Applied Energy 111 (2013) 277–287

Contents lists available at SciVerse ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

The problems and solutions of predicting participation in energy efficiency programs

Alexander L. Davis *, Tamar Krishnamurti

Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, United States

HIGHLIGHTS

- Energy efficiency pilot studies suffer from severe volunteer bias.
- We formulate an approach for accommodating volunteer bias.
- A short questionnaire and classification trees can control for the bias.

ARTICLE INFO

Article history:
Received 21 December 2012
Received in revised form 11 April 2013
Accepted 29 April 2013
Available online 30 May 2013

Keywords:
Volunteer bias
Field studies
Prediction
Human behavior

ABSTRACT

This paper discusses volunteer bias in residential energy efficiency studies. We briefly evaluate the bias in existing studies. We then show how volunteer bias can be corrected when not avoidable, using an on-line study of intentions to enroll in an in-home display trial as an example. We found that the best predictor of intentions to enroll was expected benefit from the in-home display. Constraints on participation, such as time in the home and trust in scientists, were also associated with enrollment intentions. Using Breiman's classification tree algorithm we found that the best model of intentions to enroll contained only five variables: expected enjoyment of the program, presence in the home during morning hours, trust (in friends and in scientists), and perceived ability to handle unexpected problems. These results suggest that a short questionnaire, that takes at most 1 min to complete, would allow better control of volunteer bias than a more extensive questionnaire. This paper should allow researchers who employ field studies involving human behavior to be better equipped to address volunteer bias.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The past 6 years has seen rapid growth in legislation and research on the "smart grid" in the United States. Title XIII of the U.S. Energy Independence and Security Act of 2007 (U.S. Public Law 110-140) and Title IV of the American Recovery and Reinvestment Act of 2009 (U.S. Public Law 111-5) made resources available for modernizing the electric power grid. A critical component of these laws promotes better understanding and management of electricity use among residential customers. US electric utilities are conducting studies that test programs to achieve these goals, including tariffs such as critical peak pricing, home automation technologies such as smart thermostats, and feedback approaches such as in-home displays. These studies are designed to rigorously test these programs to ensure sound conclusions, even if that rigor is costly.

The most significant barrier to the validity of these studies is volunteer bias, where a customer's decision to be in the study may be causally related to benefit in the study. This bias can occur whenever someone offered the program refuses to participate, as refusal may indicate that the person would not have benefited from the program. As a result, using an all-volunteer sample is likely to overestimate program benefits to the population from which the sample was drawn.

Consider an extreme example that illustrates the problem. Suppose that residential customers choose to participate in an energy efficiency study based only on accurate knowledge about whether they will benefit [1]. If this is true, the study sample will be comprised wholly of customers that will benefit, whereas no person outside of the study would benefit. A failure to understand this would lead researchers to incorrectly conclude that the program would greatly benefit the general population when they would actually not benefit at all. That is, volunteer bias is an especially severe problem if researchers are unaware of the difference between volunteers and non-volunteers, or if aware, unable to compensate for it.

Recruiting a representative sample with little or no volunteer bias is not easy. Some customers do not want to participate no

* Corresponding author. Address: Department of Engineering and Public Policy, Center for Climate and Energy Decision-Making, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, United States. Tel.: +1 412 216 2040.
E-mail address: alexander.l.davis1@gmail.com (A.L. Davis).

0306-2619/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apenergy.2013.04.088

matter how much they are incentivized or cajoled, and it is difficult to determine their reasons for refusing. Unless these reasons are identified and taken into account when estimating the benefit of the program, inferences from the sample to population are not warranted. Comparing the sample and population on some observable characteristics is not sufficient to reduce the problem [2], as observed characteristics may not capture the causes of both volunteering and benefit in the program.

Despite this challenge, recruiting a representative sample is a necessary condition for establishing the validity of an energy efficiency study involving residential customers. It is a critical "part of the science" of any research involving human behavior [3]. In this paper we discuss the evidence on the causes of volunteer bias in energy efficiency programs and present a method of accommodating bias when, despite best efforts, it does occur [4]. Our approach is comprised of two parts. First, we develop a simple questionnaire that can be used to predict who will volunteer in an energy efficiency trial of in-home displays. We derive questionnaire items from previous research on volunteering [5] and evaluate their psychometric properties [6–9]. Second, we develop statistical models based on this questionnaire that accurately predict intentions to enroll in an in-home display trial. To test the quality of these statistical models, we compare modern machine learning algorithms against a simpler, often better, performing alternative: the best predictor [10,11].

2. Prevalence of volunteer bias

In this section we review the history of volunteer bias in residential energy efficiency studies in North America, using in-home displays as an example.[1] Volunteer bias in these studies has predominantly occurred in two ways. In the first way, customers self-select into the study and are then randomly assigned to different treatment groups. This makes inferences from sample to population uncertain because any factor that causes both volunteering and benefit cannot be controlled.

As an example of the first type of volunteer bias, the BC Hydro PowerCost Monitor time-of-use pilot used single family dwellings in British Columbia with an opt-in design [13]. Those recruited were more educated, had higher annual household income, were more knowledgeable about electricity conservation, more active in trying to save energy, more willing to change habits, and used on average 1700 kW h less than other comparable homes in the BC area. This study fits into the first volunteer bias category because random assignment to the treatment group (a PowerCost Monitor) or control group (no PowerCost Monitor) occurred *after* participants opted into the study. Because of volunteer bias, this design does not allow one to extrapolate from the study sample to the population. Both observed (e.g., household income) and unobserved differences between the sample and the population could make the study's results not applicable to the population. However, because random assignment occurred after volunteering decisions were made, the study does allow valid comparisons between groups in the sample.

A second type of volunteer bias, where studies recruit the control and treatment groups differently, not only makes inferences from sample to population invalid, but also invalidates any comparisons between groups within the sample. Because customers recruited for the control and treatment groups may be different in unknown ways, such studies cannot separate the effectiveness of the treatment (e.g., the in-home display) from differences between how samples were obtained.

Almost every trial of in-home displays succumbs to the second, more severe, form of volunteer bias. For example, the Milton Hydro Direct Energy Smart Home Energy Conservation Kit study recruited participants using telephone, direct mail, and billing inserts [14]. Those eligible to receive an in-home display had to be at least 18 years old, must have lived in the home for at least one year, did not plan to move, and expressed a willingness to complete two surveys during the study. Eligible customers who expressed interest and registered on-line were then contacted for an installation appointment based on the order they registered, resulting in 108 homes having an in-home display installed for free.

The control group was recruited differently, consisting of 23 volunteers from a pool of 300 recruited customers who had homes that were judged to be of similar size and age to that in the treatment group, who completed a survey for a $100 gift certificate, and who lived in geographic clusters near the treatment group homes. Any factor that differed in the recruitment approach, for example the use of the gift certificate, could make the control group not comparable to the treatment group. This pattern of recruiting those in the treatment and control groups differently holds for almost every other trial of in-home displays, including the Oberlin TED5000 study [15], the Ontario Energy Board Hydro One pilot [16], the Energy Trust of Oregon PowerCost Monitor study [17], the Baltimore Gas and Electric Smart Energy Pricing Pilot with the Energy Orb [18], the Omaha Public Power study [19], and the Florida Power and Light Energy Detective study [20].

Not all studies of in-home displays, however, have been affected by volunteer bias. The first exception, the Polk's Landing study [21], had displays installed in homes before customers bought them, with no way for buyers to know which homes had the displays beforehand. The Southern California Edison study [22] used an opt-out design with an opaque opt-out procedure, resulting in no opt-outs. The Commonwealth Edison Energy Smart Pricing Pilot with Pricelight study [23,24] and PG&E's Smart-Rate Pilot [25] both explicitly modeled volunteer bias using a propensity score model, an approach that is discussed in Section 3. Lastly, studies of energy efficiency program adoption by commercial entities, such as utilities, have shown that volunteer bias is minimal [26].

3. Adjusting for volunteer bias

Even if one follows current best practices for recruitment [27–29], some proportion of those who are offered the program will not participate. Fortunately, if a statistical model can be created that accurately predicts who volunteers and who does not, then the risk of incorrect generalization from sample to population can be minimized.

One simple approach is to use *propensity score adjustment* [30,31], that explicitly models the probability of volunteering for each person offered the program. While the justification for this approach is technically sophisticated, the intuition behind it is simple: if one can accurately model the probability of each customer volunteering for the program, then by adjusting for this probability (or propensity), one can generalize from the sample to the population.

The propensity score approach was used in both PG&E's Smart-Rate Pilot [25] and the Commonwealth Edison Energy Smart Pricing Pilot [23,24]. The Commonwealth Edison propensity score model, for example, included whether customers purchased new major appliances, used a fan to reduce costs, lived in a single-family detached home, were above 65 years old, and the number and type of people living in the household. Using logistic regression, they found that those who used fans to reduce costs, as well as

[1] This does not merely apply to in-home display studies, but we focus on them here to keep the discussion shorter. Davis et al. [12] provide additional references to studies on dynamic pricing and home automation.

those who had more people in the household, were significantly less likely to enroll (t-values of 2 and 3 respectively in the model), whereas those who lived in a single-family detached home were more likely to participate ($t = 2.4$). This model correctly predicted the enrollment decision of 71% of those included in the analysis, but did not fare well on other measures of ability to discriminate between volunteers and non-volunteers (e.g., a pseudo R^2 of 0.20).

Although promising, the statistical approach to controlling volunteer bias using propensity score adjustment must address three unavoidable and seemingly intractable hurdles: (1) finding the right predictors of volunteering, (2) correctly combining the right predictors to accurately predict volunteering, and (3) extracting the necessary data by getting non-volunteers to agree to complete a survey. In the next three sections we outline our approach to dealing with these three problems.

3.1. Predictor variables

The first challenge of building a propensity score model of volunteering is knowing what predictors to include. The problem of volunteer bias has been acknowledged across the social sciences, and research on volunteering is more than 80 years old [32–34]. However, many studies discussing why people participate (both in research studies and civic engagement) have been observational, retrospective, and failed to include a range of variables broad enough to adequately predict volunteering to the degree necessary to adjust for bias.

Perhaps the best resource that exists on the topic is Rosenthal and Rosnow's now classic review [5]. They found that volunteering for psychological research experiments was associated with demographic factors such as education, socioeconomic status (with more highly educated and higher SES groups being more likely to volunteer), and sex (women volunteer more often than men). Volunteers were also more motivated to seek approval and more sociable than non-volunteers, indicating psychological factors also play a role. Other evidence suggests that those who are proficient volunteers are more likely to volunteer [35]. In the case of an in-home display trial, for example, those who have previously volunteered for energy initiatives or engaged in energy conservation behavior may also volunteer to be part of an in-home display trial.

This research is limited, however, as it largely relies on understanding why university students decide to volunteer for psychology experiments, a decision that is likely insensitive to contextual factors that matter in real world recruitment. Those factors that are linked to the specific offering, rather than fixed attributes of the volunteer (e.g., demographics), may be as good or better predictors of volunteering, even in psychological experiments.

One example of such a contextual factor is the appeal and relevance of the specific initiative being presented. In one case, inpatients in an alcoholism rehabilitation program who volunteered for a trial on sexual behavior showed greater concerns about sexual functioning and higher incidences of sexual problems than non-volunteers in the same program [36]. Similarly, college students who volunteer for research on sexual behavior tend to differ from non-volunteers on sexual characteristics, such as sexual experience and confidence, but do not differ on personality or demographic variables [37–39].

3.2. Prediction method

Second, once a set of predictors is chosen, the best combination of them must be discovered from a large number of possibilities. Even with only one variable it is possible to construct an infinite number of regression models (e.g., polynomials of n degree). The traditional approach to estimating a propensity score model is to use logistic regression [40], as was done in the Commonwealth Edison Energy Smart Pricing Pilot [23,24] and PG&E's Smart-Rate Pilot [25]. This approach is severely limited, as the predictors must combine linearly, interactions must be specified by hand, and the set of variables must be selected by trial-and-error model comparison. As a result, logistic regression usually misses relationships in the data (e.g., non-linearities and interactions), and will often include too many variables, leading to overestimation of the model's predictive validity.

Modern machine learning approaches overcome these issues by automatically selecting the variables and configurations of those variables that best fit the data. To find a method that does this well, we began by looking at the most successful multivariate machine learning approaches [41]. However, for the data we collected, most of these approaches performed poorly or at least worse than simpler methods. Instead of summarizing the performance of all the approaches, we limit our summary to what is most commonly used for multivariate classification (logistic regression) and what we found to be the best machine learning approach, the classification tree [42].

Advances in statistical techniques are exciting, making it tempting to apply the newest tool to solve prediction problems. However, this excitement should be tempered by the finding that simple (improper) statistical rules outperform human prediction [43,44] as well as more complex models that estimate optimal weights, such as multiple linear regression [45–47]. To evaluate the validity of the logistic regression and classification tree, we benchmark them against a simpler alternative, the Take the Best heuristic [11], that only uses the best univariate predictor.

3.3. Recruiting non-volunteers

The third, and most important, reason why it is difficult to develop a good model of volunteering is that it is hard to collect data on non-volunteers. Developing and validating a predictive model requires non-volunteers to respond to other measures, such as a survey or phone interview. It is likely that those who do not want to participate in the efficiency program also do not want to subsequently respond to a survey or interview. Thus, any model of volunteering must rely either on data that is indirectly collected (e.g., public records), or from short, easily completed questionnaires that are highly likely to be responded to. By comparing complex statistical approaches with a simpler alternative, we achieve the latter objective, creating the simplest questionnaire that provides the highest predictive validity.

4. Study

We present a study that develops a simple questionnaire and statistical model to address volunteer bias.[2] The study presents participants with an offer to participate in an in-home display trial along with a questionnaire that includes different questions used to predict their intentions to enroll. The content of the offer was based on best practices described elsewhere [27].[3]

4.1. Methods

4.1.1. Participants

The participants were 279 US residential bill-payers recruited on Amazon's Mechanical Turk. Their average age was 31 years old (range 18–75), 92 (42%) were women, and their self-reported average monthly bill was $118 (range $12–$600).

[2] A similar study was conducted before this one, with similar, although not identical, results. The data and statistical analyses for this study can be obtained from here http://hdl.handle.net/1902.1/19154, labeled "Study One."

[3] All materials can be found in Appendix A.

4.1.2. Materials

Our first set of variables included standard demographics, namely age, gender, employment status, education, annual household income, race, and political affiliation [48].

Next, we included questions about *constraints* or barriers to enrollment in the trial [49]. First, potential participants must trust the study sponsor, as expecting deception, manipulation, or harm would undermine the perceived worth of the study [28]. To measure this, we asked about trust in the various sponsors of the hypothetical study, including local and federal government, scientists, and the utility company. To determine whether the trust was specific to the offering, or more general, we also asked questions about trust in friends, family, and coworkers. Next, we reasoned that those who trust the study sponsors must feel competent enough to succeed in the trial and able to control their electricity consumption. If one feels unable to use the in-home display or unable to control electricity consumption, then participation in the trial would be pointless. This possibility was examined using a small set of items about self-efficacy, a construct that taps one's perceived ability to accomplish goals in the face of challenge [50]. The last barrier to participation we considered was whether participants were in the home often enough to use the display. To assess this we asked for the total hours they spent in the home during the day and whether they were in the home for each of six 4-h time periods.

In addition to demographics and constraints, we considered two additional categories of variables that have been shown to predict volunteering in other studies: *motivations* and *topical interest*. Volunteering in civic programs (e.g., at a church) has been associated with social integration, such as having close friends that one frequently talks to [51]. A second motivation is a concern for saving money, which we captured using a scale of frugality [52]. Electricity use is a constant source of spending, and being able to monitor and reduce wasteful expenses may be particularly attractive to those who are frugal. Third, some customers may find trying new things enjoyable, independent of expected financial benefit, so we included a scale that measured the degree to which people enjoy exploring new things [53].

In terms of topics, we expected the program to be construed in one of three ways: (1) a study about environmental behaviors, (2) a study about in-home displays, or (3) a study about eco-friendly technologies more generally. Interest in the first topic was measured using the New Ecological Paradigm (NEP) [54], a widely used measure of environmental attitudes. Interest in the second topic was measured using a four-item scale of their attitudes and expectations of the in-home display, such as whether they expected it to help them save money [55]. General interest in environmentally friendly technologies were measured using a separate "eco-technology" scale [56].

4.2. Results

The results are divided into three sections. In the first section we present the psychometric properties of the questionnaire items and their univariate associations with intentions to enroll. In the second section we use two multivariate statistical methods, logistic regression and the classification tree, to model intentions to enroll. In the last section, we compare these methods on measures of in-sample and generalization error against a model that uses only the single best predictor.

4.2.1. Psychometric analyses and univariate prediction

The psychometric analyses use Chronbach's alpha [57] as a measure of internal consistency (reliability), and Principal Components Analysis as a measure of dimensional structure [58]. To create a short questionnaire that non-volunteers would be inclined to complete, we limit each scale to four questions, dropping items with the lowest item-total correlations [9,6–8].

The scales had substantial internal consistency, ranging from an α of 0.73–0.93. They also had a low dimensional structure, with the first factor of the Principal Components Analysis for each scale accounting for 48–78% of the scale variance. Despite this, the scales were not better predictors of intentions to enroll than the items in the scales. As a result, we focus on the individual items.

Intentions to enroll in the trial were unassociated with demographics. Table 1 shows univariate associations between the different measures of constraints on participation and intentions to enroll. Participants who reported being home in each time period were more likely to state that they were willing to enroll in the trial than those who were not around. This was confirmed by a strong correlation between total hours spent in the home and intentions to enroll. Measures of trust in the utility and scientists were also positively associated with willingness to enroll in the trial. Scientists were the most trusted of all the 'institutions,' higher than local government ($t = 8.1$, $p < 0.01$) and federal government, with the utility company being the least trusted. Scientists were also trusted more than community and co-workers, but less than family and friends. For the last constraint, self-efficacy, participants who reported that they would not bother trying to figure out complicated tasks stated that they were less willing to enroll in the trial, indicating that feelings of difficulty and inability to use the technology were an important element of the decision to enroll.

Table 2 shows the univariate associations between intentions to enroll and the motivation items, including social integration, frugality, and exploration. None of the social integration or frugality items were statistically significant. In the former case, this may indicate that the relationship between involvement in social groups and civic engagement may apply to domains where a community can be developed around the volunteering activity, which is not likely in an energy efficiency trial. In the latter case, individual concerns about saving money may not play much of a role in the decision to volunteer in energy efficiency programs. On the other hand, those who like to work problems and explore new subjects were more willing to enroll.

Table 3 shows the three topical scales. As can be seen, the single strongest predictor of intentions to enroll in the trial was interest in the in-home display. The belief that plants and animals exist to be used by humans was associated with unwillingness to enroll. Interest in eco-friendly technology was not associated with intentions to enroll.

4.3. Multivariate prediction of enrollment intentions

4.3.1. Logistic regression

The logistic multiple regression included all of the variables that had statistically significant univariate associations with intentions to enroll. It found that participants were more likely to indicate that they would enroll in the in-home display trial when they expected to enjoy the in-home display ($t(251) = 2.8$, $p = 0.01$), expected to learn from the in-home display ($t(251) = 2.3$, $p = 0.03$), trusted the utility company ($t(251) = 1.9$, $p = 0.07$), and were less active in their community ($t(251) = -1.7$, $p = 0.09$). We repeated the logistic regression using factor scores from the Principal Components Analysis of each scale, rather than the individual items, and found that the in-home display expectations factor was the only significant predictor ($t(256) = 4.7$, $p < 0.01$). Although it may be possible to add interactions and transformations of the variables to improve the model in a trial-and-error fashion, this is likely to overfit the data and yield spurious relationships [59,60]. Instead, in the next section we use the classification tree algorithm to automatically determine the best model.

Table 1
Univariate relationships between constraints on study participation and intentions to volunteer. Principal Components Analysis and reliability are presented for each scale.[a]

Item	χ^2 (p)	Mean	SD
Time at home			
6–10 am	**10.44** (**0.01**)	0.63	0.48
10 am to 2 pm	**3.92** (**0.05**)	0.41	0.49
2–6 pm	**3.82** (**0.05**)	0.49	0.50
6–10 pm	**3.12** (**0.08**)	0.86	0.35
10 pm to 2 am	**2.92** (**0.09**)	0.88	0.33
2–6 am	2.05 (0.15)	0.88	0.32
Item	τ (Z)	Mean	SD
Aggregate time periods			
Total hours	**0.19** (**3.47**)	16.61	6.34
Morning hours (6 am to 2 pm)	**0.18** (**3.21**)	4.16	3.34
Evening hours (6 pm to 2 am)	**0.13** (**2.22**)	6.95	2.26
Item	τ (Z)	Loading	α
Trust			
Your local government.	0.07 (1.20)	0.51	
Scientists.	**0.14** (**2.57**)	0.51	
Your utility company.	**0.15** (**2.51**)	0.53	
Your co-workers.	0.07 (1.17)	0.45	
Trust factor	**0.14** (**2.87**)	53%	0.78
Self-efficacy			
If something looks too complicated I will not even bother to try it.	−**0.14** (**2.43**)	0.49	
I do not seem capable of dealing with most problems that come up in my life.	−0.05 (0.89)	0.60	
When unexpected problems occur I do not handle them very well.	0.02 (0.43)	0.58	
When I make plans, I am certain I can make them work.	0.08 (1.36)	0.23	
Self-efficacy factor	0.07 (1.37)	52%	0.74

Bolded numbers are statistically significant at the 0.10 α level.
[a] τ is Kendall's measure of rank correlation between the variable and intentions to volunteer, and Z is the associated Z statistic testing τ against the null hypothesis of no rank correlation. The χ^2 test and associated p-values test whether those in the home during that time period were equally likely to indicate that they would volunteer compared to those not in the home. The factor loading for each item in the trust and self-efficacy scales is presented in the *Loading* column, and Cronbach's α for the entire scale is in the α column.

Table 2
Univariate correlations between items and scales related to motivations to enroll and intentions to volunteer.

Item	τ (Z)	Loading	α
Social Integration			
How many close friends do you have? (meaning people that you feel at ease with, can talk to about private matters, and can call on for help)	0.09 (1.71)	0.61	
How many of these friends do you see or talk to at least once every 2 weeks?	0.05 (0.98)	0.61	
How many times have you attended a party or other social gathering in the past 2 months?	−0.01 (0.15)	0.50	
Social integration factor	0.06 (1.13)	73%	0.81
Frugality			
If you take good care of your possessions, you will definitely save money in the long run.	0.03 (0.47)	0.42	
If you can re-use an item you already have, there's no sense in buying something new.	0.03 (0.48)	0.47	
There are things I resist buying today so I can save for tomorrow.	−0.00 (0.03)	0.56	
I am willing to wait on a purchase so that I can save money.	0.04 (0.75)	0.54	
Frugality factor	0.02 (0.47)	58%	0.82
Exploration			
I like to try to solve problems that present a mental challenge.	0.08 (1.44)	0.53	
I like to work at a problem until I get it right.	**0.13** (**2.33**)	0.49	
I am always eager to know more about the universe we live in.	0.07 (1.20)	0.48	
When I hear about a new subject I like to find out more about it.	**0.14** (**2.42**)	0.50	
Exploration factor	**0.11** (**2.21**)	63%	0.92

4.3.2. Classification tree

Fig. 1 shows the best tree discovered automatically using Breiman's classification tree algorithm.[4]

Moving along the main right branch, it can be seen that 189 of 204 (93%) who expected to enjoy the IHD ($\geqslant$4 on a 5 point scale) said they would volunteer. Looking at the left main branch, intentions to enroll depended on whether the person was in the home in the morning (between 6 am and 10 am). Seven of eleven (65%) who said they trusted scientists ($\geqslant$2 on a 5 point scale), but were not in the home in the morning, indicated that they would still enroll in the study, whereas 17 of 20 (85%) who did not trust scientists indicated they would refuse.

A surprising branch emerged for those who did expect to be home in the morning. Seventeen of nineteen participants (89%) who did not trust their friends a lot (<4 on a 4 point scale) indicated they would enroll. If they did trust their friends a lot, then 6 of 7 (86%) were willing to enroll if they felt unable to handle unexpected problems ($\geqslant$2 on a 5 point scale). If they felt able to

[4] A second classification tree was also used, excluding the in-home display expectations. It was more complex, including 9 items, and more difficult to understand. We omit it here.

Table 3
Univariate correlations between items and scales related to interest in the study topic and intentions to volunteer.

Item	τ (Z)	Loading	α
New Ecological Paradigm (NEP)			
Plants and animals exist primarily to be used by humans.	−**0.14** (**2.46**)	0.45	
The balance of nature is very delicate and easily upset.	0.01 (0.25)	0.49	
There are limits to growth beyond which our industrialized society cannot expand.	0.03 (0.58)	0.53	
The earth is like a spaceship with only limited room and resources.	0.09 (1.59)	0.53	
NEP factor	0.08 (1.53)	48%	0.73
In-Home Display (IHD) expectations			
I would enjoy having an in-home display in my home.	**0.38** (**6.71**)	0.48	
An in-home display would help me save electricity each month.	**0.33** (**5.91**)	0.51	
An in-home display would help me save money each month.	**0.34** (**6.10**)	0.53	
I would learn from an in-home display.	**0.36** (**6.28**)	0.48	
IHD factor	**0.35** (**6.83**)	78%	0.93
Eco-friendly technology			
I understand the potential damage to the environment that some products can cause.	−0.05 (0.96)	0.48	
I have switched products for ecological reasons.	0.04 (0.64)	0.55	
I have purchased a household appliance because it uses less electricity than other brands.	0.03 (0.61)	0.47	
When I have a choice between two equal products, I always buy the one that is less harmful to other people and the environment.	−0.09 (1.70)	0.50	
Purchases factor	0.03 (0.55)	59%	0.78

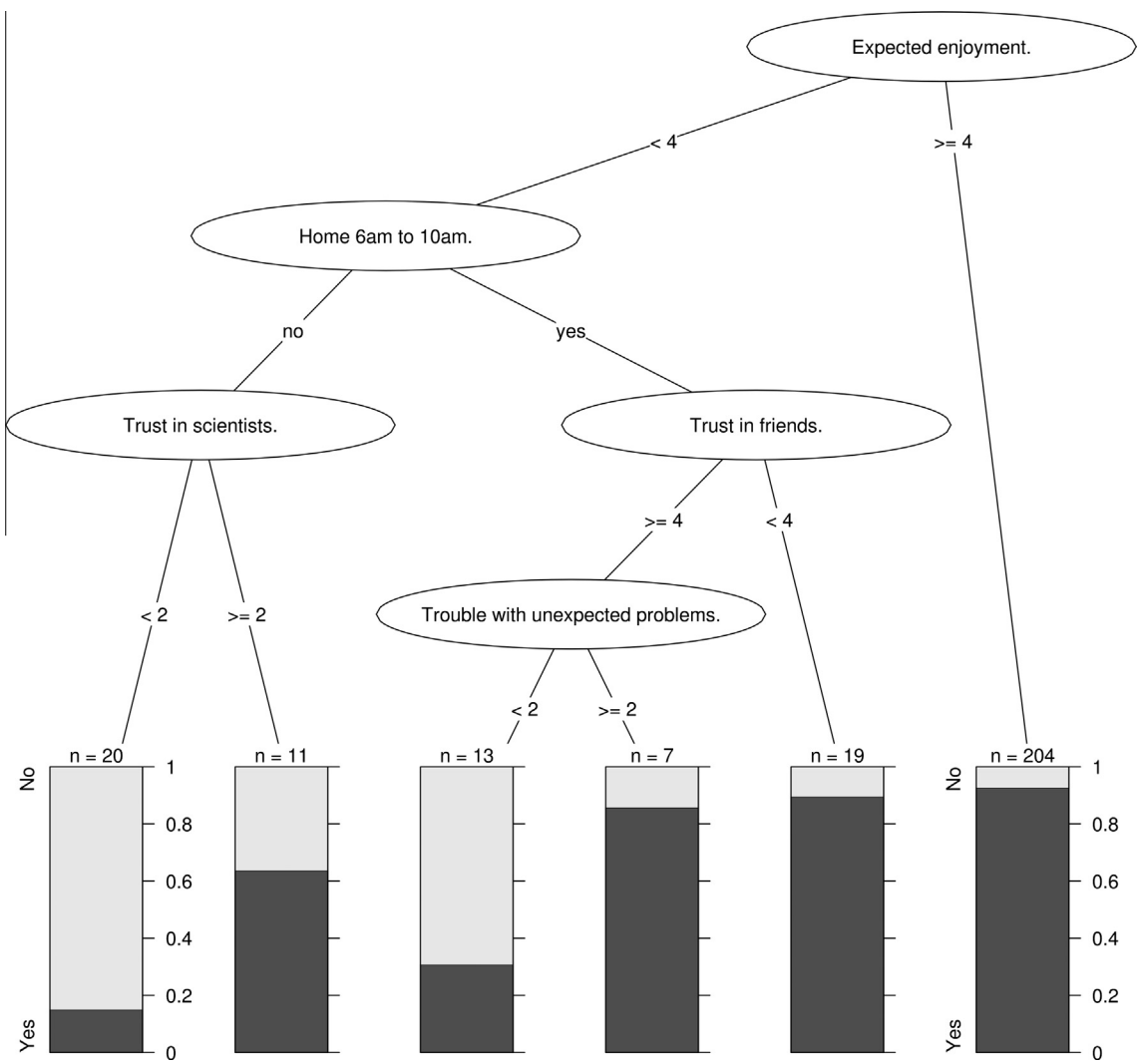

Fig. 1. *Enrollment Intentions Classification Tree.* This figure shows the classification tree that best predicted enrollment intentions. The terminal nodes at the bottom of the figure show barplots where the shaded region represents the proportion who intended to enroll. The sample size in each terminal node is shown above each barplot. The cutpoints are shown along each branch.

handle unexpected problems, 9 of 13 (69%) indicated they would not enroll.

This surprising pattern may indicate that those who do not trust their friends a lot and are home during the day feel that they are

self-reliant and can thus manage their energy consumption while they are in the home, thus making the in-home display a desirable offering. On the other hand, those who trust their friends a lot but feel less competent may rely on others to manage their energy consumption, making the in-home display less useful to them.

4.4. Comparison of methods

This section compares the complex optimized approaches (logistic regression and classification tree), to a simpler one (the best predictor), using estimates of both in-sample and generalization error [61]. Table 4 shows the performance of the different classification methods. As can be seen, the classification tree (Class Tree) performed the best on every measure. This approach was able to accurately classify 97% of those who intended to enroll, and 54% of those who said they would not enroll. As expected, logistic regression (Logistic) overfit the data, as indicated by the in-sample error underestimating the generalization error by about 20% according to both 10-fold cross-validation and the bootstrap. On the other hand, the in-sample error rate for the classification tree was slightly higher than 10-fold cross-validation, but lower than the bootstrap, indicating overfitting was minimal. Strikingly, logistic regression did not perform much better than the best predictor (Take the Best) in terms of generalization error.

4.5. Discussion

Volunteer bias undermines the ability to generalize the results of any study from the sample to the population. Previous energy efficiency studies of in-home displays with residential customers have had severe problems of volunteer bias. Although two previous studies adjusted for volunteer bias using propensity score adjustment, the models they used could be improved with a systematic approach. This paper provides this approach in three steps: (1) finding the right predictors, (2) combining them, and (3) using only few predictors.

Our ability to choose the right predictors had varying success. We divided predictors into four categories: *Demographics*, *Constraints*, *Motivations*, and *Topical Interest*. Demographics were unsuccessful at predicting enrollment intentions. This is consistent with prior research that suggests contextual factors of the program being offered matter more than fixed factors of the volunteer, and the high variability of demographics in predicting volunteering [48]. This result suggests that equating volunteers and non-volunteers on fixed attributes, such as demographics, is not sufficient to control for volunteer bias.

In contrast with demographics, constraints or barriers to participation, as a category, predicted enrollment intentions well. If people did not expect to be home during the day, especially from 6am-10am, then they saw no point in volunteering for the trial, indicating that the in-home display was seen as a device that they would check in the morning before going to work. People who distrusted scientists or the utility company, two of the three major institutions involved in the trial (the other being the federal government), also saw no point in enrolling. Because people trusted scientists more than any of the other institutions, focusing on the scientific purpose and validity of the study may be effective to encourage enrollment. The last constraint was whether participants felt they could use the device, or whether they saw it as too complicated for them to handle. Those who reported low self-efficacy, in terms of having difficulty solving complicated problems, also reported that they did not want to enroll in the study, indicating that the in-home display was seen as complicated to use. Our results show that examining constraints or barriers to participation in the study should be the first place to look when trying to predict volunteering.

Motivations to participate, as stable psychological factors or "traits," were not consistently related to intentions to enroll. Motivations to care more about the environment, be more energy friendly than others, and save money in general were not predictive at all. Like demographics, each of these factors can be considered stable traits of the participant, indicating that stable personality factors that are independent of the volunteering context are unlikely to be useful predictors of volunteering. There was some success, however, as those who enjoyed exploring new things felt as though the in-home display could satisfy that curiosity. This desire to explore may be useful for recruiting participants broadly, but could also be an anomaly of the motivations category, as enough items were tested that one or two could be statistically significant purely by chance.

Interest in the topic was also not a great category of predictors, except for interest in the specific technology being offered, the in-home display, which was the single greatest predictor of enrollment intentions. Thus, beliefs about benefit, and interest in the specific technology being offered, should be carefully measured to account for volunteer bias. Interest in the environment or new technologies were less predictive.

Addressing the second problem of prediction, the classification tree and best predictor performed well on estimates of in-sample error and generalization error. The success of the classification tree over the logistic regression suggests that the variables we used did not work independently to predict enrollment intentions, but instead interacted with each other, and the classification tree was able to automatically detect these interactions. This was particularly valuable for the unintuitive interaction between being home in the morning, trust in friends, and perceived inability to handle unexpected problems. It is unlikely that these interactions would have been specified in a logistic regression model based on a priori theory or intuition, nor would they have been discovered by trial-and-error.

The study also suggests that a survey consisting of only five questions could fully implement the best statistical model discovered by the classification tree. These five items were expectations of benefit from the in-home display, whether one would be home in the morning, whether one trusts one's friends, whether one trusts those offering the technology, and whether one feels competent enough to use the technology. A questionnaire with these items would likely take less than 1 min to complete, increasing the likelihood of non-responders returning the survey, if approached correctly and adequately incentivized. Thus, long burdensome questionnaires eliciting large amounts of psychodemographic information are unnecessary. While the five question survey would perform very well, the study also indicated that one would not perform much worse by having a single question survey, using the most predictive item.

Table 4
Comparison of logistic regression (Logistic) and the classification tree (Class Tree) against the single best variable (Take the Best) in predicting intentions to enroll.[a]

Method	In-sample error				Generalization error	
	ϕ	TPR	TNR	Error	10-fold CV	Boot
Take the Best	0.33	**0.97**	0.25	0.16	0.12	0.17
Logistic	0.51	**0.97**	0.42	0.12	0.15	0.15
Class Tree	**0.60**	**0.97**	**0.54**	**0.11**	**0.10**	**0.13**

Bold numbers indicate the statistical method that performed the best for each performance measure.

[a] ϕ is the correlation between the method's prediction and intentions to enroll. *TPR* is the proportion of people who intended to enroll correctly identified. *TNR* is the proportion of people who did not intend to enroll correctly identified. *Error* is the proportion of enrollment intentions incorrectly classified. *10-Fold CV* is the estimated generalization error rate using 10-fold cross validation. *Boot* is the bootstrap estimate of generalization error.

This research is limited in that it only evaluates intentions to enroll in a hypothetical, but not real, trial. Even though the data did not come from an actual recruitment, the approach we used is a valid guideline on how to develop a short set of questions that can be administered to customers via an electricity bill, short phone call, or mailing. Using this approach can minimize volunteer bias problems, which are costly both financially and scientifically.

5. Conclusion

Volunteer bias is a significant challenge to the rigor of energy efficiency studies involving residential customers. This paper proposed and tested a three-step approach to address this challenge, using a hypothetical in-home display trial as an example. The goal of the first step is to select the right variables that can be used to predict volunteering. We searched previous research on volunteering, as well as literature on volunteering for energy efficiency trials, and found that variables could fit into four categories: demographics, constraints, motivations, and topics. Our online study found that expected benefit in the program was the best single predictor of intentions to enroll in the trial, and constraints on participation, such was whether the person would be in the home to use the technology, was the most promising category. The second step used machine learning to discover the most predictive combination of these variables. Using the classification tree algorithm, we discovered a model with five variables that best predicted intentions to enroll: expected enjoyment in the program, presence in the home during morning hours, trust in friends, trust in scientists, and perceived ability to handle the unexpected. This model included several unintuitive interactions that would not have been discovered by using a more traditional approach, such as logistic regression. In the final step we proposed using a simple questionnaire, requiring less than 1 min to complete, to augment other variables that have previously succeeded in predicting enrollment in energy efficiency trials. By following these three steps, researchers who conduct studies of energy efficiency programs with residential customers can effectively address volunteer bias.

Acknowledgements

All materials and data, including statistical analyses, can be obtained from the first author's Dataverse.[5] We thank Jay Apt and Baruch Fischhoff for their thoughtful comments and support. This work was supported by the center for Climate and Energy Decision Making (SES-0949710), through a cooperative agreement between the National Science Foundation and Carnegie Mellon University. This material is also based upon work supported by the Department of Energy under Award Numbers DE-OE0000300 and DE-OE0000204. Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights, Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

[5] http://hdl.handle.net/1902.1/19154.

Appendix A. Materials

All participants first read the following introduction:

In this survey, we would like you to evaluate a program we are considering offering to residential electricity customers.
Please read the description on the next page carefully and judge whether you would want to participate in this program or not, if it were offered to you. Then, answer some questions about yourself in the main survey. Thank you!.

They then read the recruitment offering:

To whom it may concern,
Scientists at Carnegie Mellon University need your help to understand the best way to provide electricity information.
You have been chosen to receive a free in-home display. On the display you can see your home electricity use for 1 year.
You will evaluate the display for 1 year. At the end of the year, you can keep it if you like it or return it for a $25 gift certificate.
We will send you four short surveys (one every 3 months) over the year asking how useful you find the display and about your electricity use. If you choose to participate, you will receive the display 3 weeks from now. There will be no cost to you and your information and survey responses will be confidential, as is university policy.
We would like you to complete the attached questionnaire, even if you are unsure about participating. To thank you for your help, we have included a $2 bill. You can keep this $2 whether or not you participate in the study.
To understand if this display could benefit every resident in the area, it is very important that the people we choose to include in the study agree to do so. This helps us make sure that we have a representative set of participants, and is a critical part of the science we do.
If you want good, representative research to be conducted, please do participate.
If you would like to participate or to just ask questions, please do one of the following:
Call us: 1-800-111-1111 Email us: electricityframestudy@cmu.edu.
Return the enclosed questionnaire.
Return the enclosed postcard.
If we don't hear from you in a week, we will give you a call. If you don't want to participate, please call us, email us, or return the postcard checking the 'no thanks' box.
Thank you, Carnegie Mellon University Research Team.

The main dependent variable was their response to the following question:

Would you enroll in this offering if it were available to you? (Y/N).

Standard demographics such as age, gender, employment status (full time, part time, unemployed, student, homemaker, retired), education (less than high school, high school/GED, college, associate's degree, 4 year degree, and professional degree), annual household income, race, and political affiliation. We also asked if they were the primary billpayer, how many adults and children (under age 18) lived in the home, how many hours they spent in the home per day, and when they are usually in the home with six periods (10 am to 2 pm, 2–6 pm, 6–10 pm, 10 pm to 2 am, 2–6 am, 6–10 am).

Next was a series of questions used to predict intentions to volunteer.

Several social comparison questions were used:

To what extent do you disagree or agree with the following statements? Compared to the average household in my city...

- My household has done more to reduce its electricity consumption.
- My household cares more about the environment.
- My household recycles more consistently.
- My household is more active in the community.

The motivation questions were (strongly disagree to strongly agree):

To what extent would the following messages encourage you to join an electricity efficiency program?

- Increasing independent energy security for the US.
- Protecting the environment.
- Avoid wasting energy.
- Increasing personal control over energy use.

Next was the New Ecological Paradigm scale. They were asked:

To what extent do you disagree or agree with the following statements (strongly disagree to strongly agree):

- Plants and animals exist primarily to be used by humans.
- The balance of nature is very delicate and easily upset.
- There are limits to growth beyond which our industrialized society cannot expand.
- The earth is like a spaceship with only limited room and resources.

Several new technology and eco-purchasing questions:

To what extent do you disagree or agree with the following statements?

- I am always eager to be the first to buy a new technology.
- I understand the potential damage to the environment that some products can cause. I do not purchase these products.
- I have switched products for ecological reasons.
- I have purchased a household appliance because it uses less electricity than other brands.
- When I have a choice between two equal products, I always buy the one that is less harmful to other people and the environment.
- I only buy new products after the ones I have wear out or become obsolete.

One question asked participants to specify how much they trusted several different entities:

How much do you trust the following groups to look out for you? (not at all, somewhat, moderately, a lot).

- The federal government.
- Your local government.
- Scientists.
- Your utility company.
- Your community.
- Your family.
- Your friends.
- Your co-workers.

The behavioral volunteering items:

In the past 12 months have you done any of the following?

- Bought Compact Fluorescent Lights (CFLs).
- Used an electricity tracking device (e.g., an in-home display).
- Bought one or more energy efficient appliances.
- Insulated my home.
- Got a flu shot.
- Recycled.
- Contributed to a retirement savings (e.g., 401 k).
- Used the public library.
- Enrolled in prize drawings.
- Donated time to a charity or non-profit organization.
- Donated money to a charity or non-profit organization.
- Bought a lottery ticket.
- Signed a petition.
- Attended a protest or demonstration.
- Voted in a local or national election.
- Attended a group meeting (e.g., Neighborhood Watch, parent-teachers' association, social clubs, recreational groups, professional organizations, etc.).

They were then asked several questions about their awareness of technology:

Do you currently have a smart meter in your home (Yes/No/Don't know).
Have you heard of in-home electricity displays before this survey? (Yes/No/Don't know).
To what extent do you agree or disagree with the following statement? (strongly disagree to strongly agree).

1. An in-home display would help me save electricity each month.
2. An in-home display would help me save money each month.
3. I would enjoy having an in-home display in my home.
4. I would learn a lot from an in-home display.

The next questions were on self-efficacy:

To what extent do the following statements describe you? (not at all true, slightly true, somewhat true, very true, extremely true).

- If something looks too complicated I will not even bother to try it.
- When I decide to do something, I go right to work on it.
- I do not seem capable of dealing with most problems that come up in my life.
- When unexpected problems occur I don't handle them very well.
- When I make plans, I am certain I can make them work.

There were then several questions on curiosity and exploration:

To what extent do the following statements describe you? (not at all true, slightly true, somewhat true, very true, extremely true).

- I like to read books by writers I've not come across before.
- I like to try to solve problems that present a mental challenge.
- I like to work at a problem until I get it right.
- I like to look up new words in a dictionary.
- I am always eager to know more about the universe we live in.
- When I hear about a new subject I like to find out more about it.

The next set of questions pertained to frugality:

To what extent do you agree or disagree with the following statements?

- If you take good care of your possessions, you will definitely save money in the long run.
- There are many things that are normally thrown away that are still quite useful.
- Making better use of my resources makes me feel good.
- If you can re-use an item you already have, there's no sense in buying something new.
- I believe in being careful in how I spend my money.
- I discipline myself to get the most from my money.
- I am willing to wait on a purchase so that I can save money.
- There are things I resist buying today so I can save for tomorrow.

There was a final question about social integration:

Last of all, the following questions ask about how many people you see or talk to on a regular basis. (1, 2, 3, 4, 5, 6, 7 or more).

- How many close friends do you have? (meaning people that you feel at ease with, can talk to about private matters, and can call on for help).
- How many of these friends do you see or talk to at least once every 2 weeks?
- How many times have you attended a party or other social gathering in the past 2 months?

References

[1] Chassang S, Padro I Miquel G, Snowberg E. Selective trials: a principal-agent approach to randomized controlled experiments. Am Econ Rev 2012;102(4):1279–309.
[2] Heckman JJ, Vytlacil E. Structural equations, treatment effects, and econometric policy evaluation. Econometrica 2005;73(3):669–738.
[3] Hewison J, Haines A. Overcoming barriers to recruitment in health research. Brit Med J 2006;333(7562):300–2.
[4] Fischhoff B, Welch N, Frederick S. Construal processes in preference assessment. J Risk Uncertainty 1999;19(1):139–64.
[5] Rosenthal R, Rosnow RL. The volunteer subject. New Jersey: John Wiley & Sons; 1975.
[6] DeVellis RF. Scale development: theory and applications. Los Angeles: Sage Publications; 2011.
[7] Loehlin JC. Latent variable models: an introduction to factor, path, and structural equation analysis. New Jersey: Lawrence Erlbaum Associates; 2004.
[8] DeMars C. Item response theory. New York: Oxford University Press; 2010.
[9] Spector PE. Summated rating scale construction: an introduction. Los Angeles: Sage Publications; 1992.
[10] Dana J, Dawes RM. The superiority of simple alternatives to regression for social science predictions. J Educ Behav Stat 2004;29(3):317–31.
[11] Gigerenzer G, Goldstein DG. Betting on one good reason: take the best and its relatives. In: Gigerenzer G, Todd PM, ABC Research Group, editors. Simple heuristics that make us smart. New York: Oxford University Press; 1999. p. 75–95.
[12] Davis A, Krishnamurti T, Fischhoff B, Bruine de Bruin W. Setting a standard for electricity field studies. Unpublished Manuscript; 2012.
[13] Sulyma I, Tiedemann K, Pedersen M, Rebman M, Yu M. Power smart, and BC hydro, experimental evidence: a residential time of use pilot. In: ACEEE summer study on energy efficiency in buildings; 2008. p. 292–304.
[14] Schembri J. The influence of home energy management systems on the behaviours of residential electricity consumers: an ontario, Canada case study. Unpublished; 2008.
[15] Allen D, Janda K. The effects of household characteristics and energy use consciousness on the effectiveness of real-time energy use feedback: a pilot study. In: Proceedings of the ACEEE summer study on energy efficiency in buildings; 2006. p. 7.1–7.12.
[16] Frank S. Hydro one networks inc. time-of-use pricing pilot project results. Unpublished; 2008.
[17] Sipe B, Castor S. The net impact of home energy feedback devices. In: Energy program evaluation conference, Portland; 2009. p. 341–51.
[18] Faruqui A, Sergici S. BG&E smart energy pricing pilot (SEPP). Unpublished; 2009.
[19] Eiden JM. Investigation into the effects of real-time, in-home feedback to conserve energy. Unpublished; 2009.
[20] Parker D, Hoak D, Cummings J. Pilot evaluation of energy savings from residential energy demand feedback devices. Unpublished; 2008.
[21] McClelland L, Cook SW. Energy conservation effects of continuous in-home feedback in all-electric homes. J Environ Syst 1979;9(2):169–73.
[22] Sexton RJ, Johnson NB, Konakayama A. Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. J Consum Res 1987:55–62.
[23] Summit Blue Consulting. Evaluation of the 2005 energy-smart pricing plan-final report. Unpublished; 2006.
[24] Summit Blue Consulting. Evaluation of the 2006 energy-smart pricing plan-final report. Unpublished; 2007.
[25] George et al. PG&E load impact evaluation of Pacific Gas and Electric Company's time-based pricing tariffs. Unpublished; 2010.
[26] Väisänen H, Christensen W, Despretz H, Espergen KA, Lytras K, Reinikainen E, et al. Audit II–guidebook for energy audit programme developers; 2003.
[27] Treweek S, Pitkethly M, Cook J, Kjeldstrøm M, Taskila T, Johansen M, et al. Strategies to improve recruitment to randomised controlled trials. Coch Data Syst Rev 2010;4.
[28] Dillman DA. Mail and internet surveys: the tailored design method. New Jersey: John Wiley & Sons; 2007.
[29] Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires (review). Coch Data Syst Rev 2009;3:1–12.
[30] Wooldridge JM. Introductory econometrics: a modern approach. Ohio: South-Western Publishers; 2009.
[31] Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. New York: Cambridge University Press; 2007.
[32] Rosenzweig S. The experimental situation as a psychological problem. Psychol Rev 1933;40(4):337–54.
[33] Edgerton HA, Britt SH, Norman RD. Objective differences among various types of respondents to a mailed questionnaire. Am Sociol Rev 1947;12(4):435–44.
[34] Wallin P. Volunteer subjects as a source of sampling bias. Am J Sociol 1949;54(6):539–44.
[35] Callahan CA, Hojat M, Gonnella JS. Volunteer bias in medical education research: an empirical study of over three decades of longitudinal data. Med Educ 2007;41(8):746–53.
[36] Nirenberg TD, Wincze JP, Bansal S, Liepman MR, Engle-Friedman M, Begin A. Volunteer bias in a study of male alcoholics' sexual behavior. Arch Sex Behav 1991;20(4):371–9.
[37] Strassberg DS, Lowe K. Volunteer bias in sexuality research. Arch Sex Behav 1995;24(4):369–82.
[38] Wiederman MW. Volunteer bias in sexuality research using college student participants. J Sex Res 1999;36(1):59–66.
[39] Wolchik SA, Spencer SL, Lisi IS. Volunteer bias in research employing vaginal measures of sexual arousal. Arch Sex Behav 1983;12(5):399–408.
[40] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.
[41] Wu X, Kumar V, Ross Quinlan J, Ghosh J, Yang Q, Motoda H, et al. Top 10 algorithms in data mining. Know Inform Syst 2008;14(1):1–37.
[42] Breiman L. Classification and regression trees. Florida: Chapman & Hall/CRC; 1984.
[43] Swets JA, Dawes RM, Monahan J. Better decisions through science. Sci Am 2000;283(4):82–7.
[44] Dawes RM, Faust D, Meehl PE. Clinical versus actuarial judgment. Science 2000;243(4899):1668–74.
[45] Dawes RM. The robust beauty of improper linear models in decision making. Am Psychol 1979;34(7):571–82.
[46] Dawes RM, Corrigan B. Linear models in decision making. Psychol Bull 1974;81(2):95–106.
[47] Wilks SS. Weighting systems for linear functions of correlated variables when there is no dependent variable. Psychometrika 1938;3(1):23–40.
[48] Gannon MJ, Nothern JC, Carroll SJ. Characteristics of nonrespondents among workers. J Appl Psychol 1971;55(6):586–8.
[49] Mckenzie-Mohr D. Fostering sustainable behavior: an introduction to community-based social marketing. British Columbia: New Society; 2011.
[50] Sherer M, James EM, Mercandante B, Prentice-Dunn S, Jacobs B, Ronald WR. The self-efficacy scale: construction and validation. Psychol Rep 1982;51(2):663–71.
[51] Musick MA, Wilson J. Volunteering and depression: the role of psychological and social resources in different age groups. Soc Sci Med 2003;56(2):259–69.
[52] Lastovicka JL, Bettencourt LA, Hughner RS, Kuntze RJ. Lifestyle of the tight and frugal: theory and measurement. J Consum Res 1999;26(1):85–98.
[53] Derek CV, Hashim RR. Construct validation of a scale of academic curiosity. Psychol Rep 1974;35(1):263–6.
[54] Dunlap RE, Van Liere KD, Mertig AG, Jones RE. New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. J Soc Issues 2000;56(3):425–42.
[55] Costanzo M, Archer D, Aronson E, Pettigrew T. Energy conservation behavior: the difficult path from information to action. Am Psychol 1986;41(5):521–8.
[56] Fraj E, Martinez E. Ecological consumer behaviour: an empirical analysis. Int J Consum Stud 2006;31(1):26–33.

[57] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297–334.
[58] Pearson K. On lines and planes of closest fit to systems of points in space. Lond Edin Dublin Philos Mag J Sci 1901;2(11):559–72.
[59] Leamer EE. Let's take the con out of econometrics. Am Econ Rev 1983;73(1):31–43.
[60] Fischhoff B. For those condemned to study the past: heuristics and biases in hindsight. In: Kahneman D, Slovic P, Tversky A, editors. Judgment under uncertainty: heuristics and biases. New York: Cambridge University Press; 1982. p. 335–51.
[61] Efron B, Tibshirani R. An introduction to the bootstrap. Florida: Chapman & Hall/CRC; 1993.