Planning the Optimal Get-out-the-vote Campaign Using Randomized Field Experiments

Kosuke Imai

Princeton University

April 15, 2009
Joint work with Aaron Strauss

Get Out the Vote!

- Political scientists: Why do people vote?
- Activists and policy makers: How to increase turnout?
- Politicians: How to win an election?
- Statisticians: How to get most out of election data?

- Use of field experiments: Gosnell (1927), Eldersveld (1956)
- Resurgence of randomized experiments in social sciences
- Over 100 GOTV experiments since Gerber and Green (2000)
  - method: door-to-door, phone, mail, Internet, ...
  - timing: repeated contact, time to election, ...
  - message: reminder, civic duty, close election, ...

The Goal of this Project

- To inform GOTV campaign planners by analyzing field experiments
- Derive an optimal GOTV campaign strategy from experimental data
- Little work after Kramer (1966)'s pioneering article

- Problems of current practice:
  - Often, only the estimated *average* treatment effects are reported
  - Sometimes, the constant additive effect assumption is made
  - Not helpful for deciding whom to contact with what method
  - Some researchers report post-hoc subgroup analysis
  - Danger of overfitting and multiple testing

- Need a principled statistical approach

Overview of the Proposed Approach

- Statistical decision theory
  - Nonpartisan: increase turnout
  - Partisan: win an election
  - Learn from experiments and decide whom to contact with what method

- Modern optimization and statistical methods
  - Nonparametric method for flexible modeling
  - Cross-validation to avoid overfitting
  - Constrained optimization with a budget constraint

- Empirical evaluations
  - How well does the proposed method work in practice?
  - Use cross-validation to evaluate the performance
  - 2 nonpartisan data and 2 partisan data

Nonpartisan GOTV Campaign Planner's Decision Problem

- Target population of size N
- Known distribution of covariates (e.g., voter roll) $P(X)$
- Age, address, party registration, turnout history, etc.
- K mobilization strategies: $T = 0$ do nothing and $T \in \{0, \ldots, K-1\}$
- Known overhead and per-voter costs: $\kappa(t)$ and $\xi(t, x)$
- Overall budget constraint C
- Feasible mobilization strategy:

$$\delta(\cdot, \cdot) : (\mathcal{T}, \mathcal{X}) \mapsto [0, 1] \quad \text{s.t.} \quad \sum_{t=0}^{K-1} \delta(t, x) = 1$$

- Planner's objective function:

$$g(\delta, \rho) \;=\; N \sum_{x \in \mathcal{X}} P(X = x) \sum_{t=0}^{K-1} \delta(t, x) \rho(t, x)$$

where $\rho(t, x) = \Pr(Y_i(t) = 1 \mid X_i = x)$ is the turnout profile

The Optimal Campaign Strategy Without Uncertainty

- For the moment, assume turnout profile is known
- Optimal campaign strategy:

$$\delta^* \;=\; \operatorname*{argmax}_{\delta} \sum_{x \in \mathcal{X}} P(X = x) \sum_{t=0}^{K-1} \delta(t, x) \; \rho(t, x)$$

- Budget constraint:

$$\sum_{t=1}^{K-1} \mathbf{1}\left\{ \sum_{x \in \mathcal{X}} \delta(t, x) \neq 0 \right\} \kappa(t) + N \sum_{x \in \mathcal{X}} P(X = x) \sum_{t=1}^{K-1} \delta(t, x) \xi(t, x) \;\leq\; C,$$

- (Constrained) linear programming problem

The Bayesian Planner

- Prior: $\pi(\rho)$
- Availability of a randomized field experiment
  - A random sample from the same population
  - Random assignment of the same "treatments"
  - Stability assumption: treatment effects don't change between the experiment and the election
- Posterior: $\pi(\rho \mid D)$
- The Bayesian planner's optimal strategy:

$$\begin{aligned}\delta^* &= \underset{\delta}{\operatorname{argmax}} \int g(\delta, \pi)\, d\pi(\rho \mid D) \\ &= \underset{\delta}{\operatorname{argmax}} \sum_{x \in \mathcal{X}} P(X = x) \sum_{t=0}^{K-1} \delta(t, x)\ \tilde{\rho}(t, x)\end{aligned}$$

where $\tilde{\rho}(t, x)$ is the posterior mean of $\rho(t, x)$

Partisan GOTV Campaign Planner's Decision Problem

- Outcome variable: (1) vote for my candidate $Y_i = 1$, (2) vote for my opponent $Y_i = -1$, (3) abstain $Y_i = 0$
- Assumption: No viable third-party candidate
- Vote share differential:

$$\nu(t, x) \equiv \frac{\sum_{i \in \{i' : X_{i'} = x\}} Y_i(t)}{\sum_{i=1}^{N} \mathbf{1}\{X_i = x\}}$$

- Objective function: $0 - 1$ loss function

$$h(\delta, \nu) \equiv \mathbf{1}\left\{\sum_{x \in X} P(X = x) \sum_{t=0}^{K-1} \delta(t, x) \nu(t, x) > 0\right\}$$

- Same budget constraint as in nonpartisan case
- The Bayesian optimal strategy:

$$\delta^* = \underset{\delta}{\operatorname{argmax}} \int h(\delta, \nu)\ d\pi(\nu \mid D)$$

The Optimization Method

- Two-step approximations:
  1. Monte Carlo approximation to integral
  2. Sigmoid approximation to sum of indicator functions
- Continuously differentiable objective function:

$$\delta^* \approx \underset{\delta \in \Delta}{\operatorname{argmax}} \frac{1}{M} \sum_{m=1}^{M} s_{\sigma_M} \left(\sum_{x \in X} P(X = x) \sum_{t=0}^{K-1} \delta(t, x) \nu^{(m)}(t, x) \right),$$

  where $s_{\sigma_M}(u) = 1/\{1 + \exp(-u/\sigma_M)\}$ and $\nu^{(m)}$ is the mth Monte Carlo draw from $\pi(\nu \mid D)$
- Constrained convex programming problem
- Choose $\sigma_M > 0$ small s.t. $\lim_{M \to \infty} \sigma_M = 0$
- Almost sure convergence (Ma and Huang, 2007)

Quasi-Bayesian Approach

- Fully Bayesian method is possible
- Computationally costly, difficulty of eliciting prior on parameters
- Compromise:
  - Use a tree-based classification method to identify subgroups
  - Beta-Binomial update within each subgroup
  - Normal-Normal update when prior is specified on vote share differential or treatment effects
- Tree-based method:
  1. nonparametric
  2. yields easy-to-interpret subgroups (allows for informative prior)
  3. can effectively deal with categorical outcome and covariates
  4. relatively fast computation
- Cross-validation to avoid post-hoc subgroup analysis problem

The Details of the Model Selection Algorithm

- Relies on Gunter, Zhu and Murphy (2007)
- **Step 1:** Find predictive variables
  1. Fit a model with all covariates X and T
  2. Record the chosen variables (V) as predictive variables
- **Step 2:** Order variables by their prescriptive value
  1. For each X_j, fit a model with X_j, T, V, and $X_j \times T$
  2. Derive the expected turnout under the optimal strategy
  3. Do the same for a tree created only with T and V
  4. Order X_j by the difference between these two optimal turnout rates
- **Step 3:** Select the final model via cross-validation
  1. Fit models by adding X_j in the order of its prescriptive importance to the base model with T and V
  2. For each model, choose complexity parameters via cross-validation on the resulting (estimated) optimal turnout
  3. Choose the model with the highest resulting turnout

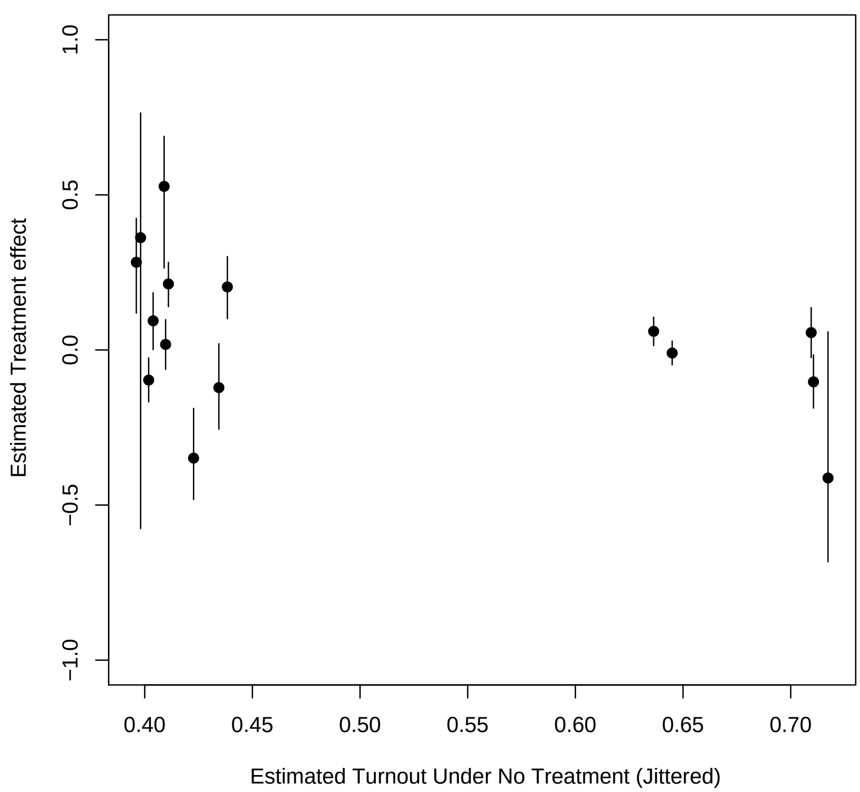
Empirical Evaluations of the Proposed Method

- **Critical question**: How well does our method work in practice?
- Ideally, we want to randomize the use of our method in real elections
- Instead, we artificially create elections from existing data
- Two nonpartisan experiments and two partisan experiments
- Procedure:
  1. Set aside a random subset of experimental data as test data
  2. Apply the proposed method to the remaining data and derive the "optimal" campaign strategy
  3. Use the test data to obtain an unbiased estimate of the resulting turnout or probability of winning under this optimal strategy
  4. Compare it with the other strategies under various budget constraints
- 10 fold cross-validation for a small sample

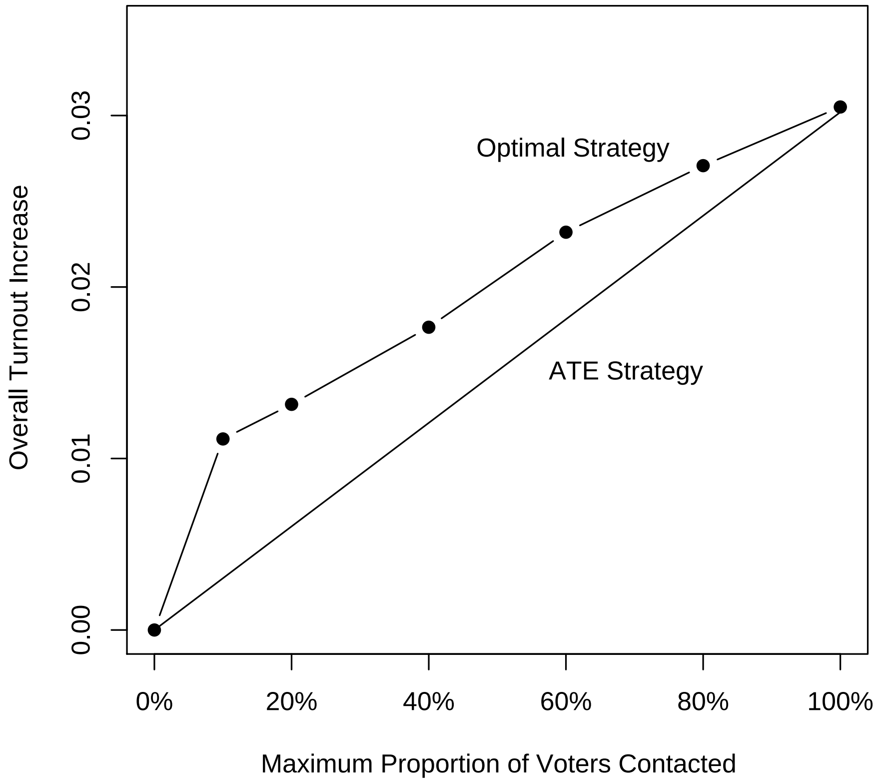
Evaluation I: Text Messaging Experiment

- Nonpartisan, Single treatment (Dale and Strauss, 2008)
- Subjects: 8,000 registrants in Aug–Oct 2006
- `A friendly reminder that TOMORROW is Election Day. Democracy depends on citizens like you - so please vote!`
- The estimated ATE of 3.0 percentage points (s.e. 1.1) increase in voter-file validated turnout
- Available covariates: gender, age, race, past voting history, county population density, and registering organization
- Age is the most prescriptive variable

Treatment Effect Heterogeneity Across Subgroups



The Evaluation Result



Evaluation II: Social Pressure Experiment

- Nonpartisan, four mailing treatments (Gerber *et al.* 2008)
- Civic duty, Hawthorne, self turnout, self+neighbors turnout
- Covariates: age, gender, household size, voting history
- Over 300,000 participants
- Training set of 200,000 and test set of 100,000
- Self+neighbors treatment had the largest estimated ATE = 9 percentage points increase!

Self+Neighbors Treatment Postcard

Dear Registered Voter:

WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

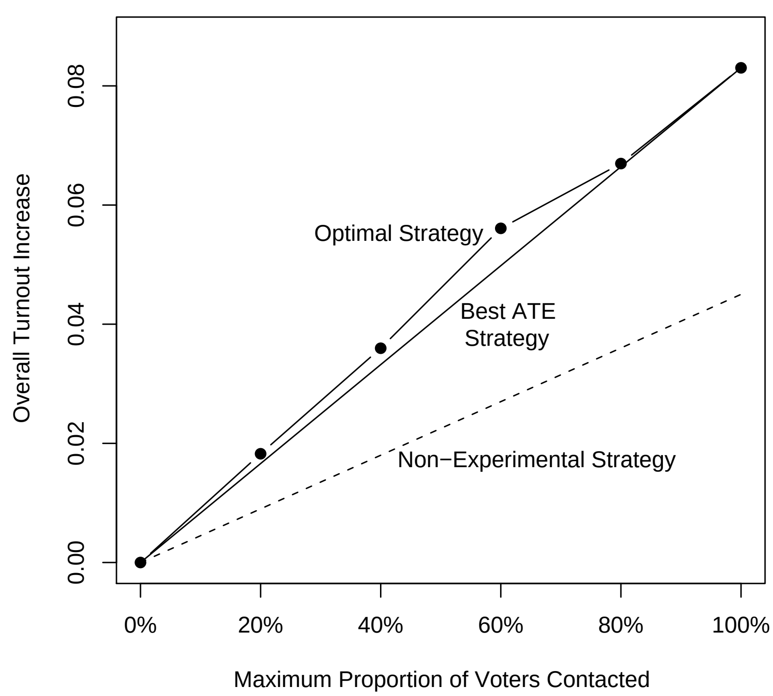
The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

DO YOUR CIVIC DUTY — VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	Voted	Voted	______
9995 JENNIFER KAY SMITH		Voted	______
9997 RICHARD B JACKSON		Voted	______
9999 KATHY MARIE JACKSON		Voted	______
9999 BRIAN JOSEPH JACKSON		Voted	______
9991 JENNIFER KAY THOMPSON		Voted	______
9991 BOB R THOMPSON		Voted	______

The Evaluation Result

- Very little heterogeneity
- The proposed method appears to do no harm



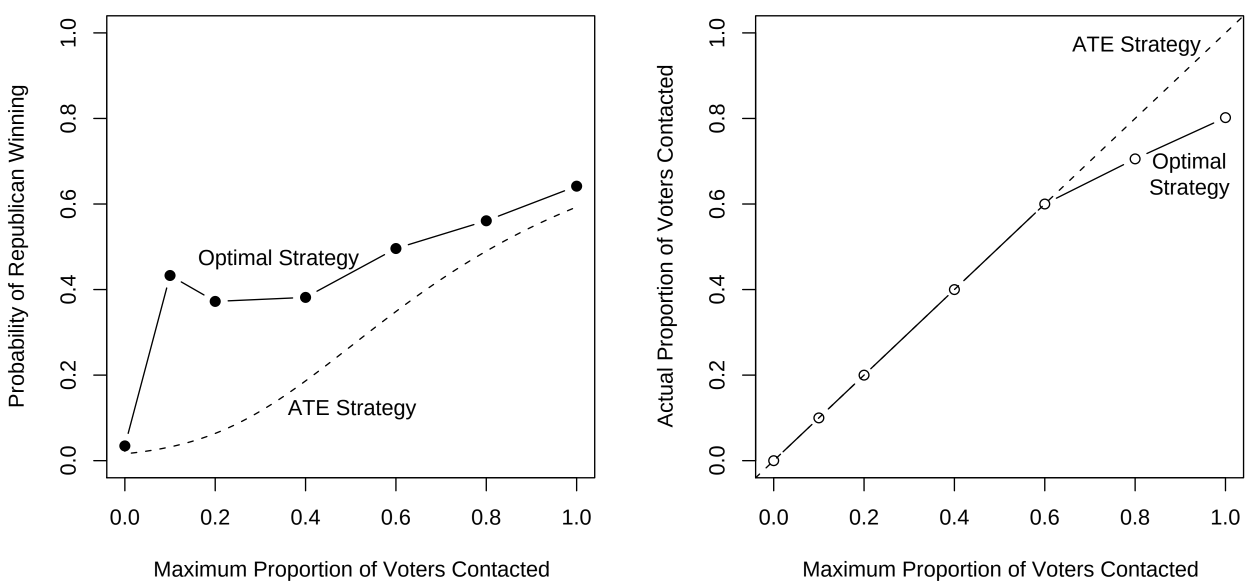
Evaluation III: Partisan Persuasion Experiment

- Partisan, single treatment (Arceneaux and Kolodny, 2007)
- A "well-known liberal activist group" endorsed two Democratic candidates
- A post-election survey of 2000 voters
- Covariates: gender, age, party identification, level of political interest, past voting history, etc.

- This mobilization failed miserably...
- The estimated ATE is negative 6 percentage points in vote margin with the standard error of 4 percentage points
- The optimal strategy: do nothing!
- We reanalyze the data from the Republican's perspective

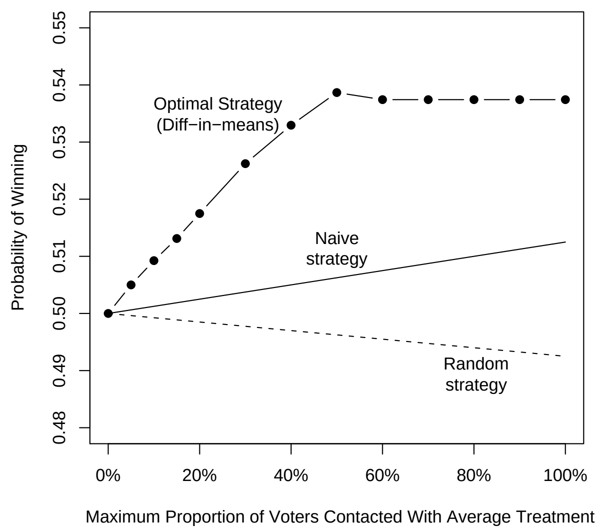
The Evaluation Result

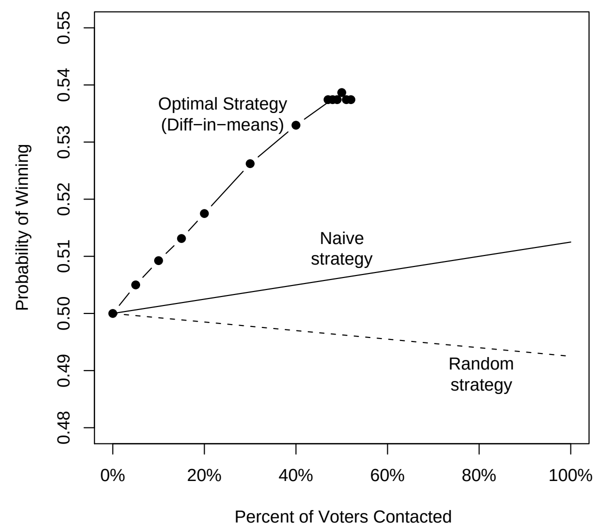
- predictive (and possibly prescriptive) variables: past vote, party id
- prescriptive variable: political interest



Evaluation IV: "Cannot-be-Published" Experiment

- Multiple treatments with differing costs
- The most successful example!





Concluding Remarks and Future Work

- Potential to use social science experiments and statistics in order to inform policy making
- In three of our four evaluations, the optimal strategy is found to be much more cost-effective

- Improving statistical methods: simplicity, transparency, theoretical properties
- Formulating the problem as the variable selection problem
- The model with many potential interaction terms
- Application of LASSO and SCAD etc.
- Challenge: How to incorporate the planner's objective function