



# **SICSS-Oxford: Advances in Computational Experimental Social Sciences**

Nuffield College Oxford

---

Ray Duch

June 22, 2019

Nuffield College Oxford Centre for Experimental Social Science - CESS

- Part I: Brief introduction to CESS and Talk
- Part II: Micro-replications and experimental measurement error
- Part III: Designing virtual experiments with post-stratified Average Treatment Effects

## **Part I: Introduction**

---

# Nuffield Centre for Experimental Social Sciences (CESS)

- CESS Centres and Labs
  - Nuffield College Oxford
  - CESS Santiago
  - Nuffield CESS FLAME India
  - CESS China Nankai University
- CESS Online
  - Nuffield CESS Online facilities
  - Virtual non-deception CESS subject pool
  - UK, Ireland, India, China, U.S., and Chile
- CESS in the field
- CESS Schools, Workshops, Visiting Post-/Pre-doc positions...

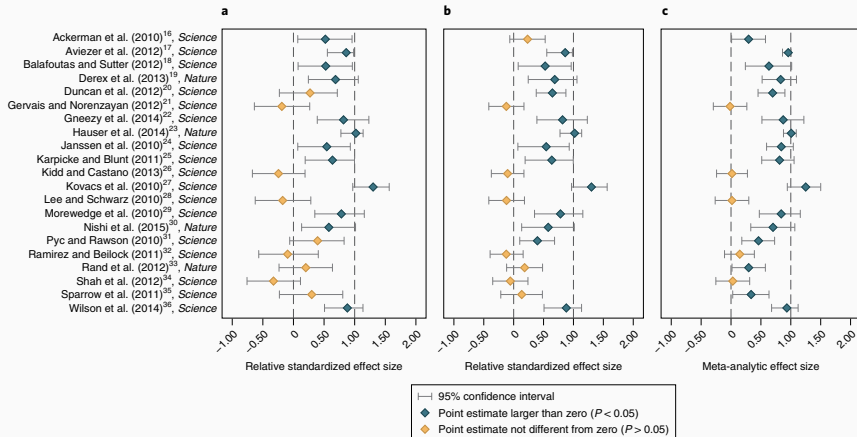
- Experimental perspective on themes you covered this week
  - Computational methods
  - Large data
  - Social media
  - Robustness/replication
- Part II
  - Detecting experimental measurement error
  - Vary the experimental context or mode
  - Machine learning for detection of heterogeneity
- Part III
  - Large-scale experimental interventions
  - Digital trace outcomes
  - Post-stratification with probability machines

## **Part II: Micro-replications**

---

- Costs declining significantly
- Convenience samples are the norm
- Proliferation of data generation modes
- Democratic

# There are Costs: Camerer et al 2018 Nature





## Some Observations

- How do you know you have this experimental measurement error?
- You typically have no clue as to whether its an issue
- Note: this has nothing to do with external validity/representative sample/etc.

# Micro-replications can help

- Maybe....
- But what micro-replication?
- In which micro-replication should you invest your research dollars?
- Multi- rather than single-mode replications are more informative of experimental measurement error

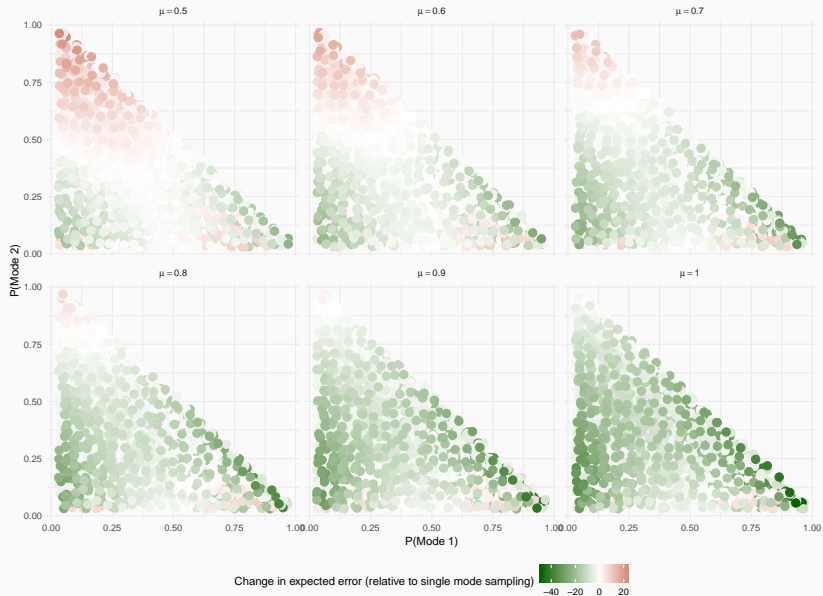
# The Experimental Mode or Context



# Modes and Experimental Measurement Error

- do modes exaggerate measurement error, i.e.,  $ME_k > 0$
- resulting in  $ATE_k^* = (ATE_T + ME_k)$
- multi-mode replication design may be informative when:
  - $ME_k \neq ME_{k'}$  and
  - there is a reasonably high probability the researcher can distinguish low from high error modes

# Multiple-mode Replication Simulation



## Illustrate: Lying Experiment (Duch Laroze Zakharov 2018)

- Outcome of interest: Lying about income from RET
- Treatment: Deduction rate that make it more expensive to lie
- Expectation: Lying declines if deduction rates rise

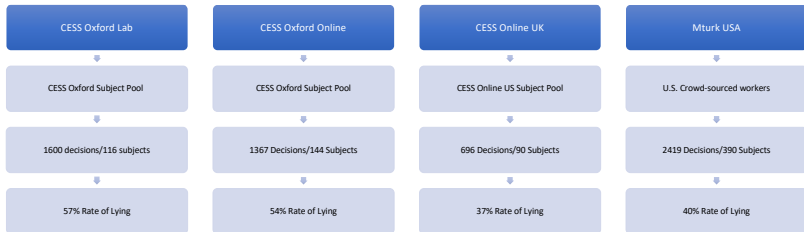
- 3 different tax rates (10%, 20% and 30%)
- Fixed at the group level
- Taxes are redistributed equally among group members
- Public good
- No excludability
- No social gains/losses
- No audits or fines
- 10 rounds
- Paid for one of them at random
- Fixed groups of 4 participants
- Random matching at the beginning

## Design: each round

- RET: solve as many additions as possible in 60 sec
- two random two-digit numbers
- Information individual gross profit (before tax)
- Declare their income (to be taxed)
- Information individual net profit (after tax and redistribution)
- Differentiated by profit, tax and redistribution



# Lying Experiments



# Conventional GLM Estimation

	Mode			
	Lab	Online Lab	Online UK	Mturk
Ability Rank	−0.500*** (0.036)	−0.163*** (0.045)	−0.163** (0.071)	−0.120*** (0.037)
20% Deduction	−0.123*** (0.024)			
30% Deduction	−0.128*** (0.025)	−0.184*** (0.025)	0.042 (0.038)	0.018 (0.021)
No Audit	−0.334*** (0.023)	−0.127*** (0.026)	−0.155*** (0.036)	0.011 (0.024)
Age	0.012*** (0.002)	0.007** (0.003)	−0.0002 (0.001)	0.002** (0.001)
Gender	0.002 (0.022)	0.100*** (0.025)	−0.022 (0.035)	−0.004 (0.020)
Constant	0.715*** (0.066)	0.476*** (0.089)	0.880*** (0.070)	0.576*** (0.043)

# BART Estimation

- Bayesian estimation strategy using tree-logic
- Highly flexible estimation strategy

To recover individual estimates of treatment effect:

- Assume binary treatment
- Run BART on experimental data (the training set) to generate both model and predicted outcomes for observed data
- Invert treatment assignment of all observations, and pass through model (test set) to generate set of counterfactual predictions
- For each individual,  $i$ ,  $CATE = Y_{i,D=1} - Y_{i,D=0}$

# BART: R Code

```
# Separate outcome and training data
y <- df$report.rate
train <- df[,-1]

# Gen. test data where those treated become untreated, for use in calculating ITT
test <- train
test$treat.het <- ifelse(test$treat.het == 1,0,ifelse(test$treat.het == 0,1,NA))

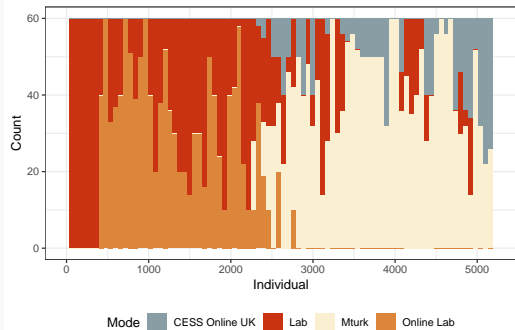
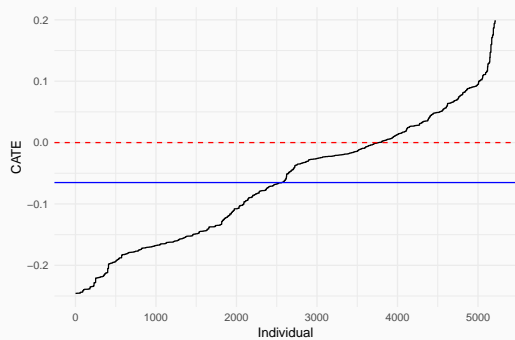
# Run BART for predicted values of observed and synthetic observations
bart.out <- bart(x.train = train, y.train = y, x.test = test)

# Recover CATE estimates and format into dataframe
CATE <- c(bart.out$yhat.train.mean[train$treat.het == 1] - bart.out$yhat.test.mean[test$treat.het == 0],
        bart.out$yhat.test.mean[test$treat.het == 1] - bart.out$yhat.train.mean[train$treat.het == 0])

CATE_df <- data.frame(CATE = CATE)
covars <- rbind(train[train$treat.het == 1,c(2:5)], test[test$treat.het==1,c(2:5)])

CATE_df <- cbind(CATE_df,covars)
CATE_df <- CATE_df[order(CATE_df$CATE),]
CATE_df$id <- c(1:length(CATE))
```

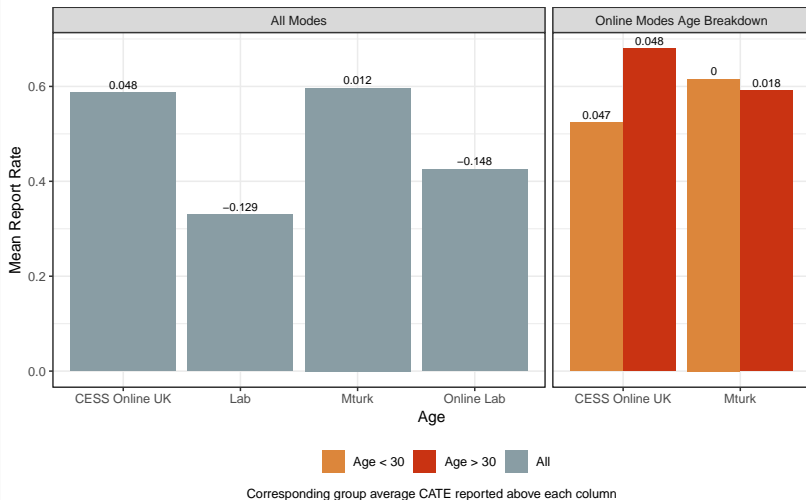
All replication code available at <https://github.com/rayduch/Experimental-Modes-and-Heterogeneity>



## Real Effort Tasks Inter-Class Correlations Across Modes

Mode	(1)	(2)	(3)	(4)
Lab	0.768 (0.018)	0.768 (0.018)	0.636 (0.039)	0.85 (0.047)
Lab Online	0.807 (0.018)	0.76 (0.017)	0.762 (0.021)	0.767 (0.047)
Online UK	0.88 (0.011)	0.827 (0.018)	0.827 (0.026)	0.752 (0.029)
MTurk	0.758 (0.015)	0.758 (0.012)	0.782 (0.024)	0.828 (0.026)
Deduction Rate	10%	30%	10%	30%
Audited?	No	No	Yes	Yes

# Comparing Percentages of Actual Earnings Reported



# India Measurement Error Experiments

Coeff	S.E.	t-statistic	p	Mode	Error	Incentivised?
-0.74	0.47	-1.57	0.12	MTurk	Control	No
-0.83	0.47	-1.76	0.08	MTurk	High	No
-3.85	0.51	-7.52	0.00	CESS Online	Control	No
-3.23	0.49	-6.64	0.00	CESS Online	High	No
-1.16	0.49	-2.35	0.02	MTurk	Control	Yes
-1.00	0.33	-3.01	0.00	MTurk	High	Yes

**Table 2:** Induced measurement error model results



- micro-replication with multiple-modes
- machine learning strategies for estimating CATE that might be mode-related
- incorporate measurement strategies in experimental design that might explain mode-related experimental measurement error

## **Part III: Designing virtual experiments with post-stratified Average Treatment Effects**

---

# The CESS Chile Audit Experiment Design

- Random Assignment of Audit “Shocks” to Municipal Voters
- Random Assignment of Information Treatments to targeted zip codes, aka Villas
- Efficient pre-and post-treatment measurement
- Post-stratification of ATE with random forest estimation
- Inspired by the Digital Trace and Vote India project that have focused on simple election outcomes rather than causal effects

# Chile 345 Municipalities



# CESS Santiago Chile Audit Experiment Design



# Facebook Sampling

---

# Facebook Ad Manager Dashboard

Felipe Torres Raposo (13... ▾)

Campaign

Ad account

Ad set

Page

Variable

Audience

Placements

Delivery optimisation

Budget & schedule

Ad A

Ad B

Test Another Ad

Close

Ad set: Create split test ad sets.

Custom Audiences ⓘ

Add a previously created Custom or Lookalike Audience

Exclude | Create new ▾

Locations ⓘ

People who live in this location ▾

Chile

📍 Calle Quebrada La Chimba, Antofagasta, Antofagasta Region,

Chile

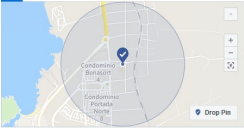
1 km ▾

📍 Include ▾ | Type to add more locations

Browse

1

Your audience location has been changed from Chile to Calle Quebrada La Chimba, Antofagasta, Antofagasta Region, Chile. Undo change



Add locations in bulk

Age ⓘ 18 ▾ - 65+ ▾

Gender ⓘ All Men Women

Languages ⓘ Enter a language...

Detailed targeting ⓘ Include people who match ⓘ

Interests > Additional interests

26

- Facebook ad manager allows us to target subjects in specific Villas
- Banner recruitment ads sent to FB subscribers in specific Villas
- 6,500 pre-treatment online qualtrics interviews completed



# Facebook Ad Manager Dashboard

Campaign

Ad account

Ad set

- Page
- Variable
- Audience
- Placements
- Delivery optimisation
- Budget & schedule**

Ad A

Ad B

Test Another Ad

Ad set: Create split test ad sets.

Use control (optional)

\$200,000

maximum bid in each auction

Facebook will aim to get the most 1,000 impressions without bidding more than \$200,000 in any auction using the bid-cap bid strategy.

Frequency cap

1

Impression every

7

days

When you are charged

Impression

View and edit

Ad scheduling

☒ Run ads all the time

☐ Run ads on a schedule

Delivery type

☒ Standard (recommended)

Get results throughout your selected schedule

☐ Accelerated

Spend your budget and get results as quickly as possible

View and edit

**Split test budget & schedule**

Define how much you'd like to spend and when you'd like your ads to appear. [Learn more](#)

Budget

Daily budget

\$2000

Estimated test power: **> 95%**

Actual amount spent per day may vary.

Your total budget will be split evenly among the ads that you're testing.

Schedule

☒ Run split test starting today

☐ Set a start and end date

Duration

3 days

☐ End the test early if a winning ad set is found

This test will run for 3 days and spend no more than \$6,000.

[Hide advanced settings](#)

Close

Back

Continue

# Facebook Ad Manager Dashboard

Campaigns

Ad account

Ad set

Ad A

- Identity
- Format
- Media
- Text

Ad B

Test Another Ad

Close

Ad name ⓘ Ad A ⚙

⚠ You've selected ads with multiple Pages. You can't edit these ads together. Deselect the ads and edit them one at a time.

Create your own Instant Experience with a custom layout and more design and interaction options.

Media

Choose one or more images, a video or a slideshow for your ad.

☒ Image

☐ Video/slideshow

💡 Boost performance with video templates

Use the Video Creation Kit to create video ads from your images, which may perform better than single image ads. You can always find templates for these ads if you select "Videoslideshow".

View case study

Use Templates

📄 Upload or select an image for your ad.

+

Upload images

Browse Library

Free Stock Images

Image recommendations

- Size: 1,080 x 1,080 pixels
- Crop: 1:1
- File type: .jpg or .png

To maximise ad delivery, use an image that contains **little or no overlaid text**. [Learn more](#)

Text

Enter the text for your ad. [Learn more](#)

Primary text ⓘ

Enter text that clearly tells people about what you're promoting

Ad preview 📄

Mobile News Feed ▾

1 of 1 ad < >

1 of 18 < >

29

## **Treatment Assignment and Sample**

---

# Treatments

- Treatment assignment and pre-treatment sample sizes are in the following Figure
- Contraloria will cooperate in randomly assigning audits to municipalities
- two X two factorial design will allow us to assess two different information treatments:
  - audit results
  - performance report
- control conditions
  - no audits
  - no information

## Treatment Summary

---

# Treatments

Block1: Ranking Level	High						Medium						Low					
Block 2: Municipal Level	Audit (15)				No Audit (15)		Audit (10)				No Audit (5)		Audit (5)				No Audit (5)	
Information treatments	No Info (40)	Audit + Perf. (40)	Audit (40)	Perf. (40)	No Info (10)	Perf. (20)	No Info (4)	Audit + Perf. (6)	Audit (6)	Perf. (4)	No Info (6)	Perf. (14)	No Info (2)	Audit + Perf. (4)	Perf. (2)	No Info (4)	No Info (4)	Perf. (6)
Pre-treatment Sample	1000	1000	1000	1000	250	500	100	150	150	100	150	350	50	100	50	100	100	150

**Total sample: 6,500 individuals**

# Information Treatments

---

# Malfeasance Information Treatment



**Chile Transparente** ✓  
@ChTransparente

- Home
- Posts
- Twitter
- Videos
- About
- Community
- Reviews
- Photos

Create a Page

Like · Reply · 1d

View 1 more comment

**Chile Transparente**  
18 June at 17:30 · 🌐

"Alcalde de Río Negro arriesga 6 años de cárcel por malversación de caudales públicos" 🗣️ "Contraloría acreditó el pago irregular de \$6 millones desde el Municipio de Río Negro a Camilo Miranda, ex jefe de gabinete del alcalde Schwalm en 2012" <https://www.biobiochile.cl/.../alcalde-de-rio-negro-arriesga-...>

 **biobiochile.cl**

Domingo 16 junio de 2019 | Publicado a las 19:49

## Alcalde de Río Negro arriesga 6 años de cárcel por malversación de caudales públicos

Por Felipe Díaz Montoya  
La información es de Soledad Fuentes



Facebook

🚩 ¿Encontraste algún error? Avisanos

👁️ 21.403 visitas

👍 16

1 comment 10 shares

Contact Us · Send Message

Open now

Suggest Edits

**Page Transparency** See More

Facebook is showing information to help you better understand the purpose of a Page. See actions taken by the people who manage and post content.

📅 Page created – 8 November 2011

**Team Members**

 Alberto Precht Morris

**Related Pages**

 **No a Piñera 2018**  
Community organisation · Like

 **Gamba.cl**  
Media/news company · Like

 **CNN Chile** ✓  
Constanza Harismendy it...  
News and media website · Like

**Pages liked by this Page** >

 **Observatorio del Gast...** · Like

 **Embassy of Canad...** · Like



- Treated Villas receive saturation information treatments
  - FB ads
  - Twitter ads
  - Instagram ads
  - Note: these are targeted at all social media in the treated Villas
- Post-Treatment Survey
  - Re-contact 6,500 initially surveyed subjects
  - Evaluations of municipal government and voting intentions

# Average Treatment Effect Estimation

---

# Average Treatment Effect Estimation

Three steps to obtaining area-estimates of ATE:

- identification of individual characteristics to use in ATE prediction model (socio-economic categories);
- estimation, at the individual-category level, of ATE;
- weighting predicted ATEs by cell counts and summing over the area of interest to recover estimates of ATE.
- Outcome of interest is change in support for municipal government incumbent as measured by the pre- and post-surveys.

## Average Treatment Effect Estimation

- Individual-level characteristics  $X$  from pre-treatment survey;
- Missing values in  $X$  are imputed with a random-forest multiple-imputation strategy implemented via the packages `ranger` and `missForest`;
- define the individual-level category  $C_g$ , for categories  $g = 1, \dots, G$  as a unique realization of the set of variables which compose  $X$ , i.e.  $C_g = \{X_1 = x_1, \dots, X_m = x_m\}$ .

## Estimation Categories for Observed Data

	Categories (C)	Audit/Info	Audit/NoInfo
Gender	Two	600	575
Education	Two	600	575
Income	Three	400	382
Age	Three	400	383
Total	36	33	32

**Table 3:** Number of cell categories and number of subjects within cells for pre-treatment survey

# Estimation of ATE: Change in Vote for Municipal Incumbent

- Pre-treatment Estimate
  - Estimate the probability of an individual-category voting for a municipal mayor prior to treatment;
  - quantity of interest is per group and is  $P_{0g}(+|C)$
- Post-treatment Estimate
  - Estimate the probability of an individual-category voting for a municipal mayor prior to treatment;
  - quantity of interest is per group and is  $P_{1g}(+|C)$

# Average Treatment Effects: High Risk 30 Municipalities

Audit Treatment Effect							
	Pre-Treatment			Post-Treatment			ATE
	Audit/No Info	No Audit/No Info	Diff/No Info	Audit/No Info	No Audit/No Info	Diff/No Info	
Male	0.6	0.5	-0.1	0.7	0.5	-0.2	-0.1
Female	0.6	0.5	-0.1	0.8	0.5	-0.3	-0.2
Audit Information Treatment Effect							
	Pre-Treatment			Post-Treatment			ATE
	No Info	Audit Info	Diff	No Info	Audit Info	Diff	
Male	0.8	0.7	-0.1	0.8	0.5	-0.3	-0.2
Female	0.6	0.5	-0.1	0.6	0.3	-0.3	-0.2

# Likelihood Estimation

---



- Train a probability machine to estimate proportion of individuals who express a positive opinion of their incumbent municipal politician before treatment ( $P_0$ ) and following the treatment ( $P_1$ );
- Make a prediction for each group before and after treatment;

$$\hat{P}_\tau(+|C) = \varphi(C); \quad \forall \tau = \{0, 1\} \quad (1)$$

- estimate the global error via MSPE1 and approximate distribution with Normal density:

$$P_g(+|C) \sim N(\varphi(C_g), (\hat{\sigma}_{\text{RMSE1}})^2); \quad (2)$$

## Area Estimation

---

## Area Estimation: Overall National ATE

- Cell counts:  $Q_g = \sum_h 1(x_h = C_g)$ ;
- Estimation of National treatment effect:

$$ATE = \frac{\sum_g (\hat{P}_{1g}(+|C) - \hat{P}_{0g}(+|C)) \times Q_g}{\sum_g Q_g}; \quad (3)$$

## Area Estimation: Specific Municipal ATE

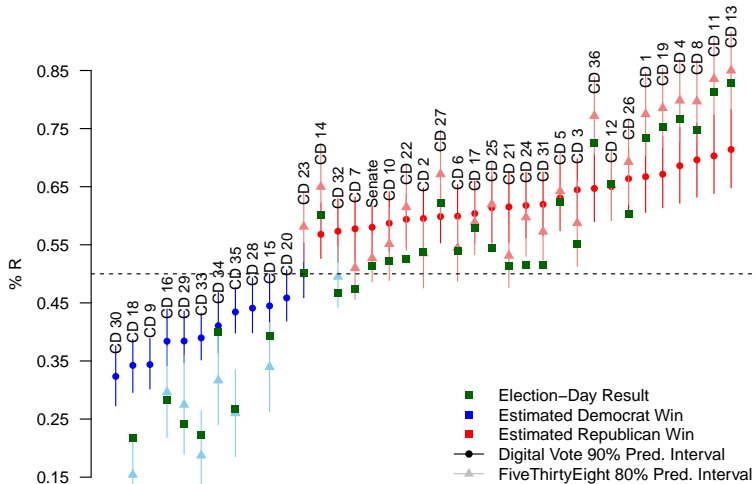
- Assume there are  $s = 1....S$  municipal units;
- Municipal cell counts:  $Q_{gs} = \sum_h 1(x_h^{muni} = C_{gs}, x^{muni} = s)$ ;
- Estimation of an  $s$  municipal treatment effect:

$$ATE_s = \frac{\sum_g (\hat{P}_{1gs} (+|C) - \hat{P}_{0gs} (+|C)) \times Q_{gs}}{\sum_{gs} Q_{gs}}; \quad (4)$$

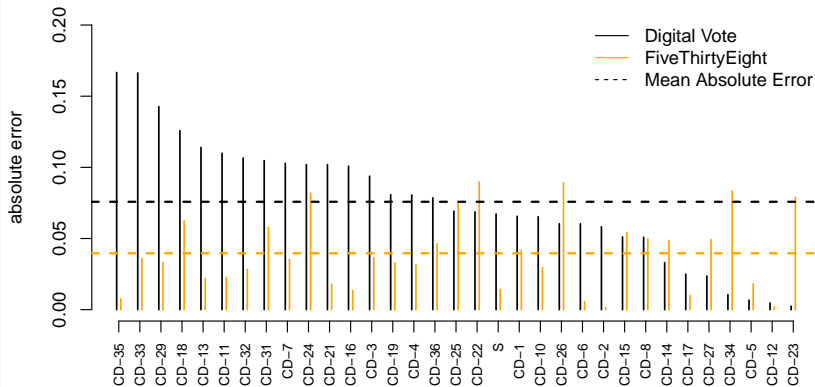
## **Illustrated Results Cerina and Duch**

---

# Results

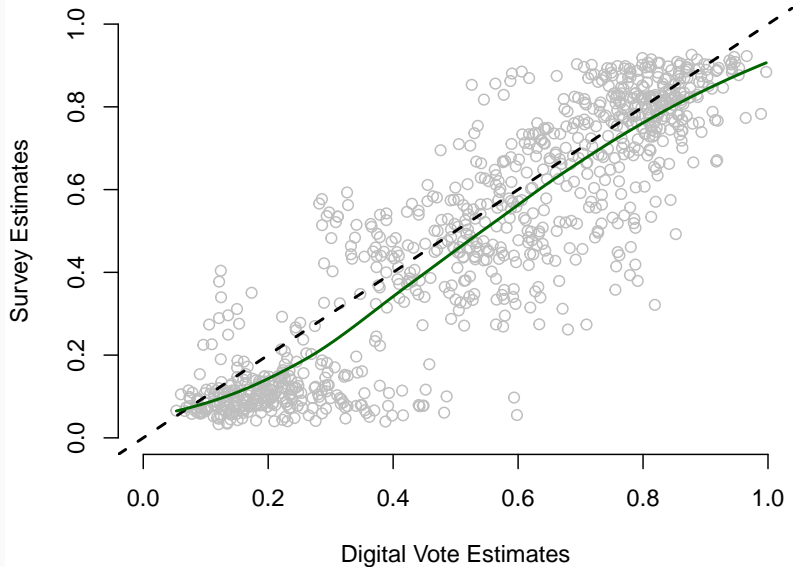


# Results



**Figure 1:** Mean Absolute Error: FiveThirtyEight and Digital Vote

# Results





## **Part III: Recommendations**

---

## Part III: Recommendations

- Virtual sampling advantage: incorporates digital trace intelligence including geo-location
- FB Ad Manager: facilitates information treatments administered to very large samples of the population
- Area Post-stratification with machine learning: convenience sample + detailed census and survey data + random forest predictions = ATE for nation and sub-national units