What Are Conjoint Experiments? Exploring Public Preferences

Todd K Hartman
Professor of Quantitative Social Science
Department of Social Statistics
The University of Manchester
https://tkhartman.netlify.com/

28 October 2021

Roadmap for Today's Talk

- Brief example to set the stage
  - UK public attitudes toward tax cheats
- Overview of conjoint experiment and analysis
  - What is a conjoint experiment?
  - What are its key assumptions?
  - How do we implement it using survey software?
  - How do we conduct conjoint analysis?
- Step-by-step walkthrough
  - US public preferences for redistricting (and partisan gerrymandering)
- Q&A Session

Brief Example: UK Attitudes toward Tax Cheats

- What does the British public think about tax cheats?
  - Do they differentiate among those at the 'bottom' vs those at the 'top' of the wealth spectrum
    - i.e., the rich vs the poor?
  - What attributes might influence their perceptions about the severity of those who cheat on their taxes?

How Do Citizens Perceive Violations to the Social Order?

Morally undeserving rich?	e.g. “Tax dodgers”, “bailed-out bankers”
Hardworking, everyday people	e.g. The “squeezed middle”, “the taxpayer”
Morally undeserving poor	e.g. “Scroungers”, “welfare queens”

Conjoint Analysis [Experiment]

- Identify causal effects from a set of features

  Green & Rao, 1971; Hainmueller, Hangartner, & Yamamoto, 2015; Hainmueller, Hopkins, & Yamamoto, 2013; Hainmueller & Hopkins, 2015

Conjoint Analysis [Experiment]

- Identify causal effects from a set of features

  Green & Rao, 1971; Hainmueller, Hangartner, & Yamamoto, 2015; Hainmueller, Hopkins, & Yamamoto, 2013; Hainmueller & Hopkins, 2015

  - Respondents choose from generated profiles that combine multiple features
    - a.k.a. attributes and levels

Conjoint Analysis [Experiment]

- Identify causal effects from a set of features

  Green & Rao, 1971; Hainmueller, Hangartner, & Yamamoto, 2015; Hainmueller, Hopkins, & Yamamoto, 2013; Hainmueller & Hopkins, 2015

  - Respondents choose from generated profiles that combine multiple features
    - a.k.a. attributes and levels
  - Realistic, simple task (choose preferred profile)

Conjoint Analysis [Experiment]

- Identify causal effects from a set of features

  Green & Rao, 1971; Hainmueller, Hangartner, & Yamamoto, 2015; Hainmueller, Hopkins, & Yamamoto, 2013; Hainmueller & Hopkins, 2015

  - Respondents choose from generated profiles that combine multiple features
    - a.k.a. attributes and levels
  - Realistic, simple task (choose preferred profile)
  - Generate comparable estimates
    - Probability increase or decrease in selecting profile
    - Allows assessment of relative strength of each attribute and level

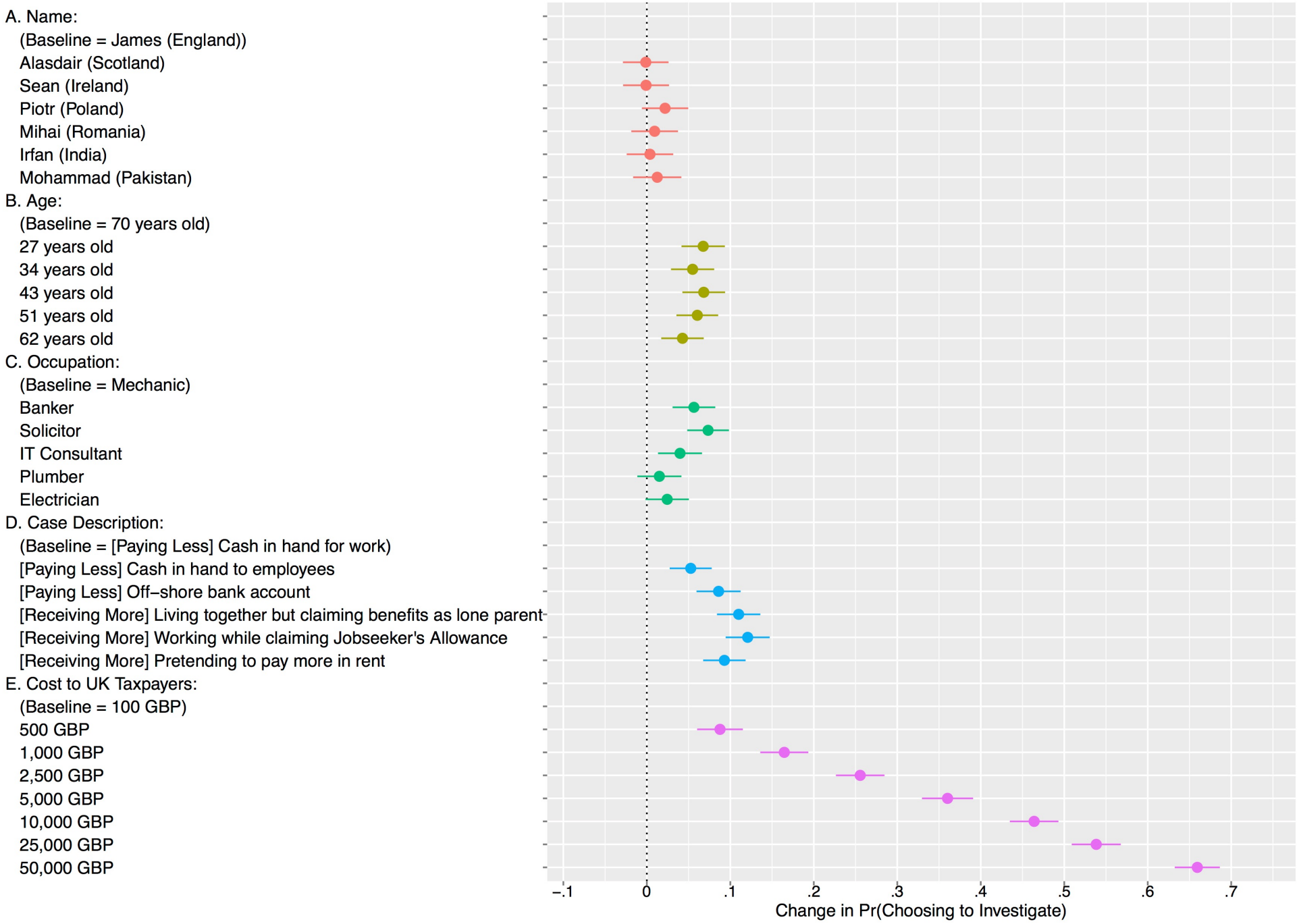
Now, imagine that you are an Investigation Officer working for the UK government. Your job is to investigate suspected cases of tax evasion and benefit fraud.

On the next few screens, you will be presented with profiles of people who have been reported to your office. For each pair, your job is to decide whom to investigate (you can only choose 1 person from each pair because of limited resources).

Task 3/4: Who would you investigate?

	Profile 1	**Profile 2**
Name (and Birthplace)	James (England)	Irfan (India)
Age	27 years old	62 years old
Occupation	Solicitor	Mechanic
Action	Paying less tax by receiving cash in hand for work	Receiving more benefit by working while claiming Jobseeker's Allowance
Cost to UK Taxpayers (GBP / year)	1,000 GBP	5,000 GBP

AMCE (Estimates)



Benefits of Conjoint Analysis

1. Realism of the task
2. Enables testing of multiple causal hypotheses
3. Estimates effects of multiple treatment components; thus, effects of each attribute can be directly compared
4. Limits social desirability concerns (allows respondents to justify choice based upon many factors)

Example: Redrawing Electoral Maps

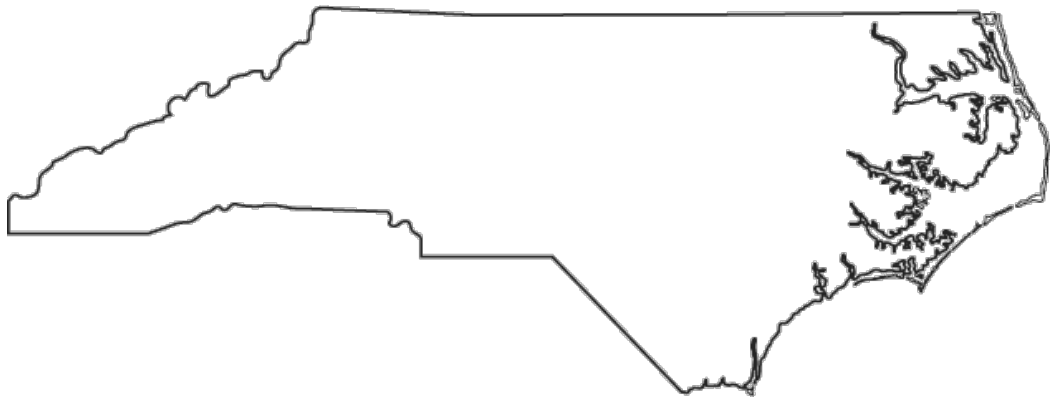
Step 1: Is your RQ suitable for conjoint analysis?

Example: Redrawing Electoral Maps

Step 1: Is your RQ suitable for conjoint analysis?

What do people want when it comes to redrawing electoral boundaries? What features of the redistricting process might influence their preferences?

- Proportionate districts (matched to partisanship)?
- Highly competitive districts?
- Compact districts (no funny business)?
- Or, are gerrymandered districts OK?



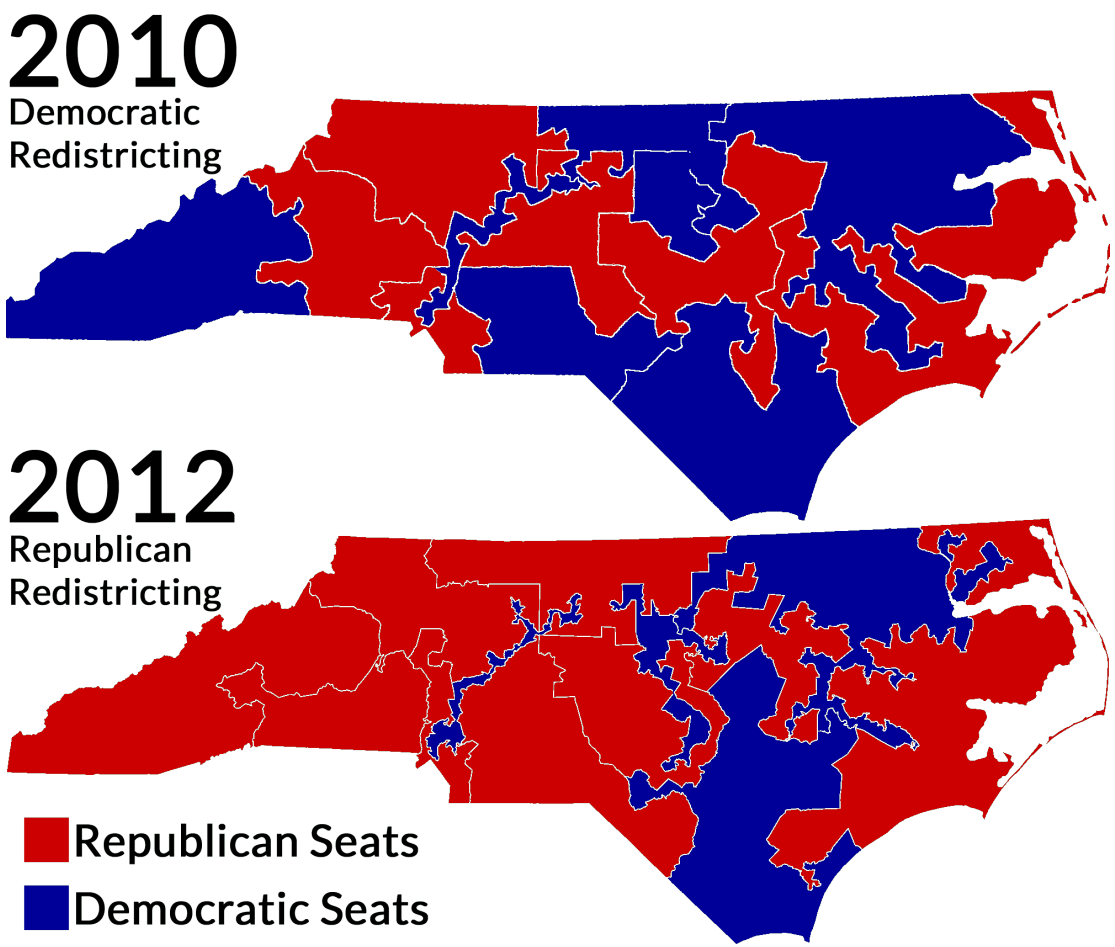
Example: Redrawing Electoral Maps

- Boundaries for electoral districts are redrawn
  - After census taken (every 10 years)
- In 75% of US states, redistricting is controlled by state legislature
- Party controlling legislature thus dictates process
  - Few rules...
  - Districts must have proportionate populations
  - No racial gerryamndering (Voting Rights Act)
- US Supreme Court: Partisan gerrymandering allowed

Consequences of Partisan Gerrymandering

Redrawing North Carolina's 13 Electoral Districts

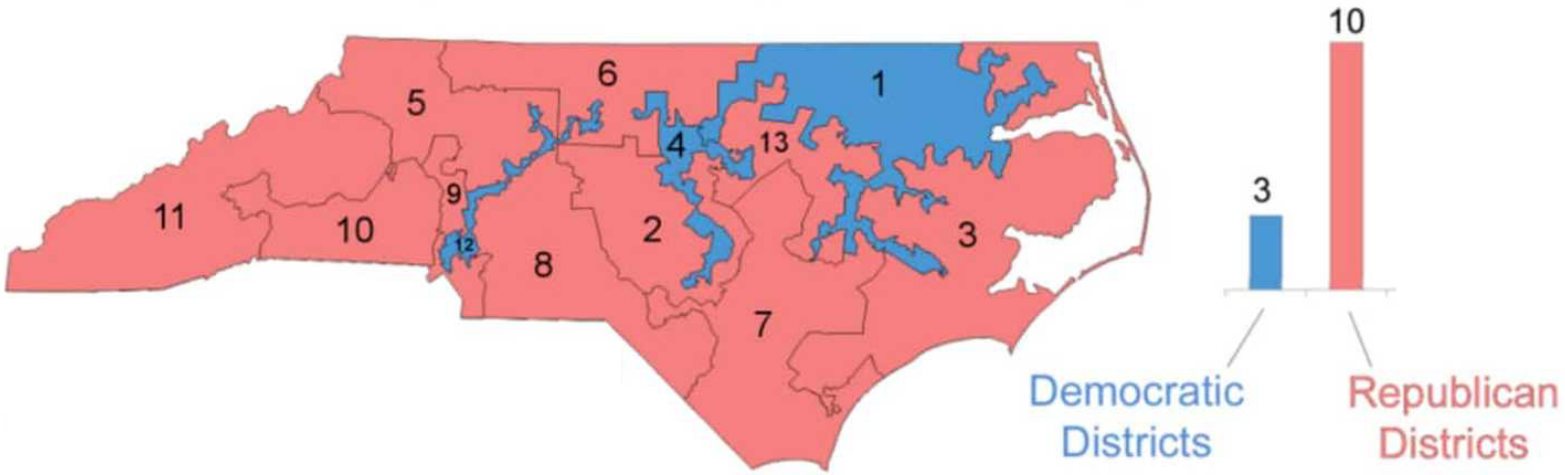
- In 2010, Reps: 54.1% (6 seats); Dems: 45.3% (7 seats)
- In 2012, Reps: 48.8% (9 seats); Dems: 50.6% (4 seats)



Partisan Gerrymandering Cont'd

Rep Chariman Lewis: "I think electing Republicans is better than electing Democrats. So I drew this map to help foster what I think is better for the country'.'

- In 2014, Reps: 55.4% (**10 seats**); Dems: 44.0% (3 seats)
- In 2016, Reps: 53.2% (**10 seats**); Dems: 46.6% (3 seats)
- In 2018, Reps: 50.4% (**10 seats**); Dems: 48.4% (3 seats)



People Seem to Dislike Partisan Gerrymandering

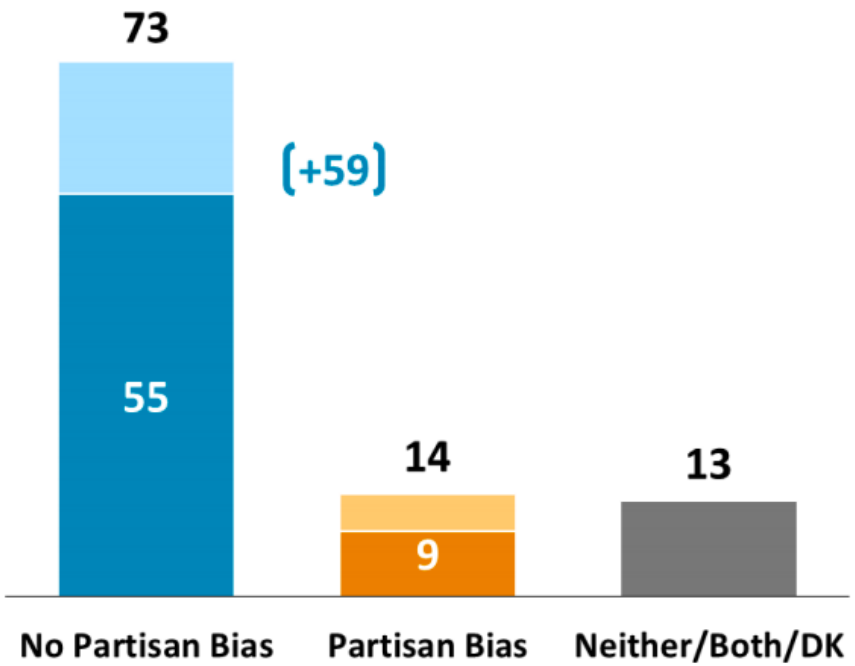
- 2006 Pew survey: Gerrymandering named as biggest reason for redistricting dissatisfaction
- 2018 Fairvote bipartisan survey concluded "everyone hates gerrymandering"

When presented both sides, voters overwhelmingly say "no" to partisan bias in drawing districts, even if that bias would help their party win more seats.

Of the two, which would you prefer?

[NO PARTISAN BIAS]
Congressional districts that are drawn with no partisan bias whatsoever, even if that means that your own preferred party would not win as many seats.

[PARTISAN BIAS]
Congressional districts that are drawn with a partisan bias, but knowing that your preferred party would likely win more seats.



Theory and Expectations

What might influence preferences for electoral boundary maps?

1. Fairness norms and inequity aversion
   (Castelli et al, 2014a/b; Falk et al., 2008; Fehr & Schmidt, 1999)
   - E1: Prefer boundary maps that are balanced or highly competitive
2. Motivated reasoning
   (Druckman, 2012; Kunda, 1990; Taber & Lodge, 2006)
   - E2: Prefer boundary maps that are gerrymandered to benefit preferred party
3. Heuristics - source cues or other cognitive shortcuts
   (Cohen, 2003; Mondak, 1993; Tversky & Kahneman, 1974)
   - E3: Prefer boundary maps endorsed by favored or trusted actors

Step 2: Design the conjoint experiment

https://github.com/astrezhnev/conjointsdt

README.md

Conjoint Survey Design Tool

The Conjoint Survey Design Tool assists researchers in creating multi-dimensional choice experiments that can be readily incorporated into any pre-existing web survey software (such as Qualtrics). Conjoint analysis is a type of survey experiment often used by market researchers to measure consumer preferences over a variety of product attributes. Hainmueller, Hopkins and Yamamoto (2014) demonstrate the value of this design for political science applications. Conjoint experiments present respondents with a choice among set of profiles composed of multiple randomly assigned attributes. This approach allows researchers to estimate the effect of each individual component on the probability that the respondent will choose a profile. This software tool is designed as a companion to Hainmueller, Hopkins and Yamamoto (2014), providing a graphical user interface for generating conjoint experiments.

Citation

Hainmueller, Jens., Hopkins, Daniel J., Yamamoto, Teppei. (2014). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments. *Political Analysis* 22(1), 1-30

Installation

There are two ways to run the Conjoint Survey Design Tool (SDT).

Additional Resources for Conjoint Analysis

Thomas Leeper
https://github.com/leeper/conjoint-example

README.md

Example Conjoint Experimental Design in Qualtrics

This is an example of conducting a conjoint experiment in Qualtrics. It benefitted heavily from Kyle Dropp's guide, "Implementing a Conjoint Analysis in Qualtrics".

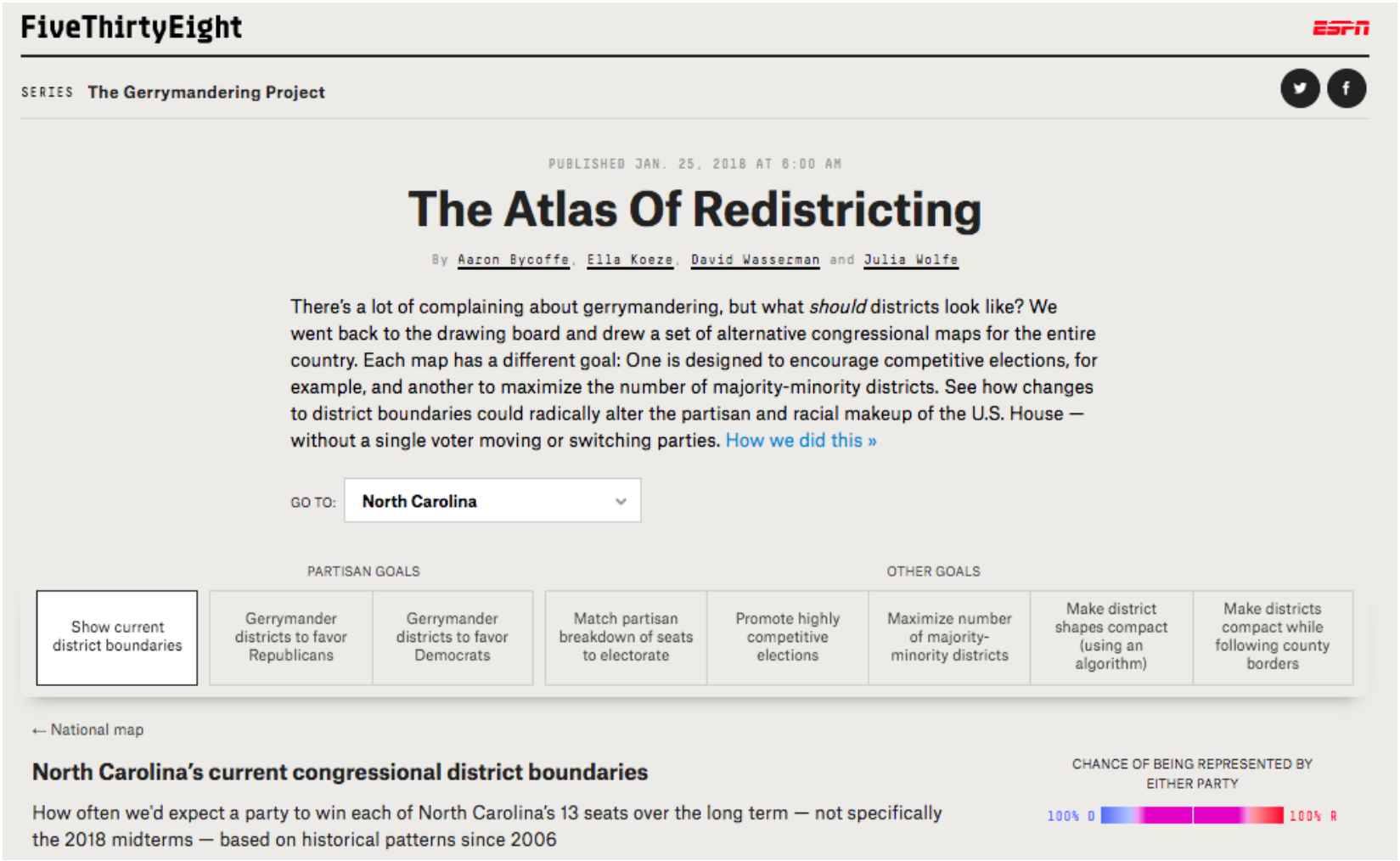
A conjoint is a fully randomized, factorial experiment involving a potentially large number of factors and factor levels. Typically, participants are exposed to a vignette wherein features of the vignette (e.g., a person description) are randomized. To gain statistical power and to avoid severe sparsity problems in such a high-dimensional design, within-subjects comparisons are typically leveraged. That is, a respondent may be exposed to 5-10 vignettes in order to multiple the statistical power of the design (under some plausible assumptions).

For example, in this example, participants are shown descriptions of two hypothetical political candidates with numerous demographic and attitudinal features:

Candidate A		Candidate B
62	**Age**	44
Male	**Sex**	Male
White	**Race**	White
Muslim	**Religion**	Mainline protestant
Member of Congress	**Occupation**	U.S. Senator
Republican	**Party**	Democrat
Did not serve	**Military Service**	Did not serve
Community college	**Education**	Ivy League university
	The candidates have reported opinions on the following issues:	
Moderately support	**Ratification of the Trans-Pacific Partnership (TPP)**	Slightly support
Strongly oppose	**Deploying ground troops to combat ISIS**	Strongly oppose

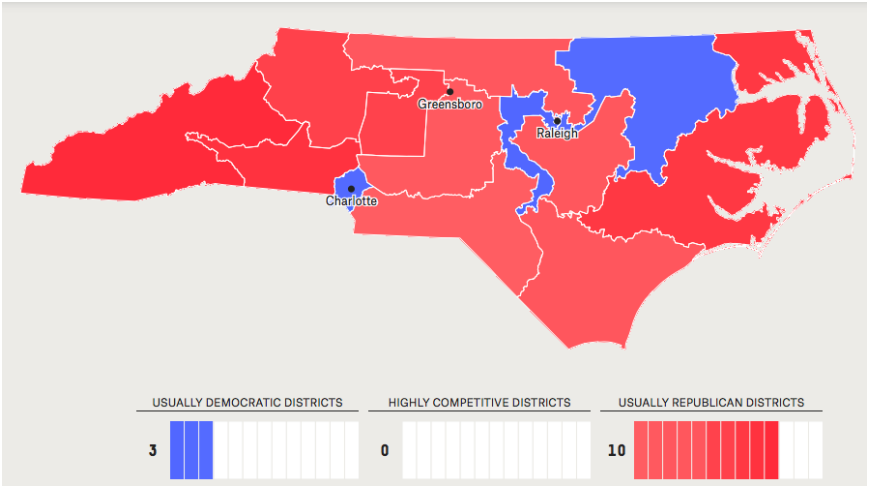
Atlas of Redistricting

FiveThirtyEight created 258 state congressional district maps

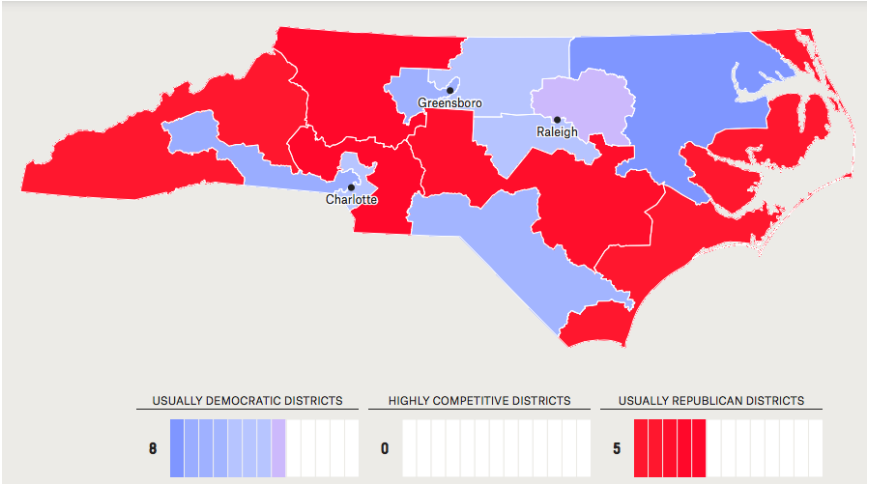


Conjoint Attributes: Proposed Map

Partisan gerrymander to benefit Republicans

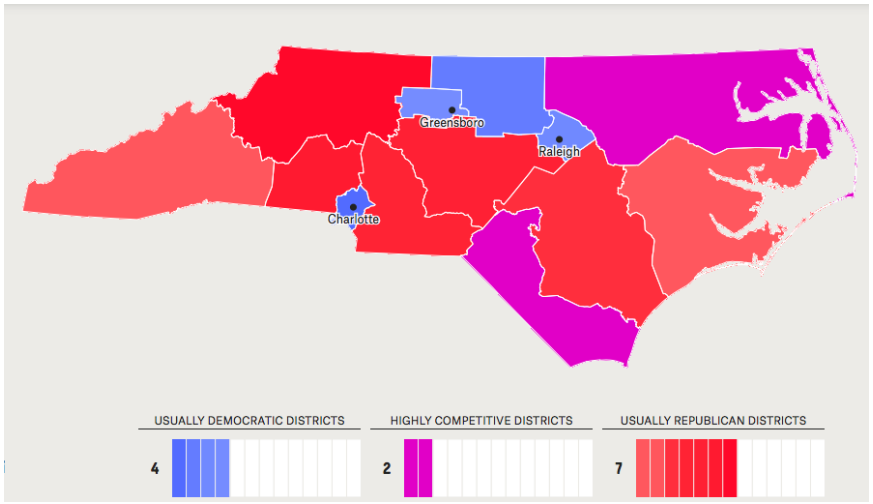


Partisan gerrymander to benefit Democrats

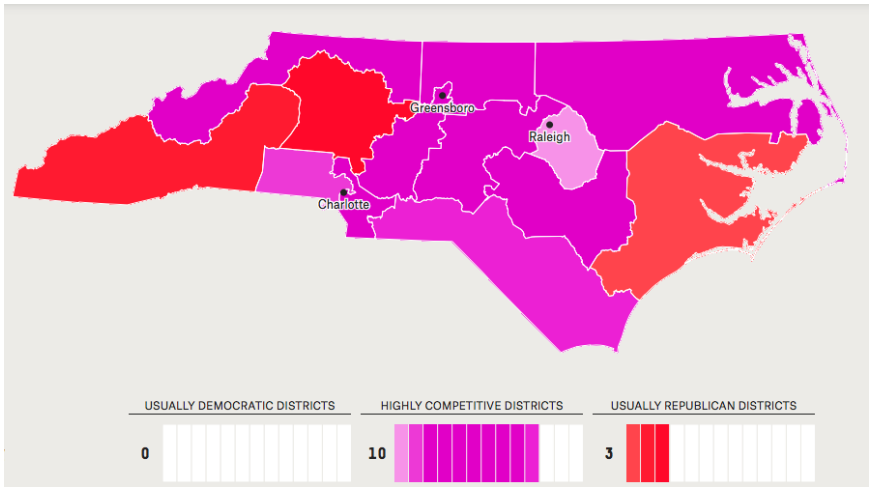


Conjoint Attributes: Proposed Map

Proportionally partisan map to match partisanship

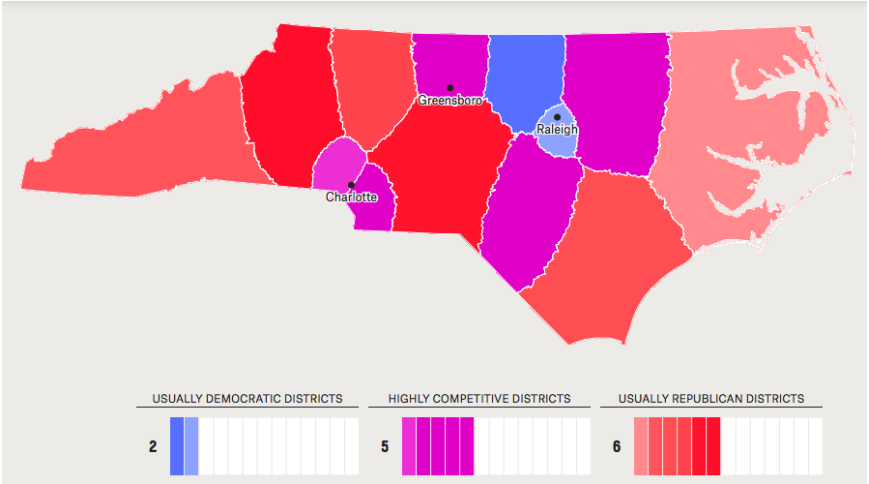


Highly competitive map

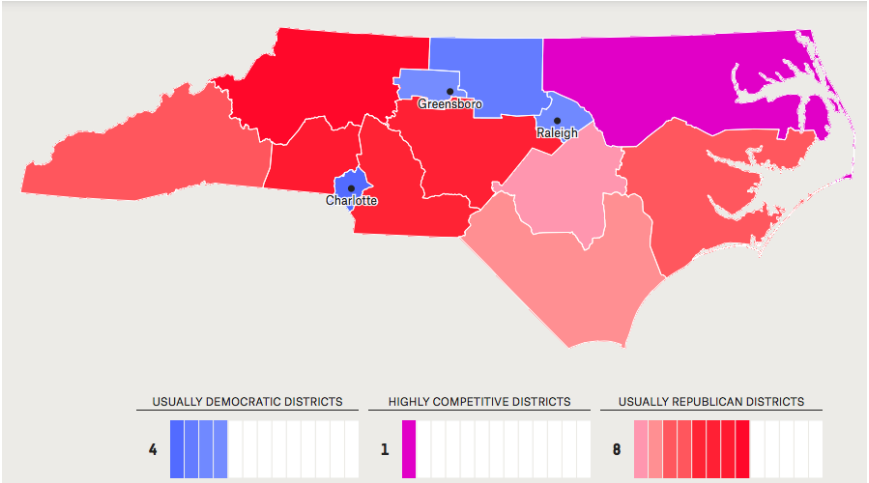


Conjoint Attributes: Proposed Map

Compact map using algorithm



Compact map drawn by hand



Other Conjoint Attributes

B. Proposed by
The baseline method is 'unknown' to provide a sense of what would happen without any information for this particular attribute. The other levels are as follows:

a. Republican state legislator
b. Consultant working for the Republican Party
c. Democratic state legislator
d. Consultant working for the Democratic Party
e. Bipartisan commission (with Republicans and Democrats)
f. Independent (nonpartisan) commission
g. Lobbyist
h. Academic researchers
j. Citizen group

C. Method
The baseline method is 'unknown' to provide a sense of what would happen without any information for this particular attribute. The other levels are as follows:

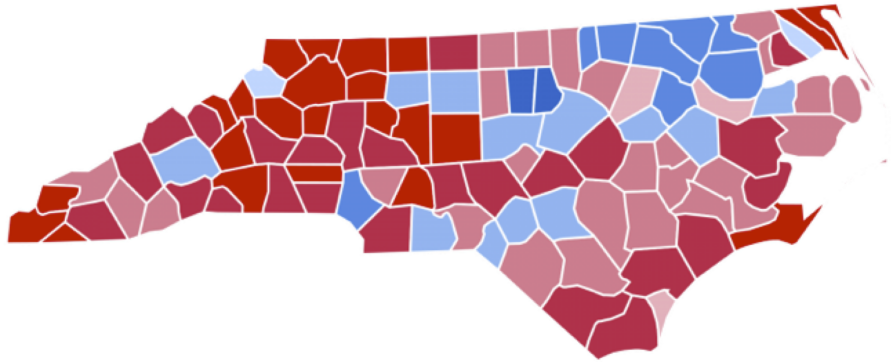
a. Human, using pencil and paper
b. Human, using mapping software
c. Computer algorithm
d. Artificial Intelligence

Design: 8 (map) x 9 (proposer) x 4 (method) = 288 combinations

Instructions - Background Information

Now you will be shown different ways that political boundaries might be drawn in the state of North Carolina. If you're not a resident of North Carolina, that's OK – just imagine that you are a voter in that state for this section.

North Carolina has a population of 10.5 million people. The figure below shows how people in North Carolina voted in the 2016 US Presidential Election. Counties in red represent more votes for Donald Trump; counties in blue represent more votes for Hillary Clinton. The darker the color, the greater the vote share for the winning candidate.



Data: Comparison of Sample Demographics

Step 3: Collect the data

	2019 Political Redistricting Study (Qualtrics)	ANES 2018 Pilot Study (YouGov)
Gender		
Male	45.2%	44.0%
Female	53.7%	56.0%
Age		
18 - 34	31.1%	24.8%
35 - 54	36.4%	32.0%
55 or older	32.4%	43.2%
Education		
High school (or less)	26.9%	30.9%
Some college	38.1%	33.8%
University degree (or higher)	35.0%	35.3%
Employment		
Full- or part-time	59.9%	46.3%
Retired	17.8%	21.6%
Not Working	22.3%	32.1%
Race / Ethnicity		
White	72.5%	74.2%
Non-white	27.5%	25.8%
Party ID		
Republican	28.2%	25.3%
Democrat	40.2%	35.7%
Independent	31.6%	39.0%
Number of US States	47	48 (plus DC)
Sample Size (N)	747	2,500

Step 4: Run analysis using R ('cjoint' package

Sample R Conjoint Script

```
## Load the conjoint data
uk.cjoint <- read.qualtrics("uk.cjoint.csv",
                            responses = c("conjoint1","conjoint2",
                            "conjoint3","conjoint4"),
                            respondentID = "id")

## Run AMCE estimator using all attributes in the design
uk.results <- amce(data = uk.cjoint,
                   selected ~ name + age + job + case + cost,
                   cluster = TRUE,
                   respondent.id = "respondentIndex",
                   baselines = baselines)
summary(uk.results)

## Plot results
plot(uk.results)
```

AMCE: Estimates from Conjoint

Formally: Average Marginal Component Effect (AMCE):

$$\bar{\pi}_l(t_1, t_0, p(\mathbf{t})) \equiv \mathbb{E}[Y_i(t_1, T_{ijk[-l]}, \mathbf{T}_{i[-j]k}) - Y_i(t_0, T_{ijk[-l]}, \mathbf{T}_{i[-j]k})]$$

$$= \sum_{[t_{[-l]}, \mathbf{t}_{[-j]}] \in \tilde{\mathcal{T}}} \mathbb{E}[Y_i(t_1, t_{[-l]}, \mathbf{t}_{[-j]}) - Y_i(t_0, t_{[-l]}, \mathbf{t}_{[-j]})] p(t_{[-l]}, \mathbf{t}_{[-j]}),$$

where

$$\begin{cases} T_{ijk(-l)} = [T_{ijk1} \cdots T_{ijk(l-1)} \; T_{ijk(l+1)} \cdots T_{ijkL}] & \text{(all components but } l\text{th)} \\ \mathbf{T}_{i(-j)k} = [T_{i1k} \cdots T_{i(j-1)k} \; T_{i(j+1)k} \cdots T_{iJk}]^\top & \text{(all profiles but } j\text{th)} \end{cases}$$

- $\tilde{\mathcal{T}}$ excludes "empty counterfactuals" (e.g. a research scientist with no formal education)
- Nonparametrically identified under Assumptions 1, 2 and 3
- Interaction effects can be similarly defined and identified

AMCE: Estimates from Conjoint

- If profiles are not restricted, attributes can be completely randomized:

$$T_{ijkl} \perp\!\!\!\perp \{T_{ijk[-l]}, \mathbf{T}_{i[-j]k}\} \quad \text{for all} \quad i, j, k.$$

- In this case, the simple difference-in-means estimator is unbiased:

$$\hat{\bar{\pi}}_l(t_1, t_0, p(\mathbf{t})) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}\mathbf{1}\{T_{ijkl} = t_1\}}{n_1} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}\mathbf{1}\{T_{ijkl} = t_0\}}{n_0},$$

  where n_1 and n_0 are the numbers of profiles in the two conditions

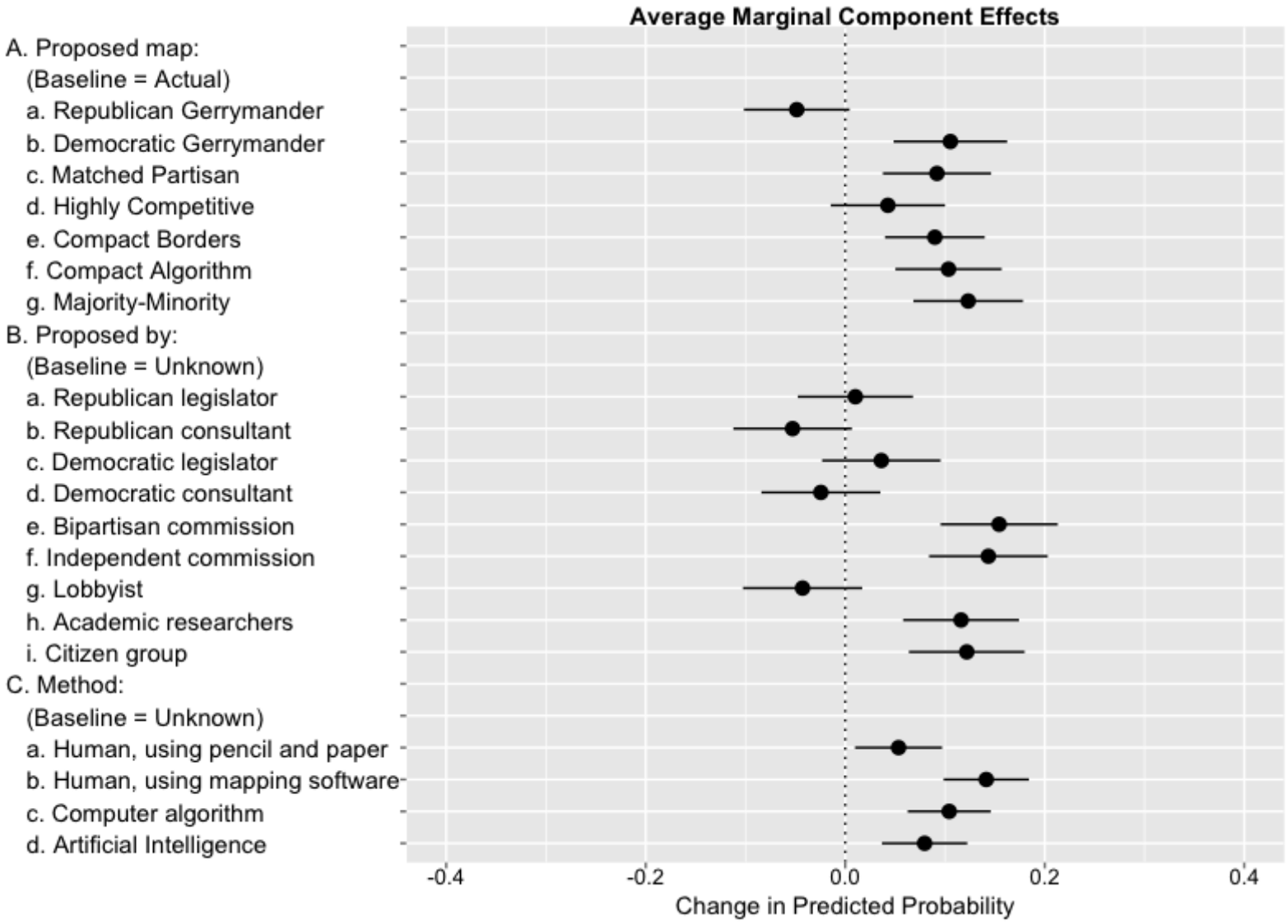
- A numerically equivalent estimate can be obtained by a regression (regress Y_{ijk} on a set of dummies for the levels of T_{ijkl})
- Works for both choice and rating outcomes

Conjoint Analysis Results - AMCEs

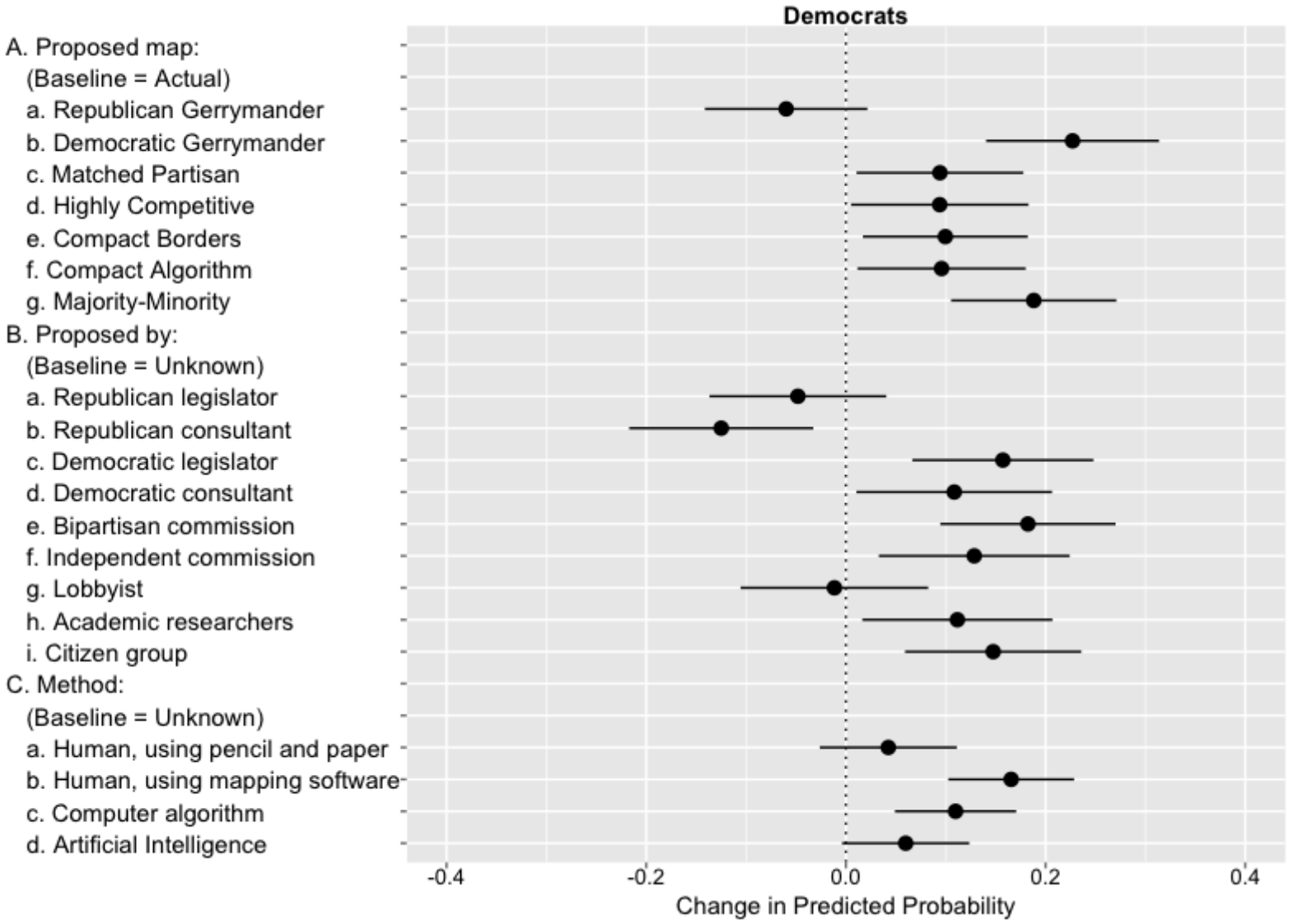
Conjoint Results

```
> summary(results[[1]])
-----------------------------------------
Average Marginal Component Effects (AMCE):
-----------------------------------------
       Attribute                              Level  Estimate Std. Err  z value    Pr(>|z|)
 A. Proposed map          a. Republican Gerrymander -0.048602 0.027035 -1.79771 7.2222e-02
 A. Proposed map          b. Democratic Gerrymander  0.105409 0.029038  3.63006 2.8335e-04 ***
 A. Proposed map                c. Matched Partisan  0.091962 0.027683  3.32190 8.9406e-04 ***
 A. Proposed map             d. Highly Competitive  0.042758 0.029244  1.46210 1.4371e-01
 A. Proposed map                e. Compact Borders  0.089861 0.025518  3.52147 4.2916e-04 ***
 A. Proposed map              f. Compact Algorithm  0.103522 0.027167  3.81056 1.3865e-04 ***
 A. Proposed map             g. Majority-Minority  0.123269 0.028038  4.39651 1.1001e-05 ***
  B. Proposed by           a. Republican legislator  0.010178 0.029478  0.34528 7.2988e-01
  B. Proposed by           b. Republican consultant -0.052771 0.030319 -1.74056 8.1762e-02
  B. Proposed by           c. Democratic legislator  0.036165 0.030302  1.19350 2.3268e-01
  B. Proposed by           d. Democratic consultant -0.024424 0.030442 -0.80231 4.2237e-01
  B. Proposed by           e. Bipartisan commission  0.154273 0.029969  5.14778 2.6359e-07 ***
  B. Proposed by         f. Independent commission  0.143494 0.030341  4.72939 2.2520e-06 ***
  B. Proposed by                      g. Lobbyist -0.042753 0.030510 -1.40130 1.6113e-01
  B. Proposed by            h. Academic researchers  0.116113 0.029637  3.91784 8.9345e-05 ***
  B. Proposed by                  i. Citizen group  0.121833 0.029587  4.11786 3.8241e-05 ***
       C. Method a. Human, using pencil and paper  0.053484 0.022298  2.39864 1.6456e-02   *
       C. Method b. Human, using mapping software  0.141363 0.021820  6.47871 9.2511e-11 ***
       C. Method              c. Computer algorithm  0.104233 0.021248  4.90559 9.3145e-07 ***
       C. Method         d. Artificial Intelligence  0.079567 0.021843  3.64267 2.6982e-04 ***
---
```

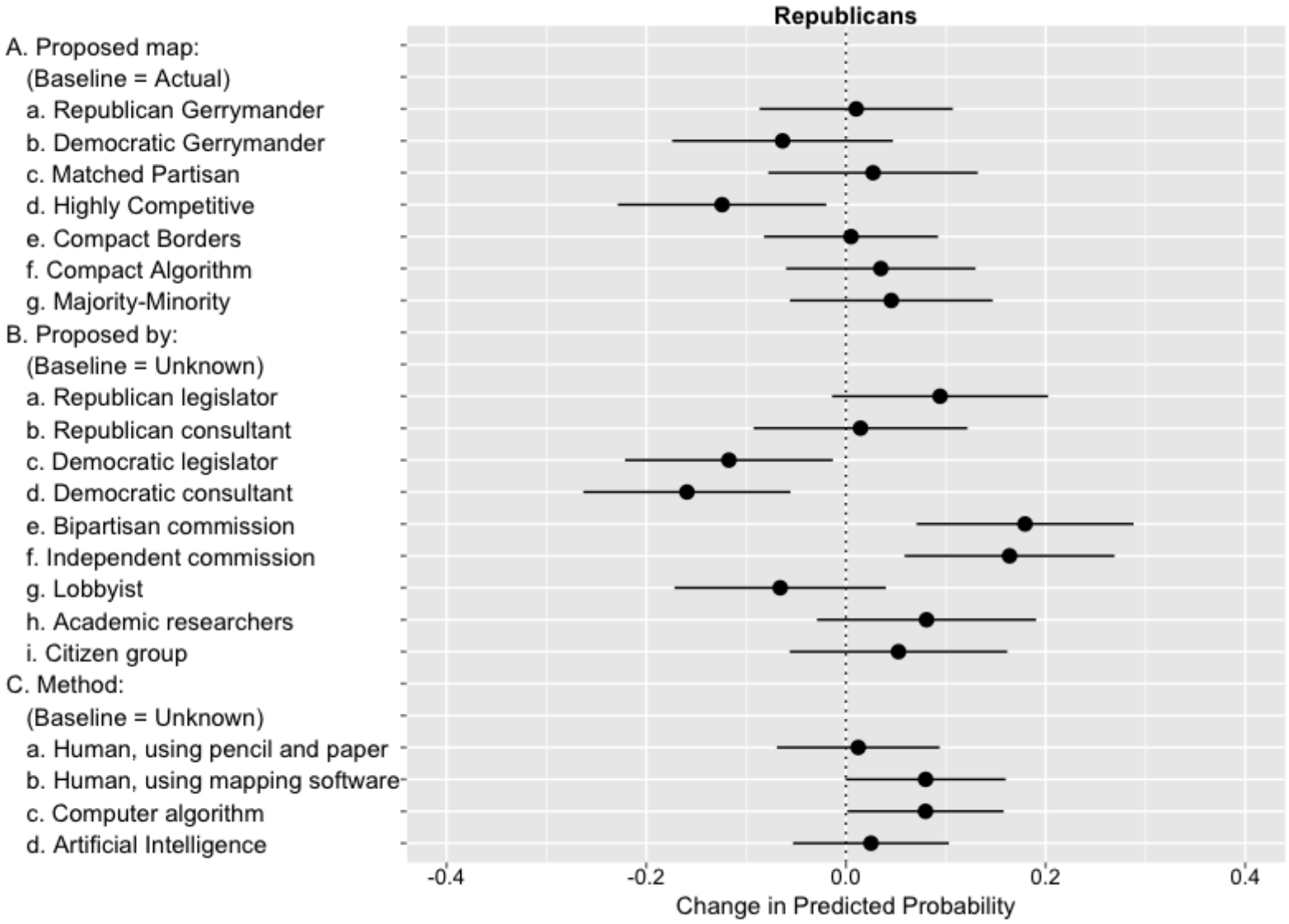
Conjoint Analysis - Full Sample AMCEs



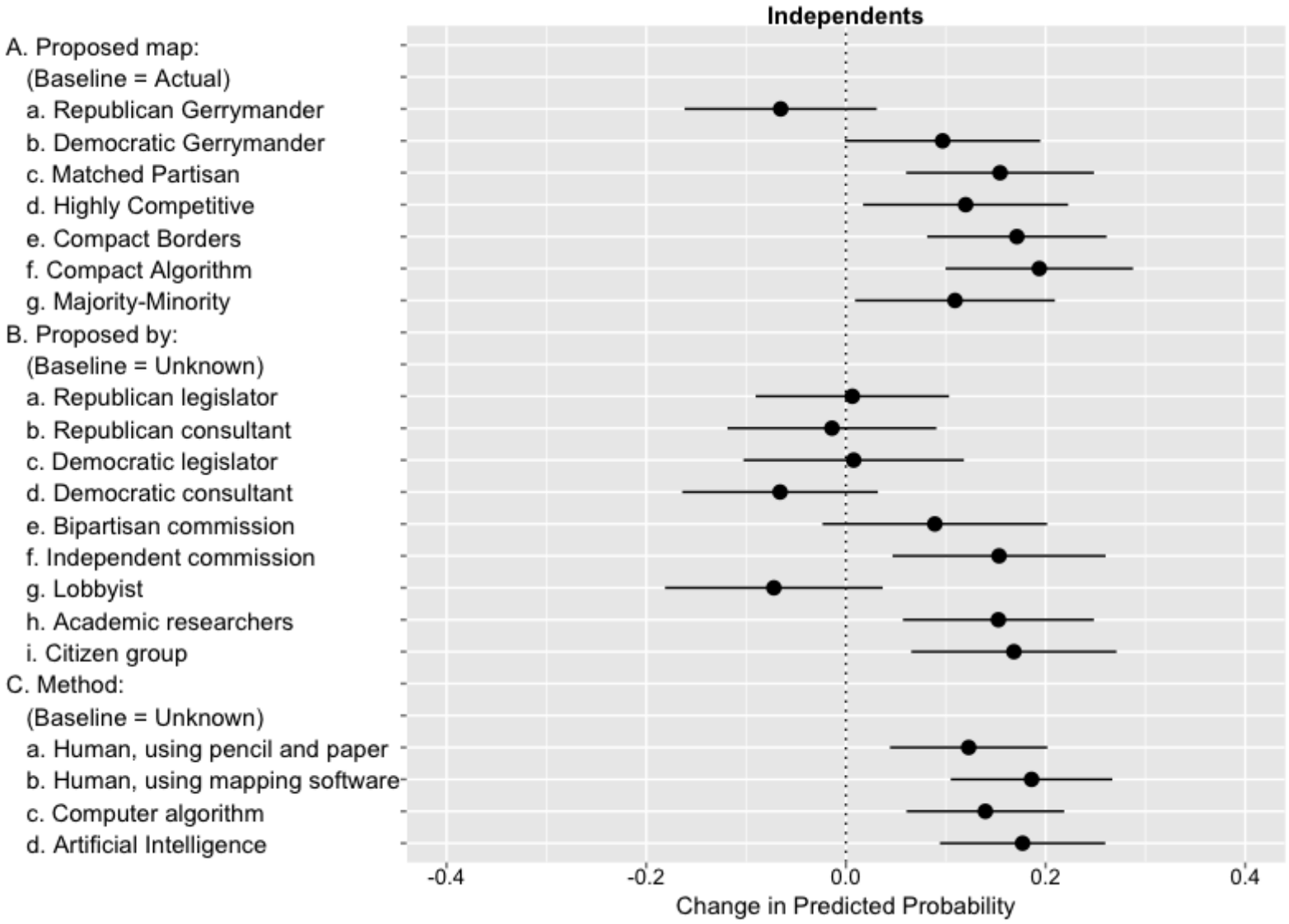
Partisan Motivated Reasoning - Democrats



Partisan Motivated Reasoning - Republicans



Nonpartisan Motivations - Independents



Conclusions

Public preferences for electoral boundary maps affected by...

- Motivated reasoning:

  Partisans prefer gerrymandered maps choose this method

- Fairness norms: Independents prefer competitive or matched maps
- Source cues: Partisans influenced by who proposed map

Fmr Rep Governor of North Carolina:
"Everyone's against gerrymandering when they're not the ones in power."

Recap: Conjoint Experiment

- Multivariate, choice-based technique originating from marketing research

  (Green & Rao, 1971)
- Respondents presented with profiles generated from randomly selected features (levels)
- Simpler decision task, thus reducing the cognitive fatigue (Louviere et al., 2008)
- Can test tens of thousands of potential combinations in relatively small samples

  (Hainmueller et al. 2013, 2015; Hainmueller & Hopkins, 2015)

More Resources for Conjoint Analysis

Automated Conjoint Analysis in R
https://github.com/tkhartman/automated-conjoint-analysis

Todd K. Hartman
tkhartman

Senior Lecturer in Quantitative Social Science at The University of Sheffield
The University of Sheffield
Sheffield, UK
Sign in to view email
https://tkhartman.netlify.com/

Block or report user

Overview | Repositories 15 | Projects 0 | Stars 4 | Followers 5 | Following 1

Popular repositories

academic-kickstart
Shell ★ 1

replication-hurdle-count-model
Replication files: Simulations in R testing negative binomial hurdle count model against alternative specifications; Stata data analysis dofile; doi:10.1007/s10940-016-9283-z
Stata

qss
Course Files: Introductory Data Analysis for Quantitative Social Scientists (QSS)
R 1

animating-election-maps
R script to animate US presidential elections, 1960 - 2016 (two-party vote share by state)
R

tkhartman.github.io
HTML

mapping-brexit