Targeting Energy Interventions

Lecture by Mateus Souza

Universidad Carlos III de Madrid

Winter/Spring 2023

Motivation

From last couple of lectures:

- Some energy interventions/policies have lower-than-expected savings
- In some cases, the interventions are not cost-effective, **on average**
- This is true even when we account for the social cost of carbon

Does this mean that we should abandon these types of policies?

- Not necessarily
- We've also seen that benefits (and sometimes costs) are highly heterogeneous
- A growing literature aims to understand if it is possible to **target** interventions to individuals/households that are associated with higher benefits
- If yes, then we could potentially make the interventions much more cost-effective

Theoretical Framework (Allcott and Kessler, 2019)

In the context of nudges:

$$\max_{x,e} U(\theta) = x + \hat{f}(e; \alpha, \gamma) + (m - \mu e)$$

$$\text{subject to: } y \geq x + ep_e$$

- x is a numeraire good; y is income
- $\hat{f}$ perceived utility from consumption of energy e
- α is consumer heterogeneity
- γ incorporates behavioral biases, inattention, or lack of information
- $(m - \mu e)$ is "moral utility"
- m is energy-independent (dis)utility from the nudges
- μ represents a "moral tax"
- p_e is price of energy

Theoretical Framework (Allcott and Kessler, 2019)

Now suppose that there are two potential scenarios:

- With nudges (θ_1), and without nudges (θ_0)

Let $V(\theta_N)$ denote consumer welfare. The effect of the nudge on consumer welfare can then be written as:

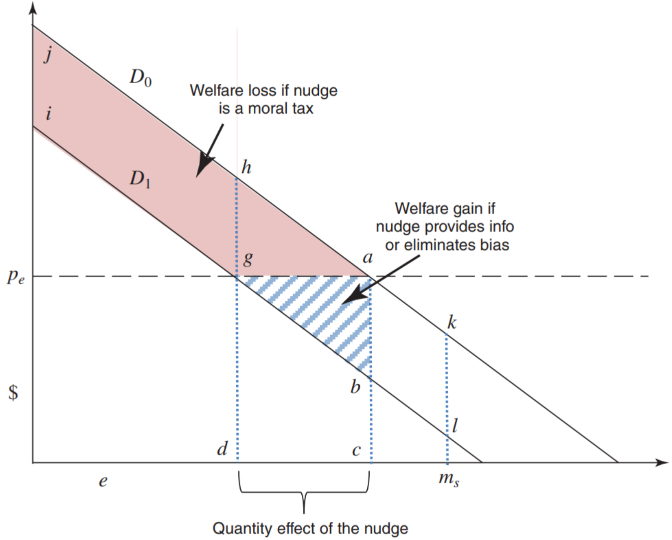
$$\Delta V = V(\theta_1) - V(\theta_0) = -\Delta \tilde{e} p_e + \Delta f + \Delta M$$

Effect on social welfare (W):

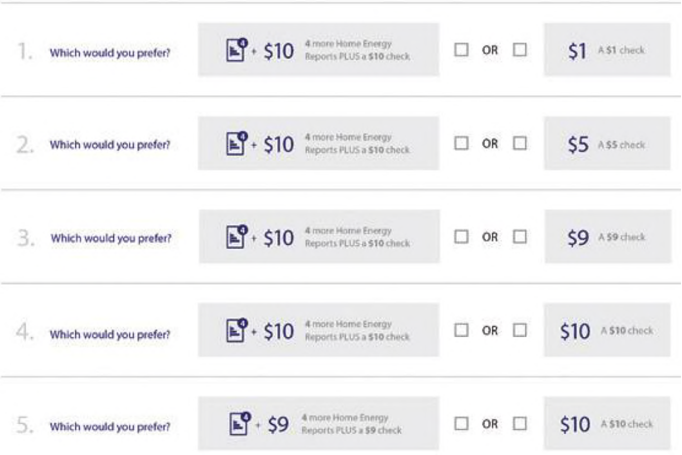
$$\Delta W = \int \Delta V - \phi_e \Delta \tilde{e} \; dF(\Theta) + \Delta \Pi - C_n$$

ϕ_e = environmental externality $\Delta\Pi$ = retailer net revenues C_n = nudge implementation cost

Consumer Welfare

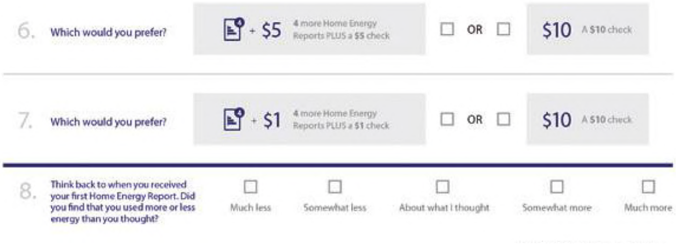


Assessing Willingness-to-Pay



1.	Which would you prefer?	+ $10 4 more Home Energy Reports PLUS a $10 check	☐ OR ☐	$1 A $1 check
2.	Which would you prefer?	+ $10 4 more Home Energy Reports PLUS a $10 check	☐ OR ☐	$5 A $5 check
3.	Which would you prefer?	+ $10 4 more Home Energy Reports PLUS a $10 check	☐ OR ☐	$9 A $9 check
4.	Which would you prefer?	+ $10 4 more Home Energy Reports PLUS a $10 check	☐ OR ☐	$10 A $10 check
5.	Which would you prefer?	+ $9 4 more Home Energy Reports PLUS a $9 check	☐ OR ☐	$10 A $10 check

Assessing Willingness-to-Pay



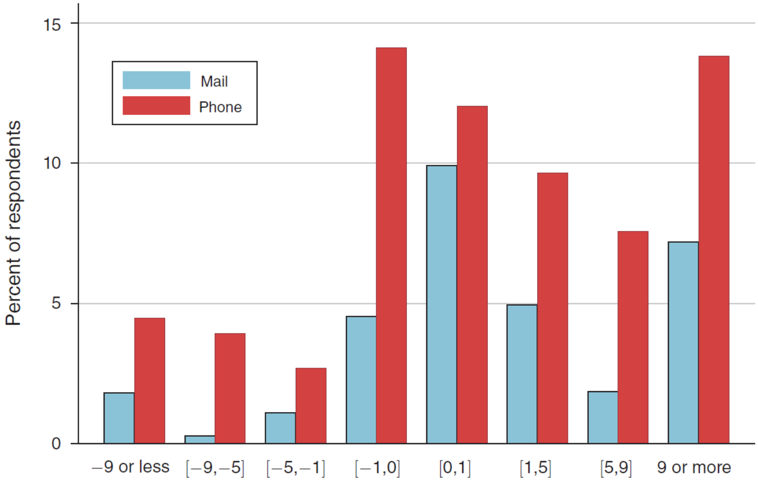
6. Which would you prefer? + $5 4 more Home Energy Reports PLUS a $5 check OR $10 A $10 check

7. Which would you prefer? + $1 4 more Home Energy Reports PLUS a $1 check OR $10 A $10 check

8. Think back to when you received your first Home Energy Report. Did you find that you used more or less energy than you thought?

Much less | Somewhat less | About what I thought | Somewhat more | Much more

Assessing Willingness-to-Pay



Change in Energy Consumption

RCT, randomizing households that receive home energy reports

Table A7: **Effects of Home Energy Reports on Natural Gas Use**

	(1)	(2)	(3)	(4)
Specification:	OLS	IV	IV	IV
Assigned to treatment ×	-0.0228	-0.0228	-0.0230	-0.0249
winter 2014-2015	(0.0117)*	(0.0117)*	(0.0118)*	(0.0128)*
Assigned to treatment×	0.00611	0.00611	0.00618	0.00836
summer 2015	(0.00807)	(0.00807)	(0.00809)	(0.00938)
Assigned to treatment ×	-0.0264			
winter 2015-2016	(0.0115)**			
Assigned to treatment×	0.00581			
summer 2016	(0.0112)			
2nd-year recipient ×		-0.0269	-0.0271	-0.0315
winter 2015-2016		(0.0117)**	(0.0117)**	(0.0121)***
2nd-year recipient ×		0.00592	0.00584	0.00399
summer 2016		(0.0114)	(0.0114)	(0.0110)
Observations	200,540	200,540	200,540	200,540
R^2	0.853	0.853	0.853	0.859
Weights	Duration	Duration	Duration × IPW for $\mathcal{P}_n$	Duration × IPW for $\mathcal{P}_s$

Targeting Home Energy Reports

- Use results from RCT and survey to calibrate model of social welfare

- Counterfactual simulations, changing the homes that receive HER
  - Baseline: 50% of homes, selected at random
  - Opt-in design: only homes that wish to participate
  - Target 50% of homes, based on energy savings
  - Target 50% of homes, based on willingness-to-pay
  - Target 50% of homes, based on welfare

Important:

For targeting, the authors rely on **predictions** of welfare

- Tried elastic net, random forests, and gradient forests
- Cross-validation to prevent overfitting
- *Errors in the predictions decrease the benefits of the targeting approach*

Allcott and Kessler (2019) Targeting Results

TABLE 8—OPT IN AND SMART DEFAULTS: RESULTS

Row	Policy	Percent of population receiving HERs (1)	Mean gas use change (ccf/ recipient-day) (2)	Mean WTP ($/ recipient) (3)	Welfare effect ($/ recipient) (4)	Total welfare effect ($000s) (5)
1	Existing opt-out program	50	−0.027	2.81	0.77	7.7
2	Opt in; zero switching cost	41	−0.027	9.78	7.42	60.7
3	Opt in; 1.5% opt-in rate	1.5	−0.027	24.5	16.93	5.1
4	Targeted on energy savings	50	−0.050	3.08	1.19	11.9
5	Targeted on WTP	50	−0.040	3.30	1.34	13.4
6	Targeted on welfare	50	−0.048	3.42	1.51	15.1

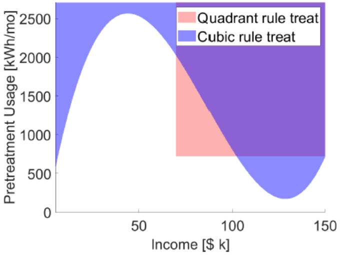
Gerarden and Yang (2022)

Still within the context of Home Energy Reports

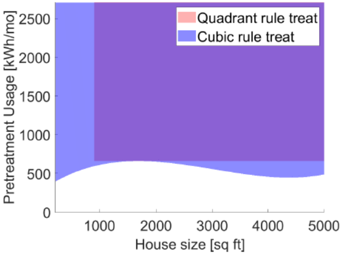
- Employ method from Kitagawa and Tetenov (2018)
- Allows for heterogeneity in savings (conditional treatment effects)
- Results in "simple" treatment rules based on observable characteristics (e.g., income, house size, mean usage)
- Relies on estimates of savings from an RCT

Treatment Rules

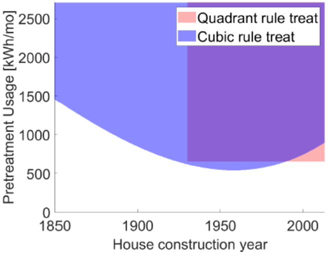
Figure 4: EWM rules for maximizing private cost savings



(a) Household income



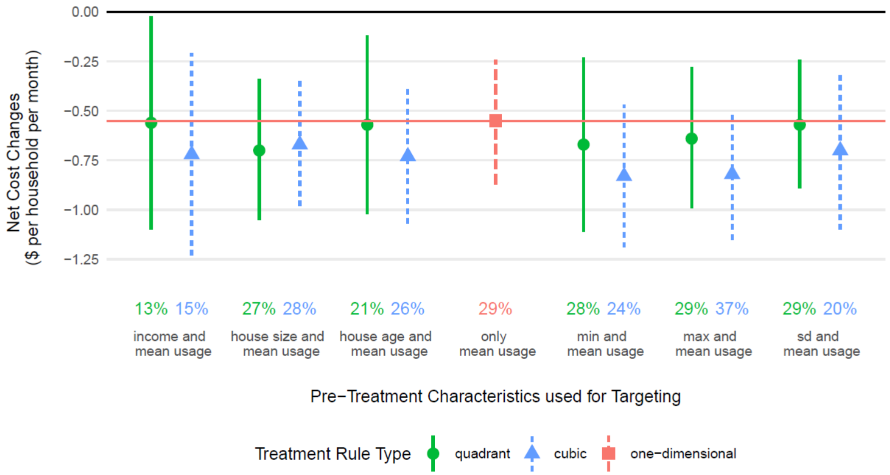
(b) House size



(c) House age

Gerarden and Yang (2022) Targeting Results

Figure 6: Comparison of the gains from targeting with and without using demographic characteristics



Christensen et al. (2021): "Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs"

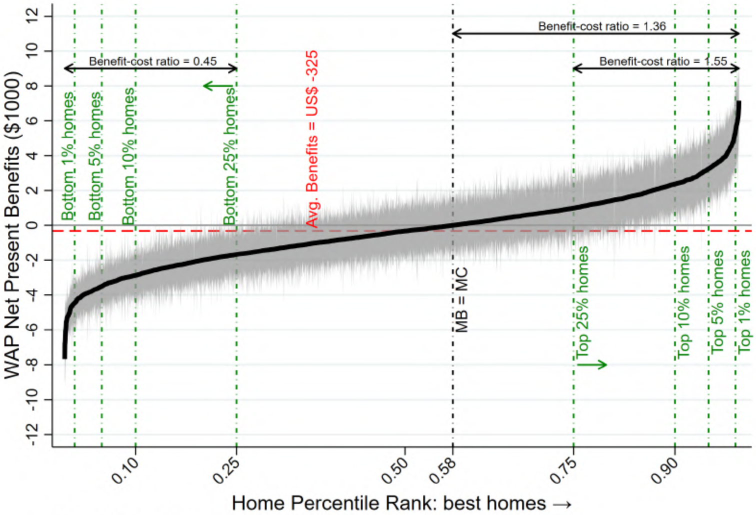
Ex-post evaluation of the Illinois Weatherization Assistance Program

We find:

1. **Low realization rates, on average** $\sim$ **50%** of expected savings
2. Substantial **heterogeneity in benefits**
3. **Modeling bias** explains much of wedge between projected and realized savings

Implication: better modeling may increase cost-effectiveness by improving allocation of funds

Heterogeneity in the Illinois Weatherization Assistance Program



Note: 95% confidence intervals based on bootstrapped standard errors.

Christensen et al. (2022) – Research Questions

Still within the context of the Illinois Weatherization Assistance Program

- Step 1: Using data from already weatherized homes, how accurately can we predict, **ex-ante**, a home's future savings?

Christensen et al. (2022) – Research Questions

Still within the context of the Illinois Weatherization Assistance Program

- Step 1: Using data from already weatherized homes, how accurately can we predict, **ex-ante**, a home's future savings?

- Step 2: What are gains from **targeting**?
  - We compare allocating funds to the most cost-effective homes according to Step 1 versus according to the status quo engineering models

Christensen et al. (2022) – Research Questions

Still within the context of the Illinois Weatherization Assistance Program

- Step 1: Using data from already weatherized homes, how accurately can we predict, **ex-ante**, a home's future savings?
- Step 2: What are gains from **targeting**?
  - We compare allocating funds to the most cost-effective homes according to Step 1 versus according to the status quo engineering models
- Step 3: How much of the gains can be realized without detailed audit data?

Weatherization Assistance Program

- Largest energy efficiency program in the US (over 8 million served since 1976)
- Qualified households: below 200% of poverty line, collect Disability or Supplemental Security Income (SSI), or Temporary Assistance for Needy Families (TANF)
- Qualified homes receive fully-subsidized "retrofits" such as:
  - Wall insulation, attic insulation;
  - Furnace repairs, or even full furnace replacements;
  - Water heater repairs;
  - Door and window replacements.

Weatherization Assistance Program

- Funds allocated using modeling tools based on a set of accepted engineering equations (e.g. US National Energy Audit Tool: NEAT)
- Successful applicant gets pre-treatment energy audit
  - Audit measures inputs to DOE-approved prioritization software
  - List of retrofits optimizes savings-to-investment ratio (SIR)
  - Performed regardless of SIR: health and safety measures, excluded from our analysis

Data and Setting

- Around 13 thousand low-income households from the Illinois Weatherization Assistance Program (WAP)
  - Program years 2009-2016
  - Rich data on: energy audits, housing structure, demographics, contractor IDs
  - Upgrades performed and their costs
  - Engineering projections of savings
- Monthly electricity/gas consumption data collected from utilities serving the whole state, excluding Chicago
- Weather data (min/max temp. and precipitation)
- Energy Prices

Step 1 – Predicting Counterfactuals

Building on the Neyman-Rubin potential outcomes framework, let:

$Y_i(0)$ = home i's energy consumption if NOT treated
$Y_i(1)$ = energy consumption if TREATED
$b_i = Y_i(1) - Y_i(0)$ = energy savings from treatment

What we do, within an **ex-ante** framework:

- Use the available data to predict both $\hat{Y}_i(1)$ and $\hat{Y}_i(0)$
  - Predict ex-ante savings: $\hat{b}_i^{EA} = \hat{Y}_i(1) - \hat{Y}_i(0)$
  - We use machine learning (ML) for prediction
    - Accounts for home/household characteristics, and weather
    - Accommodates complex interactions and nonlinearities

Benchmark for Comparison

- Compare the ex-ante savings to the ex-post estimates from Christensen et al. (2021)
- In the 2021 paper, we use event-study estimates, leveraging all the data available both pre- and post-treatment
- Data from not-yet-treated homes are used to predict untreated counterfactuals
  - Ex-post savings: $\hat{b}_i^{EP} = Y_i(1) - \hat{Y}_i(0)$
- We consider these estimates as the "best we can get" given the available information, thus we use them as the benchmark for comparison

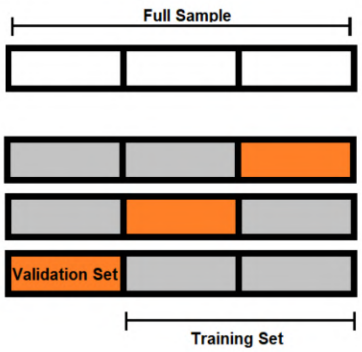
Cross-Validation Design

We use **nested cross-validation** to mimic the role of a program implementer

- They DO have outcome data from previously treated homes
- As well as home/households characteristics and predicted weather for both treated and not-yet-treated (potential target) homes
- Do NOT have outcome data for the potential target homes

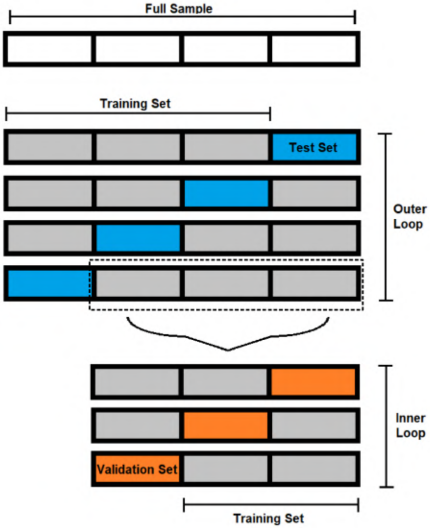
Cross-validation also reduces the bias in the estimation of **out-of-sample errors** (Andrews, Kitagawa, and McCloskey, 2021; Varma and Simon, 2006).

K-Fold Cross-Validation



- Assess the "validation set" prediction errors ($\hat{\varepsilon}_i = \hat{b}_i^{EP} - \hat{b}_i^{EA}$)
- The "best-performing" algorithm has the lowest mean squared error (MSE)

Nested Cross-Validation

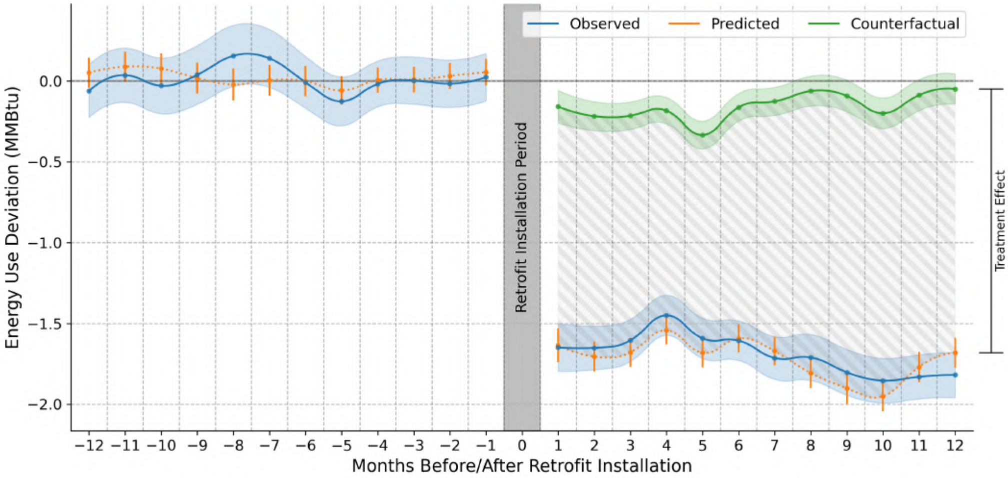


Algorithms Considered

- Lasso, Ridge, and Elastic Net
- Gradient Boosted Trees
- Random Forests
- **Neural Networks**

Algorithm Performance

Visual Inspection of Out-of-Sample Accuracy



Monetizing the Predicted Savings

Estimates of home-specific net present benefits

$$\text{NPB}_i = \sum_{t=1}^{T_i} \left[\frac{\hat{b}_i \times p_t}{(1+r)^t} \right]_i - \text{TotalCost}_i$$

$\hat{b}_i$: predicted energy savings for home i

p_t: social costs of energy in year t

r: discount rate (DOE recommended 3%)

T_i: expected lifespan of retrofits ($\sim$ 30 years)

TotalCost_i: total costs of the retrofits for home i

Monetizing the Predicted Savings

Estimates of home-specific net present benefits

$$\text{NPB}_i = \sum_{t=1}^{T_i} \left[\frac{\hat{b}_i \times p_t}{(1+r)^t} \right]_i - \text{TotalCost}_i$$

$\hat{b}_i$: predicted energy savings for home i

p_t: social costs of energy in year t

r: discount rate (DOE recommended 3%)

T_i: expected lifespan of retrofits ($\sim$ 30 years)

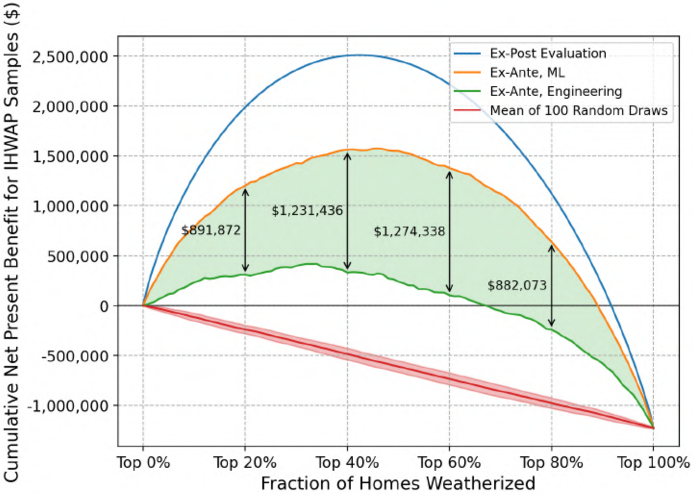
TotalCost_i: total costs of the retrofits for home i

Select projects where $\text{NPB} > 0$

Step 2 – Can We Improve on Status Quo?

- Compare targeting homes with predicted NPB > 0 versus allocation according to status quo engineering model
- Use observed or "ex-post" estimated savings to quantify realized effects (Christensen et al., 2021)
  - Informed by post-treatment data

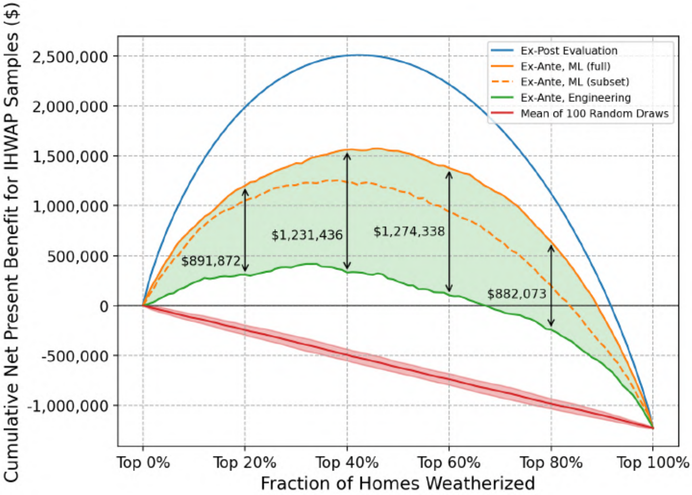
Predictions Based on Previous Homes Outperform Status Quo



Step 3 – Do we need the energy audits for better targeting?

- It might be costly to audit every home that could potentially be treated
- We ask whether targeting can be effective even if we use only a subset of "publicly available" variables
- Subset of variables:
  - income, family size, householder age
  - floor area, building vintage, number of rooms, number of stories, existence of attic, type of heating system
  - location of home (County)

Targeting Without Detailed Audit Data



Robustness Sensitivity

- Predicted versus observed weather
- Different discount rates (2%, 4%)
- Different retrofit lifespans (20 years , 40 years)

Concluding Remarks

- Targeting funds to homes predicted to be cost effective according to historical realized savings increases social net benefits of a dollar spent from \$0.93 to \$1.23.
  - We find increased benefits even when targeting with a limited set of variables

Concluding Remarks

- Targeting funds to homes predicted to be cost effective according to historical realized savings increases social net benefits of a dollar spent from \$0.93 to \$1.23.
  - We find increased benefits even when targeting with a limited set of variables

- We focus on targeting at the home level
  - Targeting measures within a given home could have larger benefits

Concluding Remarks

- Targeting funds to homes predicted to be cost effective according to historical realized savings increases social net benefits of a dollar spent from $0.93 to $1.23.
  - We find increased benefits even when targeting with a limited set of variables
- We focus on targeting at the home level
  - Targeting measures within a given home could have larger benefits
- Audit and selection software could readily accommodate predictions based on realized savings
  - Resulting predictions could be fed into the back end of already established software
  - Billing data only needed for subset of homes representative of those who qualify for the program

Concluding Remarks

- The framework presented in this paper may be useful in settings other than residential energy efficiency
- Targeting can be especially powerful when limited funds need to be allocated within programs that generate substantially heterogeneous benefits
- Other settings with recent advances in targeting:
  - Youth employment programs (Davis and Heller, REStat 2020)
  - Food assistance programs (Finkelstein and Notowidigdo, QJE 2019)
  - Occupational safety and health inspections (Johnson et al., AEJ:Applied Forthcoming)

Main References I

- Allcott, Hunt and Judd B Kessler (2019). "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons". In: *American Economic Journal: Applied Economics* 11(1), pp. 236–76.
- Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey (2021). "Inference on Winners". In: *NBER Working Paper* 25456. https://www.nber.org/papers/w25456.
- Christensen, Peter, Paul Francisco, Erica Myers, Hansen Shao, and Mateus Souza (2022). "Energy Efficiency Can Deliver for Climate Policy: Evidence from Machine Learning-Based Targeting". In: *NBER Working Paper* 30467.
- Christensen, Peter, Paul Francisco, Erica Myers, and Mateus Souza (2021). "Decomposing the Wedge Between Projected and Realized Returns in Energy Efficiency Programs". In: *The Review of Economics and Statistics*. (Forthcoming).
- Gerarden, Todd D. and Muxi Yang (2022). "Using Targeting to Optimize Program Design: Evidence from an Energy Conservation Experiment". In: *Journal of the Association of Environmental and Resource Economists*. (Forthcoming).

Main References II

- Hinton, Geoffrey, Nitish Srivastava, and Kevin Swersky (2013). "Lecture 6E : rmsprop: Divide the gradient by a running average of its recent magnitude". In: *Lecture, COURSERA*. https://www.youtube.com/watch?v=SJ48OZ_qlrc.
- Kitagawa, Toru and Aleksey Tetenov (2018). "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice". In: *Econometrica* 86(2), pp. 591–616. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA13288.
- Varma, Sudhir and Richard Simon (Feb. 2006). "Bias in error estimation when using cross-validation for model selection". In: *BMC Bioinformatics* 7(1), p. 91. URL: https://doi.org/10.1186/1471-2105-7-91.

Variables Included I

	Average	Standard Deviation	Full Model	Subset
Demographics				
Family Income ($)	16,754.27	10,091.63	X	X
Family Size	2.68	1.65	X	X
Female Householder (%)	0.68	0.47	X	X
Householder Age	53.15	15.82	X	X
Renter (%)	0.06	0.24	X	X
County ID (Categorical)	43.95	26.04	X	X
Housing Structure				
Attic R-Value	11.43	10.96	X	
Floor Area (sqft)	1450.3	622.8	X	X
Pre-Retrofit Blower Door (CFM50)	3,648.79	1,786.18	X	
Main Heat Type (Categorical)	2.25	1.15	X	X
Main Heat Age	19.44	14.6	X	
Main Heat Size (BTU)	76,735.14	41,939.71	X	
Main Heat Operational (%)	0.83	0.38	X	
Building Vintage (Categorical)	6	2.44	X	X
Has Air-Conditioning (%)	0.01	0.11	X	
Has Attic (%)	0.7	0.46	X	X
Has Multiple Stories (%)	0.32	0.46	X	X
Num. Bedrooms	2.76	0.98	X	X
Num. Windows	15.12	5.4	X	
Shielding Class (Categorical)	1.85	0.87	X	
Operational Water Heater	0.99	0.12	X	
Water Heater Setting (Categorical)	2.02	0.4	X	

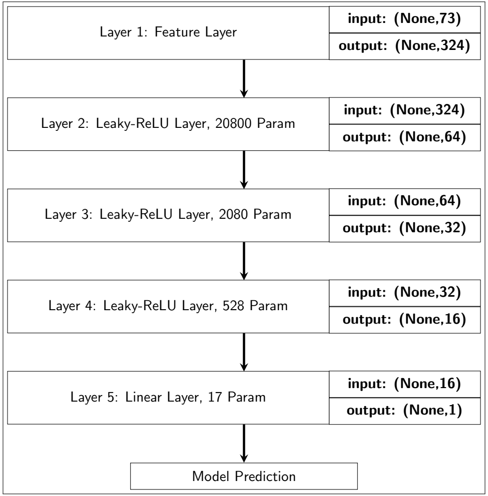
back

Variables Included II

	Average	Standard Deviation	Full Model	Subset
Administrative Variables				
Audit Month	6	3.4	X	
Audit Year	2010	2.29	X	
Retrofit Year	2011	2.21	X	X
Costs ($) per Retrofit Categories				
Air Conditioning	6.8	90.14	X	
Air Sealing	296.78	287.45	X	
Attic	930.71	714.49	X	
Baseload	175.65	232.23	X	
Door	341.58	360.11	X	
Foundation	300.73	500.35	X	
Furnace	1,352.84	1,179.08	X	
General	99.3	488.31	X	
Health and Safety	486.67	334.03	X	
Wall Insulation	274.75	622.03	X	
Window	668.82	890.98	X	
Water Heater	138.02	229.82	X	
Number of Homes in Sample	13,638	-	-	-

back

Neural Network Layers



Leaky-ReLU layers

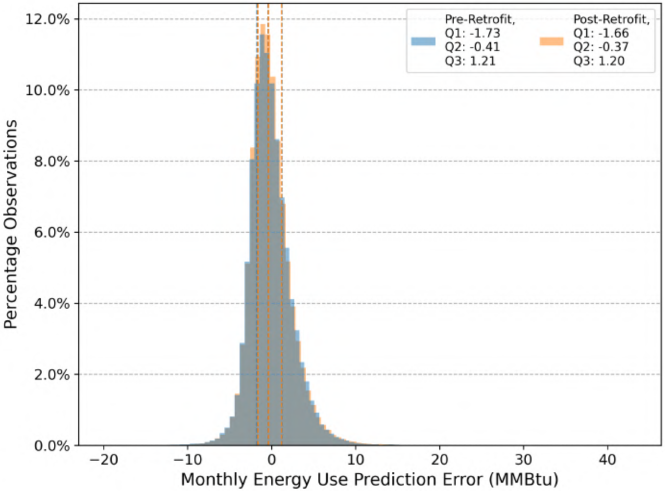
For the leaky-ReLU layers, the $f(\cdot)$ function is non-linear. Specifically, the the output of each neuron in the leaky-ReLU layer is:

$$y = f(\boldsymbol{\beta X}) = \begin{cases} \boldsymbol{\beta X} & \text{if } \boldsymbol{\beta} \cdot \boldsymbol{X} \geq 0 \\ \alpha * \boldsymbol{\beta X} & \text{otherwise.} \end{cases}$$

α was set at 0.3;
an RMSprop optimizer with learning rate equal to 0.00009 is used to find the optimal parameters of the neural network (Hinton, Srivastava, and Swersky, 2013), using mean squared error as the loss function.

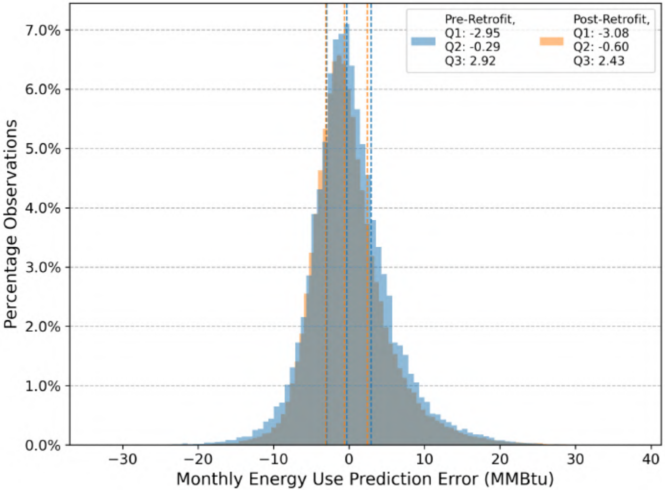
back

Distributions of Prediction Errors – Non-Winter Months



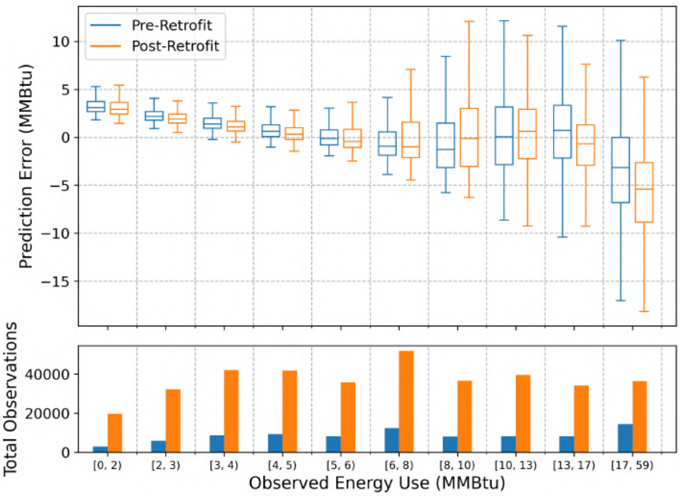
back

Distributions of Prediction Errors – Winter Months



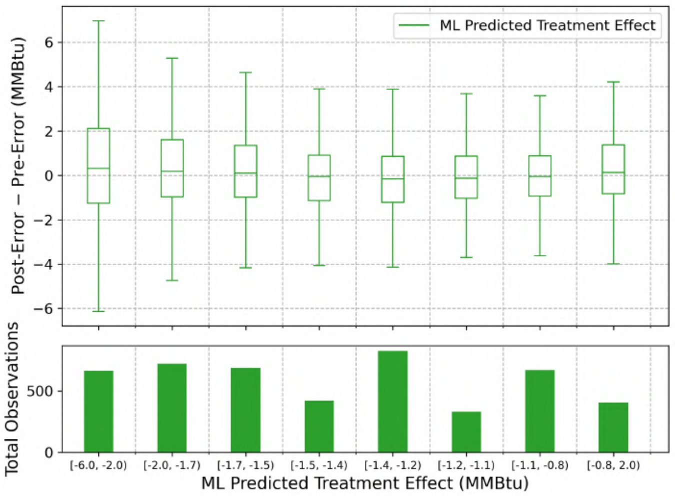
back

Prediction Errors by Usage Bins



back

Difference Between Post- and Pre-Treatment Prediction Errors



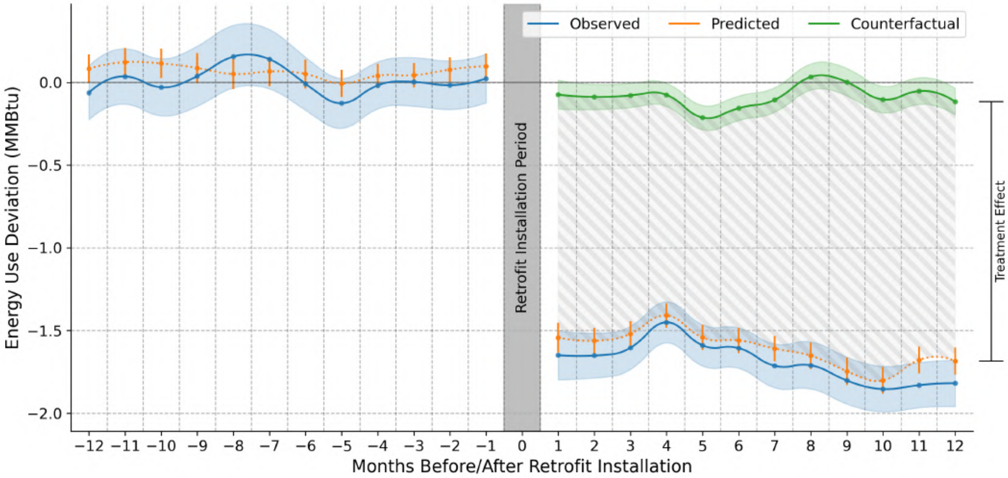
back

Status Quo Predictions

- Engineering models to predict savings
  - Equations relating energy consumption to weather, home, and household characteristics
- Difficult to project impacts
  - Multiple retrofits interacting
  - Diverse buildings
  - Often no access to energy consumption data

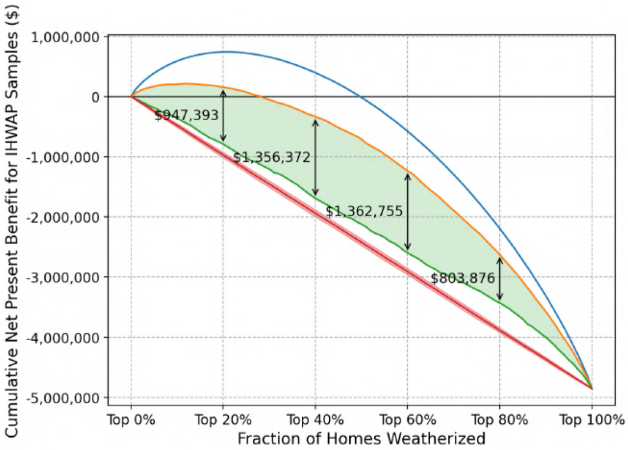
back

Results with Observed Weather



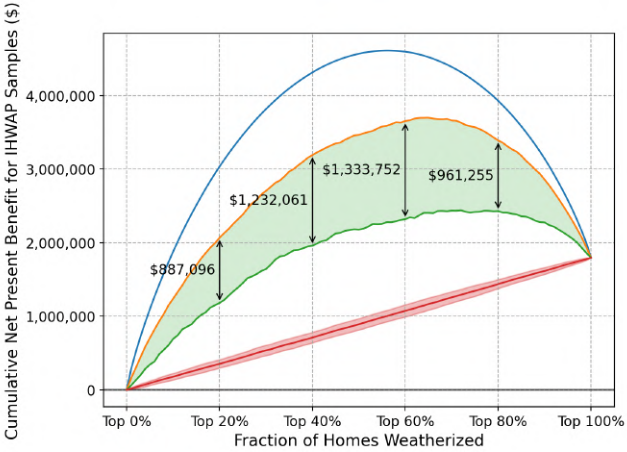
back

20-year lifespan



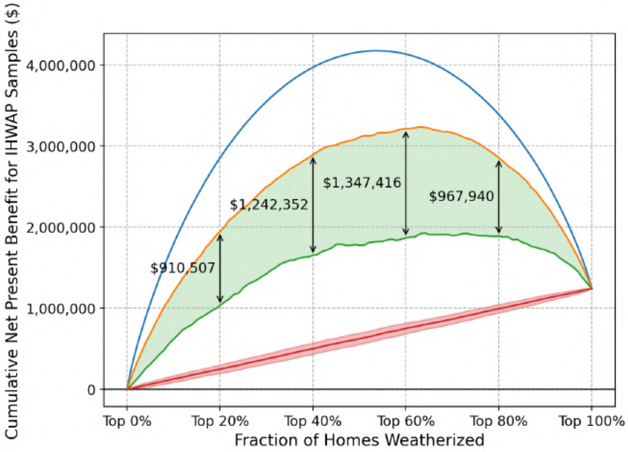
back

40-year lifespan



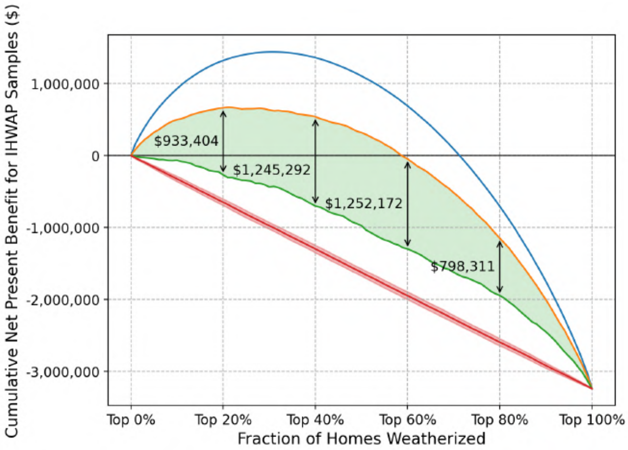
back

2% discount rate



back

4% discount rate



back

Max Benefit-Cost Ratios

	30-year lifespan, varying discount rates			3% discount rate, varying lifespans		
	2%	3%	4%	40 years	30 years	20 years
Full sample BCR (3,913 homes)	1.069	0.932	0.820	1.100	0.932	0.731
Max BCR, ex-post approach	1.457	1.362	1.306	1.477	1.362	1.252
Gains from ex-post targeting	0.388	0.430	0.486	0.378	0.430	0.522
Max BCR, ex-ante ML approach (full)	1.344	1.229	1.145	1.367	1.229	1.063
Gains from ex-ante targeting (full)	0.275	0.297	0.325	0.268	0.297	0.332
Max BCR, ex-ante ML approach (subset)	1.294	1.196	1.125	1.311	1.196	1.024
Gains from ex-ante targeting (subset)	0.225	0.264	0.305	0.211	0.264	0.293
N homes selected, ex-ante ML approach	2,175	1,685	1,189	2,303	1,685	779

back

Performance Metrics for Some Algorithms Considered

Model ID	Model Type	Hyperparameters	MSE (Treatment Effect)	MSE (MMBtu)	MSE (Pre, MMBtu)	MSE (Post, MMBtu)
1	GradientBoosting	boosting stages = 100	47.308	12.652	13.804	12.383
2	GradientBoosting	boosting stages = 120	44.685	12.526	13.639	12.266
3	RandomForest	number of trees = 20	72.722	13.323	14.336	13.086
4	RandomForest	number of trees = 30, max_depth = 4	279.469	16.171	19.468	15.400
5	RandomForest	number of trees = 30	67.992	13.122	14.078	12.898
6	Lasso	alpha = 1	365.241	19.466	25.483	18.059
7	Lasso	alpha = 0.1	109.867	15.717	18.252	15.124
8	Lasso	alpha = 0.01	58.700	14.610	16.949	14.063
9	Lasso	alpha = 0.005	59.786	14.508	16.838	13.963

back

Performance Metrics for Neural Networks

Fold 1 Test Set		Fold 2 Test Set		Fold 3 Test Set		Fold 4 Test Set	
Regularizer	MSE (Treatment Effect)	Regularizer	MSE (Treatment Effect)	Regularizer	MSE (Treatment Effect)	Regularizer	MSE (Treatment Effect)
0.9, 0.9, 0.8	40.3097	0.7, 0.9, 0.8	40.3389	0.6, 0.8, 0.8	39.4183	0.7, 0.8, 0.8	40.8598
0.9, 1.0, 0.8	40.3184	0.7, 0.8, 0.9	40.7823	0.6, 1.0, 0.8	39.1401	0.7, 0.7, 0.8	40.9488
0.9, 0.9, 0.7	40.5537	0.7, 0.8, 0.7	40.1696	0.6, 0.9, 0.9	40.1263	0.7, 0.9, 0.9	41.4700
0.9, 0.9, 0.9	40.7879	0.7, 0.8, 0.6	40.2189	0.6, 0.9, 0.7	38.6978	0.7, 0.9, 0.7	40.2058

back