Are Preconceptions Postconceptions? Evidence on Motivated Political Reasoning

Matthew Lilley & Brian Wheaton

August 6, 2021

Question

This project addresses two questions:

- To what degree are people amenable to changing their minds on contentious political issues?
- How much do individuals change their beliefs on contentious political issues when presented with relevant information?
    - Are people Bayesian?
    - Do people update asymmetrically in response to *favorable* versus *unfavorable* news?

Information Experiments

A broad literature (cf. Roth (2020)) exists featuring experiments of the following form:

- A control group has their beliefs about some issue elicited
- A treatment group is provided with pertinent information, followed by the same belief elicitation
- Implication: being provided with information $\mathcal{I}$ causes people to change their beliefs by Δy

A second literature compares the extent of belief updating in response to information to some normative benchmark of rationality.

Information Distrust

Implicit underlying question: But how much *should* individuals change their beliefs by when provided with new information?

- How much *should* people trust the information provided by the experimenter?
- How does distrust of information cause belief change to be attenuated?

Accordingly, it is extremely difficult to make definitive claims on how much it is reasonable to discount information

- This complicates attempts to measure whether people are engaging in motivated reasoning or other similar non-Bayesian behavior

Approach and Context

We attempt to study belief updating and motivated reasoning using experiments that are robust to classical forms of information distrust.

Specifically, we perform two experiments that we undertake with participants on Amazon *Mechanical Turk*.

The experiments that we present today involve beliefs and information provision about racial bias in policing.

- This is a salient and contentious political issue around which people may have substantial identity utility
- Beliefs on the issue are substantially based on underlying empirical claims, rather than merely differences in values
- I.e. "all else equal, police disproportionately kill blacks relative to whites" is a testable claim, upon which beliefs about police racism are plausibly based.

Overview

- Conceptual framework
- Experiment One - Ex Post vs. Ex Ante Hypothetical Beliefs
  - Setup
  - Empirical strategy
  - Results
- Experiment Two - Noisy Information Provision
  - Setup
  - Empirical strategy
  - Results

Conceptual Framework

Consider models of ego or identity utility where beliefs affect utility directly (ego or identity utility) and instrumentally (optimal decision making)

- E.g. Akerlof & Kranton (2010), Brunnermeier & Parker (2005), Mobius et al. (2014)

Let individuals hold beliefs $\theta \in \mathcal{R}$ about some political issue.

- Given information set $\mathcal{I}$, rational beliefs are $\theta_{\mathcal{I}}^{\mathsf{RE}}$
- Identity bliss point of θ^{ID}

Individual utility is governed by $U(c(\theta), \theta)$ where $c(\cdot)$ is decreasing in the distance between rational beliefs (or truth) and chosen beliefs

Canonical result: distortion of beliefs θ away from θ^{RE} towards θ^{ID}

- $\Delta\theta$ produces first order gain in identity utility and second order instrumental loss.

Experiment One: Ex Post vs Ex Ante Hypothetical Beliefs

Consider providing information about a variable X which is realised as x^*.

Baseline information experiment

- Information treatment: Provide people with information $X = x^*$, and then observe beliefs on related outcome measures.
- Difference in elicited beliefs between control and treatment group measures *actual* effect of providing information $X = x^*$.

To test for motivated reasoning, need a benchmark to compare ex post beliefs against.

- What *should* people answer upon learning $X = x^*$?
- **Our answer**: Ask people *hypothetically* what their beliefs would be *if* they learned that $X = x$ (for various x including x^*).

Benchmark Argument

If a rational person receives no other information, the conditional beliefs they express should be the same ex ante and ex post.

Of course, for a given individual j, we only observe either ex ante conditional or ex post conditional beliefs.

Since people are randomised into treatment arms (C, I, H), rationality implies the average elicited responses in the two arms be equal

$$E_H(\theta_j|X = x^*) = E_I(\theta_j|X = x^*) \tag{1}$$

This also holds for subgroups defined by characteristics (e.g. political party)

The variation in $E_H(\theta_j|X = x)$ as x varies reveals how much people *claim* they are amenable to changing their views in response to empirical facts

Experimental Procedure

As people with different preconceptions may update differently (and have different *favorable* news), we first elicit people's priors

- Explain that *Washington Post* collects data on officer-involved shootings in US
- Provide basic information and ask people to estimate number of white and black people fatally shot by police as measured by this database.

Police Shootings

In 2015, the Washington Post began to log every fatal shooting by an on-duty police officer in the United States. For the following questions, we will treat their data as the truth. Of course, their data may be imperfect, so to the extent you think it is flawed, you should take this into account in answering the following questions.

For context:

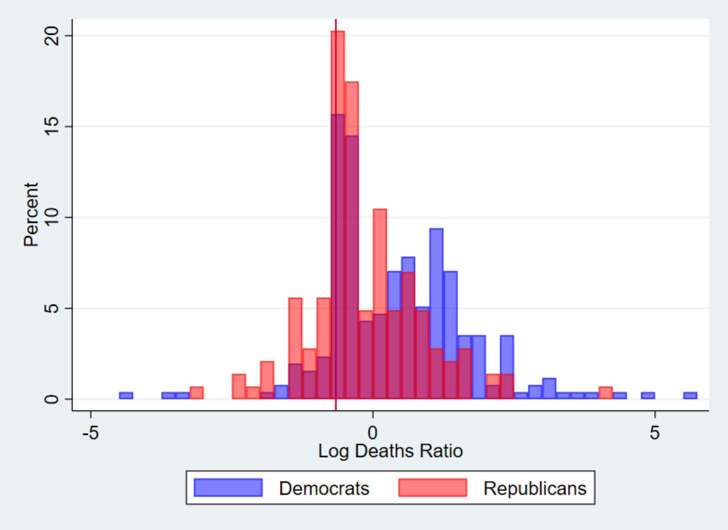
- As of 2018, the US population was approximately 72.0% White and 12.8% Black (5.6 White people for each Black person)
- In 2018, according to the FBI, among people arrested for violent crimes, approximately 58.7% were White and 37.4% were Black (1.57 White people for each Black person).

* Note that the these data include Hispanics (most of whom are White); however, the Washington Post shooting data separates Hispanics into their own category.

What is your best estimate of the number of Black people fatally shot by police since 2015?

What is your best estimate of the number of White people fatally shot by police since 2015?

Police Shootings



Experimental Procedure

Information treatment: Provide people with the true white/black ratio of police shootings, to compare against their previous guess.

Outcome: Elicit beliefs about whether police are racist (for individuals in either control or treatment arm)

- Difference in elicited beliefs about police racism between control and treatment group measures *actual* effect of providing information.

Hypothetical arm: Ask people *hypothetically* if they *would* think police are racist *if* they learned the white/black ratio of police shootings was x

- For varying values of x, including x very close to the truth.
- For hypothetical $x \approx x^*$, difference between elicited hypothetical beliefs and ex post beliefs measures degree of motivated reasoning.
- Change in hypothetical beliefs as x changes measures claimed willingness to revise beliefs about police racism.

Hypothetical Benchmark

Previously you were asked your best guess of the number of people in different groups who were fatally shot by police since 2015. You estimated that there were 0.5 times as many White (as Black) people fatally shot by police. (Alternatively stated, 2 times as many Black (as White).)

Suppose that you found out that the data collected by the Washington Post differed from your estimate. This might cause you to change you beliefs about whether police are systemically racist.

For context:

- As of 2018, the US population was approximately 72.0% White and 12.8% Black (5.6 White people for each Black person)
- In 2018, according to the FBI, among people arrested for violent crimes, approximately 58.7% were White and 37.4% were Black (1.57 White people for each Black person).

* Note that the these data include Hispanics (most of whom are White); however, the Washington Post shooting data separates Hispanics into their own category.

Would you believe that police in the US are systematically racist against black people if you discovered that, according to the Washington Post, **2 times** as many **White** (as Black) people were fatally shot by police?

Definitely yes

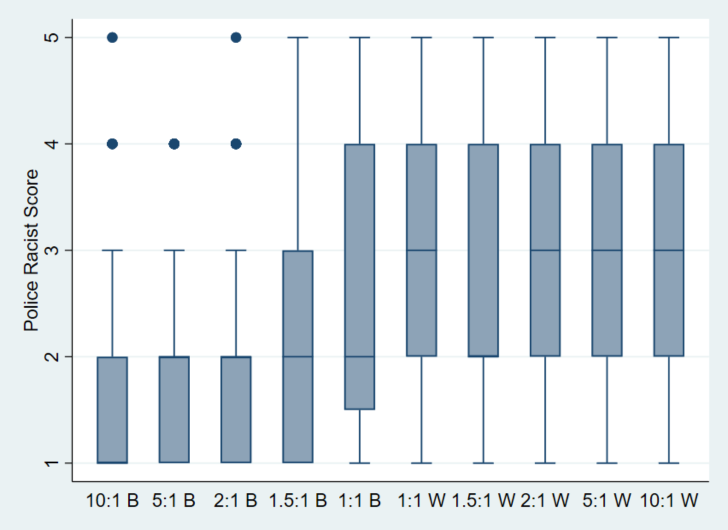
Probably yes

Might or might not

Probably not

Definitely not

Results


Police Racist Score
1
2
3
4
5
10:1 B
5:1 B
2:1 B
1.5:1 B
1:1 B
1:1 W
1.5:1 W
2:1 W
5:1 W
10:1 W

Results

Table 1: Hypothetical versus Prior Answers

	(1)	(2)	(3)	(4)
Outcome: Belief Police Racist (z-score)	Police Racist	Police Racist	Police Racist	Police Racist
Information	-0.2551**	-0.4005***		-0.2263**
	(0.1080)	(0.0980)		(0.1116)
Information × Democrat			-0.5865***	
			(0.1236)	
Information × Independent/Other			-0.3870*	
			(0.2101)	
Information × Republican			0.1124	
			(0.2133)	
Information × Death Ratio (Log)				-0.2757***
				(0.0771)
Demographics	NO	YES	NO	NO
Political Party	NO	YES	YES	NO
Death Ratio Estimate	NO	YES	NO	YES
R-squared	0.0163	0.3084	0.1792	0.0581
Observations	337	337	337	337

Robust standard errors clustered by individual. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Results

Table 2: Ex Post versus Hypothetical Answers

	(1)	(2)	(3)	(4)
Outcome: Belief Police Racist (z-score)	Police Racist	Police Racist	Police Racist	Police Racist
Information	0.3189***	0.4508***		0.2530**
	(0.1064)	(0.0951)		(0.1113)
Information × Democrat			0.6263***	
			(0.1187)	
Information × Independent/Other			0.7088***	
			(0.2062)	
Information × Republican			-0.2143	
			(0.2021)	
Information × Death Ratio (Log)				0.2917***
				(0.0790)
Demographics	NO	YES	NO	NO
Political Party	NO	YES	YES	NO
Death Ratio Estimate	NO	YES	NO	YES
R-squared	0.0255	0.3786	0.2418	0.0726
Observations	343	343	343	343

Robust standard errors clustered by individual. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Experiment Two: Noisy Information Provision

In the first experiment, we provide information about data, and ask questions about police racism as the outcome measure.

- If people distrust the accuracy of the underlying data, then they won't update about real-world beliefs in response.

Accordingly, we run a second experiment where we describe Roland Fryer's 2019 study of police shootings, and ask people to guess what it found.

- If people distrust the study, they may not update their real-world beliefs, but they should still update their beliefs about the study itself.
- This also allows us to verify the robustness of our results to a slightly different domain

Experimental Procedure

To ease belief elicitation, we discretize the possible results that Fryer could have found into a mutually exclusive and collectively exhaustive set of bins.

We then elicit participant's prior beliefs that Fryer's result fell into each of these bins.

We then randomly reveal one of the options to be false, and eliminate it, and elicit updated beliefs.

This is continued until only two options remain (one erroneous, one Fryer's actual result).

Thus, we need to construct benchmarks of how belief updating should (or may) occur under rationality versus motivated reasoning.

Bayesian Updating

Suppose a variable X takes one of the values in a finite set $\mathcal{A}$

For each element $k \in \mathcal{A}$, $p(k)$ is the prior probability that $X = k$.

Consider an individual who updates their beliefs about the the probability distribution of X as false options are eliminated at random.

Suppose option l is eliminated (that is, it is revealed that $X \neq l$). Then, according to Bayes' Rule,

$$p(k|X \neq l) = \frac{p(k)}{1 - p(l)} \tag{2}$$

The Favored Set and Motivated Reasoning

Alternately, consider an individual j who accrues identity utility and thus engages in motivated reasoning.

We wish to construct a simple heuristic of how they may update as false options are randomly eliminated.

For some option(s) $k \in \mathcal{A}$, believing $X = k$ grants the individual identity utility.

- Denote the set of such options the *favorable set* $\mathcal{A}_{\mathcal{F}}$, and the remaining options the *unfavorable set* $\mathcal{A}_{\mathcal{U}}$.

Motivated Reasoning Heuristic

We want to construct a heuristic that captures the intuition that, for an individual engaging in motivated reasoning,

- Good news (here, the elimination of unfavorable options) is unambiguously good and emphasised
- Bad news (here, the elimination of favorable options) is interpreted as ambiguous and minimised

More formally, we define the motivated reasoning heuristic as follows:

- If a unfavorable option l is eliminated, increase $p(k \in \mathcal{A}_{\mathcal{F}})$ by $p(l)$.
- If a favorable option l is eliminated, leave $p(k \in \mathcal{A}_{\mathcal{F}})$ unchanged.
- Within each of $\mathcal{A}_F$ and $\mathcal{A}_U$, assign the probabilities across the surviving options following Bayes' Rule
- If no favorable options remain, follow Bayes' Rule

Empirical Strategy

For each individual, we have elicited probabilities p_{jkr} for each (surviving) option k in round r (for $k \in \{1, 2, 3, 4, 5\}$, $r \in \{1, 2, 3, 4\}$).

Please estimate the probability that each of these options is the truth about Fryer's study.

Fryer found that, holding all else equal, police are:

More than 2.5 times as likely to shoot Black people	0
Between 1.5 and 2.5 times as likely to shoot Black people	0
About equally likely to shoot either Black/White people (factor of less than 1.5)	0
Between 1.5 and 2.5 times as likely to shoot White people	0
More than 2.5 times as likely to shoot White people	0
Total	0

Empirical Strategy

Denote the Bayesian update as $b_{j,k,r}$, and the motivating reasoning heuristic update as $m_{j,k,r}$.

To construct the motivated reasoning heuristic, we need to define the *favorable set* for each individual.

- $\mathcal{A}_{\mathcal{F}} = \{1, 2\}$ who indicate baseline support for the *Black Lives Matter* movement
- $\mathcal{A}_{\mathcal{F}} = \{3, 4, 5\}$ otherwise

We pool the data across individuals, options, and rounds, and estimate

$$p_{j,k,r} = \alpha + \beta_b * b_{j,k,r} + \beta_m * m_{j,k,r} + \varepsilon_{j,k,r} \tag{3}$$

Our interest is primarily in explicit evidence for asymmetric updating in response to news as captured by β_m

- Not merely arbitrary deviations from Bayesian behavior ($\beta_b \neq 1$)

Results

Table 3: Probability Updating Estimates

	(1) Probability	(2) Probability	(3) Probability	(4) Probability
Bayesian Posterior	0.8250***	0.8089***	0.5361***	0.5385***
	(0.0256)	(0.0353)	(0.0776)	(0.0812)
Probability (Lag)		0.0269		-0.0043
		(0.0394)		(0.0467)
Motivated Reasoning Heuristic			0.2860***	0.2861***
			(0.0700)	(0.0701)
Constant	0.0582***	0.0574***	0.0592***	0.0593***
	(0.0085)	(0.0085)	(0.0089)	(0.0091)
R-squared	0.6876	0.6877	0.6960	0.6960
Observations	1330	1330	1091	1091

Robust standard errors clustered by individual. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Results

To provide evidence that β_m is not merely picking up people using other (coincidentally correlated) heuristics, we add other heuristics as controls.

Results

Table 4: Probability Updating Estimates (Robustness)

	(1) Probability	(2) Probability
Bayesian Posterior	0.2244*	0.2245*
	(0.1186)	(0.1187)
Motivated Reasoning Heuristic	0.2255***	0.2255***
	(0.0687)	(0.0687)
Equal Levels	0.1390**	0.1400**
	(0.0545)	(0.0544)
Equal Change	0.0947	0.0929
	(0.1102)	(0.1102)
Non-Independence Heuristic	0.3137***	0.3137***
	(0.0652)	(0.0652)
Probability (Lag)		0.0018
		(0.0027)
Constant	0.0009	0.0008
	(0.0014)	(0.0015)
R-squared	0.7167	0.7167
Observations	1091	1091

Robust standard errors clustered by individual. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Thank you