solve problems of imperfect information in markets—by disclosing the unobserved costs of individual consumption decisions to consumers (8). However, because electricity demand is relatively price inelastic (9), nonprice information strategies using normative, intrinsic, or social motivations might prove effective alternatives (10, 11). In this article, we compare the effectiveness of environmental and health information disclosures on residential energy consumption to more traditional cost-based information strategies.

Public environmental and health damages from energy generation, which include premature mortality and morbidity (such as cancer, chronic bronchitis, asthma, and other respiratory diseases), have not traditionally been the focus of energy conservation policies. However, decades of research on environment and health effects of air pollution have shown electricity generation to be one of the most important sources of pollution and with recognized impacts on global health such as childhood asthma and cancer. Since the 1990s, prospective cohort studies, time-series studies, and rigorous epidemiological data have provided strong causal evidence of the associated health effects of ambient air pollution (12). These include both "somatic effects"—for example, those occurring in the persons exposed—along with "genetic effects"—those occurring in at-risk populations (12). Global health damages are by far the most prominent externalities, primarily due to air pollution from coal and natural gas, which constitute a majority of the current energy system. Health damage estimates already exceed $120 billion in 2005 US dollars (13), with electricity price structures that do not necessarily reflect these costs.

2.2 Health externalities: a missing link in consumer choice

The link between individual electricity use and the resulting impacts on human health (via energy-related industrial emissions) remains elusive for most consumers. Household electricity use is typically "invisible," meaning consumers have limited information about the external effects of their individual electricity consumption. In this article, we investigate whether information about the environmental health effects of energy consumption could impact conservation behavior.

Behavioral theory suggests that disclosing environment and health-based externalities to consumers can be effective at shifting conservation preferences and reducing the perceived costs and/or moral benefits of individual consumption (14). Prior literature also points to important differences in the effectiveness of environmental cues, according to the type of information provided and the context in which the information is communicated (15–17). In the context of energy consumption, we argue that policies that correct information asymmetries between individual consumption and pollution externalities can encourage conservation by reframing and creating new mental accounts on the perceived costs and benefits of household actions to conserve energy. In pursuing tailored information disclosures related to environment and health externalities, we examine whether moral norms and moral choice can affect how individual consumption decisions are made and subsequently evaluated by consumers.

There is a rich literature on the importance of moral payoffs and moral norms on household consumption decisions. Research in psychology (18–23), economics (24–27), marketing (28–30), sociology (31–34), philosophy (35, 36), and neuroscience (37, 38) has shown that normative strategies can motivate human behavior in the interests of the long-term benefits of the social group rather than the short-term, self-interested behavior of one person. Learning that one's marginal consumption imposes social costs on others can lead to different moral sensitivities to external health damages. However, moral sensitivity to reducing harm in others is to be distinguished from purely altruistic motivations such as in philanthropy or charitable giving, as the benefits of individual conservation actions bestow not only social benefits onto others but also private benefits on the individual (i.e., lower costs, reduced pollution, cleaner air, etc.).

We consider two psychology-based mechanisms: The first is amplification of prosocial conservation preferences that is motivated by a need to reduce harm on others (or activate behavior that aids others); the second is amplification of private benefits from reduced marginal consumption, which also provide private benefits to the individual (e.g., fewer emissions leading to known health damages). This amplification strategy serves dual purposes and could apply equally to populations with greater sensitivities to the greater good and to those households who also stand to gain from cleaner air and the reduction of health externalities, which could represent a broad segment of the population. Particular examples of such study subjects could be urban communities and, in particular, affected populations such as the elderly or families with children. Targeting urban communities and families with children, we test the effectiveness of environment/health-related social messaging on household energy conservation in a real market setting.

2.3 Experimental evidence

A large number of energy conservation studies have been conducted using various information strategies to reduce energy use (10, 39–45). These studies provide users with energy-saving tips, historical individual use, real-time energy use, and peer use, including social comparisons. Despite a growing body of literature on nonprice strategies with tailored information campaigns, researchers have not yet tested the effectiveness of consumer information disclosures based on environment and health externalities (45). Therefore, the empirical evidence of moralized consumer choice using environmental health cues remains as yet largely undetermined. Expanding the ensemble of large-scale behavioral strategies, we present experimental field evidence with residential electricity customers in a major US city. We demonstrate that nonprice-based environment and health messaging can have substantial and economically meaningful reductions in demand at the household level. Our central contribution is to test the role of information disclosure about environment and health damages as a new class of nonprice strategies for household energy conservation.

2.4 Measuring conservation behavior

In the energy conservation context, prior field studies have been limited in their ability to measure high frequency behavior and to provide residents with timely feedback about their electricity use. Prior studies often use data obtained from long or infrequent residential billing cycles, indirectly using energy modeling techniques or self-reported surveys about intentions to conserve. More generally, the lack of appliance-level energy metering data in US households and businesses has been a long-standing problem for modeling and understanding consumer behavior in residential and commercial buildings (46). In the current study, new technology developments allow us to observe kilowatt–hour (kWh) electricity behavior in real time, at the appliance level (47). A kWh is the most common unit of electricity used by electric utilities in residential and commercial billing.

Behavioral experiments in energy research are now transitioning from small-scale laboratory experiments to large-scale field studies (48–50), with randomized controlled trials (RCTs) emerging as a powerful approach for policy evaluation of information treatments. RCTs enhance the credibility of findings by modeling actual consumer behavior at scale and, under realistic settings, often in contrast to controlled laboratory studies. However, RCTs are usually more costly to conduct versus non-experimental observational studies. This is because archival data are often cheaper per unit of observation, so it is possible to have more observations for the same unit cost over a broader setting or population than might be available in a RCT, particularly in cases when there are limits to sampling, measurement error, or treatment imbalance. For a discussion of strengths and limitations of RCT, see refs. 51 and 52. Sound inference comes from triangulating multiple sources of evidence. This is why we combine RCTs with survey data, not only to provide richer evidence of the effects of a treatment before and after an intervention but also as a way to optimize the treatment itself. In the current study, we conduct a high-frequency, high time-resolution RCT study at a multiple-building, family apartment residential field site. We observe consumer behavioral responses to information treatments in real time with appliance-level metering capabilities not previously available. We integrate a behavioral science-based consumer messaging strategy, which connects the causal chain between energy use and associated environment and health consequences at the individual household level.

Our sample consists of Los Angeles Department of Water and Power (LADWP) customers who pay their electricity bills, and our experimental results represent outcomes of real-life consumption decisions in their natural settings. Our field experimental site, University Village, is a large family housing community in Los Angeles with 1,102 units. On a per capita electricity basis, University Village residents are typical of California multifamily renter populations (SI Appendix, Table S12) and are only slightly below the national average (due to the milder climate in the State of California). (For more information on the characteristics of our sample, please see SI Appendix.) Our 118 participating households consist of single, married, and domestically partnered graduate college students with and without children in the home. Residents are younger and more educated than the US population but are typical of users of information devices. Our target population represents the next generation of homeowners who are used to working with mobile electronic devices and increasingly rely on electronic communications in their consumption habits. Thus, our experimental results are indicative of how future residential electricity consumers can respond to high-frequency information, especially as electric utilities begin using smart metering data with information and communication technologies.

Building an intelligent, wireless sensor network, we gave consumers real-time access to detailed, appliance-level information about their home electricity consumption. Our results are based on a panel of 440,059 hourly kWh observations (or 3.43 million underlying appliance-level kWh observations) for 118 residences over a time span of 8 months. We also conducted the analysis at higher frequency toward the limit of the technology (metering and data processing) at 1/30 Hz—for example, one reading every 30 s—to evaluate the optimal span of inference. Our optimal unit of observation in this study is hourly, which balances several competing requirements and considerations, not the least of which are the span of decision making for conservation behavior, the technical capabilities of the metering equipment, the precision of the estimates, computational burdens, and other practical considerations. We provided treated households with high-resolution information about costs (weekly cost estimates as opposed to monthly billing) or environmental and health impacts (weekly emissions and listing of particular health consequences; e.g., childhood asthma and cancer).

Informational messages were delivered via a specialized, consumer-friendly website with monitored page views and analytics and weekly accessible emails by personal computer and portable electronic devices (SI Appendix, Fig. S1). Information feedback was specific to each consumer. Once randomly assigned to receive either cost savings or environment- and health-related information, households could not cross over between treatments. Building on previous literature and to provide all treated households with a reference point for their consumption, we compared our participants to the top 10% most energy efficient-similar neighbors in the complex. (Households were provided with factual evidence-based numbers that depended on their weekly kWh electricity consumption. Equivalent cost savings were calculated using household consumption data and the published LADWP electric rate schedules for residential customers. LADWP is the nation's largest public utility. Equivalent non–base-load emissions were calculated using emission factors from the Emissions & Generation Resource Integrated Database maintained by the US Environmental Protection Agency.) After a 6-month baseline monitoring period, the treatment period was ∼100 d, which is the typical duration of an information campaign during peak summer or winter months. Our treatment period is also greater than 60% of comparable studies from 1975 to 2012 (45).

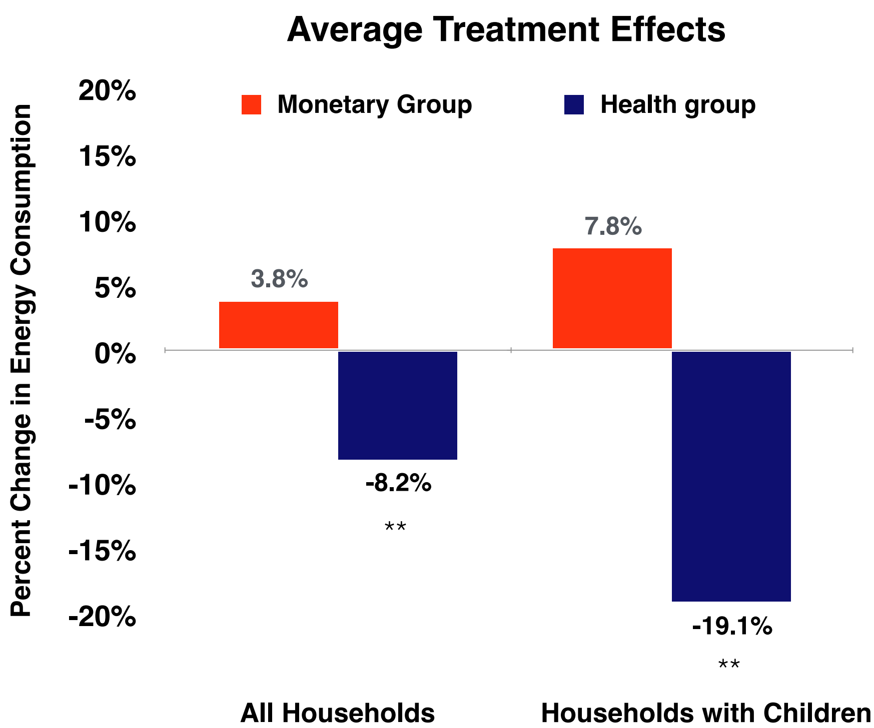


Fig. 1. *Effects of informational messages on study households (N = 490,994 hourly kWh observations, 118 apartments by random assignment into treatment and control groups). Mean treatment effects are reported versus control households before and after treatment following a 6-mo baseline monitoring period. The cost savings information group shows no significant conservation behavior after the 100-d treatment period. The health group shows significant conservation behavior of 8.2% energy savings (significant at **P < 0.05) after the 100-d experimental period. Health-related information treatments are particularly effective on families with children, achieving 19% energy savings relative to control (significant at **P < 0.05). All panel regression estimates include statistical controls for household characteristics (apartment size, apartment layout, and building floor), occupancy (number of persons living in the household), hourly weather controls (e.g., heating and cooling degree hours), time fixed effects, and environmentalist ideology (head of household reports being an active member of an environmental organization). Materials and methods are available in SI Appendix.*

2.5 Results and discussion

We find that health and environment messages, which communicate the public health externalities of electricity production such as childhood asthma and cancer, outperform monetary savings information as a driver of behavioral change in the home. Participants who received messages emphasizing air pollution and health impacts associated with energy use reduced their consumption by 8.2% over the 100-d experimental monitoring period versus control (Fig. 1 and SI Appendix, Table S4, column 1). These net energy savings, which invoke considerations of health damages as a psychological mechanism, are at the high end of prior nonprice strategies using social comparisons (39, 40). To give a practical sense for what these savings mean for a typical two-bedroom family apartment, an 8% conservation effect would be equivalent to plugging out a laptop computer for an additional 87 h/wk, plugging out a flat-screen TV for an additional 36 h/wk, or turning off one standard 60-W light bulb for an additional 72 h/wk. [For these equivalencies, we used nameplate wattages for typical household consumer appliances compiled by the US Department of Energy (available at http:// energy.gov/energysaver/articles/estimating-appliance-and-home- electronic-energy-use).] Using published price elasticities for California (53, 54), this conservation effect on the treated is equivalent to a long-run electricity price increase of 20.5% or a 60-d short-run price increase between 30% and 60%. Consistent with our predictions, health and environment messaging was particularly effective on families with children, who collectively achieved up to 19% energy savings (Fig. 1) in our target population. Our results are robust to various estimation procedures and specifications. [We estimate treatment effects by difference- in-differences panel regression. The full set of statistical controls for observable characteristics include hourly weather controls (e.g. heating and cooling degree hours), time fixed effects, apartment size, and occupancy characteristics, including a proxy for household environmental leaning. Any unobserved characteristics common to the community are captured in the control group monitoring. Supporting materials and methods and further robustness checks are available in SI Appendix.] In particular, our results are robust to sampling frequency, and we do not rely on our panel's high time dimension to achieve statistical significance (SI Appendix, Table S11). Although we expect some attenuation of these effects across larger study populations, we demonstrate the behavioral principle of using health damages and moralized consumer choice as a promising behavioral strategy for residential energy consumption. By contrast, participants who received messages informing them about monetary savings did not produce significant conservation by the end of the experimental period, net of all statistical controls (materials and methods are available in SI Appendix). This result of conservation in one group and no net conservation in another leads us to seek a deeper understanding of the underlying heterogeneity and individual behaviors driving household actions.

The lack of a significant conservation effect with cost savings information, which might initially be a surprising result, is consistent with over 35 y of experimental evidence in the behavioral literature in energy conservation (45). Although cost savings has historically been an important economic incentive for household energy conservation, in practice the actual realizable dollar savings for most US households, compared with the top 10% most energy efficient-similar neighbors, is typically small. In the current experiment, for example, household cost savings potential for a two-bedroom family apartment with an average consumption was US$5.40 to US$6.60/mo in direct kWh charges, which is roughly equivalent to a fast food combo meal or two gallons of fortified whole milk, based on the consumer price index average price data. [The consumer price index average price data, published by the Bureau of Labor Statistics, provides monthly data on prices paid by urban consumers for a representative basket of goods and services (available at www.bls.gov/ cpi/).]

On an annual basis, the savings estimate for the current multifamily residential housing complex, which is at the mid-range of national per capita electricity consumption (55), is a modest $65 to $80/y. These energy savings in dollar terms, although small relative to the US household budget, are realistic for most US households, suggesting that information about small monetary savings, especially over longer time horizons (weeks to months), may not sufficiently motivate household behavioral change and may be heavily discounted by consumers or subject to energy rebounds. Gneezy et al. (56) provide other examples on when and why monetary incentives do not work to modify behavior. Further work is needed to understand the thresholds that prompt informed consumers to change behavior, to disentangle the level of the incentive from incentive type.

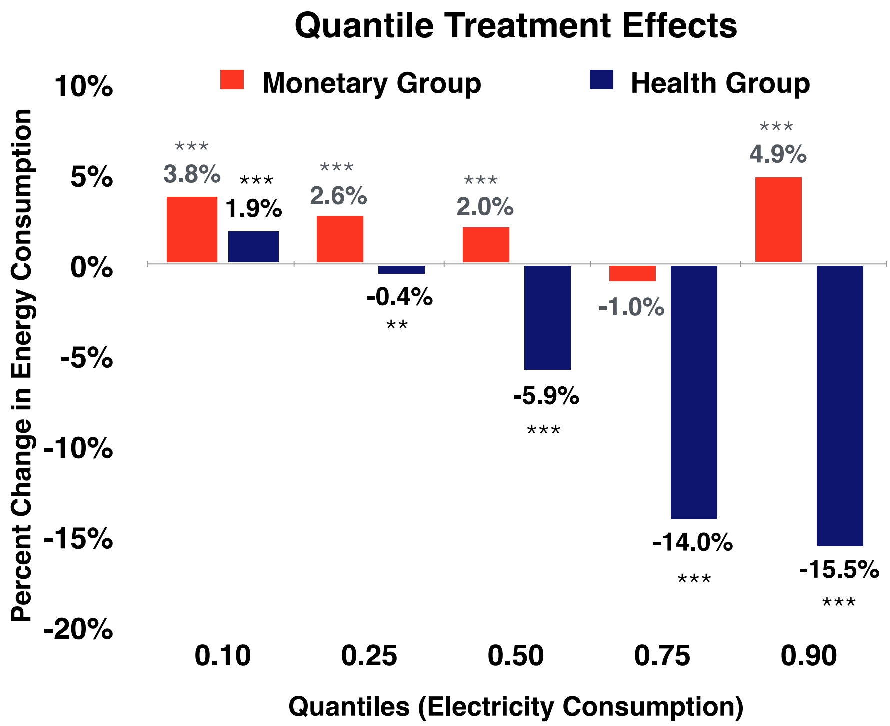


Fig. 2. *Quantile treatment effects on the treated (N = 490,994 hourly kWh observations, 118 apartments). We observe significant conservation effects in the health treatment group across all quantiles of electricity use, except for the lowest decile (most energy efficient observations). By contrast, by the end of the experiment, we observe no significant conservation effect with the monetary savings group and observe splurging behavior, particularly among the highest use quantiles. Significance levels are as follows: ***P < 0.01, **P < 0.05, *P < 0.1.*

2.5.1 Heterogeneous effects on households

Although average treatment effects vary for households with and without children (Fig. 1), we also investigated whether heterogeneous effects could be uncovered for different household use patterns. Heterogeneous responses to information treatments are well known in the behavioral literature on energy conservation. Using cross-sectional quantile regression, we evaluated the distributional impact of informational messages on treated households (Fig. 2 and SI Appendix, Table S8). We find that health and environment messaging produced statistically significant conservation effects in all but the lowest decile of household electricity use (e.g., households who are already the most energy efficient). Weekly cost savings messages, on the other hand, led to increased electricity use relative to control (Fig. 2). These deviations from mean treatment effects and positive splurging behaviors were particularly striking among families with children (Fig. 1) and the highest deciles of household electricity use (Fig. 2), whereas in contrast to health-based messages, monetary savings information was ineffective for the most energy-intensive households. To further understand what changes in behavior may be driving these results, we evaluated the experimental treatment effects by appliance and by time of day.

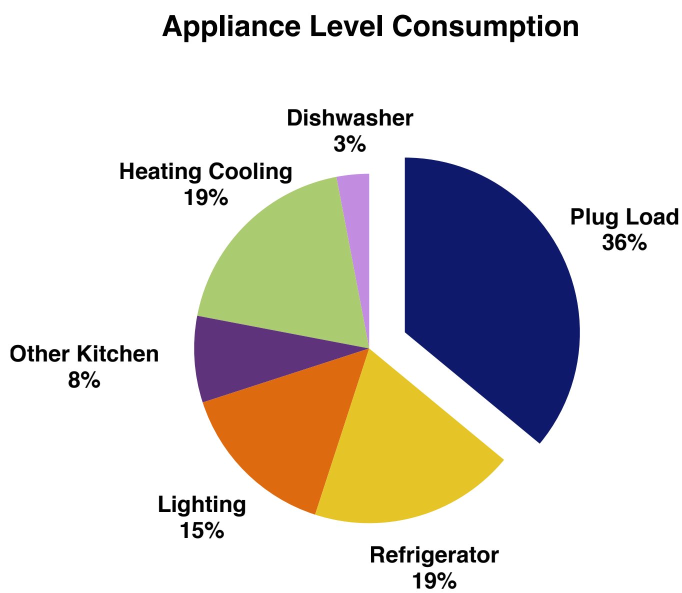


Fig. 3. *Appliance-level electricity measurements (N = 490,994 hourly kWh observations, 118 apartments). Plug load is the largest share of household electricity use. The average kWh consumption is 230.4 kWh/mo across one-, two-, and three-bedroom units ranging from 595 to 1,035 square feet. Appliance-level data for multifamily residences in this study are among the first field demonstrations of comprehensive appliance-level metering capabilities not previously available. Results above represent a weighted average of all household electricity uses obtained by direct measurement and are not based on engineering estimates by modeling.*

2.5.2 Appliance-level behavior

The average electricity consumption across all households is 0.3157 kWh/h or ~230.4 kWh/mo across one-, two-, and three-bedroom units ranging from 595 to 1,035 square feet. Because we have separately metered appliances, we can further decompose the appliance-level consumption. In Fig. 3, we provide the breakdown of the appliance-level readings for all apartments in the study. Major appliances (e.g., refrigerator, dishwasher), the plug load (e.g., charging devices, consumer electronics, etc.), and lighting make up a significant share of household direct energy use (73%). The results shown in Fig. 3 represent experimentally observed appliance-level electricity readings and are not the result of survey estimates or modeling as in traditional approaches to obtain such data. By the current state of technology, there is no centralized appliance-level metering capability in US homes or residential electricity markets (46). This study is one of the first, to our knowledge, to have experimentally measured appliance-level data in a large energy study.

For decades, heating and cooling (e.g., space conditioning) was considered to be the major source of household electricity use, based on national data from the Residential Energy Consumption Survey. Estimates from the most recent Residential Energy Consumption Survey suggest that the share of residential electricity use for heating and cooling is declining nationally in the United States, down to 48% in 2009 from 58% in 1993 (55). In California, due to the milder climate, the share of heating and cooling makes up a smaller fraction of energy use (31%), across all single and multifamily households, and only 19% in our multifamily residential field site (Fig. 3). Although space heating and cooling is declining nationally, the share of energy use for appliances and electronics continues to rise. Consistent with these estimates, by direct measurement, we show that plug load is already the largest share (36%) of appliance-level electricity consumption for residential apartments at our field site (Fig. 3).

For households randomly assigned to receive health messages, energy conservation occurs primarily through plug load and lighting behavioral changes (SI Appendix, Table S5). Whereas our environment and health strategy was most effective in reducing plug load, we observe markedly different appliance behavior with the monetary savings strategy. For households randomly assigned to receive cost savings information, we identify conservation effects at the appliance level only in lighting (SI Appendix, Table S5). However, as lighting is only a minor share of total household energy consumption (15%), any observed behavioral changes in lighting conservation are not enough to overcome observed splurging behavior in other consumption categories such as heating and cooling, resulting in no net conservation with monetary savings information by the end of the experiment, and in some cases increasing electricity use relative to control. This empirical result of conservation in one or more appliances (e.g., lighting) but no net conservation in the household aggregate energy use motivates further research into dynamic responses to information treatments and habit formation. Results from our focus group indicated that people were unclear on how to operate the refrigerator controls, for example, and we observed an 8% increase in refrigerator use (SI Appendix, Table S5), which could be an opportunity for manufacturers to improve designs. The recent work of Attari et al. highlights the importance of consumer perception and cognitive ability on the effectiveness of environmental cues (17, 57). One could ask the obvious question: Why should health-based information lead to different observed appliance-level behaviors? One explanation for this empirical result is that health-based strategies lead morally sensitized consumers to be more cognizant of household energy uses that might be perceived as "wasteful" sources of electricity—for instance, unused lights, phantom loads, or standby power sources. Consistent with this hypothesis, in post-study participant interviews, the most commonly reported behavioral changes in the health information group were turning off unused lights, unplugging electronics, and charging devices when not in use. Our metering technology has opened the possibility to study behavioral phenomena at very high resolution.

2.5.3 Implications for load shifting

We also decompose the appliance-level treatment effects by time of day to evaluate implications of our information treatments on possible load-shifting behavior. Load shifting of household electricity use from peak hours to off-peak hours is desirable for electric utilities to manage system power loads and reduce the risk of blackouts, brownouts, or overvoltages on the grid. For households randomly assigned to environment and health messages, we observe daily conservation effects, versus control households, beginning from about 12:00 AM (midnight) through 12:00 PM (noon). In-treatment energy savings persist overnight and during peak morning demand hours (SI Appendix, Table S6), where a local peak load period occurs for the community at ~9:00 AM (SI Appendix, Fig. S3). These changes in electric consumption patterns via appliance-level reductions in plug load and lighting behavior, particularly during morning peak hours, offers some evidence for habituation within treatment. Conservation treatment effects for our environment and heath group are also maintained overnight, consistent with our evidence of plug load conservation, suggesting both load-shifting behavior and conservation. By contrast, we find limited evidence of any load-shifting behavior with cost savings information treatments by the end of the experiment.

2.5.4 Attitude-behavior gap

In the conservation literature, there is often a dichotomy between what people say they do and what they actually do (58). This so-called attitude–behavior gap is uniquely revealed in this field setting. Before the study, we conducted a stated preference survey asking independent, random samples of participants to choose messages that would be most likely to change their behavior and motivate conservation in the home. When pushed to state their energy preferences, we find that consumers do state a willingness to change behavior and that financial savings are at the top of their concerns. However, when faced with decision making in an actual market setting, only our nonmonetary, environment, and health strategy produced a lasting conservation effect. This distance between what people say they would do and what they actually do is referred to as hypothetical bias. As long argued by psychologists and behavioral economists, monetary savings, which by standard accounts should motivate rational decision making in the home, can often fail with ordinary consumers (11, 14, 56). The idea that a nonmonetary, information strategy centered on environment and health could produce energy conservation without a significant change in existing economic incentives advances our understanding of the range of large-scale behavioral science-based interventions that can be carefully applied at scale. Energy conservation strategies can be guided not only by traditional economic incentives such as rebates and price-based incentives but also by nonprice-based consumer disclosures concerning environmental and health damages not necessarily reflected in prices for electricity services.

Our study shows that nonprice incentives can effectively induce energy conservation, but it is not without limitations. First, our experiment provides both novel and repeated information to participants, making it difficult to separate the effect of learning from salience. Our participants acknowledged learning about appliance-level use and indicated that the appliance-level information was the most useful piece of information provided on the website. Most of them conveyed that they were surprised by how much or little electricity specific appliances were being used. In addition, the information provided on the dashboard was updated in real time, and participants received weekly emails. Further research should seek to disentangle the effect of learning about the energy use of different appliances from the saliency of the information we provided, which reminded them repetitively about their energy consumption. This raises the important question of how often should people be reminded about their electricity use to form energy conservation habits. Our exit survey indicates that the combination of weekly emails with the possibility to access real-time data on a website was sufficient in our setting. Further research is needed to understand energy use habit-forming behavior with repeated information provision. Second, we report behavioral outcomes within the 100-d treatment period but do not study the persistence of these household behavior changes after the conclusion of the experiment. We therefore do not know whether energy conservation persisted after the end of the experiment. However, the results from the exit survey indicate that some actions undertaken during the experiment could have potential lasting effects on energy consumption. Indeed, the majority of the participants described that they achieved reduced energy use by unplugging electronics, changing the power savings settings of their computer or other electronics, or programming different temperature settings on their thermostat. This is important because it suggests that the savings resulting from these changes could persist even without taking further action.

2.6 Policy implications

The relationship between electricity use and impacts on the environment and global health remains an elusive concept for many consumers. The generation of fine particulate air pollution and its effects on health are usually removed from ordinary daily consumer decision-making. This low consumer awareness stands in contrast to strides in our scientific understanding. We show that providing consumers specific, tailored, and scientifically verifiable information about the associated environmental and health effects of their electricity consumption can influence and motivate behavioral decision making about daily electricity use. More generally, this research advances our understanding of the effectiveness of information-based policies for conservation based on the principle that making information about the external damages of activities more salient to consumers can encourage conservation through household behavioral changes (59, 60). It has been argued that given the relative price inelastic behavior of electricity consumers in both the United States and the European Union, public policies to encourage energy conservation will require more than increases in electricity retail prices (9). Consumer information strategies can inform environmental policy about conservation efforts and can be used particularly where price-based strategies may not be politically feasible or effective. We argue that behavioral strategies in household electricity markets can be complements rather than substitutes for regulatory or price-based solutions.

Energy conservation is desirable in the economy as an alternative to costly capital investments in new power generation and can help delay managerial investment decisions for new generation capacity. Although nonprice behavioral strategies can be viable alternatives to new capital projects by promoting peak load shifting and conservation, they can also be implemented immediately, at scale and at relatively low cost (11). Behavioral strategies enabled through information technologies can be an effective component of sustainable development pathways and do not require long lead times typical of new capital investments in energy generation, distribution, and storage.

References

1. Dietz T, Gardner GT, Gilligan J, Stern PC, Vandenbergh MP (2009) Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. Proc Natl Acad Sci USA 106(44):18452–18456.
2. Pacala S, Socolow R (2004) Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. Science 305(5686):968–972.
3. Jaffe AB, Stavins RN (1994) Energy-efficiency investments and public policy. Energy J (Camb Mass) 15(2):43–65.
4. Jaffe AB, Stavins RN (1994) The energy-efficiency gap what does it mean. Energy Policy 22(10):804–810.
5. Jaffe AB, Stavins RN (1994) The energy paradox and the diffusion of conservation technology. Resour Energy Econ 16(2):91–122.
6. Fisher AC, Rothkopf MH (1989) Market failure and energy policy—A rationale for selective conservation. Energy Policy 17(4):397–406.
7. Brown MA (2001) Market failures and barriers as a basis for clean energy policies. Energy Policy 29(14):1197–1207.
8. Allcott H, Greenstone M (2012) Is there an energy efficiency gap? J Econ Perspect 26(1):3–28.
9. Azevedo IML, Morgan MG, Lave L (2011) Residential and regional electricity consumption in the U.S. and E.U.: How much will higher prices reduce CO2 emissions? Electr J 24(1):21–29.
10. Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V (2007) The constructive, destructive, and reconstructive power of social norms. Psychol Sci 18(5):429–434.
11. Allcott H, Mullainathan S (2010) Energy. Behavior and energy policy. Science 327(5970):1204–1205.
12. Brunekreef B, Holgate ST (2002) Air pollution and health. Lancet 360(9341): 1233–1242.
13. National Research Council (2010) Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use (National Academy Press, Washington, DC).
14. Thaler RH, Sunstein CR (2008) Nudge: Improving Decisions About Health, Wealth, and Happiness (Yale Univ Press, New Haven, CT).
15. Dietz T, Fitzgerald A, Shwom R (2005) Environmental values. Annu Rev Environ Resour 30(1):335–372.
16. Fransson N, Garling T (1999) Environmental concern: Conceptual definitions, measurement methods, and research findings. J Environ Psychol 19(4):369–382.

17. Attari SZ, DeKay ML, Davidson CI, Bruine de Bruin W (2010) Public perceptions of energy consumption and savings. Proc Natl Acad Sci USA 107(37):16054–16059.
18. Haidt J (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. Psychol Rev 108(4):814–834.
19. Greene J, Haidt J (2002) How (and where) does moral judgment work? Trends Cogn Sci 6(12):517–523.
20. Berkowitz L, Daniels LR (1964) Affecting the salience of the social responsibility norm: Effects of past help on the response to dependency relationships. J Abnorm Psychol 68(3):275–281.
21. Schwartz SH (1973) Normative explanations of helping behavior: A critique, proposal, and empirical test. J Exp Soc Psychol 9(4):349–364.
22. Cialdini RB, Reno RR, Kallgren CA (1990) A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. J Pers Soc Psychol 58(6):1015–1026.
23. Haidt J (2007) The new synthesis in moral psychology. Science 316(5827): 998–1002.
24. Frey BS, Meier S (2004) Social comparisons and pro-social behavior: Testing conditional cooperation in a field experiment. Am Econ Rev 94(5):1717–1722.
25. Meier S (2006) A Survey of Economic Theories and Field Evidence of Pro-Social Be- havior, FRB of Boston Working Paper No. 06-6. Available at http://dx.doi.org/10.2139/ssrn.917187.
26. Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? J Econ Perspect 21(2):153–174.
27. Ariely D, Bracha A, Meier S (2009) Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. Am Econ Rev 99(1):544–555.
28. Thaler R (1985) Mental accounting and consumer choice. Mark Sci 4(3):199–214.
29. Teck H, Lim N, Camerer CF (2006) Modeling the psychology of consumer and firm behavior with behavioral economics. J Mark Res 43(3):307–331.
30. Beruchashvili M, Gentry JW, Price LL (2006) Striving to be good: Moral balance in consumer choice. European Advances in Consumer Research, eds Ekstrom KM, Brembeck H (Association for Consumer Research, Goteborg, Sweden), Vol 7, pp 303–307.
31. Homans GC (1958) Social behavior as exchange. Am J Sociol 63(6):597–606.
32. Blau P (1960) A theory of social integration. Am Sociol Rev 65(6):545–556.
33. Wilk R (2001) Consuming morality. J Consum Cult 1(2):245–260.
34. Zelizer VA (2010) Moralizing consumption. J Consum Cult 10(2):287–291.
35. Barnett C, Cafaro P, Newholm T (2005) Philosophy and ethical consumption. The Ethical Consumer, eds Harrison R, Newholm T, Shaw D (Sage, London), pp 1–16.

36. Swanton C (2003) Virtue Ethics: A Pluralistic View (Oxford Univ Press, Oxford).
37. Moll J, et al. (2002) The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. J Neurosci 22(7): 2730–2736.
38. Knutson KM, et al. (2010) Behavioral norms for condensed moral vignettes. Soc Cogn Affect Neurosci 5(4):378–384.
39. Allcott H (2011) Social norms and energy conservation. J Public Econ 95(9-10): 1082–1095.
40. Ayres I, Raseman S, Shih A (2012) Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. J Law Econ Organ 29(5):992–1022.
41. Houde S, Todd A, Sudarshan A, Flora JA, Armel KC (2013) Real-time feedback and electricity consumption: A field experiment assessing the potential for savings and persistence. Energy J (Camb Mass) 34(1):87–102.
42. Abrahamse W, Steg L, Vlek C, Rothengatter T (2005) A review of intervention studies aimed at household energy conservation. J Environ Psychol 25(3):273–291.
43. Fischer C (2008) Feedback on household electricity consumption: A tool for saving energy? Energy Efficiency 1(1):79–104.
44. Vining J, Ebreo A (2002) Emerging theoretical and methodological perspectives on conservation behavior. New Handbook of Environmental Psychology, eds Bechtel R, Churchman A (Wiley, New York).
45. Delmas MA, Fischlein M, Asensio OI (2013) Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. Energy Policy 61(1):729–739.
46. Hirst E (1980) Review of data related to energy use in residential and commercial buildings. Manage Sci 26(9):857–870.
47. Chen VL, Delmas MA, Kaiser WJ (2014) Real-time, appliance-level electricity use feedback system: How to engage users? Energy Build 70(1):455–462.
48. Harrison GW, List JA (2004) Field experiments. J Econ Lit 42(4):1009–1055.
49. Burtless G (1995) The case for randomized field trials in economic and policy research. J Econ Perspect 9(2):63–84.
50. Delmas MA, Lessem N (2014) Saving power to conserve your reputation? The effectiveness of private versus public information. J Environ Econ Manage 67(3):353–370.
51. Frank KA, Maroulis SJ, Duong MQ, Kelcey BM (2013) What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. Educ Eval Policy Anal 35(4):437–460.

52. Imai K, King G, Stuart EA (2008) Misunderstandings between experimentalists and observationalists about causal inference. J R Stat Soc [Ser A] 171(2):481–502.
53. Reiss PC, White MW (2005) Household electricity demand, revisited. Rev Econ Stud 72(3):853–883.
54. Reiss PC, White MW (2008) What changes energy consumption? Prices and public pressures. Rand J Econ 39(3):636–663.
55. Energy Information Association (2009) Residential Energy Consumption Survey (RECS) 2009 (US Department of Energy, Washington, DC). Available at www.eia.gov/consumption/residential/data/2009/.
56. Gneezy U, Meier S, Rey-Biel P (2011) When and why incentives (don't) work to modify behavior. J Econ Perspect 25(4):191–210.
57. Attari SZ (2014) Perceptions of water use. Proc Natl Acad Sci USA 111(14):5129–5134.
58. Nolan JM, Schultz PW, Cialdini RB, Goldstein NJ, Griskevicius V (2008) Normative social influence is underdetected. Pers Soc Psychol Bull 34(7):913–923.
59. Stern PC (1992) What psychology knows about energy conservation. Am Psychol 47(10):1224–1232.
60. Delmas MA, Montes M, Shimshack J (2010) Mandatory information disclosure policy: Evidence from the electric utility industry. Econ Inq 48(2):483–498.

Appendix 1

Supporting information

Materials and Methods

We outfitted 118 family apartments with wireless energy metering technology at a residential housing community in Los Angeles. We measured electricity use data in real-time 24 hours a day at the appliance level. The randomized controlled trial was conducted from October 2011 to July 2012 and weekly treatment messages were sent to participants. The first group of apartments was given detailed energy use feedback along with information about monetary savings. The second group was given feedback with a health message about emissions and its air quality impacts such as childhood asthma. The third group served as a statistical control following a six-month baseline period and random assignment. Fig. S1 shows an information graphic of the website shown to participants. No financial transfers or monetary rewards were offered for participation.

Field Site. Our field experimental site, University Village, is a large residential community located in proximity to public transportation, local businesses, parks and schools. It is a multiple building, family apartment/condo-style housing complex with 1,103 units. The community spans two census block groups serviced by the Los Angeles Department of Water and Power (LADWP), the nation's largest public utility. Although the facilities are owned and operated by the University of California, the University does not subsidize living costs for the community and offers market-based rental rates. All utilities are paid by the tenant, including electricity. While apartments vary in size and layout, all units are furnished with a common set of appliances—a refrigerator, gas stove, dishwasher, and microwave oven. This allows for standardization in the housing capital stock. We monitor direct electricity usage in each of the participant households.

Fig. S1 *Visualization of health-based messaging feedback for consumers*

Treatment Messages. Information treatments received by households contain: (i) a neighbor comparison, which provides a reference point for their household consumption, and (ii) a stated impact of electricity use, either in terms of potential cost savings or public health externalities. The specific treatment messages are listed in Table S1.

Table S1. Treatment Messages

Group	Treatment Message
Monetary Savings Group	"Last week, you used 66% more/less electricity than your efficient neighbors. In one year, this will cost you (you are saving) $34 dollars extra."*
Health Group	"Last week, you used 66% more/less electricity than your efficient neighbors. You are adding/avoiding 610 pounds of air pollutants which contribute to health impacts such as childhood asthma and cancer."*
Control Group	None.

* Efficient neighbor in this context means households in the top 10th percentile of household weekly average kWh consumption (households with the lowest electricity use) for similar size apartments in the community.

Neighbor comparisons are standardized in the following form: "Last week, you used ___% more/less electricity than your efficient neighbors" Neighbor comparisons in the energy conservation context have gained broad use in (i) small-scale lab or field studies, typically in applied social psychology, building-science and engineering, and (ii) utility-scale pilot projects, typically in economics and related fields. Impacts described were presented to households in numerical and scientifically verifiable terms. Unlike many lab studies where numerical impacts may be the subject of manipulation, we provided households with factual evidence-based numbers that depend on their weekly consumption. Equivalent cost savings were calculated using household-level consumption data and the published LADWP electric rate schedules for residential customers. Equivalent pounds of air pollutant emissions were calculated using emission factors from the Emissions & Generation Resource Integrated Database (eGRID) maintained by the U.S. EPA and based on LADWP electricity fuel mix. Treatment messages were also pre-tested in a series of questionnaires for clarity, comprehension and stated willingness-to-save energy with independent populations.

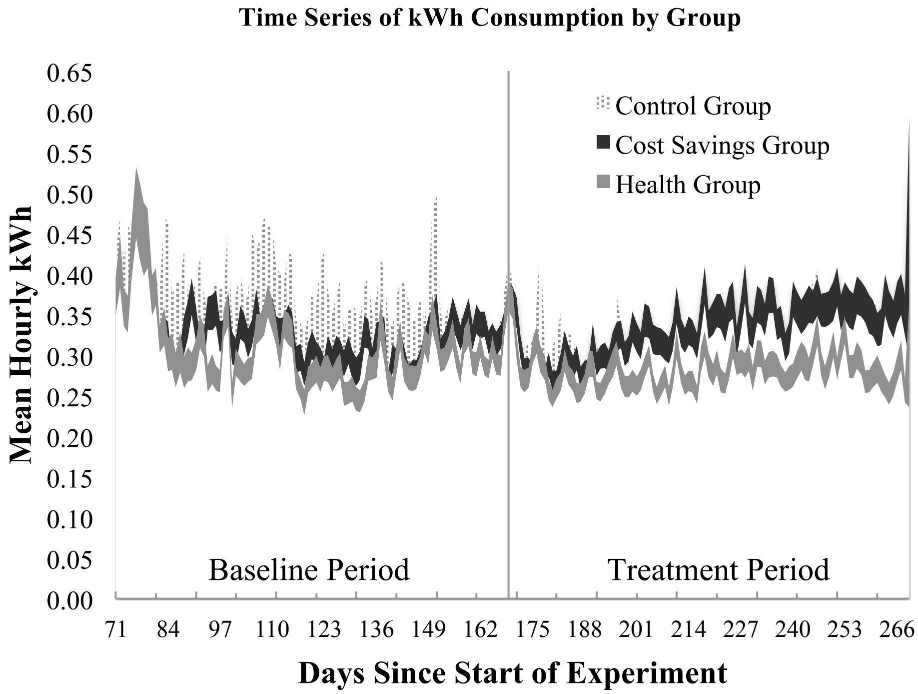


Fig. S2 *Time series of hourly community consumption. During the baseline period, the mean hourly kWh consumption is overlapping for all three groups. After treatment begins, the consumption diverges. Treatment effects are identified by difference-in-differences before-after-control-impact design.*

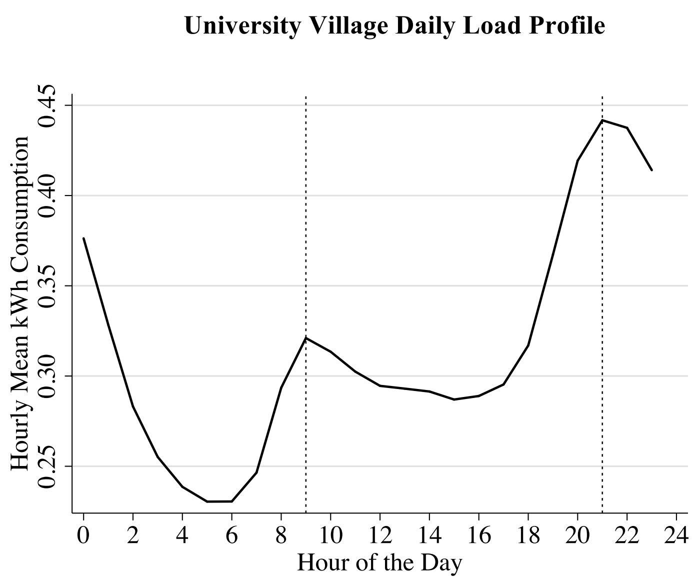


Fig. S3 *Peak daily consumption for the community occurs at 9:00am and 9:00pm*

Participant Recruitment. Households were recruited to participate in the study. In order to prevent biases in recruitment selection, no direct environmental messaging was used. The recruitment process occurred within the context of several community events and information campaigns during the summer months prior to the start of the 2011-2012 academic year. To meet all Institutional Review Board (IRB) ethics requirements regarding research with human subjects, participation was strictly voluntary and no personally identifiable information (PII) was collected or shared. We conducted an enrollment survey to capture basic apartment demographics and occupancy characteristics for the community at-large, including households who opted in and those who opted out of the study. We recruited many more willing participants than there were active equipment allotments. Among the 1,103 households at University Village, 226 households volunteered to participate and another 88 households in our entry survey chose not to participate. This equals a participation rate of 20%. We randomly selected 118 participating households from these 226 volunteers. The participating households in our experiment represent 10.7% of the population at University Village. Household assignment into treatment and control groups was then randomized. Fig. S2 shows a time series of the community consumption and Fig. S3 shows the daily load profile over a 24 hour period.

While households could at any point withdraw their consent to participate, no households dropped from the study for the entire duration of the experiment.

We tested for potential differences between the population of households at our field site and our sample of volunteer participants. We compared the monthly electricity meter readings of the entire population of University Village to those of our participants as well as other characteristics such as the size of the apartment, the number of occupants, the apartment floor and the location of the apartment in the complex. As shown in Table S2, there are no significant differences between participating and non-participating households. This analysis is based on electricity meter readings for 12 months prior to the start of the experiment.

Table S2. Comparison of Participating versus Non-Participating Households at University Village (Meter Readings Data)

	Participating Households (S.D.)	Non-Participating Households (S.D.)	Mean Difference (S.D.)	$Y_{in} - Y_{out}$ (S.E.)
Electricity Consumption [§]				
Average kWh per day	8.429	8.737	.3070	-.0004
	(15.2)	(28.7)	(32.5)	(.0004)
kWh per square foot	.2007	.2043	.0036	.0833
	(.339)	(.479)	(.587)	(.198)
kWh per person	42.53	44.72	2.18	-.0003
	(68.5)	(108.8)	(128.6)	(.0009)
Square Footage	859.79	868.83	9.04	-.0001
	(106.3)	(98.54)	(144.9)	(.0002)
Number of bedrooms	1.97	1.97	-0.003	-.0263
	(.379)	(.343)	(.511)	(.160)
Number of bathrooms	1.60	1.65	.05	.0143
	(.490)	(.474)	(.681)	(.040)
Number of occupants	4.03	4.01	-.02	-.0107
	(.566)	(.512)	(.763)	(.126)
Building Floor	2.08	2.08	.002	-.0308
	(.808)	(.786)	(1.12)	(.021)
Location in Complex	.543	.596	.053	-.041
(1 if Sawtelle, 0 if Sepulveda)	(.498)	(.491)	(.699)	(.040)
Number of Households	118	986	1,104	1,104
Number of Observations	5,533	46,184	51,718	51,718
F-test *p*-value	-	-	-	0.669

[§] *Based on 12 months of independent electricity meter readings. Coefficients for kWh per square foot and kWh per person are based on independent regressions. No significant differences are found.*

Empirical Strategy. We modeled the household behavioral outcomes as a time series of electricity consumption before and after the start of information treatments. Our general empirical strategy consists of panel regressions of total and appliance-level electricity loads on a series of treatment group indicators and important statistical controls: household and occupancy characteristics, a proxy for household environmental leaning and seasonal variables including weather and time trends. To estimate the treatment effects on the study population, we use an analytical approach by difference-in-differences (DD). We define “treatment” to mean weekly updating informational messages about household energy use defined previously. Treatments are exhaustive and mutually exclusive, meaning each household receives only one randomly assigned treatment. Once assigned, there is no crossover between treatments.

A control group is also monitored alongside the treatment groups, but receives no information. See Appendix 3 for a theoretical model.

Identification. In keeping with our identification strategy, we define treatment dummies denoting treatment *group* and event *time* status. Let $\widehat{T}_i$ be the binary *treatment group indicator*, equal to 1 if household is a member of treated $\mathbb{E}[\cdot]$ group *i,* and 0 otherwise. Let *P* be the binary *post-treatment indicator*, equal to 1 after the start of information treatments (i.e., post-treatment period), and 0 during the baseline period (i.e., pre-treatment period). Let denote the expectations operator. The behavioral response function y_j for household *j* is allowed to be heterogeneous at the household level. Conditioning on observables, we define the average treatment effect on the treated (ATET) as:

$$\tau^{DID} \equiv \left[\mathbb{E}[y_j \mid X, \widehat{T}_i = 1, P = 1] - \mathbb{E}[y_j \mid X, \widehat{T}_i = 0, P = 1] \right] \qquad \text{(post-treatment period)}$$

$$- \left[\mathbb{E}[y_j \mid X, \widehat{T}_i = 1, P = 0] - \mathbb{E}[y_j \mid X, \widehat{T}_i = 0, P = 0] \right] \qquad \text{(pre-treatment period)} \quad (1.1)$$

Treatment is identified when the group-time $\{(\widehat{T}_i = 1) \times (P = 1)\}$ interaction equals 1 over all feasible treatments $\{i = 1,2\}$ (i.e., monetary savings or health). The ATET in Equation 1.1 is the population average difference in the *control group* $\widehat{T}_i = 0$ before and after treatment minus the population average difference in the *treated group* $\widehat{T}_i = 1$ before and after treatment. We condition on household level covariates, X. Any common unobservable characteristics are also captured in the control group.

We make two identifying assumptions for the estimation of treatment effects. First, treatment selection is independent of the behavioral response function, which is given by randomization. Second, treatments are independent and mutually exclusive. See Appendix 2 for a proof of identification.

Dependent Variable. Our dependent variable and behavioral response measure is the total kilowatt-hour (kWh) electric power consumption. A kilowatt-hour (kWh) is the most common unit of electricity used by electric utilities in commercial and residential billing. We aggregate real-time electricity measurements into hourly observations. Our total kWh signal for each household is further decomposed into one of six major appliance categories. By direct measurement, the appliance-level kWh consumption categories are: (i) lighting, (ii) heating and cooling, (iii) plug load, (iv) refrigerator, (v) dishwasher, and (vi) other kitchen (including the microwave and kitchen outlets). These six appliance categories make up the complete circuit breaker distribution for all electricity uses in the household. We note that this level of granularity in kWh measurement is unique to our installed metering technology and wireless sensor network. We normalize our dependent variable by dividing by the average post-treatment control group consumption, and multiplying by 100, allowing us to interpret our regression coefficients directly as percentages versus control group. We do not use logs as monotonic transformations of the hourly kWh measurements since appliance-level electricity loads in the range $[0, R^{+})$ can frequently be equal or close to zero, for example, when the dishwasher or other appliance is off. For other examples of this normalization approach with electricity metering data, see (1). The distribution of dependent variables is shown in Table S3.

Table S3. Distributions of Dependent Variables (hourly kWh measurements)

Percentiles	Total	Heating Cooling	Lighting	Plug Load	Refrigerator	Dishwasher	Other Kitchen
1%	0.0044	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5%	0.0824	0.0007	0.0000	0.0075	0.0050	0.0000	0.0000
10%	0.1038	0.0029	0.0000	0.0160	0.0187	0.0000	0.0000
25%	0.1442	0.0083	0.0008	0.0337	0.0446	0.0000	0.0000
50%	0.2288	0.0156	0.0160	0.0616	0.0675	0.0000	0.0041
75%	0.4017	0.0239	0.0702	0.1197	0.0894	0.0060	0.0149
90%	0.6321	0.1807	0.1351	0.2200	0.1275	0.0134	0.0742
95%	0.8236	0.3704	0.1850	0.2941	0.1455	0.0487	0.1402
99%	1.3374	0.8098	0.3091	0.6311	0.1794	0.1167	0.3857
Mean	0.3157	0.0622	0.0471	0.1105	0.0701	0.0084	0.0277
Std. Dev.	0.2746	0.1622	0.0700	0.2739	0.0402	0.0297	0.1061
Observations	490,994	490,994	490,994	490,994	490,994	490,994	490,994

Independent Variables. The variables of interest are the treatment group indicators, observable household characteristics, and seasonal controls including weather and time trends. *Household occupancy* includes the number of adults (ranging from 0 to 3), and number of children (ranging from 0 to 4). *Apartment size* indicates the number of bedrooms in the unit, ranging from 1 to 3 bedrooms. *Building floor* captures apartment elevation, ranging from 1 to 3, where 1st floor indicates ground level. *Floor plan* captures differences in apartment layout, measured in nominal square footage. Because political leaning or ideology can significantly impact energy use attitudes and behaviors (2-4), we include statistical controls for household environmentalist ideology to account for the fact that greener participating households might have more proclivities toward conservation. To this end, *member environmental organization* is a proxy variable, which captures a fixed measure of household environmentalist ideology or orientation. It is equal to 1 if the head of household reports being an active member of an environmental non-governmental organization (NGO), and 0 otherwise.

Seasonality and Time Trends. Electricity demand (in kWh per unit time) exhibits seasonal fluctuations and serial correlation that depends on outside factors such as time of day or weather. Modeling electricity loads with high time-resolution data requires special consideration of seasonality and time-varying characteristics on consumption, most notably, the effects of outside temperatures on hourly energy demand. Even with the milder climate in Los Angeles, heating and cooling hours capture significant seasonal variation on electricity consumption. We calculate heating and cooling degree hours, using quality-controlled local weather data from the Santa Monica Municipal Airport weather station, as maintained by the National Climatic Data Center (NCDC). Outside dry bulb temperatures were recorded hourly at the Santa Monica Municipal Airport weather station, located less than 1 mile from the study site. Archival access was provided by the National Oceanic and Atmospheric Administration (NOAA's) Quality Controlled Local Climatological Data (QCLCD), which contains hourly, daily and monthly summaries of outside weather conditions for the specific station. Mean degree-hours are a fundamental measure in building energy management that expresses the magnitude of expected heating or cooling load at a given location. Degree-hours capture seasonal heating or cooling requirements at a finer resolution than degree-days, making our hourly kWh observations compatible with outside weather variation.

The weather vector is $\mathbf{\Psi}_t = [\Psi_t^H, \Psi_t^C]$ where:

$$\Psi_t^H = \max\left\{0,\ \sum_{h=1}^{24}\left(\theta_b - \theta_{out}\right)\right\} \quad \text{heating degree hours}$$

$$\Psi_t^C = \max\left\{0,\ \sum_{h=1}^{24}\left(\theta_{out} - \theta_b\right)\right\} \quad \text{cooling degree hours} \tag{1.2}$$

As shown in Equation 1.2, the larger the indoor heating or cooling requirement, the larger the distance between the measured mean hourly outside temperature θ_{out} and a given base temperature θ_b. By U.S. convention, the indoor base temperature θ_b is defined as 65˚F (18.3˚C) (5). When outside temperatures rise above the given indoor base temperature, cooling degree hours are strictly positive and heating degree hours are zero. Conversely, when outside temperatures fall below the base temperature, heating degree hours are strictly positive and cooling degree hours are zero. In this way, differential effects of heating and cooling load on electricity consumption are decomposed in a meaningful way over a 24-hour period. By rigorously specifying heating and cooling degree hours, we directly mitigate issues of seasonality and serial correlation in the disturbances of the regression model, addressing some methodological limitations previously identified in the literature (6).

Econometric Model. The basic econometric specification for household *j*, in treatment group *i*, at time *t*, is

$$E_{jit} = \alpha P_i + \tau(P_i \cdot T_i) + \mathbf{H}_j + \mathbf{\Psi}_t + \gamma_t + c + \varepsilon_{jt} \tag{1.3}$$

The dependent variable, E_{jit}, represents hourly panel observations of total and appliance-level electricity loads. Our main coefficient of interest, $\hat{\tau}$, indicates the *average treatment effect on the treated* and the coefficient α indicates the *post-treatment on the population*. $\mathbf{H}_j$ is the vector of household covariates and $\mathbf{\Psi}_t$ is the weather vector. We include degree hours of the study period and day of the week time dummies to control for common time trends. Time dummies offer a convenient and robust control for community-wide effects. The regression constant is denoted by c and the residual error is captured in ε_{jt}. We mitigate the effects of serial correlation—a common source of estimation bias in difference-in-differences models (7) by fully specifying important seasonal weather variables on consumption and clustering the standard errors at the household level. Our standard errors are satisfactory due to a number of important design considerations. First, we have very high-resolution measurement, down to individual appliances, in which both make and model of all appliances have been standardized across the community. This provides for more precise behavioral estimates than are otherwise available in comparable studies with monthly residential billing data. Second, we control for the impact of seasonality and time-varying characteristics on consumption by use of degree hours, which offers a finer resolution controls for weather variability than typical approaches that use heating and cooling degree-days, or that have no weather controls at all (6). In addition to seasonal degree-hours, we also specify time dummies to capture common time trends (or cycles) in the data and any calendar shocks on consumption. We estimate treatment effects in Equation 1.3 conservatively by difference-in-differences using the standard feasible generalized least squares estimator (FGLS), $\hat{\beta}_{GLS} = (\mathbf{X}\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}^{-1}\mathbf{y}$ (8). We note that GLS panel estimation is feasible because the panel's time dimension is larger than the cross-sectional dimension of N households, a characteristic of our high time-resolution data set. In the next section, we also present alternative results at various sampling frequencies and show robustness checks using OLS.

Tables S4–S8 contain the supporting regression tables. Table S4 lists treatment effects on families with and without children. Table S5 lists treatment effects by appliance and Appendix 4 shows appliance dynamics. Table S6 lists treatment effects by time of day. Table S7 contains descriptives and correlations for all variables in the study and Table S8 contains the supporting conditional quantile regressions.

Table S4. Heterogeneous Treatment Effects on Families with Children

Study Variables	(1) Total kWh	(2) Total kWh	(3) Total kWh
Experimental			
Post-Treat*Monetary Savings Group	**3.785**	**1.688**	**3.771**
	(4.391)	(5.221)	(4.391)
Post-Treat*Health Group	**-8.215****	**-8.206****	**-1.419**
	(4.120)	(4.119)	(4.862)
Post-Treat*Monetary Savings Group*Children=1 or more		**7.831**	
		(11.32)	
Post-Treat*Health Group*Children=1 or more			**-19.07****
			(8.998)
Monetary Savings Group	1.853	1.531	2.478
	(7.814)	(7.722)	(7.844)
Health Group	-1.383	-1.542	-0.844
	(8.033)	(8.022)	(8.053)
Household Characteristics			
Adults	4.003	3.705	3.400
	(8.556)	(8.419)	(8.557)
Children (1 or more)	17.91**	16.63**	21.42***
	(7.494)	(6.923)	(7.780)
Apartment Size (No. of bedrooms)	33.01*	32.44*	32.03*
	(16.95)	(16.96)	(17.06)
Floor Plan (Nominal square footage)	-0.0109	-0.00983	-0.00852
	(0.0612)	(0.0610)	(0.0613)
Building Floor	9.854***	9.732***	9.265***
	(3.400)	(3.384)	(3.426)
Ideology			
Member Environmental Organization	-7.222	-7.464	-8.908
	(9.076)	(8.937)	(8.884)
Hourly Weather Controls			
Heating Degree Hours	0.284	0.286	0.281
	(0.255)	(0.254)	(0.255)
Cooling Degree Hours	-0.811***	-0.809***	-0.813***
	(0.186)	(0.186)	(0.186)
Time Dummies			
Day-by-Week	Yes	Yes	Yes
Constant	12.94	14.59	13.83
	(33.75)	(33.45)	(33.48)
Observations	490,994	490,994	490,994
Number of Apartments	118	118	118
R^2	0.0437	0.0451	0.0454

Robust standard errors clustered at the household level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table S5. Treatment Effects by Appliance

Study Variables	(4) Heating Cooling	(5) Lighting	(6) Plug Load	(7) Refrigerator	(8) Dishwasher	(9) Other Kitchen
Experimental						
Post-Treat*Monetary Savings Group	**5.331***	**-11.46*****	**0.414**	**8.844*****	**3.260**	**0.987**
	(2.779)	(4.274)	(2.395)	(2.153)	(3.918)	(3.056)
Post-Treat*Health Group	**-2.567**	**-9.011*****	**-4.719****	**8.673*****	**-3.790**	**-1.370**
	(2.554)	(2.324)	(2.152)	(1.981)	(2.471)	(4.454)
Monetary Savings Group	3.248	-20.61	-81.42	18.37*	-16.68	-38.79
	(3.189)	(17.73)	(51.13)	(9.372)	(24.42)	(25.29)
Health Group	6.370**	-19.32	-87.15*	15.41*	-37.06*	-37.81
	(3.129)	(14.29)	(48.52)	(9.161)	(22.20)	(24.93)
Household Characteristics						
Adults	-0.839	-6.284	-2.518	16.83*	-11.89	16.16
	(3.165)	(16.27)	(18.95)	(10.13)	(14.45)	(17.74)
Children (1 or more)	3.982	0.650	-22.42	11.11*	-4.389	-1.703
	(2.909)	(14.67)	(27.07)	(6.722)	(14.02)	(14.67)
Apartment Size (No. of bedrooms)	3.792	53.26	-80.41	40.39***	28.58	39.46
	(6.700)	(43.85)	(56.84)	(14.94)	(29.88)	(28.09)
Floor Plan (Nominal square footage)	0.00887	-0.0676	0.226	-0.102*	0.00103	-0.105
	(0.0232)	(0.0958)	(0.231)	(0.0547)	(0.102)	(0.0975)
Building Floor	1.661	-8.622	-5.106	13.81***	5.492	3.162
	(1.553)	(8.345)	(22.40)	(4.016)	(8.502)	(8.908)
Ideology						
Member Environmental Organization	-6.223**	-10.97	-0.228	-3.647	-14.51	3.426
	(2.742)	(9.532)	(16.15)	(11.24)	(15.75)	(17.68)
Weather Controls						
Heating and Cooling Degree Hours	Yes	No	No	Yes	No	No
Time Dummies						
Hour-by-Day, Day-by-Week	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-10.08	128.5***	169.5	50.52*	67.38	95.00
	(12.59)	(41.70)	(141.8)	(30.67)	(53.80)	(68.10)
Observations	490,994	490,994	490,994	490,994	490,994	490,994
Number of Apartments	118	118	118	118	118	118
R^2	0.0163	0.145	0.0316	0.0964	0.0159	0.0124

Robust standard errors clustered at the household level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table S6. Treatment Effects by Time of Day

Study Variables	(10) Midnight - 3:00am	(11) 3:00- 6:00am	(12) 6:00- 9:00am	(13) 9:00- 12:00pm	(14) 12:00- 3:00pm	(15) 3:00- 6:00pm	(16) 6:00- 9:00pm	(17) 9:00- Midnight
Experimental								
Post-Treat*Monetary Savings Group	**-6.919**	**-0.351**	**4.434**	**1.965**	**8.999***	**15.46*****	**20.22*****	**5.346**
	(4.960)	(4.890)	(5.241)	(5.953)	(4.865)	(5.215)	(6.024)	(6.099)
Post-Treat*Health Group	**-17.51*****	**-12.01****	**-11.43***	**-10.13***	**-1.689**	**6.665**	**5.027**	**-5.725**
	(4.328)	(5.048)	(6.131)	(6.110)	(4.132)	(4.354)	(5.028)	(4.871)
Monetary Savings Group	2.845	-5.393	-2.503	2.788	2.180	0.467	2.420	5.663
	(9.320)	(8.854)	(8.023)	(9.561)	(9.127)	(9.267)	(10.12)	(10.34)
Health Group	-1.928	-0.428	5.402	-2.706	-3.751	-4.329	-3.850	-5.858
	(9.900)	(9.620)	(9.740)	(9.513)	(8.211)	(8.317)	(9.220)	(10.18)
Household Characteristics	2.845	-5.393	-2.503	2.788	2.180	0.467	2.420	5.663
Adults	3.251	-8.546	-9.369	-0.283	7.129	12.16	15.51	13.51
	(9.976)	(9.802)	(10.04)	(10.83)	(10.06)	(10.59)	(12.69)	(9.828)
Children	14.38*	10.79	15.81**	24.16***	18.79**	18.75**	21.73**	18.74*
	(7.690)	(6.931)	(6.373)	(9.312)	(9.292)	(9.383)	(10.32)	(9.836)
Apartment Size (No. of bedrooms)	28.36	28.26	38.97**	34.30	24.86	22.91	36.21	49.94**
	(19.41)	(17.44)	(16.59)	(20.93)	(20.01)	(20.36)	(22.79)	(20.15)
Floor Plan (Nominal square footage)	-0.0352	-0.0410	-0.0402	-0.00901	0.0143	0.0236	0.0222	-0.0177
	(0.0689)	(0.0605)	(0.0561)	(0.0697)	(0.0694)	(0.0727)	(0.0816)	(0.0782)
Building Floor	9.115**	6.130*	11.25***	8.551**	7.538*	8.820**	12.79***	14.81***
	(3.737)	(3.533)	(3.462)	(4.260)	(3.927)	(4.259)	(4.840)	(4.358)
Ideology								
Member Environmental Organization	-7.491	-4.355	-7.588	-4.960	-4.180	-2.367	-10.66	-15.94
	(9.676)	(9.151)	(9.098)	(10.14)	(9.862)	(11.07)	(12.95)	(10.92)
Hourly Weather Controls								
Heating Degree Hours	0.800***	1.251***	0.746***	1.258***	0.188	0.119	0.765**	0.579
	(0.290)	(0.269)	(0.269)	(0.241)	(0.219)	(0.219)	(0.356)	(0.395)
Cooling Degree Hours	2.662	-5.304***	3.932***	-0.245	-0.157	0.517**	-0.591	0.180
	(4.208)	(1.936)	(0.764)	(0.181)	(0.189)	(0.260)	(0.681)	(1.294)
Time Dummies								
Day-by-Week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	48.54	49.73	25.34	9.612	-5.217	-23.34	-39.93	-2.685
	(38.41)	(36.16)	(34.22)	(38.02)	(37.89)	(41.02)	(45.58)	(41.68)
Observations	60,942	60,433	61,206	61,543	61,402	61,581	61,891	61,996
Number of Apartments	118	118	118	118	118	118	118	118
R^2	0.0404	0.0521	0.0762	0.0616	0.0558	0.0542	0.0567	0.0630

Robust standard errors clustered at the household level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table S7. Means, Standard Deviations and Correlations

		Mean	S.D.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	Total kWh (normalized)	103.11	89.71	0.0	3489.1	1.00											
Experimental																	
(2)	Health Group	0.37	0.48	0.0	1.0	-0.06*	1.00										
(3)	Monetary Savings Group	0.38	0.49	0.0	1.0	0.02*	-0.61*	1.00									
(4)	Control Group	0.24	0.43	0.0	1.0	0.04*	-0.44*	-0.45*	1.00								
Household Characteristics																	
(5)	Number of Adults	1.93	0.29	1.0	3.0	-0.01*	-0.18*	0.12*	0.07*	1.00							
(6)	Number of Children	0.52	0.81	0.0	4.0	0.14*	-0.04*	-0.08*	0.13*	-0.10*	1.00						
(7)	Apartment Size (beds)	1.97	0.38	1.0	3.0	0.15*	-0.14*	0.04*	0.12*	-0.09*	0.30*	1.00					
(8)	Floor Plan (nominal sq.ft.)	862.3	104.49	595	1035	0.14*	-0.14*	0.05*	0.10*	0.06*	0.20*	0.83*	1.00				
(9)	Building Floor	2.07	0.81	1.0	3.0	0.08*	0.04*	-0.13*	0.10*	0.07*	-0.05*	0.03*	0.05*	1.00			
Ideology																	
(10)	Member Env. Organization	0.09	0.28	0.0	1.0	-0.02*	0.00	0.10*	-0.11*	-0.15*	-0.01*	0.02*	-0.03*	-0.05*	1.00		
Weather Controls																	
(11)	Heating Degree Hours	7.15	5.76	0.0	26.0	0.03*	-0.01*	-0.01*	0.02*	0.00	0.00*	0.00	0.00	-0.01	-0.01	1.00	
(12)	Cooling Degree Hours	0.6	1.94	0.0	26.0	-0.02*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.39*	1.00

N = 440,059 panel observations (118 apartments)
* p <.05

Table S8. Quantile Regression Estimates

	Quantiles				
Study Variables	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Experimental					
Post-Treat*Monetary Savings Group	**3.835*****	**2.597*****	**2.006*****	**-1.031**	**4.912*****
	(0.153)	(0.174)	(0.382)	(0.801)	(1.587)
Post-Treat*Health Group	**1.907*****	**-0.428****	**-5.898*****	**-13.98*****	**-15.52*****
	(0.170)	(0.175)	(0.339)	(0.632)	(1.200)
Monetary Savings Group	-2.551***	-2.980***	-5.655***	-2.029***	18.34***
	(0.112)	(0.159)	(0.331)	(0.667)	(1.212)
Health Group	-3.137***	-2.574***	-5.780***	-6.069***	7.288***
	(0.174)	(0.117)	(0.268)	(0.541)	(0.978)
Household Characteristics					
Adults	-0.908***	0.622***	-5.351***	-2.349***	10.28***
	(0.174)	(0.141)	(0.563)	(0.760)	(1.598)
Children	4.465***	8.248***	20.37***	26.86***	32.90***
	(0.0942)	(0.139)	(0.256)	(0.595)	(0.920)
Apartment Size (No. of Bedrooms)	6.211***	11.92***	27.41***	40.44***	36.51***
	(0.353)	(0.289)	(0.511)	(0.941)	(1.231)
Floor Plan (Nominal square footage)	0.00368***	0.00481***	-0.0121***	-0.00597*	0.0626***
	(0.00124)	(0.000914)	(0.00166)	(0.00337)	(0.00439)
Building Floor	3.760***	4.814***	6.871***	10.16***	14.93***
	(0.0810)	(0.0748)	(0.125)	(0.226)	(0.343)
Ideology					
Member Environmental Organization	0.850***	-1.488***	-0.260	-6.505***	-23.09***
	(0.135)	(0.155)	(0.377)	(0.469)	(1.182)
Hourly Weather Controls					
Heating Degree Hours	-0.0523***	-0.0566***	-0.0626***	0.687***	1.642***
	(0.00894)	(0.00895)	(0.0196)	(0.0473)	(0.0763)
Cooling Degree Hours	-0.252***	-0.380***	-0.823***	-1.129***	-0.580***
	(0.0287)	(0.0231)	(0.0507)	(0.116)	(0.223)
Time Dummies					
Day-by-Week	Yes	Yes	Yes	Yes	Yes
Constant	13.65***	9.360***	28.95***	31.59***	-4.360
	(0.687)	(0.521)	(1.369)	(2.381)	(4.052)
Observations	490,994	490,994	490,994	490,994	490,994
Number of Apartments	118	118	118	118	118
Pseudo R-squared	.0119	.0199	.0337	.0403	.0397

Quantile treatment effects with bootstrap standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table S9. Comparison of Baseline Usage Characteristics Between Treated and Control Households

	Control Group (S.D.)	Treatment Group 1: (S.D.)	Treatment Group 2: (S.D.)	Difference Treat 1- Control (S.D.)	Difference Treat 2 - Control (S.D.)	$Y_0^T - Y_0^C$ (S.E.)
Average kWh usage/Day	8.660	7.543	7.457	-1.118	-1.204	-0.000377
	(7.623)	(6.485)	(6.672)	(10.01)	(10.13)	(0.00195)
Apartment Size (bedrooms)	2.043	1.980	1.914	-0.063	-0.128	-0.153
	(0.394)	(0.339)	(0.358)	(0.520)	(0.532)	(0.205)
No. of Adults	1.968	1.970	1.847	0.002	-0.122	-0.105
	(0.175)	(0.271)	(0.360)	(0.322)	(0.401)	(0.106)
No. of Children	0.653	0.425	0.480	-0.227	-0.172	-0.0562
	(0.800)	(0.874)	(0.713)	(1.184)	(1.072)	(0.0572)
Floor Plan (Square Footage)	877.66	867.17	846.04	-10.49	-31.62	0.000203
	(97.451)	(97.019)	(108.761)	(137.51)	(146.03)	(0.000674)
Building Floor	2.163	1.919	2.103	-0.244	-0.060	-0.0494
	(0.861)	(0.813)	(0.760)	(1.184)	(1.148)	(0.0501)
Member Environmental Organization	0.024	0.119	0.082	0.096	0.058	0.157*
	(0.152)	(0.324)	(0.274)	(0.358)	(0.313)	(0.0835)
Number of Observations	119,609	187,684	183,701	307,293	426,902	371,385
Number of Households	33	43	42	76	75	118
F-test *p*-value						0.2485

6 month baseline period (no electricity use feedback) *** p<0.01, ** p<0.05, * p<0.1

Baseline Characteristics. Table S9 shows descriptive statistics for both treated and control households during the 6-month baseline period. As shown in Table S9, the covariates and electricity consumption are reasonably balanced between treated and control households. In particular, the average electricity consumption is statistically indistinguishable between groups along with other important household fixed effects. The last column in Table S9 shows the results of a regression testing for significant differences between groups. As given by the F-test p-value of 0.2485, we reject a hypothesis of imbalance between groups. One exception is the variable representing membership of an environmental organization, which is significant at the 10 percent level. We note that households who report membership in an environmental organization represent a very minor share (~8%) of households in the study. In separate results, we computed the effect of belonging to an environmental organization as a proxy for green behavior. These results show no significant interaction with either treatment (results available from the authors upon request). This indicates that environmentalist households are not driving the study's main results.

Robustness checks. Table S10 shows the ATE specifications using OLS. Table S10 lists results of standard protocols with robust standard errors clustered at the household level, starting with a simple comparison between treatment and control groups and subsequently adding covariates.

Table S10. ATE Specifications, OLS (Hourly Sampling)

	I	II	III	IV	V
	Total kWh	Total kWh	Total kWh	Total kWh	Total kWh
Post-Treat*Cost Savings Group	**5.210**	**3.917**	**3.915**	**3.822**	**5.297**
	(5.019)	(4.966)	(4.968)	(4.972)	(4.533)
Post-Treat*Health Group	**-9.958****	**-9.694****	**-9.682****	**-9.833****	**-8.192***
	(4.656)	(4.648)	(4.647)	(4.652)	(4.306)
Treat Cost Savings	-7.302	2.801	2.797	2.902	2.238
	(8.488)	(7.303)	(7.303)	(7.298)	(7.382)
Treat Health Group	-8.469	-0.157	-0.173	-0.035	-0.795
	(8.870)	(8.085)	(8.086)	(8.090)	(8.060)
Degree-hour bins	No	No	No	No	Yes
Apartment fixed effects	No	Yes	Yes	Yes	Yes
Day x Week time dummies	No	No	Yes	Yes	Yes
Hour x Day time dummies	No	No	No	Yes	No
Observations	490,994	490,994	490,994	490,994	490,994
R^2	0.005	0.043	0.044	0.094	0.044
F-statistic	2.549	3.627	9.117	27.480	8.985
Number of households	118	118	118	118	118

Robust standard errors clustered at the household level *** p<0.01, ** p<0.05, * p<0.1

Column I shows a simple comparison between treatment and control groups in the post-period, without adjustment for the covariates. We obtain a -9.9% point estimate of the treatment effect in the health treatment group, and no significant conservation result in the cost savings group. We then add covariates to reduce standard errors. Specifications II to V present the estimates with covariates, which are robust to different configurations of fixed effects and controls. In Column V in Table S10, we include heating and cooling degree hours in addition to hourly fixed effects. As described above, degree hours capture both the magnitude and direction of heating and cooling loads on electricity consumption due to outside weather variation. We note that our use of degree-hour bins instead of hourly dummies leads to more conservative estimates of the treatment effects -8.2% treatment effect (Table S10, column V) versus -9.8% (Table S10, column IV). Here we confirm why usage of rigorous degree hours might be preferable to usage of time dummies alone.

We also carefully considered the impact of a large effective sample size for this case given a fixed N and large T dimension across households. The issue of autocorrelation, cross-sectional correlation and finding appropriate controls has been in the household energy consumption literature for some time (9). As robustness checks on our estimates, we considered both a range of sampling intervals and clustering options in order to distinguish statistically trivial from substantively important treatment effects. First, we compared results based on different frequencies. Second, we evaluated some of the pitfalls of panel data analysis identified in Bertrand, Duflo, and Mullainathan (7), particularly autocorrelation variance estimation. Third, we implemented multi-way clustering as described by Cameron, Gelbach and Miller (10) and Thompson (11) to account for dependence in both group and time dimensions.

Table S11. ATE Estimates at Various Sampling Frequencies, OLS

	I	II	III	IV	V	VI
	Monthly	Weekly	Daily	Hourly	Minute	30 sec
Post-Treat*Cost Savings Group	**5.669**	**4.111**	**3.914**	**2.962**	**2.961**	**2.972**
	(4.808)	(4.696)	(4.720)	(4.455)	(4.455)	(4.471)
Post-Treat*Health Group	**-8.673***	**-9.131****	**-9.474****	**-10.54****	**-10.54****	**-10.58****
	(4.409)	(4.376)	(4.429)	(4.177)	(4.176)	(4.191)
Treat Cost Savings	-0.314	0.4611	0.580	0.994	0.994	0.998
	(7.377)	(7.478)	(7.499)	(7.542)	(7.542)	(7.569)
Treat Health Group	-2.431	-2.186	-1.991	-1.523	-1.523	-1.529
	(7.85)	(7.859)	(7.879)	(7.864)	(7.864)	(7.892)
Apartment fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Degree-hour bins	Yes	Yes	Yes	Yes	Yes	Yes
Day by Week time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	855	3,320	21,437	490,994	26,718,555	53,437,110
R^2	0.003	0.023	0.023	0.048	0.024	0.024
F-statistic	3.176	4.319	11.30	12.93	12.93	12.93
Number of households	118	118	118	118	118	118

Robust standard errors clustered at the household level *** p<0.01, ** p<0.05, * p<0.1

In order to check for the potential effects of large sample size on our estimates, Table S11 shows OLS estimates at various sampling frequencies. To do this, we re-sampled our electricity time series at monthly, weekly, daily, hourly, minute, and 30-second intervals. As expected, our clustered standard errors decrease as the sampling frequency increases, and we show that our ATE estimates are robust even at the lower-frequency sampling rate. While the precision of our estimates is improved by our panel's time dimension, we do not rely on high T to demonstrate statistical significance. As such, we differentiate statistically trivial from substantively important effects, particularly for the health group in which the ATE estimates range from 8-11%. We report the most conservative ATE estimates in this study.

Comments on External Validity. Our sample population consists of Los Angeles Department of Water and Power (LADWP) customers who pay their electricity bills. They are a California multi-family renter population with typical housing characteristics and demographics (age, income, household composition, per capita electricity usage, etc.). Table S12 compares the per capita residential electricity consumption between University Village residents to LADWP utility customers, California population and general U.S. population. Our population has been described as one of five recognizable U.S. lifestyle consumers: young urban families –new baby, new car, smaller unit, newer appliances, fast food, frozen food, travel for commuting, shopping and visiting (12). Importantly, our participants are part of the information generation of consumers who regularly use Internet-based devices in their consumption habits.

Table S12. Per capita residential electricity consumption

Region	2010 Population (in thousands)	Annualized kWh	kWh per capita
United States*	308,746	3,749,985 x 10^6	12,146
California*	37,254	250,384 x 10^6	6,721
LADWP*	1400	8017.65 x 10^6	5,726
University Village	0.518	2910.782	5,619

Source: California Energy Commission data, 2010

Here we compare the housing characteristics of our multi-family renter community with broader populations. For example, 42.1% of housing units in Los Angeles County and 30.9% of housing units in California are in multi-unit housing structures, making the multi-unit housing communities meaningful to study (U.S. Census, 2014). More generally, there are 28.1 million multi-family housing units in the United States (Residential Energy Consumption Survey 2013, 2009 data) and 24.3 million of these housing units are renter occupied. According to data from the American Community Survey 2013, 52.7% of American housing units are renter-occupied. Among these renter-occupied households nationally, the average number of occupants was 2.84 persons, which falls very close to the average occupancy of 2.42 persons in our sample at University Village. We also note that 90% of all multi-family housing units in the United States are 1-, 2- and 3-bedroom units (Residential Energy Consumption Survey 2009), with the most common type being 2-bedrooms (there are 12.7 million 2-bedroom units in the U.S.). In our sample at University Village, all multi-family apartments are 1-,2- and 3-bedroom units, with 2-bedroom units being the most common type (N=101 households, 86% of all units in the study). In terms of square footage, the average size of multi-family homes in the U.S. (with 5 or more units) is 811 sq. ft. (Residential Energy Consumption Survey 2009). In our sample, the average sq. footage of multi-family homes at University Village is 835 (ranges from 595-1035 sq. ft.).

In terms of sample demographics, we also compared the age range of our sample participants to a broader population. For example, the median age in our sample of participants (heads of household) is 31 (ranges from 22 to 47); while the median age in California is 35.2 and in the U.S. is 37.2 (U.S. Census 2010). We note that persons aged 18 to 44, who are the most common age span of our sample participants, make up 38.7% of the entire population in California (14.4 MM people), and 36.5% of the U.S. population (112.8 MM people) based on Census data. In terms of their educational attainment status, our participants at University Village are more highly educated than the general U.S. population, having all received a bachelors degree or higher. We note however, that this is still a population of interest. Persons with a bachelor's degree or higher (age 25+) represent about 1/3 of the population: 29.5% of the population in Los Angeles county, 30.5% of the population in California and 31.7% of the population in the U.S. as a whole (U.S. Census 2010).

Our sample population is also a fast growing demographic in the. Between 2003 and 2013, there has been a 28% increase in the population of males seeking advanced degrees and 52.2% increase in females seeking advanced degrees. Thus, while educational attainment represents about 1/3 of the population in the U.S., our sample participants who are seeking advanced degrees, are also a growing demographic.

Our final demographic variable we consider is family income. Because income disclosure was voluntary, we had very few respondents (N=46, or 38% of population) who provided family income information. Among those participants who chose to disclose the information: the median annual household income for University Village participants is \$50,000 to \$74,000 (ranging from under 25,000 to 100,000 or more). By comparison, the median household income in the U.S. was \$51,017 in 2012 (US Census 2014), which places our sample participants in the middle range of income in the U.S. Because our self-reported income data is a biased sample due to nonresponse, we report the average household income for the two nearest Census block groups. The average income for University Village block group 1 is \$51,182 (U.S. Census 2010) and the average income for University Village block group 2 is \$61,467 (U.S. Census 2010), which also places our sample in the mid-range of earners in the U.S.

SI References

1. Allcott H (2011) Social norms and energy conservation. J. Pub. Econ. 95(9-10):1082-1095.
2. Gromet DM, Kunreuther H, Larrick RP (2013) Political ideology affects energy efficiency attitudes and choices. Proc. Natl. Acad. Sci. U.S.A. 110(23): 9314-9319.
3. Dietz T, Leshko C, McCright AM (2013) Politics shapes individual choices about energy efficiency. Proc. Natl. Acad. Sci. U.S.A. 110(23): 9191-9192.
4. Costa DL, Kahn ME (2013) Energy conservation “nudges” and environmentalist ideology: evidence from a randomized residential field experiment. *Journal of the European Economic Association* 11(3): 680-702.
5. Day AR, Karayiannis TG (1998) Degree-days: Comparison of Calculation Methods. *Building Services Engineering Research and Technology* 19(1): 7-13.
6. Delmas MA, Fischlein M, Asensio OI (2013) Information strategies and energy conservation behavior: a meta-analysis of experimental studies from 1975 to 2012. *Energy Policy* 61: 729-739.
7. Bertrand M, Duflo E, Mullainathan S (2004) How Much Should We Trust Differences-In-Differences Estimates? *Quarterly Journal of Economics* 119(1): 249-275.
8. Cameron AC, Trivedi PK (2005) Microeconometrics: Methods and Applications: (Cambridge University Press, Cambridge, UK) pp.82-83.
9. Dietz T, Vine EL (1982) Energy impacts of a municipal conservation policy. *Energy* 7(9):755-758.
10. Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multiway clustering. *Journal of Business and Economic Statistics* 29(2): 238-249.
11. Thompson SB (2011) Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99(1): 1-10.
12. Lutzenhiser L (1992) A cultural model of household energy consumption. *Energy* 17(1): 47-60.

Appendix 2 Proof of Identification

Following Manski (1996), we extend the proof of identification for classical experiments in the case of two independent, randomly assigned treatments with known treatment shares (e.g. the fraction of households receiving treatment) and high compliance to treatment. Let $\Re$ denote the randomization treatment rule which specifies treatments received by all households *j*=1,2,...N. For each household, let ξ_j denote the *received* information treatment by each household in the study population. The population outcome for each household in the study $y_j^{\Re}$ can be represented as,

$$y_j^{\Re} \equiv \sum_{t \in T} y_j(t) \cdot 1[\xi_j = t] \tag{2.1}$$

where 1 is an indicator function, equals 1 if $\xi_j = t$ and 0 otherwise. Under rule $\Re$, the outcome distribution conditional on observed covariates X is:

$$P(y_j^{\Re} \mid X) = \sum_{t \in T} P[y_j(t) \mid X, \xi_j = t] \cdot P(\xi_j = t \mid X), \qquad \forall t \in T \tag{2.2}$$

The first term on the RHS in Eq. 2.2 is the experimentally observed outcome distributions, conditioning on X and received information treatments ξ_j. The second term on the RHS is the conditional probability of receiving treatment. Elements of the set of feasible treatments $t \in T$ are independent of one another. Write

$$P[y_j^{\Re}(t)] = P[y_j^{\Re}(t) \mid X, \xi_j = t] \cdot P(\xi_j = t \mid X)$$

$$+ P[y_j^{\Re}(t) \mid X, \xi_j \neq t] \cdot P(\xi_j \neq t \mid X) \quad \text{(by law of Total Probability)} \tag{2.3}$$

But $P[y_j^{\Re}(t) \mid X,\ \xi_j = t] \equiv P[y_j^{\Re}(t) \mid X]$ by conditional independence, which maintains the statistical independence of the household's behavioral response $y_j(\cdot)$ function from treatment selection.

For feasible treatments, $T = \{t_0, t_1, t_2\}$, the outcome distribution becomes:

$$P[y_j^{\Re}(t) \mid X] \quad = P[y_j^{\Re}(t_0) \mid X] \cdot P(\xi_j = t_0) \text{ (control)}$$

$$+P[y_j^{\Re}(t_1) \mid X] \cdot P(\xi_j = t_1) \text{ (treatment 1)}$$

$$+P[y_j^{\Re}(t_2) \mid X] \cdot P(\xi_j = t_2) \text{ (treatment 2)} \tag{2.4}$$

By design, known fractions p_1, p_2 of the study population receive treatments t_1 and t_2 respectively. The ex-ante treatment shares are $P(\xi_j = t_1) \equiv p_1$ for treatment 1, $P(\xi_j = t_2) \equiv p_2$ for treatment 2, and its complement, $P(\xi_j = t_0) \equiv 1 - p_1 - p_2$ for the non-treated control group.

Since treatment shares are known, the probabilites $P(\xi_j = t)$ are identified for each $t \in T$. Household outcomes $P[y_j^{\Re}(t_i) \mid X]$ are also experimentally observed for each treatment group *i*. Therefore, each term on the RHS of Eq. 2.4 is point-identified for the study population.

It follows that the the identification region, *H* for all treated households is:

$$H\{P[y_j^{\Re}(t_{1,2}) \mid X]\} \quad \text{(monetary savings or health group)}$$

$$= [0,1] \cap \Big[\{P[y_j^{\Re}(t_1,t_2) \mid X, \xi_j = (t_1,t_2)] - (1 - p_1 - p_2)\} / p_1 + p_2,\ P[y_j^{\Re}(t_1,t_2) \mid X / (1 - p_1 - p_2)] \Big] \tag{2.5}$$

The identification region, *H* for non-treated control households is:

$$H\{P[y_j^{\Re}(t_0) \mid X]\} \text{ (control group)}$$

$$= [0,1] \cap \Big[\{P[y_j^{\Re}(t_0) \mid X, \xi_j = (t_0)] - (p_1 + p_2)\} / (1 - p_1 - p_2),\ P[y_j^{\Re}(t_0) \mid X / (p_1 + p_2)] \Big] \tag{2.6}$$

Appendix 3 Theoretical model

A Simple Mathematical Model of Energy Saving Bursts

Consider a continuous series of energy savings over time versus a reference consumption level. The series *S(t)* can represent a negative shock in consumption as percentage savings or kilowatt-hours per unit time. The change in relative consumption can be expressed in continuous time as the product of two exponential functions. The first exponential models the immediate burst in energy savings behavior in response to an information campaign, $-A\exp(-\tau_r / t)$, which we refer to as the novelty effect. The second exponential models the gradual decay in energy savings over time, $-A\exp(-t/\tau_d)$ with repeated information provision, which is analogous to exponential or hyperbolic discount models. A is the amplitude of the savings and τ_r, τ_d are non-zero temporal parameters that govern the rise and rate of decay analogous to discount functions. By convention, negative values of $S(t)$ means energy savings. We assume the rise in savings is fast relative to the rate of decay such as $\tau_r < \tau_d$. A functional form for energy saving bursts satisfies the following:

$$S(t) = -A\exp\left(-\frac{\tau_r}{t} - \frac{t}{\tau_d}\right) \qquad t > 0 \ , \tag{3.1}$$

where $S(t)$ is the energy savings achieved versus the baseline consumption or reference control group over time t. In Eq. 3.1, novelty effects initially dominate after information provision, and then persistence effects dominate after the peak conservation (or rebound) occurs. With no consumption rebounds, the effects of repeated information provision are assumed to gradually descend back to the reference level such as $\lim_{t\to\infty} S(t) = 0$. We do not assume consumers maximize consumption utility; instead we model psychology-based mechanisms of novelty and persistence using a value function that captures mental accounting over time. The objective function above is not the only feasible functional form available, but provide the best empirical fits to the earliest available field data with high resolution energy metering.

Peak Conservation. With complete or partial compliance to a given information strategy, peak energy saving occurs when $dS/dt = 0$. We refer to this as *peak conservation*.

$$dS/dt = \underbrace{-A\exp\left(\frac{-\tau_r}{t} - \frac{t}{\tau_d}\right)}_{\neq 0}\left(\frac{\tau_r}{t^2} - \frac{1}{\tau_d}\right) = 0$$

$$\Rightarrow t^* = \sqrt{\tau_r \cdot \tau_d} \qquad \text{(peak conservation time)} \qquad (3.2)$$

The maximum energy savings achieved $S(t^*)$ at peak conservation is:

$$S(t^*) = -A\exp\left(-\frac{\tau_r}{\sqrt{\tau_r \cdot \tau_d}} - \frac{\sqrt{\tau_r \cdot \tau_d}}{\tau_d}\right)$$
$$= -A\exp\left(-2\sqrt{\frac{\tau_r}{\tau_d}}\right).$$

Before we characterize the savings behavior in greater detail, let us parameterize the model to allow for model calibration and statistical fits. Consider peak conservation along the time axis as $\tau = \sqrt{\tau_r \cdot \tau_d}$ and the resulting peak shape as $\upsilon = \sqrt{\tau_r / \tau_d}$, which gives a relative measure of learning to decay rate of energy savings. By direct substitution, we can express Eq. (1) as:

$$S(t \mid A, \upsilon, \tau) = -A\exp\left[-\upsilon\left(\frac{t}{\tau} + \frac{\tau}{t}\right)\right], \qquad (3.3)$$

In this model of dynamic behavior, time is strictly positive $t > 0$ and our response function is continuously differentiable.[3] Under this equivalent representation in Eq. 3.3, we can fully characterize the consumption shock with just three intuitive parameters governing the rise, peak and decay of energy

---

3 Continuous differentiability is reasonable because we have real-time electricity metering.

savings bursts. First, the A parameter adjusts the amplitude of the savings. Second, the υ parameter defines the peak shape, where generally $\tau_r < \tau_d$ implies the onset of energy savings will be fast relative to the decay. Third, the τ parameter stretches the timescale, where higher τ values mean longer response times to reach peak energy savings. Figure 1 shows the effects of individual parameters on this theoretical model of energy saving bursts expressed by Eq. 3.3.

Feedback Delays. When feedback delays are short, we implicitly set the initial time to receive information, t_o, as zero. Examples of consumer feedback delays in energy information might include, for example, the time delay to receive home energy reports in the mail, or the time delay to receive e-mail alerts or view billing records online. Including an initial feedback delay at t_0, the expression for energy savings can be expressed more generally as,

$$S(t \mid A, \upsilon, \tau) = -A \exp\left[-\upsilon\left(\frac{t-t_0}{\tau} + \frac{\tau}{t-t_0}\right)\right] \quad \text{for } t > 0, \tau \neq 0 \tag{3.4}$$

A note on exponential response functions. Observations that rapid behavioral responses to information stimuli change exponentially over time was initially formulated in early behavioral theories of impulsive behavior and impulse control (Ainslie 1975). Psychologists have also used exponential fits to model impulsive behaviors in animal studies for instance in pigeons and other species (Chung 1965; Killeen 1968; Chung and Hernstein 1967) and clinical studies in humans (Mowrer and Ullman 1945; Ainslie 1975). These studies generally show highly concave response curves that either exponentially rise or decline over time. These parametric choice experiments have generally been unidirectional (either exponentially increasing or decreasing) and have generally focused on social learning or time discounting. We describe a new application in mental modeling with two distinct temporally separable mechanisms. The following figures illustrate the temporal fits and numerical simulation obtained using the model of Eq. 3.4.

Numerical simulations

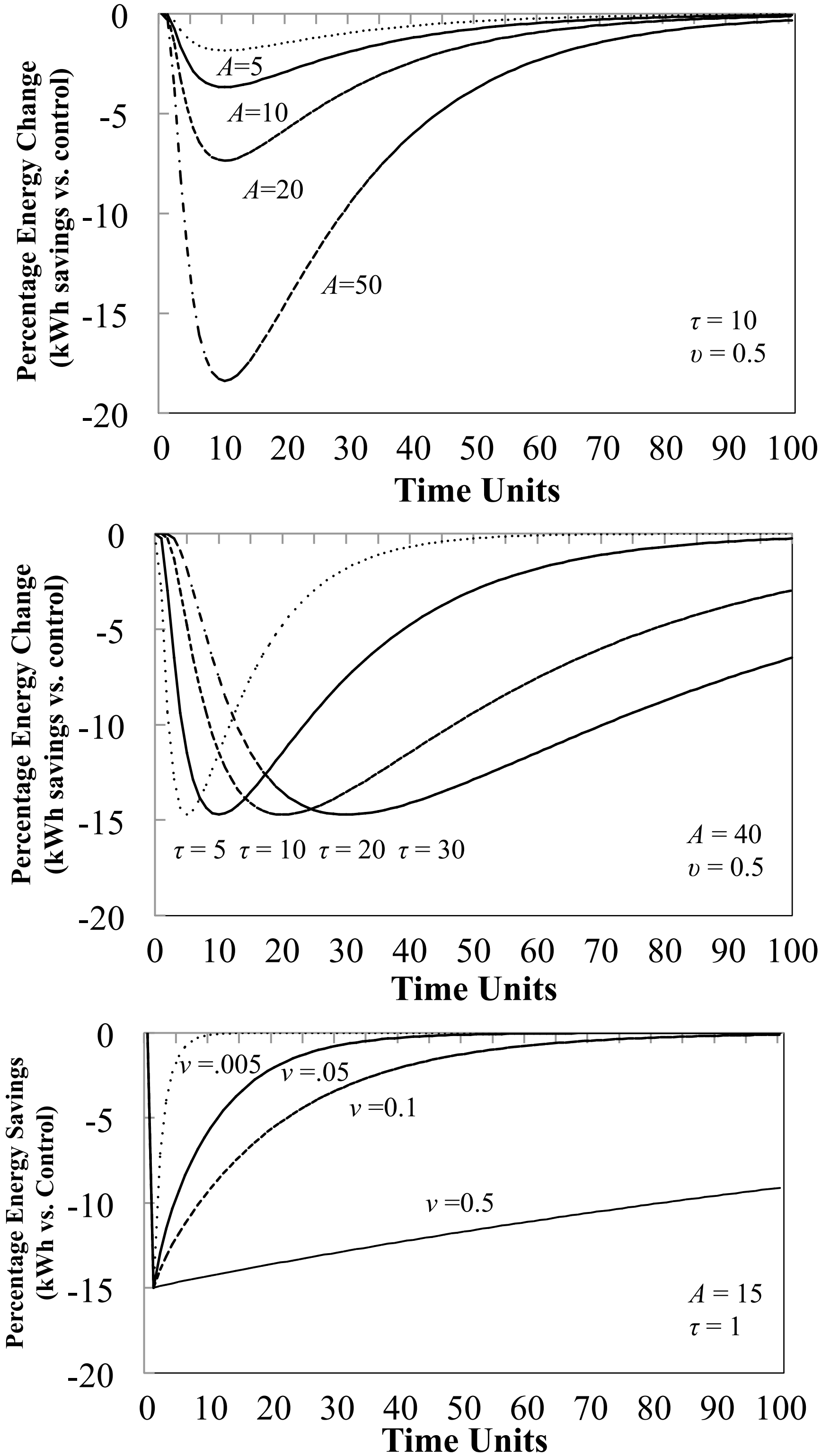


Fig. S4. *Effects of parameters on hypothetical model of energy saving bursts*

SI References

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* 82(4):463-496.

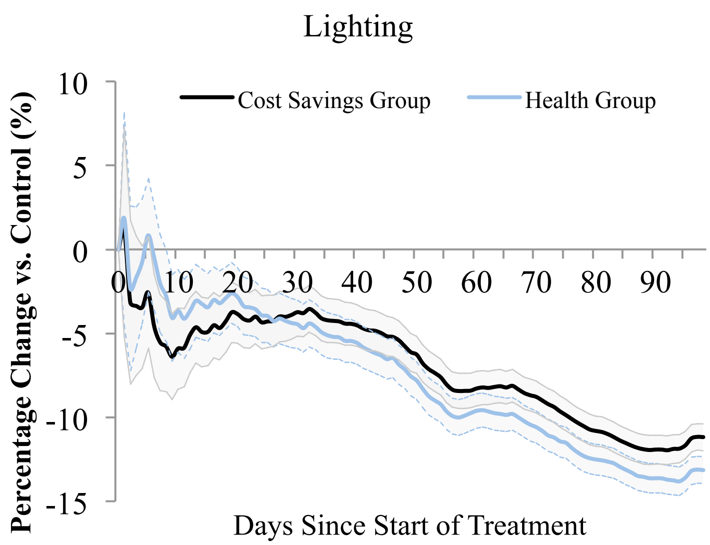
Chung SH (1965) Effects of delayed reinforcement in a concurrent situation. *Journal of Experimental Analysis of Behavior* 8(6):439-444.

Chung SH, Hernstein RJ (1967) Choice and delay reinforcement. *Journal of Experimental Analysis of Behavior* 10(1):67-74.

Manski CF (1996) Learning about treatment effects from experiments with random assignment of treatments. *Journal of Human Resources* 31(4):709-733.

Mowrer OH, Ullman AD (1945) Time as a determinant in integrative learning. *Psychological Review* 52:61-90.

Appendix 4 Supplement on appliance dynamics



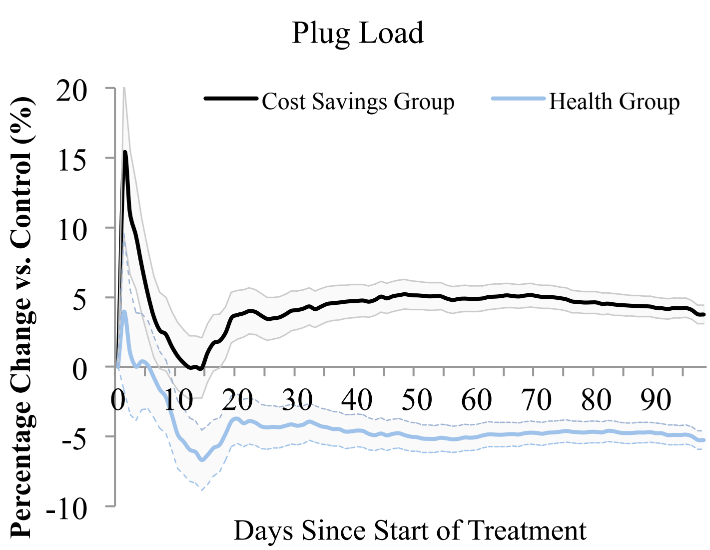
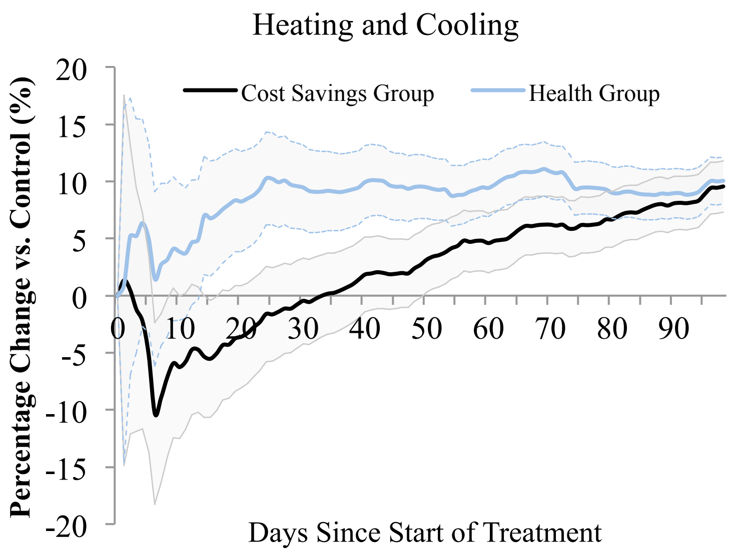


Fig. S5. *Dynamic treatment effects at the appliance level (Continued on next page)*



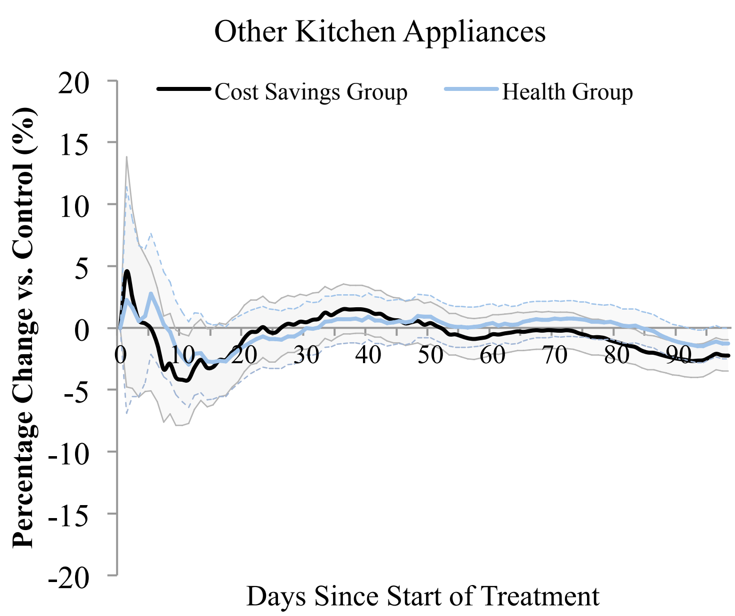


Fig. S5. *(Continued) Dynamic treatment effects at the appliance level.*

Appendix 5 Supplement on Conservation Messaging

This chapter supplement summarizes results of a series of flash survey questionnaires administered at UCLA from October 2011 to January 2012. The purpose of these "Flash Surveys" was to evaluate and select treatment messages for the 2012 University Village field experiments in residential electricity use. The challenge posed was how to design messages to motivate conservation at the individual level. The approach outlined here emphasizes personal relevance for unobserved external impacts.

This supplement analyzes the effectiveness of behavioral messaging in promoting stated intent, and not actual behavior, which is covered in the main text in Chapter II. I will describe the questionnaire, the audience and results. I use ordered choice models to identify willingness-to-save estimates and thresholds. The approach is easily scalable to larger study populations or consumer segments.

The Messages. In this section, I describe several candidate messages, designed to connect the individual with one or more dimensions of externalities (social costs) associated with individual electricity use. The communication strategy begins with the recognition that household electricity is largely 'invisible' to the consumer and disconnected from the impacts of marginal electricity production, particularly emissions and health outcomes as a result of repeated air exposure to pollutants in the environment. We appeal to six distinct treatment messages, which are randomized and presented to respondents as self-administered prompts. One message describes the most obvious incentive to conserve energy—saving money. The remaining five messages make up other classes of nonmonetary incentives, which can appeal to higher order associations and cognition. Each message contains a peer comparison (how much energy is consumed relative to a similar neighbor or some reference level), and an impact described (e.g. what additional impacts their individual consumption may be causing).

Peer Comparisons. In all cases, treatment messages are standardized with peer comparisons of the following form: *"Last month, you used 66% more electricity than your efficient neighbor…"* This type of language is commonly referred to as comparative or normative feedback, and has gained broad use in (1) small-scale conservation lab or field studies, typically psychology and decision sciences and (2) utility-scale conservation pilot projects, typically economics and related fields. Whereas the peer feedback component between messages is standardized, the second half of each message features a distinct end use impact or equivalence.

Impacts Described. Each impact described in the second half of each conservation message is framed in such a way to maintain independence of tested concepts. For example, if conserving energy helps reduce emissions, only the emissions component of the impact is described in turn, regardless of whether collective conservation actions might also along other dimensions save money or reduce cases of childhood asthma. More importantly, each impact described is presented in *numerical and scientifically verifiable terms.* We use publicly available databases for calculating effect sizes and emissions equivalencies for all hypothetical electricity reductions. In particular, we derive our numbers using emission factors from the Emissions & Generation Resource Integrated Database (eGRID) maintained by the U.S. EPA.[4]

It is important to note that household conservation actions are not generally assumed to affect baseload emissions (that is, emissions from continuously running power plants), but rather, non-baseload emissions (power plants that are brought online as necessary to meet excess demand). From a marketing perspective, these reductions in kWh must then be converted into familiar, contextually relevant and accessible consumer language. For instance, we can appeal to the number of additional smoke stacks fired, additional cars on the road, or the equivalent number of trees planted. The job of effective communication strategies therefore, is to translate this verifiable scientific data into salient dimensions that can be easily understood, have a high degree of credibility, and that can be individually relevant for inducing appropriate behavioral responses (e.g. conservation actions) broadly within the consumer base. Note these simple principles apply to both economic incentives like cost savings (are 100 pennies worth

4 eGRID is a national database of the environmental characteristics of almost all electric power generated in the United States. Environmental characteristics include air emissions for common greenhouse gases, NOx, SOx, CO2, CH4; emissions rates; net generation; resource mix; and other attributes. eGRID was accessed on 2/1/2012 here: http://www.epa.gov/cleanenergy/energy-resources/egrid

the trouble?) and non-economic incentives (do I care about saving trees?). Both types of messages are tested in this supplement.

One important challenge in communicating to consumers is that effect sizes can be small at the margin (0.5% of a car for example) and so, often times, the expected costs and benefits must be aggregated over large distances, populations or time periods for salience at the consumer level. In this survey questionnaire, we do not vary the effect sizes within a message, as we are only interested in comparing conservation preferences for these emissions equivalencies. We fix the relative consumption to represent realistic scenarios and numbers.

The Survey. Table S13 lists the specific conservation messages tested. The flash surveys were administered online using Qualtrics survey software. Self-administered web-based data collection is particularly fitting because e-mail and website communications are the main communication channels in the University Village study. Three iterations of the questionnaire were conducted online over a span of four months.[5] Questionnaires were sent out anonymously via private link invitations distributed throughout the UCLA community. Survey invitations were not sent to residents of the UCLA University Village apartments to prevent contamination in the experimental study. Respondents were eligible to participate in the survey questionnaire only if they pay electricity bills at home. No other screening questions were used, although this could be easily be modified in a larger survey with broader consumer targeting or segmentation. Closed questions were administered on 7-point Likert scales and were interpreted as ordinal data to allow for rankings. The results summarized here correspond to the third iteration of the questionnaire—Flash3 following various rounds of pre-testing.

No cash incentives or payouts were offered for participation in the survey questionnaire.

---

5 Flash1 pre-test was in the field October 2011, Flash2 in November 2011 and Flash3 in January 2012. Each iteration incorporated minor messaging changes and additions in response to feedback from respondents. Each survey maintained independent sampling populations.

Table S13. Treatment messages, flash survey

Impact	Message
Financial savings	"Last month you used 66% more electricity than your efficient neighbors. In one year, this will cost you $34 dollars extra."
Health	"Last month you used 66% more electricity than your efficient neighbors. You are adding 610 pounds of air pollutants which contribute to health impacts such as childhood asthma and cancer."
Trees	"Last month you used 66% more electricity than your efficient neighbors. Over a year, your extra emissions are equivalent to removing 7 trees in your community."
Cars	"Last month you used 66% more electricity than your efficient neighbors. Over a year, this is equivalent to adding 1 car to the road."
Emissions	"Last month you used 66% more electricity than your efficient neighbors. Over a year, this is an additional 609 pounds of CO_2 emissions from a coal-fired power plant."
Developing Country	"Last month you used 66% more electricity than your efficient neighbors. Over a year, the extra energy would be enough to provide power to 3 Kenyan citizens."

Table S14. Respondents' stated preference for willingness to save (strongly disagree=1, strongly agree=7)

Dimension	I would save energy. (% respondents who Agree or Strongly Agree)	(% respondents who Disagree or Strongly Disagree)	Mean	S.D.	Min	Max
Health	60.0	6.4	5.454	1.542	1	7
Financial savings	48.2	8.2	5.318	1.602	1	7
Trees	46.4	7.3	5.127	1.527	1	7
Emissions	42.7	8.2	5.054	1.5621	1	7
Cars	40.2	10.7	4.795	1.709	1	7
Developing Country	38.7	10.8	4.802	1.683	1	7

N=110

Table S14 shows descriptive statistics, the percentages, means, standard deviations and min/max for all tested messages. Responses were coded by level of agreement from negative to positive (Strongly Disagree=1, Strongly Agree=7). The 'Health' and 'Financial savings' strategies are our top two ranking messages by stated preference. Figure S6 lists the results by top 2 boxes, those percentages of respondents who agree or strongly agree that they would save energy after reading the randomly assigned message. It turns out that 'Cars on the Road' and 'Developing Country' were the lowest scoring messaging strategies with approximately 11% of respondents rating these messages in the bottom 2 levels of agreement in willingness-to-save.

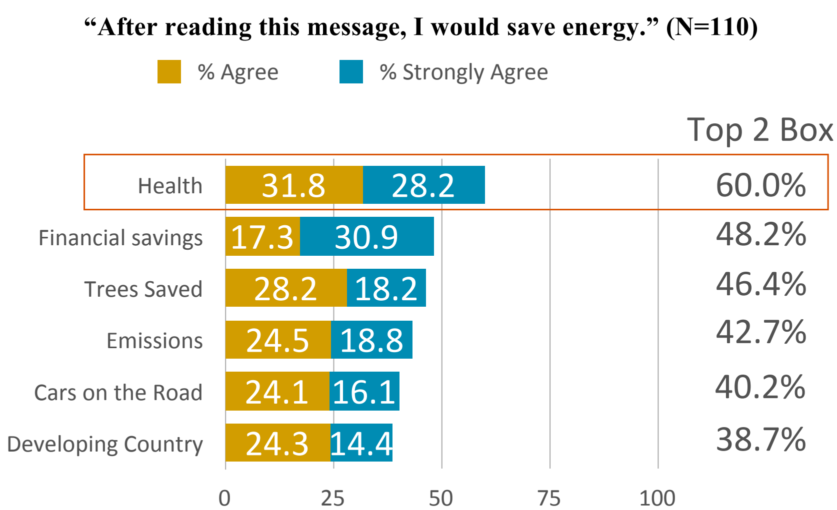


Fig. S6. *Health and Financial savings are the top ranking messages*

In the next sections, we present the results of our stated preference survey scored along three dimensions, namely, *Comprehension, Believability, and Relevance.*

Comprehension. If study subjects do not easily understand or are confused by our treatment messages in any way, maybe because of poor wording, poorly defined terms or unduly complicated language—they will not easily be motivated to change behavior. This might be a considerable challenge, given the fact that the language surrounding energy and conservation can be highly technical. For example, what is the difference between saving power in watts and saving energy in kilowatt-hours? For ordinary consumers who do not work in the electric power industry or may not be as familiar with the technical language, some additional education or specialized communication may be necessary to get an intuitive understanding. What does saving an additional 10 kilowatt-hours per month mean anyway? The cognitive aspects of conservation messaging can be considerable. I argue that comprehension at the consumer level is a barrier that can crowd out statistical identification of treatment effects in experimental studies. This barrier is not widely discussed in the literature, although the issue is sometimes bundled with problems of consumer inattention or imperfect information.

We tested each of our six conservation messages for stated *Comprehension*. Respondents were simply asked to rate their level of agreement (1=strongly disagree to 7=strongly agree) with the following statement: *"I understand this message."* Figure S7-A contains the tabulated results. Respondents who rated in the Top 2 boxes (Agree=6 or Strongly Agree=7) were in the range of 78-93% and all tested messages show a high level of stated comprehension, although the standalone emissions and developing country messages could be improved. Such a favorable response rate to self-administered prompts is surprisingly high for conservation related messaging, and is at a level consistent with rule-of-thumb in marketing practice. We expect respondents who state a high level of comprehension to be more likely to state a willingness to save energy as a necessary but perhaps not sufficient condition.

Believability. The next dimension we tested is *Believability*. Here we test the perceived credibility of the stated impacts. This turns out to be an incredibly important dimension not just for conservation related messaging, but also for other types of green marketing studies generally. 'Believability' or credibility might also be an important communication challenge to consider for other reasons. Conservation messaging is a special type of communication in the sense that behavioral responses require collective action. How credible are the external impacts from individual use? Can any one individual actually make a difference? This is a complex behavioral challenge because even if consumers understand the importance of conservation, if the stated impacts are not fully credible—in other words, if I don't believe that through my actions I can *actually* save 100 pennies or 100 pounds of air pollution—I may not be motivated to act. We expect 'Believability' to be positively correlated with respondents' stated willingness to save energy. How should we test for this in our treatment messages? After each conservation message, we asked for ratings on the level of agreement with the following statement: *"The impacts described are realistic."* [6] Figure S7-B shows the results. The cost savings message had the highest level of agreement on believability (Top 2 boxes were 78.2%), although not all respondents felt the energy savings were fully credible. By contrast, the developing country message had the lowest level of agreement on believability. Only 53% of respondents agreed or strongly agreed that the impacts described were realistic.

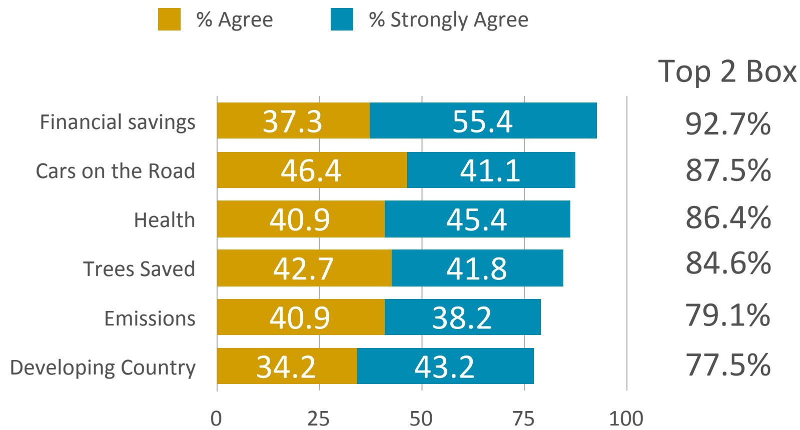
In earlier iterations of the questionnaire, the Saving Trees message was ranked lowest in this category: *"…Over a year, your extra emissions are equivalent to removing 7 trees."* We added the qualifying words "in your community" at the end of the impact message, so that the final communication stated: *"…Over a year, your extra emissions are equivalent to removing 7 trees <u>from your community</u>."* This subtle word choice had the effect of increasing the importance of the trees message and the wording change boosted its relative ranking versus the other messages. One important insight is that this subtle word change effectively made the perceived social costs more personal, as we invariably do with the health messaging approach. In our statistical analysis, the word change also increased the likelihood of our respondents' stated willingness to save energy. **Our *"7 Trees in your community"* example highlights the fact that subtle word choice and framing can have an enormous influence on stated conservation preferences and the perceived value of conservation actions by individuals.** To motivate conservation behavior, it is not simply enough to give consumers feedback—by conserving energy one can save 100 pennies—it's also very important *how* those 100 pennies are described.

---

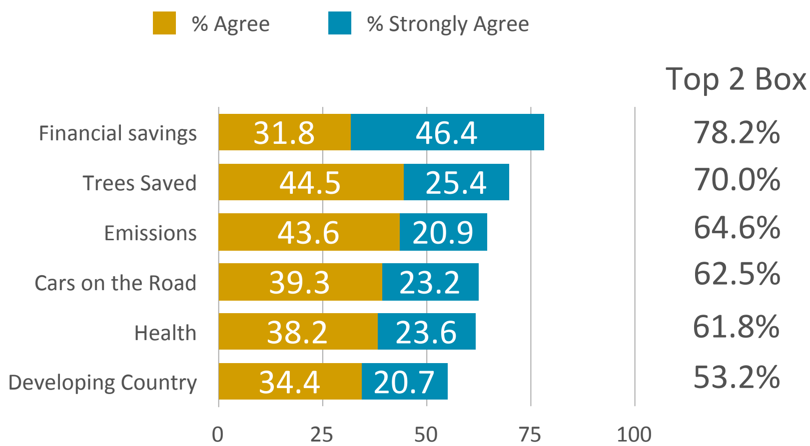
6 We also added the following qualifying language on each screen: "Realistic means the impacts described are believable."

Ranking Conservation Messages.

S7-A. "I understand this message."



S7-B. "The impacts described are realistic."



S7-C. "The message is relevant to me."

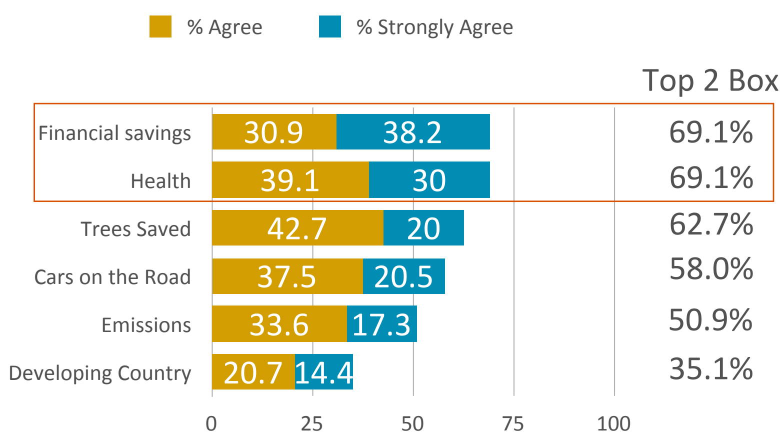


Fig. S7. *Rank ordering messages on comprehension, believability, and relevance (N=110)*

Relevance. Our third and final dimension we consider is Relevance. By relevance, we mean the relative importance the stated impacts have on an individual's disposition toward the message. Relevance in this context is very closely related to the marketing concept known as "salience" which often stands for perceived differences in consumer attitudes. If measured correctly, relevance or salience can be quite an effective measure for what psychologists call predictive validity. Respondents were asked to rate their level of agreement (1=strongly disagree to 7=strongly agree) with the following statement: *"This message is relevant to me."* Figure S7-C contains the tabulated results. Our Financial savings and Health message were our top two ranking messages, (Top 2 boxes, 69.1%). We predict high scores in Relevance ratings to be positively correlated with respondents' stated willingness to save.

For simplicity in communication, we classify our 3 measures as follows: "I understand this message" means Comprehension; "The impacts described are realistic," means Believability; and "The message is relevant to me," means Relevance. Table III shows results of these 3 dimensions on our respondents' stated willingness to save energy.

Survey results. We run ordered logit models for each message type, and we report coefficients in log odds and the odds ratios. From Table S15, we see that all significant odds ratios are greater than 1, which implies higher levels of agreement with our tested dimensions are positively correlated with a stated willingness to save. For the cost savings message for example, a 1-unit Likert scale increase in Believability ratings implies respondents were 2.06 times as likely to state a willingness to save energy, holding all else equal. Similarly, for a 1-unit increase in Relevance ratings, respondents were 1.51 times as likely to state a willingness to save energy. For the Health message, estimates from odds ratios are 1.98 and 2.08; so, high Believability and Relevance ratings on Health implies that survey respondents are twice as likely to state a willingness to save energy (See Table S15 for the full results). Even at the current sample size (N=110), simply 'understanding' the message, especially for the standalone emissions message, is not a sufficient promoter of Willingness-to-save.

In summary, these results show that survey respondents are 1.5 to 2.8 times as likely to state a willingness-to-save energy when the treatment message is "Relevant to me" and/or when "Impacts described are believable." Our discrete choice data have allowed us to extract revealed preference orderings for subsequent predictions about which message(s) might be most effective in an experimental setting.

Table S15. Ordered logit regressions of respondents' willingness to save energy. "After reading this message, I would save energy."

Dimension	(1) Cost Savings		(2) Health		(3) Trees		(4) Emissions		(5) Cars		(6) Developing country	
	Coef.	Odds ratio	Coef.	Odds ratio	Coef.	Odds ratio	Coef.	Odds ratio	Coef.	Odds ratio	Coef.	Odds ratio
I understand this message.	.0761 (.2332)	1.079 (.2516)	-.0166 (.2347)	.9835 (.2308)	.2190 (.2275)	1.244 (.2832)	.1464 (.1964)	1.157 (.2274)	.2537 (.2112)	1.288 (.2722)	.0305 (.1829)	1.031 (.1886)
The impacts described are realistic.	.7242** (.2353)	2.063** (.4853)	.6867*** (.1802)	1.9872*** (.3581)	.5040* (.2215)	1.6554* (.3666)	.5676* (.2328)	1.764* (.4107)	.5444** (.1977)	1.723** (.3408)	.6421*** (.1791)	1.900*** (.3403)
This message is relevant to me.	.4133* (.1666)	1.511* (.2519)	.7357** (.2552)	2.0871** (.5326)	1.0542*** (.2184)	2.869*** (.6266)	.7888*** (.1498)	2.200*** (.3296)	.5983*** (.1502)	1.819*** (.2731)	.7460*** (.1624)	2.108*** (.3424)
χ^2 (d.f.=3)	54.16		57.15		89.08		70.18		63.76		85.06	
Log likelihood	-155.41		-150.96		-141.43		-154.32		-170.96		-154.35	
Pseduo R^2	0.1484		0.1592		0.2395		0.1853		0.1572		0.2160	
N	110		110		110		110		110		110	

Standard errors in parentheses. Significant to * p<.05, ** p <.01, *** p<.001

In summary, we conducted a stated preference survey to identify willingness to save estimates prior to the experiment. We identified three important dimensions—*Comprehension, Believability and Relevance*—and used these to reveal a rank ordering of preferences for conservation messages in our target population. While all candidate messages scored high in stated Comprehension, survey respondents were 1.5 to 2.8 times more likely to state a willingness-to-save energy when the treatment message is "Relevant to me" and when "Impacts described are believable." These ratings were strongest for health and cost savings. Simply 'understanding' the message or the impacts described, especially for the standalone emissions messaging, is not a sufficient motivator of willingness-to-save. Our analysis suggests that cost savings and health are the two best messaging alternatives. These two approaches were therefore carried into the experimental study.

III

EVALUATING PERFORMANCE IN SOCIETAL GRAND CHALLENGES: THE UNITED STATES DEPARTMENT OF ENERGY BETTER BUILDINGS CHALLENGE

The diffusion of energy efficient technologies is increasingly important for firm competitiveness and productivity. Yet our ability to estimate causal relationships between targeted policies for technology adoption and performance outcomes has been limited. In this chapter, we use quasi-experimental techniques to investigate the performance of information-based programs unified under the U.S. Department of Energy's Better Buildings Challenge. Using a dataset of 178,777 commercial buildings performance data resulting from a public-private partnership in the City of Los Angeles, we analyze energy savings and economic returns to technological investments. We use evolutionary search algorithms to match commercial buildings on observable characteristics that define credible counterfactual consumption scenarios, which serve as the basis for performance comparisons. Our results suggest energy savings in the range of 18-25% across 35 million square feet of commercial real estate space for the buildings that implemented the energy audit program. We observe an energy efficiency project implementation rate of 42% for projects identified through government sponsored energy audits. We discuss program level barriers and drivers of success.

Keywords: grand challenges, innovation, energy efficiency, matching algorithms

3.1 Introduction

Grand challenges have emerged in recent years as a way to inspire and organize research and innovation toward major societal problems. Grand challenges serve to capture the public imagination on major unresolved problems facing society such as global climate change, pandemics and communicable disease (Omenn, 2006). These mission-led approaches in research and innovation policy can span a wide array of fields from engineering and computer science, to medicine, global health and the environment. How effective are grand challenges in stimulating new technical or behavioral solutions? Assessing the degree to which grand challenges will have a measurable impact on the trajectory of technological progress will require deeper thinking about counterfactual scenarios and outcomes.

The National Research Council defines grand challenges as “major scientific thrusts…that offer potential for major breakthroughs on the basis of recent advances, and are feasible with current capabilities, given a serious infusion of resources” (NRC, 2001). For many societal grand challenges, establishing causal links between specific challenge-based interventions and changes in performance or technological trajectories will be nontrivial for impact evaluation. This is because the costs and benefits of technological investments will accrue to many actors, which will require access to large amounts of disparate data on costs and benefits in new areas where public statistics have been scant. Technological progress on grand challenges will also require both public and private spending, especially for implementing new innovations. This stands in contrast with prior mission-oriented programs such as the Apollo mission or the Manhattan project, which were primarily publicly funded and in which government was the primary customer (Mowery, Nelson, & Martin, 2010). This chapter provides a basis for understanding performance in the context of the U.S. Department of Energy (DOE) *Better Buildings Challenge*, a national policy directive aimed at modernizing existing buildings and infrastructure.

We evaluate demand side innovation in response to information-based programs in the commercial buildings sector. Using building-level micro data from the City of Los Angeles Better Buildings Challenge (LABBC), a public-private partnership aimed at driving energy efficiency in buildings, we estimate energy savings and economic returns to technological investments. However, participating buildings in the program are not randomly assigned and we therefore use matching techniques to evaluate impacts versus a reference set of buildings with comparable characteristics. We use evolutionary search algorithms to match buildings on observable characteristics, which serve as the basis for performance comparisons. Our results suggest energy savings in the range of 18-25% across 35 million square feet of commercial real estate. We observe a project implementation rate of 42% for projects identified through government sponsored audits. We discuss program level drivers and barriers of success.

3.2 Grand challenges: the demand side

Several interrelated grand challenges have been announced on restoring and improving urban infrastructure, particularly approaches that consider the long-run sustainability of energy, water and transportation systems that fundamentally support communities (NRC 2001, NAE 2008). On resource use in the existing buildings, the U.S. Department of Energy announced the *Better Buildings Challenge* in 2009 with an aim to enhance innovation and drive greater energy efficiency in the economy. The program supports U.S. commercial and industrial building owners to voluntarily commit to reducing energy consumption 20 percent or more over 10 years. The innovation challenge is to increase the rate, quality and effectiveness of building renovation to reduce energy intensity with large expected social benefits in reducing energy use externalities.

The Better Buildings Challenge can be understood more generally in the context of demand side policies, which have grown in popularity in recent years in both the U.S. and E.U. (Edler & Georghiou, 2007; Foray, Mowery, & Nelson, 2012). The idea is that when private and public actors express a concentrated need, it can thus trigger innovation in response to market demand (Mowery, & Rosenberg, 1979; von Hippel, 1986; Edquist, & Hommen, 1999). Demand side policies address both the need for market implementation of new innovations (e.g. responsive demand) and also signal new functional needs to producers and providers of technology (e.g. triggering demand).

In stimulating demand for energy efficient technologies, there is an extensive literature describing under-investment in the economy due to structural, informational and other market barriers (Jaffe & Stavins, 1994a, 1994b; Gillingham, Newell & Palmer, 2009; Dietz, 2010). This so-called energy efficiency gap has stirred empirical debates over the existence and magnitude of this investment gap along with theoretical discussions about the role of incentives, information and other market-based technology policies (Allcott & Greenstone, 2012; Gillingham & Palmer, 2014). Informing agents of profitable investment opportunities and offering technical assistance with their implementation alleviates asymmetries of information, thereby increasing opportunities for conservation and reducing pollution externalities (Newell, Jaffe & Stavins, 2006). Search costs and imperfect information can impede sellers from effectively conveying the benefits to potential buyers of technologies (Howarth & Sanstad, 1995). The economic rationale for reducing these information asymmetries lies in the public goods aspects of information programs. Many energy efficiency programs work by providing subsidies to households and firms to adopt new technologies. The U.S. Department of Energy estimates that $34 billion dollars were spent by electric utilities on subsidies between 1995-2013, with spending increases every year from 2004 to 2012 (DOE, 2013). However, designing incentives to spur investment and making causal inferences when these investments do occur are two very different tasks. Despite the importance of identifying linkages between incentives and behavior in meeting policy goals, there is surprisingly little direct evidence on private returns to energy efficiency investments (Allcott & Greenstone, 2013). Here we evaluate the effects of voluntary and information-based programs targeting social returns from energy efficiency investments. We investigate measures of cost effectiveness in strategic investment decisions by firms.

3.3 Impact Evaluation

What would technological progress be like in the absence of mission-oriented policies? The difficulty arises in the fact that participating firms in grand challenges rarely constitute random samples.

We also never directly observe the alternative states among participants in which an investment or participation decision is not made. These issues, non-random selection and how to construct a valid control group for counterfactual analysis have been an ongoing challenge in the evaluation literature (Klette, Moen, & Griliches 2000). Often, the more aggregate the level of analysis (e.g. policies, countries, regions), the harder it is to identify a proper control group. On the other hand, the more disaggregated the level of analysis (i.e. individual technologies or projects) the harder the data is to obtain. Other important challenges include the ability to establish an appropriate baseline, the inadequacy of public statistics to capture unobserved group and time varying characteristics, and blurred boundaries between desired implementation outcomes and measured impacts (Edler, Georghiou, Blind, & Uyarra, 2012). Further, cross-sectional analyses of programs can be misleading because changes in outcomes over time can be endogenous with technology adoption, pricing and consumer preferences, all of which limit our ability to make causal inferences. Prior evaluation studies have largely been based on non-experimental, observational data. These are primarily case studies of a select group of firms or industries; or interviews with program managers and field experts (Jaffe, 2002). This kind of expert testimony from professional evaluators often leads to upward biases in evaluation efforts because of the potential for favorable evaluations to lead to continued programmatic funding (Klette et al., 2000).

Recent advances on the use of econometric techniques with firm-level data have been developed to mitigate some of these issues (Imbens & Wooldridge, 2009; Angrist & Pischke, 2010). Scholars have argued that the most rigorous approaches in the evaluation of energy efficiency programs for example is to use randomized controlled trials (Allcott, 2011; Allcott & Rogers 2014; Ayres, Raseman, & Shih, 2013) and natural or quasi-experimental techniques (Boomhower & Davis, 2014; Ito, 2015). For example, Asensio and Delmas (2015) designed experimental framing interventions to motivate significant energy savings behavior in a California residential field site using nonmonetary health information disclosures. When randomization in treatment assignment is not possible, quasi-experimental techniques such as matching strategies and regression discontinuity (RD) designs can overcome some criticisms on the ability of prevailing methods to effectively control for confounding variables in regressions based on non-experimental, observational data (Imbens & Woolridge, 2009; Heckman, Ichimura, & Todd, 1998). For a discussion of strengths and limitations of experimental research in policy analysis, see surveys by Imai, King and Stuart (2008); Barrett and Carter (2010). Recently, scholars have called for better research designs that can allow for the study on the effectiveness of incentives based on ex-post evaluation of data (Allcott & Greenstone, 2012; Mathew, Dunn, Sohn, Mercado, Custudio, & Walter, 2015). In the next section we use micro level data and several matching algorithms to evaluate the performance of technological investments by private building owners and managers. We evaluate Better Buildings Challenge improvements in energy intensity (kWh per square foot) and building performance upgrades in commercial buildings. These impacts can also be translated into other performance measures such as economic returns, emissions reductions, or avoided pollution or health costs. We also benchmark performance outcomes versus complementary voluntary green building certification programs, LEED and Energy Star to evaluate strategic choices between competing approaches for efficiency goals.

3.4 Los Angeles Better Buildings Challenge

The City of Los Angeles was among the first and largest metropolitan cities selected to participate in the Better Buildings Challenge program. The LABBC is managed through a utility funded public-private partnership and is believed to be a model for other cities in stimulating regional innovation. The LABBC offers building owners and managers a number of services including: technical assistance, subsidized energy audits, benchmarking, project development support, energy measurements and verification, and hands-on assistance on financing, rebates and other available incentive programs. Beginning in 2011, eight independent auditing firms conducted subsidized ASHRAE Level II energy audits in participating buildings totaling 35 million square feet of commercial real estate across the City. A central concern with third party auditing is that independent auditors may invariably face a conflict of interest between providing favorable reports and maintaining business with client firms (Duflo, Greenstone, Pande, &

Ryan, 2012). Indeed, when firms can directly hire their auditing firm, they can bid for a favorable report and there is some prior evidence that this can undermine information provision. LABBC differs from other energy efficiency programs in that the entire program, including assignment of auditing firms to participating buildings owners was managed through external agents. The public private partnership acts as a conduit between technology providers, and the project engineering and project finance communities; with a primary objective to grow the pool of motivated building owners and managers, by reducing informational barriers and search costs. This includes standardization of all audit reports from engineering-based technical recommendations into financial investment language for proposed energy savings measures and investment risk profiles.

3.5 Data Sources

Our data include monthly electric utility data for all commercial accounts in the Los Angeles Department of Water and Power (LADWP) service territory from 2005-2012. This corresponds to a universe of 178,777 commercial buildings (16.5 million building-month panel observations) including pre- and post- challenge program participation. We also acquired detailed building stock characteristics from Costar, the premier commercial real estate database, which provides information such as physical and location characteristics, occupancy rates, building certifications, and other measures of building quality. We also obtained weather station data from NOAA, which allow us to calculate heating and cooling degree-days to the nearest weather station. See Delmas, Fischlein, & Asensio (2013) for a discussion of the importance of weather and seasonality controls in modeling electricity consumption data. Finally, we have detailed project data for 91 LABBC participating commercial buildings that received subsidized energy audits as part of the LABBC program. These total approximately 35 million square feet.[7] We also benchmark energy performance of Energy Star and LEED certified buildings, which

[7] For a detailed listing of participating buildings see http://la-bbc.com/projects/ (Accessed April 8, 2015)

have also invested in energy efficiency throughout this time. We conduct the analysis at the building-level.

3.6 Results & Discussion

Figure 1 presents maps of the sampling distributions of participating buildings across the City for the LABBC audits, LEED and Energy Star certified buildings. We observe program participation in each grid across all city council districts. The largest density of participating buildings by square footage is concentrated in areas of major commercial activity. This corresponds to the east city center for the largest concentration of LABBC audits and West Los Angeles for the largest concentration of LEED and Energy Star buildings. An important observation from Figure 1 is that given the underlying spatial heterogeneity, a simple linear or radial distance from city center would be a relatively ineffective matching strategy. Buildings close together are not necessarily similar in observable characteristics or in energy intensity. Also, given the diverse physical and occupant characteristics, it is necessary to test for imbalance in pre-treatment covariates in order to reduce any model dependence and potential for bias in the statistical estimation of treatment effects.

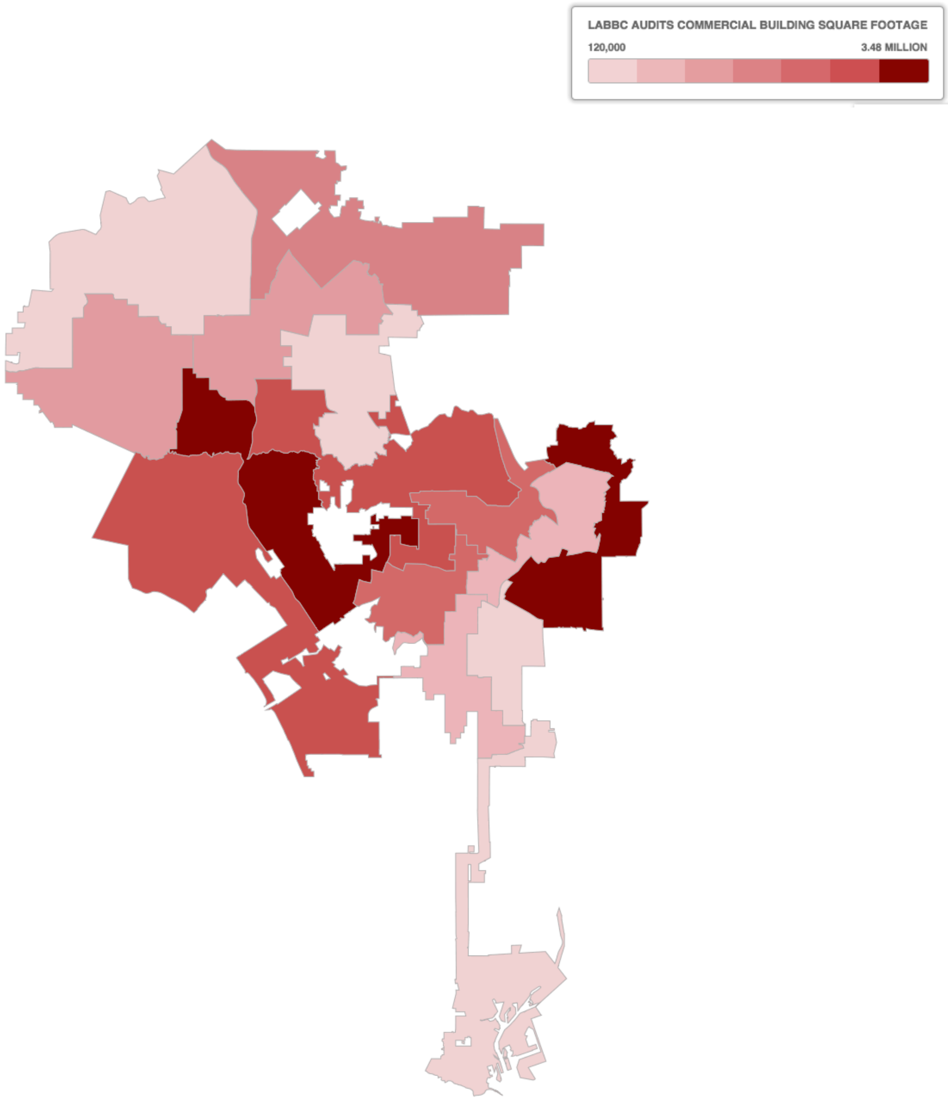
Table 1. Comparison of participating and non-participating buildings

LABBC Audit Indicator	kWh per month	kWh per Sq. Ft.	Building Class (1=A, 2=B, 3=C, 4=Z)	Sq. Ft.	Taxes per Sq. Ft.	Year Built	Renovated 1= Yes	Climate Zone (1=cool zone, 2= hot zone)
0	5,601	0.906	2.742	23,307	3.12	1959.0	0.00901	1.57302
1	178,932	1.687	1.548	262,584	3.17	1969.3	0.28365	1.55646
Number of Observations	16,580,737	2,345,642	1,316,797	2,345,645	2,279,141	2,296,775	16,618,680	16,580,737

Table 1 presents a descriptive comparison of both participating and non-participating buildings. The reference group for this initial comparison is a general commercial utility customer. We see significant differences in observable characteristics. Audited buildings are generally larger (by square footage), more energy intensive (in kWh per month and kWh per square foot), and are more likely to have been renovated. In terms of building vintage, the average year of construction for participating buildings is 1969 and for an average non-participating building is 1959. While the groups are reasonably balanced by billing climate zone and average taxes paid per square foot, participating and non-participating buildings are also significantly different in energy intensity during historical baseline years 2006-2009. This implies that a reference group comparison based on the average commercial building in the City would prove an ineffective reference group for counterfactual analysis, even after controlling for observable characteristics.

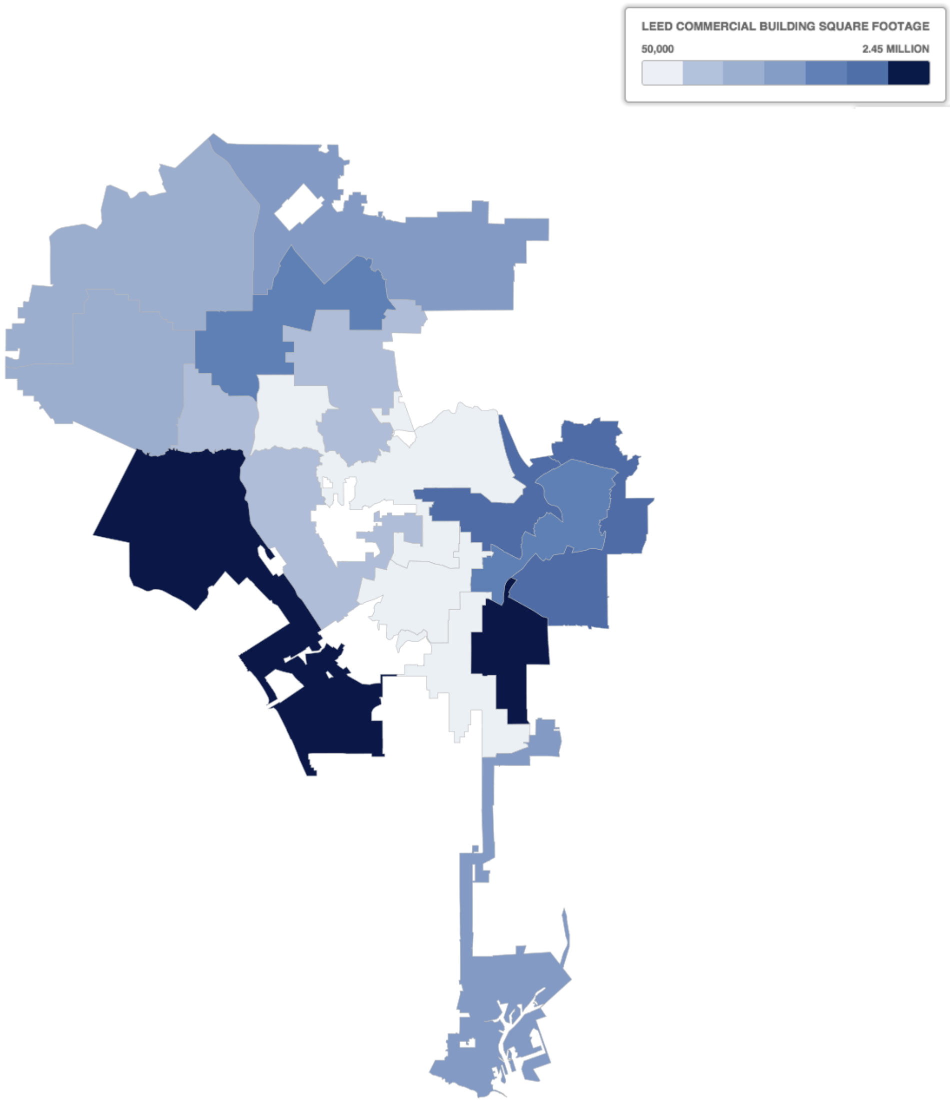
These baseline differences are not too surprising and likely reveal targeting by program managers and administrators, and self-selection into the challenge program. An OLS regression that includes building fixed effects would therefore be severely biased because after challenge participants have adopted new energy efficient technologies following program audits, the treatment effect estimates would also reflect bias in the underlying distributions.

Figure 1.
Distribution of Commercial Building Types (continued on the following page)



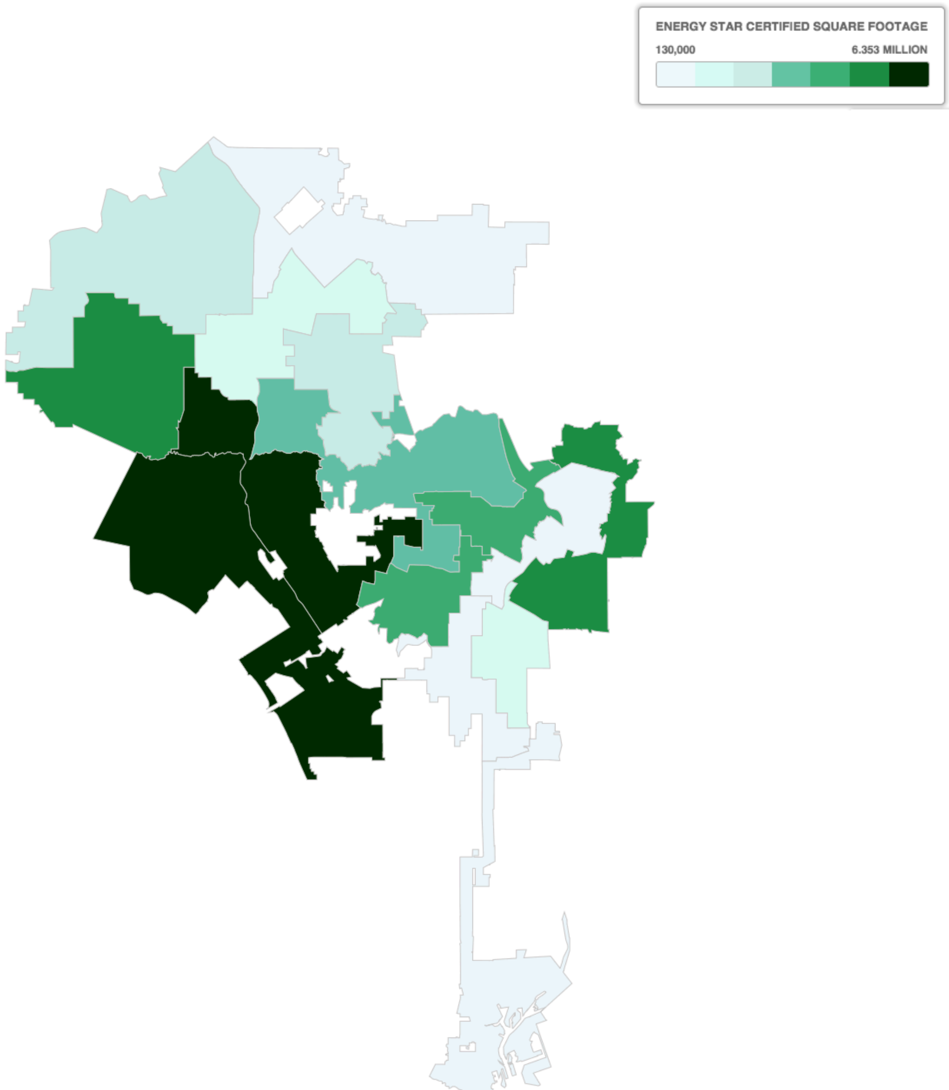
A. LABBC Building Audits

Figure 1. (Continued)
Distribution of Commercial Building Types



B. LEED Certified Buildings

Figure 1. (Continued)
Distribution of Commercial Building Types



C. Energy Star Certified Buildings

An important insight behind the use of matching techniques is to analyze observational data by approximating as closely as possible a completely randomized experiment. This is done by examining observable explanatory variables to adjust for differences in outcomes unrelated to the treatment (Heckman & Navarro-Lozano, 2004). Implementing matching estimators is most effective in the case when there is a large available pool of possible matches and when the underlying covariate distributions reveal overt bias due to systematic differences between treatment and the reference group (Dehejia & Wahba, 2002) as shown in the previous section.

Table 2 lists the most important observable building characteristics from Costar that have been merged with the monthly consumption data. If the selection process or behavioral response function depends on other unobserved factors, this could be problematic for the estimation of causal effects (Heckman, Ichimura, Smith, & Todd, 1996). For this reason, we also include the Costar building rating (scored from 1-5), which stands as a general proxy for building quality and captures a number of unobservable characteristics.

Table 2. List of Balancing Characteristics

Observable Building Characteristic
Physical Building Characteristics
Year Built
Year Renovated
Building Location/Climate
Climate Zone
Occupancy Characteristics
Rentable Building Area (Sq. footage)
Property Type
Occupancy Rate (Percent leased)
Building Quality
Building Class
CoStar Rating *
Industry Characteristics
SIC Industry Code
Utility Customer Class
Building Operating Expenses
Average Rent
Taxes per Sq. Ft.

** The CoStar building rating system is a national rating system for commercial buildings, which captures a number of characteristics including architectural attributes, structural and systems specifications, amenities, site and landscaping treatments and detailed property type specifics. Ratings reflect commercial real estate quality as valued by investors.*

The CoStar rating is a national rating system for commercial buildings. It captures a number of characteristics including architectural attributes, structural and building system specifications, amenities, site and landscaping treatments and other detailed property type specifics. Ratings reflect commercial real estate quality as valued by investors. In observational studies, variables that affect the response outcome can be distributed unequally across treatment and reference groups.

We implement the Genetic Matching algorithm using procedures described in Sekhon (2011) and Diamond and Sekhon (2013). Genetic Matching is an evolutionary search algorithm that automatically finds the optimal balance optimization for covariates by automatically determining the weight each covariate is given. We evaluate its performance versus other commonly used matching strategies based on logit estimation of propensity scores. The initial theory for use in economic applications and its subsequent implementation from machine learning was motivated in Mebane and Sekhon (1998). Increasing the sample size improves the performance of the genetic matching algorithm, which is favorable in this application with relatively high ratio of non-participating to participating buildings. Importantly, genetic matching avoids the manual process of checking covariate balance in post-matched samples and then re-estimating propensity scores. By using an automated process to search the available reference data for the best possible matches, the genetic matching algorithm is able to obtain better levels of covariate balance without requiring the evaluator to correctly specify the propensity score, which can be a strong assumption in many empirical settings.

When using matching models based on estimated propensity scores, decisions have to be made regarding distance measure and the algorithm used. The propensity score is often considered an ancillary statistic for estimating the average treatment effects given the assumption that treatment assignment is ignorable, conditional on observed covariates (Hahn, 1998). Matching on correctly specified propensity scores will achieve balance conditional on the observed covariates (Rosenbaum and Rubin, 1983, 1984). For a given set of covariates, X, indexed by i, the propensity score, $\pi(X_i)$ is the conditional probability of assignment to treatment:

$$\pi(X_i) \equiv \Pr(\text{Treat}_i = 1 \mid X_i) = E(\text{Treat}_i \mid X_i) \tag{1}$$

The observed covariates are conditionally independent to the assigned treatment, given the true propensity score (Rosenbaum and Rubin, 1983), such as,

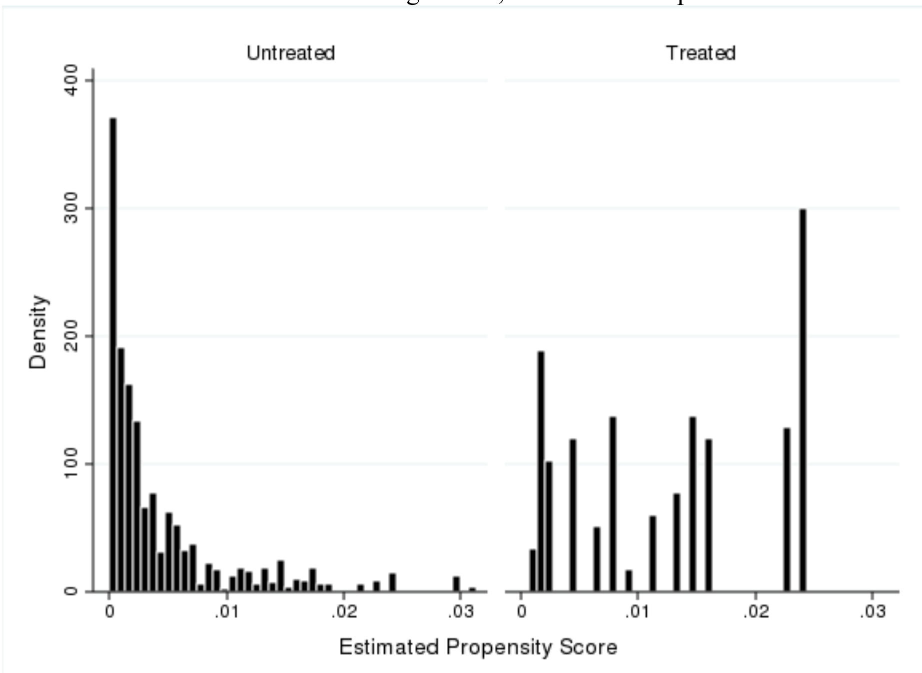
$$X \perp \text{Treat} \mid \pi(X) \tag{2}$$

One of the central benefits of PS matching is that outcome data is not used in the estimation of propensity scores. However, covariate imbalance after matching can still be a concern. When there is a lack of sufficiently sampled data on the common support, this presents a challenge for achieving covariate balance and bias reduction. Prior to treatment, buildings with the same set of *X* observable characteristics should have a positive probability of being both participants and nonparticipants, $0 < \Pr(\text{Treat} = 1 \mid X) < 1$ in the area of common support. There is often a tension between the quality and quantity of available matches in a tradeoff of removing residual bias versus the efficiency of the estimates (Caliendo, Kopeinig, 2008). In data balanced enough to approximate complete randomization, recent evidence has shown that PS matching can actually increase imbalance (King and Nielson, 2015; Rubin & Thomas, 1992). Therefore, it is important to assess covariate balance in the matched sample particularly the extent to which residual bias in the underlying distributions have been mitigated in finite samples.

Figure 2 reports the estimated propensity scores for our unmatched sample of commercial buildings using nearest neighbor matching (NN). Table 3 contains the first stage logit results used to estimate the propensity scores using standard procedures. The left hand side in Panel A refers to all non-participating commercial buildings and the right side refers to the LABBC participating commercial buildings, which defines the region of common support. In Panel B of Figure 2, we see that there is a sufficient sampling density in the region of common support in the matched sample. We therefore maintain the common support restriction in the matching algorithm, using only those observations that lie in the common support. This ensures that only the closest building matches are used when estimating treatment effects.

Figure 2. Estimated Propensity Scores

Panel A: Building Audits, unmatched sample



Panel B: Building Audits, matched sample

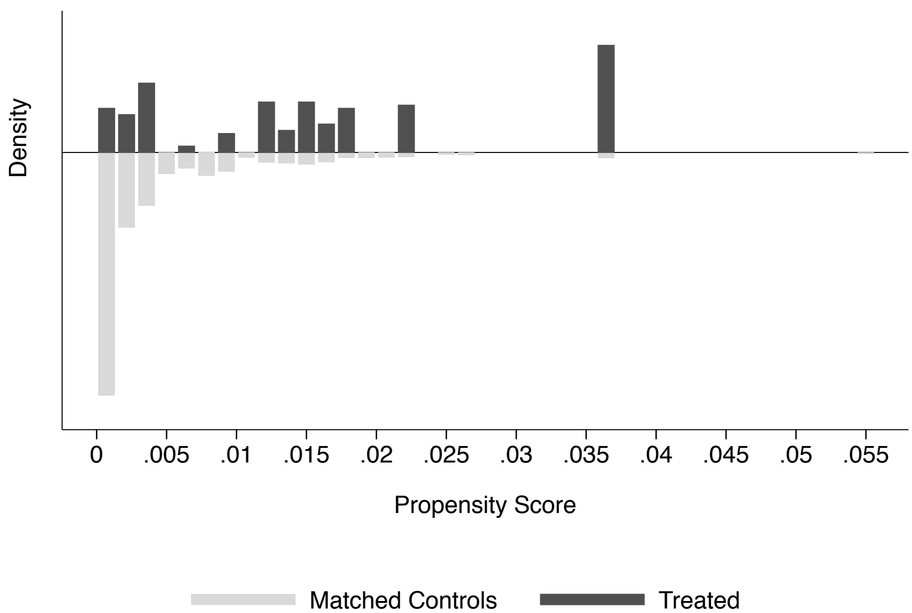


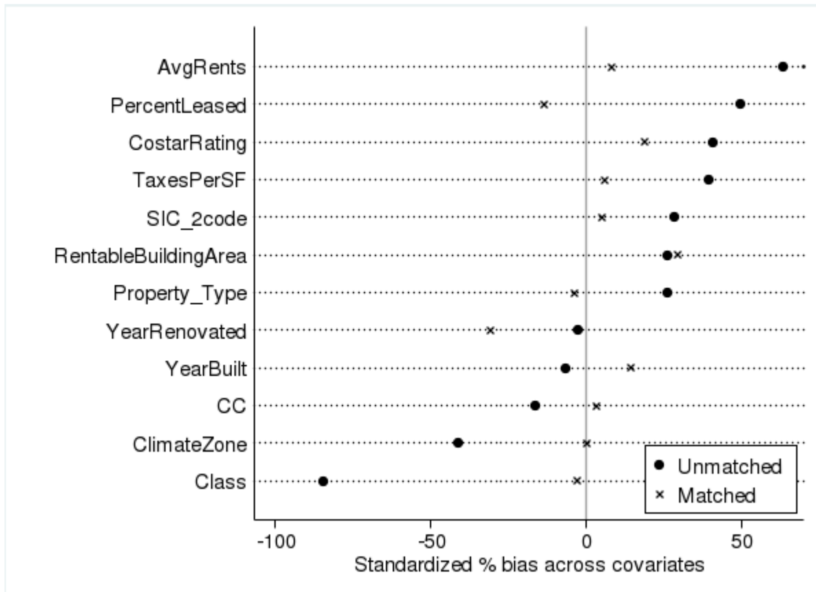
Table 3. Estimation of propensity score with logit

Building Characteristics	Coefficients	Std. Err.	p-values	95% Conf. Interval	
Building Class	-2.32161	0.2582874	0.000	-2.827844	-1.815376
Year Built	-0.0276775	0.0053746	0.000	-0.0382116	-0.0171434
Year Renovated	-0.0084937	0.010705	0.428	-0.0294751	0.0124877
Rentable Building Area	-4.17E-09	3.86E-07	0.991	-7.61E-07	7.52E-07
Property Type	0.0491568	0.0681534	0.471	-0.0844214	0.1827349
Percent Leased	0.0342052	0.0077865	0.000	0.018944	0.0494664
Average Rent	0.0164088	0.0091641	0.073	-0.0015525	0.0343702
Taxes Per SF	0.1439159	0.0493309	0.004	0.047229	0.2406027
SIC 2-digit industry code	0.2440616	0.1208359	0.043	0.0072275	0.4808956
Utility Climate Zone	-0.7324131	0.1704593	0.000	-1.066507	-0.398319
Utility Customer Class	-1.898097	1.39845	0.175	-4.639009	0.8428144
Costar Rating	-0.7147464	0.190264	0.000	-1.087657	-0.3418359
Constant	69.69619	21.16018	0.001	28.22301	111.1694
Observations	39,120				
Pseudo R^2	.17				
Prob > chi2	0.000				

In Figure 3, we report the reductions in standardized percent bias across our covariates after matching. We allow for replacement in the matching algorithm, meaning that matched buildings can be reused in the pool of available matches. We demonstrate in Panel B in Figure 3 that we have effectively removed large systematic differences in the underlying distributions and achieved a high degree of balancing in observable characteristics after matching. Here we show that data-driven algorithms can be quite effective in removing common sources of bias. In a survey of the evaluation literature, Diamond and Sekhon (2013) find that very few studies in the empirical literature verify the extent of covariate balance and hence the reliability of resulting treatment effects.

Figure 3. Bias Reduction in Propensity Score Matched Samples

Panel A: Nearest neighbor without replacement



Panel B: Nearest neighbor with replacement

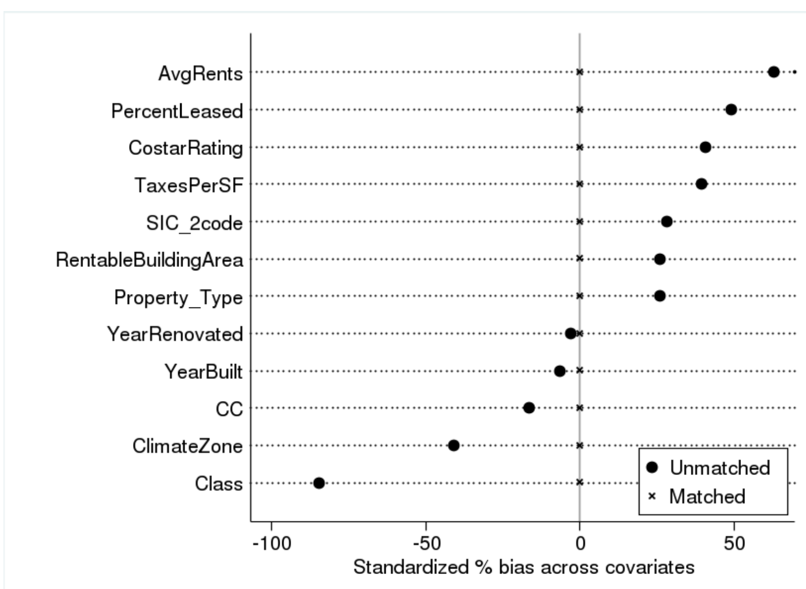


Table 4. Building Energy Performance using various matching algorithms without weather controls

Matching Algorithm	Treatment	Control	ATT	St. Err. (Abadie-Imbens)	T-stat
Unmatched	38.498	63.554	**-25.056**	6.613	-3.79
Nearest Neighbor with replacement (k=1)	38.498	57.522	**-19.024**	3.298	-5.77
Nearest Neighbor with replacement (k=2)	38.498	58.128	**-19.630**	3.978	-4.94
Nearest Neighbor with replacement (k=3)	38.498	57.862	**-19.365**	4.161	-4.65
Nearest Neighbor without replacement (k=1)	35.381	98.538	**-63.157**	4.183	-15.1
Radius	38.498	63.554	**-25.056**	4.073	-6.15
Radius with Caliper (0.05)	38.498	63.554	**-25.056**	2.655	-9.44
Kernel Gaussian	38.498	63.663	**-25.165**	4.097	-6.14
Kernel Epanechnikov	38.498	63.792	**-25.295**	4.097	-6.17
Genetic Matching (Heuristic: minimize pvals)*	38.498	57.522	**-19.024**	7.9876	-2.38
Genetic Matching (Heuristic: minimize QQ distance)*	38.498	57.522	**-19.024**	9.0335	-2.10

Table 5. Los Angeles Better Building Challenge energy savings

Strategy	Average Treatment Effect	Std. Err.	T-stat	p-value
LABBC	-18.69	10.95	-1.71	0.088
LEED certified	-29.99	12.06	-2.49	0.055
Energy Star	-19.31	5.81	-3.32	0.021

Estimated energy savings includes matching and additional weather and time controls

Table 4 lists the detailed results of estimated program level energy savings using the most common matching algoritms including nearest neighbor, radius, and kernel matching methods. These methods rely on researchers to manually check the balancing properties. See Caliendo and Kopenig (2008) for a review of these widely applied matching strategies in the empirical literature. The energy savings (average treatment effects on the treated, ATT) are in the range of 19-25%. Not all matching algorithms are created equal. In Table 4, we observe that nearest neighbor matching without replacement, also referred to as greedy matching, leads to a gross over-estimation of the energy savings due to poorer quality of matches and this matching with replacement is preferred. Both nearest neighbor (k=1) and genetic matching lead to the optimal balancing characteristics and estimates of program level the energy savings without time based weather controls are 19%, although the standard errors are more conservative with genetic matching (see Diamond and Sekhon, 2013). By removing residual bias in the underlying covariate distributions, genetic matching achieves automatic load balancing, effectively weighting control group observations to mimic randomization in finite samples. While matching in this application is effective in simulating group level effects, we are additionally concerned about time-varying seasonal effects due to outside weather variation and other changes over time.

In Table 5, we report the final program level energy savings by difference-in-differences after including weather controls and time fixed effects, using the estimated propensity scores as weights. The final program level energy savings for the LABBC audited commercial buildings is -18.69%, significant at the 10% level. We also report the estimated program level savings for LEED and Energy Star green building certification programs as -29.99% and -19.31%, respectively. These results show that policies that reduce information barriers and target action in individual buildings can be used to accelerate new investment and yield significant energy savings via capital improvements.

Conclusions

We show that information-based programs can be effective at driving new energy efficiency investments as part of larger grand challenge goals. We analyzed technology adoption decisions in response to government-sponsored audits and voluntary certification programs. Our results indicate that information programs that include targeted building level information can lead to adoption of profitable, but previously unimplemented technologies. The energy savings of 18.7% achieved for the LABBC program imply a cost effectiveness ratio of 5.5 cents per square foot. This includes public expenditures of $4.2MM ($3.5MM for the audits and $700M in administrative costs) and private expenditures of $74MM in energy efficiency upgrades. For every public grand challenge dollar invested, this yielded a return of

$17.6 dollars in private infrastructure spending. Overall, 42% of recommendations outlined in the audits were implemented by building owners, which is consistent with implied hurdle rates in prior literature (Anderson & Newell, 2004). In response to community based grand challenges, individual owners and investors appear to be responsive to implementation costs rather than to actual energy savings (Anderson and Newell, 2004; Blass, Corbett, Delmas & Muthulingham, 2014). On restoring and improving urban infrastructure, finding opportunities to link challenge-based interventions with performance outcomes means we will have to devise new methods such as experimental paradigms for causal analysis. Having a unifying framework with community driven goals can be an important strategy to mobilize resources for grand challenges. Given the limits to public finance, public private partnerships can serve to extend traditional boundaries of the public sector and increase directed innovation toward meeting societal goals.

References

Angrist, J. D., & Pischke, J. 2010. The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. ***Journal of Economic Perspectives***, 24(2): 3-30.

Anderson, S. T., & Newell, R. G. 2004. Information programs for technology adoption: the case of energy efficiency audits. ***Resource and Energy Economics***, 26: 27-50.

Allcott, H. & Greenstone, M. 2012. Is there an energy efficiency gap? ***Journal of Economic Perspectives***, 26(1): 3-28.

Asensio, O. I., & Delmas, M. A. 2015. Nonprice incentives and energy conservation. ***Proceedings of the National Academy of Sciences***, 112(6): E510-E515.

Blass, V., Corbett, C. J., Delmas, M. A., Muthulingham, S. 2014. Top management and the adoption of energy efficiency practices: Evidence from small and medium-sized manufacturing firms in the US. ***Energy***, 65: 560-571.

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

Boomhower, J., & Davis, L. 2014. A credible approach for measuring inframarginal participation in energy efficiency programs. ***Journal of Public Economics***, 11: 67-79.

Barrett, C. B., & Carter, M. R. 2010. The power and pitfalls of experiments in development economics: some non-random reflections. ***Applied Economic Perspectives and Policy***, 32(4): 515-548.

Foray, D., Mowery, D. C., & Nelson, R. R. 2012. Public R&D and social challenges: what lessons from mission R&D programs? ***Research Policy***, 41(10): 1697-1702.

Caliendo, M, & Kopenig, S. 2008. Some practical guidance for the implementation of propensity score matching. ***Journal of Economic Surveys***, 22(1): 31-72.

Dehejia, R. H., & Wahba, S. 2002. Propensity score-matching methods for nonexperimental causal studies. ***Review of Economics and Statistics***, 84(1): 151-161.

Delmas, M. A., Fischlein, M., & Asensio, O.I. 2013. Information strategies and energy conservation behavior: a meta-analysis of experimental studies from 1975-2012. ***Energy Policy***, 61: 729-739.

Diamond, A., & Sekhon, J. S. 2013. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. ***Review of Economics and Statistics***, 95(3): 932-945.

Dietz, T. 2010. Narrowing the US energy efficiency gap. ***Proceedings of the National Academy of Sciences***, 107(37): 16007-16008.

Edler, J., & Georghiou, L. 2007. Public procurement and innovation—resurrecting the demand side. ***Research Policy***, 37(7): 949-963.

Edler, J., Georghiou, L., Blind, K., & Uyarra, E. 2012. Evaluating the demand side: new challenges for evaluation. ***Research Evaluation***, 21(1): 33-47.

Edquist, C., & Hommen, L. 1999. Systems of innovation: theory and policy for the demand side. ***Technology in Society***, 21: 63-79.

Gillingham, K., Newell, R. G., & Palmer, K. 2009. Energy efficiency economics and policy. ***Annual Review of Resource Economics***, 1: 597-619.

Gillingham, K. & Palmer, K. 2014. Bridging the energy efficiency gap: policy insights from economic theory and empirical evidence. ***Review of Environmental Economics and Policy***, 8(1): 18-38.

Hahn, J. 1998. On the role of the propensity score in efficient estimation of average treatment effects. ***Econometrica***, 66(2): 315-331.

Howarth, R. B., & Sanstad, A. H. 1995. Discount rates and energy efficiency. ***Contemporary Economic Policy***, 13(3): 101-109.

Heckman, J. J., Ichimura, H., Smith, J., Todd, P. 1996. Sources of selection bias in evaluating social programs: an interpretation of conventional measures and evidence on the effectiveness of matching as a program evaluation method. ***Proceedings of the National Academy of Sciences***, 93(23): 13416-13420.

Heckman, J., & Navarro-Lozano, S. 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. ***Review of Economics and Statistics***, 86(1): 30-57.

Imai, K, King, G., & Stuart, E. A. 2008. Misunderstandings between experimentalists and observationalists about causal inference. ***Journal of the Royal Statistical Society: Series A***, 171(2): 481-502.

Imbens, G. W., & Wooldridge, J. M. 2009. Recent developments in the econometrics of program evaluation. ***Journal of Economic Literature***, 47(1): 5-86.

Ito, K. 2015. Asymmetric incentives in subsidies: evidence from a large-scale electricity rebate program. ***American Economic Journal: Economic Policy***, forthcoming

Jaffe, A. B., Stavins, R. N. 1994a. The energy-efficiency gap what does it mean? ***Energy Policy***, 22(10): 804-810.

Jaffe, A. B., Stavins, R. N. 1994b. The energy paradox and the diffusion of energy conservation technology. ***Resource and Energy Economics***, 16: 91-122.

Jaffe, A. B. 2002. Building programme evaluation into the design of public research-support programmes. ***Oxford Review of Economic Policy***, 18(1): 22-34.

King, G., & Nielsen, R. 2015. Why propensity scores should not be used for matching. Harvard University working paper.

Klette, T. J., Moen J, & Griliches, Z. 2000. Do subsidies to commercial R&D reduce market failures? microeconometric evaluation studies. ***Research Policy***, 29: 471-495.

Mathew, P. A., Dunn, L. N., Sohn, M. D., Mercado, A., Custudio, C., & Walter, T. 2015. Big-data for building energy performance: lessons from assembling a very large national database of building energy use. ***Applied Energy***, 140: 85-93.

Mebane, W. R., & Sekhon, J. S. 1998. GENetic Optimization Using Derivatives (GENOUD). Software Package. http://sekhon.berkeley.edu/rgenoud/.

Mowery, D. C., Nelson, R. R., & Martin, B. R. 2010. Technology policy and global warming: Why new policy models are needed (or why putting new wine in old bottles won't work). ***Research Policy***, 39: 1011-1023.

Mowery, D., & Rosenberg, N. 1979. The influence of market demand upon innovation: a critical review of some recent empirical studies. ***Research Policy,*** 8(2): 102-153.

National Research Council, Grand Challenges in Environmental Sciences (National Academies Press, Washington, DC, 2001).

National Academy of Sciences. 2008. Grand Challenges for Engineering (National Academies Press, Washington, DC, 2008).

Omenn, G. S. 2006. Grand challenges and great opportunities in science, technology and public policy. ***Science***, 314: 1696-1704.

Rosenbaum, P. R., & Rubin, D. B. 1983. The central role of the propensity score in observational studies for causal effects. ***Biometrika***, 70(1): 41-55.

Rosenbaum, P. R., & Rubin, D. B. 1984. Reducing bias in observational studies using subclassification on the propensity score. ***Journal of the American Statistical Association***, 79(387): 516-524.

Rubin, D. B., & Thomas, N. 1992. Affinely invariant matching methods with ellipsoidal distributions. ***Annals of Statistics***, 20(2): 1079-1093.

Sekhon, J. S. 2011. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. ***Journal of Statistical Software***, 42(7): 1-52.

Von Hippel, E. 1986. Lead users: a source of novel product concepts. ***Management Science***, 32(7): 791-805.