Models All the Way Down

Presented at the MZES Social Science Data Lab Colloquium

Asya Magazinnik (Hertie School)
with Scott Abramson (University of Rochester) and Korhan Kocak (NYUAD)

February 21, 2024

Goals of the Project

We want to place some theoretical scaffolding under commonly used statistical methodologies for measuring (political) preferences.

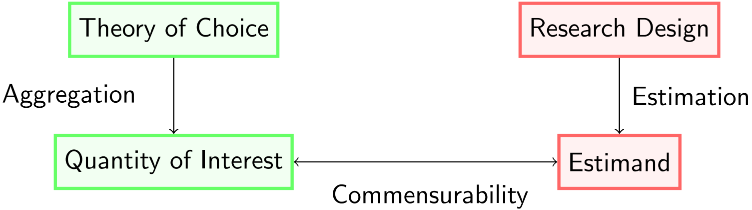
- By political, I really mean multidimensional

Key Takeaway: Causally identified quantities of interest $\neq$ statements about preferences. Today, I will:

1. Discuss what we might mean by preferences
2. Illustrate why they may not align with an intuitive causal quantity
3. Propose a research design that more tightly connects the two

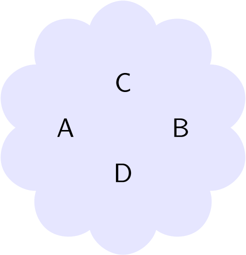
Framework

See: *Scott Ashworth, Christopher R. Berry, and Ethan Bueno de Mesquita. 2021. Theory and Credibility. Princeton, NJ: Princeton University Press.*



1. What Is a Preference?

What is a Preference?



For any two members of the set, we can define a preference relation:

$$A \succ B \textit{ or}$$

$$B \succ A \textit{ or}$$

$$A \sim B$$

What is a Preference?

Donald Trump Joe Biden

What is a Preference?

Tax Plan A Tax Plan B

We Seek to Generalize

As social scientists, we are not usually interested in preferences over *specific* candidates or tax plans so much as preferences over *general attributes*:

- Is there a systematic bias against women or racial minorities?
- Do voters seek *descriptive representation*, i.e. do they want representatives in office who are somehow like themselves?
- How do voters trade off different characteristics of candidates, e.g. policy positions vs. personal attributes?

These questions follow from the multidimensionality of real-world political choices.

Preferences over Attributes

Donald Trump

$$\begin{bmatrix} \textbf{Age} = 77 \\ \textbf{Race} = \textit{White} \\ \textbf{Party} = \textit{Republican} \\ \textbf{Gender} = \textit{Male} \\ \vdots \end{bmatrix}$$

How do we make general statements about a voter's preferences over these attributes?

1. We can define primitives over candidates and summarize
2. We can define primitives over attributes and build a bridge to candidates

1. Primitives over Candidates

$$
\begin{array}{cc}
\text{Donald Trump} & \text{Nikki Haley} \\
\begin{bmatrix} \textbf{Age} = 77 \\ \textbf{Race} = \textit{White} \\ \textbf{Party} = \textit{Republican} \\ \textbf{Gender} = \textit{Male} \\ \vdots \end{bmatrix} &
\begin{bmatrix} \textbf{Age} = 52 \\ \textbf{Race} = \textit{White} \\ \textbf{Party} = \textit{Republican} \\ \textbf{Gender} = \textit{Female} \\ \vdots \end{bmatrix}
\end{array}
$$

$$\text{Donald Trump} \succ \text{Nikki Haley}$$

1. Primitives over Candidates

$$
\begin{array}{cc}
\text{Donald Trump} & \text{Alternative Candidate} \\
\begin{bmatrix} \textbf{Age} = 77 \\ \textbf{Race} = \textit{White} \\ \textbf{Party} = \textit{Republican} \\ \textbf{Gender} = \textit{Male} \\ \vdots \end{bmatrix} & \begin{bmatrix} \textbf{Age} = 52 \\ \textbf{Race} = \textit{Hispanic} \\ \textbf{Party} = \textit{Republican} \\ \textbf{Gender} = \textit{Female} \\ \vdots \end{bmatrix}
\end{array}
$$

$$\text{Donald Trump} \succ \text{Alternative Candidate}$$

1. Primitives over Candidates

$$\begin{array}{cc} \text{Candidate 1} & \text{Candidate 2} \\ \begin{bmatrix} \textbf{Age} = x_1 \\ \textbf{Race} = y_1 \\ \textbf{Party} = z_1 \\ \textbf{Gender} = \textit{Male} \\ \vdots \end{bmatrix} & \begin{bmatrix} \textbf{Age} = x_2 \\ \textbf{Race} = y_2 \\ \textbf{Party} = z_2 \\ \textbf{Gender} = \textit{Female} \\ \vdots \end{bmatrix} \end{array}$$

A preference for male candidates may be understood as the proportion of binary comparisons in which men are selected over women, taken over all possible values of the other attributes, minus one-half.

- We call this the Feature Choice Probability (FCP) (Abramson, Kocak, Magazinnik and Strezhnev 2024).
- Like the marginal component effect (MCE), this quantity is sensitive to the distributions of the other attributes (de la Cuesta et al. 2022).

2. Primitives over Attributes

1. The voter has rank-orderings of preferred values for every attribute:

$$\begin{bmatrix} \textbf{Gender} \\ 1.\textit{Female} \\ 2.\textit{Male} \end{bmatrix} \begin{bmatrix} \textbf{Party} \\ 1.\textit{Democrat} \\ 2.\textit{Republican} \end{bmatrix} \begin{bmatrix} \textbf{Age} \\ 1.\textit{Middle-aged} \\ 2.\textit{Young} \\ 3.\textit{Old} \end{bmatrix}$$

as well as a rank-ordering that prioritizes attributes:

$$Gender, Party, Age$$

which allows this voter to evaluate any two candidates and form a preference (Abramson, Kocak and Magazinnik 2022).

2. The voter has a utility function of the form:

$$\begin{aligned} U(Candidate) = \; & u_0 + w_{gender} * u(female) * I(\text{Candidate is female}) \; + \\ & w_{party} * u(Democrat) * I(\text{Candidate is Democrat}) + ... \end{aligned}$$

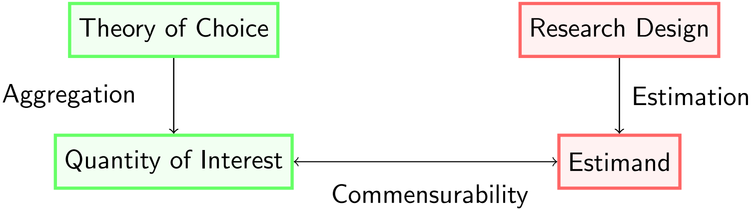
Quantities of Interest

With this framework in place, we can come up with theoretically meaningful aggregate quantities of interest:

- If primitives over candidates:
  - The average feature choice probability
  - The all-else-equal analog of the average feature choice probability
- If primitives over attributes:
  - The proportion of the sample that prefers male to female candidates
  - The average utility difference for male vs. female candidates

2. Pathologies of Causal Inference (or: Commensurability)

Framework


Theory of Choice
Research Design
Aggregation
Estimation
Quantity of Interest
Estimand
Commensurability

Example: The Audit Study

Research Question: Is there a bias against female job applicants in the labor market?

Research Design:

- Construct identical resumes except for the gender of the applicant and randomly send the female version to 50% of employers and the male version to 50% of employers
- **Estimand:** Difference in proportion of sent resumes that get a callback between employers in the male condition and employers in the female condition

(see, e.g., Bertrand and Mullainathan 2003 and many others)

Stephen Nichols

nichols.stephenw@gmail.com

Education: Wesleyan University, B.A. in Computer Science, 2017 G.P.A. 3.6

Languages:	C++, C#, Java, Python, Ruby, JavaScript, TypeScript, Go
Font-End:	React, Unity, Node, Swift, PHP, HTML/CSS
Cloud Computing:	Google Cloud Platform, AWS
Infrastructure/Scalability:	Docker, Kubernetes, MongoDB, MySQL, jQuery

EXPERIENCE

Blue Slate Solutions Albany, NY
Software Engineer 2018-Present

- Built a platform to analyze 50K+ Skype call recordings daily using Spark
- Led a 5-person team to create an interactive KPI platform to build & share dashboards
- Developed a cloud app to review performance of IT services daily
- Trained new hires in agile methodology, documentation protocols, & Git procedures
- Built RESTful API endpoints and tests, using Java and SAP HANA SQL database

Astro Technology (acquired by Slack) Palo Alto, CA
Software Development Engineer 2017-2018

- Integrated customer metrics data flow and improved data visualization protocols to give product and marketing teams weekly consumer experience summaries
- Started "Bug Busters," led bi-weekly code review sessions, fixing 100s of bugs
- Deployed TensorFlow jobs to convert Kafka streams of email & voice messages to dynamic calendar events

eBay San Jose, CA
Software Engineer Intern June-August, 2016

- Created a scoring system for advertising unit to evaluate advertiser goals
- Collaborated with data scientists to improve insight reports for 1M+ clients

PROJECTS

Stephanie Nichols

nichols.stephenw@gmail.com

Education: Wesleyan University, B.A. in Computer Science, 2017 G.P.A. 3.6

Languages:	C++, C#, Java, Python, Ruby, JavaScript, TypeScript, Go
Font-End:	React, Unity, Node, Swift, PHP, HTML/CSS
Cloud Computing:	Google Cloud Platform, AWS
Infrastructure/Scalability:	Docker, Kubernetes, MongoDB, MySQL, jQuery

EXPERIENCE

Blue Slate Solutions Albany, NY
Software Engineer 2018-Present

- Built a platform to analyze 50K+ Skype call recordings daily using Spark
- Led a 5-person team to create an interactive KPI platform to build & share dashboards
- Developed a cloud app to review performance of IT services daily
- Trained new hires in agile methodology, documentation protocols, & Git procedures
- Built RESTful API endpoints and tests, using Java and SAP HANA SQL database

Astro Technology (acquired by Slack) Palo Alto, CA
Software Development Engineer 2017-2018

- Integrated customer metrics data flow and improved data visualization protocols to give product and marketing teams weekly consumer experience summaries
- Started "Bug Busters," led bi-weekly code review sessions, fixing 100s of bugs
- Deployed TensorFlow jobs to convert Kafka streams of email & voice messages to dynamic calendar events

eBay San Jose, CA
Software Engineer Intern June-August, 2016

- Created a scoring system for advertising unit to evaluate advertiser goals
- Collaborated with data scientists to improve insight reports for 1M+ clients

PROJECTS

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$\frac{1}{3}$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$\frac{1}{3} + 0$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$\frac{1}{3} + 0 + \frac{1}{3}$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$\frac{1}{3} + 0 + \frac{1}{3} = \frac{2}{3}$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$\frac{1}{3} + 0 + \frac{1}{3} = \frac{2}{3}$$

And what is the causal estimate from our experiment?

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$1/3 + 0 + 1/3 = 2/3$$

And what is the causal effect from our experiment?

$$\text{ATE} = Pr(\text{Call}|\text{Male condition}) - Pr(\text{Call}|\text{Female condition})$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$1/3 + 0 + 1/3 = 2/3$$

And what is the causal effect from our experiment?

$$\begin{aligned} \text{ATE} &= Pr(\text{Call}|\text{Male condition}) - Pr(\text{Call}|\text{Female condition}) \\ &= 1/3 \end{aligned}$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$1/3 + 0 + 1/3 = 2/3$$

And what is the causal effect from our experiment?

$$\begin{aligned}\text{ATE} &= Pr(\text{Call}|\text{Male condition}) - Pr(\text{Call}|\text{Female condition}) \\ &= 1/3 - 2/3\end{aligned}$$

Example: The Audit Study

Rank	Type 1	Type 2	Type 3
1	Male Candidate	Female Candidate	Outside Option
2	Female Candidate	Outside Option	Male Candidate
3	Outside Option	Male Candidate	Female Candidate

Table 1: Preference orderings for three employer types, each constituting 1/3 of the sample.

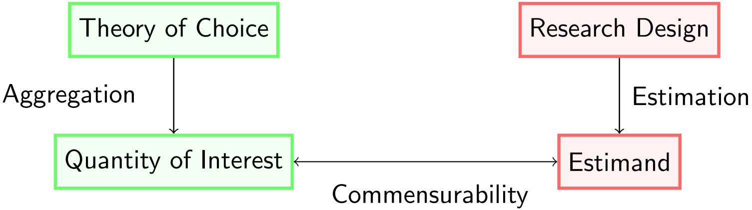
What proportion of employers prefer our experimentally generated male candidates over the female candidates?

$$1/3 + 0 + 1/3 = 2/3$$

And what is the causal effect from our experiment?

$$\begin{aligned} \text{ATE} &= Pr(\text{Call}|\text{Male condition}) - Pr(\text{Call}|\text{Female condition}) \\ &= 1/3 - 2/3 = -1/3 \end{aligned}$$

What we have here is a commensurability problem



Wait, why did we estimate this causal quantity?

We seek causal quantities to get around the fundamental problem of causal inference: we can't simultaneously observe two states of the world, but we wish to know how outcomes differ between those two worlds:

$$Y_i(T_i = 1) - Y_i(T_i = 0)$$

And what about the audit study?

- Preferences are fixed
  - They are not directly manipulated by the researcher
  - They are not affected by any manipulation
- What is manipulated? The choice set.
- Why is this manipulated?
  - Not the usual "fundamental problem" – social desirability bias, feasibility

From Causal Effects to Statements About Preferences

In a standard randomized control trial, the individual-level quantity of interest is:

$$Y_i(T_i = 1) - Y_i(T_i = 0)$$

Whereas in this audit study, it is:

$$Y_i(Preferences_i, T_i = 1) - Y_i(Preferences_i, T_i = 0)$$

and we are trying to make descriptive statements about *Preferences*!

The causal effect of your manipulation on the observed choice is not a direct measure of the fixed, latent object we call preferences. You have to build a bridge to them.

- Measuring preferences is not inherently a causal exercise. It is a descriptive exercise for which causal inference may sometimes be useful.

So where do we go from here?

1. Know your model
    - What are your primitives?
    - Mapping between attributes and candidates
    - Is the causal effect of first-order interest?
    - More structure $\rightarrow$ more meaningful estimates
2. Think more carefully about aggregation
    - What is the ideal quantity *in asymptopia*?
    - Import insights from social choice theory
3. If possible, develop research designs to directly target theoretical parameters of interest (EITM)
    - But note estimation-interpretability trade-offs
4. Under what assumptions can existing experimental designs recover theoretically meaningful quantities? (TIEM)

3. Example: Directly Targeting Parameters of Interest

Model Setup

- Unidimensional policy space, voter i's ideal point is $x_i \in \mathbb{R}$
- Let $d_{ij} := |x_i - x_j|$ denote ideological distance between voter i and candidate j
- Candidate j also has personal attributes $v_j \in \mathbb{R}^\tau$ (e.g. gender, race, party)
- Voter i assigns weights r_i to ideology and k_i to personal attributes
- Unobserved error term ε_{ij}
- Then, voter i's payoff from having candidate j in office is given by:

$$u_{ij} = -d_{ij} r_i + v_j k_i + \varepsilon_{ij}$$

Instrumental vs. Expressive Concerns

Instrumental: by voting for j, voter i increases the probability j wins

But voters also derive expressive utility from voting for a candidate they like:

- Warm glow,
- Signaling.

We can build in instrumental (r_i, k_i) and expressive (ρ_i, κ_i) concerns over ideology and personal attributes:

$$u_{ij} = -d_{ij}(r_i + \rho_i) + v_j(k_i + \kappa_i) + \varepsilon_{ij}$$

Instrumental and Expressive Concerns are Most Distinguishable in Primary Elections

General elections: both concerns require voting for preferred candidate.

Primary elections: might be optimal to vote for a less favored candidate:

- e.g. Top choice has a lower chance of winning in general election.

Let q_{jc} denote probability primary candidate j wins general election against the other party's challenger, c.

Voter i's payoff from voting for primary candidate j is given by:

$$u_{ij} = \underbrace{-d_{ij}(q_{jc}r_i + \rho_i)}_{\text{policy term, } j} + \underbrace{v_j(q_{jc}k_i + \kappa_i)}_{\text{personal attributes, } j} + \underbrace{(1 - q_{jc})(-d_{ic}r_i + v_c k_i)}_{\text{challenger, } c} + \varepsilon_{ij}$$

Estimating Equation

Assume $\varepsilon_{ij} \sim$ Type I Extreme Value.

Then, the log odds ratio for voter i's probability of voting for some primary candidate j over another primary candidate j' is:

$$\ln\left(\frac{p_{ij}}{p_{ij'}}\right) = -r_i\left(q_{jc}(d_{ij} - d_{ic}) - q_{j'c}(d_{ij'} - d_{ic})\right) - \rho_i(d_{ij} - d_{ij'}) + \\ b_i(q_{jc} - q_{j'c}) + k_i\left(q_{jc}(v_j - v_c) - q_{j'c}(v_{j'} - v_c)\right) + \kappa_i(v_j - v_{j'}).$$

Estimating Equation

Assuming $\epsilon_{ij} \sim$ Type I Extreme Value.

Then, the log odds ratio for voter i's probability of voting for j over j' is:

$$\ln\left(\frac{p_{ij}}{p_{ij'}}\right) = -r_i\left(q_{jc}(d_{ij} - d_{ic}) - q_{j'c}(d_{ij'} - d_{ic})\right) - \rho_i(d_{ij} - d_{ij'}) +$$
$$b_i(q_{jc} - q_{j'c}) + k_i\left(q_{jc}(v_j - v_c) - q_{j'c}(v_{j'} - v_c)\right) + \kappa_i(v_j - v_{j'}).$$

Now, we can create an intentional randomized survey design where:

- We observe p, q, d, and v
    - We randomize the primary candidates' electability, ideology, and personal attributes
    - And we ask respondents for their own ideological positions and their candidate choice

Estimating Equation

Assuming $\epsilon_{ij} \sim$ Type I Extreme Value.

Then, the log odds ratio for voter i's probability of voting for j over j' is:

$$\ln\left(\frac{p_{ij}}{p_{ij'}}\right) = -r_i\left(q_{jc}(d_{ij} - d_{ic}) - q_{j'c}(d_{ij'} - d_{ic})\right) - \rho_i(d_{ij} - d_{ij'}) + b_i(q_{jc} - q_{j'c}) + k_i\left(q_{jc}(v_j - v_c) - q_{j'c}(v_{j'} - v_c)\right) + \kappa_i(v_j - v_{j'}).$$

Now, we can create an intentional randomized survey design where:

- We observe p, q, d, and v
    - We randomize the primary candidates' electability, ideology, and personal attributes
    - And we ask respondents for their own ideological positions and their candidate choice
- and we estimate r, ρ, b, k, and κ (via OLS).

Quantities of Interest

Willingness-to-pay for personal attribute v_j in terms of ideology:

$$WTP_i^d(v_j) = \frac{q_{jc}k_i + \kappa_i}{q_{jc}r_i + \rho_i}.$$

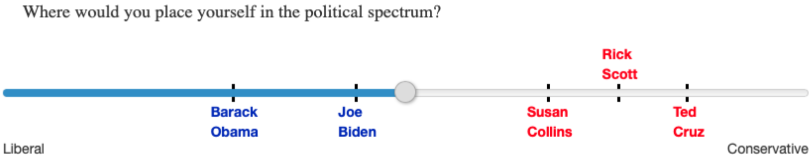
Willingness-to-pay for personal attribute v_j in terms of electability:

$$WTP_i^q(v_j) = \frac{q_{jc}k_i + \kappa_i}{r_i(d_{ij} - d_{ic}) - (k_i + b_i)}.$$

We estimate these quantities and calculate measures of uncertainty using block bootstrap.

Experimental Design

Respondents place themselves on an ideological spectrum:



This gives us our x_i for every survey respondent.

Experimental Design

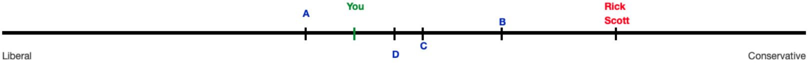
We randomize q_{jc} (probability of beating c), and three personal attributes: gender, race, and political experience. Respondents give us p_{ij} for every randomized candidate.

For each candidate below, you are given estimates of their ideology on the political spectrum, the probability they will win in the general election, their gender, race, and their experience in politics. Please indicate your probability of voting for each of the candidates.

Scenario 1

	A	B	C	D	Rick Scott
Probability of beating Rick Scott	57%	32%	51%	35%	NA
Gender	Female	Male	Female	Male	Male
Race	White	White	White	Black	White
Years in politics	5	29	15	11	10

This is how your ideology compares with the candidates:



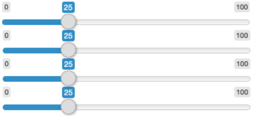
Question 1: What is the probability you would vote for each Democratic primary candidate?

Candidate A

Candidate B

Candidate C

Candidate D



‹ Previous › Next

Results: WTP in Ideology for Candidate Features

	Median	1st quartile	3rd quartile	Prop. prefer
Female vs. male	0.02	-0.14	0.18	0.54
	[0.01, 0.02]	[-0.16, -0.12]	[0.16, 0.20]	[0.52, 0.56]
Black vs. white	0.01	-0.25	0.26	0.51
	[-0.00, 0.02]	[-0.29, -0.22]	[0.22, 0.30]	[0.49, 0.53]
Hispanic vs. white	0.01	-0.22	0.22	0.52
	[-0.00, 0.02]	[-0.25, -0.18]	[0.20, 0.24]	[0.50, 0.54]
Asian vs. white	0.02	-0.29	0.32	0.48
	[0.00, 0.03]	[-0.33, -0.25]	[0.29, 0.36]	[0.46, 0.50]
Experience	-0.00	-0.01	0.01	0.50
	[-0.00, 0.00]	[-0.01, -0.01]	[0.01, 0.01]	[0.48, 0.51]

Results: Preference Intensity for Candidate Features in Ideology

	Outcome: $|WTP\ Ideology|$		
	Median	1st quartile	3rd quartile
Female vs. male	0.162	0.065	0.413
	[0.149,0.177]	[0.058,0.071]	[0.376,0.451]
Black vs. white	0.254	0.087	0.692
	[0.227,0.283]	[0.08,0.098]	[0.628,0.738]
Latinx vs. white	0.221	0.083	0.567
	[0.208,0.234]	[0.074,0.09]	[0.519,0.631]
Asian vs. white	0.306	0.109	0.874
	[0.286,0.338]	[0.097,0.12]	[0.779,0.95]
Experience	0.01	0.004	0.026
	[0.009,0.011]	[0.003,0.004]	[0.023,0.028]

In Summary

Beginning with an explicit model of how preferences translate into choices lets us design bespoke experiments to structurally estimate the theoretical parameters you actually care about.

- The more assumptions you make on the way in, the more you get on the way out. But obviously you have to believe your model.
- When we don't assume any model on the way in, we often end up assuming one on the way out — whether we realize it or not (Abramson, Kocak, and Magazinnik 2022).
- We encourage researchers to tailor the model to their electoral context: political setting, voters' incentives, system of voting & aggregating votes (beware pathologies of aggregation — a talk for another time).