Experimental Methods: Lecture 5

Machine Learning

Raymond Duch

May 26, 2021

University of Oxford

Road Map

- Machine Learning
- Machine Learning: BART
- BART Heterogeneous Treatment Effects
- Treatment adaptive design
- Post-stratified Treatment Effects

Machine learning

How Machine Learning Works

- Prediction: produce predictions of y from x
  - Supervised – use subset of "known" y values to train model e.g. LASSO, random forest, BART etc.
  - Unsupervised – self-organised learning e.g. k-means clustering
- Discovers complex structure not specified in advance
- Fits complex and very flexible functional forms to the data
  - without simply overfitting; and
  - functions that work well out-of-sample

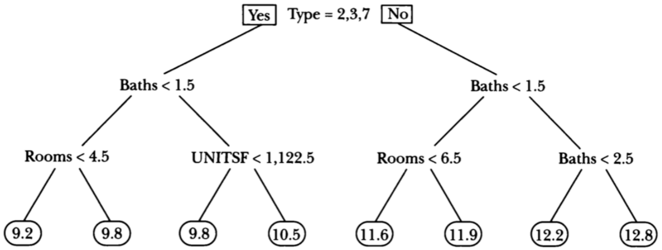
Performance of different machines

Table 1
Performance of Different Algorithms in Predicting House Values

	Prediction performance (R^2)		Relative improvement over ordinary least squares by quintile of house value				
Method	*Training sample*	*Hold-out sample*	1st	2nd	3rd	4th	5th
Ordinary least squares	47.3%	41.7% [39.7%, 43.7%]	–	–	–	–	–
Regression tree tuned by depth	39.6%	34.5% [32.6%, 36.5%]	–11.5%	10.8%	6.4%	–14.6%	–31.8%
LASSO	46.0%	43.3% [41.5%, 45.2%]	1.3%	11.9%	13.1%	10.1%	–1.9%
Random forest	85.1%	45.5% [43.6%, 47.5%]	3.5%	23.6%	27.0%	17.8%	–0.5%
Ensemble	80.4%	45.9% [44.0%, 47.9%]	4.5%	16.0%	17.9%	14.2%	7.6%

Regression tree

A Shallow Regression Tree Predicting House Values



Note: Based on a sample from the 2011 American Housing Survey metropolitan survey. House-value predictions are in log dollars.

Improving ML-estimation: ensemble methods

Bagging:

- "Bootstrap aggregation", random samples with replacement as training data
- Reduces variance, as used in *random forest* methods

Boosting:

- Sequential weak-learner models, data weighted by misclassification
- Reduces bias, as used in *gradient tree* methods

Stacking:

- Run different estimation strategies and weight super-learner by each model's predictive capacity
- Increases predictive capacity, see Grimmer et al. (2017)

Predicting counterfactual outcomes in experimental contexts

Suppose we have 8 observations of an outcome, treatment assignment and two covariates:

y	d	Gender	Education
12	1	Female	High
13	1	Female	Low
5	0	Female	High
6	0	Female	Low
7	1	Male	High
8	1	Male	Low
7	0	Male	High
6	0	Male	Low

Table 1: Observed

y	d	Gender	Education
?	0	Female	High
?	0	Female	Low
?	1	Female	High
?	1	Female	Low
?	0	Male	High
?	0	Male	Low
?	1	Male	High
?	1	Male	Low

Table 2: Unobserved counterfactual

$$\text{ATE}_{\text{Observed}} = 10 - 6 = 4$$

Random Forest Estimation

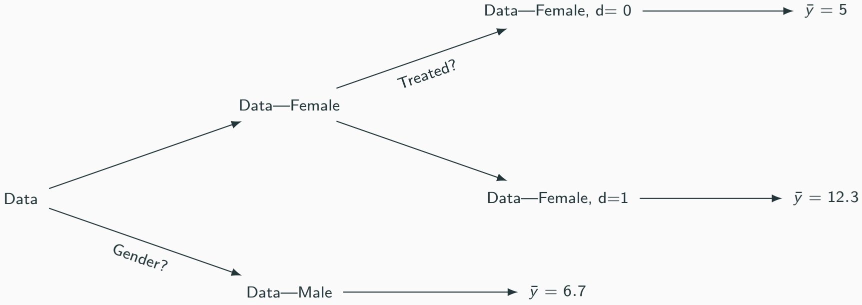
- Random sample of data *and* predictors
- Take bootstrap samples with replacement of training data:

y	d	Gender	Education
12	1	Female	High
13	1	Female	Low
5	0	Female	High
5	0	Female	High
12	1	Female	High
7	0	Male	High
7	0	Male	High
6	0	Male	Low

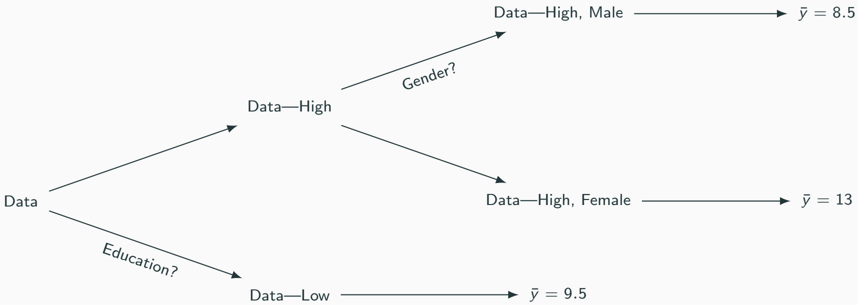
Table 3: Sampled observations (with replacement)

- Construct tree from sample, using random selection of predictor variables

Tree-based logic – Tree #1



Tree-based logic – Tree #2



Random Forest: Estimating the CATE

$$
\underbrace{\begin{pmatrix}
\hat{y}_{i,d=1,t=1} & \hat{y}_{i,d=0,t=1} \\
12.3 & 5 \\
12.3 & 5 \\
12.3 & 5 \\
12.3 & 5 \\
6.7 & 6.7 \\
6.7 & 6.7 \\
6.7 & 6.7 \\
6.7 & 6.7
\end{pmatrix}}_{\text{Tree \#1}}
\overbrace{\begin{pmatrix}
\hat{y}_{i,t=2} \\
13 \\
9.5 \\
13 \\
9.5 \\
8.5 \\
9.5 \\
8.5 \\
9.5
\end{pmatrix}}^{\text{Tree \#2}}
=
\underbrace{\left(\begin{array}{cc|c}
\hat{y}_{i,d=1} & \hat{y}_{i,d=0} & \textbf{CATE} \\
12.7 & 9 & 3.7 \\
10.9 & 7.3 & 3.6 \\
12.7 & 9. & 3.7 \\
10.9 & 7.3 & 3.6 \\
7.6 & 7.6 & 0 \\
8.1 & 8.1 & 0 \\
7.6 & 7.6 & 0 \\
8.1 & 8.1 & 0
\end{array}\right)}_{\text{Average over trees}}
$$

Treatments $d \in \{0, 1\}$, trees $t \in \{0, 1\}$, no. of trees $= T$

Predicted outcome given treatment assignment $d = \hat{y}_{i,d} = \frac{1}{T}\sum_t^T \hat{y}_{i,d}$

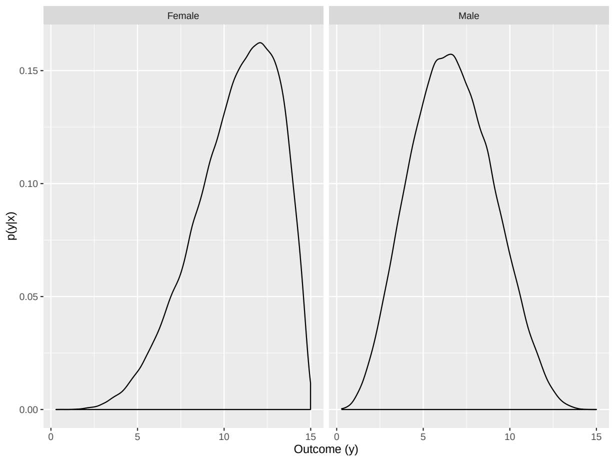
Machine Learning: BART

BART Estimation strategy

- Estimate $f(x) = E(Y|x)$
- Fit a *sequence* of “weak” tree-based regression models
- Each tree contributes a “a small and different portion of f” (Chipman et al 2010)[1]
- Iterative application of sum-of-trees effectively generates a posterior probability distribution of outcomes, given covariate vector X
- From which you can recover $E(Y|x)$ and uncertainty intervals

[1] *BART: Bayesian Additive Regression Trees*, The Annals of Applied Statistics, 2010, Vo.4, No.1

Altered posterior probabilities given covariate values



$$\hat{y} = \frac{1}{K}\sum_{k=1}^{K} k \leftarrow f(x)$$

Estimating the CATE - overall strategy

- BART model estimation generates posterior function of $f(x)$
- Averaging repeat draws from posterior density generates mean outcome for each observation given its vector of predictors x_i
- x_i contains treatment assignment plus other covariates
- Predict $\hat{y}_i$ for two matrices:
    1. Actual observed treatment values (plus covariates)
    2. Counterfactual matrix of reversed treatment assignment $(1 \leftrightarrow 0)$ (plus same covariates)
- For each observation i, we recover two estimates: $y_{i,d=1}$ and $y_{i,d=0}$
- CATE $= y_{i,d=1} - y_{i,d=0}$

Estimating the CATE - generate two test matrices

- Predictions are made using two matrices[2]
- Second matrix is the test dataset in the R code
- Matrices are identical except treatment assignment is reversed in second matrix

$$
\left[\begin{array}{cccc|c}
D_{\textbf{Obs.}} & \textbf{Gender} & \textbf{Education} & y_{i,d} \\
1 & \textit{Female} & \textit{High} & 14 \\
1 & \textit{Female} & \textit{Low} & 12 \\
0 & \textit{Female} & \textit{High} & 4 \\
0 & \textit{Female} & \textit{Low} & 6 \\
1 & \textit{Male} & \textit{High} & 7 \\
1 & \textit{Male} & \textit{Low} & 7 \\
0 & \textit{Male} & \textit{High} & 8 \\
0 & \textit{Male} & \textit{Low} & 6
\end{array}\right]
\left[\begin{array}{ccc|c}
D_{\textbf{Counter.}} & \textbf{Gender} & \textbf{Education} & y_{i,d} \\
0 & \textit{Female} & \textit{High} & 7 \\
0 & \textit{Female} & \textit{Low} & 7 \\
1 & \textit{Female} & \textit{High} & 12 \\
1 & \textit{Female} & \textit{Low} & 13 \\
0 & \textit{Male} & \textit{High} & 8 \\
0 & \textit{Male} & \textit{Low} & 6 \\
1 & \textit{Male} & \textit{High} & 8 \\
1 & \textit{Male} & \textit{Low} & 6
\end{array}\right]
$$

[2] NB: The first, observed matrix is implicitly generated by BART since it is the initial training data (excluding observed outcome)

Estimating the CATE - rearrange matrices

- Matrices can be rearranged such that all observations in matrix 1 are $d = 1$ and *vice versa* for matrix 2
- Covariate information is constant across both matrices

$$
\left[\begin{array}{cccc|c}
D_{\textbf{Obs.}} & \textbf{Gender} & \textbf{Education} & y_{i,d=1} \\
1 & \textit{Female} & \textit{High} & 14 \\
1 & \textit{Female} & \textit{Low} & 12 \\
\color{red}{1} & \textit{Female} & \textit{High} & \color{red}{12} \\
\color{red}{1} & \textit{Female} & \textit{Low} & \color{red}{13} \\
1 & \textit{Male} & \textit{High} & 7 \\
1 & \textit{Male} & \textit{Low} & 7 \\
\color{red}{1} & \textit{Male} & \textit{High} & \color{red}{8} \\
\color{red}{1} & \textit{Male} & \textit{Low} & \color{red}{6}
\end{array}\right]
\left[\begin{array}{ccc|c}
D_{\textbf{Counter.}} & \textbf{Gender} & \textbf{Education} & y_{i,d=0} \\
\color{red}{0} & \textit{Female} & \textit{High} & \color{red}{7} \\
\color{red}{0} & \textit{Female} & \textit{Low} & \color{red}{7} \\
0 & \textit{Female} & \textit{High} & 4 \\
0 & \textit{Female} & \textit{Low} & 6 \\
\color{red}{0} & \textit{Male} & \textit{High} & \color{red}{8} \\
\color{red}{0} & \textit{Male} & \textit{Low} & \color{red}{6} \\
0 & \textit{Male} & \textit{High} & 8 \\
0 & \textit{Male} & \textit{Low} & 6
\end{array}\right]
$$

Estimating the CATE - recover CATE

- CATE $= \hat{y}_{i,d=1} - \hat{y}_{i,d=0}$
- To check for treatment heterogeneity, append covariate information since this is constant across two matrices[3]

$$
\begin{pmatrix} \hat{\mathrm{y}}_{\mathrm{i,d=1}} \\ 14 \\ 12 \\ \color{red}{12} \\ \color{red}{13} \\ 7 \\ 7 \\ \color{red}{6} \\ \color{red}{7} \end{pmatrix} - \begin{pmatrix} \hat{\mathrm{y}}_{\mathrm{i,d=0}} \\ \color{red}{7} \\ \color{red}{7} \\ 4 \\ 6 \\ \color{red}{8} \\ \color{red}{6} \\ 8 \\ 6 \end{pmatrix} = \left(\begin{array}{c|cc} \textbf{CATE} & \textbf{Gender} & \textbf{Education} \\ 7 & \textit{Female} & \textit{High} \\ 5 & \textit{Female} & \textit{Low} \\ 8 & \textit{Female} & \textit{High} \\ 7 & \textit{Female} & \textit{Low} \\ -1 & \textit{Male} & \textit{High} \\ 1 & \textit{Male} & \textit{Low} \\ -2 & \textit{Male} & \textit{High} \\ 1 & \textit{Male} & \textit{Low} \end{array}\right)
$$

[3] NB: all observations are predicted from posterior draws; red numbers indicate predictions using counterfactual treatment assignment

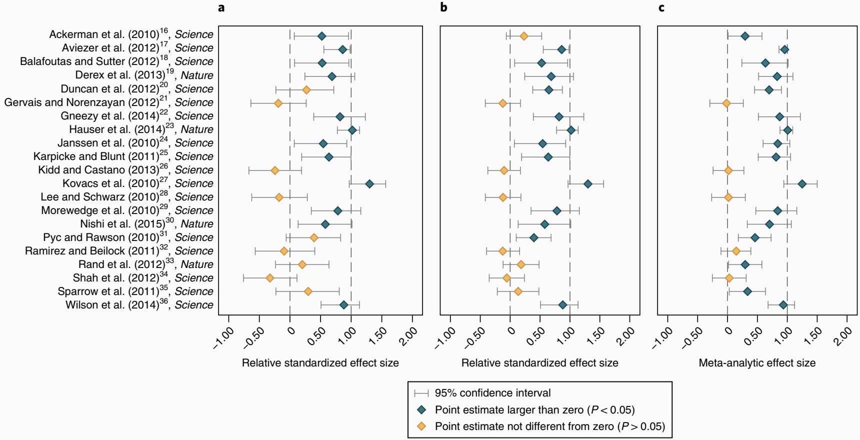
BART: Heterogeneous Treatment Effects

Machine learning, heterogeneity and experimental measurement error

Data Generation

- 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

amazon mechanical turk

LUCID

respondi

KNOWLEDGE NETWORK

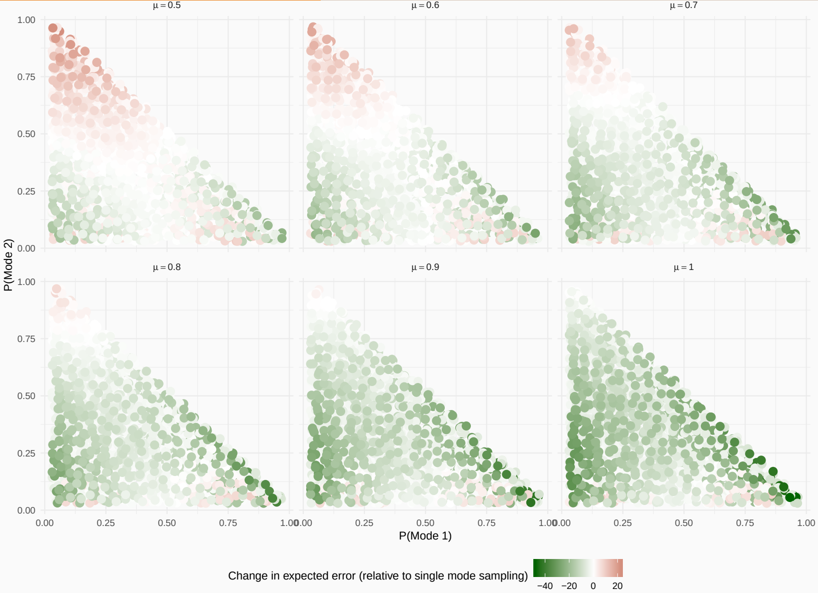
CENTRE FOR EXPERIMENTAL SOCIAL SCIENCES

ANES
American National Election Studies

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

Lying Experiment Design (Duch Laroze Zakharov 2018

- 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

CESS Oxford Lab	CESS Oxford Online	CESS Online UK	Mturk USA
CESS Oxford Subject Pool	CESS Oxford Subject Pool	CESS Online US Subject Pool	U.S. Crowd-sourced workers
1600 decisions/116 subjects	1367 Decisions/144 Subjects	696 Decisions/90 Subjects	2419 Decisions/390 Subjects
57% Rate of Lying	54% Rate of Lying	37% Rate of Lying	40% Rate of Lying

Conventional GLM Estimation

	Mode			
	Lab	Online Lab	Online UK	Mturk
Ability Rank	-0.500^{***}	-0.163^{***}	-0.163^{**}	-0.120^{***}
	(0.036)	(0.045)	(0.071)	(0.037)
20% Deduction	-0.123^{***}			
	(0.024)			
30% Deduction	-0.128^{***}	-0.184^{***}	0.042	0.018
	(0.025)	(0.025)	(0.038)	(0.021)
No Audit	-0.334^{***}	-0.127^{***}	-0.155^{***}	0.011
	(0.023)	(0.026)	(0.036)	(0.024)
Age	0.012^{***}	0.007^{**}	-0.0002	0.002^{**}
	(0.002)	(0.003)	(0.001)	(0.001)
Gender	0.002	0.100^{***}	-0.022	-0.004
	(0.022)	(0.025)	(0.035)	(0.020)

BART Estimation

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

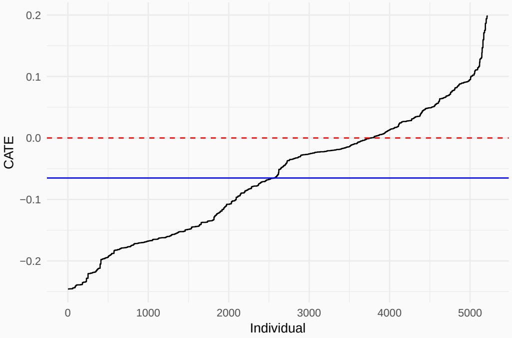
To recover individual estimates of treatment effect:

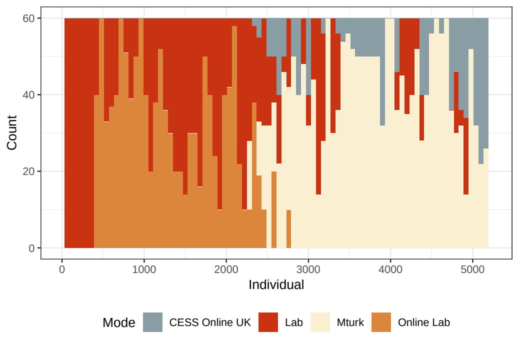
- Assume binary treatment
- Run BART on experimental data (the training set) to generate both model and predicted outcomes for observed data
- Invert treatment assignment of all observations, and pass through model (test set) to generate set of counterfactual predictions
- For each individual, i, $CATE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

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`


CATE
0.2
0.1
0.0
−0.1
−0.2
0
1000
2000
3000
4000
5000
Individual


Count
60
40
20
0
0
1000
2000
3000
4000
5000
Individual
Mode
CESS Online UK
Lab
Mturk
Online Lab

Treatment-adaptive designs

Types of Sequential Randomised Experiments

- Non-adaptive - assignment probabilities fixed
- Treatment-adapative - change based on number of subjects in treatment
- Covariate-adaptive - change based on covariate profiles of new and previous subjects
- Responsive-adaptive - change as function of previous units' outcomes

Treatment-adaptive designs

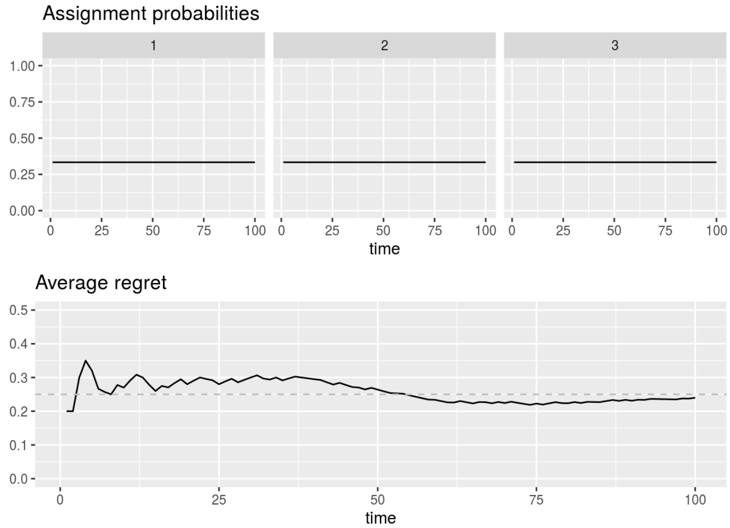
- *ATE* is not always quantity of interest
- Particularly online firms such as Google, Tiktok, FB, etc.
    - Randomly assign sampled users to different arms and dynamically re-orient sample based on which is more successful/more informative
    - Identify which of many will get the most clicks
- But also of interest to political scientists: Ballot initiatives and malfeasance information
- How do adaptive multi-arm trials work?

Regret

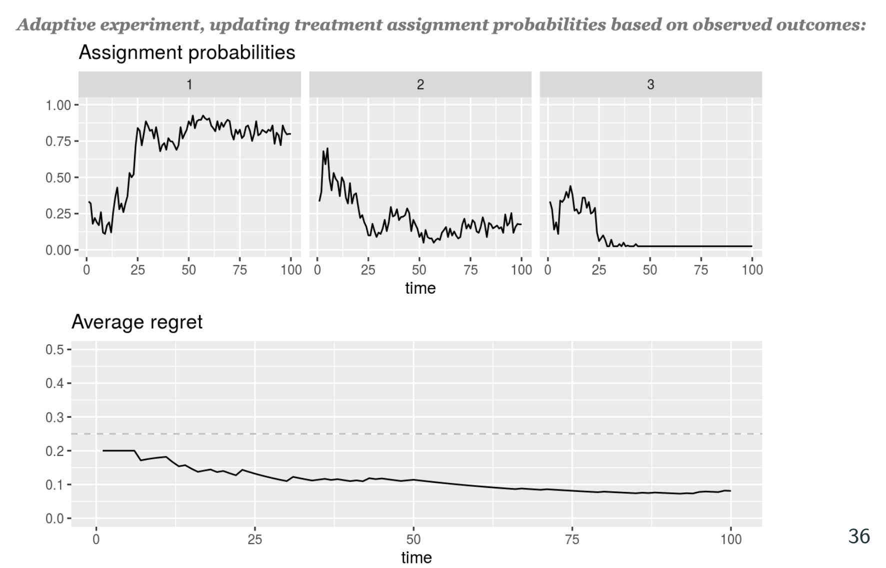
- the difference between the average outcomes we would have observed under optimal assignment and the average outcomes we actually observe under a given assignment algorithm
- Example
  - if the best prototype gives us a 90% click-through rate on average
  - a different prototype gives us a 40% rate on average
  - the regret from assigning the sub-optimal arm is 0.5

Regret: True arms 1 (.8) 2 (.6) 3 (.3)

Random experiment, with balanced assignment probabilities throughout:



Regret: True arms 1 (.8) 2 (.6) 3 (.3)



Molly Offer-Westort et al 2020

Table 2: Study One, Treatments and Outcome Measures

	Minimum Wage	Right to Work
Question Text	Imagine that the following ballot measure were up for a vote in your state. [**ballot measure text**]. If this measure were on the ballot in your state, would you vote in favor or against? [I would vote in favor of this measure; I would vote against this measure]	Imagine that the following ballot measure were up for a vote in your state. [**ballot measure text**]. If this measure were on the ballot in your state, would you vote in favor or against? [I would vote in favor of this measure; I would vote against this measure]
Proposal 1	The measure would: increase the minimum wage [**from {current}**] to {current + 1} per hour, adjusted annually for inflation, and provide that no more than $3.02 per hour in tip income may be used to offset the minimum wage of employees who regularly receive tips.	The measure would [**amend the State Constitution to**]: prohibit, as a condition of employment, forced membership in a labor organization (union) or forced payments of dues or fees, in full or pro-rata ("fair-share"), to a union. The measure will also make any activity which violates employees' rights provided by the bill illegal and ineffective and allow legal remedies for anyone injured as a result of another person violating or threatening to violate those employees' rights. The measure will not apply to union agreements entered into before the effective date of the measure, unless those agreements are amended or renewed after the effective date of the measure.
Proposal 2	The measure would: raise the minimum wage [**from {current}**] to {current + 1} per hour effective September 30th, 2021. Each September 30th thereafter, minimum wage shall increase by $1.00 per hour until the minimum wage reaches {current + 5} per hour on September 30th, 2026. From that point forward, future minimum wage increases shall revert to being adjusted annually for inflation starting September 30th, 2027.	The measure [**reads / would amend the State Constitution to read**]: The right of persons to work may not be denied or abridged on account of membership or nonmembership in any labor union or labor organization, and all contracts in negation or abrogation of such rights are hereby declared to be invalid, void, and unenforceable.
Proposal 3	The measure reads: Shall the minimum wage for adults over the age of 18 be raised [**from {current}**] to {current + 1} per hour by January 1, 2019?	The measure would [**amend the State Constitution to**]: ban any new employment contract that requires employee to resign from or belong to a union, pay union dues, or make other payment to a union. Required contributions to charity or other third party instead of payments to union are also banned. Employees must authorize payroll deduction to unions. Violations of the section is a misdemeanor.
Proposal 4	The measure would: raise the minimum wage [**from {current}**] to {current + 1} per hour worked if the employer provides health benefits, or {current + 2} per hour worked if the employer does not provide health benefits.	The measure [**reads / would amend the State Constitution to read**]: No person shall be deprived of life, liberty or property without due process of law. The right of persons to work shall not be denied or abridged on account of membership or nonmembership in any labor union, or labor organization.
Proposal 5	The measure would: raise the State minimum wage rate [**from {current}**] to at least {current + 1} per hour, and require annual increases in that rate if there are annual increases in the cost of living.	

Boldface text indicates randomly varied elements.

Treatment-adaptive designs

- When researchers are initially agnostic about the relative performance of the K arms, priors are distributed uniformly over parameter space, i.e. $\beta_{1,1}$
- In each t, treatment is randomly assigned according to probability of arms being best (= highest success rate)

$$P\left[\Theta_k = \max_k\{\Theta_1, ..., \Theta_K\} | (X_1^{n_{1,t}}, ..., X_K^{n_{K,t}})\right]$$

for K arms, vector of responses under treatment arm k observed up until and including t $X_k^{\{n_{k,t}\}}$ and Θ_k distributions of success rates

best_binomial_bandit

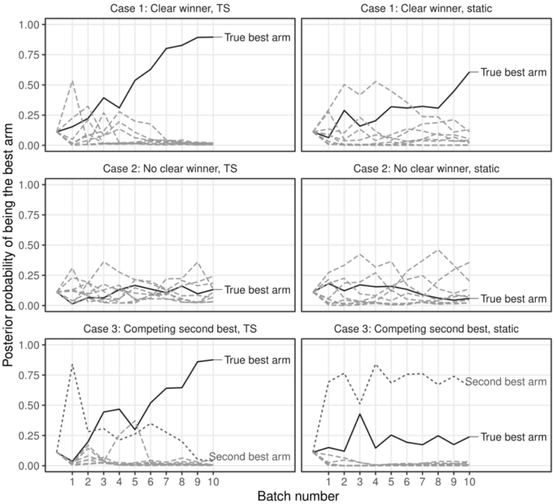
```
x=c(10,20,30,50)
n=c(100,102,120,130)
arm_probabilities =
best_binomial_bandit(x,n) print(arm_probabilities)
[1] 1.611266e-07 8.048293e-04 1.142867e-02 9.877663e-01

sum(arm_probabilities)
[1] 1
```

Treatment-adaptive designs

- Simulations to illustrate design and estimation
- Sample 100 observations for each of 10 periods, updating posterior probability of being best after each period, and assign treatment probabilities in the subsequent period accordingly
- In the first case, one arm has a true 0.20 probability of success, and the remaining 8 arms have a 0.10 probability of success

Treatment-adaptive designs

Figure 1: Simulated Posterior Probabilities Over Time, Thompson Sampling and Static Designs



Ballot Initiative Ex. from Offer-Westort et al 2020

scifile.tex | rankings.R | Chile Corruption MRP.Rmd | Corona_Italy_cleaning.R | Analysis.R | Corona_italy_anal

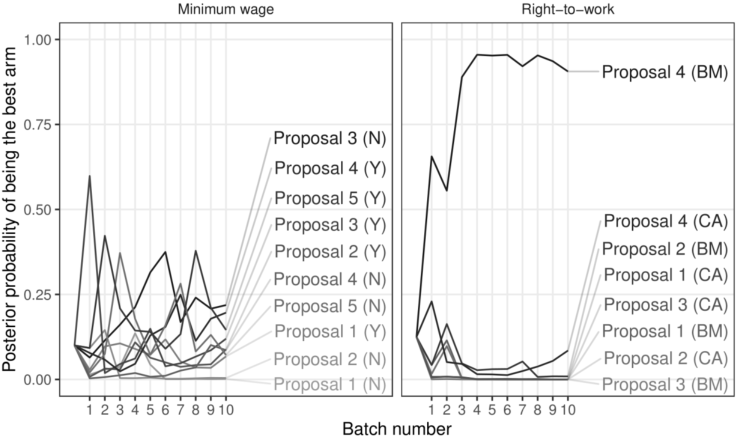
Filter

	id	Y_mw	Y_rtw	Z_mw	Z_rtw	weights_rtw	prob_rtw	weights_mw	prob_mw	batch
1	1	1	1	Proposal 4 (N)	Proposal 2 (CA)	8.000000	0.12500	10.000000	0.10000	1
2	2	1	1	Proposal 4 (N)	Proposal 4 (BM)	8.000000	0.12500	10.000000	0.10000	1
3	3	1	0	Proposal 3 (N)	Proposal 1 (CA)	8.000000	0.12500	10.000000	0.10000	1
4	4	1	0	Proposal 2 (N)	Proposal 3 (BM)	8.000000	0.12500	10.000000	0.10000	1
5	5	1	1	Proposal 2 (N)	Proposal 2 (BM)	8.000000	0.12500	10.000000	0.10000	1
6	6	1	1	Proposal 3 (N)	Proposal 4 (BM)	8.000000	0.12500	10.000000	0.10000	1
7	7	1	1	Proposal 1 (N)	Proposal 3 (BM)	8.000000	0.12500	10.000000	0.10000	1
8	8	1	1	Proposal 2 (Y)	Proposal 3 (BM)	8.000000	0.12500	10.000000	0.10000	1
9	9	1	1	Proposal 4 (Y)	Proposal 3 (BM)	8.000000	0.12500	10.000000	0.10000	1
10	10	1	1	Proposal 3 (N)	Proposal 1 (CA)	8.000000	0.12500	10.000000	0.10000	1
11	11	1	1	Proposal 3 (Y)	Proposal 1 (CA)	8.000000	0.12500	10.000000	0.10000	1
12	12	1	1	Proposal 5 (Y)	Proposal 4 (CA)	8.000000	0.12500	10.000000	0.10000	1
13	13	0	1	Proposal 5 (Y)	Proposal 2 (BM)	8.000000	0.12500	10.000000	0.10000	1
14	14	1	1	Proposal 5 (Y)	Proposal 3 (CA)	8.000000	0.12500	10.000000	0.10000	1
15	15	1	1	Proposal 1 (Y)	Proposal 1 (BM)	8.000000	0.12500	10.000000	0.10000	1
16	16	1	1	Proposal 3 (Y)	Proposal 4 (CA)	8.000000	0.12500	10.000000	0.10000	1
17	17	1	0	Proposal 1 (Y)	Proposal 2 (CA)	8.000000	0.12500	10.000000	0.10000	1
18	18	1	1	Proposal 5 (Y)	Proposal 1 (BM)	8.000000	0.12500	10.000000	0.10000	1

Showing 1 to 18 of 1,000 entries, 10 total columns

Treatment-adaptive designs

Figure 3: Study One, Overtime Posterior Probabilities



Adaptive experimentation tutorial

- Molly Offer-Westort, Vitor Hadad, Susan Athey
- https://mollyow.shinyapps.io/adaptive/

Small Area Estimation and Post-stratified Treatment Effects

Economic Preferences

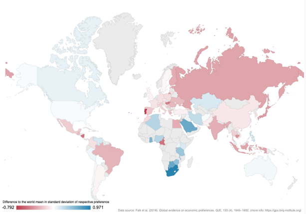


Figure 1: Risk

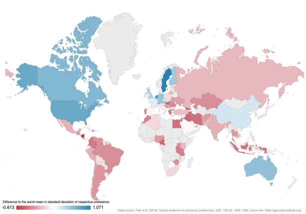


Figure 2: Patience

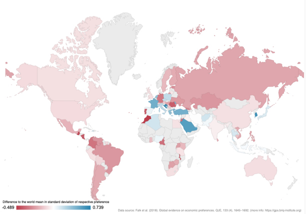


Figure 3: Neg. Reciproc.

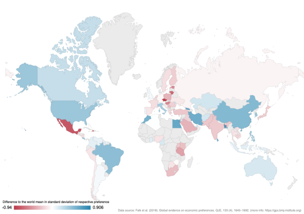
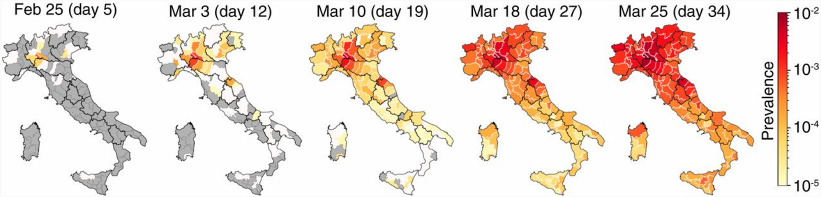
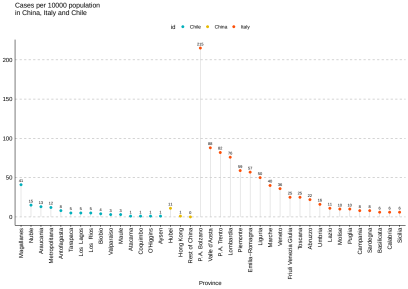


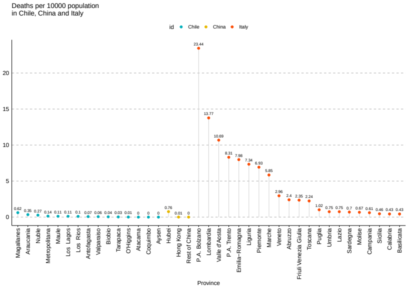
Figure 4: Altruism

LOCAL Shock: COVID-19 Community Infections Italy



Design


Cases per 10000 population
in China, Italy and Chile
id Chile China Italy
0
50
100
150
200
215
88 82 76 59 57 50 40 36 25 25 22 16 11 10 10 8 8 6 6 6
41 15 13 12 8 5 5 5 4 3 3 1 1 1 1
11 1 0
Magallanes
Nuble
Araucania
Metropolitana
Antofagasta
Tarapaca
Los Lagos
Los Rios
Biobio
Valparaiso
Maule
Atacama
Coquimbo
O'Higgins
Aysen
Hubei
Hong Kong
Rest of China
P.A. Bolzano
Valle d'Aosta
P.A. Trento
Lombardia
Piemonte
Emilia−Romagna
Liguria
Marche
Veneto
Friuli Venezia Giulia
Toscana
Abruzzo
Umbria
Lazio
Molise
Puglia
Campania
Sardegna
Basilicata
Calabria
Sicilia
Province


Deaths per 10000 population
in Chile, China and Italy
id Chile China Italy
0
5
10
15
20
23.44
13.77 10.69 8.31 7.98 7.34 6.93 5.85 2.96 2.4 2.35 2.24 1.02 0.75 0.75 0.7 0.67 0.61 0.46 0.43 0.43
0.62 0.35 0.27 0.14 0.11 0.11 0.1 0.07 0.06 0.04 0.03 0.01 0 0 0
0.76 0.01 0
Magallanes
Araucania
Nuble
Metropolitana
Maule
Los Lagos
Los Rios
Antofagasta
Valparaiso
Biobio
Tarapaca
O'Higgins
Atacama
Coquimbo
Aysen
Hubei
Hong Kong
Rest of China
P.A. Bolzano
Lombardia
Valle d'Aosta
P.A. Trento
Emilia−Romagna
Liguria
Piemonte
Marche
Veneto
Abruzzo
Friuli Venezia Giulia
Toscana
Puglia
Umbria
Lazio
Sardegna
Molise
Campania
Sicilia
Calabria
Basilicata
Province

COVID Awareness: Chile, China, Italy

	Severity City			People Died (Log)		
	Chile	*China*	*Italy*	*Chile*	*China*	*Italy*
Infection	0.398***	0.057***	0.048***	0.005	0.001***	0.010***
	(0.072)	(0.008)	(0.005)	(0.004)	(0.000)	(0.001)
Education	0.483**	-0.289**	-0.479***	-0.006	0.000	0.039
	(0.179)	(0.110)	(0.046)	(0.015)	(0.007)	(0.106)
Age	0.017***	0.007	0.017*	0.000	-0.001***	0.009***
	(0.004)	(0.012)	(0.008)	(0.001)	(0.000)	(0.000)
Male	0.024	-0.117	-0.406***	-0.011	-0.008	0.040
	(0.162)	(0.109)	(0.034)	(0.021)	(0.006)	(0.052)
GDP	-0.152*	-0.138	-0.475***	0.022*	-0.006***	-0.001
	(0.071)	(0.183)	(0.034)	(0.011)	(0.002)	(0.013)
(Intercept)	5.563***	6.186***	4.876***	0.065*	0.050***	-0.219***
	(0.279)	(0.492)	(0.114)	(0.027)	(0.012)	(0.062)
N	820	1500	773	820	1500	773
Country FE	No	No	No	No	No	No

Table 5: Time Preferences: Chile, China, Italy

	Patience (GPS)			
	Pooled	*Chile*	*China*	*Italy*
Infection Rate	0.024	0.029	0.055***	0.022***
	(0.013)	(0.179)	(0.015)	(0.004)
Education Medium	-1.807***	-2.760***	-1.567***	-1.817*
	(0.359)	(0.620)	(0.467)	(0.736)
Age	-0.025	-0.006	0.088	-0.117**
	(0.030)	(0.021)	(0.062)	(0.039)
Male	-0.076	0.586	-1.448*	1.924***
	(0.575)	(1.319)	(0.635)	(0.363)
GDP	0.382	0.193	0.733	-0.449***
	(0.259)	(0.481)	(0.381)	(0.130)
(Intercept)	9.740***	9.024***	12.318***	18.752***
	(1.051)	(0.565)	(2.368)	(1.541)
Observations	2881	780	1426	675
Country FE?	Yes	No	No	No

***$= p < 0.001$, **$= p < 0.01$, *$= p < 0.05$.

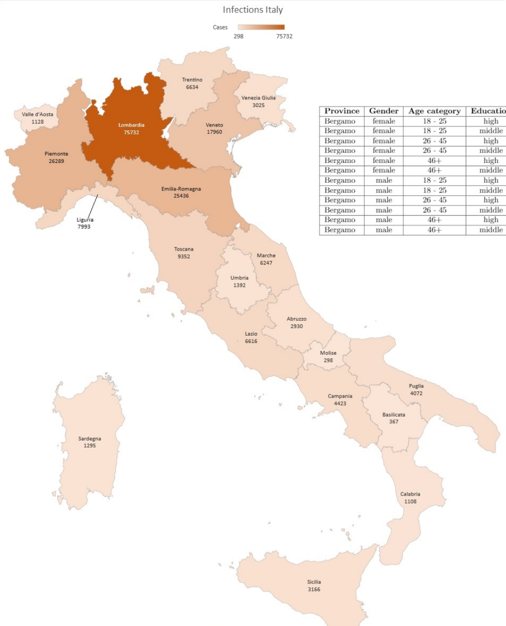
Consequential Infection Effects? Small Area Estimation

- *experiment training data*: individual-level results that allow for demographic profiles to be correlated with decisions/treatment effect;

- *stratification frame*: a vector of counts for a set of *cells*, or mutually-exclusive and exhaustive subject categories, in the population;

- *learner* - an algorithm which can best summarize the relationships implied by the individual-level dataset and coherently project these onto subject-categories we have not yet seen in the sample (‘out-of-sample’ predictions)

Stratification Frame



Province	Gender	Age category	Education	N	Infection category	GDP category
Bergamo	female	18 - 25	high	11	medium	308.4 - 446.6
Bergamo	female	18 - 25	middle	94	medium	308.4 - 446.6
Bergamo	female	26 - 45	high	82	medium	308.4 - 446.6
Bergamo	female	26 - 45	middle	269	medium	308.4 - 446.6
Bergamo	female	46+	high	62	medium	308.4 - 446.6
Bergamo	female	46+	middle	801	medium	308.4 - 446.6
Bergamo	male	18 - 25	high	11	medium	308.4 - 446.6
Bergamo	male	18 - 25	middle	109	medium	308.4 - 446.6
Bergamo	male	26 - 45	high	61	medium	308.4 - 446.6
Bergamo	male	26 - 45	middle	275	medium	308.4 - 446.6
Bergamo	male	46+	high	67	medium	308.4 - 446.6
Bergamo	male	46+	middle	727	medium	308.4 - 446.6

Learner

Model the response for each individual i:

$$Y_i = \beta^0 + \alpha^{gender}_{j[i]} + \alpha^{age}_{k[i]} + \alpha^{education}_{l[i]} + \alpha^{province}_{p[i]} \tag{1}$$

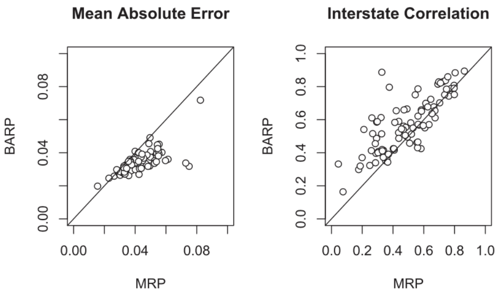
$$\begin{aligned} \alpha^{gender}_{j} &\approx N(0, \sigma^2_{gender}), \text{for } j = 1, 2 \\ \alpha^{age}_{k} &\approx N(0, \sigma^2_{age}), \text{for } k = 1, ..., 4 \\ \alpha^{educate}_{l} &\approx N(0, \sigma^2_{educate}), \text{for } k = 1, 2 \end{aligned} \tag{2}$$

The effects associated with the Italian provinces:

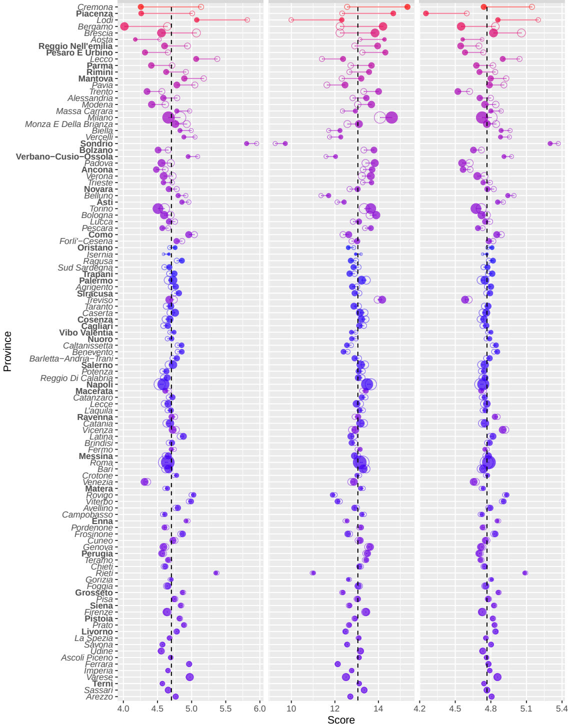
$$\alpha^{province}_{p} \approx N(\beta^{\mathrm{GDPpc}} \cdot \mathrm{GDPpc} + \beta^{\mathrm{COVID}} \cdot \mathrm{COVID\ Infection}, \sigma^2), \text{for } p \tag{3}$$

BART Estimation: Bisbee 2020

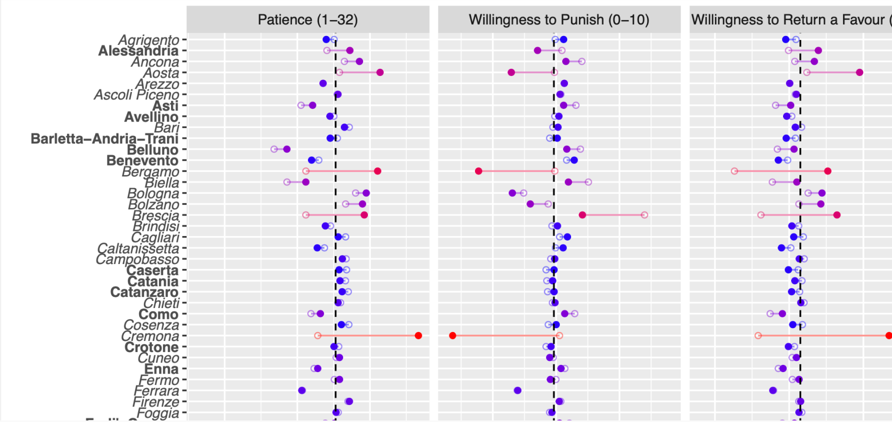
FIGURE 1. Predictive Accuracy



Notes: Predictive accuracy of BARP (y-axes) versus MRP (x-axes) across 89 surveys as measured by mean absolute error (left panel) and interstate correlation (right panel).


Province
Cremona
Piacenza
Lodi
Bergamo
Brescia
Aosta
Reggio Nell'emilia
Pesaro E Urbino
Lecco
Parma
Rimini
Mantova
Pavia
Trento
Alessandria
Modena
Massa Carrara
Milano
Monza E Della Brianza
Biella
Vercelli
Sondrio
Bolzano
Verbano−Cusio−Ossola
Padova
Ancona
Verona
Trieste
Novara
Belluno
Asti
Torino
Bologna
Lucca
Pescara
Como
Forli'−Cesena
Oristano
Isernia
Ragusa
Sud Sardegna
Trapani
Palermo
Agrigento
Siracusa
Treviso
Taranto
Caserta
Cosenza
Cagliari
Vibo Valentia
Nuoro
Caltanissetta
Benevento
Barletta−Andria−Trani
Salerno
Potenza
Reggio Di Calabria
Napoli
Macerata
Catanzaro
Lecce
L'aquila
Ravenna
Catania
Vicenza
Latina
Brindisi
Fermo
Messina
Roma
Bari
Crotone
Venezia
Matera
Rovigo
Viterbo
Avellino
Campobasso
Enna
Pordenone
Frosinone
Cuneo
Genova
Perugia
Teramo
Chieti
Rieti
Gorizia
Foggia
Grosseto
Pisa
Siena
Firenze
Pistoia
Prato
Livorno
La Spezia
Savona
Udine
Ascoli Piceno
Ferrara
Imperia
Varese
Terni
Sassari
Arezzo
4.0
4.5
5.0
5.5
6.0
10
12
14
4.2
4.5
4.8
5.1
5.4
Score

Small Area Estimation: Italy



Italy Population Affected