Survey Experiments

Research Programs and Conjoint Analysis

Ryan T. Moore

2024-08-22

The Margin of Error and Sample Size

Margin of Error and Sample Size

In survey sampling, we sometimes refer to the *margin of error* (MoE). This is a component of the confidence interval calculation:

$$\left[\text{Estimate } - \underbrace{\text{Critical Value} \cdot SE}_{\text{Margin of Error}}, \quad \text{Estimate } + \underbrace{\text{Critical Value } \cdot SE}_{\text{Margin of Error}}\right]$$

Margin of Error and Sample Size

That is,

$$\text{MoE} = \text{Critical Value } \cdot SE$$

Margin of Error and Sample Size

That is,

$$\text{MoE} = \text{Critical Value } \cdot SE$$

We can use this to find the minimum sample size for a certain level of precision in a survey.

Margin of Error and Sample Size

That is,

$$\text{MoE} = \text{Critical Value } \cdot SE$$

We can use this to find the minimum sample size for a certain level of precision in a survey.

Suppose we have a survey asking whether Scottish voters support Brexit, and we want it to be precise to within 0.03 (three percentage points), with 95% confidence.

Margin of Error and Sample Size

That is,

$$\text{MoE} = \text{Critical Value } \cdot SE$$

We can use this to find the minimum sample size for a certain level of precision in a survey.

Suppose we have a survey asking whether Scottish voters support Brexit, and we want it to be precise to within 0.03 (three percentage points), with 95% confidence.

Margin of Error and Sample Size

SE for sample proportion:

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Margin of Error and Sample Size

SE for sample proportion:

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

To be conservative, use largest possible SE

Margin of Error and Sample Size

SE for sample proportion:

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

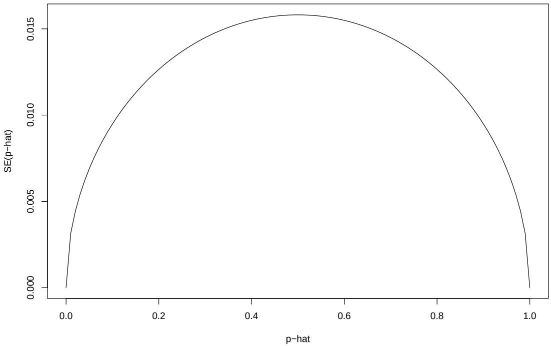
To be conservative, use largest possible SE – occurs at $\hat{p} = 0.5$.

Margin of Error and Sample Size

SE for sample proportion:

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

To be conservative, use largest possible SE – occurs at $\hat{p} = 0.5$.



Margin of Error and Sample Size

So, find n given other parameters:

$$\text{MoE} \quad = \quad \text{Critical Value} \ \cdot SE$$

Margin of Error and Sample Size

So, find n given other parameters:

$$\begin{aligned}\text{MoE} &= \text{Critical Value } \cdot SE \\ 0.03 &= 1.96 \cdot \sqrt{\frac{.5(1-0.5)}{n}}\end{aligned}$$

Margin of Error and Sample Size

So, find n given other parameters:

$$\begin{aligned}
\text{MoE} &= \text{Critical Value } \cdot SE \\
0.03 &= 1.96 \cdot \sqrt{\frac{.5(1-0.5)}{n}} \\
0.03^2 &= 1.96^2 \cdot \frac{.5(1-0.5)}{n}
\end{aligned}$$

Margin of Error and Sample Size

So, find n given other parameters:

$$\begin{aligned}
\text{MoE} &= \text{Critical Value} \cdot SE \\
0.03 &= 1.96 \cdot \sqrt{\frac{.5(1-0.5)}{n}} \\
0.03^2 &= 1.96^2 \cdot \frac{.5(1-0.5)}{n} \\
n &= 1.96^2 \cdot \frac{.5(1-0.5)}{0.03^2}
\end{aligned}$$

Margin of Error and Sample Size

So, find n given other parameters:

$$
\begin{aligned}
\text{MoE} &= \text{Critical Value } \cdot SE \\
0.03 &= 1.96 \cdot \sqrt{\frac{.5(1-0.5)}{n}} \\
0.03^2 &= 1.96^2 \cdot \frac{.5(1-0.5)}{n} \\
n &= 1.96^2 \cdot \frac{.5(1-0.5)}{0.03^2} \\
n &\approx 3.8416 \cdot 277.8 \\
n &\approx 1067
\end{aligned}
$$

Power and Sample Size

The MoE can help guide sampling, but ignores the treatment effect.

Power and Sample Size

The MoE can help guide sampling, but ignores the treatment effect.

If a survey experimental treatment effect $= 0.15$, MoE of 0.03 will likely detect it.

Power and Sample Size

The MoE can help guide sampling, but ignores the treatment effect.

If a survey experimental treatment effect $= 0.15$, MoE of 0.03 will likely detect it.

If a survey experimental treatment effect $= 0.02$, MoE of 0.03 *not* likely to detect it.

Power and Sample Size

Power is probability of detecting a particular TE, if it in fact exists.

Power and Sample Size

Power is probability of detecting a particular TE, if it in fact exists.

E.g., what's our probability of detecting effect of 0.02, with total sample size 1067, from base rate of 0.5?

Power and Sample Size

Power is probability of detecting a particular TE, if it in fact exists.

E.g., what's our probability of detecting effect of 0.02, with total sample size 1067, from base rate of 0.5?

```
power.prop.test(n = 1067 / 2, p1 = 0.5, p2 = 0.52)
```

```
##
##      Two-sample comparison of proportions power calc
##
##               n = 533.5
##              p1 = 0.5
##              p2 = 0.52
##       sig.level = 0.05
##           power = 0.09564109
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Power and Sample Size

For continuous outcomes, everything matters that is in

$$SE(\widehat{ATE}) = \sqrt{\frac{1}{N-1}\left[\frac{m\text{Var}(Y_i(0))}{N-m} + \frac{(N-m)\text{Var}(Y_i(1))}{m} + 2\text{Cov}(Y_i(0), Y_i(1))\right]}$$

Power and Sample Size

For continuous outcome,

```r
power.t.test(n = 100, delta = 1.5, sd = 1)
```

Power and Sample Size

For continuous outcome,

```r
power.t.test(n = 100, delta = 1.5, sd = 1)
```

```
##
##      Two-sample t test power calculation
##
##               n = 100
##           delta = 1.5
##              sd = 1
##       sig.level = 0.05
##           power = 1
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

Power and Sample Size

(See `code/03-power.R`.)

Power and Sample Size

(We often use simulation to get power, sample size, MDE for complex assignments or estimation strategies.)

List Experiments

Split the class!

- I have read George Orwell's *1984*.
- I am currently enrolled in a doctoral program (PhD).
- I purchased groceries this week.

- I have read George Orwell's *1984*.
- I am currently enrolled in a doctoral program (PhD).
- I purchased groceries this week.
- I have ridden a bicycle in the month of August.

- I have read George Orwell's *1984*.
- I am currently enrolled in a doctoral program (PhD).
- I purchased groceries this week.
- I have ridden a bicycle in the month of August.

Let X be the number of items agreed with.

$X_{\text{Group1}} = \_\_\_\_$
$X_{\text{Group2}} = \_\_\_\_$

1991 US National Race and Politics Survey

Now I'm going to read you three things that sometimes make people angry or upset. After I read all three, just tell me HOW MANY of them upset you. (I don't want to know which ones, just how many.)

(1) the federal government increasing the tax on gasoline
(2) professional athletes getting million-dollar-plus salaries
(3) large corporations polluting the environment

How many, if any, of these things upset you?

1991 US National Race and Politics Survey

Now I'm going to read you four things that sometimes make people angry or upset. After I read all four, just tell me HOW MANY of them upset you. (I don't want to know which ones, just how many.)

(1) the federal government increasing the tax on gasoline

(2) professional athletes getting million-dollar-plus salaries

(3) large corporations polluting the environment

(4) a black family moving next door to you

How many, if any, of these things upset you?

Assumptions

Blair and Imai (2012) formalise analysis.

Assumptions

Blair and Imai (2012) formalise analysis.

Assumptions:

1. No design effects.
2. No liars.

Assumptions

Blair and Imai (2012) formalise analysis.

Assumptions:

1. No design effects.
2. No liars.

Under these two, difference in means is *unbiased*.

Assumptions

Blair and Imai (2012) formalise analysis.

Assumptions:

1. No design effects.
2. No liars.

Under these two, difference in means is *unbiased*.

Let τ = true ATE. Let

$$\hat{\tau} = \frac{1}{n_{\text{Tr}}} \sum_{i=1}^{n} T_i Y_i - \frac{1}{n_{\text{Co}}} \sum_{i=1}^{n} (1 - T_i) Y_i$$

Assumptions

Blair and Imai (2012) formalise analysis.

Assumptions:

1. No design effects.
2. No liars.

Under these two, difference in means is *unbiased*.

Let τ = true ATE. Let

$$\hat{\tau} = \frac{1}{n_{\mathrm{Tr}}} \sum_{i=1}^{n} T_i Y_i - \frac{1}{n_{\mathrm{Co}}} \sum_{i=1}^{n} (1 - T_i) Y_i$$

Then

$$E(\hat{\tau}) = \tau$$

Assumptions

1. No design effects.

Assumptions

1. No design effects.

Count of (3) control items is constant:

Assumptions

1. No design effects.

Count of (3) control items is constant:

$$\sum_{j=1}^{J} Z_{ij}(0) = \sum_{j=1}^{J} Z_{ij}(1)$$

Assumptions

1. No design effects.

Count of (3) control items is constant:

$$\sum_{j=1}^{J} Z_{ij}(0) = \sum_{j=1}^{J} Z_{ij}(1)$$

or, (count under Tr) = (count under Co) + (0/1 for sensitive item):

$$Y_i(1) = Y_i(0) + Z_{i,J+1}(1)$$

Assumptions

2. No liars. Response to sensitive item is true.

Assumptions

2. No liars. Response to sensitive item is true.

$$Z_{i,J+1}(1) = Z^*_{i,J+1}$$

Interpretation

The unbiasedness of the difference in means implies

$$E(\hat{\tau}) = Pr(Z^*_{i,J+1} = 1)$$

Interpretation

The unbiasedness of the difference in means implies

$$E(\hat{\tau}) = Pr(Z^*_{i,J+1} = 1)$$

Interpret difference-in-means as
"probability sensitive item is true."

Interpretation

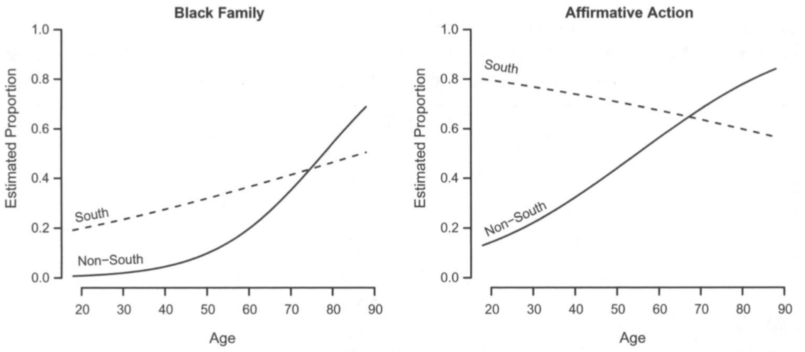
The unbiasedness of the difference in means implies

$$E(\hat{\tau}) = Pr(Z^*_{i,J+1} = 1)$$

Interpret difference-in-means as
"probability sensitive item is true."

As with prior diff-in-means, can calculate it via regression.

Interpretation



Interpretation

What is *social desirability bias* in surveys?

Interpretation

What is *social desirability bias* in surveys?

$$S(x) = Pr(Z_{i,J+1}(0) = 1|X_i = x) - Pr(Z^*_{i,J+1} = 1|X_i = x)$$

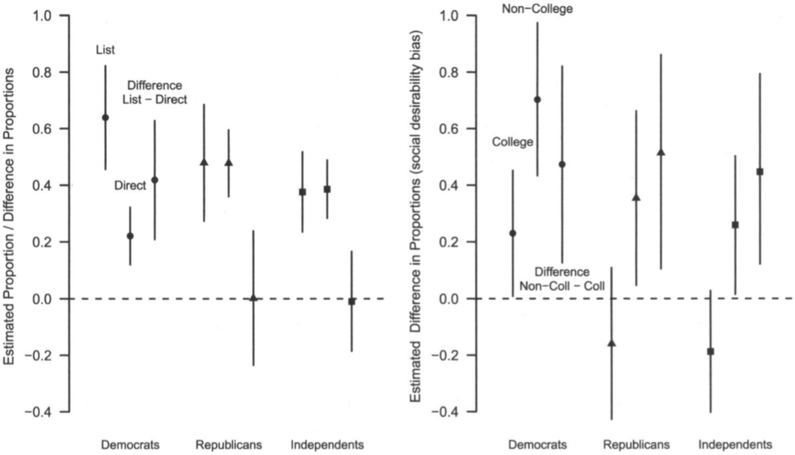
Interpretation

What is *social desirability bias* in surveys?

$$S(x) = Pr(Z_{i,J+1}(0) = 1 | X_i = x) - Pr(Z^*_{i,J+1} = 1 | X_i = x)$$

(First term: shown control, then asked *directly*)

Interpretation



Conjoint Experiments

Conjoint Experiments

Conjoint experiments . . .

- ask participants to select between hypothetical profiles, with their attributes randomized;

Conjoint Experiments

Conjoint experiments ...

- ask participants to select between hypothetical profiles, with their attributes randomized;
- are a way to address multidimensionality – many factors, but only one vote;

Conjoint Experiments

Conjoint experiments ...

- ask participants to select between hypothetical profiles, with their attributes randomized;
- are a way to address multidimensionality – many factors, but only one vote;
- can randomly assign attributes, or randomly assign them conditional on some restrictions;

Conjoint Experiments

Conjoint experiments ...

- ask participants to select between hypothetical profiles, with their attributes randomized;
- are a way to address multidimensionality – many factors, but only one vote;
- can randomly assign attributes, or randomly assign them conditional on some restrictions;
- can be a little tricky to interpret.

Please read the descriptions of the potential immigrants carefully. Then, please indicate which of the two immigrants you would personally prefer to see admitted to the United States.

	Immigrant 1	**Immigrant 2**
Prior Trips to the U.S.	Entered the U.S. once before on a tourist visa	Entered the U.S. once before on a tourist visa
Reason for Application	Reunite with family members already in U.S.	Reunite with family members already in U.S.
Country of Origin	Mexico	Iraq
Language Skills	During admission interview, this applicant spoke fluent English	During admission interview, this applicant spoke fluent English
Profession	Child care provider	Teacher
Job Experience	One to two years of job training and experience	Three to five years of job training and experience
Employment Plans	Does not have a contract with a U.S. employer but has done job interviews	Will look for work after arriving in the U.S.
Education Level	Equivalent to completing two years of college in the U.S.	Equivalent to completing a college degree in the U.S.
Gender	Female	Male

	Immigrant 1	Immigrant 2
If you had to choose between them, which of these two immigrants should be given priority to come to the United States to live?	○	○

```r
load("../data/03-candidate.RData")
cand <- x
head(cand)
```

```
##             resID    atmilitary          atreligion
## 1 A2NEN4NSNS208S Did Not Serve              Mormon Commu
## 2 A2NEN4NSNS208S        Served                None
## 3 A2NEN4NSNS208S Did Not Serve            Catholic     S
## 4 A2NEN4NSNS208S        Served Mainline protestant     S
## 5 A2NEN4NSNS208S        Served                None   Bap
## 6 A2NEN4NSNS208S        Served Mainline protestant
##                atprof atinc          atrace atage atmale
## 1          Car dealer  5.1M           White    75 Female
## 2 High school teacher   65K  Asian American    60   Male
## 3              Farmer   32K Native American    68 Female
## 4              Doctor   54K           White    75   Male
## 5              Doctor  5.1M Native American    45 Female
## 6              Lawyer   54K           White    52   Male
```

```r
load("../data/03-immigrant.RData")
immig <- x
head(x)
```

```
##   CaseID contest_no
## 1      4          1       Equivalent to completing high
## 2      4          1                                   No
## 3      4          2 Equivalent to completing a graduate
## 4      4          2      Equivalent to completing fourth
## 5      4          3       Equivalent to completing high
## 6      4          3  Equivalent to completing a college
##   FeatGender FeatCountry
## 1       male        Iraq                        Seek be
## 2     female      France                        Seek be
## 3     female       Sudan        Escape political/relig
## 4     female     Germany Reunite with family members al
## 5     female Philippines                        Seek be
## 6       male       Sudan                        Seek be
##                FeatJob
## 1                Nurse More than five years of job traini
```

Notation

Indices

- i respondent
- j alternative (candidate 2)
- k task (3rd task)
- l component of profile

Notation

Treatments

- T_{ijkl} a component shown
- T_{ijk} a profile shown (from $\mathcal{T}$)
- $\mathbf{T}_{ik}$ all profiles shown for task k
- $\bar{\mathbf{T}}_i$ all profiles hypothetically shown $(J \cdot K)$
- $\mathbf{t}$ all profiles actually shown, in sequence

Notation

Potential outcomes

- $Y_{ik}(\bar{\mathbf{t}})$ pot outcomes **observed** under full seq of profiles (J-dim)
- $Y_{ik} \equiv Y_{ik}(\bar{\mathbf{T}}_i)$ pot outcomes under hypothetical full seq of profiles (J-dim)
- $Y_{ijk}(\bar{\mathbf{T}}_i)$ component of $Y_{ik}(\bar{\mathbf{T}}_i)$
- $Y_i(\mathbf{t})$ pot outcomes under observed seq of profiles (given Assumption 1)

Assumptions

1. Stability, no carryover: if $\mathbf{T}_{ik} = \mathbf{T}'_{ik'}$,

$$Y_{ijk}(\bar{\mathbf{T}}_i) = Y_{ijk'}(\bar{\mathbf{T}}'_i)$$

   If treatments in task 1 same as treatments in task 6, same potential outcomes.

Assumptions

1. Stability, no carryover: if $\mathbf{T}_{ik} = \mathbf{T}'_{ik'}$,

$$Y_{ijk}(\bar{\mathbf{T}}_i) = Y_{ijk'}(\bar{\mathbf{T}}'_i)$$

   If treatments in task 1 same as treatments in task 6, same potential outcomes.

2. No profile-order effects: if $T_{ijk} = T'_{ij'k}$ and $T_{ij'k} = T'_{ijk}$,

$$Y_{ij}(\mathbf{T}_{ik}) = Y_{ij'}(\mathbf{T}'_{ik})$$

   Treatment A-B has same potential outcomes as treatment B-A.

Assumptions

1. Stability, no carryover: if $\mathbf{T}_{ik} = \mathbf{T}'_{ik'}$,

$$Y_{ijk}(\bar{\mathbf{T}}_i) = Y_{ijk'}(\bar{\mathbf{T}}'_i)$$

   If treatments in task 1 same as treatments in task 6, same potential outcomes.

2. No profile-order effects: if $T_{ijk} = T'_{ij'k}$ and $T_{ij'k} = T'_{ijk}$,

$$Y_{ij}(\mathbf{T}_{ik}) = Y_{ij'}(\mathbf{T}'_{ik})$$

   Treatment A-B has same potential outcomes as treatment B-A.

3. Randomization:

$$Y_i(\mathbf{t}) \text{ indep of } T_{ijkl}$$

   Profiles don't disprop go to those who like them, e.g.

An Estimand: ATE

With Holland (1986), “the effect of a cause is always relative to another cause”

An Estimand: ATE

With Holland (1986), "the effect of a cause is always relative to another cause"

$$\bar{\pi}(\mathbf{t}_1, \mathbf{t}_0) = E\left[Y_i(\mathbf{t}_1) - Y_i(\mathbf{t}_0)\right]$$

An Estimand: ATE

With Holland (1986), "the effect of a cause is always relative to another cause"

$$\bar{\pi}(\mathbf{t}_1, \mathbf{t}_0) = E\left[Y_i(\mathbf{t}_1) - Y_i(\mathbf{t}_0)\right]$$

Can this be observed?

An Estimand: ATE

With Holland (1986), "the effect of a cause is always relative to another cause"

$$\bar{\pi}(\mathbf{t}_1, \mathbf{t}_0) = E\left[Y_i(\mathbf{t}_1) - Y_i(\mathbf{t}_0)\right]$$

Can this be observed?

The Fundamental Problem of Causal Inference

An Estimand: AMCE

$$\begin{aligned}\bar{\pi}_l(t_1, t_0, p(\mathbf{t})) \quad &\equiv \quad E\left[Y_i\left(t_1, T_{ijk[-l]}, T_{i[-j]k}\right)\right.\\ &\left.-Y_i\left(t_0, T_{ijk[-l]}, T_{i[-j]k}\right) \mid \left(T_{ijk[-l]}, T_{i[-j]k}\right) \in \widetilde{\mathcal{T}}\right]\end{aligned}$$

An Estimand: AMCE

$$\begin{aligned}\bar{\pi}_l(t_1, t_0, p(\mathbf{t})) \quad &\equiv \quad E\Big[Y_i\left(t_1, T_{ijk[-l]}, T_{i[-j]k}\right) \\ &\quad -Y_i\left(t_0, T_{ijk[-l]}, T_{i[-j]k}\right) \mid \left(T_{ijk[-l]}, T_{i[-j]k}\right) \in \widetilde{\mathcal{T}}\Big]\end{aligned}$$

Can this be observed?

An Estimand: AMCE

$$
\begin{aligned}
\bar{\pi}_l(t_1, t_0, p(\mathbf{t})) \quad \equiv \quad & E\Big[Y_i\left(t_1, T_{ijk[-l]}, T_{i[-j]k}\right) \\
& -Y_i\left(t_0, T_{ijk[-l]}, T_{i[-j]k}\right) \mid \left(T_{ijk[-l]}, T_{i[-j]k}\right) \in \widetilde{\mathcal{T}}\Big]
\end{aligned}
$$

Can this be observed?

The Fundamental Problem of Causal Inference

The AMCE in 2×2 Designs

- Collapsing into 2 groups $\rightsquigarrow$ AMCE

The AMCE in 2×2 Designs

- Collapsing into 2 groups $\rightsquigarrow$ AMCE
- Weight by probs from joint dist'ns of other attribute

Estimation: Calculating the AMCE

See `code/03-amce-simple.R`.

Related Findings

- How many *attributes*?

Related Findings

- How many *attributes*?
- A lot!

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: $0.198 \rightarrow 0.147$

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: $0.198 \rightarrow 0.147$
  - Effect of SSM position: $0.228 \rightarrow 0.190$

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: $0.198 \rightarrow 0.147$
  - Effect of SSM position: $0.228 \rightarrow 0.190$
  - Effect of health care position: $0.146 \rightarrow 0.090$

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: $0.198 \rightarrow 0.147$
  - Effect of SSM position: $0.228 \rightarrow 0.190$
  - Effect of health care position: $0.146 \rightarrow 0.090$
  - Effect of age 72 penalty nearly disappears

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: 0.198 $\rightarrow$ 0.147
  - Effect of SSM position: 0.228 $\rightarrow$ 0.190
  - Effect of health care position: 0.146 $\rightarrow$ 0.090
  - Effect of age 72 penalty nearly disappears

Related Findings

- How many *attributes*?
- A lot!
- Using 4 core attributes vs 25 or 35 (!) fillers
  - Effect of PID: $0.198 \rightarrow 0.147$
  - Effect of SSM position: $0.228 \rightarrow 0.190$
  - Effect of health care position: $0.146 \rightarrow 0.090$
  - Effect of age 72 penalty nearly disappears

See Bansak et al. (2021).

Related Findings

- How many *tasks*?

Related Findings

- How many *tasks*?
- Dozens!

Related Findings

- How many *tasks*?
- Dozens!
- (But biggest effects on first task)

Related Findings

- How many *tasks*?
- Dozens!
- (But biggest effects on first task)

Related Findings

- How many *tasks*?
- Dozens!
- (But biggest effects on first task)

See Bansak et al. (2018).

Strange things can happen ...

Strange things can happen ...

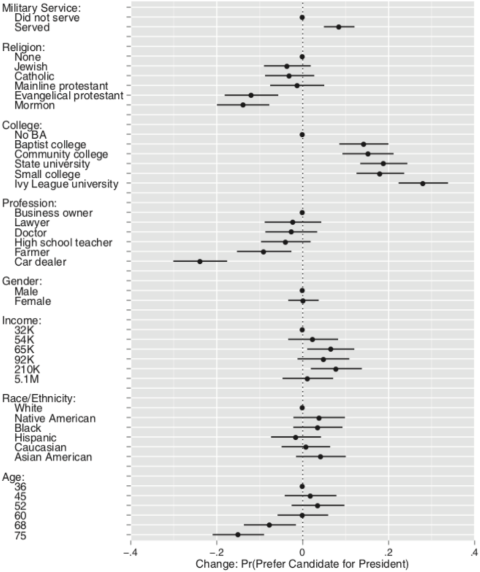
```r
cand |>
  filter(atprof == "High school teacher") |>
  count(atinc)
```

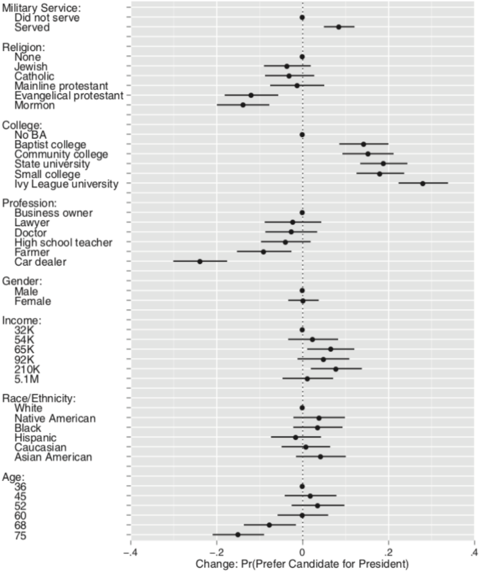
```
##   atinc   n
## 1   32K 107
## 2   54K  77
## 3   65K 111
## 4   92K  97
## 5  210K  99
## 6  5.1M 100
```

Strange things can happen ... (“atypical profiles”)

```r
cand |> filter(atprof == "High school teacher",
               atinc == "5.1M",
               ated == "No BA") |>
  dim()
```

```
## [1] 16 11
```


Military Service:
Did not serve
Served
Religion:
None
Jewish
Catholic
Mainline protestant
Evangelical protestant
Mormon
College:
No BA
Baptist college
Community college
State university
Small college
Ivy League university
Profession:
Business owner
Lawyer
Doctor
High school teacher
Farmer
Car dealer
Gender:
Male
Female
Income:
32K
54K
65K
92K
210K
5.1M
Race/Ethnicity:
White
Native American
Black
Hispanic
Caucasian
Asian American
Age:
36
45
52
60
68
75
-.4
-.2
0
.2
.4
Change: Pr(Prefer Candidate for President)


Military Service:
Did not serve
Served
Religion:
None
Jewish
Catholic
Mainline protestant
Evangelical protestant
Mormon
College:
No BA
Baptist college
Community college
State university
Small college
Ivy League university
Profession:
Business owner
Lawyer
Doctor
High school teacher
Farmer
Car dealer
Gender:
Male
Female
Income:
32K
54K
65K
92K
210K
5.1M
Race/Ethnicity:
White
Native American
Black
Hispanic
Caucasian
Asian American
Age:
36
45
52
60
68
75
−.4
−.2
0
.2
.4
Change: Pr(Prefer Candidate for President)

Atypical Profiles

- "groups of respondents exposed to different numbers of atypical profiles"

Atypical Profiles

- “groups of respondents exposed to different numbers of atypical profiles”
- (A reasonable conditional-effect diagnostic)

Atypical Profiles

- "groups of respondents exposed to different numbers of atypical profiles"
- (A reasonable conditional-effect diagnostic)
- Alternative: how atypicality affects AMCEs/ACIEs

Atypical Profiles

- "groups of respondents exposed to different numbers of atypical profiles"
- (A reasonable conditional-effect diagnostic)
- Alternative: how atypicality affects AMCEs/ACIEs
- E.g., if *high inc* usually appears in atypical profiles, then AMCE of *high inc* vs. *low inc* not internally valid (?)

Conjoint Interpretation

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

“prefer female to male candidates”.

But – AMCE does **not** give …

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give …

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"
- "`female` candidates tend to beat `male` candidates"

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"
- "`female` candidates tend to beat `male` candidates"

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"
- "`female` candidates tend to beat `male` candidates"

Why? AMCE averages over *intensity* of prefs

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"
- "`female` candidates tend to beat `male` candidates"

Why? AMCE averages over *intensity* of prefs

Some who *strongly* prefer `female` to `male` can create AMCE > 0, even though *minority*.

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., `male` to `female` $\rightsquigarrow$ positive AMCE in Teele, Kalla, and Rosenbluth (2018):

"prefer female to male candidates".

But – AMCE does **not** give ...

- "majority prefer `female` to `male`"
- "median voter prefers `female` to `male`"
- "`female` candidates tend to beat `male` candidates"

Why? AMCE averages over *intensity* of prefs

Some who *strongly* prefer `female` to `male` can create AMCE > 0, even though *minority*.

(Abramson, Koçak, and Magazinnik 2022)

Interpreting the AMCE

AMCE is avg (mean) effect of varying attribute l in profile.

E.g., from l_0 to l_1.

But – AMCE does **not** give …

- "majority prefer l_1 to l_0"
- "median voter prefers l_1 to l_0"
- "l_1 candidates tend to beat l_0 candidates"

Why? AMCE avgs over *intensity* of prefs

Some who *strongly* prefer l_1 to l_0 can create AMCE > 0, even though *minority*.

(Abramson, Koçak, and Magazinnik 2022)

What can we learn from the AMCE?

- AMCE is average of both *intensity* and *direction* of preferences

What can we learn from the AMCE?

- AMCE is average of both *intensity* and *direction* of preferences
- "Average of voter ideal points"

What can we learn from the AMCE?

- AMCE is average of both *intensity* and *direction* of preferences
- "Average of voter ideal points"
- Greater

$$cor(\text{intensity}, \text{direction})$$

  more misleading AMCE is.

What can we learn from the AMCE?

- AMCE is average of both *intensity* and *direction* of preferences
- "Average of voter ideal points"
- Greater

  $$cor(\text{intensity}, \text{direction})$$

  more misleading AMCE is.
- If voters with *strong* prefs all prefer same *direction*, then AMCE misleading. (E.g., if strong gender pref is always for `female`, AMCE misleading.)

V1	V2	V3	V4	V5
$M \succ F$	$M \succ F$	$M \succ F$	$F \succ M$	$F \succ M$
$R \succ D$	$R \succ D$	$R \succ D$	$D \succ R$	$D \succ R$

Table 1—: Preferences over attributes

Rank	V1	V2	V3	V4	V5
1.	MR	MR	MR	FD	FD
2.	FR	FR	FR	FR	FR
3.	MD	MD	MD	MD	MD
4.	FD	FD	FD	MR	MR

Table 2—: Preferences over candidate profiles

Abramson, Koçak, and Magazinnik (2022)

Comparison	V1	V2	V3	V4	V5	Tally
MR,FR	MR	MR	MR	FR	FR	3, 2
MR,FD	MR	MR	MR	FD	FD	3, 2
MR,MD	MR	MR	MR	MD	MD	3, 2
MD,**FR**	FR	FR	FR	FR	FR	0, 5
MD,FD	MD	MD	MD	FD	FD	3, 2
FR,FD	FR	FR	FR	FD	FD	3, 2

Table 3—: Aggregate preferences over candidate profiles

Abramson, Koçak, and Magazinnik (2022)

Comparison	V1	V2	V3	V4	V5	Tally
MR,FR	MR	MR	MR	FR	FR	3, 2
MR,FD	MR	MR	MR	FD	FD	3, 2
MR,MD	MR	MR	MR	MD	MD	3, 2
MD,**FR**	FR	FR	FR	FR	FR	0, 5
MD,FD	MD	MD	MD	FD	FD	3, 2
FR,FD	FR	FR	FR	FD	FD	3, 2

Table 3—: Aggregate preferences over candidate profiles

But, AMCE for `M` is negative!

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences
- AMCE is proportional to difference in Borda scores (number of candidates i prefers candidate j to)

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences
- AMCE is proportional to difference in Borda scores (number of candidates i prefers candidate j to)
- Exclusion of atypical profiles can alter AMCE, too.

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences
- AMCE is proportional to difference in Borda scores (number of candidates i prefers candidate j to)
- Exclusion of atypical profiles can alter AMCE, too.
- Need homogeneity of pref *intensity* for majoritarian interpretations

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences
- AMCE is proportional to difference in Borda scores (number of candidates i prefers candidate j to)
- Exclusion of atypical profiles can alter AMCE, too.
- Need homogeneity of pref *intensity* for majoritarian interpretations
- Suggest small number of binary attributes …

- Fundamentally, AMCE cannot disentangle *intensity* and *direction* of preferences
- AMCE is proportional to difference in Borda scores (number of candidates i prefers candidate j to)
- Exclusion of atypical profiles can alter AMCE, too.
- Need homogeneity of pref *intensity* for majoritarian interpretations
- Suggest small number of binary attributes …
    - but still have IIA concerns

Bansak et al. (2022)

- Incorporating direction *and* intensity is realistic (*ceteris paribus* is not)

Bansak et al. (2022)

- Incorporating direction *and* intensity is realistic (*ceteris paribus* is not)
- AMCE is an effect on candidate/party's *expected vote share*

Bansak et al. (2022)

- Incorporating direction *and* intensity is realistic (*ceteris paribus* is not)
- AMCE is an effect on candidate/party's *expected vote share*
- Can translate this into effect on probability of winning (given voting system)

Another Critique of the AMCE

Ganter (2023) argues for the *average component preference*:

- AMCE suited for "selection-process" questions: how would outcome obtain?
  - How likely is a male immigrant to get a visa?
  - Are female immigrants more or less likely to get a visa than male immigrants?
  - Comparing (A v. B) to (A' v. B)
- ACP suited for "preference-related" questions: which is preferred?
  - Do people prefer male or female immigrants?
  - Is gender more determinant than countries of origin in people's choices?
  - How do these preferences differ across subgroups?
  - Comparing pref for A v. A'

What should we do tomorrow?

What should we do tomorrow?

Let's vote!

What should we do tomorrow?

Let's vote!

PollEv.com/rmoore952

Select ≤ 3 topics to inform our discussion for tomorrow! (You may "Skip" registration.)

Next:
Multiarm Bandits

Abramson, Scott F, Korhan Koçak, and Asya Magazinnik. 2022. "What Do We Learn about Voter Preferences from Conjoint Experiments?" *American Journal of Political Science* 66 (4): 1008–20.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2018. "The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments." *Political Analysis* 26: 112–19.

———. 2022. "Using Conjoint Experiments to Analyze Election Outcomes: The Essential Role of the Average Marginal Component Effect." *Political Analysis*, 1–19. https://doi.org/10.1017/pan.2022.16.

Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins, and Teppei Yamamoto. 2021. "Beyond the Breaking Point? Survey Satisficing in Conjoint Experiments." *Political Science Research and Methods* 9 (1): 53–71.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20 (1): 47–77.

Ganter, Flavien. 2023. "Identification of Preferences in Forced-Choice Conjoint Experiments: Reassessing the Quantity of Interest." *Political Analysis* 31 (1): 98–112. https://doi.org/10.1017/pan.2021.41.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.

Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth. 2018. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112 (3): 525–41.