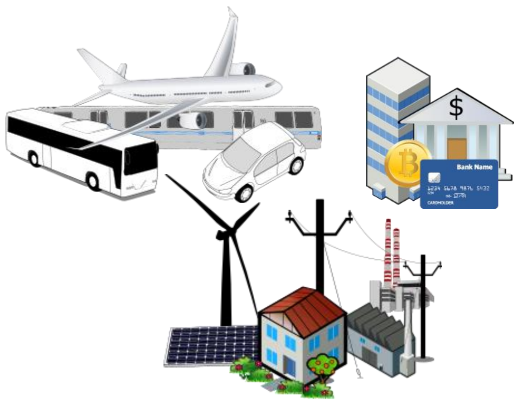
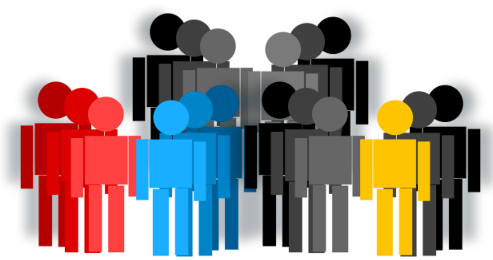
Quantifying User Engagement in Residential Demand Response Programs

Maximilian Balandat, Datong Zhou and Claire Tomlin

Analyzing Responses to Economic Incentives



- Effect of human behavior on engineered systems
  - use incentives to induce changes in behavior

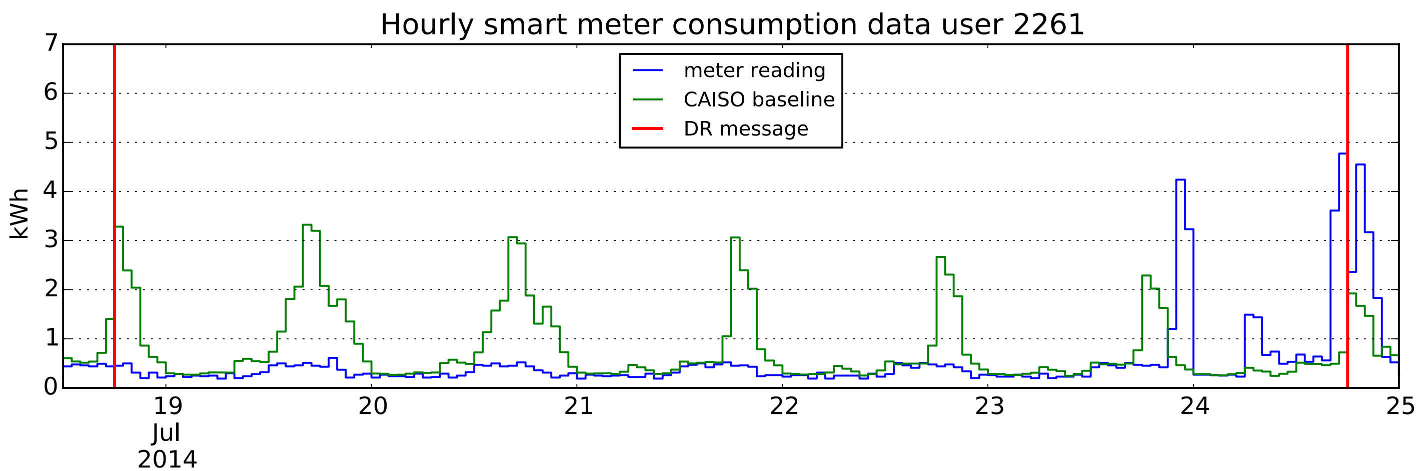


- Econometrics: Measure effect of “treatments” (policy changes, incentive schemes, etc.)
  - ideally: randomized controlled trial
  - between-group vs. within-subject experiment design

- Availability of massive amounts of data
  - detailed, high freqency, from various sources
  - typically not experimental, but observational
  - Goal: combine tools from Machine Learning and Econometrics to analyze behavior of individuals

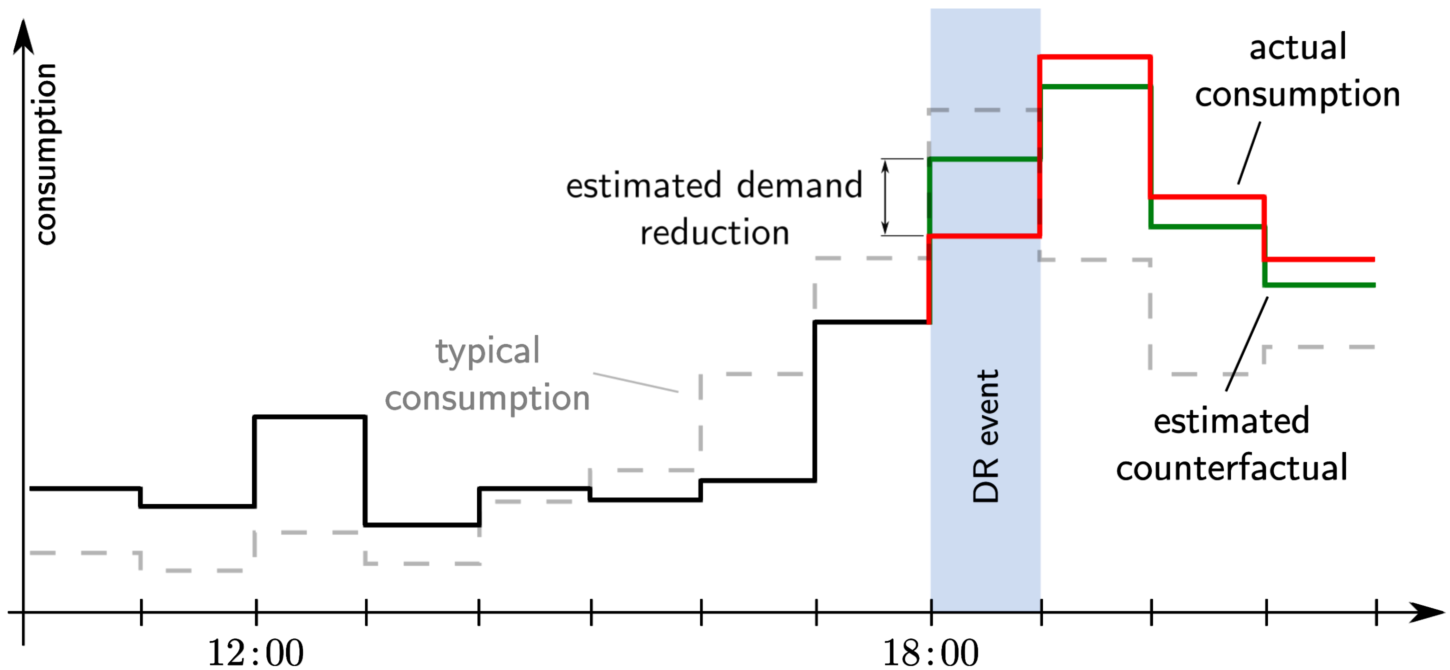
Residential Demand Response

* Customers receive incentives to reduce their power consumption during Demand Response (DR) events
  * How to quantify the demand reduction?
  * *Baselining* is typically used to estimate counterfactual consumption



  * meaningful for customers that exhibit *consistent* consumption
  * but: user consumption may still be *predictable*

Estimating Demand Reduction



- Estimated demand reductions are very noisy. How to make sense of them?

The ohmconnect data set

ohmconnect

* Founded August 2013, launched February 2014
* Only third-party residential Demand Response Provider in California

Data for each user in initial data set of 500 users:

* smart meter data (15min / hourly)
* "#OhmHour" DR messages (time stamped)
* approx. location (ZIP code)
* weather data (outside temperature, humidity, wind chill, etc.)
* web site page views and social media posts (daily aggregate)
* number of automated devices (e.g. EVs, smart thermostats)
* indoor temperature and temperature setpoint (for some users)
* electricity tariff

Methodology: Individual Treatment Effects

Clean data

- remove outliers
- normalize
- scrape weather data
- get weekends / holidays

Split data into non-DR and DR components

- Assumption: User reverts to regular behavior within 12h after a DR event

Train Machine Learning algorithms on non-DR data using cross-validation

- consumption (auto-regressive)
- temperature / humidity
- time of day, day of week

Predict counterfactual consumption during DR events

- Normalized by users's mean consumption
- determine prediction errors

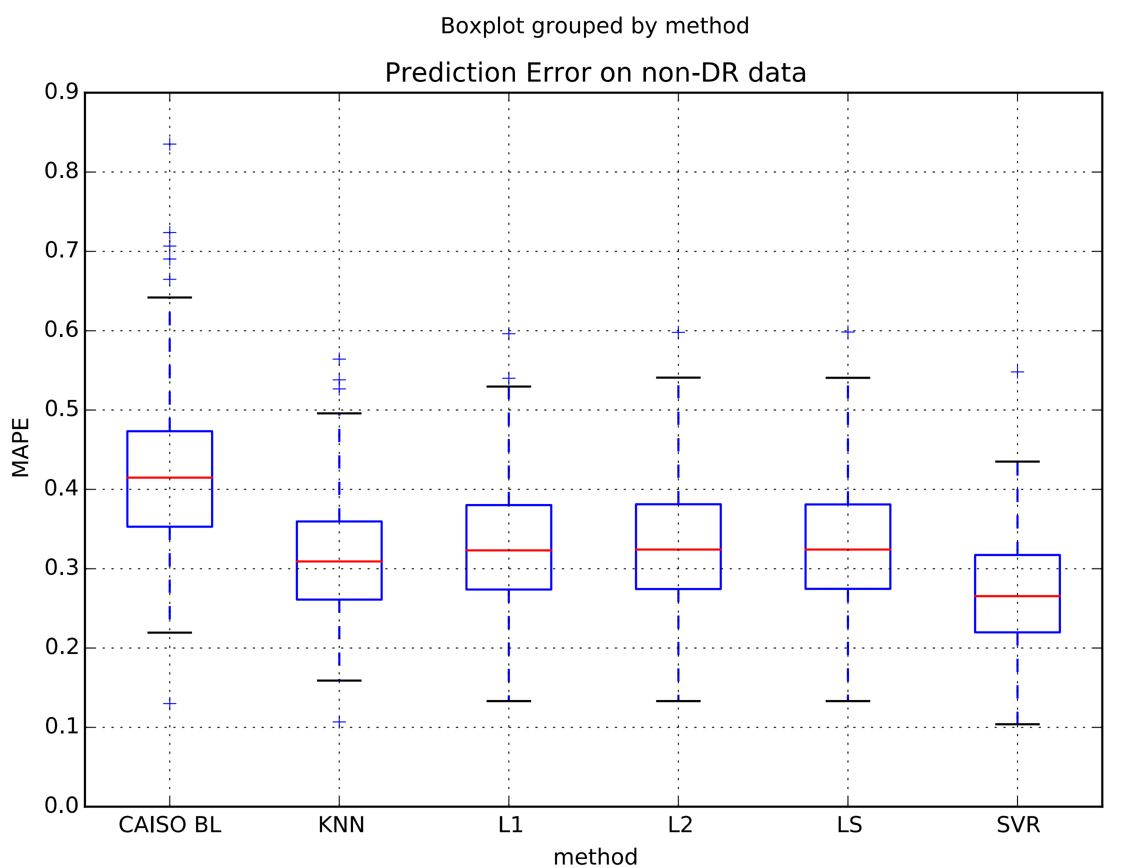
Perform non-parametric hypothesis test

- Null Hypothesis: Samples come from same distribution
- Alternative Hypothesis: location parameter shifted by Δ during DR events

Obtain estimate $\hat{\Delta}$ and non-parametric confidence interval for Δ

Prediction Quality

* Results based on 321 users for which sufficient trainig data is available

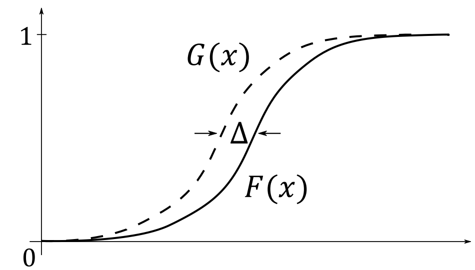


Non-Parametric Estimates of Demand Reduction

y / y_{DR} : vectors of realized consumption readings during non-DR / DR events

$\hat{y}$ / $\hat{y}_{DR}$: vectors of predicted counterfactuals during non-DR / DR events

* *Assumption:* The location parameters of distributions of $\hat{y}_{DR} \sim F$ and $y_{DR} \sim G$ differ by Δ, i.e. $G(x) = F(x + \Delta)$
* *Null Hypothesis:* $\Delta = 0$



First-differences (FD) specification:

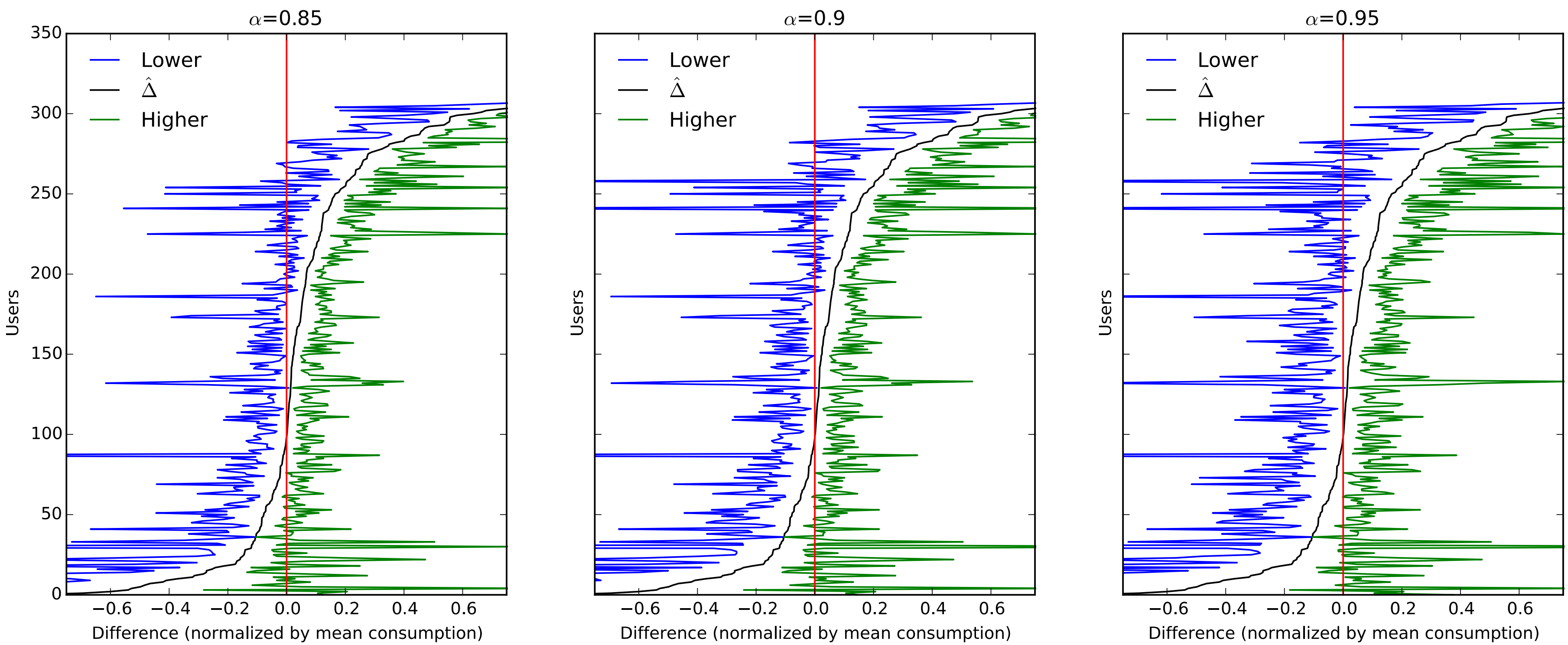
* compare paired samples $(\hat{y}_{DR}, y_{DR})$ using a singed rank test
* estimate median of the difference between a sample from $\hat{y}_{DR}$ and a sample from y_{DR}

Difference-in-differences (DID) specification:

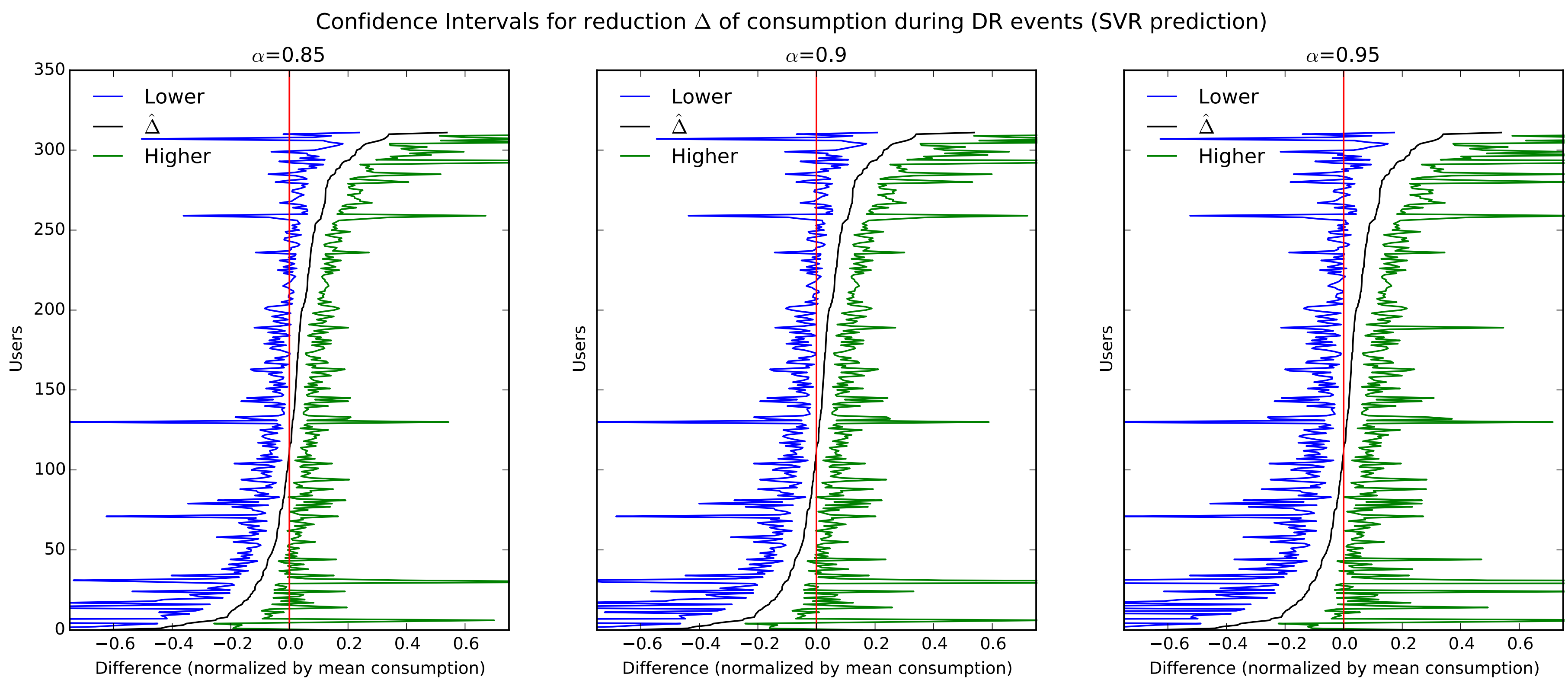
* compare prediction errors $e = \hat{y} - y$ and $e_{DR} = \hat{y}_{DR} - y_{DR}$ using a rank sum test
* estimate median of the difference between a sample from e and a sample from e_{DR}

Results: FD Specification

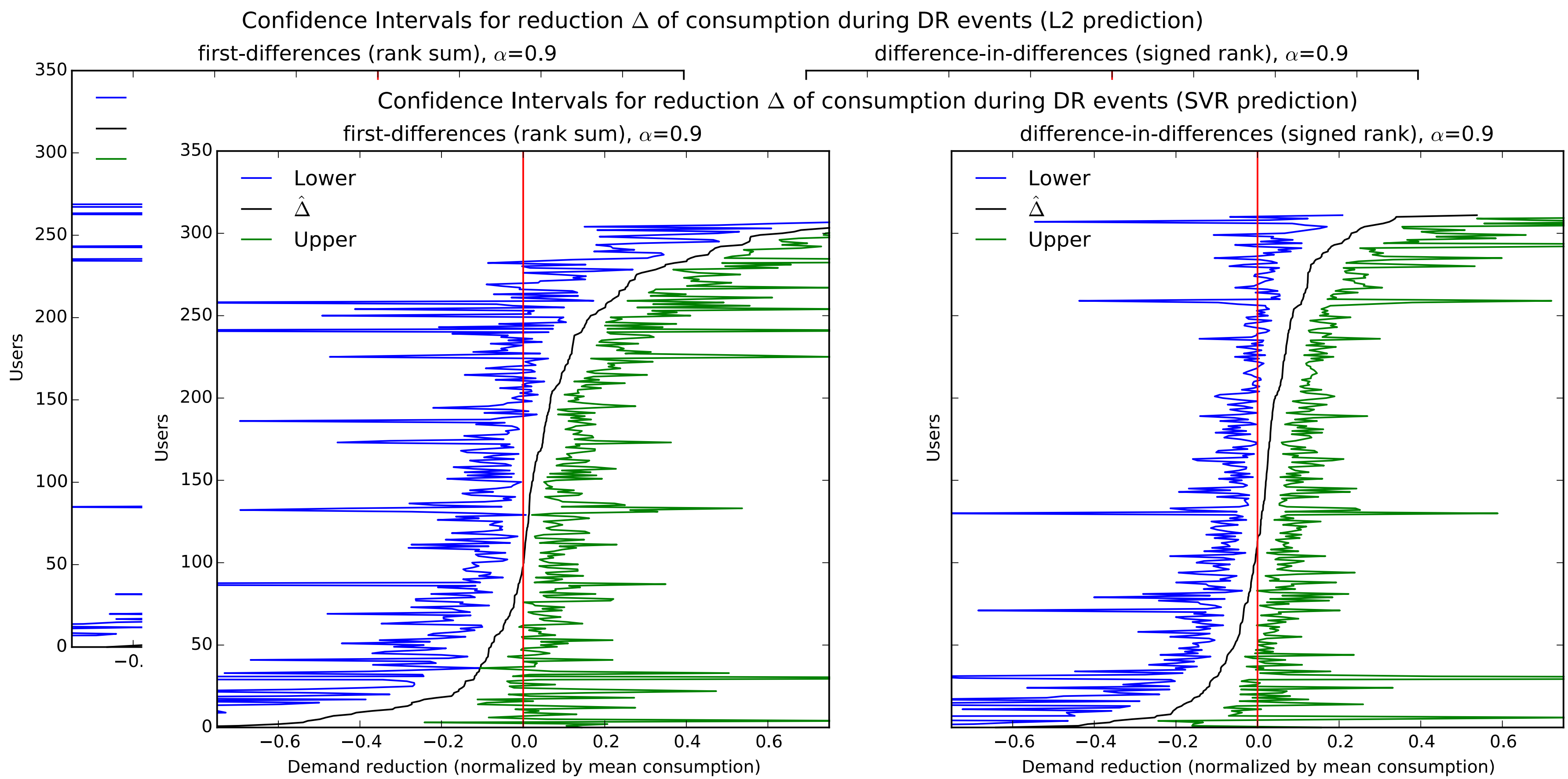
Confidence Intervals for reduction Δ of consumption during DR events (SVR prediction)



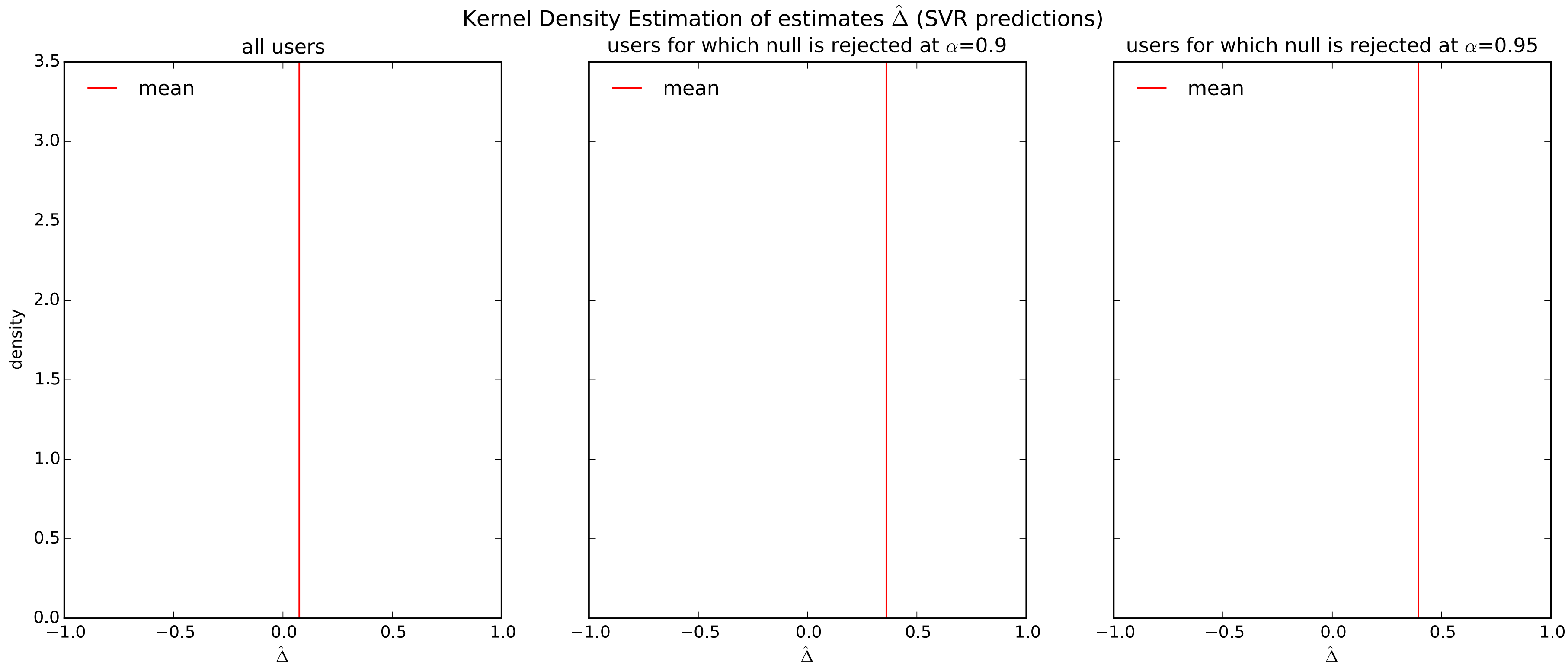
Results: DID Specification



Effects of Bias in the Prediction Model



Distribution of Reductions during DR events



Additional Challenges:

- Effect of prediction method on the test's power:

α	# users null rejected: DiD, SVR	# users null rejected: DiD, baseline	# users in common
85%	81	94	39
90%	63	70	27
95%	42	47	14

  - *Potential issue:* Endogeneity in choosing the DR events

- Where are the automated users?

α	% automated (DID, SVR)	% automated (DID, baseline)
85%	16.0	16.0
90%	11.1	14.3
95%	9.5	14.9

Conclusion and Future Work

Conclusion

- Combining Machine Learning and non-parametric statistics provides powerful econometric tools for user-level analysis
  - allows to estimate user engagement based on observational data only
- Additional care must be taken in handling potential pitfalls arising from generic Machine Learning algorithms (biased estimators, overfitting, etc.)

Future Work

- Analyze and correct for various potential biases in the methodology:
  - self-selection bias, endogeneity of DR events, omitted variable bias, etc.
  - validate results against randomized field experiment
- Perform analysis on larger data set with more DR events per user
- Predict users engagement based on additional high-level data on users