The Impact of Behavioral Science Experiments on Energy Policy

Robert Hahn and Robert Metcalfe*

June 9, 2016

Abstract

One of the most exciting areas of research today is the use of experiments informed by behavioral science to understand how to change energy consumption decisions of consumers. This article provides a survey and synthesis of experiments and focuses on general principles that can be gleaned from these experiments to date. We identify four general insights from the literature. First, the "law of demand" is typically satisfied in experimental settings, but responsiveness to energy price changes can vary dramatically in different contexts. Second, information provision can help promote reductions in energy use, but it does not always work. Third, the use of social norms can change energy use. Finally, the economic welfare impacts of behavioral interventions aimed at promoting either energy conservation or energy efficiency are not well understood, but initial research suggests that some people want nudges and some do not. The essay also identifies a number of areas for future research.

*Robert Hahn is Professor and Director of Economics at the Smith School of Enterprise and the Environment, University of Oxford and a Senior Fellow at the Institute for New Economic Thinking. Robert Metcalfe is Research Scholar at the Becker Friedman Institute, University of Chicago. We would like to thank Paul Joskow and Catherine Wolfram for valuable feedback on this article. Brian Campbell and Julia Schmitz provided excellent research assistance. The authors alone are responsible for any opinions and errors. Support for this research was provided by the Smith School and the Institute for New Economic Thinking at the University of Oxford and the Becker Friedman Institute at the University of Chicago.

The Impact of Behavioral Science Experiments on Energy Policy

Robert Hahn and Robert Metcalfe

1. Introduction

More than fifty years ago, Edward Chamberlin demonstrated that economists could use experiments productively to study markets in a laboratory setting (Chamberlin, 1948). These experiments eventually provided very useful information on the relative performance of different market rules and regulations on economic efficiency (Smith, 1976). More recently, there has been an explosion in the use of "field" experiments to test and understand how humans respond to behavioral interventions in real world settings (Harrison and List, 2004; List and Metcalfe, 2014). These experiments have provided a wealth of information on topics including how to improve education, understand charitable giving, increase health quality and productivity, and reduce poverty.

One of the most exciting areas of research today is the use of behavioral science experiments to better understand how various behavioral interventions change the energy consumption decisions of consumers and businesses. Many policy makers face severe constraints, and are unable to alter markets, regulation, or taxes, so that the price of energy reflects its full social cost. Understanding and changing the pattern of energy consumption is critical for a variety of reasons, including meeting the challenge of climate change by reducing greenhouse gas emissions.

This article provides a survey and synthesis of behavioral experiments that may affect energy consumption. We focus primarily on experiments examining how alternative behavior interventions affect residential electricity consumption because that is where most experiments have been done.

When economists started conducting experiments in the energy sector in the 1970s, the focus was on the importance of economic variables, such as price, in affecting consumption decisions. The "law of demand" tells us as the price of something increases, people generally consume less of it, and this is generally borne out in traditional econometric studies of energy demand. However, the story is much richer than that. We now know there are a host of ways in which behavior can be affected in addition to traditional price and income variables (Kahneman and Tversky, 1979; Thaler and Sunstein, 2008; DellaVigna, 2009; Dolan *et al.*, 2012). These include supplying people with additional information relevant for making decisions, and changing the way information is presented and framed. Thus, it will be useful to explore experiments that focus on price and non-price interventions, as well as experiments that focus on both simultaneously.

In this review, we summarize *general principles* that can be gleaned from the behavioral science experiments aimed at examining energy consumption. Many of these were discussed in the excellent review by Price (2014). We build on Price's review in several ways. First, we highlight general insights that will be of use to policy makers. Second, there is a growing empirical literature in this area, so we focus on some of the more recent empirical results. Third, there are some areas, such as welfare analysis associated with "nudges", on which insights are only beginning to emerge.

We will focus on "field experiments" that try to understand the likely impact of different treatments in a real-world setting. If subjects are unaware they are in a field experiment, we refer to this as a natural field experiment, following Harrison and List (2004). By not telling subjects they are participating in an experiment, the experimenter reduces or eliminates the chance the subject changes her behavior because they know they are in an experiment. In a typical version of a natural field experiment, subjects are randomly selected from the population, and randomly assigned to a treatment group (or groups) and a control group. Randomly selecting agents from the population creates a representative sample that is similar to the population itself. Random assignment to treatment and control groups assures that the two groups are virtually identical, at least from a statistical point of view. The experimenter can then compare the treatment with the control to see, for example, if the average effect for the treatment was different from the average effect for the control.[1]

A natural field experiment can include various kinds of designs. One design that is sometimes used is a randomized encouragement design. This design can help measure a treatment's impact by, say, randomly changing the probability a customer participates in a program. A second type of natural field experiment design is to phase in a program or treatment over time, such as the installation of free smart meters. The program can be randomized so that each new group receiving the treatment is randomly selected. Those that have not received the treatment yet serve as the control group – i.e., they have been denied the treatment for a given period of time. The point is that there are many approaches to designing field experiments, which allow the experimenter to compare treatments with the control using rigorous measures.[2]

Sometimes, subjects may need to volunteer for an experiment, or at least be told they are part of an experiment. Following Harrison and List (2004), we call this a "framed field experiment."[3,4] Individuals, thus, may opt in to a framed field experiment. Once part of the experimental population, they are typically allocated to treatments and the control group randomly. This experimental approach is often used in testing the efficacy

---

[1] Selecting participants randomly from the target population and allocating them randomly to the control group and treatment groups are key parts of a randomized controlled trial (RCT).
[2] See Gandhi et al. (2016) for a useful discussion of different types of field experiments as they apply to energy efficiency programs, and Al-Ubaydli and List (2012) for a general discussion of the design of field experiments.
[3] A framed field experiment might have no selection into the experiment, but the subjects still might know they are part of the experiment. Generalizability in that context needs to take account of any behavioral consequences as a result of the subject knowing that they are being watched and randomized into an experimental group.
[4] A third kind of field experiment is an artefactual field experiment where market participants' behavior is studied in the laboratory (Harrison and List, 2004). We will not focus on that type of field experiment here.

of new drugs. A key difference between natural and framed field experiments is the ease with which results can be generalized to the population of interest. If, for example, subjects opt in to a field experiment, there are typically selection effects, which often limit the generalizability of the results.[5] For both of these different types of field experiments, the variable of interest is randomized across people (Harrison and List, 2004; List and Metcalfe, 2014). Much of the literature reviewed in this paper consists of framed and natural field experiments.

Many of the experiments in our review attempt to reduce energy demand. This demand reduction can be separated into two parts: energy conservation and energy efficiency. “Energy conservation” is used to mean reductions in energy use, holding the energy use per unit of output constant. “Energy efficiency” measures energy use per unit of output. A decrease in the use of energy per unit of output means energy efficiency has increased.

Here are some of general insights from the literature. First, the “law of demand” is typically satisfied, but responsiveness to price changes can vary dramatically in different contexts. Second, information provision can help reduce energy use, but it does not always work in a cost effective way or at all. Third, the use of social norms can reduce energy consumption in many contexts.[6] Finally, the economic welfare impacts of behavioral interventions aimed at promoting energy conservation are not well understood, but initial research suggests that some people want nudges and some do not. More generally, it is important to distinguish between policies that reduce demand and policies that increase overall economic welfare or efficiency. We measure the latter by net changes in producer and consumer surplus. Thus, policies that promote conservation may or may not increase economic welfare as defined here.

Section 2 of the paper provides a review of the literature. We divide our treatment into four parts: price interventions, information interventions, social norm interventions and welfare impacts. Section 3 explores policy implications. Finally, Section 4 concludes and offers some suggestions for future research that could be useful to policy makers.

2. Literature review

This literature review is not meant to be exhaustive. We searched the literature for papers published in peer-reviewed journals, papers in established working paper series, and papers by authors doing research in this area.

We review many recent papers that use either natural or framed field experiments. In addition, we contrast this with some earlier work. Our intent is not to give a detailed

[5] Both randomized encouragement design and opt-in and deny design can be classed as a framed field experiment if the participants know they are part of an experiment.

[6] A social norm is defined an expected type of behavior in a particular situation. For example, saying that nine out of ten of your neighbors turn down their thermostat at night, and asking you to do the same thing, would be considered a social norm. See Cialdini and Trost (1998) and section 2C for a more detailed discussion.

critique of methodology, but rather to give the reader a sense for the kinds of issues that have been addressed, as well as future research opportunities.

A. The law of demand generally works, but details matter

There have been many experiments aimed at understanding the price elasticity of demand for energy. The key question that economists have focused on is the impact of pricing peak and non-peak usage differently, sometimes referred to as time-of-use or dynamic pricing for electricity. The idea behind these experiments was to raise the price of energy for customers during periods of high demand, so that it more closely approximated the marginal cost of production during these periods.

One of the earliest published framed field experiments on electricity consumption was by Battalio *et al. (1979).* The authors tested the impact of rebates, feedback and information on energy consumption, but the sample was only about 100 households who receive monthly bills, divided into four treatment groups and one control group. The authors use an opt-in design for their framed field experiment, where households voluntarily choose to be in the experiment. Because people volunteer, the sample may not be representative of the general population, and their conclusions may not generalize. All four treatment groups in the study received information on their energy use in different ways; three groups received feedback on their energy use relative to a previous baseline; and two of those groups received rebates for reducing their electricity use, increasing the marginal price per KWh by 50%, and 235%, respectively. The authors found that the two groups receiving rebates significantly reduced their consumption, while groups receiving only feedback increased their usage for one month after the experimenting started.

In the 1970s, there were a number of time-of-use framed field experiments in the United States. For example, Caves and Christensen (1980) estimated price elasticities for a time-of-use experiment for residential electricity consumption using data from Wisconsin in 1977. The authors studied roughly 600 customers in a field experiment who faced ten different pricing treatments. The control group paid the same price for electricity all day. The experimenters varied the length of the peak and the ratio of the peak price to the off peak price. Their research suggests that peak shifting is possible, but the sample sizes were small making inference difficult.

In another time-of-use assessment, Caves, Christensen, and Herriges (1984) examined five of the early U.S. time-of-use field experiments. The experiments were conducted by Carolina Power and Light, Connecticut Light and Power, Los Angeles Department of Water and Power, Southern California Edison, Wisconsin Public Service to study the impact of mandatory time-of-use rates. The experiments varied greatly in design. The authors note that the price difference between peak and off-peak electricity consumption is the primary factor affecting demand response; factors such as customer characteristics, appliance holdings, and climate matter, but are less important in predicting the demand response. Their results were consistent with the law of demand.

We now consider more recent work in this area. In line with both laboratory experiments and other field experiments, we find that later research on the effects of price interventions on energy consumption expands the scope of inquiry, adds more statistical rigor, and frequently uses larger sample sizes than were used in early experiments.

Our discussion of later papers begins by focusing on those addressing the retail pricing of electricity. Then we review papers that address other forms of energy consumption, and papers that address both price and non-price interventions. Faruqui and George (2005) estimated the impact of varying electricity prices by time of day. They discussed the implications of the California Statewide Pricing Plan in 2003 and 2004, which studied about 2,500 residential and small commercial and industrial customers in California. The participants were divided into treatment groups with different rates and price schedules, including a time-of-use pricing group, and groups of critical peak pricing schedules that faced either fixed or variable lengths of critical periods. The control group paid the standard rate. The authors found peak period reductions of 5% for the time-of-use treatment that had a peak-to-standard price ratio of about 2 to1, while the peak pricing groups with price ratios of 5:1 and 10:1 reduced peak usage by about 8% and 15%, respectively.

To investigate how customers respond to dynamic pricing, Wolak (2006) sampled about 100 residential customers in Anaheim, California. Between June and October 2005, the treatment participants were subject to critical peak price days between noon and 6 p.m., while the control participants paid according to the standard increasing-block, fixed-price schedule. The treatment participants received a rebate for decreasing their use during these times relative to their base usage. Wolak found the treatment group responded strongly to the rebate by decreasing their usage by 12% during these times.

Allcott (2011a) analyzed data from a framed field experiment on real-time pricing conducted by ComEd in Chicago in which about 700 households opted in to the study. ComEd set hourly electricity rates for the treatment group each day (based on the supply costs of energy for ComEd). Control group households remained on the standard ComEd residential tariff (8.275 cents per kWh in the summer and 6.208 cents/kWh in other seasons). For the treatment group, prices were set so that the expected total electricity bills would be slightly lower compared to the standard tariff for a typical household. The retail prices ranged from 4.6 to 16.0 cents/kWh, and there were 30 hours on nine days in which the wholesale component of price exceeded the High Price Alert cutoff of 10 cents/kWh. Customers in the treatment group also had a small plastic globe that changed color in real time on a continuous spectrum from blue to red, indicating low to high electricity prices, to help the transition to the new prices. Allcott found a low price elasticity (around -0.1). He also found that there was no net load shifting from the peak to off-peak, and that households reduced their consumption during peak times and slightly during non-peak times. He estimated that welfare increased by hundreds of millions of dollars, but noted that this was a small fraction of a household’s total electricity expenditures.

There are a number of studies that explore the impact on energy consumption of both price and non-price interventions experimentally. Jessoe and Rapson (2014) conducted a framed field experiment where they compared the impact of price increases and information. They studied around 400 households in Bridgeport Connecticut in 2011. All households experienced price increases between 200% and 600% in selected periods. In addition, some were given in-home displays that provide real-time information. They found that those households that experienced price increases alone decreased demand by 0% to 7%, and those that received addition information feedback decreased usage from 8% to 22%, depending on when they received the information. The study highlights the positive impact of providing consumers with decision-relevant information (addressed below) and the need to examine the impact of the timing of information.

LaRiviere *et al.* (2014) studied the impact of prices and nudges on investments in energy efficiency by residential households, and attempt to understand their impact on energy use. The authors conducted a natural field experiment in a medium-sized metropolitan area in the US with a sample of about 100,000 households. They sent out ads for an in-home energy audit and randomized information received on the ads and the price of an in-home energy audit. They then measured electricity consumption effects of the various treatments. The authors found that behavioral nudges can affect the consumer responses to energy audits, but have less of an impact on actually installing equipment than changing the rebates. The authors did not have enough statistical power to detect actual electricity consumption differences as a result of the installations arising from the energy audits. However, they found that the ads for in-home audits that provided information on personalized electricity expenditures or emissions reduced electricity consumption by 1% for households with greater than average electricity consumption.

While most of the experiments in the energy literature were done in the United States, there is increasing interest outside the U.S., Ito, Ida, and Tanaka (2015) conducted a framed field experiment in Kyoto, Japan, in which they tried to understand the impact of moral suasion and economic incentives on electricity consumption. Approximately 700 households opted in to an experiment with a control and two treatments, studied during the summer of 2012 and winter of 2013. The moral suasion treatment group was told about the importance for the greater good of decreasing electricity usage during critical peak demand days due to relatively limited supply. On peak demand days, the economic incentive group was charged a critical peak price higher than the baseline price paid by the moral suasion and control groups.

The authors found that moral suasion had an impact on energy consumption, but the impact was sustained for only a few days. In contrast, the price increases led to a marked and sustained reduction during peak periods -- from 14% to 17% depending on the critical peak price. The implied price elasticity estimates for summer and winter were both around -0.14 – i.e. a ten per cent increase led to a 1.4% decrease in use. They also found a reduction in the non-peak hours on treatment days for the economic incentive groups, implying that the economic incentive encouraged the treatment group to reduce usage throughout the day. Furthermore, they found that the impact of the economic incentive treatment persisted even after the incentives were eliminated, suggesting that people may form habits in response to the economic treatment. They did not find similar habit formation for the moral suasion group.

Pellerano *et al.* (2015) completed a natural field experiment in Quito, Ecuador, where they tested the impact of financial incentives, information, and social comparisons on changing energy consumption. The experiment involved a randomized trial of about 28,000 households. The main findings were that the social comparison messages reduced consumption by about 1%. However, financial incentives added to a social comparison does not further decrease electricity usage, and may even backfire. Indeed, they found that such incentives may be counterproductive, calling into question the law of demand in this context. Maybe the price is not salient to these consumers or the incentives are not large enough.

A brief summary of the literature on experiments involving the impacts of pricing suggests several lines for future research. First, most of the work is on electricity consumption by residential customers in selected places in the U.S. The impacts of price changes on other forms of energy consumption, such as consumption of gasoline, heating oil, and natural gas have not been explored extensively. Similarly, non-price treatments have not been considered for the consumption of other types of energy. The academic focus on research with utilities may be explained by the fact that this sector is regulated and some state regulators encourage experimentation. Because most of the work on prices has been done in the U.S. we don't know the extent to which it generalizes to other countries, though we are inclined to believe the law of demand still holds. It would be useful to extend the analysis beyond households to include businesses.

B. Information can help improve outcomes, but not always

There are a smaller number of experimental studies that focus solely on the impact of providing additional information to consumers about their energy consumption. Here we review three studies that are based on providing consumers with additional information and the effects of such interventions on residential electricity consumption. The upshot is that these studies all show a decline in electricity use with the information treatments used in these experiments. We then consider two studies that focus on consumer product choices. One showed that the information provided leads consumers to increase their purchases of more energy-efficient light bulbs; a second experiment suggests that information provided by the Energy Star label does not result in an increase of purchases of products with the Energy Star label.

Kahn and Wolak (2013) worked with two California utilities in 2011 to study how customers responded to information about the nonlinear price schedules that define how much they would pay at different consumption levels. Their design was a framed field experiment, where they sent invitations to a potential treatment group of roughly 18,000 customers, offering a range of Amazon gift cards if the customer would participate in the study. Roughly 2,000 customers accepted and took an on-line educational course about their own energy use and pricing. Kahn and Wolak found that customers changed their electricity demand based on the marginal cost that they face six months after the treatment. Those who faced a higher marginal cost reduced their consumption while those who faced a lower marginal cost increased theirs.

Ivanov *et al.* (2013) reported on a smart meter framed field experiment in Minnesota with a 1,000 participants. The treatment group had programmable thermostats, but the control did not. The authors found that the treatment group used about 15% less energy on peak days. In a natural experiment, Gans, Alberini, and Longo (2013) found that the programmable thermostats, which provide immediate feedback on electricity consumption, can make a positive difference. Examining a sample of residential customers in Northern Ireland, the authors suggested that information provision was associated with a decline in usage of 11% to 17%. There is also some evidence that suggests that once customers know they are being watched, they change their behavior. In the field experiment by Schwartz *et al.* (2013), the authors randomly notified approximately 2,800 households five times over a course of a month and told them that their energy was being monitored for a study. They found a 2.7% reduction in use through this monitoring effect. So some of the benefits arising from smart meters might be due to a monitoring effect. Bollinger and Hartman (2015) randomize different in-home automatic technologies and different price signals across households who opted in for a demand response program. While they demonstrate significant reductions in energy use during peak days, they vary both technology adoption and price at the same time, limiting the ability to identify whether one or both factors are critical.

We review two experiments that examine the effects of various behavioral interventions on the purchase of more energy-efficient products. Allcott and Taubinsky (2015) examined the market for light bulbs. Customers were given cash and asked to make decisions between energy-efficient bulbs and conventional bulbs. The authors found that information about relative electricity use and replacement costs of the different bulbs increased the market share of energy-efficient bulbs by about 12%. Allcott and Sweeney (2014) completed a natural field experiment in the U.S. to examine the effects of providing additional information on the choice of products with Energy Star labels, such as water heaters. They found that providing information was not helpful. They conjectured that sales agents do not try to market these products because customers are not generally interested. They did find that financial incentives matter. For example, a $100 customer rebate along with an agent incentive increased purchases of Energy Star products by up to 22%.

Taken together, these experiments show that information can be helpful in in helping consumers to make more energy-efficient consumption and product choice decisions, but it is not always helpful. However, the context matters and the details matter. More natural field experiments will be needed to develop generalizations about what works.[7]

[7] In their review on the impact of information on energy use, Buchanan, Russo, and Anderson (2015) find surprisingly little causal evidence that it matters, and that there are many poorly designed energy studies out there.

C. The use of social norms can increase energy conservation and the adoption of energy-efficient technologies

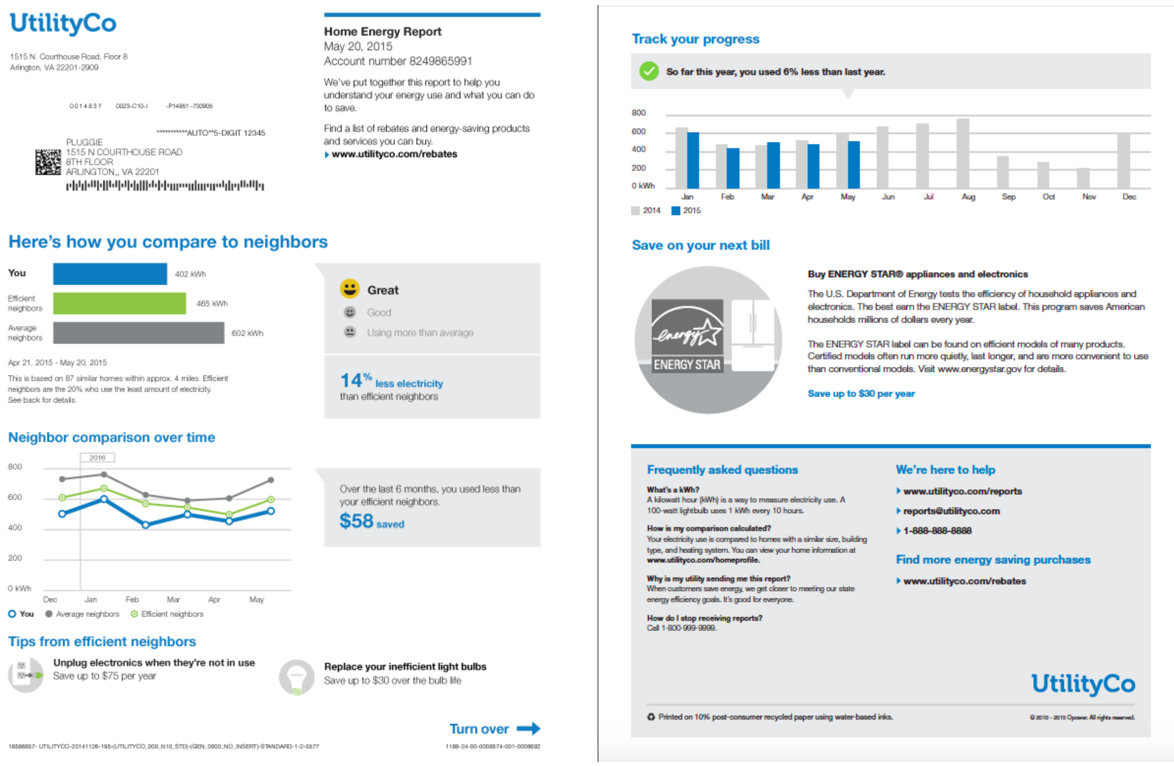
An increasing number of studies have suggested that reference to social norms can change a whole range of individual behaviors (Kluger and DeNisi, 1996; Sherif, 1936). The exogenous impact of social norms has been identified by economists mainly in charitable giving (Frey and Meier, 2004), voting (Gerber and Rogers, 2009), retirement savings (Beshears *et al.*, 2015), water consumption (Ferraro and Price, 2013), and tax compliance (Hallsworth *et al.*, 2014). There is also a growing literature that shows that social norms can impact energy consumption behavior. Social norms have many different definitions. We define it as an expected type of behavior in a particular situation (Cialdini and Trost, 1998). Economists and psychologists doing field experiments often use a “descriptive” social norm (e.g., describing how a subject’s energy consumption compares with her neighbors) or an injunctive social norm (*e.g.*, describing how other people think about the subject’s energy consumption).

In the energy world, the norm is typically assumed to be average behavior of people in similar markets. However, if the average is not standardized for key variables, such as the number of household members or the number of hours occupied, it is not clear that the population will in fact interpret “average” as a meaningful social norm. For example, the Opower Home Energy Report (HER) notifies people of their monthly energy use, and then states the average monthly use for similar sized properties in their neighborhood (see figure 1 below). This is an example of a descriptive social norm, which has been show to be effective. The HER also describes whether the consumption of the household is good or bad from a moral standpoint. It does that by providing an injunctive statement with an emoticon (see figure 1 below). The first example of descriptive and injunctive norms in the energy literature comes from the pioneering work of Schultz *et al.* (2007).

Figure 1

Panel A: HER Front Page

Panel B: HER Back Page



The most comprehensive account of how social norms affect energy consumption is found in Allcott (2011b). He evaluated the Opower HER in many natural field experiments around the U.S.[8] Examining approximately 600,000 treatment and control households, he found that the HER on average reduces energy consumption by 2.0%. He estimated that the effect is equivalent to that of a short-run electricity price increase of 11% to 20%, and the cost effectiveness of using norms compares favorably to that of traditional energy conservation programs. In these data, he found that those customers who are the largest energy consumers are the most likely to reduce their energy demand. From all of the field experiments completed by Opower until the end of 2015, they have estimated a $1 billion saving for utility customers, which amounts to about 8 TWh of energy saved.[9]

The work by Ayers *et al.* (2009) and Costa and Kahn (2012) support these results. However, the work of Costa and Kahn suggests that there is some heterogeneity in the results. They took Opower data from one deployment in California and examined the role that political ideology and environmentalism play in determining a change in energy behavior. They collected data on “green households” by examining whether a household had an individual political party registration, had donations to environmental organizations, had participation in renewable energy programs, and data on the characteristics of the local residential communities where the households lived. They categorized these “green households” types as environmentalists, and found that these types of households are more likely to respond to the HER.

[8] Residents can opt out of treatment but very few actually do.
[9] Taken from the Opower website.

These results tie in nicely with the work of Allcott (2015), who examined whether the early utility clients of the Opower HER differed from later clients. He found that the treatment effects from the HER in the first ten deployments were larger than the treatment effects in the later deployments. This result is important for the generalizability of such behavioral interventions. For instance, he found that utilities in areas more supportive of environmental issues are more likely to adopt the program, and their customers are more responsive to the treatment. Also, because utilities initially targeted treatment at higher-usage consumer subpopulations, efficacy dropped as the program later expanded. The results from this paper demonstrated that continuous field experiments are needed to understand the efficacy of behavioral interventions over time and location. This is a point raised in other research on field experiments (see the overview in List and Metcalfe, 2014).

The Opower HER had an assumed social norm and also had a great deal of information on hints and rebates to reduce energy use. Dolan and Metcalfe (2013) examined the impact of descriptive social norms separately from information on energy efficiency in a natural field experiment. Dolan and Metcalfe demonstrated that social norms reduce consumption even in the absence of information on efficiency. They randomized their interventions on approximately 600 households in London that had subsidized energy prices. They also then extended their study to examine how social norms interact with financial incentives. In a second natural field experiment, they found that financial incentives lead to very large reductions in use, but combining social norms with financial incentives nullifies any of the conservation benefits of financial incentives. These findings in Dolan and Metcalfe are supported by the field experiments in Pellerano *et al.* (2015). Understanding how social norms interact with prices remains an important area of future research.

In an extension of these papers, Allcott and Rogers (2014) attempted to understand how persistent the social norm HER effect is on energy conservation. For those customers in the treatment group, a random subset had their HER withdrawn. The field experiment took place in the service territory of a large West Coast utility, and comprised of approximately 79,000 households. The first home energy reports were mailed in October 2008. Approximately 11,600 households were randomly selected to stop receiving reports after September 2010 – the 'dropped group.' The remainder of the treatment group, which we call the 'continued group,' is still receiving reports at their original frequency. Allcott and Rogers found that when the HERs were discontinued after two years of duration, about two-thirds of the 2% treatment effect remained two years later. The energy reductions in energy consumption decline over time, but they do not disappear.

All of these studies have focused on the effects on energy consumption of giving consumers information on how their energy consumption compares to average energy consumption in their neighborhood. They do not focus on particular behaviors, such as getting consumers to adopt energy-efficient lightbulbs, or turn the lights off at night.

We next turn to studies that have focused on technology adoption and demand response. Herberich, Price, and List (2011) focused on the effects of several interventions on individuals' decisions to adopt energy-efficient technologies. They evaluated the relative importance of altruism and social pressures on the adoption decision for energy efficient lightbulbs in the context of a large-scale door knocking natural field experiment in Chicago. They explored the effect of (i) a subsidy that serves to lower the price of purchasing the lightbulbs and (ii) the use of a descriptive "social norm statement" detailing the proportion of households in their neighborhood that have adopted the lightbulbs. Their theoretical model predicted that social norms would influence behavior along the adoption margin, but would have no impact on the number of lightbulbs purchased. In contrast, subsidies (or changes in prices) are shown to impact behavior on both the adoption decision and the number of lightbulbs decision.

They found that households in the social norm group were approximately 3 to 5 percentage points more likely than counterparts in their neutral frame treatment to purchase at least one packet of energy efficient lightbulbs. Prices had a similar impact on rates of purchase. Conditioned on opening the door, households were approximately 8 percentage points (or 65%) less likely to purchase a package of lightbulbs at their high ($5) price level. Importantly, they found that prices and descriptive norms influence behavior along different margins. Social norm statements affect the decision about whether or not to buy a lightbulb, but had no impact on the number of packets of lightbulbs the consumer purchased at the door. Prices, in contrast, primarily affect the decision of how many lightbulbs to buy. Households that were offered lightbulbs at $5 per package were 70% less likely to purchase a second package than counterparts who could have purchased lightbulbs at $1 per package. The authors found that the implicit price of a descriptive social norm is substantial, with the demand adjustments due to the social norm statements equivalent to between a 30% and 70% price reduction.

LaRiviere *et al.* (2014), discussed above, also focused on using social norms to increase technology adoption. This study attempted to use different types of norms. They sent approximately 55,000 households information on how their electricity usage related to other households in their neighborhood over a twelve month period. They then offered the households the opportunity to sign up for an in-home energy audit in addition to tracking their subsequent energy consumption.[10] They conveyed the usage information to households in one of four ways: their relative use in KWhs, their relative use in monthly expenditures, their relative use in CO_2 emissions and their absolute use in KWhs with no comparison. For audit uptake, they found that only the KWh comparison significantly increased the likelihood of audits (by 400-500%). Subsidies also increased the likelihood of uptake. The value of a KWh signal was roughly $20. Put another way, the subsidy required to increase audit uptake by the same probability as a KWh comparison was $20.

[10] In home energy audits are very common programs offered by electricity utilities and wholesalers in which firms enter a household to identify the best ways the household can reduce their electricity usage. The auditor also provides engineering estimates of monthly savings.

There have also been studies on using moral appeals to reduce consumption on peak days, when the marginal costs of supplying energy are typically high. The Ito, Ida, and Tanaka (2015) study, discussed above, used moral suasion to attempt to curtail peak consumption. The moral suasion group received a message requesting voluntary energy conservation with no economic incentives. They find that moral suasion group showed a usage reduction of 8% for the first few treatment days. However, their usage became statistically indistinguishable from that of the control group for the remaining interventions. It is interesting that social norms in terms of a HER have a persistent effect on reducing energy consumption (see Allcott and Rogers above), but here they have no long-run impact on energy behavior. There are many differences between the two studies that make it difficult to identify what is driving the difference.

This section reviewed the key evidence that suggests social norms encourage households to reduce energy consumption. There is still much to learn in terms of behavioral motivations for conserving energy. One key issue is whether it is possible to go beyond the average 2 to 3% treatment effect with social norms using other behavioral interventions. A second key issue is whether social norms could be used effectively in a business setting to reduce energy consumption. A third key issue relates to understanding the relative cost-effectiveness of changing habits and technologies. In reducing energy consumption, for example, do households tend to change their habits, use new technologies, or both? This question speaks to the work on persistence and mechanisms. If behavioral interventions trigger the adoption of energy-efficient products and technologies, then the benefits of these interventions are likely to be longer lasting than if they only change habits. Understanding the black box of household decision-making will be important for understanding persistence and mechanisms.

D. Estimates of economic welfare impacts are important, and there is little research to date on nudges

Many of the interventions in the previous sections change prices, information, and/or assumed social norms. There have been relatively few studies of the economic welfare impacts associated with changes in information, and social norms. The latter can impose moral costs and benefits on consumers. To understand the economic welfare consequences of many energy products and technologies, we would ideally like to estimate the net consumer surplus of the product for consumers.[11] There is increasing interest in understand the welfare consequences of energy products and interventions. The two studies to date in this area have been focused on the value of social pressure in the adoption of energy-efficient lightbulbs (Herberich, Price, and List, 2011) and the value of the social norm from the Opower HER (Allcott and Kessler, 2015).

In the Herberich, Price, and List (2011) study, as discussed in subsection C, the authors assessed the welfare effects of lightbulb purchases by structurally estimating the parameters of their theoretical model. They found that the customers negatively value

[11] This is welfare in an economics sense – i.e. preference satisfaction. There are other types of welfare that might interest the reader, such as objective wellbeing and subjective wellbeing.

social pressure. When assuming a one year life span and a conservative estimate that households only replace one incandescent light bulb per purchase decision, the resulting welfare represented a loss in welfare for an average household between $3.08 and $4.43. So by applying social pressure to the consumers to buy lightbulbs, they impose a negative $3-$4.50 value on the consumers. However, this range is very sensitive to both the assumed life span of a lightbulb and how many lightbulbs are adopted in the household.

In the Allcott and Kessler (2015) study, they attempted to measure welfare changes of an energy efficiency intervention more directly. They provided a welfare evaluation of the Opower HERs. They used a multiple-price-list tool to elicit the price that people are willing to pay to receive the HERs, or money they are willing to accept to stop receiving the HERs. For most of those individuals who self-selected into the multiple-price-list survey, they had a positive value for the HERs, but 35% of those taking the survey had weakly negative willingness to pay. In the case of these 35%, the HER was reducing their economic welfare. On average, their estimates suggested that the second year of this HER program increases social welfare by $0.70 per household. The HER's welfare effects were driven down by the fact that 60% of its recipients are not willing to pay the social marginal cost of the product, including many who had a negative willingness to pay. On the other hand, more than 30% of recipients were willing to pay more than twice the social marginal cost.

These two studies demonstrated that we do not know enough about the welfare consequences of energy nudges or interventions. There has been some preliminary survey evidence that suggests that people might prefer price or information nudges over nudges that use social norms or framing from policymakers (see Sunstein, 2016). However, this has not been tested in the field. Furthermore, it is unclear whether social norms provide value in terms of extra information on what other people are doing, or whether they impose a social pressure cost on the individual to conform. It is difficult to know what part of the brain social norms lie, but it is not clear there is a consensus. Moving forward, we encourage utilities and policymakers to think harder about how to measure the welfare consequences of their products and policies, with and without social norms.

3. Policy Opportunities and Challenges

In this section, we examine policy opportunities that flow directly from this research. We begin with the observation that it is frequently not politically feasible to price energy at its full marginal social cost. This full cost would include the direct production costs of the energy plus the costs of externalities, such as carbon dioxide.

Much has been written about the political difficulties of raising energy taxes (Davis and Muehlegger, 2010; Kinittel, 2013). For our purposes, the key insight is that behavioral nudges can help change behavior. This explains the success of companies, such as Opower, which provide home energy reports that often use assumed social norms to reduce consumption. There are still many opportunities for using behavioral science to reduce energy consumption cost-effectively in the United States. In the utility sector, many utilities still do not make use of the latest research for residential customers. In the rest of the world, the opportunities appear to be even larger.

From a behavioral point of view, there are many underlying mechanisms that have not been fully explored as a way to increase both energy conservation and efficiency. For instance, one of the main behavioral theories, prospect theory (Kahneman and Tversky, 1979), suggests that people's consumption is anchored at a particular reference point, and around that point losses are weighted more heavily than gains (i.e. called loss aversion). We know very little in energy research about how loss aversion impacts energy demand. Moreover, hyperbolic discounting has been provided as a theory as to why people's behavior is very myopic (see Laibson, 1997). Again, we know very little about how myopia drives energy behavior and whether this behavioral bias leads to large inefficient investments in energy. Indeed, many of the behavioral mechanisms surveyed by DellaVigna (2009) have not been fully established in the energy policy arena.[12] Thus, researchers have an opportunity to generate new insights on the actual impact of many of the behavioral mechanisms on energy use. We have found that forging lasting partnerships with innovative practitioners is one of the best ways to develop these insights.

Behavioral science research has the potential to improve our understanding of the so-called energy efficiency gap or paradox. The basic idea is that people may not be investing in technologies, such as insulation, that have a high rate of return. The energy efficiency gap (Allcott and Greenstone, 2012) and the energy efficiency paradox (Gillingham *et al.*, 2013; Gillingham *et al.*, 2016) will remain active research areas. Furthermore, we need more research on how technology adoption affects energy consumption and economic efficiency. An example of this is the work of Fowlie *et al.* (2015) on weatherization programs. With a sample of approximately 30,000 households in Michigan, they found that the upfront investment costs were about twice the actual energy savings, and that the engineering estimates of the savings were roughly 2.5 times the actual savings. They attempted to rule out any rebound effect, and even found that when accounting for the broader societal benefits of energy efficiency investments, the costs still substantially outweigh the benefits.

Another area of policy relevance is whether policymakers should subsidize energy efficiency measures in the commercial and industrial sector. Traditional theories of industrial organization assume that firms in perfect competition are already using energy inputs efficiently. Even if we assume imperfect competition, firms should still want to minimize operational costs, and therefore energy costs, to maximize profit. While there

[12] For example, policymakers sometimes have a choice as to whether they will ask customers to opt into or opt out of a program, such as one promoting energy efficiency or renewable energy. Ebeling and Lotz (2015) report on a natural field experiment conducted by a German utility to test whether an opt-out mechanism for a more expensive green tariff increases adoption relative to an opt-in mechanism. In this case, the tariff is more expensive because it reflects the cost of renewable energy. The authors found that setting the default choice to opt out for the green energy tariff (i.e., consumers have to actively opt out if they do not want the green energy option) increased adoption of this tariff nearly tenfold.

is much debate among economists on how efficient companies and markets are with respect to cost-minimization, there is actually little empirical work to understand whether behavioral science can improve the efficiency of production. There are two pieces of recent evidence to suggest that greater productive efficiency can improve energy efficiency, and that firms may not be on their production possibilities frontier. The first is research on attempting to understand management styles, and how different types of management can affect the amount of energy used. Bloom *et al*. (2010) estimate the management practices in over 300 manufacturing firms in the UK and match this management data to production and energy usage information for establishments owned by these firms. They find that better-managed firms are significantly less energy intensive – e.g., going from the 25th to the 75th percentile of management practices is associated with a 17.4% reduction in energy intensity. Interestingly, better-managed firms are also significantly more productive, so more productive firms tend to use less energy per unit of output.

Second, the recent work of Gosnell *et al*. (2016) shows that lowering energy use in the aviation sector can benefit workers, shareholders and the environment. They partnered with Virgin Atlantic Airways to construct a framed field experiment to understand how behavioral interventions can improve the fuel-efficient behaviors of their senior airline captains. They found that by simply telling the captains that they are being monitored on fuel use dramatically changed their fuel use. They also found that non-pecuniary targets and prosocial incentives (i.e., hit the target and money is donated to charity) increased the fuel-efficiency behaviors of their captains.[13] Interestingly, the marginal abatement cost for a ton of carbon dioxide estimated from this study was -$250 per ton – this estimate is currently the most cost-effective way to abate a ton of carbon dioxide. The experiment is the first known study to change airline and transport behaviors more generally in an experimental setting, and demonstrates the potential to improve energy efficiency within the commercial sector.

4. Conclusion

This paper reviewed a large number of experimental studies that rely on behavioral science to inform energy policy. The paper demonstrated that behavioral science is beginning to offer some useful insights on how to design better energy policies.

We have several suggestions for future research. One is to do more on assessment of welfare impacts, because we think that is important for policy makers to understand the efficiency implications of policy choices. Related to welfare, policy makers could benefit if they had more solid research on the cost-effectiveness of different kinds of programs, such as rebates. That way, they could select programs that provide the most “bang for the buck”.

A second research area is to learn more about what really drives consumer decisions. For example, when do consumers respond to average prices and when do they respond

[13] The interventions used were similar to those interventions advocated by Bloom and van Reenen (2007) to improve management practices.

to marginal prices? And if consumers do respond to average prices in some settings, are there ways to move their attention to marginal prices. A third area that could be useful for researchers and policy makers is to test the *interactions* of different kinds of policies. We conjecture based on initial research that such interactions may yield results that are not always consistent with simple economic models of behavior. As discussed above, the one case that was found that violated the law of demand involved a policy interaction. A fourth area where more work could be undertaken is with business customers. We do not have a good understanding, for example, of how businesses in supply chains respond to different kind of interventions. A fifth area of interest to policy makers would be to assess whether interventions have different impacts across different income groups, so that equity and efficiency considerations could be taken into account. Finally, we are only beginning to understand the potential for changing energy consumption through habit formation versus getting customers to adopt and use new technology.

Finally, we think it is useful to explore the full range of nudges (or interventions) to establish an optimal mix. We include here as a "nudge" anything that can change behavior in a desired direction. Thus, price changes would be included in this group. We know very little about the relative efficacy of different kinds of interventions, so we are a long way from determining the optimal mix.

A key motivation for doing behavioral science experimentation on energy use is to generate new insights for the purpose of changing policy. We think it would be useful to do research on the extent to which specific policies were changed as a result of the experiments that were done. Admittedly, this assessment is hard to do. We sent out a number of emails to the first authors of many of the studies cited in this paper; six responded, giving us information on 14 studies. Few were aware of whether their findings were actually implemented in some way. This may not be particularly surprising, given that academics do not typically get high payoffs for monitoring whether their work is actually implemented.[14] Still, we think studies that try to link changes in specific energy policies to experiments would be useful to get a better sense of the value of such research policy makers.

References

Allcott, H. (2011a), "Rethinking Real-time Electricity Pricing", *Resource and Energy Economics* 33(4): 820-842.

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

[14] The responding authors collectively knew of seven studies in which their findings were at least discussed for implementation by utilities and government agencies. Four studies' findings were discussed by utility companies, often the company that partnered on the study. Three were debated in public forums, both at state and national levels.

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

Allcott, H. and M. Greenstone (2012). “Is There an Energy Efficiency Gap?” *Journal of Economic Perspectives* 26(1): 3-28.

Allcott, H. and J. B. Kessler (2015). "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." NBER Working Paper 21671.

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

Allcott, H. and D. Taubinsky (2015). “Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market.” Working Paper, New York University, pp. 1-80.

Allcott, H. and R. Sweeney (2014). “Can Retailers Inform Consumers about Energy Costs? Evidence from a Field Experiment.” Working Paper, New York University, pp 1-52.

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

Ayres, I., S. Raseman, and A. Shih (2009). “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.

Battalio, R. C., J. H. Kagel, R. C. Winkler, and R. A. Winett (1979). "Residential Electricity Demand: An Experimental Study." *The Review of Economics and Statistics* 62(2): 180-89.

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

Bloom, N., C. Genakos, R. Martin, R. Sadun (2010). “Modern Management: Good for the Environment or Just Hot Air?” *Economic Journal*, 120, 544, 551-572.

Bloom, N. and J. Van Reenen (2007). “Measuring and explaining management practices across firms and countries.” Quarterly Journal of Economics 122 (4), 1351–1408.

Bollinger, B. and W.R. Hartman (2015). “Welfare Effects of Home Automation Technology with Dynamic Pricing.“ Working Paper.

Buchanan, K., R. Russo, and B. Anderson. (2015). “The question of energy reduction: The problem(s) with feedback.” *Energy Policy* 77: 89-96.

Caves, D., and L. Christensen (1980), “Econometric Analysis of Time-of-use Electricity Pricing Experiments.” *Journal of Econometrics*, 14(3): 287–306.

Caves, D., L. Christensen, and J. Herriges (1984), “Consistency of Residential Customer Response in Time-of-use Electricity Pricing Experiments”, *Journal of Econometrics 26(1-2)*: 179–203.

Chamberlin, E. (1948). "An Experimental Imperfect Market," *Journal of Political Economy* 56(2): 95–108.

Cialdini, R., and T. Trost (1998). Social influence: Social norms, conformity and compliance. In *The Handbook of Social Psychology* (Eds: Daniel Gilbert, Susan Fiske, Gardner Lindzey), pp. 151-192.

Costa, D. and M. Kahn (2013). “Energy Conservation “Nudges” and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment.” *Journal of European Economic Association* 11(3): 680-702.

Davis, L.W. and E. Muehlegger (2010). Do Americans consume too little natural gas? An empirical test of marginal cost pricing. RAND Journal of Economics, 41(4), 791-810.

DellaVigna, S. (2009). "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315-72.

Dolan, P. and R. Metcalfe (2013). “Neighbors, knowledge, and nuggets: Two natural field experiments on the role of incentives on energy conservation.” LSE CEP Discussion Paper No. 1222.

Dolan *et al.*, (2012). “Influencing behaviour: The mindspace way.” *Journal of Economic Psychology* 33(1): 264-277.

Ebeling, F., and S. Lotz (2015). Domestic uptake of green energy promoted by opt-out tariffs. *Nature Climate Change*, 5, 868-871.

Faruqui, A., and S. George (2005). “Quantifying Customer Response to Dynamic Pricing.” *The Electricity Journal* 18(4): 53–63.

Ferraro, Paul J. and Michael K. Price (2013). “Using Non-Pecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment,” *Review of Economics and Statistics* 95 (1): 64-73.

Fowlie, M., M. Greenstone, and C. Wolfram. (2015). “Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program.” E2e Working Paper 020.

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

Gans, W., A. Alberini, and A. Longo (2013). “Smart Meter Devices and the Effect of Feedback on Residential Electricity Consumption: Evidence from a Natural Experiment in Northern Ireland.” *Energy Economics 36*: 729-43.

Gerber, A., and T. Rogers (2009). “Descriptive Social Norms and Motivation to Vote: Everybody’s Voting and So Should You.” *Journal of Politics* 71(1): 178-191.

Gandhi, R., C. R. Knittel, P. Pedro, and C. Wolfram. (2016). Running Randomized Field Experiments for Energy Efficiency Programs: A Practitioner’s Guide, draft.

Gillingham, K., D. Rapson, and G. Wagner (2016). “The Rebound Effect and Energy Efficiency Policy.” *Review of Environmental Economics and Policy* 10 (1): 68-88.

Gillingham, K., M. Kotchen, D, Rapson, and G. Wagner (2013). “The Rebound Effect is Overplayed.” *Nature* 493: 475-476.

Gosnell, G.K., J.A. List, R.D. Metcalfe (2016). “A New Approach to an Age-Old Problem: Solving Externalities by Incenting Workers Directly.” Working paper.

Hallsworth, M., J. List, R. Metcalfe, I. Vlaev (2014). “The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance.” NBER Working Paper No. 20007.

Harrison, G.W. and J.A. List (2004). “Field Experiments.” *Journal of Economic Literature* 42(4): 1009–1055

Herberich, D.H., M. Price, and J.A. List (2011). “How Many Economists Does it Take to Change a Light Bulb? A Natural Field Experiment on Technology Adoption.” Working Paper.

Ito, K., T. Ida, and M. Tanaka (2015). "The Persistence of Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *National Bureau of Economic Research*. Working Paper, Kyoto University, pp. 1-43.

Ivanov, C., L. Getachew, S. Fenrick, and B. Vittetoe (2013). "Enabling Technologies and Energy Savings: The Case of EnergyWise Smart Meter Pilot of Connexus Energy." *Utilities Policy* 26: 76–84.

Jessoe, K., and D. Rapson (2014). "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review* 104(4): 1-42.

Kahn, M. E. and F. Wolak (2013). “Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment.” Working Paper, Stanford University.

Kahneman, D. and A. Tversky (1979) “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47(2): 263-291.

Kluger, A., and A. DeNisi (1996). “The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory.” *Psychological Bulletin* 119(2): 254-284.

Knittel, C. (2013). The Energy-Policy Efficiency Gap: Was There Ever Support for Gasoline Taxes? NBER Working Paper No. 18685.

Laibson, D. (1997) “Golden Eggs and Hyperbolic Discounting.” *Quarterly Journal of Economics*, 112 (2), 443-478.

LaRiviere, J., S. Holladay, D. Novgorodsky, and M. K. Price (2014) “Prices vs Nudges: A Large Field Experiment on Energy Efficiency Fixed Cost Investments.” Working Paper, University of Tennessee, pp.1-48.

List, J.A. and R. Metcalfe (2014). “Introduction to Field Experiments.” *Oxford Review of Economic Policy* 30(4): 585-596.

Pellerano, J., M. K. Price, S. L. Puller, and G. E. Sanchez (2015). "Do Extrinsic Incentives Undermine Social Norms? Evidence from a Field Experiment in Energy Conservation." Working Paper, Universidad Iberoamericana, pp 1-24.

Price, M. K. (2014). “Using Field Experiments to Address Environmental Externalities and Resource Scarcity: Major Lessons Learned and New Directions for Future Research.” *Oxford Review of Economic Policy* 30(4): 621–638.

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

Schwartz, D., B. Fischhoff, T. Krishnamurti, and F. Sowell (2013). “The Hawthorne effect and energy awareness.” Proc. Natl. Acad. Sci.110(38):15242–15246.

Sherif, M. (1936). *The psychology of social norms.* New York: Harper.

Smith, V, (1976). "Experimental Economics: Induced Value Theory." *American Economic Review* 66(2): 274-79.

Sunstein, C. (2016). “People Prefer System 2 Nudges (Kind Of).” *Duke Law Journal* 66*.*

Thaler, R. and C. Sunstein (2008) *Nudge*. New Haven, CT: Yale University Press.

Wolak, F. (2006). “Residential Customer Response to Real-time Pricing: The Anaheim Critical Peak Pricing Experiment.” *Center for the Study of Energy Markets*. UC Berkeley.