Mental Models and Learning: The Case of Base-Rate Neglect

Ignacio Esponda
UC Santa Barbara

Emanuel Vespa
UC San Diego

Sevgi Yuksel
UC Santa Barbara

Econometric Society Meetings, January 2022

Motivation

A large literature documents biases in decision-making.

- Base-rate neglect (Kahneman Tversky 73), overconfidence (Mobius et al. 10), sunk-cost effect (Thaler 80), Gamber's fallacy (Rabin 02), correlation neglect (Enke Zimmermann 19), failures of cont. reasoning (Esponda Vespa 14 19, Martinez-Marquina et al. 19), etc.

Motivation

A large literature documents biases in decision-making.

- Base-rate neglect (Kahneman Tversky 73), overconfidence (Mobius et al. 10), sunk-cost effect (Thaler 80), Gamber's fallacy (Rabin 02), correlation neglect (Enke Zimmermann 19), failures of cont. reasoning (Esponda Vespa 14 19, Martinez-Marquina et al. 19), etc.

Less is known on whether such biases interact with feedback.

Initial misconceptions can impact learning from feedback.

Proof of concept: Base-Rate Neglect

(Kahneman Tversky 73)

- Disease prevalence in the population: $p = 0.15$.
- Reliability of the test: $q = 0.80$, i.e. test is positive (negative) with 80% chance if the person is sick (not sick).
- What is the likelihood that the person is sick conditional on a positive test?

Proof of concept: Base-Rate Neglect

(Kahneman Tversky 73)

- Disease prevalence in the population: $p = 0.15$.
- Reliability of the test: $q = 0.80$, i.e. test is positive (negative) with 80% chance if the person is sick (not sick).
- What is the likelihood that the person is sick conditional on a positive test?

Bayesian Benchmark:
$\Pr(\text{sick}|\text{positive}) = \frac{pq}{pq+(1-p)(1-q)} = 0.41.$

Proof of concept: Base-Rate Neglect

(Kahneman Tversky 73)

- Disease prevalence in the population: $p = 0.15$.
- Reliability of the test: $q = 0.80$, i.e. test is positive (negative) with 80% chance if the person is sick (not sick).
- What is the likelihood that the person is sick conditional on a positive test?

Bayesian Benchmark:
$\Pr(\text{sick}|\text{positive}) = \frac{pq}{pq+(1-p)(1-q)} = 0.41.$

Base-Rate Neglect (BRN):
(incorrect mental model) $= 0.80$.

Preview of Results

- Are people Bayesian in the long run?

Preview of Results

- Are people Bayesian in the long run?

  No, convergence to the benchmark is slow and partial.

Preview of Results

- Are people Bayesian in the long run?

  No, convergence to the benchmark is slow and partial.

- Does BRN hinder learning from feedback?

Preview of Results

» Are people Bayesian in the long run?

No, convergence to the benchmark is slow and partial.

» Does BRN hinder learning from feedback?

Yes, learning is faster in the absence of it.

Preview of Results

- Are people Bayesian in the long run?

  No, convergence to the benchmark is slow and partial.

- Does BRN hinder learning from feedback?

  Yes, learning is faster in the absence of it.
  Those who initially display BRN are less attentive to feedback.

Preview of Results

- Are people Bayesian in the long run?

  No, convergence to the benchmark is slow and partial.

- Does BRN hinder learning from feedback?

  Yes, learning is faster in the absence of it.
  Those who initially display BRN are less attentive to feedback.

- Can BRN be corrected?

Preview of Results

- Are people Bayesian in the long run?

  No, convergence to the benchmark is slow and partial.

- Does BRN hinder learning from feedback?

  Yes, learning is faster in the absence of it.
  Those who initially display BRN are less attentive to feedback.

- Can BRN be corrected?

  Yes, when feedback is presented in a way that unequivocally challenges BRN.

Experimental Design

Sec. 1 **Introducing the main updating task.**

Subjects update beliefs on a binary state using a binary signal.

(Kahneman Tversky 73)

» There are 100 projects.

» 15 projects are successes; 85 are failures.

» Task: assess the chance that a randomly selected project is a success conditional on a test result.

» The test result is either positive or negative and has a reliability of 80%.

Experimental Design

Sec. 1 **Introducing the main updating task.**

Subjects update beliefs on a binary state using a binary signal.

(Kahneman Tversky 73)

» There are 100 projects.

» 15 projects are successes; 85 are failures.

» Task: assess the chance that a randomly selected project is a success conditional on a test result.

» The test result is either positive or negative and has a reliability of 80%.

Sec. 2 **Repetition of the task for 200 rounds.**

Round 1 Feedback

Round 5 Feedback

Experimental Design

2nd goal: Study how initial misconceptions can impact learning.

Two between-subjects treatments that *only* differ in how the main updating task is introduced.

Primitives (P): Primitives (p and q) provided.

Experimental Design

2nd goal: Study how initial misconceptions can impact learning.

Two between-subjects treatments that *only* differ in how the main updating task is introduced.

Primitives (P): Primitives (p and q) provided.

NoPrimitives (NoP): Primitives (p and q) <u>not</u> provided.

Experimental Design

- There are 100 projects.

  Primitives: 15 projects are successes; 85 are failures.

  NoPrimitives: A certain number of projects are successes and the remaining ones are failures.

- The computer runs a test on the selected project.

  Primitives: The test result either positive or negative and has a reliability of 80%.

  NoPrimitives: The test result either positive or negative and has a reliability of q%.

Feedback is structurally the same in both treatments, but BRN is possible only in Primitives.

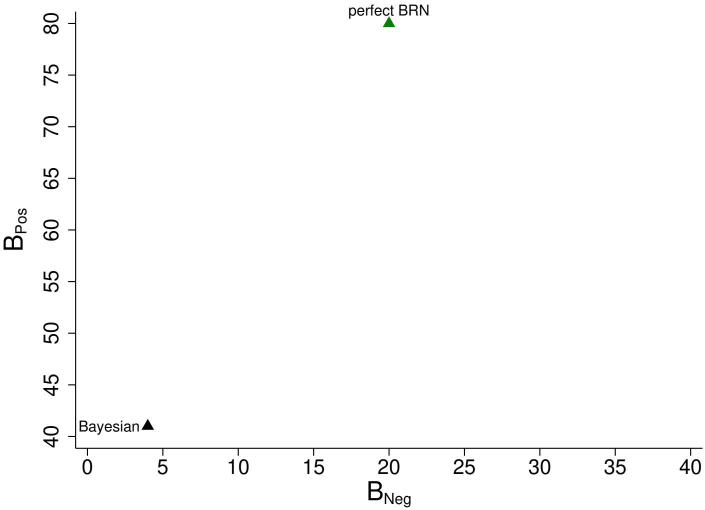
Experimental Design

Other details:

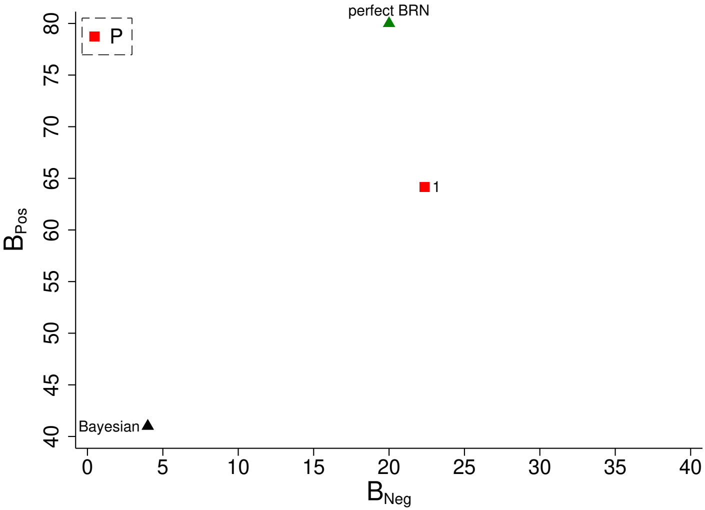
- Participants: 128 UCSB students.
- Subjects paid for one choice in one part (using BDM).
- ~90 min sessions (subjects move at own pace).
- Average payoffs: $22.5 (either $10 or $35).

RESULTS

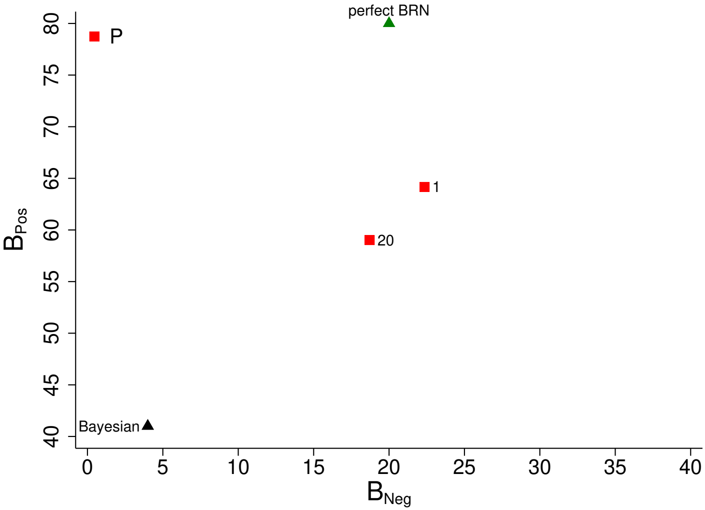
Theoretical Benchmarks



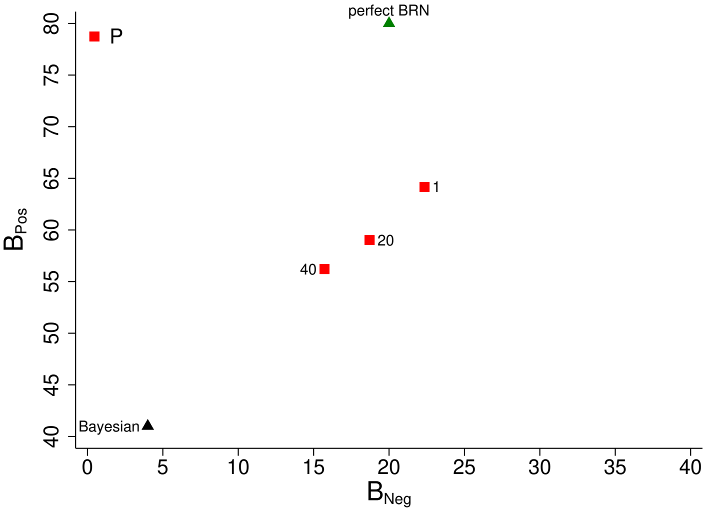
Average Beliefs: Primitives



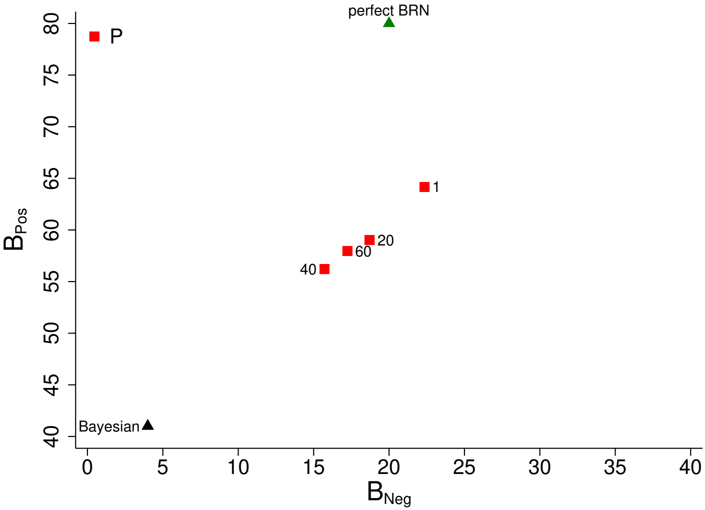
Average Beliefs: Primitives



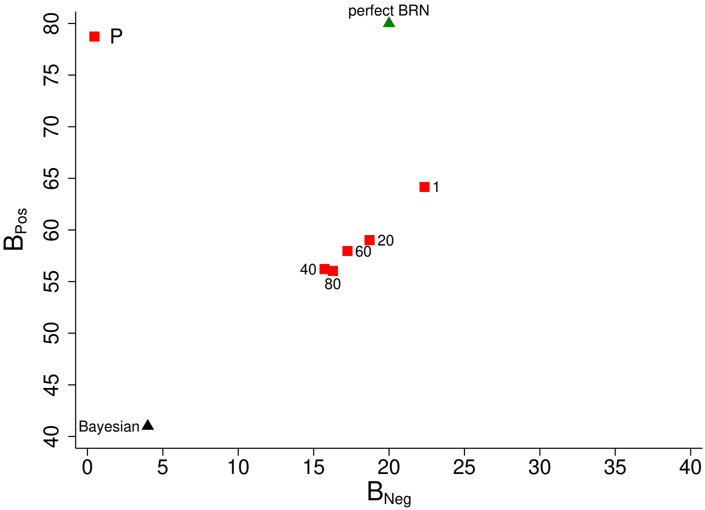
Average Beliefs: Primitives



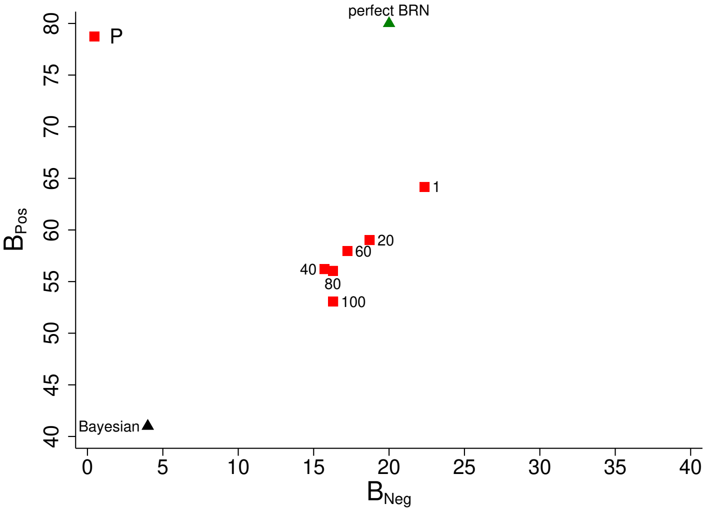
Average Beliefs: Primitives



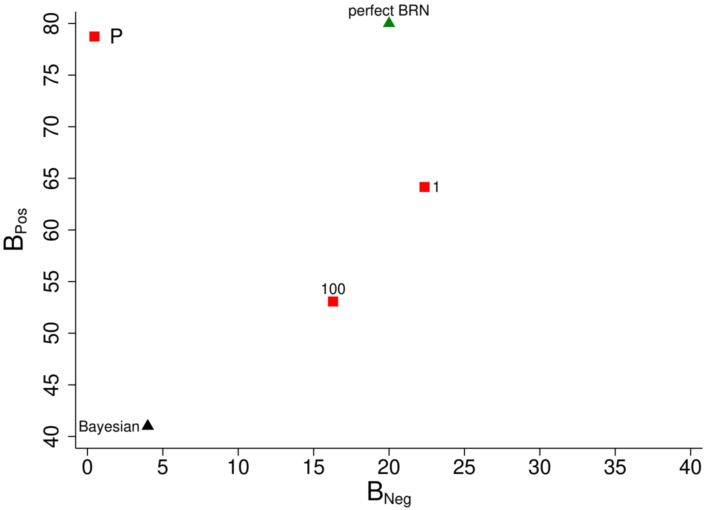
Average Beliefs: Primitives



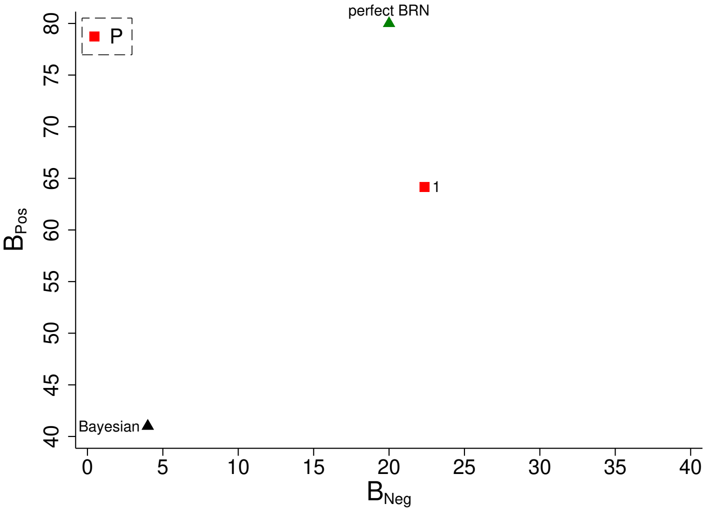
Average Beliefs: Primitives



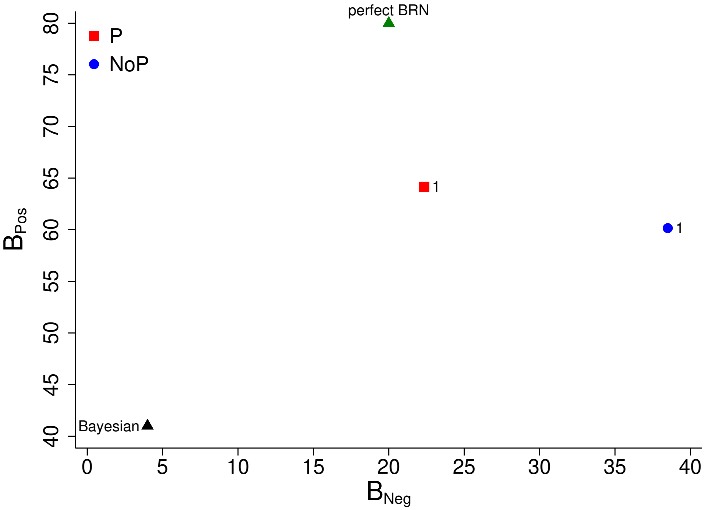
Beliefs in Primitives: Rounds 1 & 100



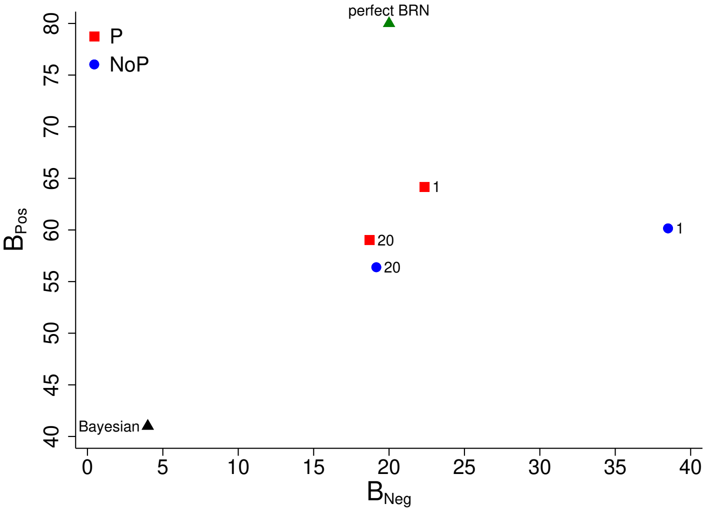
Average Beliefs: Round 1



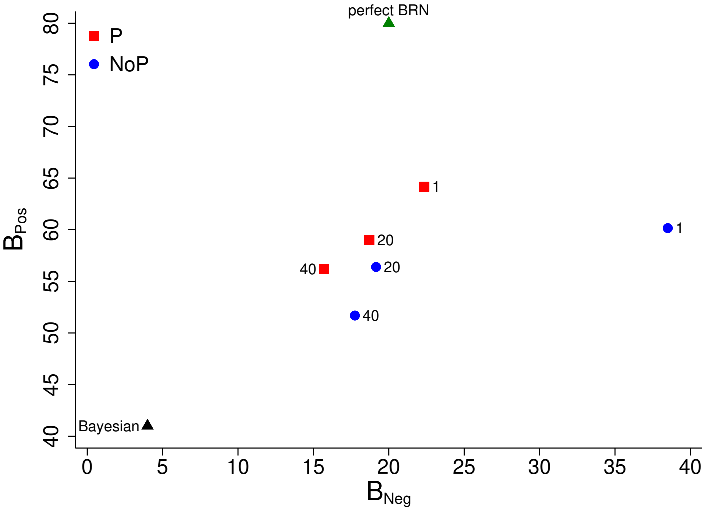
Average Beliefs: Round 1



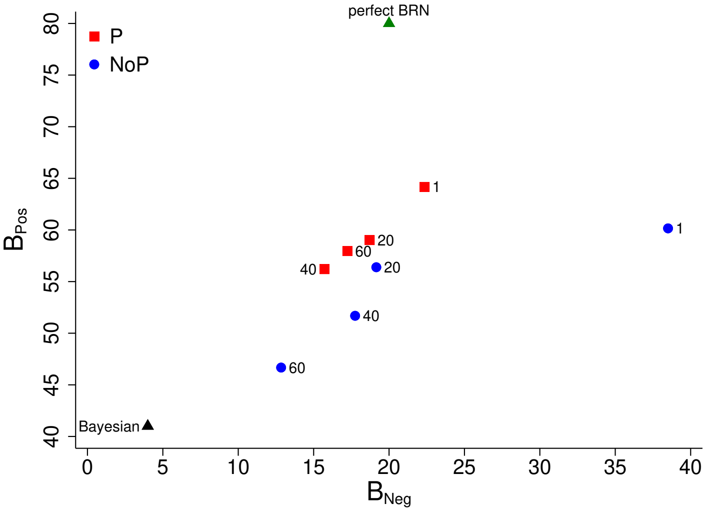
Evolution of Average Beliefs



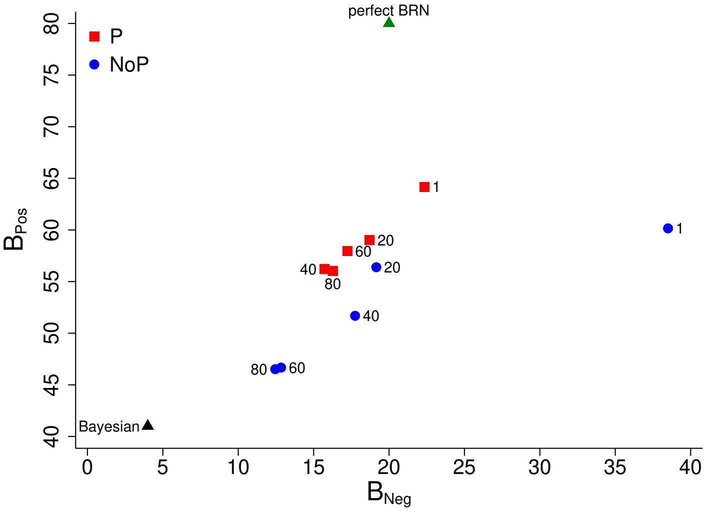
Evolution of Average Beliefs



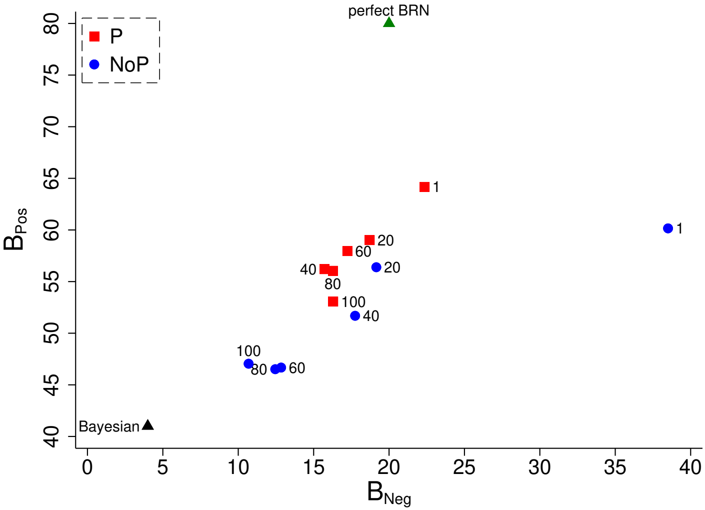
Evolution of Average Beliefs



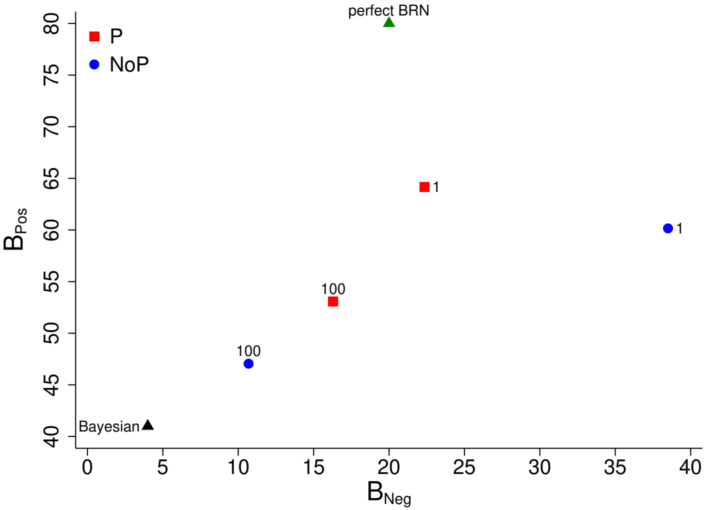
Evolution of Average Beliefs



Evolution of Average Beliefs

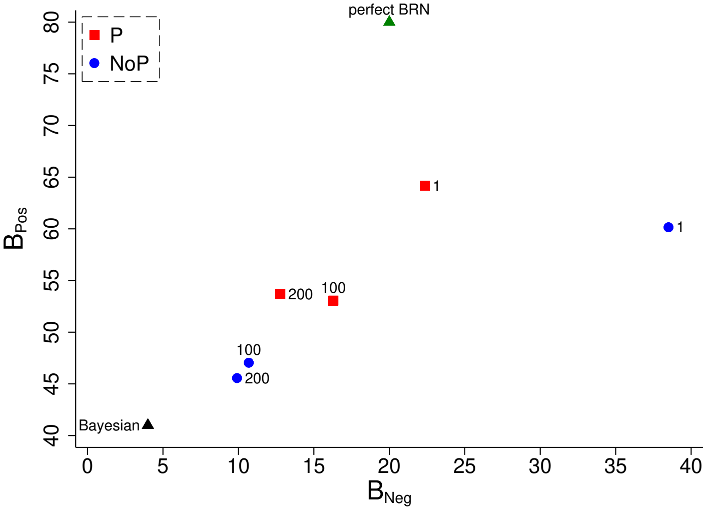


Evolution of Beliefs: Rounds 1 & 100



Evolution of Average Beliefs: Round 200

Statistical analysis



Result #1

Beliefs in both treatments move closer to the Bayesian benchmark from round 1 to 200.

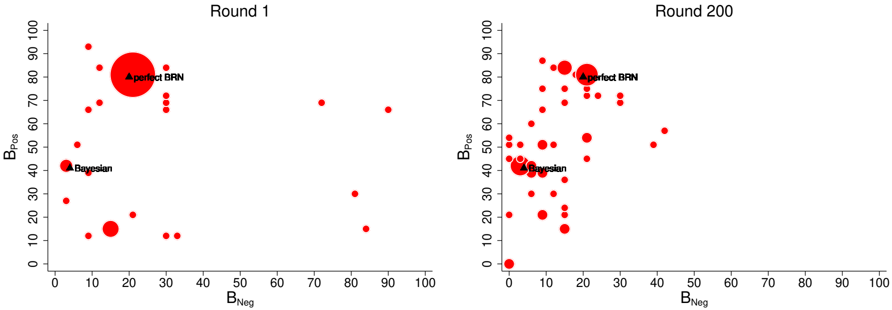
By round 200,

» Beliefs are significantly different between Primitives and NoPrimitives.

» Beliefs closer to the Bayesian benchmark in NoPrimitives.

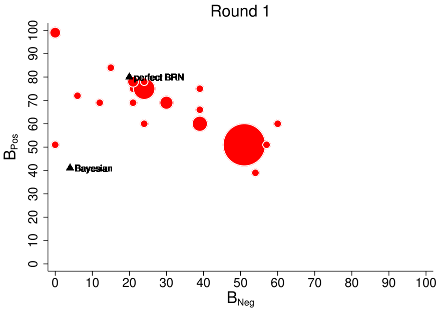
Distribution of Beliefs

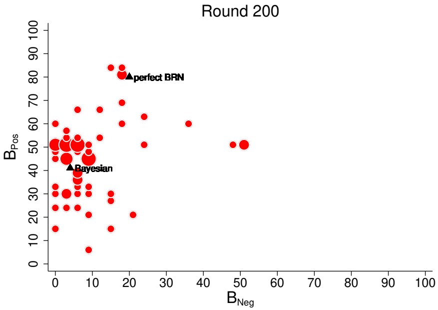
Primitives Treatment



Distribution of Beliefs

NoPrimitives Treatment





Result #2

Subjects in the Primitives treatment who display BRN in round 1 are driving the treatment effect in round 200.

Main Findings

Are beliefs Bayesian in the long run?

» **Adjustment is slow and partial.**

Does starting with BRN hinder learning?

» **Beliefs closer to the Bayesian benchmark in NoPrimitives.**

Why is learning slower those who initially display BRN?

» **Results suggest they are less attentive to feedback.**

Can BRN be corrected?

» **If feedback is provided in summary form to challenge the BRN model.**

Some Implications

Biases can be persistent because they impact how agents learn from feedback.

- Literature studying implications of misspecified models. Esponda Pouzo 16, Fudenberg Romanyuk Strack 17, Bohren Hauser 17, Heidhues Koszegi Strack 18.
- Literature on endogenous attention. Sims 03, Caplin Dean 15, Schwartzstein 14, Gagnon-Bartsch Rabin Schwartzstein 18.

Some Implications

Policy implications:

- Biases can be persistent even in information rich environments.
- Interventions need to influence how agents engage with information.
- Withholding payoff relevant information can improve long run behavior.

Some Implications

Policy implications:

- Biases can be persistent even in information rich environments.
- Interventions need to influence how agents engage with information.
- Withholding payoff relevant information can improve long run behavior.

Proof of concept: design can be used to study persistence of biases in other settings.