Take the Load Off:
Effort and Technology as Determinants of Electricity Demand Response

Megan Bailey (University of Calgary)
David Brown (University of Alberta)
Blake Shaffer (University of Calgary)
Frank Wolak (Stanford University)

Toulouse, June 2026

The electricity system of the past involved forecasting demand and dispatching supply; the grid of the future will increasingly involve forecasting supply and dispatching demand.

- Jeff Nagle, Pacific NW National Lab

How to unlock flexible demand?

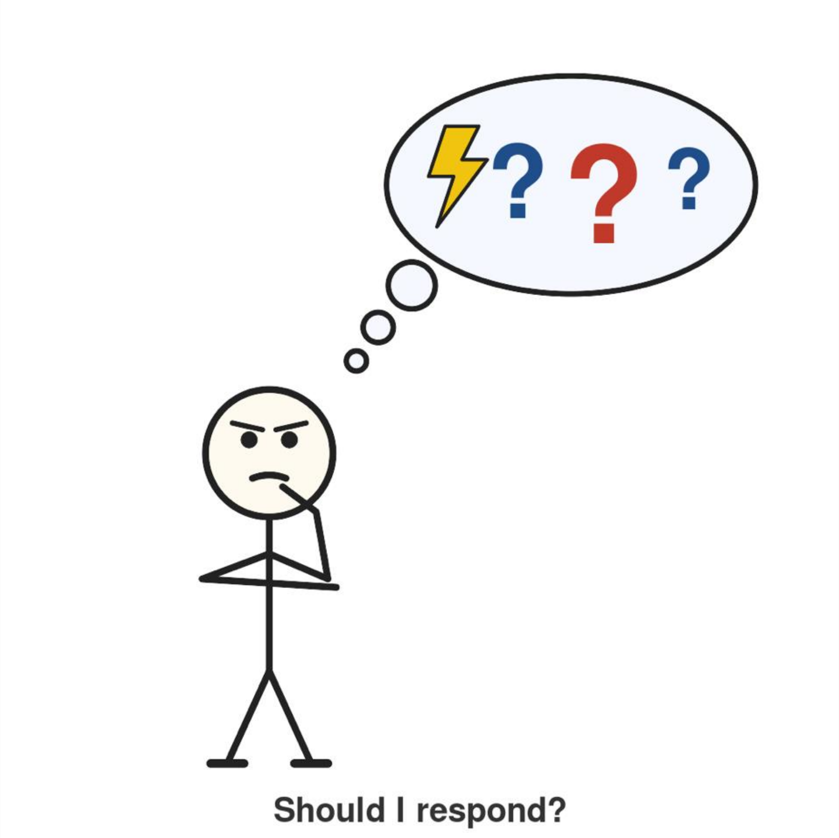
- **Benefits** of time-varying pricing are well-known (Boiteux, 1960; Borenstein and Holland, 2005)
- Some evidence households **respond** to time-varying rates (Faruqui et al., 2014; Harding and Sexton, 2017)
- But there are **barriers** to responsiveness:
  - Adjustment costs and information acquisition (Fabra et al, 2021)
  - Inattention and behavioural biases (Sallee, 2014; Schneider and Sunstein, 2017)
  - Evidence that households have a difficult time understanding marginal prices and complexity (Ito, 2014; Shaffer, 2020; Jacobsen and Steward, 2022)
- **Technology** should help! But it might not... (Brandon et al, 2022)
- Evidence **automation** can assist in demand response when combined with pricing (Blonz et al., 2025; Bollinger and Hartman, 2020)

Our Paper...

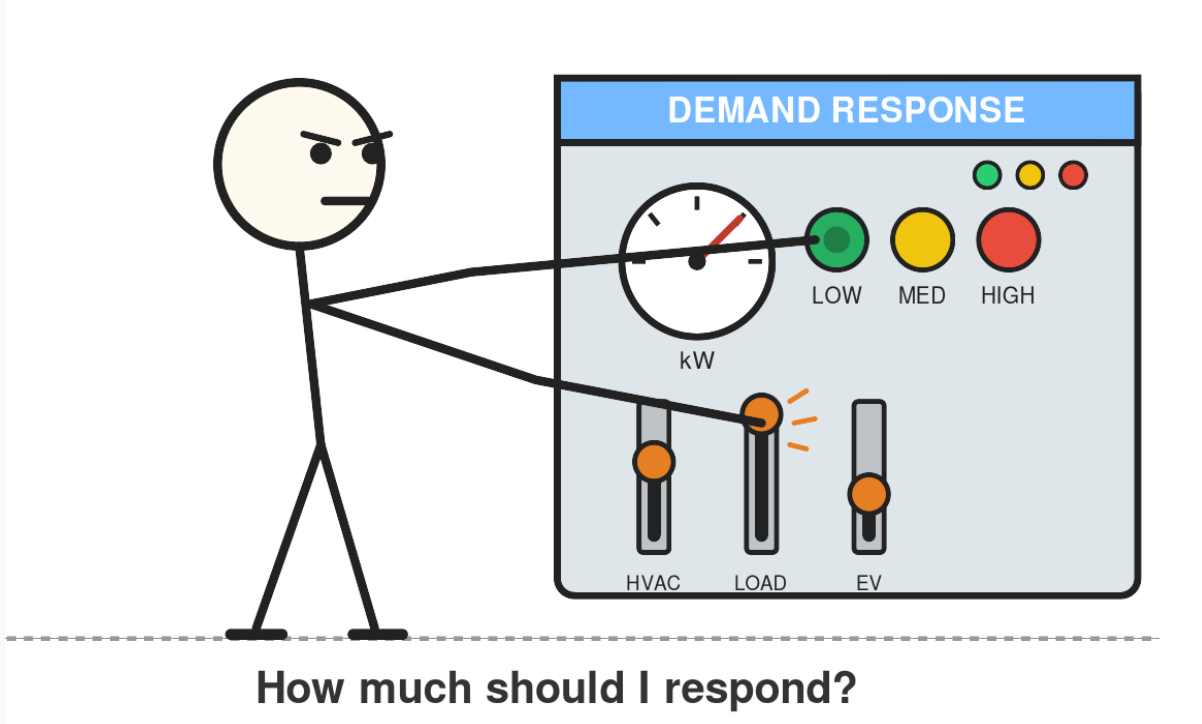
- **How much of the barrier to demand response is *effort*, and can *technology* overcome it?**
- **Randomized control trial (framed field experiment)**
- **Within-subject design**
  - *Individually* randomize “peak events”
  - Estimate **household-specific** treatment effects
- Various programs to better understand the **barrier of effort** and **benefit of technology**:
  1. Behavioural with **manual** response (*active*)
  2. Behavioural with **load control tech** (*active*)
  3. Automated with **centralized** load control (*passive*)

Two stages of consumer response

1. *Should I respond to this event?*



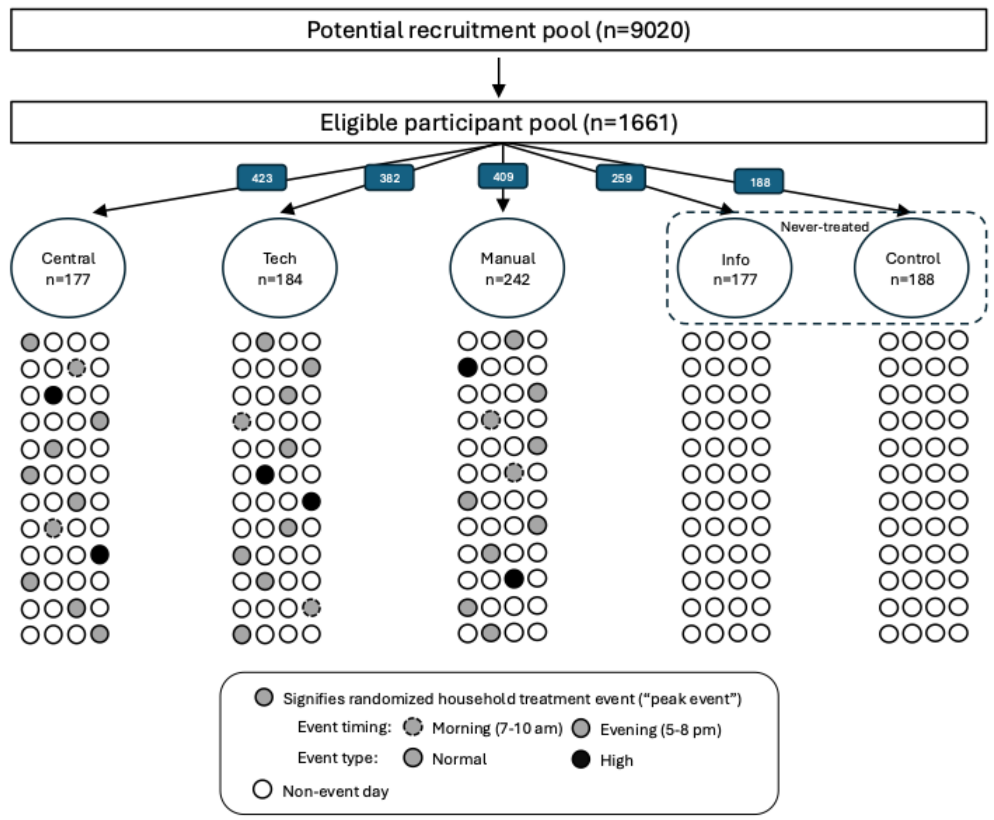
2. *Conditional on responding, by how much should I respond?*



The Experiment...

What we do

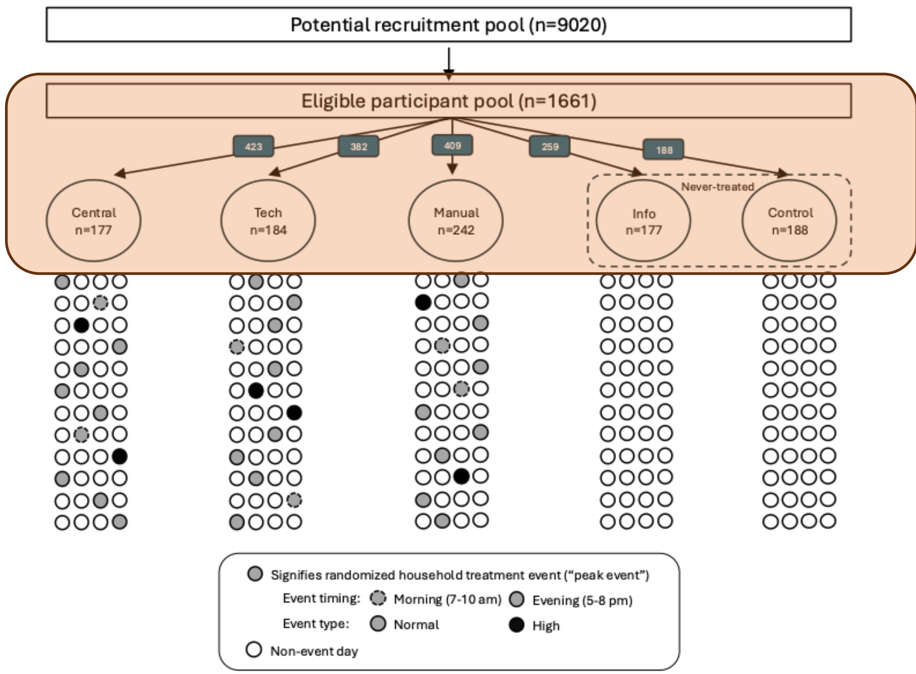
- Partnered with a large Canadian electric utility
- Recruited ~1700 households from utility's App
- Step 1: Randomize assignment to demand response programs
- Step 2: Randomize peak events to individual participants



Step 1: Assign to DR Programs

- Randomize offers to eligible HHs into several programs:

	DR Initiator	Load Control Tech	Price Incentive	Usage Info
Central	Utility	✓	✓	✓
Tech	HH	✓	✓	✓
Manual	HH		✓	✓
Info	HH			✓
Control	HH			



- Central vs Tech: Effect of **passive vs active** response
- Tech vs Manual: Effect of **enabling technology**

Step 1: Program Acceptance

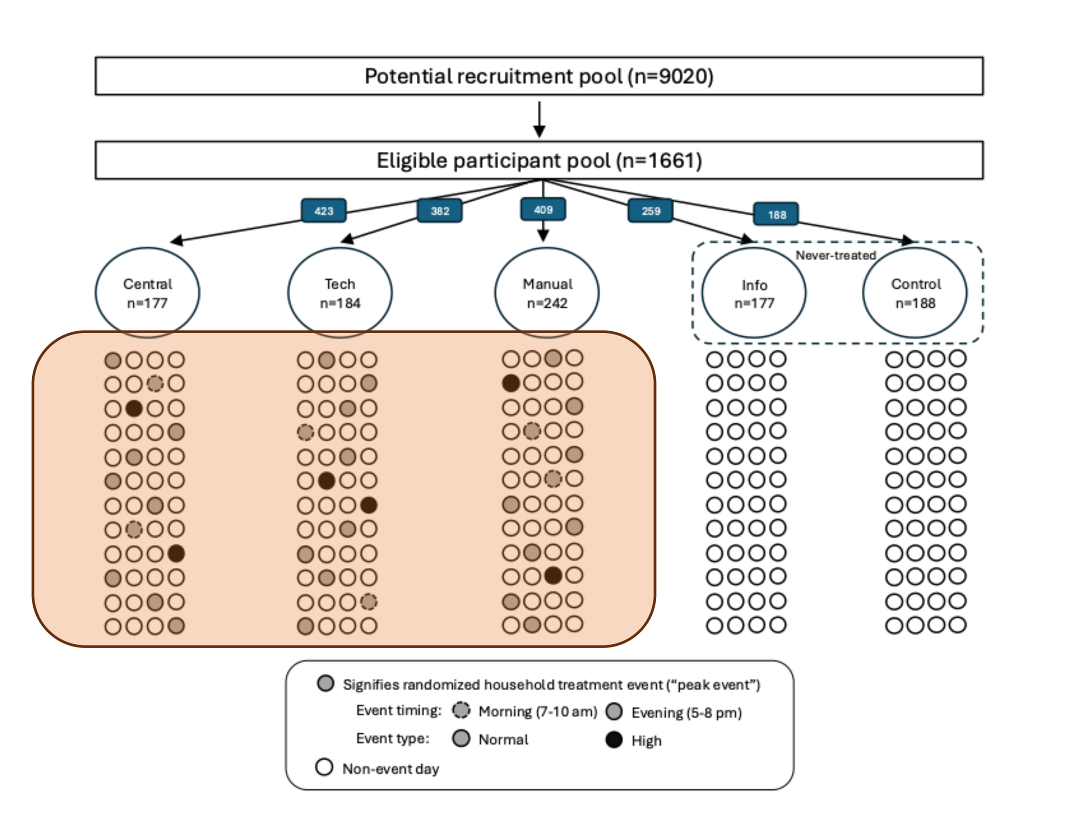
	Central	Tech	Manual	Info	Control
Invited	423	382	409	259	188
Accepted	177	184	242	177	188
	(42%)	(48%)	(59%)	(68%)	(100%)

- Central and Tech take-up rates roughly similar
- Manual take-up rate is higher

Step 2: Randomize peak events

- ***Individually*** randomize notification of "peak events"
- Every participant gets a different schedule

- Times: Morning (7-10am), Evening (5-8pm)
- Incentives:
  - Measured as reductions vs personalized baselines
  - "Regular": (10%, 30%, 50%) ? ($1, $2, $3)
  - "High": (10%, 30%, 50%) ? ($1, $3, $6)
- Frequency: ~3 events per month over 17 months
- Notifications: via App (21 and 2 hours prior to event)



User Experience: Peak Event Messaging

Central group 21hr notification

32m ago

PM High Peak Event Monday, 5pm-8pm
[redacted] will adjust your connected devices to reduce peak electricity usage.

Reduce by 10%, Earn $1
Reduce by 30%, Earn $3
Reduce by 50%, Earn $6

Tech/Manual group 21hr notification

32m ago

PM High Peak Event Monday, 5pm-8pm
Take action to reduce electricity usage by turning off/down appliances.

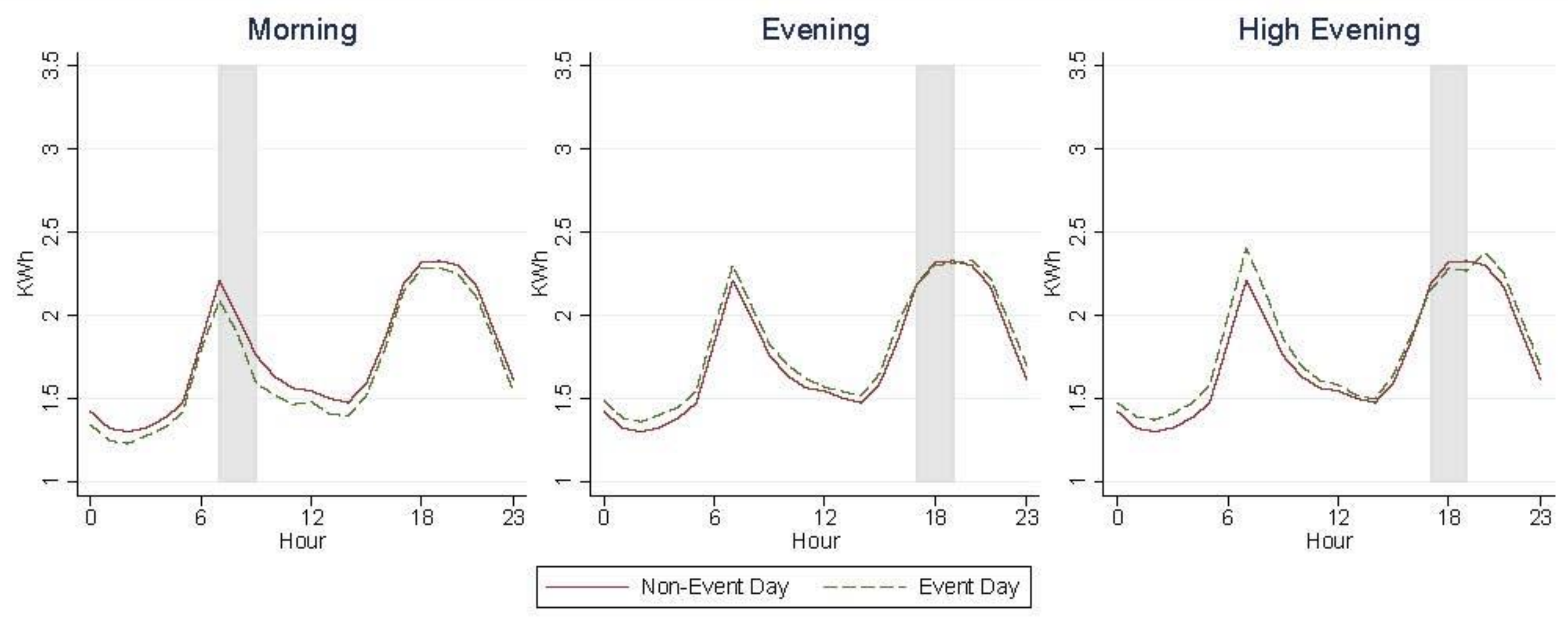
Reduce by 10%, Earn $1
Reduce by 30%, Earn $3
Reduce by 50%, Earn $6

Data

- Household-level **hourly meter consumption** data
  - September 2021 – June 2023
  - **Over 15 million observations**
- Time-stamped **App interaction** data
- Survey data on household **demographics**
- Hourly **weather** data (temp, humidity)

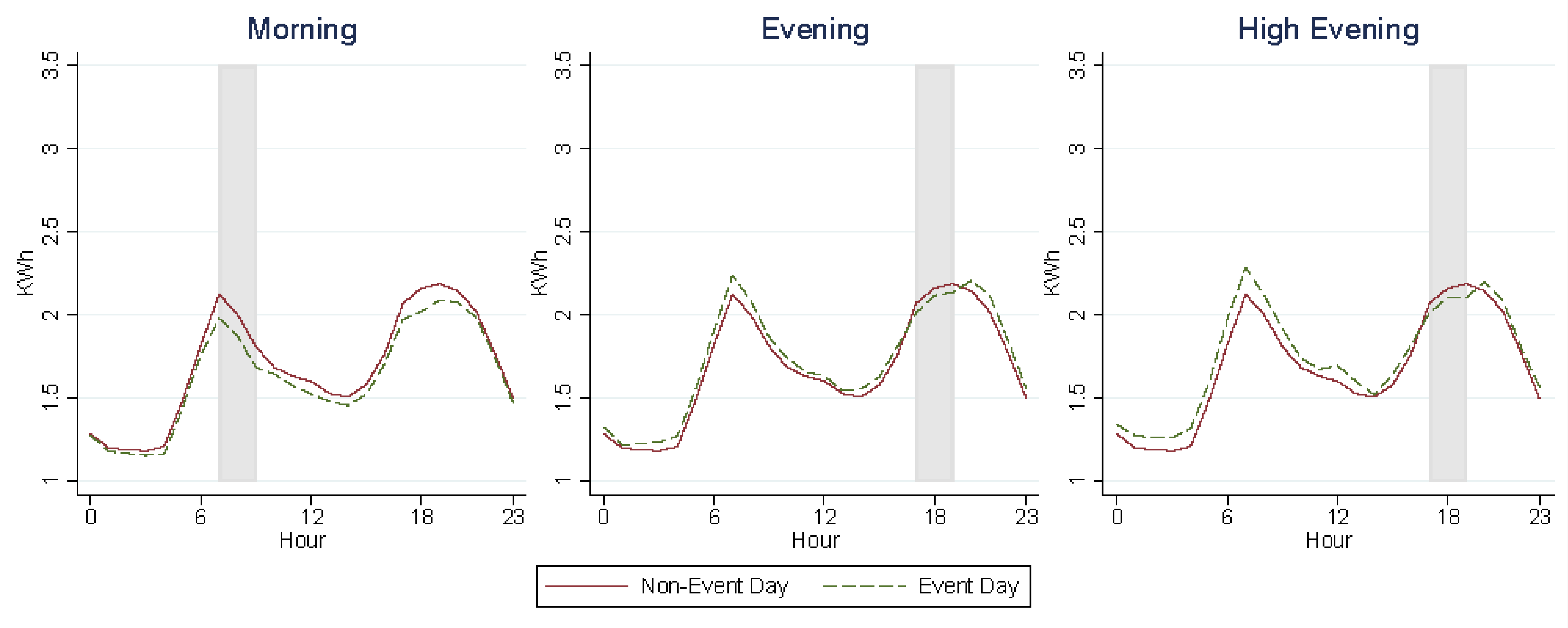
Descriptive Analysis

Manual group: Non-event vs Event days



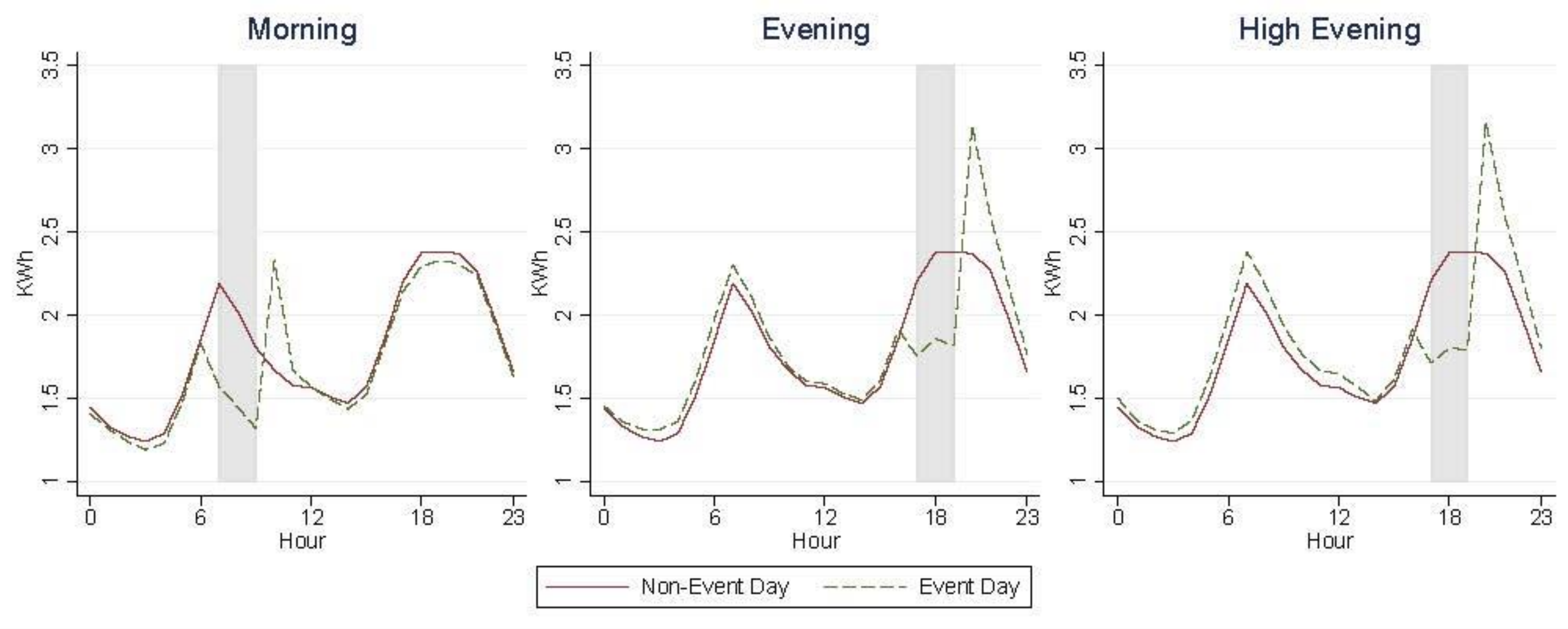
- Minimal difference between Event and Non-Event

Tech group: Non-event vs Event days



- Minimal difference between Event and Non-Event

Central group: Non-event vs Event days



- Large difference between Event and Non-Event; evidence of snapback

Regression Analysis

Program-level estimation

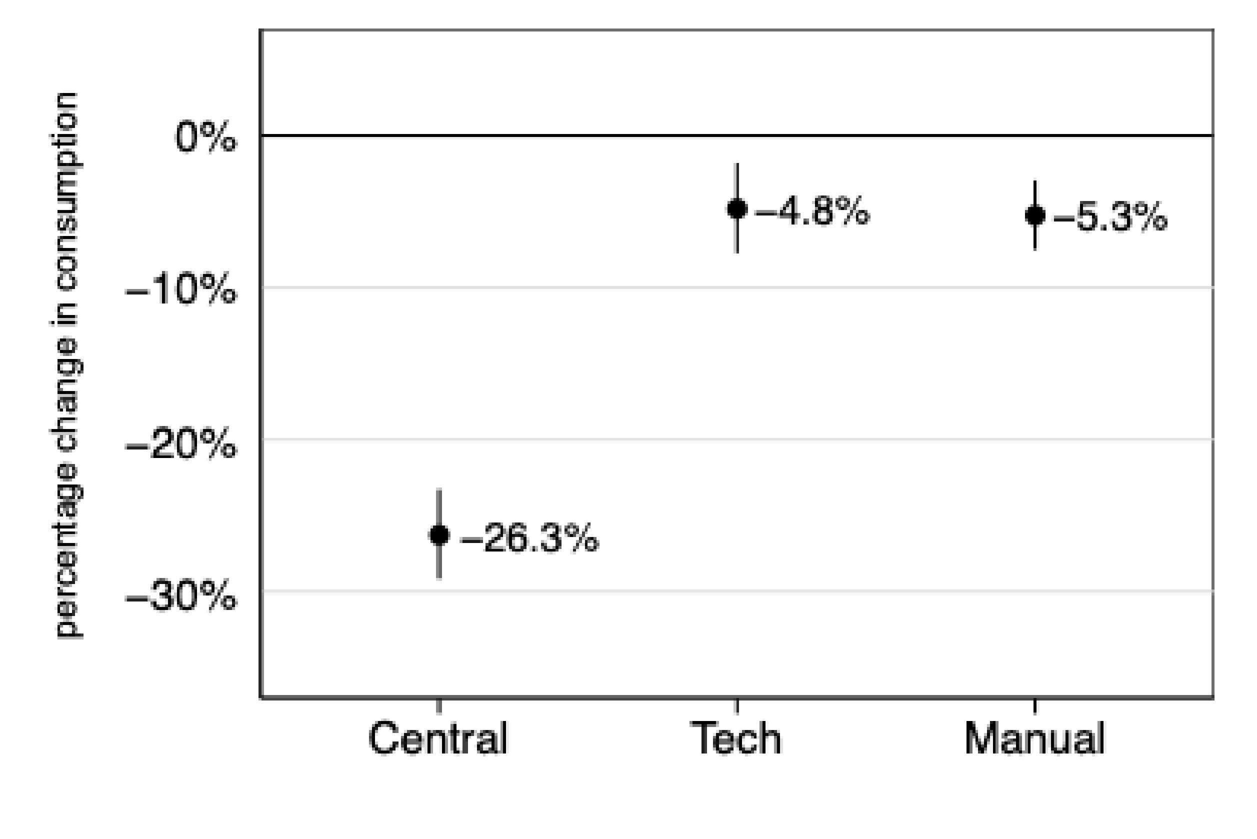
- Compare event times to non-event, controlling for household and time fixed effects, weather variables, and heterogeneity by program

$$\ln(c_{it}) = \sum_{j \in \{C,T,M\}} \beta_j E_{it} Program_{ji} + \alpha_i + \tau_t + \delta \boldsymbol{X_{it}} + \epsilon_{it}$$

- c_{it}: hourly household consumption (in kWh)
- E_{it}: event indicator variable
- $Program_{ji}$: program indicator variables
- $\boldsymbol{X_{it}}$: hourly weather variables (CDD, HDD, humidity, incl. non-linear terms)
- α_i and τ_t: Fixed effects for households and hour-of-sample
- Cluster SEs at HH level

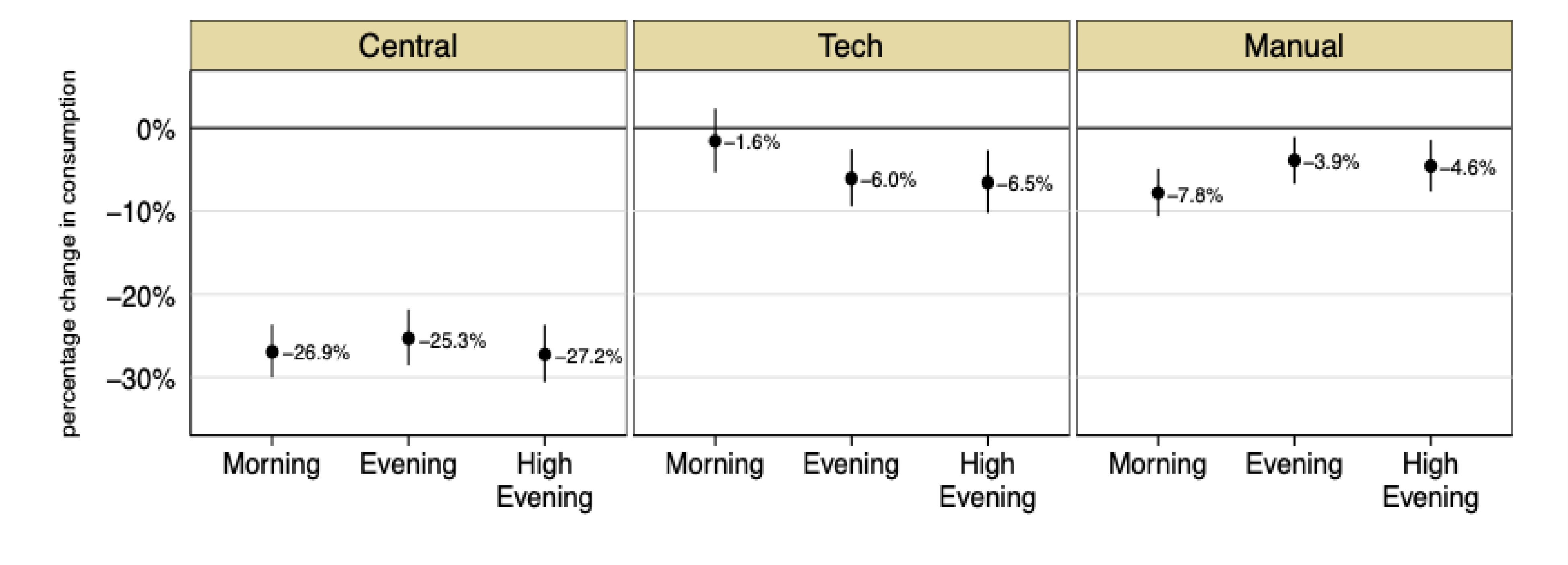
- Various control groups for robustness:
  - Within-program non-event
  - Other-program non-event
  - Pure control groups

Results: Consumption changes by program



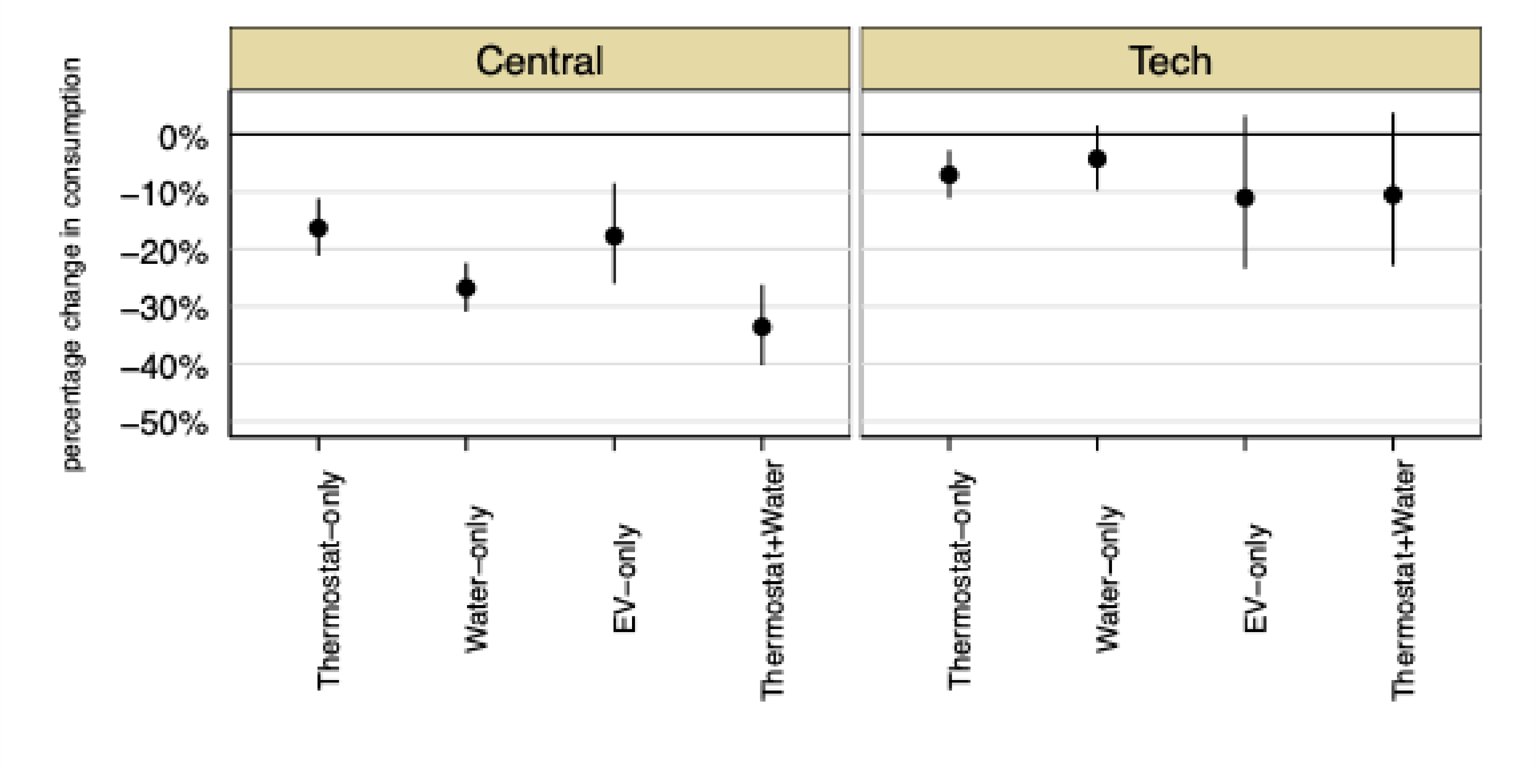
- Modest reductions in manual group
- Tech no better than manual
- Central performs 5x better

Results: Consumption changes by event type



- More money doesn't increase responsiveness
- Consistent with (Prest, 2020)

Results: Consumption changes by device composition



- Central performs across the board
- Biggest difference is hot water heaters
- EV households show some responsiveness in Tech
- More devices = more response

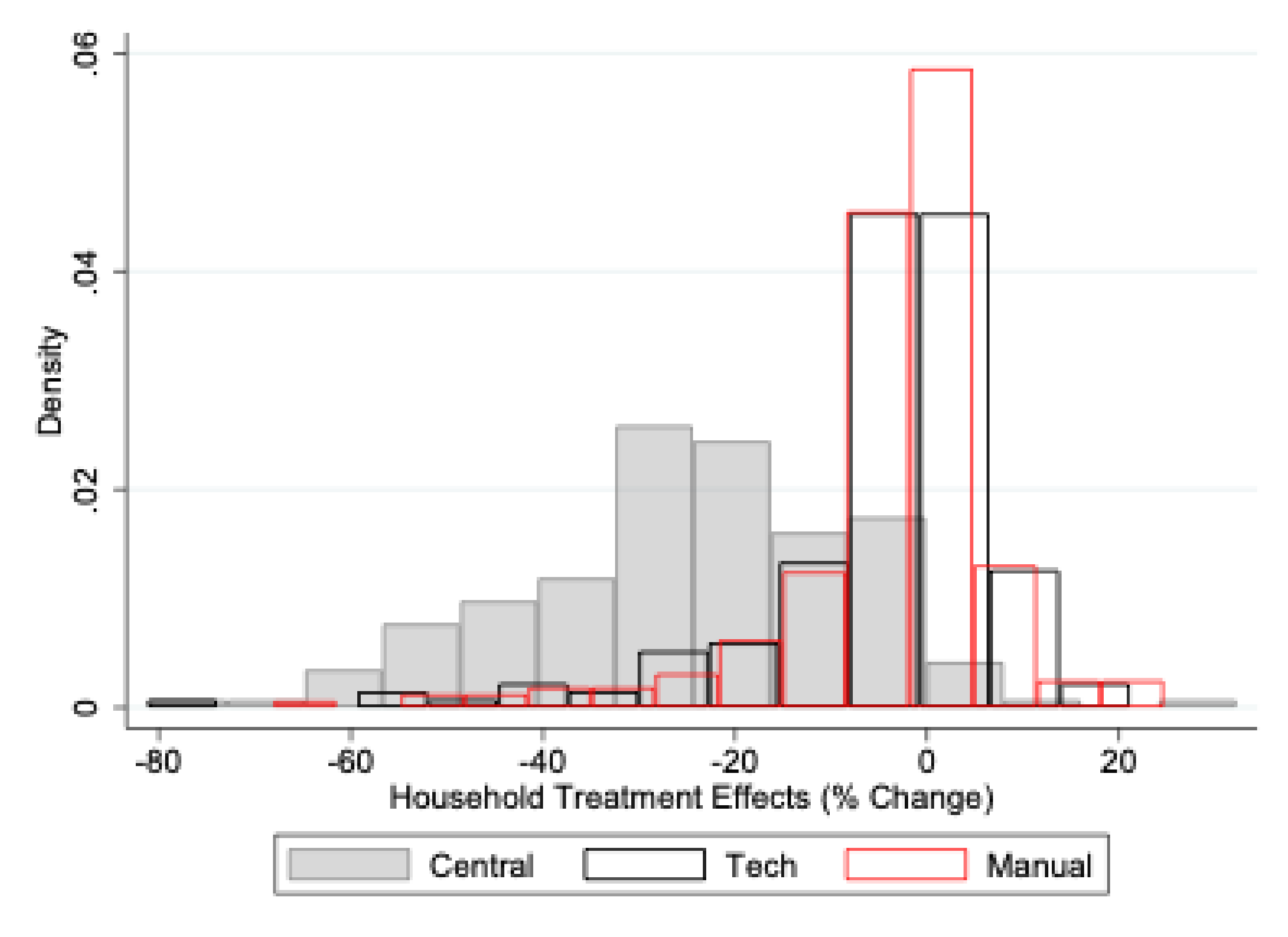
Household-level estimation

- Time-series regression, leverages *within-subject* variation only
- Compare event times to non-event within a household controlling for time fixed effects and weather:

$$\ln(c_{it}) = \gamma_i + \beta_i E_{it} + T_t + \delta_i \boldsymbol{X}_{it} + \eta_{it}$$

- $\hat{\beta}_i$: household-specific treatment effects
- **Estimate the full distribution of peak event responses**

Results: Household-level estimates



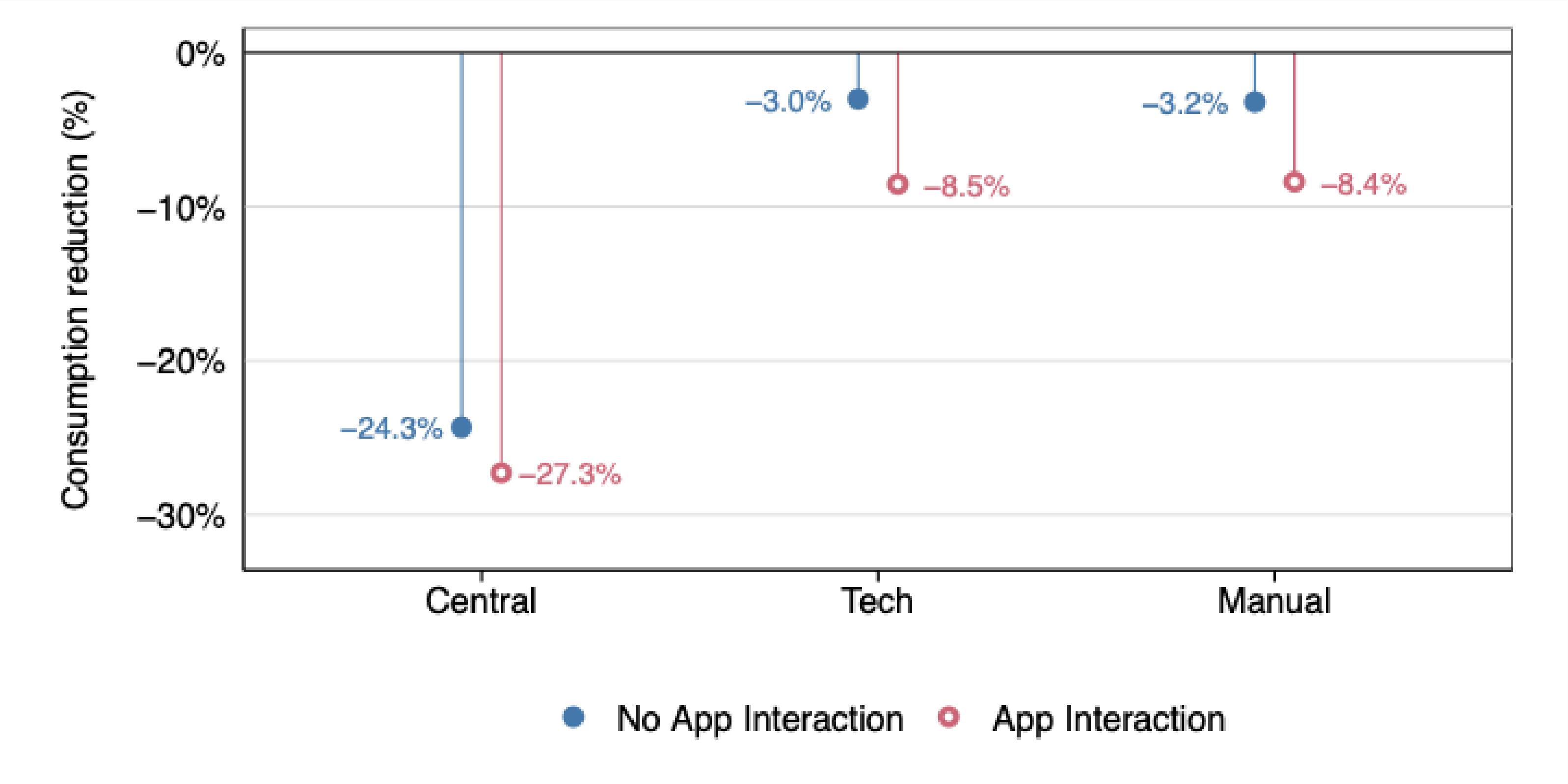
- **Tech and Manual:**
  - Modal response near 0
  - Mean response 4.6% and 4%
  - Some large outliers (big responders)
  - 20% of HH estimates significant
- **Central:**
  - Mean response 24%
  - 80% of HH estimates significant

Mechanisms: Attention and Effort

Mechanisms: Attention and Effort

- What is driving differences both *within* programs and *across* programs?
- Question: ***Within a household, do we see a differential response by whether or not they interact with the App?***

Results: App Interaction by Program



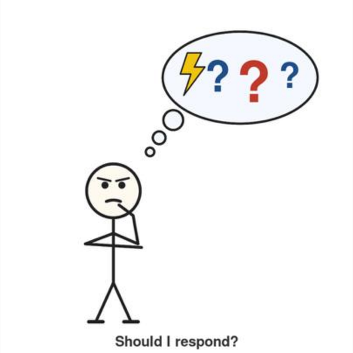
- When households interact with App → larger response
- **Central group has considerable “head start” regardless of App Interaction**

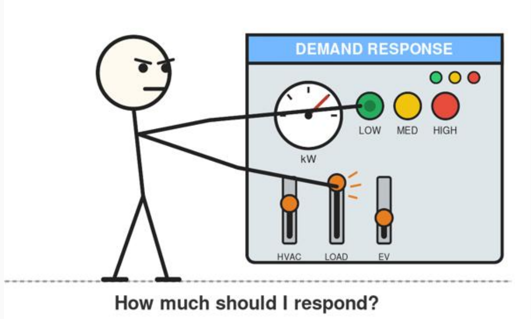
Decomposing Responsiveness

Decomposing responsiveness

$$Overall\ Demand\ Reduction =$$

$$(Likelihood\ of\ responding) * (Magnitude\ of\ Response \mid Responding)$$





Step 1: Identifying responders

- Approach: Use HH-level estimates to find HHs with estimated treatment effects that are negative (i.e. reductions) with p-value < 0.01
  - i.e. these are “always responding”

Program	Count “responders”	Participants in program	Proportion of “responders”
Manual	36	242	**15%**
Tech	32	184	**17%**
Central	136	177	**77%**

Step 2: Find the conditional effect

- Use the HH-level estimates and calculate the mean treatment effect of those flagged as responders, by group

Program	Count "responders"	Participants in program	Proportion of "responders"	Mean treatment effect ($\hat{\beta}$)	Overall % reduction in demand
Manual	36	242	15%	**-0.26**	-3.8%
Tech	32	184	17%	**-0.31**	-5.2%
Central	136	177	77%	**-0.38**	-25.5%

- Central group not only increases the *likelihood of responding*, it also increases the conditional <u>*magnitude of response*</u>!

Conclusion

Summing up

→ We ran a large field experiment to explore the role of effort and technology in demand response

- **We find:**
  1. Take-up rates similar between Central and Tech (both lower than manual)
  2. Central group dominates in responsiveness (5x greater)
  3. Tech group no better than Manual
  4. Water heaters are largest beneficiary of automation
  5. Some evidence of EV responsiveness without central control
  6. App interactions are correlated with responsiveness
- **Key takeaways:**
  1. **Central boosts *likelihood of responding*** **(confirms Attention and Effort are key barriers)**
  2. **Conditional on responding, Central group *also* performs better! (larger magnitude)**
- **Next question:** Willingness to hand over control???

Thank you!

blake.shaffer@ucalgary.ca
www.blakeshaffer.ca

Appendix slides

Balance table

Table: Comparison of Means by Group - Final Accepted Households

	Central	Tech	Manual	Info	Control	ANOVA (p-value)
Cumul. KWhs						
Winter	5,507	5,302	5,422	5,037	5,265	0.71
	(2,706)	(2,737)	(3,240)	(2,768)	(2,950)	
Spring	3,900	3,739	3,797	3,642	3,712	0.85
	(1,934)	(1,791)	(1,939)	(2,159)	(1,974)	
Summer	2,851	2,672	2,766	2,702	2,729	0.93
	(1,869)	(1,759)	(1,547)	(1,849)	(1,710)	
Fall	3,754	3,550	3,677	3,547	3,623	0.86
	(1,733)	(1,659)	(1,992)	(1,788)	(1,860)	
Load Factor						
Winter	24.62	25.56	24.93	24.93	24.67	0.90
	(8.68)	(8.21)	(9.04)	(7.76)	(8.63)	
Spring	19.33	20.48	19.80	19.59	19.91	0.72
	(7.43)	(6.30)	(6.45)	(7.05)	(7.41)	
Summer	16.33	16.87	16.95	16.61	16.32	0.91
	(8.54)	(6.02)	(5.95)	(7.80)	(8.29)	
Fall	18.17	18.97	19.11	18.53	19.06	0.65
	6.27	5.78	6.29	6.11	6.50	
Electric Vehicle	0.25	0.21	0.30	0.34	0.27	0.07*
	(0.43)	(0.41)	(0.46)	(0.47)	(0.45)	
Baseboard Heating	0.68	0.70	0.60	0.59	0.63	0.12
	(0.47)	(0.46)	(0.49)	(0.49)	(0.48)	
Air Conditioning	0.46	0.46	0.51	0.51	0.54	0.41
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	
Electric Hot Water	0.75	0.74	0.68	0.65	0.72	0.16
	(0.43)	(0.44)	(0.47)	(0.48)	(0.45)	
House/Duplex	0.82	0.77	0.89	0.84	0.84	0.02**
	(0.39)	(0.42)	(0.32)	(0.37)	(0.37)	
Median Income	84,978	88,274	86,718	89,504	85,948	0.23
	(19,647)	(20,432)	(19,494)	(21,079)	(21,541)	

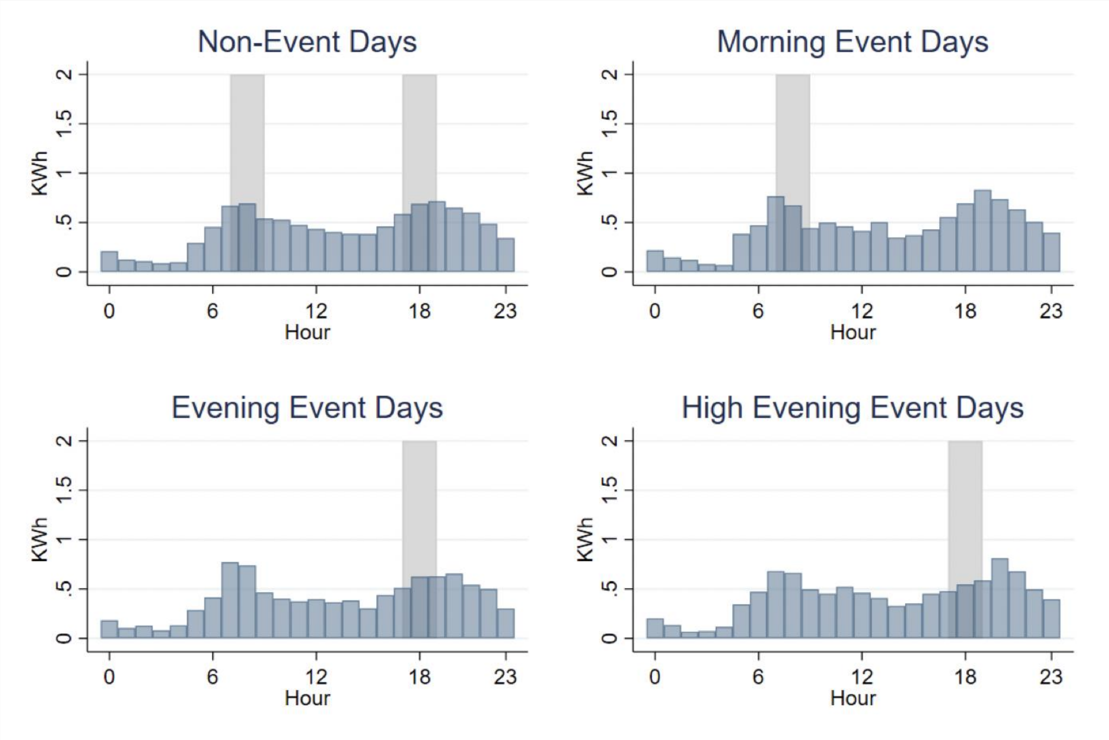
App Interaction and Performance

Program	Performance Quartile	Household Count	General Interact
Central	1	112 (63%)	0.32
	2	42 (24%)	0.20
	3	16 (9%)	0.20
	4	7 (4%)	0.15
Tech	1	20 (11%)	0.54
	2	45 (25%)	0.30
	3	52 (28%)	0.23
	4	67 (36%)	0.13
Manual	1	19 (8%)	0.60
	2	64 (26%)	0.18
	3	83 (34%)	0.15
	4	76 (31%)	0.14

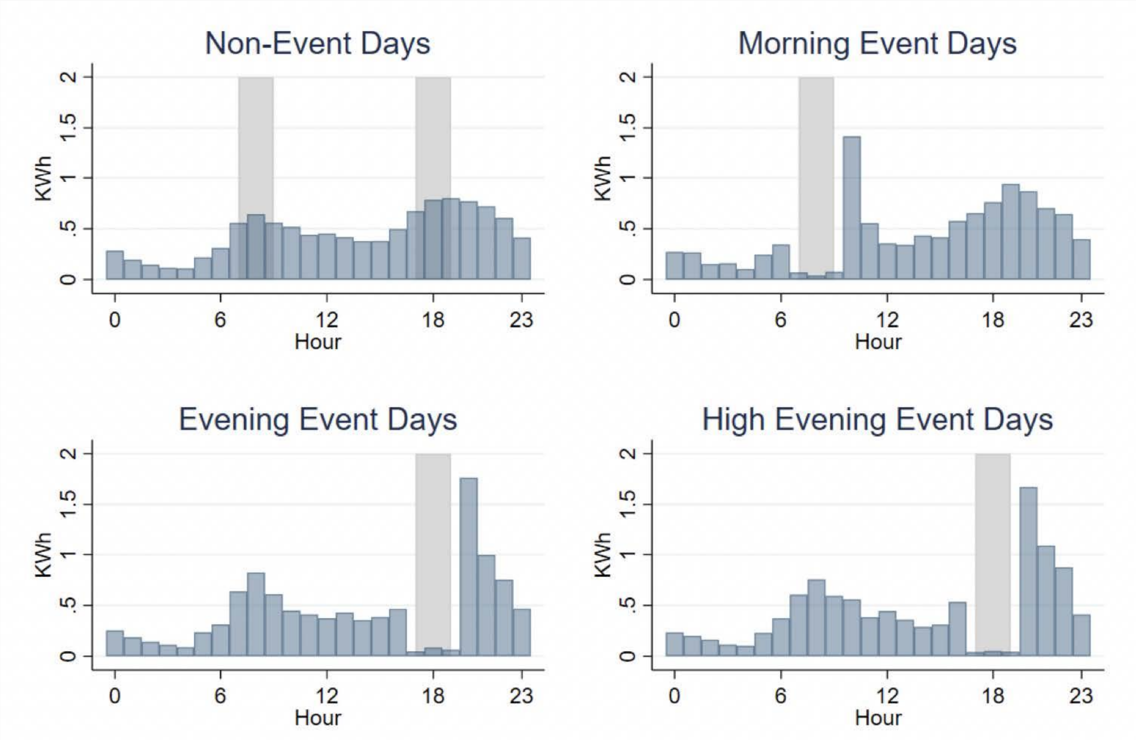
- Far more “top performers” in the Central group
- Top performers in Tech and Manual groups *need to interact far more*

Device data: Hot Water Heaters

Tech

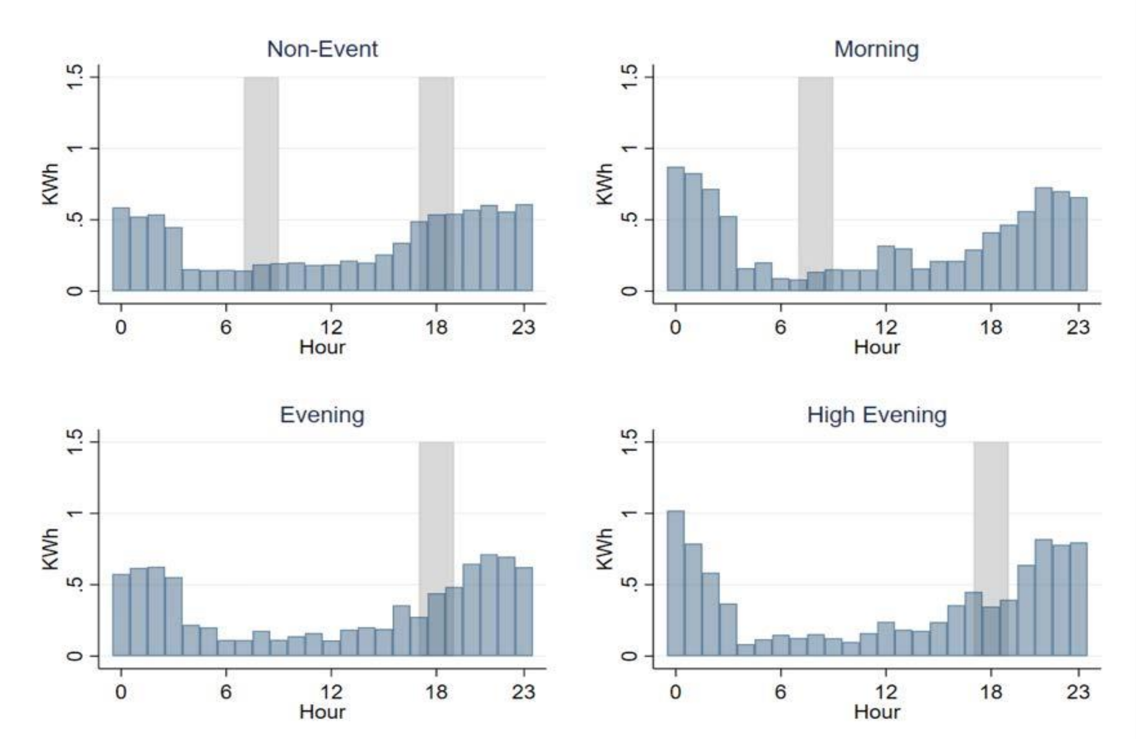


Central

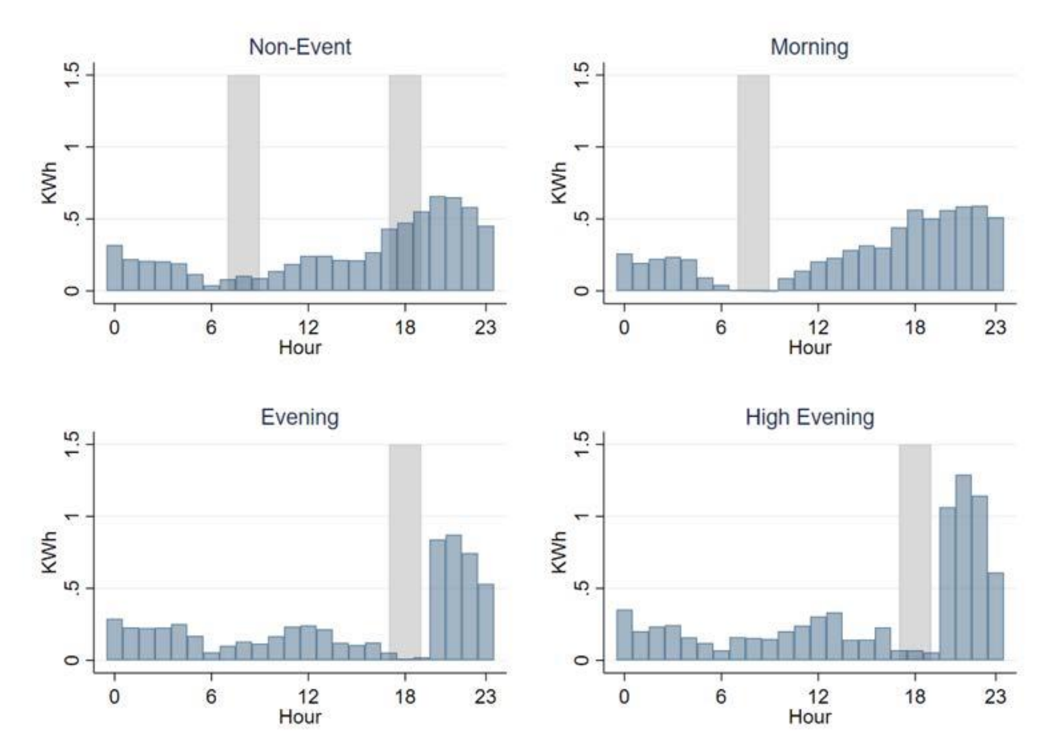


Device data: Level 2 EV Chargers

Tech

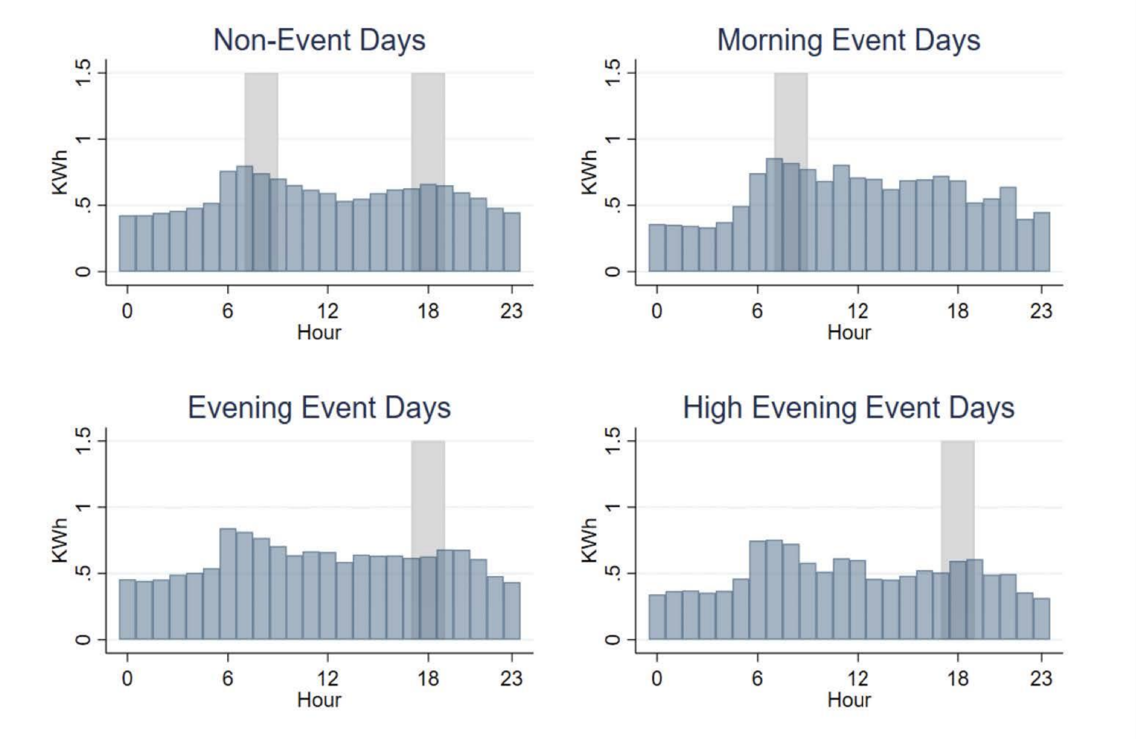


Central

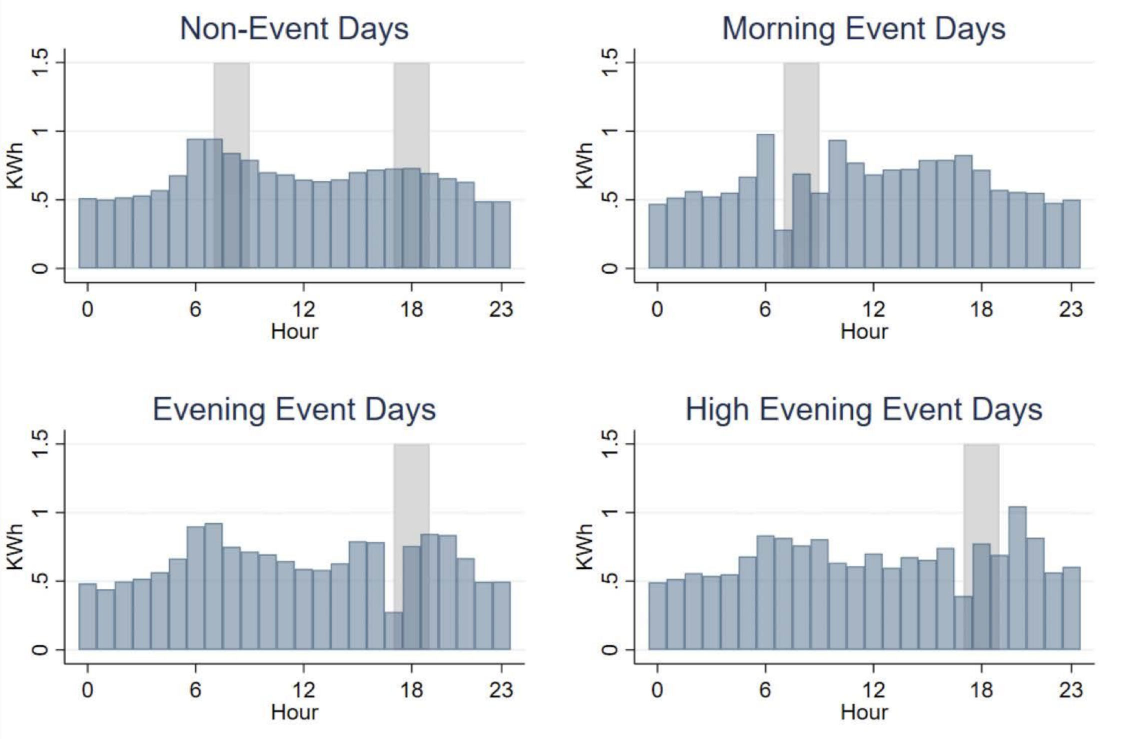


Device data: Baseboard Heaters

Tech



Central



Spillover estimates

Table C3. Treatment Effects of Participants by Program - Spillover Effects

Program	Pre-Event	Event	Post-Event
Central	0.0056	-0.3036***	0.2063***
	(0.0123)	(0.0204)	(0.0195)
Tech	0.0163	-0.0496***	0.0034
	(0.0118)	(0.0160)	(0.0121)
Manual	0.0030	-0.0542***	-0.0031
	(0.0102)	(0.0123)	(0.0101)

Notes. The reported results are program-specific treatment effect coefficients in (1), with the addition of demand response group-specific Pre-Event and Post-Event covariates. Pre-Event (Post-Event) is an indicator that equals 1 for two hours prior to (after) peak events, and zero otherwise. Standard errors are reported in the parentheses and clustered at the household level. All specifications include fixed effects at the household and hour-of-sample levels. Statistical Significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Intent-to-treat estimates

Table C7. ITT Estimates by Program

	Central	Tech	Manual
Assigned × Event	-0.1479***	-0.0377***	-0.0287***
	(0.0140)	(0.0102)	(0.0089)

Notes. The reported results are program-specific ITT coefficients from equation (5). For each demand response program, the sample includes households from their own treatment program and the never-treated groups. Standard errors are reported in the parentheses and clustered at the household level. All specifications include fixed effects at the household and hour-of-sample levels. Statistical Significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Robustness to control group

Table C1. Treatment Effects of Participants by Program

Program	(1)	(2)	(3)
Central	-0.3151***	-0.3007***	-0.3047***
	(0.0206)	(0.0204)	(0.0204)
Tech	-0.0661***	-0.0475***	-0.0495***
	(0.0132)	(0.0155)	(0.0159)
Manual	-0.0507***	-0.0459***	-0.0540***
	(0.0092)	(0.0117)	(0.0122)
Comparisons			
Own Program	Y	Y	Y
Other Treated			Y
Never Treated		Y	Y

No carry over

For each demand response program, we run separate regressions:

$$\ln(c_{it}) = \beta D_i EventWindow_{it} + \alpha_i + \tau_t + \delta X_{it} + \epsilon_{it}$$

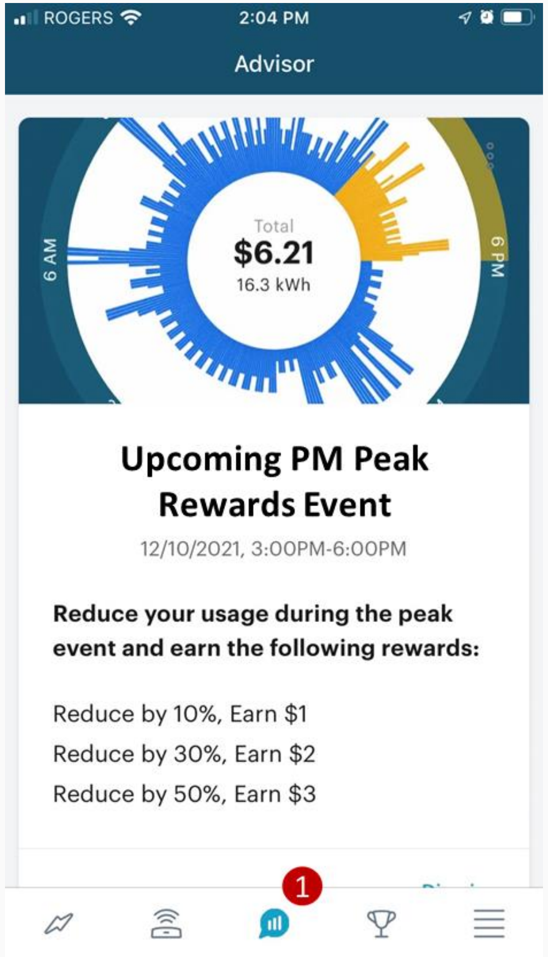
- Excludes event days
- Includes never-treated HHs
- $EventWindow_{it}$ equals 1 in the post-treatment period for hours where morning or evening events occur, 0 otherwise
- D_i equals 1 if the household is in a demand-response group (i.e., C, T, or M), 0 otherwise

Table C3. Carry Over DID Estimates by Program

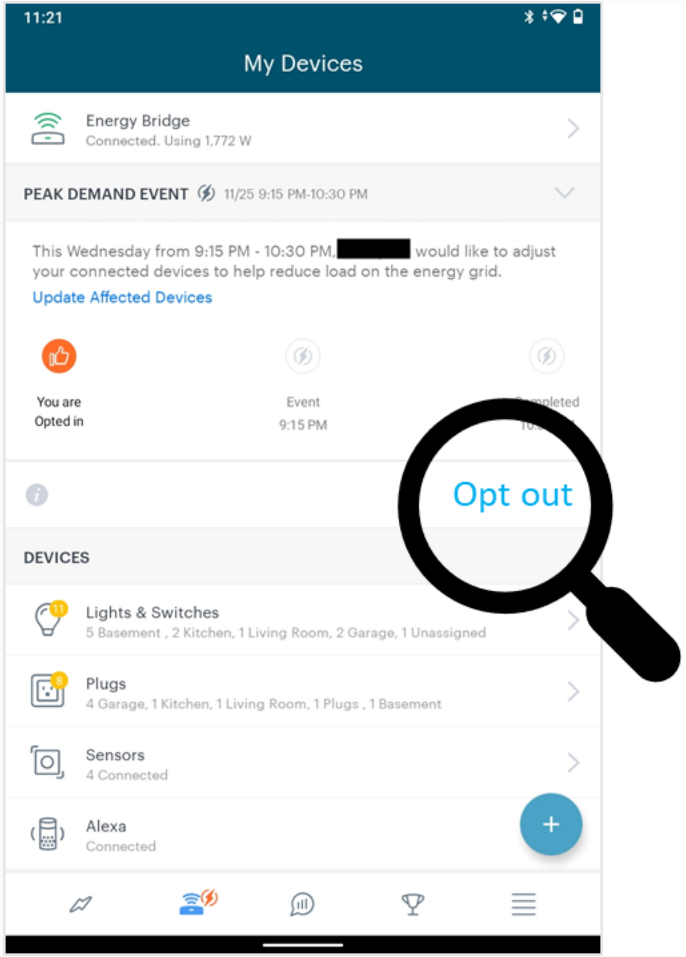
	Central	Tech	Manual
Event Window	0.0241	0.0302	-0.0014
	(0.0169)	(0.0181)	(0.0155)

User Experience: In-App Message & Central Opt-out

In-App notification



Central Opt-out



User Experience: Tech Opt-in

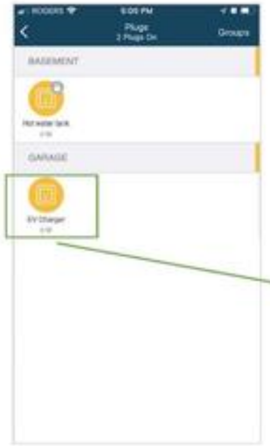
All smart devices, including smart plugs and load controllers can be set up and controlled through the “My Devices” page in the [redacted] app.

Select this icon at the bottom of your app to go to the “My Devices” page.

On the “My Devices” page your plugs should appear here.

Select “Plugs”

Turning Plugs On or Off



By pressing the icon you can turn a plug on or off.

Orange is on

White is off