Randomization Inference with Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election

Kosuke Imai

Department of Politics
Princeton University

Joint work with
Daniel E. Ho, Yale Law School

June 20, 2005

Overview

Overview

- The Impact of Ballot Format on Voting:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.

- Randomized Natural Experiments:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.

- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.

Overview

- The Impact of Ballot Format on Voting:
  - Butterfly ballot in 2000 Florida.
  - Voting equipment.
  - Partisan labels.
  - Ballot order.

- Randomized Natural Experiments:
  - Treatment is physically randomized.
  - Experiment is conducted in real elections.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his “Lady Tasting Tea” experiment.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his “Lady Tasting Tea” experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model: Probability measure generated by random assignment.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model: Probability measure generated by random assignment.
  - ⋆ No large-sample approximation:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model: Probability measure generated by random assignment.
  - ⋆ No large-sample approximation: *Exact* test.

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model: Probability measure generated by random assignment.
  - ⋆ No large-sample approximation: $Exact$ test.
  - ⋆ Confidence intervals with correct coverage:

Overview

- The Impact of Ballot Format on Voting:
  - ⋆ Butterfly ballot in 2000 Florida.
  - ⋆ Voting equipment.
  - ⋆ Partisan labels.
  - ⋆ Ballot order.
- Randomized Natural Experiments:
  - ⋆ Treatment is physically randomized.
  - ⋆ Experiment is conducted in real elections.
  - ⋆ California alphabet lottery: randomization-rotation procedure.
  - ⋆ 2003 recall election to estimate the ballot page effect.
- Randomization Inference:
  - ⋆ R. A. Fisher used it first in his "Lady Tasting Tea" experiment.
  - ⋆ No distributional assumption: Nonparametric method.
  - ⋆ No probability model: Probability measure generated by random assignment.
  - ⋆ No large-sample approximation: $Exact$ test.
  - ⋆ Confidence intervals with correct coverage: Inverting the test.

The California Alphabet Lottery

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.

- A random alphabet is drawn for every statewide election that applies to all statewide offices

The California Alphabet Lottery

- Started in 1975: “[B]oth the ‘incumbent first’ and ‘alphabetical order’ procedures are constitutionally impermissible.” *Gould v. Grubb*, 14 Cal. 3d 661, 676.

- A random alphabet is drawn for every statewide election that applies to all statewide offices

- California Elections Code 13112(a) holds:

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.

- A random alphabet is drawn for every statewide election that applies to all statewide offices

- California Elections Code 13112(a) holds:

  > Each letter of the alphabet shall be written on a separate slip of paper, each of which shall be folded and inserted into a capsule.

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.

- A random alphabet is drawn for every statewide election that applies to all statewide offices

- California Elections Code 13112(a) holds:

  > Each letter of the alphabet shall be written on a separate slip of paper, each of which shall be folded and inserted into a capsule. **Each capsule shall be opaque and of uniform weight, color, size, shape, and texture.**

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.

- A random alphabet is drawn for every statewide election that applies to all statewide offices

- California Elections Code 13112(a) holds:

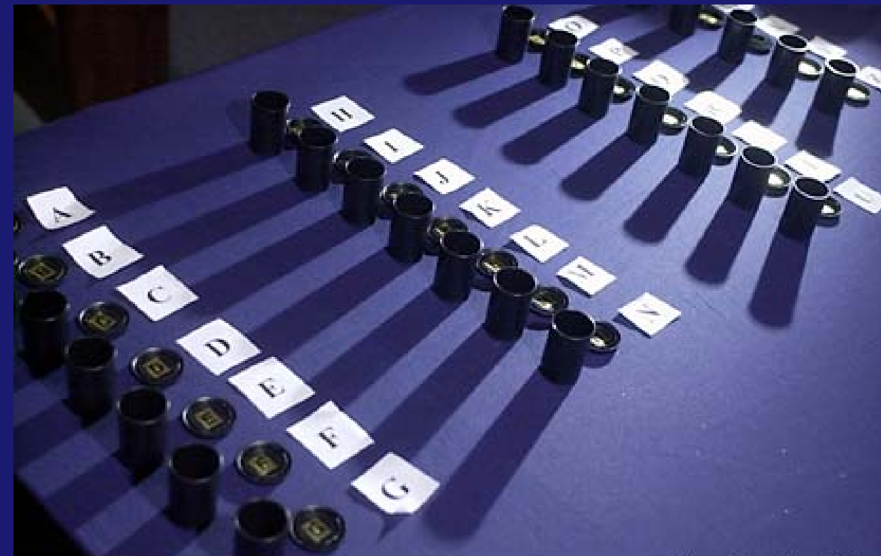
> Each letter of the alphabet shall be written on a separate slip of paper, each of which shall be folded and inserted into a capsule. Each capsule shall be opaque and of uniform weight, color, size, shape, and texture. The capsules shall be placed in a container, which shall be shaken vigorously in order to mix the capsules thoroughly.

The California Alphabet Lottery

- Started in 1975: "[B]oth the 'incumbent first' and 'alphabetical order' procedures are constitutionally impermissible." *Gould v. Grubb*, 14 Cal. 3d 661, 676.
- A random alphabet is drawn for every statewide election that applies to all statewide offices
- California Elections Code 13112(a) holds:

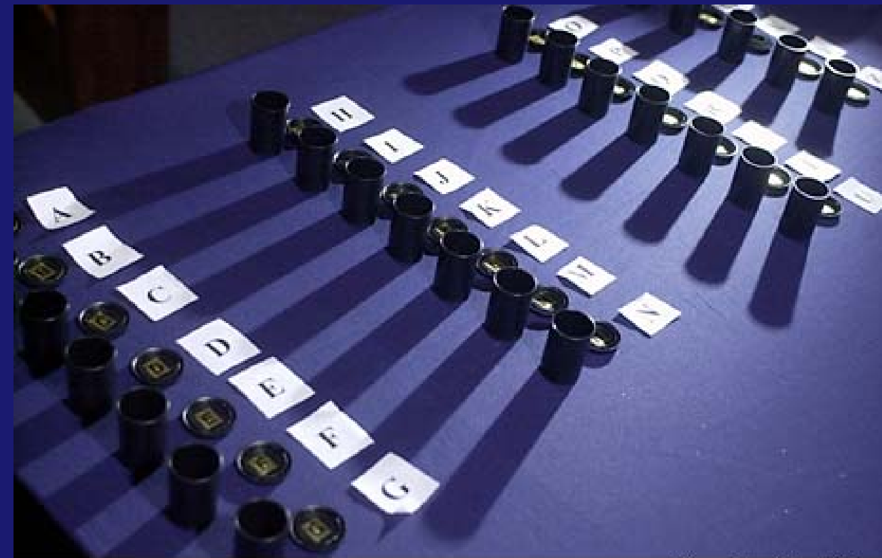
> Each letter of the alphabet shall be written on a separate slip of paper, each of which shall be folded and inserted into a capsule. Each capsule shall be opaque and of uniform weight, color, size, shape, and texture. The capsules shall be placed in a container, which shall be shaken vigorously in order to mix the capsules thoroughly. The container then shall be opened and the capsules removed at random one at a time. As each is removed, it shall be opened and the letter on the slip of paper read aloud and written down. The resulting random order of letters constitutes the randomized alphabet, which is to be used in the same manner as the conventional alphabet in determining the order of all candidates in all elections. For example, if two candidates with the surnames Campbell and Carlson are running for the same office, their order on the ballot will depend on the order in which the letters M and R were drawn in the randomized alphabet drawing.


Chronicle / Paul Chinn



Chronicle / Paul Chinn

Chronicle / Paul Chinn

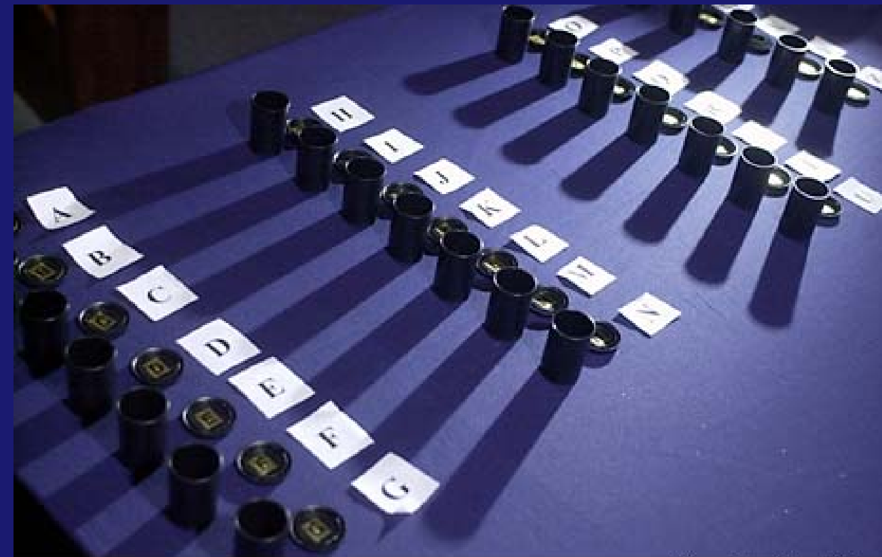


Chronicle / Paul Chinn

- Random alphabet for 2003 recall election:
  R W Q O J M V A H B S G Z X N T C I E K U P D Y F L

- The random alphabet is drawn roughly six weeks before the election, after certification of candidates has been completed.

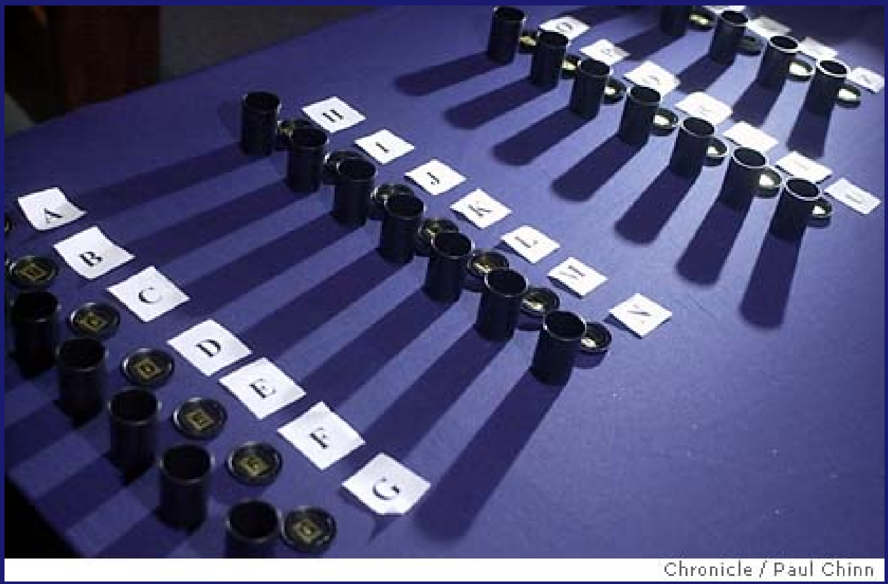

Chronicle / Paul Chinn


Chronicle / Paul Chinn

- Random alphabet for 2003 recall election:
  R W Q O J M V A H B S G Z X N T C I E K U P D Y F L

- The random alphabet is drawn roughly six weeks before the election, after certification of candidates has been completed.

- Three-step randomization-rotation procedure:

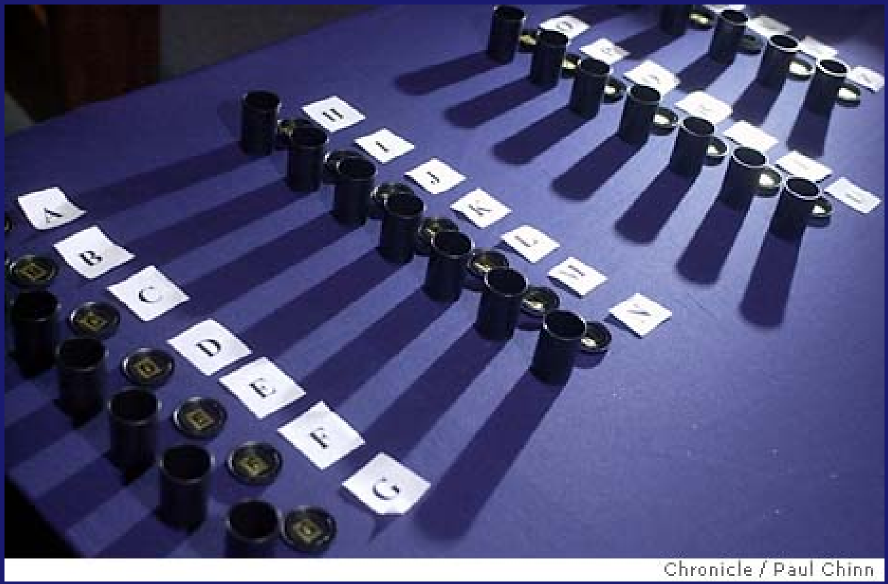


- Random alphabet for 2003 recall election:
  R W Q O J M V A H B S G Z X N T C I E K U P D Y F L

- The random alphabet is drawn roughly six weeks before the election, after certification of candidates has been completed.

- Three-step randomization-rotation procedure:

1. Secretary of State randomly draws alphabet letters.

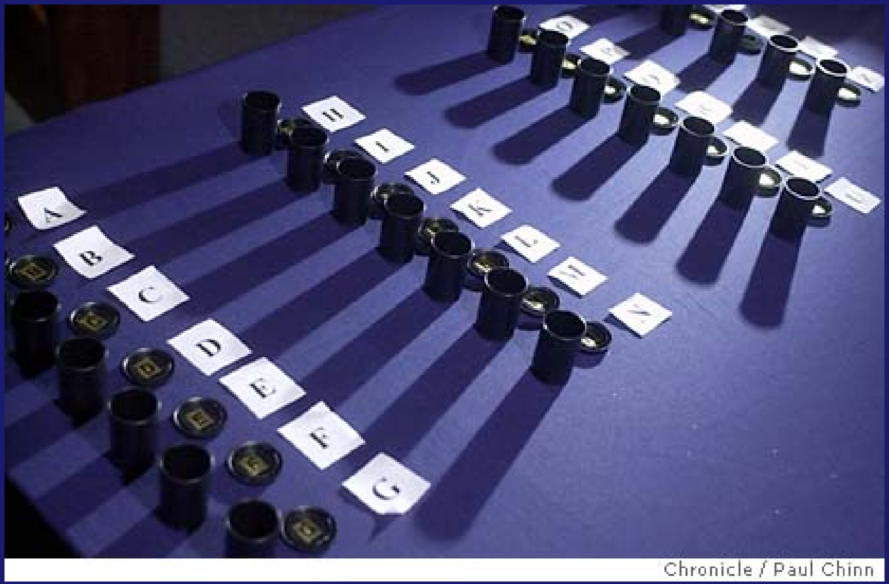


- Random alphabet for 2003 recall election:
  R W Q O J M V A H B S G Z X N T C I E K U P D Y F L

- The random alphabet is drawn roughly six weeks before the election, after certification of candidates has been completed.

- Three-step randomization-rotation procedure:

  1. Secretary of State randomly draws alphabet letters.
  2. Names of candidates are ordered by the randomized alphabet for the first of 80 assembly districts.



- Random alphabet for 2003 recall election:
  R W Q O J M V A H B S G Z X N T C I E K U P D Y F L

- The random alphabet is drawn roughly six weeks before the election, after certification of candidates has been completed.

- Three-step randomization-rotation procedure:

  1. Secretary of State randomly draws alphabet letters.
  2. Names of candidates are ordered by the randomized alphabet for the first of 80 assembly districts.
  3. Candidate names are systematically rotated for the rest of assembly districts.

The 2003 Gubernatorial Recall Election

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.
- Page placement of each candidate depends on:

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.
- Page placement of each candidate depends on:
  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

- New York Times ballot data:

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

- New York Times ballot data:

  ⋆ Page placement of 135 candidates for 158 unique "districts."

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

- New York Times ballot data:

  - ⋆ Page placement of 135 candidates for 158 unique “districts.”
  - ⋆ We focus on 121 districts with multiple-page ballot.

The 2003 Gubernatorial Recall Election

- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

- New York Times ballot data:

  - ⋆ Page placement of 135 candidates for 158 unique "districts."
  - ⋆ We focus on 121 districts with multiple-page ballot.
  - ⋆ Schwarzenegger was on the first page in 75 districts; Robinson in 43 districts.

The 2003 Gubernatorial Recall Election

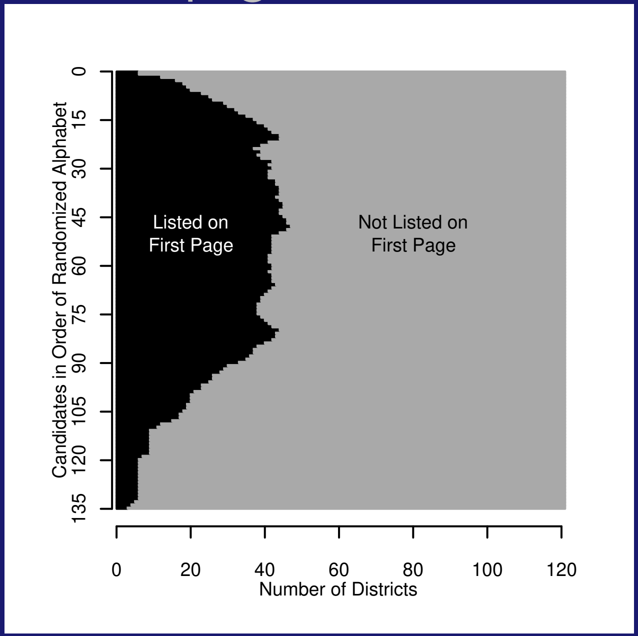
- 135 candidates running for the office.

- Page placement of each candidate depends on:

  1. Ballot format decided by each of 58 county registrars.
  2. Ballot order in each of 80 assembly districts determined by randomized alphabet and systematic rotation.

- New York Times ballot data:

  - ⋆ Page placement of 135 candidates for 158 unique "districts."
  - ⋆ We focus on 121 districts with multiple-page ballot.
  - ⋆ Schwarzenegger was on the first page in 75 districts; Robinson in 43 districts.



- Additional data:

- Additional data:
  - ⋆ Election returns for each candidate in every district.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).
- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.
- Statistical challenges:
  1. Variance identification problem:

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.
  2. Unequal probability assignment across districts for a given candidate:

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.
  2. Unequal probability assignment across districts for a given candidate:
     - ⋆ A large number of candidates relative to the number of districts.

- Additional data:

  ⋆ Election returns for each candidate in every district.
  ⋆ Registration data for each party in every district.
  ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     ⋆ Standard variance estimators require knowledge of population distribution.
     ⋆ The order of California districts is non-random (from Northwest to Southeast).
     ⋆ The distribution of district characteristics is unknown.

  2. Unequal probability assignment across districts for a given candidate:
     ⋆ A large number of candidates relative to the number of districts.
     ⋆ Possible confounding effects due to heterogeneous districts.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:

  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.
  2. Unequal probability assignment across districts for a given candidate:
     - ⋆ A large number of candidates relative to the number of districts.
     - ⋆ Possible confounding effects due to heterogeneous districts.
  3. Clustering of candidates:

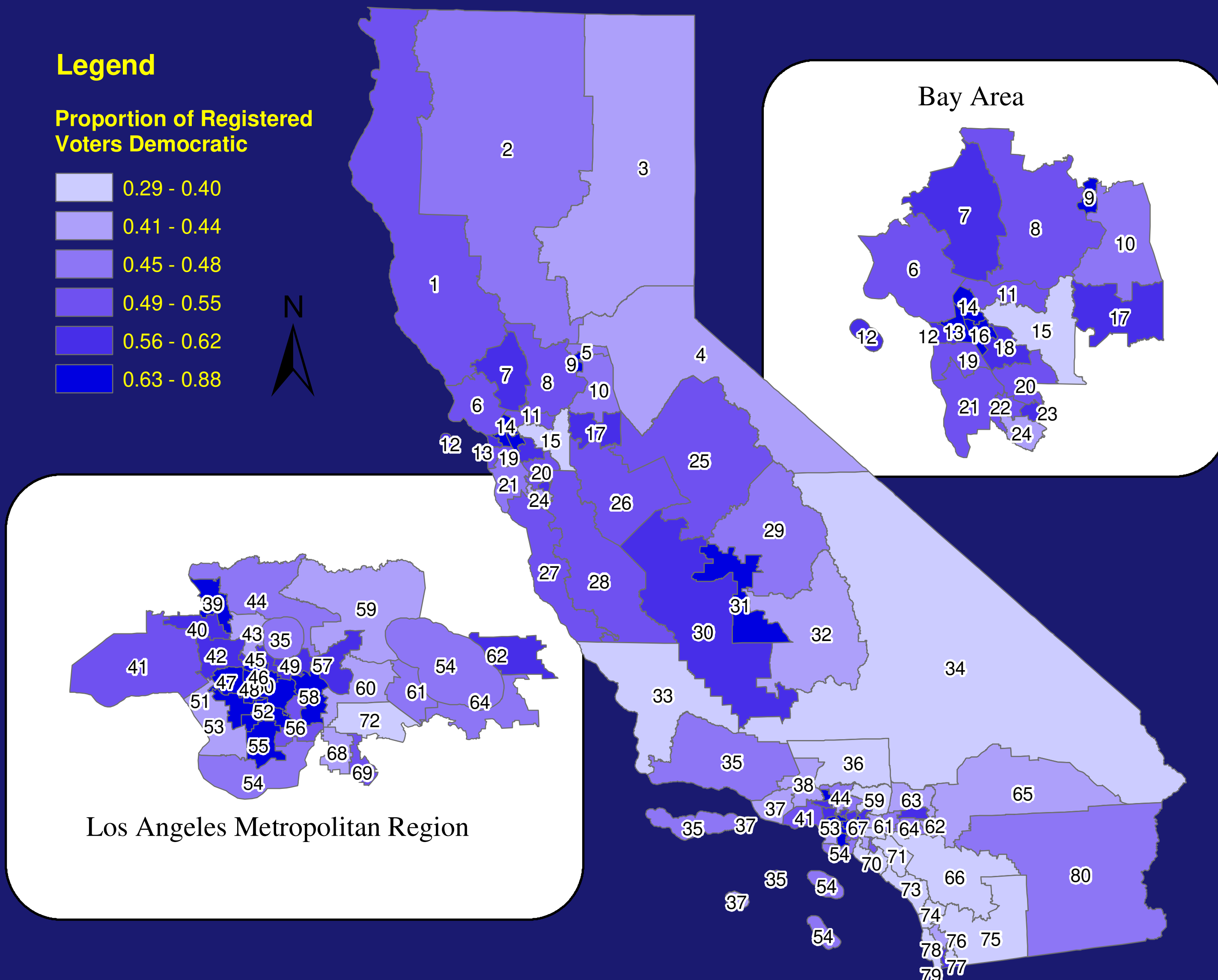
- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).

- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.

- Statistical challenges:
  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.
  2. Unequal probability assignment across districts for a given candidate:
     - ⋆ A large number of candidates relative to the number of districts.
     - ⋆ Possible confounding effects due to heterogeneous districts.
  3. Clustering of candidates:
     - ⋆ Alphabet randomization rather than candidate randomization is used.

- Additional data:
  - ⋆ Election returns for each candidate in every district.
  - ⋆ Registration data for each party in every district.
  - ⋆ Income, education, race for assembly districts (2000 Census).
- We estimate the causal effect of being placed on the first page of the ballot on each candidate's vote share.
- Statistical challenges:
  1. Variance identification problem:
     - ⋆ Randomization-rotation procedure resembles systematic sampling in surveys.
     - ⋆ Standard variance estimators require knowledge of population distribution.
     - ⋆ The order of California districts is non-random (from Northwest to Southeast).
     - ⋆ The distribution of district characteristics is unknown.
  2. Unequal probability assignment across districts for a given candidate:
     - ⋆ A large number of candidates relative to the number of districts.
     - ⋆ Possible confounding effects due to heterogeneous districts.
  3. Clustering of candidates:
     - ⋆ Alphabet randomization rather than candidate randomization is used.
     - ⋆ Many candidate names start with the same letter of the alphabet.

California Assembly Districts by Percentage of Registered Democrats



Randomization Inference via Fisher's Exact Test

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$

    $Y_i(1)$: potential vote share when the candidate is placed on the first page.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$

    $Y_i(1)$: potential vote share when the candidate is placed on the first page.

    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:

1. **Formulate a (sharp) null hypothesis**:

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:

  1. **Formulate a (sharp) null hypothesis**:
     - ⋆ $H_0 : \tau_i = 0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121$

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:
  1. **Formulate a (sharp) null hypothesis**:
     - ⋆ $H_0 : \tau_i = 0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121$
     - ⋆ Unit ballot effect is zero for all districts; i.e., $Y_i(1) = Y_i(0) = y_i \ \forall i$.

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:

1. **Formulate a (sharp) null hypothesis**:
   - ⋆ $H_0 : \tau_i = 0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121$
   - ⋆ Unit ballot effect is zero for all districts; i.e., $Y_i(1) = Y_i(0) = y_i \; \forall i$.

2. **Choose a test statistic**:

Randomization Inference via Fisher's Exact Test

- Notations (for each candidate in the ith district):
  - ⋆ $T_i = 1$ if the candidate is listed on the first page, and $T_i = 0$ otherwise.
  - ⋆ $Y_i = Y_i(1)T_i + (1 - T_i)Y_i(0)$
    $Y_i(1)$: potential vote share when the candidate is placed on the first page.
    $Y_i(0)$: potential vote share when the candidate is not placed on the first page.
  - ⋆ t_i and y_i: observed values of T_i and Y_i.
  - ⋆ Unit ballot page effect: $\tau_i \equiv Y_i(1) - Y_i(0)$.

- Hypothesis Testing Procedure:

1. **Formulate a (sharp) null hypothesis**:
   - ⋆ $H_0 : \tau_i = 0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121$
   - ⋆ Unit ballot effect is zero for all districts; i.e., $Y_i(1) = Y_i(0) = y_i \ \forall i$.
2. **Choose a test statistic**:
   - ⋆ Sample average ballot effect:

$$W^D(T) = \frac{\sum_{i=1}^{121} T_i y_i}{N_1} - \frac{\sum_{i=1}^{121} (1 - T_i) y_i}{N_0},$$

corresponding to the difference-in-means estimator.

- ⋆ Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top M T\right)^{-1} T^\top M y$$

corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1}X^\top$, and X is the matrix of the observed pretreatment covariates.

- Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top M T\right)^{-1} T^\top M y$$

corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1} X^\top$, and X is the matrix of the observed pretreatment covariates.
- Under H_0, distribution of $W^D(T)$ depends *only* on T.

- ⋆ Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top MT\right)^{-1} T^\top My$$

  corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1}X^\top$, and X is the matrix of the observed pretreatment covariates.
- ⋆ Under H_0, distribution of $W^D(T)$ depends *only* on T.

3. **Compute the exact p-value**:

- ⋆ Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top M T\right)^{-1} T^\top M y$$

corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1} X^\top$, and X is the matrix of the observed pretreatment covariates.

- ⋆ Under H_0, distribution of $W^D(T)$ depends *only* on T.

3. **Compute the exact p-value**:
   - ⋆ California alphabet lottery determines T_i.

- ⋆ Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top M T\right)^{-1} T^\top M y$$

  corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1} X^\top$, and X is the matrix of the observed pretreatment covariates.

- ⋆ Under H_0, distribution of $W^D(T)$ depends *only* on T.

3. **Compute the exact p-value**:
   - ⋆ California alphabet lottery determines T_i.
   - ⋆ Distribution of T_i is known *exactly* (No distributional assumption).

- Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top MT\right)^{-1} T^\top My$$

corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1}X^\top$, and X is the matrix of the observed pretreatment covariates.
- Under H_0, distribution of $W^D(T)$ depends *only* on T.

3. **Compute the exact p-value**:
   - California alphabet lottery determines T_i.
   - Distribution of T_i is known *exactly* (No distributional assumption).
   - One-tailed exact p-value: $p^D \equiv \Pr(W^D(T) \geq W^D(t))$.

- ⋆ Covariance-adjusted statistic:

$$W^L(T) = \left(T^\top M T\right)^{-1} T^\top M y$$

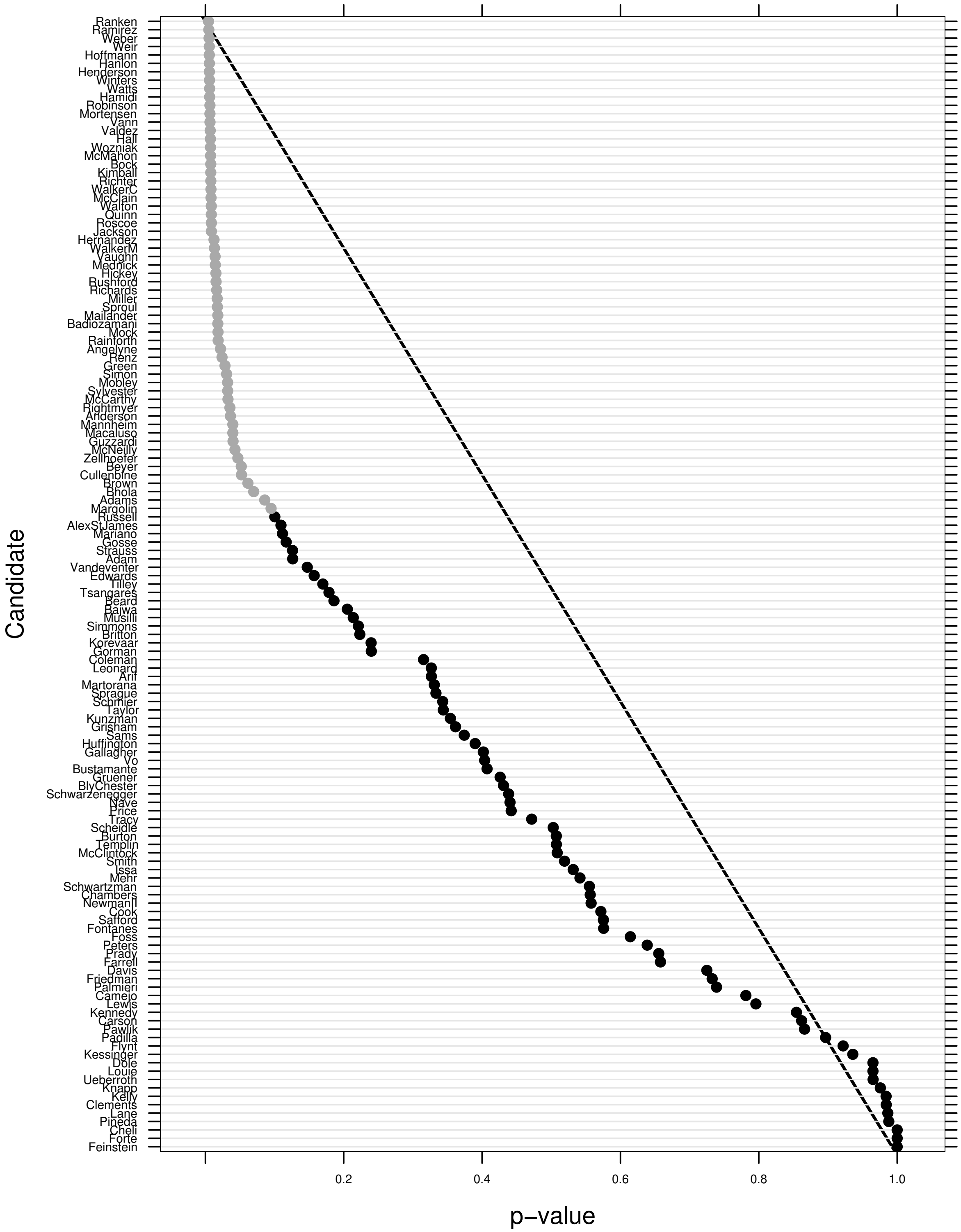
  corresponding to the linear least squares estimator, where $y = (y_1, y_2, \ldots, y_{121})$, $M = I - X(X^\top X)^{-1} X^\top$, and X is the matrix of the observed pretreatment covariates.
- ⋆ Under H_0, distribution of $W^D(T)$ depends *only* on T.

3. **Compute the exact p-value**:
   - ⋆ California alphabet lottery determines T_i.
   - ⋆ Distribution of T_i is known *exactly* (No distributional assumption).
   - ⋆ One-tailed exact p-value: $p^D \equiv \Pr(W^D(T) \geq W^D(t))$.

- Since the number of permutations is large, we use Monte Carlo approximation,

$$\Pr(W^D(T) \geq W^D(t)) \approx \frac{1}{m} \sum_{j=1}^{m} I(W^D(T^{(j)}) \geq W^D(t)),$$

with $m = 10,000$.



- No significant effect on major candidates.
- Positive ballot effect on 40% of minor candidates.

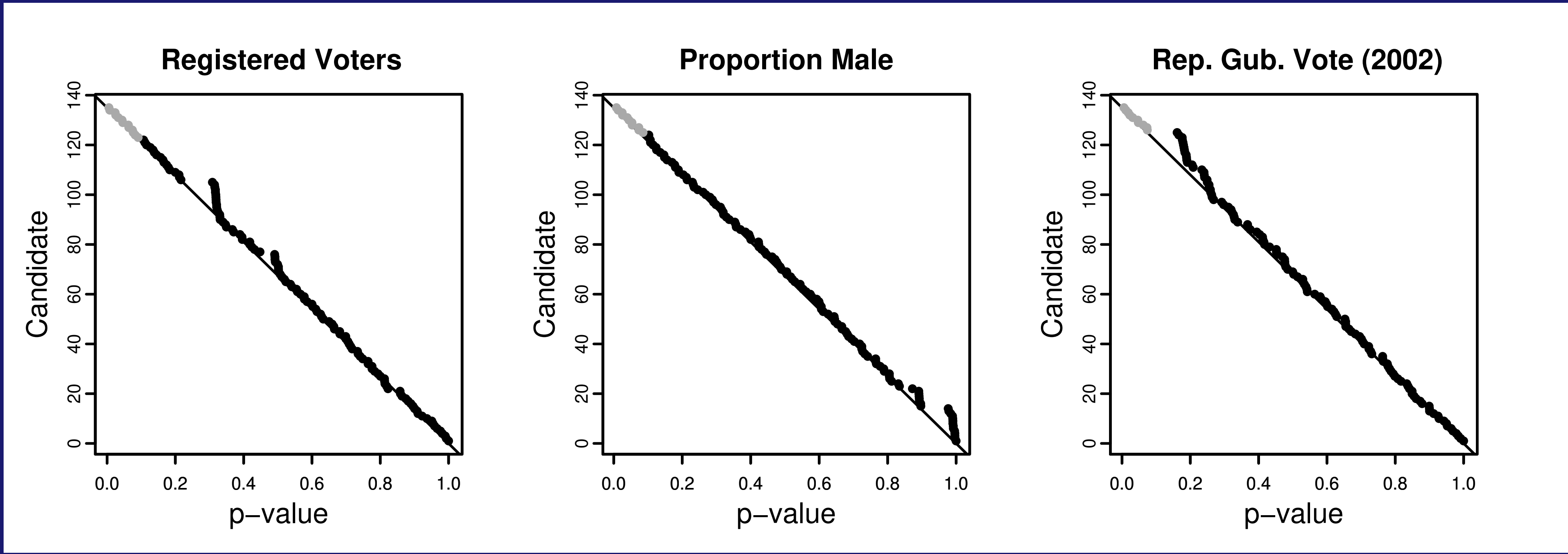
No Effect on Pretreatment Variables

No Effect on Pretreatment Variables

- If the treatment is random, it should not affect pretreatment variables.

No Effect on Pretreatment Variables

- If the treatment is random, it should not affect pretreatment variables.



Estimating Nonparametric Confidence Intervals

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all}\, i = 1, \ldots, 121.$

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121.$
   - Constant additive treatment effect assumption.

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121.$
   - Constant additive treatment effect assumption.

2. Given the null value τ_0, the test statistic is given by

$$W^D_{\tau_0}(T) \ = \ \frac{\sum_{i=1}^{121} T_i\{y_i + (1 - t_i)\tau_0\}}{\sum_{i=1}^{121} T_i} - \frac{\sum_{i=1}^{121}(1 - T_i)(y_i - t_i\tau_0)}{\sum_{i=1}^{121}(1 - T_i)},$$

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121.$
   - Constant additive treatment effect assumption.

2. Given the null value τ_0, the test statistic is given by

$$W^D_{\tau_0}(T) \; = \; \frac{\sum_{i=1}^{121} T_i\{y_i + (1 - t_i)\tau_0\}}{\sum_{i=1}^{121} T_i} - \frac{\sum_{i=1}^{121} (1 - T_i)(y_i - t_i\tau_0)}{\sum_{i=1}^{121} (1 - T_i)},$$

   or its covariance-adjusted analogue

$$W^L_{\tau_0}(T) \; = \; \left(T^\top M T\right)^{-1} T^\top M y^*,$$

   where each element of y^* is $y_i^* = T_i\{y_i + (1 - t_i)\tau_0\} + (1 - T_i)(y_i - t_i\tau_0)$.

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121.$
   - Constant additive treatment effect assumption.

2. Given the null value τ_0, the test statistic is given by

$$W^D_{\tau_0}(T) \;=\; \frac{\sum_{i=1}^{121} T_i\{y_i + (1 - t_i)\tau_0\}}{\sum_{i=1}^{121} T_i} - \frac{\sum_{i=1}^{121}(1 - T_i)(y_i - t_i\tau_0)}{\sum_{i=1}^{121}(1 - T_i)},$$

   or its covariance-adjusted analogue

$$W^L_{\tau_0}(T) \;=\; \left(T^\top M T\right)^{-1} T^\top M y^*,$$

   where each element of y^* is $y^*_i = T_i\{y_i + (1 - t_i)\tau_0\} + (1 - T_i)(y_i - t_i\tau_0)$.

3. Two-tailed level α test;

Estimating Nonparametric Confidence Intervals

Generalizing Fisher's exact test:

1. Sharp null hypothesis: $H_0 : \tau_i = \tau_0 \quad \text{for} \quad \text{all } i = 1, \ldots, 121.$
   - Constant additive treatment effect assumption.
2. Given the null value τ_0, the test statistic is given by

$$W^D_{\tau_0}(T) \;=\; \frac{\sum_{i=1}^{121} T_i\{y_i + (1 - t_i)\tau_0\}}{\sum_{i=1}^{121} T_i} - \frac{\sum_{i=1}^{121}(1 - T_i)(y_i - t_i\tau_0)}{\sum_{i=1}^{121}(1 - T_i)},$$

   or its covariance-adjusted analogue

$$W^L_{\tau_0}(T) \;=\; \left(T^\top M T\right)^{-1} T^\top M y^*,$$

   where each element of y^* is $y^*_i = T_i\{y_i + (1 - t_i)\tau_0\} + (1 - T_i)(y_i - t_i\tau_0)$.

3. Two-tailed level α test; accept H_0 if

$$t \in A_\alpha(\tau_0) = \left\{u : \frac{\alpha}{2} \leq \Pr(W^D_{\tau_0}(T) \geq W^D(u)) \leq 1 - \frac{\alpha}{2}\right\},$$

   and reject H_0 otherwise.

Inverting the test:

Inverting the test:

- The $(1-\alpha)$ confidence set; $C_{\alpha}(t) = \{\tau : t \in A_{\alpha}(\tau)\}$.

Inverting the test:

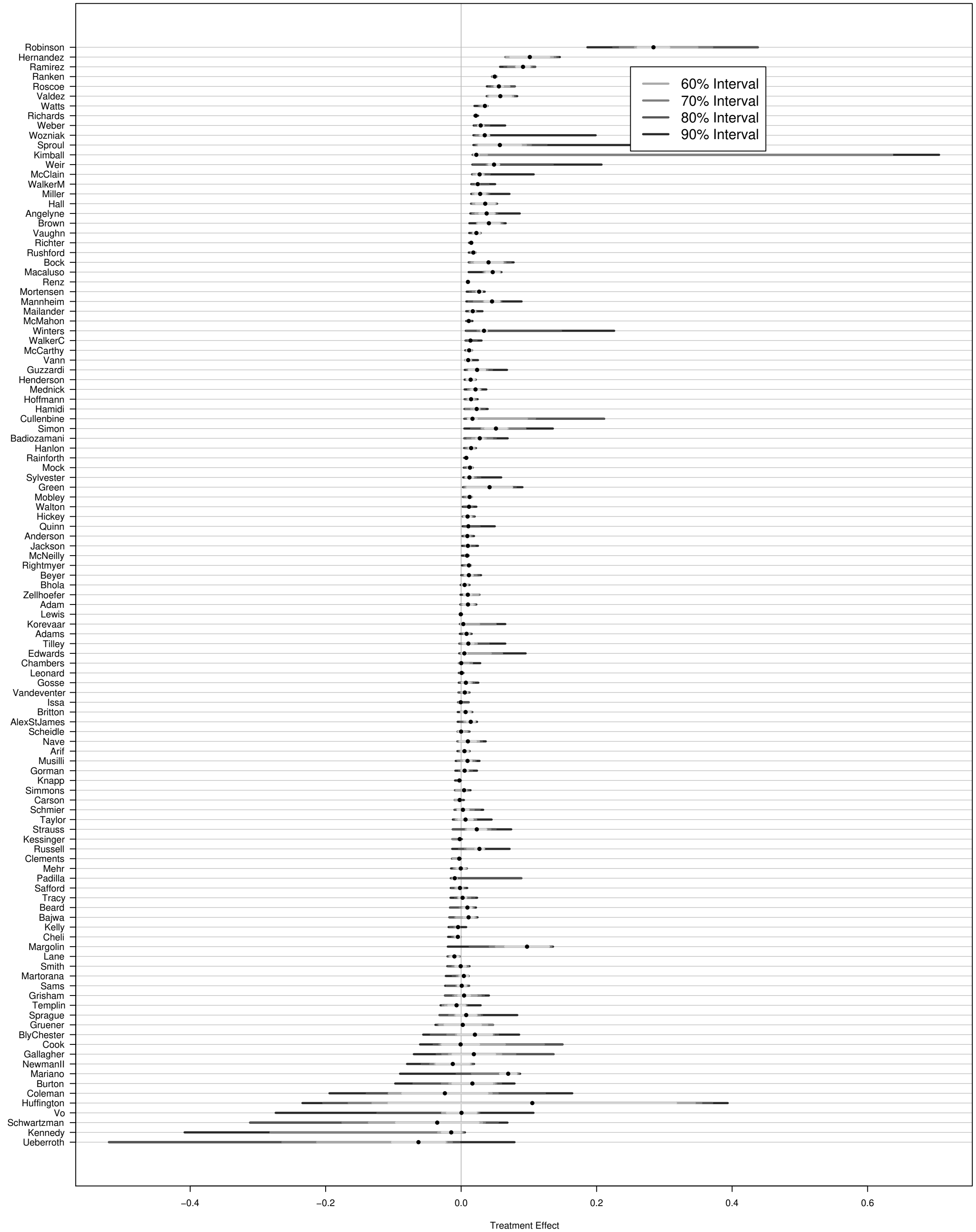
- The $(1-\alpha)$ confidence set; $C_\alpha(t) = \{\tau : t \in A_\alpha(\tau)\}$.
- Confidence interval defined as the shortest closed interval in the confidence set.

Inverting the test:

- The $(1-\alpha)$ confidence set; $C_\alpha(t) = \{\tau : t \in A_\alpha(\tau)\}$.
- Confidence interval defined as the shortest closed interval in the confidence set.
- Identify the upper and lower bounds, $\tau_L = \sup_\tau \; A_\alpha(\tau)$ and $\tau_U = \inf_\tau \; A_\alpha(\tau)$, via a (Monte Carlo) bisection algorithm.

Inverting the test:

- The $(1-\alpha)$ confidence set; $C_\alpha(t) = \{\tau : t \in A_\alpha(\tau)\}$.
- Confidence interval defined as the shortest closed interval in the confidence set.
- Identify the upper and lower bounds, $\tau_L = \sup_\tau A_\alpha(\tau)$ and $\tau_U = \inf_\tau A_\alpha(\tau)$, via a (Monte Carlo) bisection algorithm.
- Nonparametric estimates of CDF (for the sampling distributions of causal effect estimators) can also be obtained by estimating τ_U and τ_L for different values of $\alpha \in [0, 0.5]$.


Robinson
Hernandez
Ramirez
Ranken
Roscoe
Valdez
Watts
Richards
Weber
Wozniak
Sproul
Kimball
Weir
McClain
WalkerM
Miller
Hall
Angelyne
Brown
Vaughn
Richter
Rushford
Bock
Macaluso
Renz
Mortensen
Mannheim
Mailander
McMahon
Winters
WalkerC
McCarthy
Vann
Guzzardi
Henderson
Mednick
Hoffmann
Hamidi
Cullenbine
Simon
Badiozamani
Hanlon
Rainforth
Mock
Sylvester
Green
Mobley
Walton
Hickey
Quinn
Anderson
Jackson
McNeilly
Rightmyer
Beyer
Bhola
Zellhoefer
Adam
Lewis
Korevaar
Adams
Tilley
Edwards
Chambers
Leonard
Gosse
Vandeventer
Issa
Britton
AlexStJames
Scheidle
Nave
Arif
Musilli
Gorman
Knapp
Simmons
Carson
Schmier
Taylor
Strauss
Kessinger
Russell
Clements
Mehr
Padilla
Safford
Tracy
Beard
Bajwa
Kelly
Cheli
Margolin
Lane
Smith
Martorana
Sams
Grisham
Templin
Sprague
Gruener
BlyChester
Cook
Gallagher
NewmanII
Mariano
Burton
Coleman
Huffington
Vo
Schwartzman
Kennedy
Ueberroth
60% Interval
70% Interval
80% Interval
90% Interval
−0.4
−0.2
0.0
0.2
0.4
0.6
Treatment Effect

Sensitivity Analyses

Sensitivity Analyses

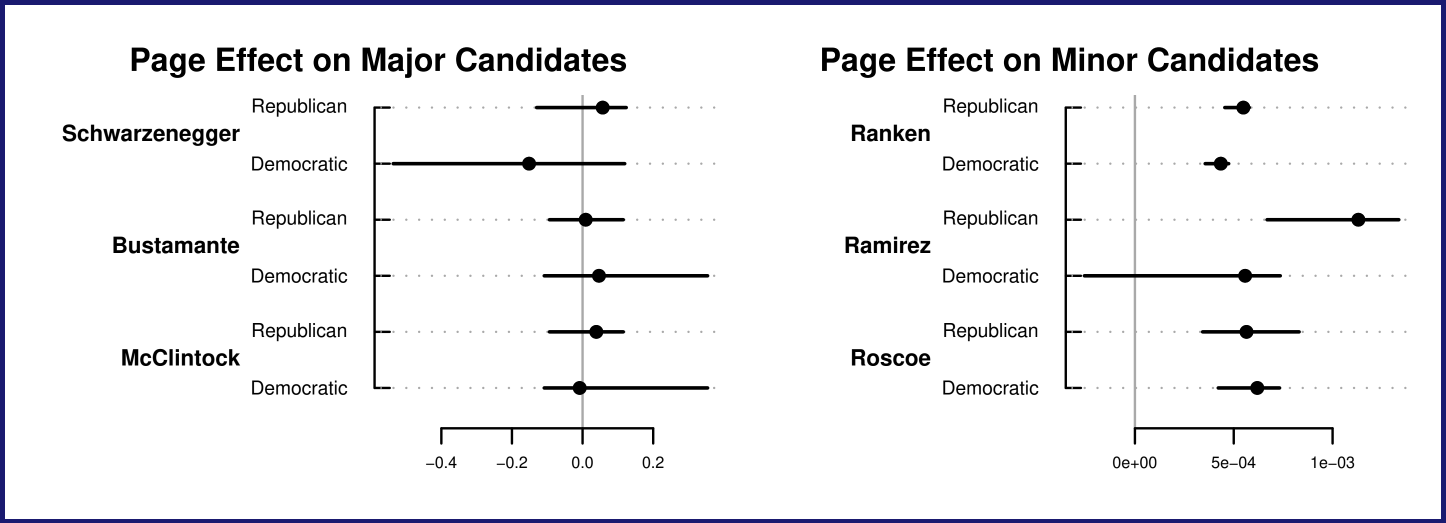
- Constant additive treatment effect assumption:

Sensitivity Analyses

- Constant additive treatment effect assumption:
  - In principle, one can define a vector of null values τ_0.

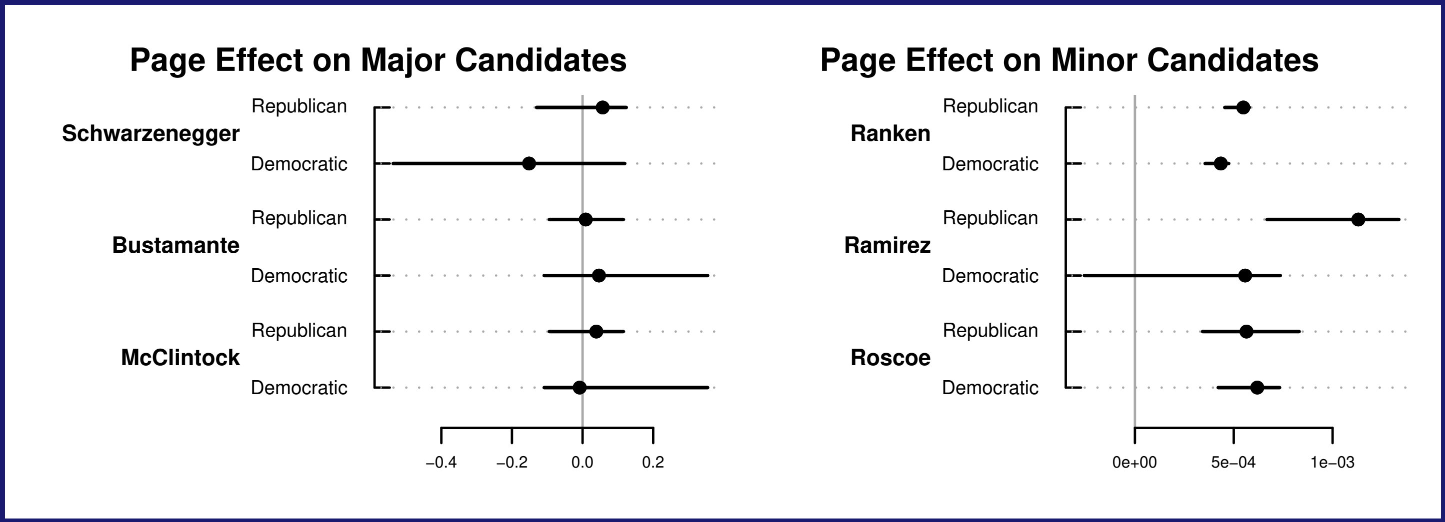
Sensitivity Analyses

- Constant additive treatment effect assumption:
  - ⋆ In principle, one can define a vector of null values τ_0.



Sensitivity Analyses

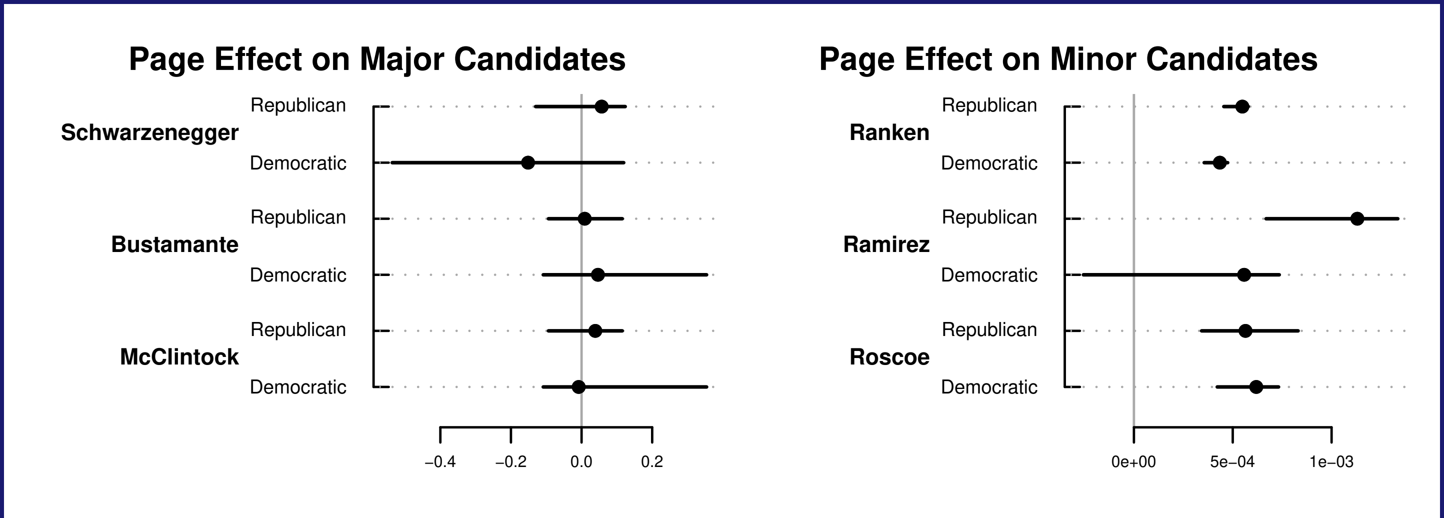
- Constant additive treatment effect assumption:
  - ⋆ In principle, one can define a vector of null values τ_0.



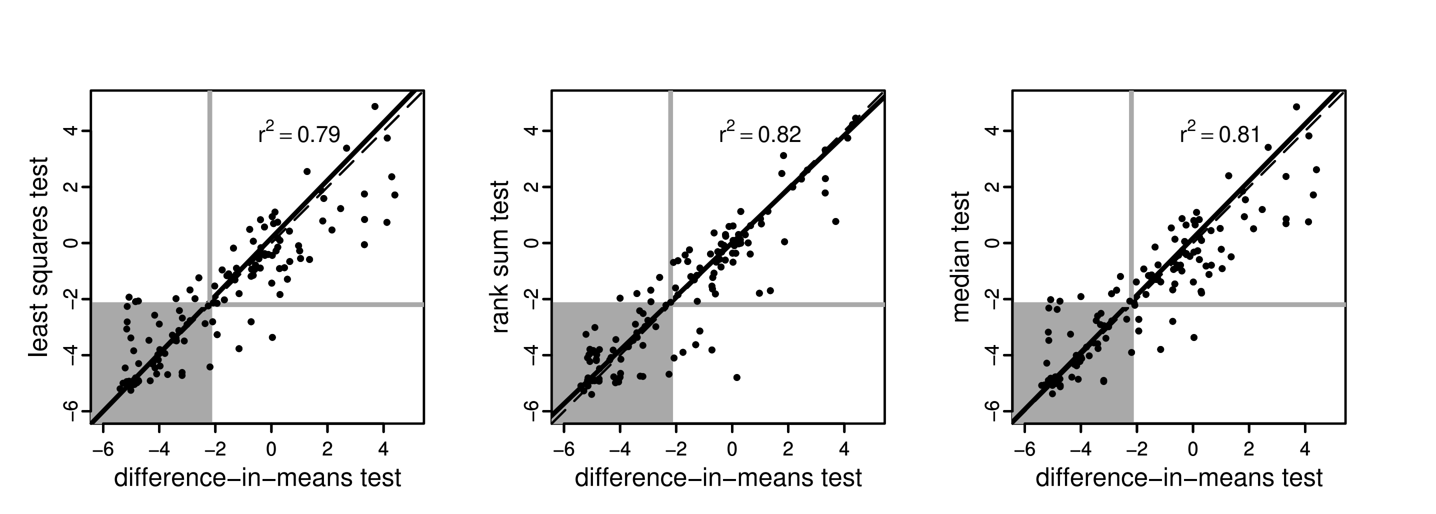
- Choice of test statistics:

Sensitivity Analyses

- Constant additive treatment effect assumption:
  - ⋆ In principle, one can define a vector of null values τ_0.



- Choice of test statistics:



Comparison with Conventional Estimators

Comparison with Conventional Estimators

- Two frequently used parametric estimators:

Comparison with Conventional Estimators

- Two frequently used parametric estimators:

1. Linear least squares (with and without covariates).

Comparison with Conventional Estimators

- Two frequently used parametric estimators:

  1. Linear least squares (with and without covariates).
     based on the same statistics as those used in randomization inference.

Comparison with Conventional Estimators

- Two frequently used parametric estimators:

  1. Linear least squares (with and without covariates).
     based on the same statistics as those used in randomization inference.
  2. Binomial GLM with logit link and overdispersion.

Comparison with Conventional Estimators

- Two frequently used parametric estimators:
  1. Linear least squares (with and without covariates).
     based on the same statistics as those used in randomization inference.
  2. Binomial GLM with logit link and overdispersion.

- Results are appreciably different.

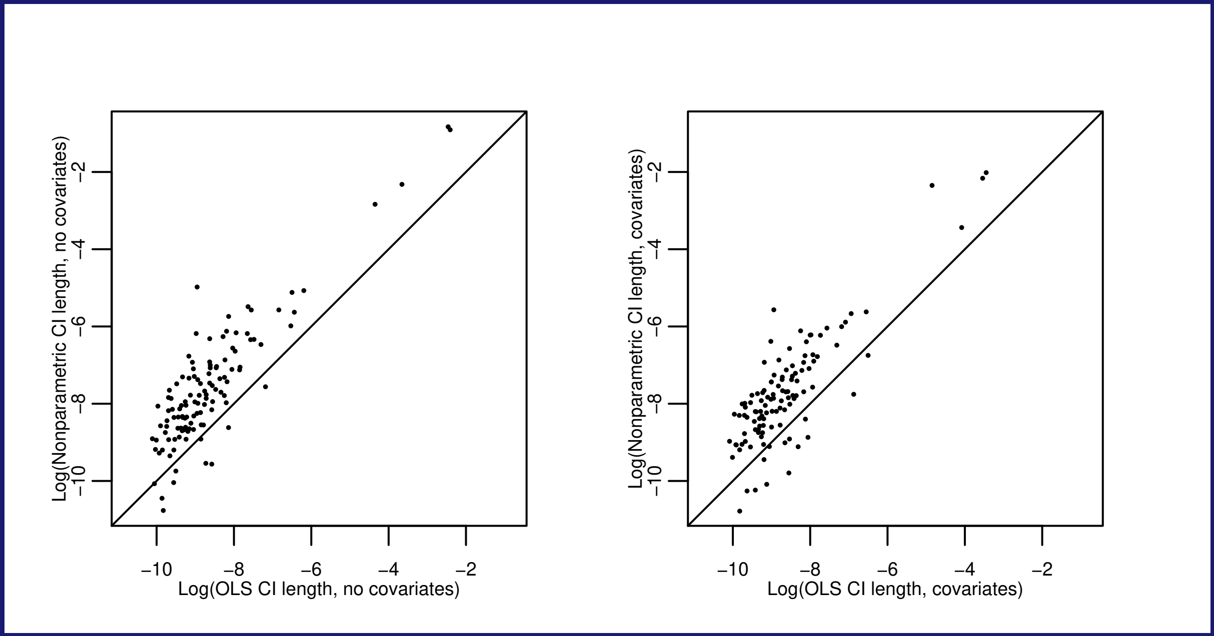
Comparison with Conventional Estimators

- Two frequently used parametric estimators:
  1. Linear least squares (with and without covariates).
     based on the same statistics as those used in randomization inference.
  2. Binomial GLM with logit link and overdispersion.
- Results are appreciably different.

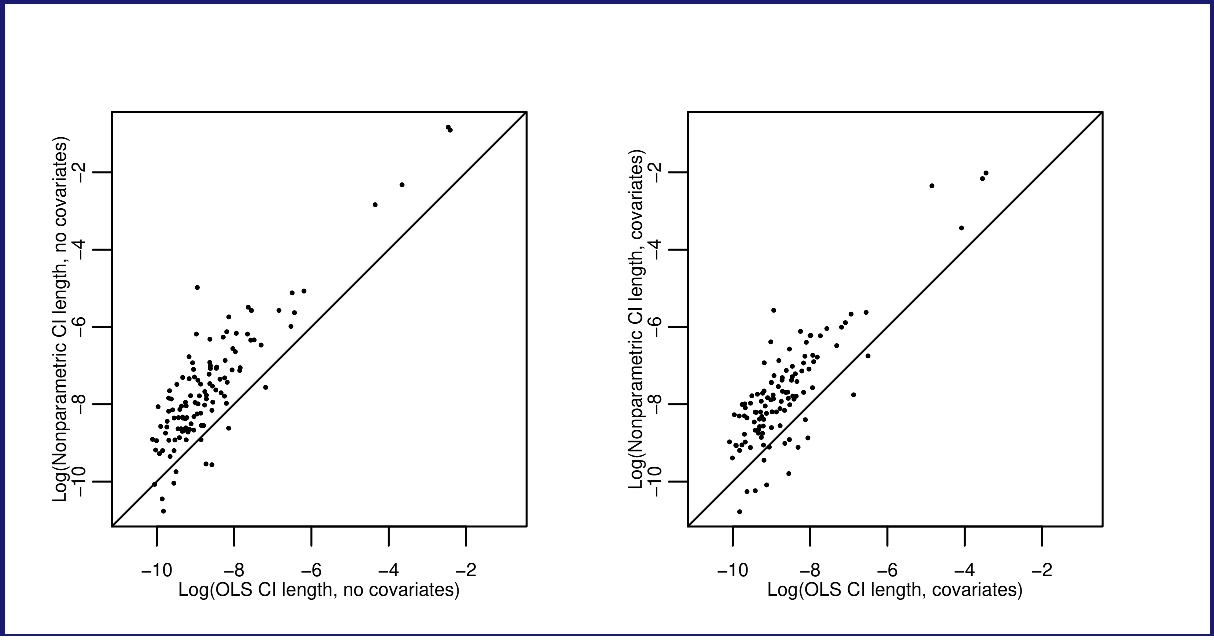
	Without Covariates						With Covariates					
	OLS		GLM logit		RI Fisher		OLS		GLM logit		RI Fisher	
Major Candidates												
Schwarzenegger	1.09	9.63	−1.23	7.53	−23.72	19.90	−2.97	0.21	−4.98	−2.05	−6.44	6.87
Bustamante	−8.46	0.54	−5.37	4.04	−20.07	20.31	−1.12	1.78	0.96	3.01	−5.86	5.64
McClintock	0.50	3.09	−1.10	1.24	−3.47	6.36	1.56	3.25	0.29	2.05	0.36	3.57
All Candidates												
Positive effects	56(41%)		63(47%)		55(41%)		50(37%)		59(44%)		47(35%)	
Negative effects	11 (8%)		8 (6%)		4 (3%)		8 (6%)		17(13%)		2 (1%)	
Null effects	68(50%)		64(47%)		59(44%)		77(57%)		59(44%)		64(47%)	
Unidentified	0 (0%)		0 (0%)		17(13%)		0 (0%)		0 (0%)		22(16%)	
Comparison with Randomization Inference												
Agreement	89(66%)		87(64%)		108(80%)		88(65%)		74(55%)		108(80%)	

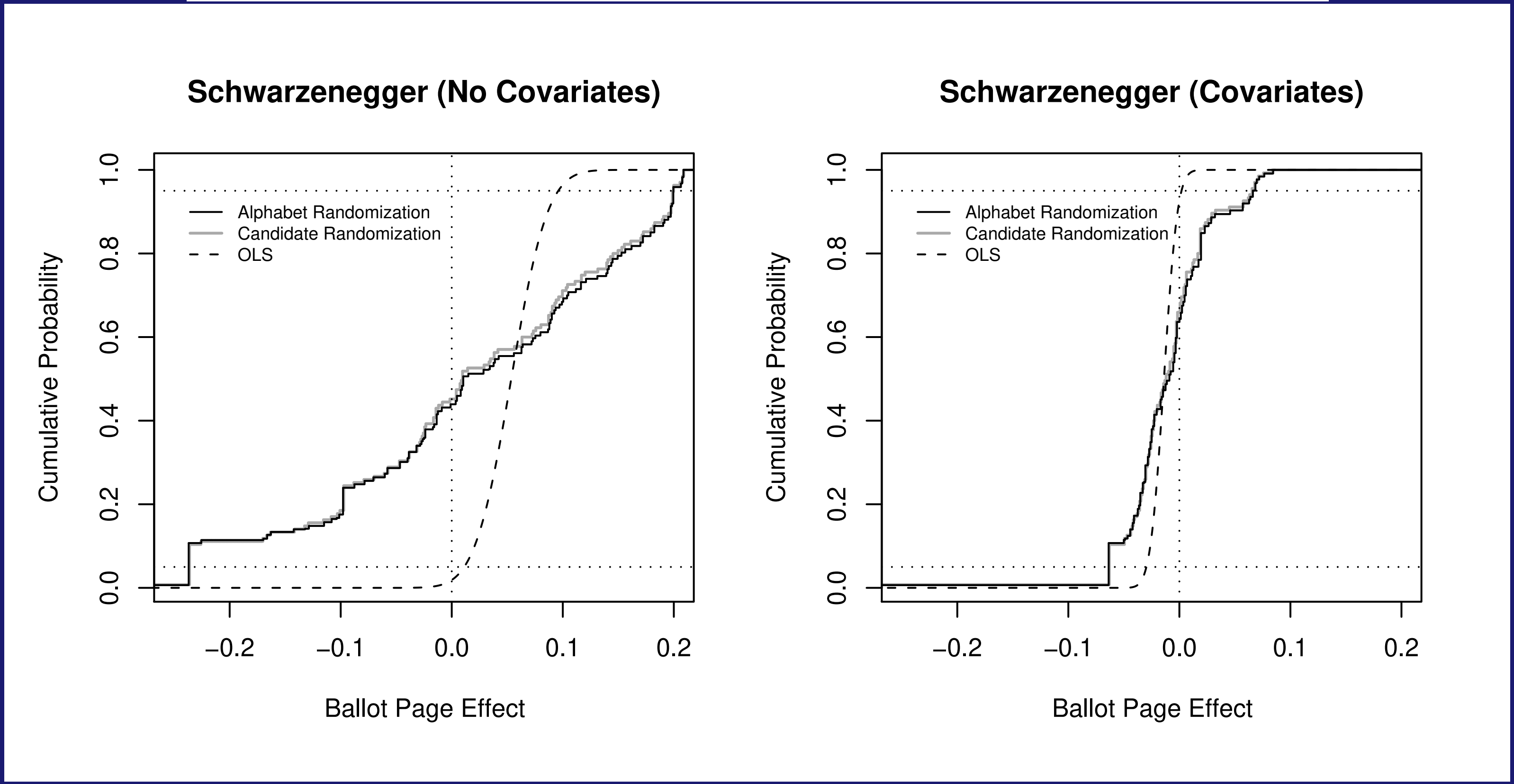
Comparisons of Confidence Intervals and CDF

Comparisons of Confidence Intervals and CDF



Comparisons of Confidence Intervals and CDF





Concluding Remarks

Concluding Remarks

- Randomization inference provides a general framework for robust causal inference in randomized experiments.

Concluding Remarks

- Randomization inference provides a general framework for robust causal inference in randomized experiments.
- Parametric inferences can be sensitive to modeling and other assumptions.

Concluding Remarks

- Randomization inference provides a general framework for robust causal inference in randomized experiments.
- Parametric inferences can be sensitive to modeling and other assumptions.
- Randomized natural experiments provide social scientists with rare opportunities to draw valid causal inferences.

Concluding Remarks

- Randomization inference provides a general framework for robust causal inference in randomized experiments.
- Parametric inferences can be sensitive to modeling and other assumptions.
- Randomized natural experiments provide social scientists with rare opportunities to draw valid causal inferences.
- The randomization inference framework can directly incorporate complex randomization schemes in natural experiments.