Nudges in the marketplace: Using peer comparisons and incentives to reduce household electricity consumption

By Anant Sudarshan*

Behavioral interventions are increasingly used to change consumer behavior. This paper analyzes an electricity demand management experiment in urban India to compare the effectiveness of three instruments designed to reduce electricity consumption: (i) behavioral nudges using peer comparisons; (ii) nudges augmented with financial incentives and (iii) price changes. Peer comparisons are found to reduce household electricity consumption by over 8 percent. However they become ineffective when electricity prices are higher and when paired with additional financial incentives. Price increases remain effective in reducing consumption. These results suggest that behavioral instruments may interact in complicated ways with incentives and market prices.

JEL: D03, Q41, Q50

Keywords: electricity demand management, social norms, incentives, price elasticity

Electricity utilities around the world are designing and implementing demand side management programs to reduce electricity consumption. In the United States many

* Harvard Kennedy School and IFMR, anant_sudarshan@hks.harvard.edu. I thank the Sustainability Science Program (Harvard University) for financial support and Karthik Dinne and Nazuk Kumar for field assistance. I also thank Rohini Pande, Rohini Somanathan, Nava Ashraf as well as various seminar participants for many helpful comments. All errors are my own.

such programs have been explicitly linked to efforts to increase energy efficiency[1]. Demand reduction efforts may also be motivated by overconsumption concerns due to unpriced global and local environmental externalities (see (Muller et al., 2011) for a quantification of such externalities). Beyond interventions that aim to reduce total energy consumption, utilities have also introduced programs to reduce peak hour consumption through dynamic and real time pricing (Allcott, 2011b). Such programs seek to minimize the inefficiencies introduced by using static electricity prices when the cost of supply varies with the time of day.

Although much of the published literature on electricity demand management focuses on the United States (Sudarshan, 2013, Wolak, 2011, Arimura et al., 2009, Allcott, 2011a), nevertheless the rationale for such programs is arguably even stronger in developing countries such as India where several additional factors motivate efforts to change consumption patterns. These include distortionary tariff cross-subsidies (Singh, 2006) and frequent blackouts owing to crippling supply shortfalls[2]. The policy importance of demand side management in India has therefore been growing, underscored by the passing of an ambitious Energy Conservation Act in 2001 and the initiation of a National Mission for Enhanced Energy Efficiency by India's Ministry of Power in 2010-11. In its most recent five year plan, India's Planning Commission proposed allocating over 50 million dollars to enhance the capacity of state utilities to implement demand side management programs.

Given the objective of reducing electricity consumption, an immediate challenge involves identifying which instruments are best suited to achieve this goal? The most

[1]Gillingham et al. (2006) describe some of these interventions. Jaffe and Stavins (1994) and Allcott and Greenstone (2013) bookend over 20 years of debate on the 'Energy Efficiency Gap' hypothesis which holds that consumers make sub-optimally low investments in energy efficiency.

[2]Peak demand in India exceeded supply by 18% in 1996, 13% in 2002 and 13% in 2011 leading to widespread rationing and outages. A striking reminder of this problem arrived on July 30-31, 2012, when over 600 million people in North India simultaneously found themselves without electricity.

straightforward means to modify consumption would appear to involve changing the prices consumers pay for the energy they use (including the introduction of time of day pricing regimes). Yet while changing electricity tariffs is an economically attractive solution on paper, it is often neither politically nor administratively easy to implement. In practice therefore, a wide range of more indirect tools have been used all over the world. These include conditional financial incentives, short term and targeted subsidies, behavioral techniques, outreach and education programs, appliance subsidies and appliance exchange schemes (see Gillingham et al. (2006) for a review of examples from the United States).

Of these different instruments, two that are increasingly advocated (including in contexts other than electricity demand) are (i) behavioral interventions and (ii) conditional financial incentives. In particular, behavioral interventions such as peer group comparisons have attracted significant academic and policy interest (Thaler and Sunstein, 2008, Allcott and Mullainathan, 2010). Part of the attraction of behavioral techniques is that they are often relatively inexpensive to implement since they typically rely only on providing carefully framed information to consumers designed to 'nudge' them towards desired behaviors.

Yet notwithstanding the enthusiasm around 'nudges', we know relatively little about the mechanisms through which such instruments influence behavior nor how reliably useful they can be to policy makers (especially compared to traditional approaches such as raising prices)[3]. For instance, most of the evidence on the effectiveness of peer comparisons in changing electricity consumption behaviors comes from a set of programs implemented in the United States by the company O-Power (Allcott, 2011a). Less is known about consumer responses in different cultural settings,

[3]These concerns also underpin the response by Stern et al. (2010) to an influential article (Allcott and Mullainathan, 2010) recommending the use of behavioral nudges as tools to reduce household energy consumption.

an omission that may be important given the role of social norms in this context (Cialdini, 2007, Schultz et al., 2007).

Similarly while conditional financial incentives are widely used in electricity demand management, evidence on their effectiveness is mixed (Ito, 2013) . We also know little about how behavioral interventions and financial incentives interact and whether or not they might work to complement each other? This seems important because in practice consumers do find themselves simultaneously responding to a variety of programs and incentives.

This paper provides insight into these questions through an experimental evaluation of a 'real-world' electricity demand management program implemented in a community of urban middle class households in India. I first exploit a unique natural experiment providing quasi-experimental variation in tariffs to estimate short run price elasticities for households. I thus establish that households do adjust consumption in response to tariff changes although the price elasticity is relatively low (-0.13). This response to electricity prices provides a benchmark against which to compare the effectiveness of other instruments.

Next, I utilize the experimental design of the demand management program and find that providing households information on the average electricity consumption of their peers relative to their own (a behavioral nudge) resulted in an approximately 11 percent reduction in electricity consumption averaged over the entire summer season. This suggests that replicating the mean effect of the nudge through tariff changes alone would require an approximately 65 percent increase in the price. This result provides evidence underscoring the potential of social comparisons as tools to change behavior and suggests that household responses to social comparisons of electricity use, as observed in the United States (Allcott, 2011a), can be replicated in an entirely different cultural and economic setting.

At the same time I also find evidence suggesting that this potential may have important limitations. I exploit the mechanics of how electricity is supplied in the experimental setting[4] in order to examine the effectiveness of the behavioral nudge on electricity consumption both when the underlying commodity (electricity) is expensive, as well as when it relatively cheap. I find that households respond to peer comparisons only when electricity is priced low. For electricity that is priced high the 'nudge' no longer influences behavior.

Similarly, when peer comparisons are 'strengthened' by coupling them with conditional financial incentives that are equivalent to an increase in the marginal price, electricity consumption *increases* relative to households relative to the 'pure' behavioral intervention. These results suggest that monetary incentives may crowd out instead of complementing a non-monetary behavioral intervention. Put another way, in the presence of small financial incentives, nudges appear to become ineffective.

These results provide perhaps the first evidence from a directly policy relevant population[5], on the effectiveness of peer comparisons in the presence of other incentives. The experiment confirms the potential of nudges but also underscores the need for caution in using these instruments to change electricity consumption behaviors. While households do respond to peer comparisons, the effectiveness of this instrument seems to depend strongly on what else is happening in the 'marketplace'. The outcomes I observe are consistent with a model of behavior where people are influenced by behavioral cues only when the economic stakes are low and revert to resembling classical rational agents when the stakes are raised. This suggests that the choice of behavioral tools to manage electricity demand needs to be made with

[4]Owing to the frequency of grid supply shortages, households use two separate supply sources that are separately metered and priced differently. See II for details.

[5]The residential sector is the fastest growing electricity sector in India (CEA, 2010) driven largely by the consumption of a growing middle class.

care and that responses to non-monetary interventions may be negatively affected by traditional monetary incentives.

The remainder of this paper is structured as follows. In Section I I locate this study within the context of existing theory and laboratory evidence from the psychology and behavioral economics literature. Next, in Section II I describe the design of the field experiment I study. In Section III I review empirical results and I conclude in Section IV.

I. Theory

Laboratory and field experiments have shown that people may modify behavior when exposed to comparisons with their peers (the definition of a peer varying by context). Applications have ranged from charitable giving (Frey and Meier, 2004, Alpizar et al., 2008) to modifying retirement savings behavior (Beshears et al., 2011). In the context of electricity consumption, large randomized trials have been implemented in the United States by the company O-Power, using peer comparisons to try and reduce household electricity consumption (Allcott, 2011a). Unfortunately little evidence exists on whether peer comparisons can be effective in changing energy behaviors in developing country contexts where incomes are lower and cultural norms very different. One contribution of this paper is to provide such evidence.

We also know relatively little about the mechanisms through which peer comparisons change energy behaviors. One explanation originating from the psychology literature is that people seek to adhere to 'social norms', which are prescriptions of behavior regarded as socially desirable or normal. On this basis, when households are informed that they are consuming electricity at a level very different from their peers, they incur a psychological cost. In order to minimize this cost, they shift behavior (Cialdini, 2007, Schultz et al., 2007).

Interestingly, social norm mechanisms do not rule out the possibility that low consuming households might increase their usage to approach the group average. Whether this happens in practice depends on whether the underlying socially desirable behavior involves consuming similarly to peers (so called 'descriptive' norms) or consuming less than the average ('injunctive' norms). In the electricity setting, previous empirical evidence indicates that peer comparisons do lead to reductions at all levels of consumption(Allcott, 2011a) and this has contributed to nudges being advocated as a useful tool.

Alternative mechanisms have also been proposed including the idea that peer comparisons change behavior because of straightforward learning effects(Conley and Udry, 2010, Cai et al., 2009). Consider a simple model where households attempt to produce an optimal amount of energy services, but are uncertain about parameters of the household production function. They may choose to invest resources in learning but the value of such information is bounded and therefore some uncertainty is likely to remain in the status quo. Provided households regard their peers as having production function parameters drawn from the same family, information about the average behavior of the population can be used to learn about the true value of one's own type[6]. In turn, this can enable households to re-optimize and over time learn their true production function.

A. *The Behavioral Effects of Monetary Contracts*

Peer comparisons represent a tool to change behavior that relies on information provision alone. One might reasonably ask whether it is possible to augment such nudges by adding financial incentives?

[6]As a practical example, households might learn about the unobservable efficiencies of their own appliances by comparing their electricity use to the peer average

The answer to this question likely depends on the mechanisms through which peer comparisons influence behavior. If psychological costs (such as shame or guilt) drive the effectiveness of behavioral nudges then observed outcomes might change in unpredictable ways depending on whether the commodity or behavior in question is available at no cost, priced low or priced high. For example any psychological guilt households feel about consuming more than their peers may be diminished if they are aware that they have already paid heavily for having done so. On the other hand, if households use peer comparisons to learn new information, changes in external market prices should not change the fact that once a household is better informed, it will consume differently.

Although the interaction of behavioral and financial instruments has not been directly explored, evidence from laboratory experiments leaves open the possibility that financial incentives and prices can affect behavior in unpredictable ways. In particular, studies have shown (Heyman and Ariely, 2004, Ariely et al., 2009b) that people respond quite differently to monetary markets and contracts versus non-monetary or social contracts. These differences can result in reductions in effort in response to financial incentives especially where these incentives are relatively low powered (Kamenica, 2012, Gneezy and Rustichini, 2000). Different mechanisms have been proposed as explanations. These include theories of crowding out of so-called 'intrinsic motivation' by extrinsic rewards (Sansone and Harackiewicz, 2000, Frey and Jegen, 2002) as well as more formal models relying on information asymmetries and the expectations of contracted agents (Benabou and Tirole, 2003), or on reputation effects (Ariely et al., 2009a, Benabou and Tirole, 2006).

The relevance of this evidence to the use of nudges in policy settings is more difficult to assess. Much of what we know about the behavioral effects of monetary incentives comes from lab experiments or framed field experiments (to use the taxonomy of Harrison and List (2004))[7]. There remain legitimate concerns with extrapolating from lab experiments conducted largely with developed country participants to settings in other cultures (Henrich et al., 2010). Evidence also suggests that short run effects in such settings may disappear over even slightly longer time periods (Gneezy, 2006). Therefore measuring outcomes in real-world program contexts seems essential.

A key contribution of this paper is to provide cleanly measured field evidence over an extended period of time. I study a 'natural field experiment' (Harrison and List, 2004), involving an actual conservation program implemented by a private estate management firm (see Section II). This minimizes the risk that household responses are driven by experimenter influence or that measured impacts are short run effects. This study also deals explicitly with the question of how different monetary and non-monetary instruments interact with each other.

B. Synthesis

Before describing the experiment in detail, it is helpful to pull together what we have discussed into a single framework. Let household demand for electricity Y_e be represented as

$$Y_e = \beta_t Z - \theta_p \cdot p_e + \epsilon \qquad (1)$$

where p_e is the price of electricity, $\theta_p \in R^+$ is an elasticity parameter associated with price and $\beta_t Z$ represents other parameters where β_t is indexed by t to denote a dependence of demand function parameters on a households belief about its own

[7]The field studies that do exist have largely focused on the special case of pro-social behavior with reputation benefits, such as charitable giving (Gneezy and Rustichini, 2000, Ariely et al., 2009a).

type t.

I now let peer comparisons influence utility and thence demand through two channels. First, I assume that households have imperfect knowledge about their type and can learn about their type t based on information about their electricity consumption Y_e compared to the peer average $\bar{Y}_e$. I capture this idea by letting

$$t = \begin{cases} t_o & \text{when } \bar{Y}_e \text{ is not observed} \\ t(\bar{Y}_e) & \text{when } \bar{Y}_e \text{ is observed} \end{cases}$$

The second mechanism through which peer comparisons might matter is via psychological costs faced by the households when violating a social norm. To model this I introduce a psychological cost term $P_S(\bar{Y}_e, p_e)$ which depends on both the average consumption of peers, $\bar{Y}_e$ as well as the price of electricity p_e. This assumption then suggests testing for the presence of psychological costs by observing whether the impact of peer comparisons depends on the price p_e.

Thus the demand function for a household subjected only to peer comparisons and market prices might be written as

$$Y_e = \beta_{t(\bar{Y}_e)} Z - \theta_p \cdot p_e - \theta_{ps} \cdot P_S(\bar{x_e}, p_e) + \epsilon \tag{2}$$

Now assume that peer comparisons are augmented with additional financial incentives rewarding the same behavior. In the field experiment I study, the incentives offered are effectively an increase in the marginal price of electricity (see Section II for a description). Thus to describe the impact of augmenting peer comparisons with financial rewards r, we can rewrite the electricity price p_e as $p_e = p_m + r$. That is, the net price p_e is the sum of the market price p_m and a financial incentive per unit consumption r.

The demand for electricity can now be written

$$Y_e = \beta_{t(\bar{Y}_e)} Z - \theta_p \cdot (p_m + r) - \theta_{ps} \cdot P_S(\bar{Y}_e, p_m + r) + \epsilon \tag{3}$$

The incentive r thus influences demand through two channels. The first is the direct, conventional impact of a changing marginal price. The second is the impact of changing economic stakes on the effectiveness of peer comparisons through the channel of psychological costs $P_S(\bar{Y}_e, p_e)$[8]. The functional forms here are of course only illustrative but Equation 3 helps clarify how different instruments might change electricity consumption. In Section III I focus attention on the four questions below to which I seek to find empirical answers to complement model predictions.

1. Do households respond to changes in market price?

$\frac{\partial Y_e}{\partial p_e} = -\theta_p < 0 \implies$ Consumption decreases as market price increases.

2. Does consumption change when provided peer comparisons (P)?

$\Delta Y_e = Y_e|P - Y_e|!P = [\beta_{t(\bar{Y}_e)} - \beta_{t_0}]Z - \theta_{ps} \cdot P_S(\bar{Y}_e, p_e) \implies$ Direction of change depends on net effect of psychological costs and learning. Prior empirical evidence is consistent with $P_S \geq 0$ and $\Delta x_e < 0$.

3. Do market prices change the effectiveness of peer comparisons?

$\frac{\partial \Delta Y_e}{\partial p_e} = -\theta_{ps} \frac{\partial P_S(\bar{Y}_e, p_e)}{\partial p_e} \implies$ The market price may modify the effect of peer comparisons if psychological costs are present. Additionally if the following hold true

1) $P_S > 0$ (psychological costs are positive therefore reducing consumption)

[8] In this formulation the impact on psychological costs due to an independently offered incentive is the same as due to an equivalent change in price. This assumption is not strictly necessary but simplifies exposition.

2) $\frac{\partial P_S}{\partial p_e} < 0$ (psychological costs reduce as economic stakes increase)

Then we would expect $\frac{\partial \Delta Y_e}{\partial p_e} \geq 0$. That is nudges may become less effective as the price increases.

4. What is the effect of adding incentives (I) to peer comparisons (P)?

$$Y_e|P, C - Y_e|P, !C = \underbrace{-\theta_p \cdot r}_{\text{price effect}} - \underbrace{\theta_{ps}[P_S(\bar{Y}_e, p_m + r) - P_S(\bar{Y}_e, p_m)]}_{\text{behavioral effect}}$$

Under the assumptions in 3 above these two terms act in opposite directions so that

1) When r is large, the price effect dominates and consumption decreases.

2) If r is small, the behavioral effect may dominate and consumption may increase

II. Experimental Context and Design

Much of the growth in residential electricity demand in India has come from the boom in new urban construction[9]. The capital city of New Delhi is a good example of this nationwide trend. The city and its surrounding suburbs form the second largest urban agglomeration in the world. Over 22 million people live in the so called National Capital Region (Delhi and satellite cities) a figure that is expected to rise to 32 million by 2025 (UN, 2012).

The dominant form of new middle-class housing in the National Capital Region (NCR) has been gated apartment complexes constructed by private real estate firms. These developments range from a few hundred to a couple of thousand living units in size. Individual units are sold or rented to families. A lack of good public infrastructure, police and transport services and the fact that many of these areas have

[9]The McKinsey Global Institute (MGI, 2010) estimates that the urban population of India is likely to rise by over 200 million over the next 20 years.

relatively high crime rates[10], make this type of gated community a popular choice for both real estate developers and consumers. Within such apartment complexes, internal security and various services can be privately arranged.

The experiment described here involved households located in just such an apartment complex in one of the urban satellite cities of Delhi (Indirapuram in Ghaziabad district)[11]. The development consisted of over 700 separate housing units. Four types of living units were available - 2 bedroom units, 3 bedroom units, 4 bedroom units and a few larger penthouse apartments. The majority were two or three bedroom designs and the experiment targeted this subset of homes. Notwithstanding the absence of heterogeneity in climate and housing stock, significant variance existed in baseline electricity consumption levels (1), underlining the importance of individual behavior in determining electricity consumption.

A. *Household Electricity Supply*

The mechanics of how electricity is supplied to many new apartments in the NCR provides a unique opportunity to investigate the questions motivating this study. In particular, the billing and metering of electricity in these developments has been increasingly left to private apartment management companies. This is partly due to the inability of capacity constrained state electricity utilities to keep up with a rapidly growing consumer base.

Thus for the residential community I study here, the estate management provided metered electricity at the household level sourced from *two* supply sources. The first

[10]In a June 2011 article, The New York Times (Jim Yardley) wrote about the city of Gurgaon in the NCR: "Gurgaon... would seem to have everything except... a functioning citywide sewer or drainage system; reliable electricity or water; and public sidewalks, adequate parking, decent roads or any citywide system of public transportation."

[11]On May 2 2012 the leading business daily in India, The Economic Times, reported that over 500,000 housing units were under construction in the NCR alone in May 2012. A third of these were located in Ghaziabad district.

was regular grid power supplied by the state utility to the estate management at the public tariff (with internal metering and payments left to the private agency). The second source was via captive diesel power generating units owned and operating by the estate management company. The use of captive power has become increasingly common - a consequence of crippling grid electricity shortages and regular blackouts which have together created a robust market for diesel generation across India[12]. Regular grid outages tend to be unscheduled and occur at different times and for different durations. Figure 1 shows the distribution of the number of hours in a day that households in our study had no supply of grid electricity.

The four on-site captive power units were used to provide a complete backup to apartments when grid power was not available[13]. The self generated power thus provided a perfect substitute for grid electricity, with the single exception of the tariff. The two power sources were separately metered at the household level and billed using two separate per unit rates with no non-linear pricing. Backup diesel power was about 4 times as expensive as grid power (INR 12.10/KWh versus INR 3.2/KWh where INR 10 was approximately equal to 18 cents at the exchange rate during the study period in May-Aug 2012). Because the price difference was significant, households were provided with a red warning light which turned on when backup power was being used[14]. These tariffs were determined at the start of the year based on the utility price of power and diesel generation costs. In other words, thanks to captive power and poor grid supply, the households I study were paying for electricity at two very different prices with unscheduled (but observed) switching between rates.

[12]Blackouts are much more frequent in satellite cities of the capital city than in Delhi itself and are exacerbated during the intense summer season

[13]Load rationing under captive power was not a concern, since the diesel generators produced sufficient power to allow all households to run multiple two tonne split airconditioning units.

[14]This warning may have been somewhat redundant since when the large captive diesel units were running they were also audible.

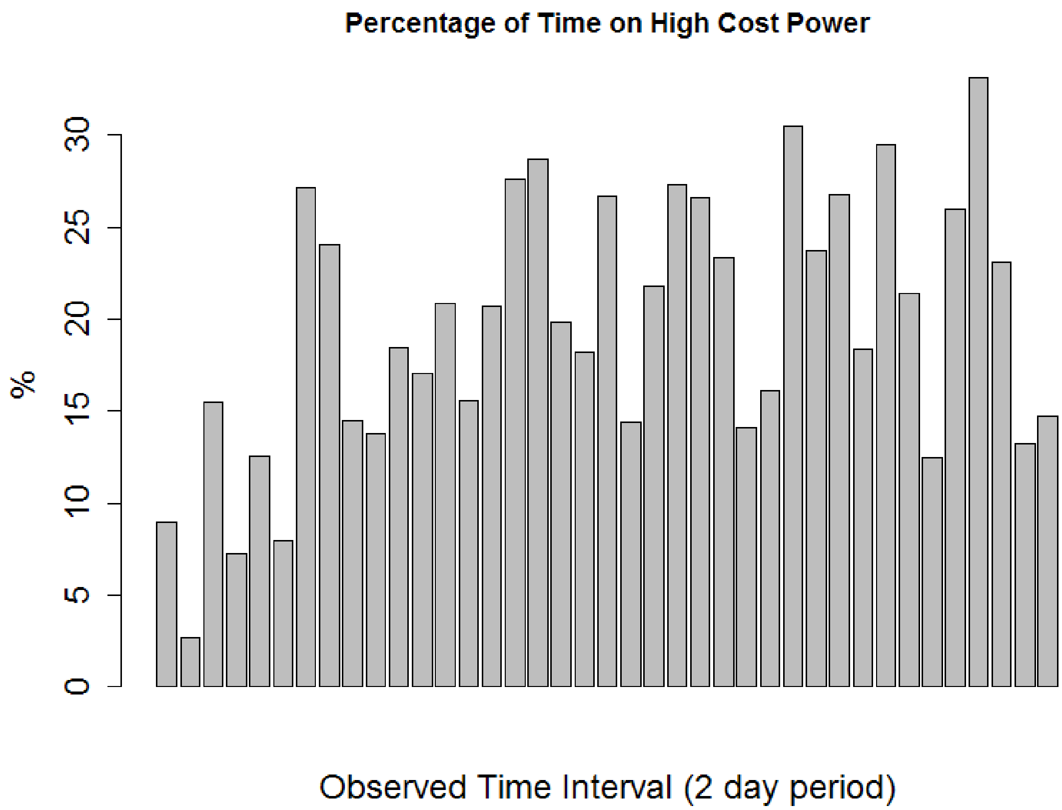


FIGURE 1. PERCENTAGE OF TIME WITHOUT GRID ELECTRICITY (MAY TO AUGUST). EACH BAR COVERS AN APPROXIMATELY TWO DAY INTERVAL OVER EACH OUTAGES ARE AGGREGATED.

B. *Experimental Intervention*

From the point of view of the estate management, the cost of supplying grid power was predictable because it was billed to them at a fixed public tariff rate. However captive power was generated using diesel fuel purchases. Since changing tariffs frequently was not feasible under the terms of operation the private company was directly exposed to shocks in supply costs owing to the increasing market price of diesel. This formed part of the motivation for exploring ways to reduce household electricity consumption. Two experimental interventions were therefore piloted by the estate management company and I analyze these in this paper.

A subset of households in the residential community were selected for the demand management pilot. All occupied two and three bedroom apartments were intended to be chosen. However an administrative error in drawing up occupancy lists meant that of a total of 534 occupied units only 484 were initially identified as potentially eligible for the program. These 484 units were then randomly assigned to two different program conditions (treatments). 124 households were placed in what I refer to as the 'Nudge Treatment' (denoted by N). 240 households were placed in what I will refer to as the 'Nudge+Incentives' Treatment (denoted as NI). 124 homes formed controls (denoted as C). Randomization was stratified by household size (two or three bedrooms).

The remaining 50 homes that were not part of random assignment also did not receive any treatment, just like the control. I refer to these homes as the 'Default Control' (denoted as DC). Although I have no reason to believe these subjects differ from other households in the sample (see also Table 1), I do not include them in the main evaluation of treatment effects. However they are used as a robustness check.

Time of day consumption data was not available at any point during the experiment and our consumption measure consists of periodic measures of aggregate household consumption over 2 or 3 day periods from April through August obtained by periodically querying meter readings thrice a week. Thus the definition of a 'time period' in the analysis of this paper refers always to a 2 or 3 day interval.

The treatment and control groups are summarized below. The pilot as a whole began at the start of May 2012 and continued through the end of August, 2012 (4 months covering the summer season). Baseline consumption data was collected from April 10, 2012 through till the beginning of May.

Group C (Controls). Control households were never contacted but their electricity consumption was recorded over the course of the project.

Group DC (Default Controls). Households that were not part of the randomization pool owing to an administrative error and were therefore assigned to no treatment by default.

Group N (Nudge Treatment). This group was sent a weekly report card delivered by the estate management to their apartment mailboxes. The report card detailed their own electricity consumption for the past week for both grid electricity (low priced) and diesel backup (high priced) power. This was compared to the average consumption in other households with the same number of bedrooms[15].The weekly report card also contained a general set of tips on how to save energy on the back and an injunction to conserve on the front. The appendix contains a sample of the front and back of the complete letter sent to these households and Figure 2 extracts part of the letter, namely the peer comparison and surrounding explanatory text.

Group NI (Nudge+Incentives). This group of households was sent a report card similar to Group N detailing their electricity consumption and the average in other similar households. In addition however these households were also enrolled in a reward scheme which functioned as follows. Every household in the incentive group was provided a starting reward balance of 750 Indian Rupees (about 13 US dollars) funded by the billing agency (later reimbursed to the company by the researchers). Thereafter, for the duration of the program, this reward balance could increase (or decrease) depending on the difference between household electricity consumption and

[15]The reported averages were based only on similarly sized households in the same treatment condition although the reports themselves did not indicate that other treatments existed

the peer average. When consuming less (consuming more) than the group average, the reward balance was increased (reduced) at the rate of INR 2.00 per unit for grid electricity (where the per unit tariff is INR 3.2) and INR 4.00 per unit for diesel electricity (where the per unit tariff is INR 12.10). In other words the information in the peer comparison was monetized with an additional financial incentive to reduce consumption equal to the per unit reward rates.

As an example, if a household were to consume 10 units of grid electricity and the average consumption that week were 20 units, a reward $r = 2 \cdot (20 - 10) = 20$ would accrue. Households could not lose more than their starting reward balance over the course of the experiment, meaning that by the end of the program a fraction of homes had entered a region of zero marginal financial incentives (while still receiving the same information and still remaining eligible for positive rewards). This corner case excepted, this incentive design is thus theoretically equivalent to an increase in the marginal price of electricity by the per unit reward rate[16]. The appendix contains a sample of the front and back of the complete letter sent to these households and Figure 3 extracts part of the letter, namely the peer comparison and surrounding explanatory text.

A useful property of the financial incentive was that rewards and penalties were linked to exactly the same comparisons as were provided to the Nudge Treatment. Therefore no additional reference points or new information, that might complicate the interpretation of results, were introduced here.

Although these incentives are significant as a fraction of the per unit price, nevertheless actual transfers were quite small even taken over the course of the entire summer. Ignoring households who lost some or all of their initial reward allocation, the average transfer to net gainers was only 8 USD per month. The 90th percentile

[16]A similar incentive is analyzed in Falkinger (2000)

transfer was about 15 dollars a month. Therefore this treatment probably does not create high powered financial incentives. This partly reflects the fact that the incentive applied only to the difference between household consumption and the peer average but is largely a consequence of electricity expenditures being fairly low relative to incomes and other expenditures (typical rental payments for households over this period were of the order of 500 dollars per month). This weakness is unfortunately quite common since the economics of efficiency and demand management incentives mean that financial rewards are often relatively low powered - a general challenge for program implementation in this setting[17].

The assignment to each of the treatments was optional with households able to drop out at any time over the period of the study. A negligible amount of attrition occurred. Only three households chose to opt out (in the first two weeks) and no others thereafter. A few other homes were dropped from the pilot even before delivering the first report because baseline electricity consumption data indicated they were not occupying their apartments although they had administratively taken possession. Eventually 119 (out of 124) households remained in the information treatment, 233 (out of 240) in the incentives treatment and 121 (out of 124) households were controls.

Table 1 provides a summary of t-tests comparing baseline electricity consumption measured in treatment and control groups. A baseline survey was not carried out to minimize the possibility of contamination due to experimenter effects and to stay as close to possible to a natural field experiment setting. However an end-line survey was carried out on a subsample of homes distributed equally across each experimental condition. Table 1 therefore also includes a comparison of certain observable charac-

[17]See Du Pont (1998) for a discussion of how low financial benefits may influence the design of efficiency labels.

teristics measured during the end-line survey that could be expected to remain constant over the treatment (such as household composition and expensive appliances). These comparisons serve to confirm that the randomization across treatments was properly carried out. I also compare daily electricity consumption (averaged over the experiment duration) at low prices (grid) and high prices (diesel backup) for the randomized experimental control and the additional 50 default control households (Group DC) and confirm that the two are are similar.

III. Results

As the description in Section II makes clear, the households I study responded to three different types of instruments. These are (i) tariff changes, (ii) behavioral nudges, (iii) nudges with conditional incentives. In this section I analyze electricity consumption data and household response with a view to understanding how these different instruments work in the field. Specifically I consider the four questions posed at the end of (Section I). I begin by estimating the response of households to changes in the electricity price.

A. Do households respond to changes in market price?

The price elasticity estimates that seem most relevant to this setting (to the best of my knowledge) are the summer season price elasticities for urban Indian households reported in Filipini and Pachauri (2002) (elasticity = -0.16, based on a cross-section study using nationwide sample surveys).

Nevertheless it is hard to form strong priors regarding price response from the existing literature. A challenge in estimating electricity price elasticities cleanly is that price variation is rarely random. Thus most elasticity estimates in the literature tend to rely either on structural models of behavior (Reiss and White, 2005)

		Control (C)	Treatment (N)	Treatment (NI)	$p_{C,N}$	$p_{N,NI}$
1	Baseline KWh/day (Grid)	0.45 (0.29)	0.49 (0.26)	0.47 (0.33)	0.36	0.60
2	Baseline KWh/day (Diesel)	0.50 (0.42)	0.52 (0.34)	0.49 (0.41)	0.62	0.40
3	HH Members (0-15)	1.26 (0.60)	1.49 (0.68)	1.55 (0.61)	1.00	1.00
4	HH Members (15-60)	2.47 (0.94)	2.41 (0.89)	2.48 (0.78)	1.00	1.00
5	HH Members (60-100)	1.48 (0.73)	1.20 (0.77)	1.52 (0.60)	1.00	1.00
6	LCD TV	1.23 (0.88)	1.15 (0.87)	1.10 (0.70)	1.00	1.00
7	CRT TV	0.31 (0.52)	0.30 (0.49)	0.25 (0.46)	1.00	1.00
8	All AC	2.48 (1.11)	2.42 (0.87)	2.31 (0.99)	1.00	1.00
9	Refrigerators	0.97 (0.24)	0.99 (0.27)	0.90 (0.30)	1.00	0.85
10	Coolers	0.08 (0.28)	0.06 (0.23)	0.10 (0.34)	1.00	1.00
11	Washing Machine	0.87 (0.38)	0.90 (0.30)	0.88 (0.33)	1.00	1.00
12	Microwave	0.83 (0.48)	0.82 (0.39)	0.80 (0.40)	1.00	1.00
				Control (C)	(DC)	p
13	Mean KWh/day (Grid)			15.42 (7.81)	15.60 (8.37)	0.88
14	Mean KWh/day (Diesel)			3.38 (1.78)	3.57 (2.04)	0.47

Notes:
1. *p<0.1; **p<0.05; ***p<0.01
2. Cluster robust standard errors (Arellano-Bond)
3. $p_{C,N}$, $p_{N,NI}$ are p values for 'control vs. nudge' and 'nudge only vs. nudge+incentives' respectively.
4. Rows 3-5 and 6-12 compare household composition and appliance ownership from end-line survey.
5. Rows 1-2 compare baseline (pre-treatment) electricity consumption for grid and diesel power.
6. Rows 13-14 compare electricity consumption of untreated groups (C and DC) over experiment duration.
7. Rows 1, 2, 13, 14 based on all households.
8. Rows 3-12 based on a randomly surveyed sub-sample (70 per condition).

Table 1—Randomization check of control and treatment groups.

(which may not always describe behavior very well (Ito, 2012)) or on cross-sectional comparisons which require the often unrealistic assumption of exogenous price variation and tend to average over a great deal of heterogeneity. These challenges are compounded by the absence of price elasticity estimates specifically applicable to middle-class consumers in India.

In order to quantify the responsiveness of consumption to price for households in the experiment I proceed as follows. Consider a reduced form equation relating consumption to price,

$$Y_{i,T} = 1 + \gamma_T + \delta_i + \theta P + \epsilon_{i,T} \tag{4}$$

where T represents a period of time over which we observe cumulative electricity consumption $Y_{i,T}$ for household i. γ_T, δ_i are fixed effects for every time period and every household controlling for household specific heterogeneity and time shocks to consumption (including temperature, rainfall etc) that affect the experimental population as a whole. P represents the number of hours (or minutes) in period T that high priced power was being consumed. θ is the change in electricity consumption for every additional hour paying the high tariff. A random source of variation in P (hours on high priced electricity) would allow us to estimate θ without bias and thence the price elasticity.

At first glance, it is hard to see why any random variation in P should exist. Using day to day temporal variation in P is likely to be misleading because while load-shedding is unscheduled, it is not uncorrelated with other temporal determinants of consumption. For example P might rise on hot days which would bias elasticity toward zero. Similarly comparing the hourly intensity of electricity consumption

(KW) during low and high priced supply periods within households may also be biased upwards because outages are likely less frequent at night and thus intra-day hours spent using the two supply sources are not perfectly comparable.

Thus in order to estimate price elasticities specific to the population I study, I exploit a unique natural experiment that induces a third source of variation in P that is plausibly random.

Recall that households in the experiment were subject to frequent and unscheduled blackouts (see Figure 1) and during periods of time when grid power is unavailable, households consumed higher priced diesel generated backup electricity. Since metering for the two electricity sources was separated, it was possible to observe consumption separately for both types of power. Grid power was billed at a lower per unit rate (3.2 INR/KWh) and diesel based captive power at a higher price (12.10 INR/KWh).

Power blackouts in the area occur across the hundreds of substations operated by the state utility. Therefore even for two distribution substations located in the same area, it is not necessarily the case that blackout periods are synchronized. As it happens, grid power supplied to apartments in the experiment was sourced from two separate distribution substations. Therefore supply to about half the experimental population was associated with one node of the distribution grid while the remainder of apartments were linked to another.

As a consequence, over any period of time, the time spent by half the households in our population paying for high priced backup power could differ from the other half. In effect, these two groups of households were *as if randomly* assigned to two different regimes differing in the length of time on high priced power. For the most part, aggregated over a two day period, the differences time spent without power are insignificant but on occasion they may be substantial. I observe this variation in grid supply because the estate management kept records of total power outage durations for both grid substations separately in order to supervise diesel fuel consumption.

This variation allows me to obtain an unbiased estimate of price response θ. I estimate θ in 4 using the set of those time periods where the difference in outage hours between the two groups was at least two hours in total. I find that $\hat{\theta}$=-0.336 (s.d 0.14, p=0.015) which is equivalent to a price elasticity of about -0.13 (using the two flat tariffs of 3.2 and 12.10 INR). The coefficient is negative and significant although it is also relatively small suggesting that price responsiveness for this relatively well-off population is limited. This is also a short run elasticity which seems an appropriate measure to use as a comparison to the other behavior change instruments I examine[18].

The bias from alternative estimation methods is also easy to see. Calculating price elasticity by comparing consumption intensities under diesel and grid power yields an upward biased insignificant estimate of 0.003 [0.009]. Similarly using temporal variation by regressing the weekly average consumption of electricity against the weekly effective price (grid and diesel tariffs weighted by weekly hours of supply) yields an estimate of 0.06 [0.08] (an upward bias would occur here if high temperature weeks have higher demand but also longer grid outages and a higher effective price).

B. How do households respond to peer comparisons?

In section III.A I described the degree to which experiment households respond to tariff changes. Against this benchmark, I now examine the demand management treatments (both nudges and monetary incentives).

Treatment effects in their most general form may be estimated based on a simple

[18]Survey evidence in Table III.E suggests households did not respond to the demand management treatments by investing in new appliances (a common long run adaptive response).

reduced form equation of the following type

$$Y_{i,t}^{E,D} = 1 + \beta_1\mathbf{N} + \beta_2\mathbf{NI} + \gamma_t + \delta_i + \epsilon_{i,t} \tag{5}$$

where $Y_{i,t}^{E,D}$ is the grid electricity (E) or diesel backup (D) consumed in time period t by household i, $\beta_{1,2}$ are the treatment effects associated with receiving either a nudge using peer comparisons (**N**) or a combination of nudges and incentives (**NI**). Correspondingly $\mathbf{N}, \mathbf{NI}$ are dummy variables which take the value 1 when a household belongs to Treatment Group **N** or **NI** and 0 when the household belongs to the Control (**C**). γ_t, δ_i are time and individual fixed effects (to allow for general heterogeneity in household specific intercepts and time shocks) and and $\epsilon_{i,t}$ are unobserved household specific shocks. Electricity consumption data was recorded aggregated over two or three day intervals (each of which forms a time period t).

Results are reported in Table 2. I report estimates using the linear model described here as well as using models using $log(Y_{i,t}^{E,D})$ as the dependent variables (this allows interpretation of $\beta_{1,2}$ as percentage changes). As a robustness check I also compare results using both the original control group and expanding the control using the 50 de-facto control households for which electricity consumption data is available. I separately estimate effects on both the low priced electricity (grid power, 3.2 INR/KWh) and the high priced electricity (backup diesel power, 12.10 INR/KWh).

Across both linear and log-linear models (and robust to the addition of the 50 de-facto controls mentioned in Section II) households who were provided only the nudge treatment (peer comparisons and generic energy saving tips on the back page) reduced grid electricity consumption (their primary source of electricity that is priced relatively low) by an amount that is both statistically and economically significant.

		Low Price (Grid)			High Price (Diesel Backup)		
	Dependent Variable	I KWh	II ln(KWh)	III KWh	IV KWh	V ln(KWh)	VI KWh
1	Peer Comparisons (N)	-3.44**	-0.154**	-4.42***	0.005	-0.069	-0.245
		[1.72]	[0.074]	[1.66]	[0.50]	[0.072]	[0.472]
2	Comparisons+Incentive (NI)	0.03	-0.036	-0.942	0.562	0.082	0.312
		[1.48]	[0.058]	[1.43]	[0.46]	[0.059]	[0.428]
3	HH Fixed Effects	Y	Y	Y	Y	Y	Y
4	Time Fixed Effects	Y	Y	Y	Y	Y	Y
5	Include Default Controls?	N	N	Y	N	N	Y
6	% of Daily KWh (N)	-8.34**	-15.4**	-10.71***	0.05	-6.9	-2.4
		[4.17]	[7.4]	[4.02]	[4.8]	[7.2]	[4.5]
7	% of Daily KWh (NI)	0.06	-3.6	-2.28	5.4	8.29	3.02
		[3.58]	[5.8]	[3.44]	[4.5]	[5.9]	[4.1]
8	Mean Daily KWh	15.80	15.80	15.81	3.95	3.95	3.96
		[11.62]	[11.62]	[11.72]	[3.81]	[3.81]	[3.84]
9	Mean Intensity (KW)	0.86	0.86	0.86	0.93	0.93	0.93
10	No. of Households	466	466	516	466	466	516
11	Duration (days)	120	120	120	120	120	120
13	Incremental Effect of Incentives	3.47**	0.19***	3.47**	0.55	0.15**	0.55
		[1.54]	[0.065]	[1.54]	[0.45]	[0.061]	[0.45]

Notes: 1. *p<0.1; **p<0.05; ***p<0.01
2. Cluster robust standard errors (Arellano-Bond)
3. Model III and VI club together non-treatment groups D and DC
4. Rows 6-7 provide average treatment effects as a percentage of daily consumption.
5. Row 13 presents the interaction of incentives and the nudge.

Table 2—Impact of a nudge and incentive treatments on low and high priced electricity consumption

The log-linear models report larger treatment effects likely because the log dependent variable is sensitive to reductions at the lower end of the electricity consumption distribution (which may have modest absolute reductions in consumption but large percentage reductions).

My preferred estimates are therefore the more conservative estimates from the linear model (Columns I, III)[19]. The treatment effect estimate from model I for example implies a reduction of over 8 percent of the average daily consumption.

It is also interesting to separately look at the treatment effect on households who consume above and below the peer average. The literature on social norm effects does not rule out the possibility that low consuming households may increase consumption to approach the peer average although empirically this does not seem to occur for energy consumption behaviors in the US (Allcott, 2011a). To test whether reductions in response to the nudge are restricted to high consuming households I estimate quantile treatment effects for the nudge treatment. While quantile estimates are fairly imprecise owing to the small sample size involved, it seems clear from Figure that if norms are at work here they are likely 'injunctive', leading to a reduction in electricity consumption across quantiles.

C. Does the market price change the effectiveness of peer comparisons?

The evidence in Section III.B replicates the outcomes from Allcott (2011a) in a developing country setting. This suggests nudges using social norms may have significant potential across different cultural norms and economic contexts. However we might wonder whether these effects are influenced by the price of electricity or more generally by the economic stakes involved? The possibility that the usefulness

[19]In estimating average treatment effects I drop all time periods where a household reports no consumption of electricity (to control for unoccupied apartments / days).

of nudges may depend on other market incentives is central to evaluating their use as policy tools.

The setting in which these households consume electricity allows us to separately examine outcomes on electricity consumption from each of the two supply sources (grid and diesel power). This compares the effect of the same intervention (peer comparisons) at two different price levels, while holding the population and targeted behaviors (electricity consumption) constant.

I find that households provided a behavioral nudge (peer comparisons) do not reduce their consumption of high priced electricity (diesel backup, columns IV-VI of Table 2) although the use of low priced electricity changes significantly. Point estimates of the treatment effect on diesel power are both close to zero and not significant. Nevertheless treatment effects for grid electricity are negative and significant. While both effects are estimated with some uncertainty, I can statistically reject the hypothesis that these two point estimates are equal. To do so I first normalize the dependent variable (KWh of electricity) for both low and high priced electricity and then re-estimate treatment effects (to make coefficients comparable). The treatment effect estimates from these two normalized models are then tested for equality using a Wald Test (cluster robust errors) and the null hypothesis is rejected at 95 percent significance.

This result is consistent with a model of behavior where peer comparisons act through the imposition of psychological costs and where these costs diminish when the economic stakes are raised (perhaps because market prices increase). Note that this difference in treatment effects occurs even though the intensity of consumption of diesel power is similar to grid electricity (Table 1)[20]. In other words the lack

[20]This observation is also consistent with the relatively low price elasticity that we estimate in Section III.A.

of response to peer comparisons cannot be explained by differences in the baseline intensity of use.

This outcome also provides some indication on the mechanisms through which peer comparisons act and suggests that their effectiveness may not be purely due to a learning or information effect. It seems reasonable to assume that if households gain new information from peer comparisons and simply re-optimize consumption accordingly, then this effect should persist even if prices change. On the other hand, if peer comparisons act through imposing *psychological costs* when deviating from a social norm (the standard explanation in the psychology literature (Schultz et al., 2007)) then it is plausible that impacts will depend on whether the commodity in question is priced low or high. For instance households who consume above the norm may feel that they already pay a heavy price for this behavior and therefore no further psychological costs are imposed by the peer comparison.

D. What is the effect of coupling incentives with a nudge?

The economic stakes associated with energy behaviors are a function of both market prices and other monetary contracts offered to agents. In the case of electricity demand management it is not uncommon for consumers to be subjected to a suite of different programs - monetary and non-monetary - which are generally assumed to act independently. I find that while the peer comparison treatment causes significant reductions in consumption of grid electricity this is no longer true when the same information is augmented by financial incentives encouraging reductions in electricity use (Table 2, Model III and VI).

We may directly test whether the addition of financial incentives to the vanilla information intervention results in a statistically significant increase in electricity consumption by re-framing the reduced form equation we estimate as follows

$$Y_{i,t} = 1^J + \beta_1^J T + \beta_2^J I + \gamma_t^J + \delta_i^J + \epsilon_{i,t}^J \tag{6}$$

Here T is a dummy variable that is 1 for all households treated with any intervention (that is for groups P and PI pooled together) and I the interaction of the treatment dummy with a dummy representing the offer of financial rewards/penalties linked to the peer comparisons. If β_2, the coefficient on I is found to be positive and significant we may conclude that adding financial incentives to peer comparisons reduces household effort and leads to a deterioration in treatment effects consistent with motivation crowding out. The last row of Table 2 reports the coefficient β_2 when estimating an equation of this form.

One concern might be whether income effects explain what we find since households in the incentives group stand to gain some money[21]. However this does not seem to be a plausible explanation in this setting for various reasons.

First, the literature on electricity consumption has tended to find near zero income effects at this consumption level (see for instance Reiss and White (2005)). The level at which reward balances were initialized was modest (750 INR, less than 15 USD) and is over two orders of magnitude smaller than the average total amount that we would expect to be spent on rent payments alone for the four month period (about 2000 dollars at rental rates prevailing at the time of the study). For income effects to explain our results would require households to have an implausibly high income elasticity.

A more direct way of evaluating this concern is to re-run the estimation in Equation 6 dropping all observations where transfers are high. These should be households

[21]Although no transfers were actually made until the end of the experiment.

where income effects might be most important. I therefore drop observations where the reward balance exceeds a specified cutoff in the treatment with incentives (Group PI). To maintain a comparable population in the treatment group without rewards (Group P) I also simulate gains that would have accrued to them had they been eligible and drop these observations. Setting a cut-off as low as INR 1000 (less than 20 USD) still produces the same results, with the addition of financial incentives producing a statistically significant increase in electricity use (relative to the information only condition) even though over 30 percent of our observational data is dropped in the process.

I also run a specification retaining only those households whose *baseline* consumption levels were above the mean baseline electricity consumption. This ex-ante identifies households who would be expected to lose money once enrolled in the incentives program. Once again I obtain similar results.

E. Survey Evidence

A little more insight into these results (particularly the effectiveness of peer comparisons) can be had by examining responses to a short survey that was administered to a subset of households following program completion (70 in each condition). In addition to collecting basic demographic and appliance ownership data (Table 1), households were also asked to estimate the average number of hours per day they used different appliances over the summer. Responses are summarized in Table III.E.

Cooling accounts for easily the major share of electricity use in the summer season. Households in the Nudge Treatment group (N) reported a statistically significant reduction in hours of use of air-conditioners[22]. Households given a combination of

[22]One way to reduce electricity consumption in a hot, dry summer is to switch from using air conditioners to air coolers part of the day. In addition air conditioners can be switched off in the morning or switched off in unoccupied rooms

nudges and incentives report relatively higher air conditioning usage compared to the nudge only group though this difference is not statistically significant[23].

IV. Conclusions

Policymakers sometimes find it necessary to intervene in markets with a view to changing consumer behaviors. One such example involves demand side management of electricity consumption, especially in the residential sector.

In order to change energy behaviors various instruments have been recommended by both economists and psychologists. These include changing market prices (sometimes through taxes or subsidies), introducing conditional or unconditional financial incentives and, more recently, using non-monetary behavioral interventions or 'nudges'. Both nudges and conditional incentives have become popular for in electricity demand management (Ito, 2013) and peer comparisons in particular have been recommended as a useful tool with high potential(Allcott and Mullainathan, 2010).

Yet while these instruments have significant potential, much remains to be learned about how people respond to nudges and when they might fail to work. In particular, we know little about the effectiveness of nudges in the presence of other determinants of behavior such as price changes and financial incentives.

This paper provides cleanly measured field evidence from a real world program implemented over an extended period of time that helps us understand some of these questions. The experimental context involves urban upper-middle class households in India, an under-studied but growing population that is centrally important to determining both global and Indian energy consumption.

The results from this experiment suggest that behavioral nudges (in this case using

[23]Refrigerator use also declines slightly for households provided the peer comparisons. This measure includes secondary units that may not always be in use.

		Control (C)	Treatment (P)	Treatment (PI)	p (C,P)	p (P,PI)
1	Bought CFL lamps	0.48	0.54	0.53	0.51	0.90
2	Bought efficient AC	0.13	0.17	0.19	0.48	0.77
3	Bought Air-cooler	0.01	0.03	0.03	0.56	0.90
4	Bought efficient refrigerator	0.15	0.15	0.23	1.00	0.27
5	LCD TV Hours Used	4.58	4.49	5.16	0.86	0.14
6	CRT TV Hours Used	4.65	3.24	3.53	0.19	0.75
7	Room AC Hours Used	5.73	4.24	4.93	0.04**	0.30
8	Refrigerator Use	23.84	22.04	21.99	0.02**	0.96
9	Air Cooler Use	4.17	2.25	2.42	0.34	0.88

Notes: 1. *p<0.1; **p<0.05; ***p<0.01
2. Rows 1-4 present fraction of households reporting a new purchase during the summer.
3. Rows 5-9 present self reported hours of use per day.
4. Bonferroni adjusted p-values to account for multiple hypothesis tests.
5. All comparisons based on randomly selected subsample of 70 homes per condition.

Table 3—Purchase and use of appliances reported by surveyed households.

peer comparisons) can change energy behaviors. This provides the first field evidence on the effectiveness of these instruments from a developing country cultural and economic setting. Furthermore the effect of nudges on electricity consumption is significant even compared with price response. Indeed for the population studied in this experiment, replicating the mean effect of the nudge through tariff changes alone would require an approximately 65 percent increase in the price (see Sections III.A and III.B).

At the same time these nudges also have important limitations, at least within the context I study. The effectiveness of peer comparisons is significantly reduced when market prices for electricity are higher. This suggests that behavioral nudges in this context may be more effective when the underlying economic stakes are relatively low (see Section III.C).

The fact that the effectiveness of the nudge depends on prices also opens up the possibility that these techniques may interact with other economic cues in the marketplace, especially monetary incentives. Section I reviews experimental evidence from the psychology and behavioral economics literature which suggests that offering agents financial incentives to undertake a given behavior may not always enhance effort.

I find that peer comparisons becomes much less effective when coupled with financial incentives leading to a net increase in electricity consumption. Furthermore, because these incentives are relatively low powered, and because the (independently estimated) price elasticity is low, the negative impact of incentives on the effectiveness of the nudge outweighs any reduction in consumption from their role in increasing the marginal price of electricity (see Sections I.B and III.D).

These outcomes suggest a challenge for policy makers seeking to use nudges to change energy behaviors. Demand side management programs are rarely implemented in isolation and in practice consumers are likely to be subjected to a variety of different interventions. The dependence of nudges on market prices, as well as the potentially negative effect of conditional financial incentives on existing behavioral cues, also raises important design questions for the application of these instruments more generally[24].

Lastly it is necessary to acknowledge that this experiment also leaves some important questions unanswered.Although sustained for a season, the results of this pilot do not tell us whether peer comparisons would continue to be effective over a longer period. In addition, while I look at some factors that complicate the use of nudges, there are other considerations that might be relevant in practice. These include (i) Do incentives that are non-monetary (gifts for instance) perform better than cash rewards and work better with peer comparisons? (ii) Does the entity implementing a specific treatment matter (e.g utility, government or a local residents associations)?

These questions provide useful avenues for future research and their answers might help us understand better when nudges can reliably change behavior and when they might get crowded out by other forces in the marketplace.

REFERENCES

H. Allcott. Social norms and energy conservation. *Journal of Public Economics*, 95: 1082–1095, 2011a.

H. Allcott and S. Mullainathan. Behavior and Energy Policy. *Science*, 327(5970):

[24]Conditional cash transfers and nudges have found application in various arenas including health and education. Conditional cash transfers have had many successes (Gneezy et al., 2011, Fiszbein et al., 2009) but there are also examples (such as this experiment) where they have had little impact on incentivized behaviors or even had negative effects. (Benhassine et al., 2013, Gneezy and Rustichini, 2000)

1204–1205, March 2010. ISSN 0036-8075. doi: 10.1126/science.1180775. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1180775`.

Hunt Allcott. Rethinking real-time electricity pricing. *Resource and Energy Economics*, 33:820–842, 2011b.

Hunt Allcott and Michael Greenstone. Is there an energy efficiency gap? *Journal of Economic Perspectives*, 26(1):3–28, 2013.

Francisco Alpizar, Fredrik Carlsson, and Olof Johansson-Stenman. Anonymity,reciprocity, and conformity: Evidence from voluntary contributions to a national park in costa rica. *Journal of Public Economics*, 92:1047–1060, 2008.

D. Ariely, A. Bracha, and S. Meier. Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99 (1):544–555, 2009a.

Dan Ariely, Uri Gneezy, G Lowenstein, and Nina Mazar. Large stakes and big mistakes. *The Review of Economic Studies*, 76:451–469, 2009b.

Toshi Arimura, Richard Newell, and Karen Palmer. Cost-Effectiveness of Electricity Energy Efficiency Programs. Technical report, Resources for the Future, Washington D.C., 2009. URL `www.rff.org/documents/RFF-DP-09-48.pdf`.

Roland Benabou and Jean Tirole. Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520, 2003.

Roland Benabou and Jean Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, 2006.

Najy Benhassine, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. Turning a shove into a nudge? a "labeled cash transfer" for education. Technical report, National Bureau of Economic Research, Jul 2013. URL `http://dx.doi.org/10.3386/w19227`.

John Beshears, James Choi, David Laibson, Brigitte Madrian, and Katherine L. Milkman. The effect of providing peer information on retirement savings decisions. In *NBER Working Paper No. 17345*, 2011.

Hongbin Cai, Yuyu Chen, and Hanming Fang. Observational learning: Evidence from a randomized natural field experiment. *American Economic Review*, 99(3): 864–882, 2009.

Robert B. Cialdini. Descriptive Social Norms as Underappreciated Sources of Social Control. *Psychometrika*, 72(2):263–268, June 2007. ISSN 0033-3123. doi: 10.1007/s11336-006-1560-6. URL `http://www.springerlink.com/index/10.1007/s11336-006-1560-6`.

Timothy Conley and Christopher Udry. Learning about a new technology: Pineapple in ghana. *American Economic Review*, 100(1):35–69, 2010.

Peter Du Pont. *ENERGY POLICY AND CONSUMER REALITY: THE ROLE OF ENERGY IN THE PURCHASE OF HOUSEHOLD APPLIANCES IN THE U.S. AND THAILAND.* Phd, University of Delaware, 1998. URL `http://www.iiec.org/pdf/Chapter_1_8.pdf`.

Josef Falkinger. A simple mechanism for the efficient provision of public goods: Experimental evidence. *American Economic Review*, 90(1):247–264, 2000.

Massimo Filipini and Shonali Pachauri. Elasticities of electricity demand in urban indian households. Technical report, Center for Energy Policy and Economics Working Paper (16), 2002.

Ariel Fiszbein, Norbert Schady, F.H.G Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto, and Emmanuel Skoufias. Conditional cash transfers: Reducing present and future poverty. Technical report, The International Bank for Reconstruction and Development / The World Bank, 2009.

B. Frey and R. Jegen. Motivation crowding theory. *Journal of Economic Surveys*, 15(5):589–611, 2002.

Bruno Frey and Stephan Meier. Social comparisons and pro-social behavior: Testing 'conditional cooperation' in a field experiment. *American Economic Review*, 94 (5):1717–1722, 2004.

Kenneth Gillingham, Richard Newell, and Karen Palmer. Energy Efficiency Policies: A Retrospective Examination. *Annual Review of Environment and Resources*, 31(1):161–192, November 2006. ISSN 1543-5938. doi: 10.1146/annurev.energy.31.020105.100157.

List J. Gneezy, U. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5):1365–1384, 2006.

Uri Gneezy and Aldo Rustichini. Pay enough or dont pay at all. *Quarterly Journal of Economics*, 115(3):791–810, 2000.

Uri Gneezy, Stephan Meier, and Pedro Rey-Biel. When and why incentives dont work to modify behavior. *Journal of Economic Perspectives*, 25(4):1–21, 2011.

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

J. Henrich, S. Heine, and A Norenzayan. The weirdest people in the world? *Behavioral and Brain Sciences*, 33:61–83, 2010.

James Heyman and Dan Ariely. Effort for payment: A tale of two markets. *Psychological Science*, 15(11):787–793, 2004.

K. Ito. Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate program. Working Paper 19485, NBER, 2013. URL $http://people.bu.edu/ito/Ito_Rebate.pdf$.

Koichiro Ito. Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. Technical Report 18533, NBER Working Paper, Berkeley, 2012.

A Jaffe and R Stavins. The energy paradox and the diffusion of conservation technology. *Resource and Energy Economics*, 16(2):91–122, May 1994. ISSN 09287655. doi: 10.1016/0928-7655(94)90001-9. URL `http://linkinghub.elsevier.com/retrieve/pii/0928765594900019`.

Emir Kamenica. Behavioral economics and psychology of incentives. *Annual Review of Economics*, 4:427–452, 2012.

Roger Koenker. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89, 2004.

MGI. India's urban awakening: Building inclusive cities, sustaining economic growth. Technical report, Mckinsey Global Institute, 2010.

Nicholas Z. Muller, Robert Mendelsohn, and William Nordhaus. Environmental accounting for pollution in the united states economy. *American Economic Review*, 101(5):1649–1675, 2011.

Peter C Reiss and Matthew W White. Household Electricity Demand Revisited. *Review of Economic Studies*, 72:853–883, 2005.

C. Sansone and J. Harackiewicz. *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*. Academic Press: San Diego, CA, 2000.

Wesley Schultz, Jessica Nolan, Robert Cialdini, Noah Goldstein, and Vladas Griskevicius. The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18:429–434, 2007.

Anoop Singh. Power sector reform in india: current issues and prospects. *Energy Policy*, 34:2480–2490, 2006.

P. C. Stern, Thomas Dietz, G. Gardner, J. Gilligan, and M.P. Vandenbergh. Energy efficiency merits more than a nudge. *Science*, 328(5976):308–309, 2010.

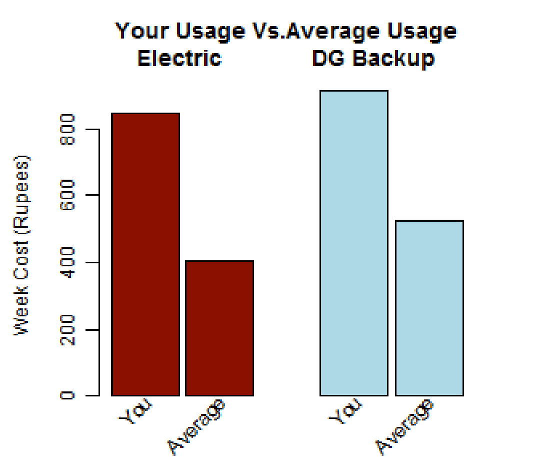
Anant Sudarshan. Deconstructing the Rosenfeld curve: Making sense of Californias low electricity intensity. *Energy Economics*, 39:197–207, Sep 2013. doi: 10.1016/j.eneco.2013.05.002. URL `http://dx.doi.org/10.1016/j.eneco.2013.05.002`.

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

UN. World urbanization prospects, the 2011 revision. Technical report, United Nations, Department of Economic and Social Affairs, Population Division, 2012.

Frank Wolak. Do residential customers respond to hourly prices? evidence from a dynamic pricing experiment. *American Economic Review: Papers and Proceedings*, 101(3):83–87, 2011.

Q. What is the use of this report card?
Every week we will show your weekly electricity consumption and your weekly DG backup power consumption. This report also shows you the average electricity costs of OTHER [redacted] households (of the same size as your flat). This will help you see if there are opportunities to reduce your own bills. Don't use more than you need to and investigate your electrical appliances and usage habits if you are using more than others!



	Your Costs/ Usage	Average Costs/ Usage
Electricity	Rs 849.6 (265.49 KVA)	Rs 405.2 (126.61 KVA)
DG Backup	Rs 917.2 (75.8 KVA)	Rs 525.3 (43.41 KVA)

There is no cost to these services. If you do not want to receive letters or have your usage data calculated you can drop out anytime by emailing [redacted] with Tower and Flat number. You will no longer receive any reports.

Figure 2. Exhibit sent to Peer Comparison households. Blacked out text contains location identifying information.

There is no cost to being in the reward programme. After October you will receive a cheque with your reward balance. If you do not want to receive letters or have your usage calculated you can drop out anytime by emailing [redacted] with Tower and Flat number. You will no longer receive any rewards or reports.

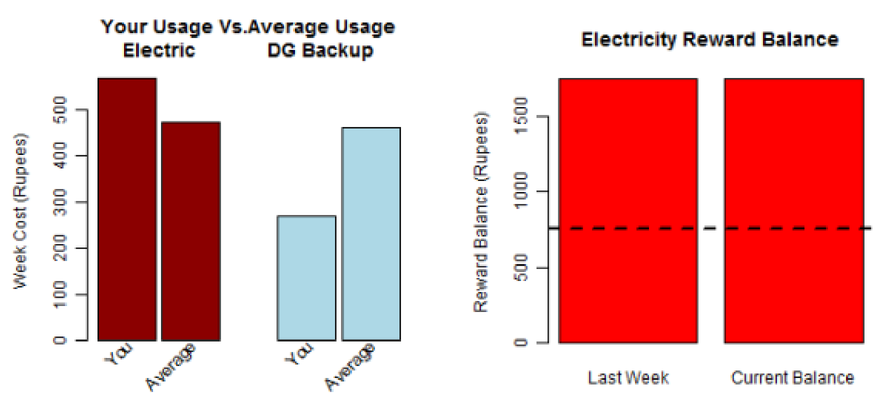
Q. What is the use of this report card?

Every week we will show your weekly electricity consumption and your weekly DG backup power consumption. This report also shows you the average electricity costs of OTHER [redacted] households (of the same size as your flat). This will help you see if there are opportunities to reduce your own bills. Don't use more than you need to and investigate your electrical appliances and usage habits if you are using more than others!

Q. What is the electricity rewards programme?

It's simple. **As a selected participant you begin with a balance of Rs 750!** Every week, if you use LESS electricity or backup power than the average of other residences your reward balance will increase (up to a maximum balance of Rs 5000.00). If you use MORE than average your reward balance will go down (minimum balance of Rs 0.00).

For electricity you gain Rs 2.0 for every unit that you are BELOW average (and lose Rs 2.0 for every unit ABOVE average). For DG backup you gain Rs. 4.0 for every unit you are below average and lose Rs 4.0 when above average. eg: If average electricity used is 50 KVA and your use is 30 KVA your reward balance will increase by 2*20=40 Rs.



	Your Costs/Usage	Average Costs/Usage	Reward Balance (Starting Balance: Rs. 750)	
Electric	Rs 568.6 (177.68 KVA)	Rs 470.8 (147.12 KVA)	Last Week Balance	Current Balance
Backup	Rs 270.2 (22.33 KVA)	Rs 459.4 (37.97 KVA)	Rs 1747.3	Rs 1748.74

Figure 3. Report card format sent to the group of households provided Peer Comparisons and associated Incentives. Blacked out text contains location identifying information.

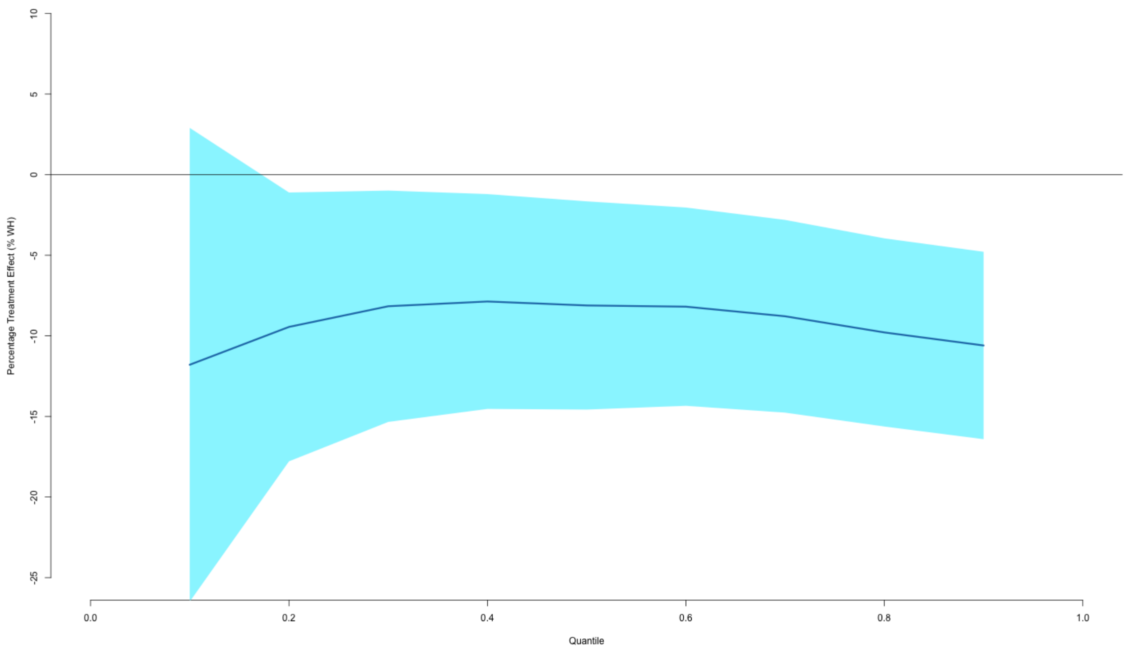


Figure 4. Quantile treatment effects for nudge only treatment with bootstrapped 90% confidence interval (Koenker, 2004).