The Problems of Predicting Participation in Energy Efficiency Programs

Alexander L. Davis
Tamar Krishnamurti

Department of Engineering and Public Policy
Center for Climate and Energy Decision Making
Carnegie Mellon University

January 6, 2013

Volunteer Bias

Volunteer Bias

Definition

Volunteer Bias

Definition

A volunteer sample provides a biased estimate of the treatment effect in a population.

Volunteer Bias

Definition

A volunteer sample provides a biased estimate of the treatment effect in a population.

- Most program evaluations depend on all-volunteer samples, usually early adopters.

Why Volunteer Bias?

Why Volunteer Bias?

My Reasons

Why Volunteer Bias?

My Reasons

- Pervasive problem without a solution or advocate.

Why Volunteer Bias?

My Reasons

- Pervasive problem without a solution or advocate.
- Widely applicable: energy, medicine, education, etc.

Why Volunteer Bias?

My Reasons

- Pervasive problem without a solution or advocate.
- Widely applicable: energy, medicine, education, etc.
- Subtle problem that many won't recognize.

Why Volunteer Bias?

My Reasons

- Pervasive problem without a solution or advocate.
- Widely applicable: energy, medicine, education, etc.
- Subtle problem that many won't recognize.
- Deceptive, allowing the public to be fleeced.
  - (this really annoys me)

Effect on you

Effect on you

Researchers

Effect on you

Researchers

- Exacerbates the challenge of recruitment in your behavioral studies.

Effect on you

Researchers

- Exacerbates the challenge of recruitment in your behavioral studies.

Decision-Makers

Effect on you

Researchers

- Exacerbates the challenge of recruitment in your behavioral studies.

Decision-Makers

- Cost-benefit estimates will be skewed toward favorability.

Effect on you

Researchers

- Exacerbates the challenge of recruitment in your behavioral studies.

Decision-Makers

- Cost-benefit estimates will be skewed toward favorability.
- Unwarranted inferences from studies with biased samples can waste money and delay the discovery of effective alternatives.

Effect on you

Researchers

- Exacerbates the challenge of recruitment in your behavioral studies.

Decision-Makers

- Cost-benefit estimates will be skewed toward favorability.
- Unwarranted inferences from studies with biased samples can waste money and delay the discovery of effective alternatives.
- PUC, DOE, and Utility may endorse, subsidise, or provide energy efficiency programs and technologies that are ineffective.

Example: In-Home Displays

Example: In-Home Displays

Example: In-Home Displays

- Primary Goal: Reduce overall household energy consumption.

Example: In-Home Displays

- Primary Goal: Reduce overall household energy consumption.
- Secondary Goal: Promote learning about causes of use.

Volunteer Bias

Volunteer Bias

Prior Evidence

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).
  - This is really, really, really-really-really, bad.

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).
  - This is really, really, really-really-really, bad.
- 4 had low risk of volunteer bias (24%).

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).
  - This is really, really, really-really-really, bad.
- 4 had low risk of volunteer bias (24%).
  - 1 study had no opt-outs.

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).
  - This is really, really, really-really-really, bad.
- 4 had low risk of volunteer bias (24%).
  - 1 study had no opt-outs.
  - 1 study participants didn't know about the study.

Volunteer Bias

Prior Evidence

- 17 studies in the US/Canada evaluated 11 different displays.
- 13 studies included only volunteers (76%).
- 12 recruited treatment and control group differently (70%).
  - This is really, really, really-really-really, bad.
- 4 had low risk of volunteer bias (24%).
  - 1 study had no opt-outs.
  - 1 study participants didn't know about the study.
  - 2 studies used propensity score adjustment.

Risks

Risks

Policy

Risks

Policy

- PUC mandates an “energy speedometer.”

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

"Hydro One, Ontario's largest utility, indicated that the presence of real time feedback devices has a measurable and positive impact on energy conservation."

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

"Hydro One, Ontario's largest utility, indicated that the presence of real time feedback devices has a measurable and positive impact on energy conservation."
"Save up to 20% off your electricity bill!"

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

"Hydro One, Ontario's largest utility, indicated that the presence of real time feedback devices has a measurable and positive impact on energy conservation."
"Save up to 20% off your electricity bill!"

- Actual *unadjusted* overall conservation was 6.5%.

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

"Hydro One, Ontario's largest utility, indicated that the presence of real time feedback devices has a measurable and positive impact on energy conservation."
"Save up to 20% off your electricity bill!"

- Actual *unadjusted* overall conservation was 6.5%.
- This study had every bias we assessed in our meta-analysis:[1]

[1] Davis, A., Krishnamurti, T., Fischhoff, B., and Bruine de Bruin, W. (2012). Setting a standard for electricity field studies. *Unpublished Manuscript*

Risks

Policy

- PUC mandates an "energy speedometer."
- Utility may buy technology, passing costs to customers.

Industry

"Hydro One, Ontario's largest utility, indicated that the presence of real time feedback devices has a measurable and positive impact on energy conservation."
"Save up to 20% off your electricity bill!"

- Actual *unadjusted* overall conservation was 6.5%.
- This study had every bias we assessed in our meta-analysis:[1]
  - Volunteers, Lack of randomization, Lack of blinding, Attrition

[1] Davis, A., Krishnamurti, T., Fischhoff, B., and Bruine de Bruin, W. (2012). Setting a standard for electricity field studies. *Unpublished Manuscript*

Solutions

Solutions

Study Design

Solutions

Study Design

- Randomization after recruitment.

Solutions

Study Design

- Randomization after recruitment.
- Opt-out, as non-responders are more procrastinators/misinformed than rejectors.[2]

[2] Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Serv Res*, 7:59–69

Solutions

Study Design

- Randomization after recruitment.
- Opt-out, as non-responders are more procrastinators/misinformed than rejectors.[2]
- Recruitment best practices (can be kept under 30%).

[2] Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Serv Res*, 7:59–69

Solutions

Study Design

- Randomization after recruitment.
- Opt-out, as non-responders are more procrastinators/misinformed than rejectors.[2]
- Recruitment best practices (can be kept under 30%).

Adjustment

[2] Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Serv Res*, 7:59–69

Solutions

Study Design

- Randomization after recruitment.
- Opt-out, as non-responders are more procrastinators/misinformed than rejectors.[2]
- Recruitment best practices (can be kept under 30%).

Adjustment

- Sample non-respondents.

[2] Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Serv Res*, 7:59–69

Solutions

Study Design

- Randomization after recruitment.
- Opt-out, as non-responders are more procrastinators/misinformed than rejectors.[2]
- Recruitment best practices (can be kept under 30%).

Adjustment

- Sample non-respondents.
- Extrapolate based on temporal information.[3]

[2] Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007). When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. *BMC Health Serv Res*, 7:59–69

[3] Armstrong, J. and Overton, T. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, 14:396–402

Propensity Score

Propensity Score

Propensity Score

Propensity Score

Propensity Score

- If we can confidently predict who will volunteer, then average treatment effects will be unbiased.

Requirements of the model

Requirements of the model

Model Requirements

Requirements of the model

Model Requirements

- Input variables to accurately predict enrollment (ignorability).

Requirements of the model

Model Requirements

- Input variables to accurately predict enrollment (ignorability).
- Estimation method that discovers the right configuration.

Requirements of the model

Model Requirements

- Input variables to accurately predict enrollment (ignorability).
- Estimation method that discovers the right configuration.
- Recruit non-volunteers to train the model.

Model Inputs

Model Inputs

Model Inputs

Model Inputs

Model Inputs

- Previous propensity score models.

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]
  - Commonwealth Edison Energy Smart Pricing Pilot.[5]

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

[5] Summit Blue Consulting (2007). Evaluation of the 2006 energy-smart pricing plan-final report

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]
  - Commonwealth Edison Energy Smart Pricing Pilot.[5]
    - Purchase of new major appliances, using fans to reduce costs, age, single-family detached home, # and type of people in household.

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

[5] Summit Blue Consulting (2007). Evaluation of the 2006 energy-smart pricing plan-final report

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]
  - Commonwealth Edison Energy Smart Pricing Pilot.[5]
    - Purchase of new major appliances, using fans to reduce costs, age, single-family detached home, # and type of people in household.
- Prior research on volunteering (survey research, psychology, medicine).[6]

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

[5] Summit Blue Consulting (2007). Evaluation of the 2006 energy-smart pricing plan-final report

[6] Rosenthal, R. and Rosnow, R. (1975). *The Volunteer Subject.* John Wiley & Sons

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]
  - Commonwealth Edison Energy Smart Pricing Pilot.[5]
    - Purchase of new major appliances, using fans to reduce costs, age, single-family detached home, # and type of people in household.
- Prior research on volunteering (survey research, psychology, medicine).[6]
- Those generated from open-ended, mental-models interviews.

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

[5] Summit Blue Consulting (2007). Evaluation of the 2006 energy-smart pricing plan-final report

[6] Rosenthal, R. and Rosnow, R. (1975). *The Volunteer Subject.* John Wiley & Sons

Model Inputs

Model Inputs

- Previous propensity score models.
  - PG&E's Smart-Rate Pilot.[4]
  - Commonwealth Edison Energy Smart Pricing Pilot.[5]
    - Purchase of new major appliances, using fans to reduce costs, age, single-family detached home, # and type of people in household.
- Prior research on volunteering (survey research, psychology, medicine).[6]
- Those generated from open-ended, mental-models interviews.
- Our expert judgment given the context.

[4] George et al. (2010). PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs

[5] Summit Blue Consulting (2007). Evaluation of the 2006 energy-smart pricing plan-final report

[6] Rosenthal, R. and Rosnow, R. (1975). *The Volunteer Subject.* John Wiley & Sons

Survey

Survey

we conducted a survey

Prediction Task

Prediction Task

Your Prediction

Prediction Task

Your Prediction

"You have been chosen to receive a free in-home display. On the display you can see your home electricity use for one 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."

Prediction Task

Your Prediction

"You have been chosen to receive a free in-home display. On the display you can see your home electricity use for one 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."

- What do you think will predict volunteering? (write down).

Demographics

Demographics

Will demographics predict?

Age, Gender, Employment Status, Education, Annual Household Income, Race, Political Affiliation, Adults/Children in the home.

Yes or No?

Demographics

Will demographics predict?

Age, Gender, Employment Status, Education, Annual Household Income, Race, Political Affiliation, Adults/Children in the home.

Yes or No?

No

Demographics

Will demographics predict?

Age, Gender, Employment Status, Education, Annual Household Income, Race, Political Affiliation, Adults/Children in the home.

Yes or No?

No

These are standard demographics, some of which have been shown to predict volunteering previously.

Number of hours in the home

Number of hours in the home

Will number of hours in the home predict?

6am-10am, 10am-2pm, 2pm-6pm, 6pm-10pm, 10pm-2am, 2am-6pm

Yes or No?

Number of hours in the home

Will number of hours in the home predict?

6am-10am, 10am-2pm, 2pm-6pm, 6pm-10pm, 10pm-2am, 2am-6pm

Yes or No?

Yes

Number of hours in the home

Will number of hours in the home predict?

6am-10am, 10am-2pm, 2pm-6pm, 6pm-10pm, 10pm-2am, 2am-6pm

Yes or No?

Yes

Participants need to be in the home to expect that the trial would be worthwhile.

Trust

Trust

Will trust predict?

- Local government
- Scientists
- Utility Company

Yes or No?

Trust

Will trust predict?

- Local government
- Scientists
- Utility Company

Yes or No?

Yes

Trust

Will trust predict?

- Local government
- Scientists
- Utility Company

Yes or No?

Yes

Survey literature suggests trust in the source matters.

($\tau = .\mathbf{14}$, Factor Loading $= 53\%$, Cronbach's $\alpha = .78$)

Self-Efficacy

Self-Efficacy

Will self-efficacy predict?

- "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."

Yes or No?

Self-Efficacy

Will self-efficacy predict?

- "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."

Yes or No?

No

Self-Efficacy

Will self-efficacy predict?

- "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."

Yes or No?

No

The General Self-Efficacy Scale[7]

($\tau = .07$, Factor Loading $= 52\%$, Cronbach's $\alpha = .74$)

[7] Sherer, M., James, E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., and Ronald, W. The self-efficacy scale: Construction and validation. *Psychological Reports*, 51

Social Networks

Social Networks

Will integration in social networks predict?

- "How many close friends do you have?"
- "How many of these friends do you see or talk to at least once every 2 weeks?"

Yes or No?

Social Networks

Will integration in social networks predict?

- "How many close friends do you have?"
- "How many of these friends do you see or talk to at least once every 2 weeks?"

Yes or No?

No

Social Networks

Will integration in social networks predict?

- "How many close friends do you have?"
- "How many of these friends do you see or talk to at least once every 2 weeks?"

Yes or No?

No

The Social Network Index[8]

($\tau = .06$, Factor Loading $= 73\%$, Cronbach's $\alpha = .81$)

[8] Cohen, S., Doyle, W., Skoner, D., Rabin, B., and Gwaltney Jr, J. (1997). Social ties and susceptibility to the common cold. *JAMA: the Journal of the American Medical Association*, 277(24):1940–1944

Frugality

Frugality

Will frugality predict?

- "There are many things that are normally thrown away that are still quite useful."

Yes or No?

Frugality

Will frugality predict?

- "There are many things that are normally thrown away that are still quite useful."

Yes or No?

No

Frugality

Will frugality predict?

- "There are many things that are normally thrown away that are still quite useful."

Yes or No?

No

Lastovicka Frugality Scale[9]

($\tau = .02$, Factor Loading $= 58\%$, Cronbach's $\alpha = .82$)

[9] Lastovicka, J., Bettencourt, L., Hughner, R., and Kuntze, R. (1999). Lifestyle of the tight and frugal: theory and measurement. *Journal of Consumer Research*, 26(1):85–98

Exploration and Curiosity

Exploration and Curiosity

Will exploration and curiosity predict?

- "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."

Yes or No?

Exploration and Curiosity

Will exploration and curiosity predict?

- "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."

Yes or No?

Yes

Exploration and Curiosity

Will exploration and curiosity predict?

- "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."

Yes or No?

Yes

Curiosity Index[10]

($\tau = .\mathbf{11}$, Factor Loading $= 63\%$, Cronbach's $\alpha = .92$)

[10] Derek, C. and Hashim, R. (1974). Construct validation of a scale of academic curiosity. *Psychological Reports*, 35(1):263–266

Concern for the Environment

Concern for the Environment

Will concern for the environment predict?

- "The balance of nature is very delicate."
- "Plants and animals exist primarily to be used by humans."

Yes or No?

Concern for the Environment

Will concern for the environment predict?

- "The balance of nature is very delicate."
- "Plants and animals exist primarily to be used by humans."

Yes or No?

No

Concern for the Environment

Will concern for the environment predict?

- "The balance of nature is very delicate."
- "Plants and animals exist primarily to be used by humans."

Yes or No?

No

The New Ecological Paradigm[11]

($\tau = .08$, Factor Loading $= 48\%$, Cronbach's $\alpha = .73$)

[11] Dunlap, R., Van Liere, K., Mertig, A., and Jones, R. (2000). New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. *Journal of Social Issues*, 56(3):425–442

Expectations of the In-home display

Expectations of the In-home display

Will expectations of the in-home display predict?

- "I expect to save money using the in-home display."
- "I expect to enjoy using the in-home display."

Yes or No?

Expectations of the In-home display

Will expectations of the in-home display predict?

- "I expect to save money using the in-home display."
- "I expect to enjoy using the in-home display."

Yes or No?

Yes

Expectations of the In-home display

Will expectations of the in-home display predict?

- "I expect to save money using the in-home display."
- "I expect to enjoy using the in-home display."

Yes or No?

Yes

Our intuition that expected benefit from the specific offering is a strong motivator of volunteering.

($\tau = .\mathbf{35}$, Factor Loading $= 78\%$, Cronbach's $\alpha = .93$)

Making Eco-Friendly Purchases

Making Eco-Friendly Purchases

Will the desire to make eco-friendly purchases predict?

- "I have switched products for ecological reasons."
- "I have purchased a household appliance because it uses less electricity than other brands."

Yes or No?

Making Eco-Friendly Purchases

Will the desire to make eco-friendly purchases predict?

- "I have switched products for ecological reasons."
- "I have purchased a household appliance because it uses less electricity than other brands."

Yes or No?

No

Making Eco-Friendly Purchases

Will the desire to make eco-friendly purchases predict?

- "I have switched products for ecological reasons."
- "I have purchased a household appliance because it uses less electricity than other brands."

Yes or No?

No

Ecological Purchasing[12]

($\tau = .03$, Factor Loading $= 59\%$, Cronbach's $\alpha = .78$)

[12] Fraj, E. and Martinez, E. (2006). Ecological consumer behaviour: an empirical analysis. *International Journal of Consumer Studies*, 31(1):26–33

Environmental Social Comparisons

Environmental Social Comparisons

Will environmental social comparisons predict?

- "My household has done more to reduce its electricity consumption."
- "My household cares more about the environment."

Yes or No?

Environmental Social Comparisons

Will environmental social comparisons predict?

- "My household has done more to reduce its electricity consumption."
- "My household cares more about the environment."

Yes or No?

No

Environmental Social Comparisons

Will environmental social comparisons predict?

- "My household has done more to reduce its electricity consumption."
- "My household cares more about the environment."

Yes or No?

No

Social norms and social behavior[13]

[13]Reno, R., Cialdini, R., and Kallgren, C. (1993). The transsituational influence of social norms. *Journal of Personality and Social Psychology*, 64(1):104

Items Generated from Mental Models Interviews

Items Generated from Mental Models Interviews

Will items generated from mental models interviews predict?

- "I would participate to promote energy independence in the US."
- "I would participate to avoid wasting energy."

Yes or No?

Items Generated from Mental Models Interviews

Will items generated from mental models interviews predict?

- "I would participate to promote energy independence in the US."
- "I would participate to avoid wasting energy."

Yes or No?

No

Items Generated from Mental Models Interviews

Will items generated from mental models interviews predict?

- "I would participate to promote energy independence in the US."
- "I would participate to avoid wasting energy."

Yes or No?

No

These predicted acceptance of smart-meters well, but not volunteering intentions.

Evaluation of Inputs

- we took stuff from previous propensity models in this area; we took stuff from volunteeringliteratures across various disciplines; we asked people in mental models; we brainstormed and use expert judgment; the hit rate was low
- talk about how unintuitive this stuff is; even for the person making the decision, they have difficulty with insight into what makes them work
- even when we ask people in mental models, they can't figure it out
- we can't rely on expert judgment, people's own theories, etc.

Accuracy

Accuracy

Univariate Prediction

Accuracy

Univariate Prediction

- 15/72 individual items successful (21%).

Accuracy

Univariate Prediction

- 15/72 individual items successful (21%).
- 4/12 constructs successful (PCA).[14]

[14] Loehlin, J. (2004). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. Lawrence Erlbaum

Accuracy

Univariate Prediction

- 15/72 individual items successful (21%).
- 4/12 constructs successful (PCA).[14]
- Best predictor: 12-17% generalization error.[15]

[14] Loehlin, J. (2004). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. Lawrence Erlbaum

[15] Error rate estimated w/ 10-fold CV and bootstrap.
Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall/CRC

Accuracy

Univariate Prediction

- 15/72 individual items successful (21%).
- 4/12 constructs successful (PCA).[14]
- Best predictor: 12-17% generalization error.[15]
  - Whether they thought they would enjoy the display.

[14] Loehlin, J. (2004). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. Lawrence Erlbaum

[15] Error rate estimated w/ 10-fold CV and bootstrap.
Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall/CRC

Accuracy

Accuracy

Logistic Regression

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).
  - Expected enjoyment of the in-home display.

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).
  - Expected enjoyment of the in-home display.
  - Expected learning from the in-home display.

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).
  - Expected enjoyment of the in-home display.
  - Expected learning from the in-home display.
  - Whether they trust the utility.

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).
  - Expected enjoyment of the in-home display.
  - Expected learning from the in-home display.
  - Whether they trust the utility.
  - Whether they are active in their community.

Accuracy

Logistic Regression

- 4/72 items used in the model (1/12 constructs).
  - Expected enjoyment of the in-home display.
  - Expected learning from the in-home display.
  - Whether they trust the utility.
  - Whether they are active in their community.
- 15% generalization error.

Accuracy

Accuracy

Classification Tree

Accuracy

Classification Tree

- 5/72 items used in the model.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.
  - In the home from 6am-10am.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.
  - In the home from 6am-10am.
  - Whether they trust scientists.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.
  - In the home from 6am-10am.
  - Whether they trust scientists.
  - Whether they trust friends.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.
  - In the home from 6am-10am.
  - Whether they trust scientists.
  - Whether they trust friends.
  - Whether they handle unexpected problems well.

Accuracy

Classification Tree

- 5/72 items used in the model.
  - Expected enjoyment of the in-home display.
  - In the home from 6am-10am.
  - Whether they trust scientists.
  - Whether they trust friends.
  - Whether they handle unexpected problems well.
- 10-13% generalization error.

Recruiting non-volunteers

Recruiting non-volunteers

Recruiting non-volunteers

Recruiting non-volunteers

Recruiting non-volunteers

- All participants receive a 5-item survey with their recruitment offer.

Recruiting non-volunteers

Recruiting non-volunteers

- All participants receive a 5-item survey with their recruitment offer.
- Send two postcard follow-ups with at most 5 items.

Recruiting non-volunteers

Recruiting non-volunteers

- All participants receive a 5-item survey with their recruitment offer.
- Send two postcard follow-ups with at most 5 items.
- Sample non-responders for interviews.[16]

[16] Clausen, J. and Ford, R. (1947). Controlling bias in mail questionnaires. *Journal of the American Statistical Association*, 42(240):497–511

Checklist

Checklist

Checklist

Checklist

Checklist

Checklists have been very effective in promoting correct completion of routinized but complicated actions.[17]

[17] Gawande, A. (2010). *The Checklist Manifesto: How to Get Things Right*. Profile Books

Checklist

Checklist

Item	Yes	No

Checklist

Item	Yes	No
Recruitment		
Have recruitment best practices been used?		
Has opt-out been used?		
Have 2 follow-up post-cards/letters been used?		
Do these postcards include 1-5 questions about non-response?		
Has a small up-front incentive ($2) been used?		

Checklist

Item	Yes	No
Recruitment		
Have recruitment best practices been used?		
Has opt-out been used?		
Have 2 follow-up post-cards/letters been used?		
Do these postcards include 1-5 questions about non-response?		
Has a small up-front incentive ($2) been used?		
Inputs		
Have inputs been derived from volunteering research in general?		
Do inputs include constraints, topical interest, expected benefit, and exploratory personality?		
Do inputs include lay theories about volunteering in this context?		

Checklist

Item	Yes	No
Recruitment		
Have recruitment best practices been used?		
Has opt-out been used?		
Have 2 follow-up post-cards/letters been used?		
Do these postcards include 1-5 questions about non-response?		
Has a small up-front incentive ($2) been used?		
Inputs		
Have inputs been derived from volunteering research in general?		
Do inputs include constraints, topical interest, expected benefit, and exploratory personality?		
Do inputs include lay theories about volunteering in this context?		
Estimation		
Has a classification tree been used to discover the necessary questions?		
Has the best model been selected based on generalization error (CV and Bootstrap)?		

Checklist

Item	Yes	No
Recruitment		
Have recruitment best practices been used?		
Has opt-out been used?		
Have 2 follow-up post-cards/letters been used?		
Do these postcards include 1-5 questions about non-response?		
Has a small up-front incentive ($2) been used?		
Inputs		
Have inputs been derived from volunteering research in general?		
Do inputs include constraints, topical interest, expected benefit, and exploratory personality?		
Do inputs include lay theories about volunteering in this context?		
Estimation		
Has a classification tree been used to discover the necessary questions?		
Has the best model been selected based on generalization error (CV and Bootstrap)?		
Sampling Non-Responders		
Do follow-up postcards include 1-5 questions about reasons for non-response?		
Is the reason for non-response survey very short (less than 1 minute)?		
Has a subsample of non-respondents been interviewed?		

Summary

Summary

Summary

Summary

Summary

- Volunteer bias is a serious and unsolved problem.

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...
  - Expectations of benefit specific to the program.

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...
  - Expectations of benefit specific to the program.
  - Constraints on participation (home during the day).

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...
  - Expectations of benefit specific to the program.
  - Constraints on participation (home during the day).
  - Trust in scientists.

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...
  - Expectations of benefit specific to the program.
  - Constraints on participation (home during the day).
  - Trust in scientists.
- Classification Trees perform well, but not much better than the best predictor.

Summary

Summary

- Volunteer bias is a serious and unsolved problem.
- A propensity-type adjustment can yield unbiased results.
- It is difficult to select the right inputs to the model, but so far...
  - Expectations of benefit specific to the program.
  - Constraints on participation (home during the day).
  - Trust in scientists.
- Classification Trees perform well, but not much better than the best predictor.
- Short follow-up postcards and interviews may need only 1 question.

Future Directions

Future Directions

Three Future Directions

Future Directions

Three Future Directions

- Improving theory and the psychology of prediction.

Future Directions

Three Future Directions

- Improving theory and the psychology of prediction.
- Modern alternative to propensity score.

Future Directions

Three Future Directions

- Improving theory and the psychology of prediction.
- Modern alternative to propensity score.
- Actual implementation in a real trial.

Future Directions

Future Directions

Improving theory

Future Directions

Improving theory

- A review paper on the validity of different methods/predictors of volunteering is needed.

Future Directions

Improving theory

- A review paper on the validity of different methods/predictors of volunteering is needed.
- Theory is lacking, especially with respect to more basic psychological models.

Future Directions

Future Directions

The psychology of prediction

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]
- Other-knowledge is limited.[19]

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

[19] Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, 10:173–220

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]
- Other-knowledge is limited.[19]
- Complex models rarely better than simpler ones.[20]

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

[19] Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, 10:173–220

[20] Dana, J. and Dawes, R. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, 29(3):317–331

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]
- Other-knowledge is limited.[19]
- Complex models rarely better than simpler ones.[20]
- Experts (us) do not perform better than laypeople.[21]

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

[19] Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, 10:173–220

[20] Dana, J. and Dawes, R. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, 29(3):317–331

[21] Armstrong, J., Du, R., Graefe, A., Green, K., and House, A. (2012). Predictive validity of evidence-based advertising principles

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]
- Other-knowledge is limited.[19]
- Complex models rarely better than simpler ones.[20]
- Experts (us) do not perform better than laypeople.[21]
  - Advertising, Clinical Psychology, stock market forecasting

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

[19] Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, 10:173–220

[20] Dana, J. and Dawes, R. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, 29(3):317–331

[21] Armstrong, J., Du, R., Graefe, A., Green, K., and House, A. (2012). Predictive validity of evidence-based advertising principles

Future Directions

The psychology of prediction

We've known this type of prediction is difficult for a while.

- Self-knowledge is limited.[18]
- Other-knowledge is limited.[19]
- Complex models rarely better than simpler ones.[20]
- Experts (us) do not perform better than laypeople.[21]
  - Advertising, Clinical Psychology, stock market forecasting
- "suckers will pay for the existence of seers"

[18] Nisbett, R. and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231

[19] Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, 10:173–220

[20] Dana, J. and Dawes, R. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, 29(3):317–331

[21] Armstrong, J., Du, R., Graefe, A., Green, K., and House, A. (2012). Predictive validity of evidence-based advertising principles

Future Directions

Future Directions

A citizen-scientist alternative approach

Future Directions

A citizen-scientist alternative approach

- Mine lay intuition to complement the usual "expert" hypothetico-deductive approach.

Future Directions

A citizen-scientist alternative approach

- Mine lay intuition to complement the usual "expert" hypothetico-deductive approach.
- This allows:

Future Directions

A citizen-scientist alternative approach

- Mine lay intuition to complement the usual "expert" hypothetico-deductive approach.
- This allows:
  1. Understanding the psychology of "folk theories" and prediction.[22]

[22] Keil, F. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, 7(8):368–373

Future Directions

A citizen-scientist alternative approach

- Mine lay intuition to complement the usual "expert" hypothetico-deductive approach.
- This allows:
  1. Understanding the psychology of "folk theories" and prediction.[22]
  2. Possible discovery of the "right theory" by a layperson.[23]

[22] Keil, F. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, 7(8):368–373

[23] Khatib, F., Cooper, S., Tyka, M., Xu, K., Makedon, I., Popovi, Z., Baker, D., and Players, F. (2011). Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, 108(47):18949–18953

Future Directions

A citizen-scientist alternative approach

- Mine lay intuition to complement the usual "expert" hypothetico-deductive approach.
- This allows:
  1. Understanding the psychology of "folk theories" and prediction.[22]
  2. Possible discovery of the "right theory" by a layperson.[23]
  3. Understanding two-way communication between scientists and citizen-scientists.

[22] Keil, F. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, 7(8):368–373

[23] Khatib, F., Cooper, S., Tyka, M., Xu, K., Makedon, I., Popovi, Z., Baker, D., and Players, F. (2011). Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, 108(47):18949–18953

Future Directions

Future Directions

Modern estimation alternative: Semi-supervised learning

Future Directions

Modern estimation alternative: Semi-supervised learning

- Simple example: Self-training[24]

[24] Zhu, X. and Goldberg, A. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers

Future Directions

Modern estimation alternative: Semi-supervised learning

- Simple example: Self-training[24]
  1. Estimate average treatment effect (ATE) based on predictors of volunteering.

[24] Zhu, X. and Goldberg, A. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers

Future Directions

Modern estimation alternative: Semi-supervised learning

- Simple example: Self-training[24]
  1. Estimate average treatment effect (ATE) based on predictors of volunteering.
  2. Find nearest match among non-volunteers in the control and treatment group.

[24] Zhu, X. and Goldberg, A. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers

Future Directions

Modern estimation alternative: Semi-supervised learning

- Simple example: Self-training[24]
  1. Estimate average treatment effect (ATE) based on predictors of volunteering.
  2. Find nearest match among non-volunteers in the control and treatment group.
  3. Add this participant with the estimate from the ATE model.

[24] Zhu, X. and Goldberg, A. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers

Future Directions

Modern estimation alternative: Semi-supervised learning

- Simple example: Self-training[24]
  1. Estimate average treatment effect (ATE) based on predictors of volunteering.
  2. Find nearest match among non-volunteers in the control and treatment group.
  3. Add this participant with the estimate from the ATE model.
  4. Repeat until treatment effect is estimated for all non-volunteers.

[24] Zhu, X. and Goldberg, A. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers

Future Directions

Future Directions

Conducting a real trial

Future Directions

Conducting a real trial

- The approach needs to be validated with real data.

Future Directions

Conducting a real trial

- The approach needs to be validated with real data.
- The approach has the potential to greatly improve the validity of any behavioral trial.

The contents of this talk, including data, analyses, and slides are available:

"Predicting Volunteering for Energy Efficiency Programs,"
`http://hdl.handle.net/1902.1/19154` V2 [Version]

References

Armstrong, J., Du, R., Graefe, A., Green, K., and House, A. (2012).
Predictive validity of evidence-based advertising principles.

Armstrong, J. and Overton, T. (1977).
Estimating nonresponse bias in mail surveys.
Journal of Marketing Research, 14:396–402.

Clausen, J. and Ford, R. (1947).
Controlling bias in mail questionnaires.
Journal of the American Statistical Association, 42(240):497–511.

Cohen, S., Doyle, W., Skoner, D., Rabin, B., and Gwaltney Jr, J. (1997).
Social ties and susceptibility to the common cold.
JAMA: the Journal of the American Medical Association, 277(24):1940–1944.

Dana, J. and Dawes, R. (2004).
The superiority of simple alternatives to regression for social science predictions.
Journal of Educational and Behavioral Statistics, 29(3):317–331.

Davis, A., Krishnamurti, T., Fischhoff, B., and Bruine de Bruin, W. (2012).
Setting a standard for electricity field studies.
Unpublished Manuscript.

Derek, C. and Hashim, R. (1974).
Construct validation of a scale of academic curiosity.
Psychological Reports, 35(1):263–266.

Dunlap, R., Van Liere, K., Mertig, A., and Jones, R. (2000).

New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale.
Journal of Social Issues, 56(3):425–442.

Efron, B. and Tibshirani, R. (1993).
An Introduction to the Bootstrap.
Chapman & Hall/CRC.

Fraj, E. and Martinez, E. (2006).
Ecological consumer behaviour: an empirical analysis.
International Journal of Consumer Studies, 31(1):26–33.

Gawande, A. (2010).
The Checklist Manifesto: How to Get Things Right.
Profile Books.

George et al. (2010).

PG&E load impact evaluation of pacific gas and electric company's time-based pricing tariffs.

Keil, F. (2003).
Folkscience: Coarse interpretations of a complex reality.
Trends in Cognitive Sciences, 7(8):368–373.

Khatib, F., Cooper, S., Tyka, M., Xu, K., Makedon, I., Popovi, Z., Baker, D., and Players, F. (2011).
Algorithm discovery by protein folding game players.
Proceedings of the National Academy of Sciences, 108(47):18949–18953.

Lastovicka, J., Bettencourt, L., Hughner, R., and Kuntze, R. (1999).
Lifestyle of the tight and frugal: theory and measurement.
Journal of Consumer Research, 26(1):85–98.

Loehlin, J. (2004).
Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis.
Lawrence Erlbaum.

Nisbett, R. and Wilson, T. (1977).
Telling more than we can know: Verbal reports on mental processes.
Psychological Review, 84(3):231.

Reno, R., Cialdini, R., and Kallgren, C. (1993).
The transsituational influence of social norms.
Journal of Personality and Social Psychology, 64(1):104.

Rosenthal, R. and Rosnow, R. (1975).
The Volunteer Subject.
John Wiley & Sons.

Ross, L. (1977).
The intuitive psychologist and his shortcomings: Distortions in the attribution process.
Advances in Experimental Social Psychology, 10:173–220.

Sherer, M., James, E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., and Ronald, W.
The self-efficacy scale: Construction and validation.
Psychological Reports, 51.

Summit Blue Consulting (2007).
Evaluation of the 2006 energy-smart pricing plan-final report.

Williams, B., Irvine, L., McGinnis, A., McMurdo, M., Crombie, I., et al. (2007).
When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of

individuals refusing participation in a health-related research project.
BMC Health Serv Res, 7:59–69.

Zhu, X. and Goldberg, A. (2009).
Introduction to semi-supervised learning.
Morgan & Claypool Publishers.