Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Big-Data Mechanisms and Energy-Policy Design

Ankit Pat, Kate Larson and S. Keshav
Cheriton School of Computer Science
University of Waterloo
Waterloo, ON, Canada
{apat, klarson, keshav}@uwaterloo.ca

Abstract

A confluence of technical, economic and political forces are revolutionizing the energy sector. Policy-makers, who decide on incentives and penalties for possible courses of actions, play a critical role in determining which outcomes arise. However, designing appropriate energy policies is a complex and challenging task. Our vision is to provide tools and methodologies for policy makers so that they can leverage the power of big data to make evidence-based decisions. In this paper we present an approach we call *big-data mechanism design* which combines a mechanism design framework with stakeholder surveys and data to allow policy-makers to gauge the costs and benefits of potential policy decisions. We illustrate the effectiveness of this approach in a concrete application domain: the peaksaver PLUS program in Ontario, Canada.

Introduction

A confluence of technical, economic, and political factors are revolutionizing the energy sector around the world. In this transformation, policy-makers, who decide on incentives and penalties for possible courses of action, play a critical role. For example, the *Energiewende* in Germany was essentially driven by a policy incentive of high feed-in tariffs for solar and wind generators. Developing appropriate energy policies is a complex and challenging task. A policy has to find a balance between many competing forces and must take into account a range of potential outcomes. As a case in point, the German feed-in tariff had to balance the interests of decentralized generators and traditional grid operators. Outcomes included substantial deployment of renewable technologies (40 GW in the last eight years), but also significant drops in revenues of traditional electricity generators, a stressed transmission grid, increased reliance on fast-ramp but high-carbon generators, and the transfer of billions of Euros from consumers to solar farm owners (Energiewende 2014). These unanticipated consequences have enormous societal and political implications. There is, therefore, an urgent need for tools that allow policy-makers to make informed decisions.

Our vision is for policy-makers to leverage the power of data to make evidence-based decisions. In this work, we make this vision concrete using a specific approach and instantiating it in a specific problem domain. Our approach is to use stakeholder surveys coupled with formalisms from game theory and mechanism design to allow policy-makers to gauge the benefits of potential policy decisions, that, when balanced with implementation costs, allows evidence-based policy design.

Our problem domain is the peaksaver PLUS program in Ontario, Canada. This program provides homeowners with a programmable thermostat for free; in return, the homeowner allows the electricity utility to increase the temperature setpoint in their homes. The question we ask is: *What policy can be used to increase the penetration of peaksaver PLUS thermostats in Ontario?* One obvious approach is to give homeowners cash incentives to deploy thermostats. However, these incentives are likely to be fairly small in comparison to household incomes and research has shown that they are not very effective (Delmas, Fischlein, and Asensio 2013). Instead, we study the use of *non-cash* incentives to promote adoption of the peaksaver PLUS program. Since these non-cash incentives also have a monetary cost, there is a need to estimate the relative costs and benefits of different types of non-cash incentives. This forms the focus of our work.

The contributions of this paper are two-fold:

- We provide a methodology for policy design that couples agent-data elicitation with formalisms from game theory and mechanism design in order to provide support for evidence-based policy design. We illustrate the feasibility of our approach through an energy-policy design example.
- We illustrate that non-cash incentives can be an effective tool for promoting user behavior, and that it is possible to instantiate and reason about stakeholders' preferences and attitudes about non-cash incentives.

Motivation and Methodology

We envision an approach to energy-policy design which uses the formalism of game-theoretic and mechanism-design based models coupled with data collected via low-cost high-volume surveys (or other data collection methods).

Game theory provides us with a precise way of reasoning as to how rational agents will behave in strategic settings, whereas mechanism design focuses on describing how policies should be designed so as to ensure the incentives are in place so that desirable outcomes (often measured in terms of agents' preferences) arise. Despite its power and widespread usage, we see several limitations with these formal frameworks when it comes to energy-policy design.

First, at its core, there is a strong assumption that the agents in question are fully rational, whereas there is significant evidence from behavioral economics that this assumption has limited merit (Ariely 2008; Camerer 2003; Wright and Leyton-Brown 2010). We, on the other hand, are interested in cognitive biases and decision-making supports used by agents and how these could be leveraged when designing incentives.

Second, in much of the economics (and game-theoretic) literature agents' beliefs and preferences are treated as abstractions (Parkes and Wellman 2015), whereas with the advent of easy access to data, we argue that one should use this data to better model and instantiate the agent population, thus improving the analysis, predictions and robustness of models.

In this paper, we present a methodology for designing and evaluating policies using available data to model the underlying agent population and emphasizing the use of non-cash incentives (by leveraging cognitive biases and other decision-making supports). We describe the policy space, utility functions and goals of the different players using game-theoretic formalisms, but then use low-cost high-volume surveys to gather information about the effectiveness of non-cash incentives and use this information to model a population of agents with utility functions derived from the survey data. We test and evaluate different policies on this population, allowing the policy-maker the flexibility to experiment and determine the marginal benefit of alternative policy-design choices.

Our work complements several research themes. First, the use of surveys for policy evaluation is not a new idea in the energy domain (see, for example (DeCicco et al. 2014)). However, these surveys tend to be labor intensive, expensive to conduct, and often focus on post hoc analysis. We, instead, envision the use of low-cost high-volume data collection (through surveys or other means) that would allow for the modeling of agents' preferences within a population, on which alternative policies could be evaluated.

Econometric analysis of strategic settings is an interesting recent research direction (see, for example (Bajari, Hong, and Nekipelov 2013; Nekipelov, Syrgkanis, and Tardos 2015; Chawla, Hartline, and Nekipelov 2014)). Much of this work focuses on learning preference information and other game parameters via the strategic interactions of agents. Instead, we use data to learn agents' attitudes towards different incentive schemes and incorporate this into utility functions for analysis. Our work arguably falls within a preference elicitation framework (*e.g.* (Goldsmith and Junker 2009)) though to the best of our knowledge we are the first to use our proposed methodology for policy design, particularly in the energy sector. Finally, we see parallels with empirical mechanism design (Vorobeychik, Kiekintveld, and Wellman 2006). In the empirical mechanism design framework, learning and search techniques are used to estimate outcome features of interest as a function of mechanism parameters. While our methods are different, we also envision a framework where policy designers can experiment with alternative policy-features.

peaksaver PLUS

To illustrate our methodology, we use the peaksaver PLUS program, run by the Ontario Power Authority in Canada as an example. In this program, eligible participants (*i.e.* homeowners with a central air conditioner) are offered a free programmable thermostat. In exchange, on hot summer days, when electricity demand is at its highest, a signal is sent to the thermostats to increase the temperature setting by up to 2 degrees Celcius, thus reducing the central air conditioner's energy demands. If there are enough participants, then the program is able to reduce load on the system as a whole, thus ensuring wide adoption of the program is of interest to the corporation. Participants are provided with a free thermostat and should see some savings on their electricity bills, however will experience some discomfort as their thermostat settings are increased during certain periods of the summer.

Formal Model

In this section we present our formal model for the eligible participants (who we call *agents*) and the corporation (which we call the *principal*). At a high-level, the principal's goal is to reduce energy load across the system and can do this by increasing the temperature setting of agents' thermostats. The challenge, however, is that the agents have the freedom to participate in the program or not, and so the principal must offer incentives so as to encourage participation. This induces a tension where the principal must determine which incentives to offer so as to encourage agents to participate, while at the same time keeping its own costs as low as possible, whereas the agents prefer to experience no change in the temperature setting, but might be willing to accept certain (non-cash) incentives in exchange for experiencing some thermal discomfort.

The Principal

Let $\mathbb{I}$ be the set of (non-cash) incentives the principal may offer, and let $\theta_i \in \{0, 1\}$ be a variable where $\theta_i = 1$ if the principal offers incentive $i \in \mathbb{I}$ and $\theta_i = 0$ otherwise. We additionally assume that when the principal offers some incentive $i \in \mathbb{I}$, it incurs a cost C_i. An *incentive policy* is specified by a vector $P^{\text{in}} = (\theta_1, \ldots, \theta_{|\mathbb{I}|})$ indicating which incentives are offered by the principal. The cost associated with a particular incentive policy is $\text{cost}(P^{\text{in}}) = \sum_{i \in \mathbb{I}} C_i \theta_i$. A *policy*, P, is an incentive policy coupled with a (homogeneous) temperature increase, ΔT, the principal wishes to implement. That is, $P = \langle P^{\text{in}}, \Delta T \rangle$.

We assume that there is some minimal level of agent participation the principal requires before it will implement a policy, and we denote this minimum threshold as λ^*. Given a policy, P, we let $\lambda(P)$ be the number of agents willing to participate in the program given the policy. Given these conditions, we can specify the principal's optimization problem:

$$\begin{array}{ll} \text{maximize}_P & \lambda(P)\Delta T - \sum_i C_i\theta_i \\ \text{subject to} & \lambda(P) \geq \lambda^*. \end{array} \tag{1}$$

The Agents

We assume there is a set of N agents (*i.e.* eligible participants) who each derive some utility from a policy

$$P = \langle(\theta_1, \ldots, \theta_{|\mathbb{I}|}), \Delta T\rangle,$$

offered by the principal. In particular, for each incentive $i \in \mathbb{I}$ we assume agent j derives utility w_i^j when the incentive is provided. Similarly, we assume that each agent has a preferred temperature setting and experiences discomfort $d(\Delta T)$ for any temperature change ΔT. An agent's total utility, given policy P, is thus the linear combination

$$U_j(P) = \sum_{i \in \mathbb{I}} w_i^j \theta_i - d(\Delta T). \tag{2}$$

No agent can be forced to participate in the program, and so the action space of agent j is $A_j = \{\text{opt in}, \text{opt out}\}$ where the utility of *opting out* is 0. Thus, when presented with a policy P, agent j will only decide to *opt in* if

$$U_j(P) \geq 0. \tag{3}$$

Otherwise, agent j will *opt out*. Knowing this, we can rewrite $\lambda(P)$, defined previously, as

$$\lambda(P) = |\{j | U_j(P) \geq 0\}|.$$

Instantiating the Model: The Policy Space

One of our main interests is using non-cash incentives in (energy) policy design and, in particular, we argue that *cognitive biases* and other decision-making supports can be leveraged in this domain, like in other domains (Ariely 2008; Tversky and Kahneman 1974). In our study we included four well-studied non-cash incentives in set $\mathbb{I}$, which we describe in the following.

Status Quo Bias: The status quo bias is that an agent prefers to maintain the current situation or decision (Samuelson and Zeckhauser 1988). In our application domain, this could be implemented by enrolling agents into the program by default, for example, by legislating that all new homes must be enrolled in peaksaver PLUS.

Commitment Devices: A commitment device is a choice made by agents to restrict their number of future choices, and is often used as a way to reduce impulsive behavior (Bryan, Karlan, and Nelson 2010). In our application domain, this could be implemented by using longer-term contracts or making it possible to publicly acknowledge one's commitment to the peaksaver PLUS program, such as a sticker on one's door.

Social Norm Bias: The social norm bias simply states that agents prefer to conform to the rules or a behavior that is deemed acceptable by the larger group (Young 2008). In our application domain, a carefully crafted advertising campaign could be used to provide the message that the program is heavily adopted and has positive environmental impact, along with other desirable features the general population may care about.

Gamification: This refers to the use of game design elements in non-game contexts (Deterding et al. 2011). In our application domain this could include giving out points depending on energy savings and including such information on leader boards or through badges.

In concrete terms, a policy for the principal is a combination of the above incentives, coupled with a temperature increase ΔT.

Instantiating the Model: Using Surveys to Elicit Preferences

To elicit agents' preferences we conducted a survey to learn how agents react to the different non-cash incentives and to derive information about agents' thermal discomfort and preferred baseline temperature. We used CrowdFlower[1] for recruiting participants (with the geographical restriction that they must be in Canada or the United States), SurveyMonkey[2] to host the survey, and paid 1 USD to each participant. We had 990 responses and after a data cleaning process where we removed incomplete responses, duplicates based on IP filtering, responses that failed to follow directions, and those with very low variance across all questions, we were left with 425 responses.

The survey was divided into multiple sections. We used qualifier questions to gauge whether a respondent was a member of our target population (*i.e.* a homeowner with control of a thermostat). We then had a series of sections where we asked questions designed to elicit respondents' reactions to cognitive biases and non-cash incentives used in our model, before concluding with sections designed to elicit temperature preferences and relative preference-weights.

We used the survey responses to formulate a utility function for each respondent, which resulted in a population of 425 agents with different utility functions, where each utility function took the form presented in Equation 2. In the rest of this section we describe how we determined the utility function parameters.

Agents' Utilities: Non-Cash Incentives

For each incentive $i \in \mathbb{I}$ we assume that each agent j assigned a weight w_i^j indicating how strongly the presence of that type of incentive influenced their overall utility. Furthermore, we assume that w_i^j is separable such that $w_i^j = t_i^j \cdot S_i^j$ where S_i^j is the strength of the incentive for the agent and t_i^j scales the incentive strength by the relative extent to which incentive i was effective to overcome thermal discomfort, as compared to other incentives.

Our survey was structured into separate sections for each non-cash incentive. In each section we asked a set

[1] http://www.crowdflower.com

[2] http://www.surveymonkey.com

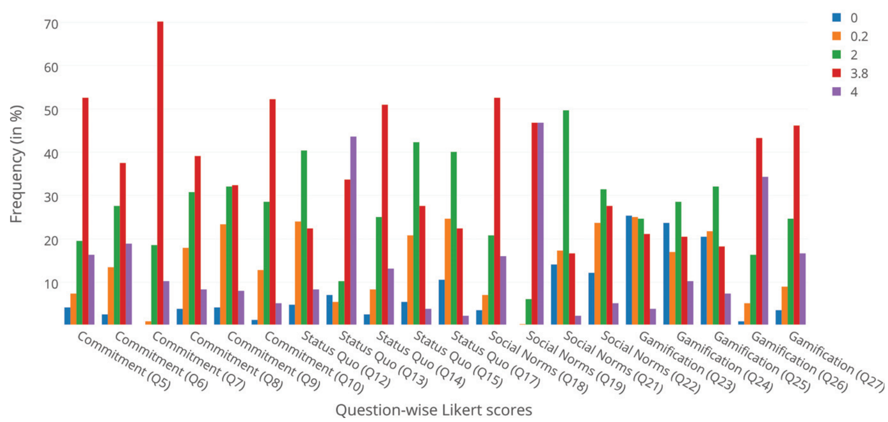

Figure 1: *The distribution of responses for every question eliciting non-cash incentive preferences.*

of questions designed to elicit respondents' attitudes towards the different incentives, where each respondent's reply to a question was scored using a five point adjusted Likert scale (Likert 1932; Lantz 2013) with scoring set $\{0, 0.2, 2, 3.8, 4\}$ to overcome the *middle-of-the-scale* effect (Lantz 2013). Figure 1 shows the raw survey results. To determine S_i^j we computed the average psychometric score for each bias.

We applied a scaling factor to each (average) psychometric score for each bias by asking each respondent the temperature increase, ΔT_i^j they might be willing to accept for each incentive class $i \in \mathbb{I}$. We then set

$$t_i^j = \frac{\Delta T_i^j}{\sum_{x \in \mathbb{I}} \Delta T_x^j}.$$

Agents' Utilities: Thermal Discomfort

For our model we leverage the significant body of research on determining thermal comfort for humans. Fanger proposed the Predicted Mean Vote (PMV) model to predict human comfort in different scenarios (Fanger 1970). The full model depends on several factors including air temperature, mean radiant temperature, air speed, humidity, metabolic rate of the person and their clothing level. Since we are interested in the comfort level of humans inside homes, during the summer, with a thermostat, we make a number of assumptions that reflect this. In particular, we assume humidity is equal to 50%, the mean radiant temperature is equal to the air temperature, the airspeed is zero, the metabolic rate is 1.2 *met*, and clothing level is 0.5 *clo* which represents typical summer indoor clothing (Hoyt et al. 2013). Using these assumptions, we simplify the PMV model using a linear regression model. Under this specific criteria, the regressed line fits

$$PMV(\Delta T) = 0.306|\Delta T| - 0.48$$

with mean square error 0.02%, where ΔT is the temperature change in Celsius.

A PMV of 0 is considered ideal, with the recommended limits of PMV for ideal comfort being in the range of [-0.5,0.5] according to the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Standard 55 (American Society of Heating, Refrigerating and Air-Conditioning Engineers 2013). This allows us to specify the thermal discomfort function as follows:

$$d(\Delta T) = \begin{cases} 0 & \text{if } |\Delta T| \leq 1.6 \\ PMV(\Delta T) & \text{otherwise} \end{cases}$$

Results

Using a population of agents with utility functions extracted from our survey data, we ran a series of experiments to determine how different policies changed participation rates and what temperature changes we could support with which incentive policies. Unless noted otherwise, we assumed that the costs incurred by the principal for implementing each incentive is zero.

In our first experiment we were interested in understanding how participation rates changed as a function of the policy, and how this influenced the principal's actual goal of maximizing the product $\lambda(P)\Delta T$. Figure 2 shows the results for the incentive policy $P^{\text{in}} = (1, 1, 1, 1)$ (*i.e.* where all incentives were deployed). As expected, participation decreased as ΔT increased, but, from the perspective of the principal, drops in participation were offset by higher temperature increases. Given this particular incentive policy, we observed that the optimal ΔT for maximizing the product $\lambda(P)\Delta T$ was approximately two degrees Celsius.[3]

For each possible incentive policy (assuming zero cost) we computed the optimal ΔT as shown in Figure 3. We observe that ΔT increased as more incentives were provided to the agents. However this increase was not uniform and different combinations of incentives were able to support different ΔT.

Finally, for each incentive policy we computed the optimal ΔT, and then plotted the participation rate under that temperature increase along with the principal's objective. The results are presented in Figure 4. We observe that participation rates decreased, but this is offset by participating

[3]We note that the current peaksaver PLUS policy is to also allow temperature increases of 2 degrees Celsius.

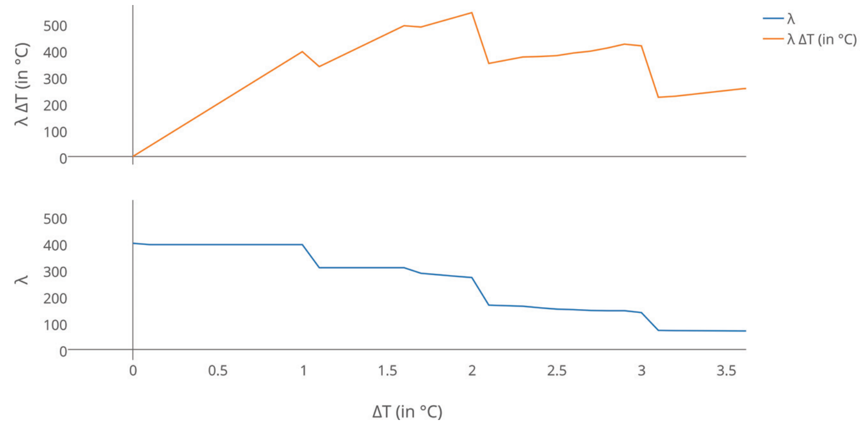

Figure 2: *Change in participation rate (bottom graph) and principal's utility (top graph) as a function of ΔT when all incentives are used.*

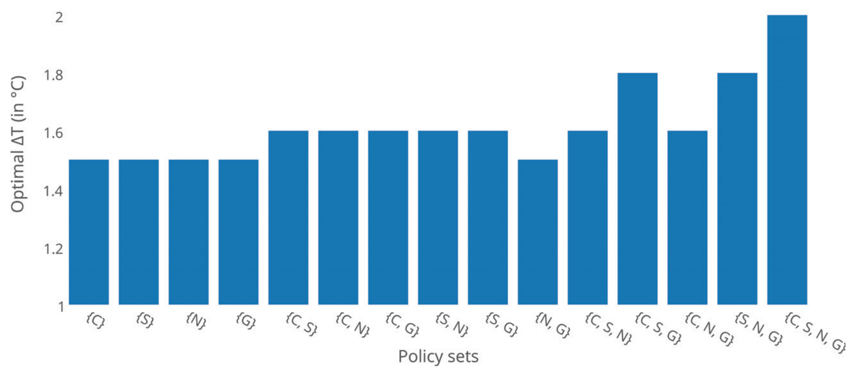

Figure 3: *Optimal ΔT for each incentive policy. We use the notation C for commitment device, S for status quo bias, G for gamification and N for social norm.*

agents accepting a higher ΔT in response to the provided incentives.

While it is unrealistic to assume the principal has no costs associated with providing different incentives, we also wish to make no claims as to what those costs might actually be, and thus did not instantiate the cost parameters C_i. However, we can compute the *marginal benefit* of adding different incentives to a policy and thus can provide the principal with information that allows it to determine how best to use its own resources. For example, given the population we surveyed, Table 1 describes the marginal benefit, in terms of $\lambda(P)\Delta T$, of adding additional incentives to policies. We note that adding any incentive had a benefit compared to not offering anything at all (*i.e.* as seen in the first row of Table 1). However, after that, given our survey population, we see that the marginal benefit of including a status quo and social norm bias in the incentives was high compared to other incentive structures.

Conclusion

Designing appropriate energy policies is a complex and challenging task. Our research goal is to provide tools and methodologies for policy makers so that they can leverage the power of (big) data to make evidence-based decisions. In this paper, using a concrete application of policy design for energy-use reduction, we showed how it is possible to combine formal game-theoretic models and elicit preferences about non-cash incentives through low-cost high-volume surveys, in order to provide policy-makers the ability to gauge the costs and benefits of potential policy decisions.

We see numerous directions for future research. For example, we used a simple weighted-linear utility model and assumed that there were no interdependencies between the cognitive biases. One future research question would be to see if richer utility models would provide additional insights and benefits, given the additional challenges of elicitation and instantiation they pose. We also used a model where the principal applied a single policy across the entire population. It is conceivable that there would be significant benefits to

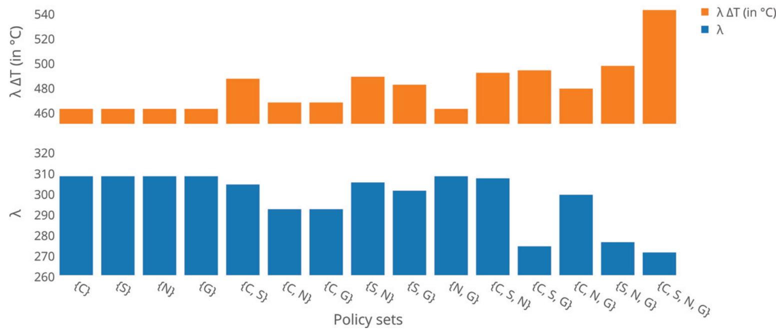

Figure 4: *Participation rate (bottom graph) and principal's objective as a function of the optimal ΔT for each incentive policy (top graph). We used the notation C for denoting commitment device, S for denoting status quo bias, G for denoting gamification and N for denoting social norm.*

Current Policy	Marginal Benefit (C)	Marginal Benefit (S)	Marginal Benefit (N)	Marginal Benefit (G)
$\{\}$	**462**	**462**	**462**	**462**
$\{C\}$		**24.4**	5.2	5.2
$\{S\}$	24.4		**26**	19.6
$\{N\}$	5.2	**26**		0
$\{G\}$	5.2	**19.6**	0	
$\{C, S\}$			4.8	**6.8**
$\{C, N\}$		**24**		11.2
$\{C, G\}$		**26**	11.2	
$\{S, N\}$	3.2			**8.8**
$\{S, G\}$	11.6		**15.2**	
$\{N, G\}$	16.4	**34.8**		
$\{C, S, N\}$				50.8
$\{C, S, G\}$			48.8	
$\{C, N, G\}$		63.6		
$\{S, N, G\}$	45.2			

Table 1: *Marginal contributions of adding a particular incentive given a current incentive policy. The notation C represents commitment device, S represents status quo bias, N represents social norm and G represents gamification.*

allowing the principal to tailor different policies for different subsets of agents. One challenge to this approach, however, would be identifying interesting agent subpopulations. As part of our research we experimented with a variety of machine learning methods to learn utility models. However, we had limited success. Further investigation in this direction is warranted to understand why this was so, and to determine what features are important to elicit from surveys or other data sources so that machine learning can be added as an effective part of the policy-design toolkit. For example, it would be interesting to try to learn what features make an individual react favorably to a particular incentive and thus design more fine-grained, targeted policies. Finally we are interested in illustrating our approach to policy-design in other domains, both energy-based and beyond.

References

American Society of Heating, Refrigerating and Air-Conditioning Engineers. 2013. Thermal environmental conditions for human occupancy. ANSI/ASHRAE Standard 55-2013.

Ariely, D. 2008. *Predictably irrational: The hidden forces that shape our decisions*. HarperCollins.

Bajari, P.; Hong, H.; and Nekipelov, D. 2013. Game theory and econometrics: A survey of some recent research. In *Advances in Economics and Econometrics*. 3–52.

Bryan, G.; Karlan, D.; and Nelson, S. 2010. Commitment devices. *Annual Review of Economics* 2:671–698.

Camerer, C. 2003. *Behavorial Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Chawla, S.; Hartline, J.; and Nekipelov, D. 2014. Mechanism design for data science. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC 2014)*, 711–712.

DeCicco, J.; Yan, T.; Keusch, F.; Munoz, D. H.; and Neidert, L. 2014. University of Michigan energy survey: Year one report. University of Michigan.

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

Deterding, S.; Dixon, D.; Khaled, R.; and Nacke, L. 2011. From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. ACM.

Energiewende, A. 2014. The German Energiewende and its climate paradox. Report 038/04-A-2014/EN.

Fanger, P. O. 1970. *Thermal comfort: Analysis and applications in environmental engineering*. Danish Technical Press.

Goldsmith, J., and Junker, U. 2009. Preference handling for artificial intelligence. *AI Magazine* 29(4):9.

Hoyt, T.; Schiavon, S.; Piccioli, A.; Moon, D.; and Steinfeld, K. 2013. CBE Thermal Comfort Tool. http://http://comfort.cbe.berkeley.edu/.

Lantz, B. 2013. Equidistance of Likert-type scales and validation of inferential methods using experiments and simulations. *Electronic Journal of Business Research Methods* 11(1):16–28.

Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology* 22(140):1–55.

Nekipelov, D.; Syrgkanis, V.; and Tardos, E. 2015. Econometrics for learning agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation (EC 2015)*, 1–18.

Parkes, D., and Wellman, M. 2015. Economic resasoning and artificial intelligence. *Science* 349(6245):267–272.

Samuelson, W., and Zeckhauser, R. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1:7–59.

Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–1131.

Vorobeychik, Y.; Kiekintveld, C.; and Wellman, M. 2006. Empirical mechanism design: Methods, with application to a supply-chain scenario. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC 2006)*, 306–315.

Wright, J. R., and Leyton-Brown, K. 2010. Beyond equilibrium: Predicting human behavior in normal-form games. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2010)*, 901–907.

Young, H. P. 2008. Social norms. In Durlauf, S. N., and Blume, L. E., eds., *New Palgrave Dictionary of Economics*. Palgrave Macmillan.