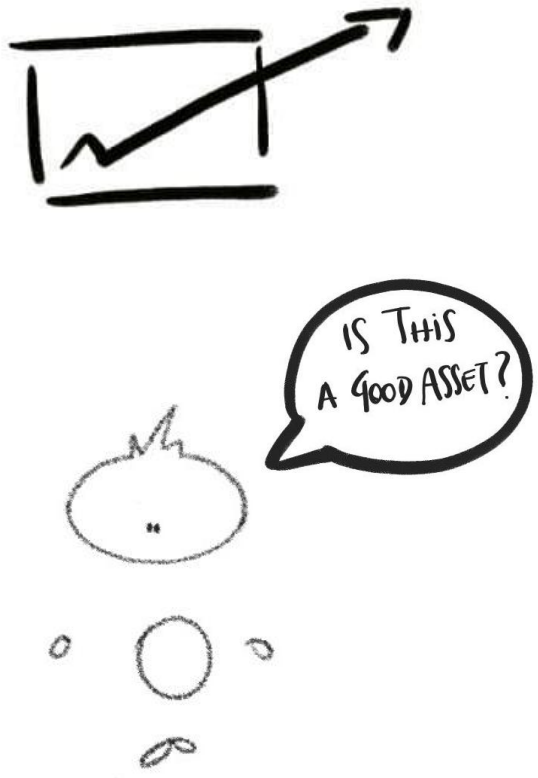
Weighting Competing Models

Chiara Aina
Universitat Pompeu Fabra

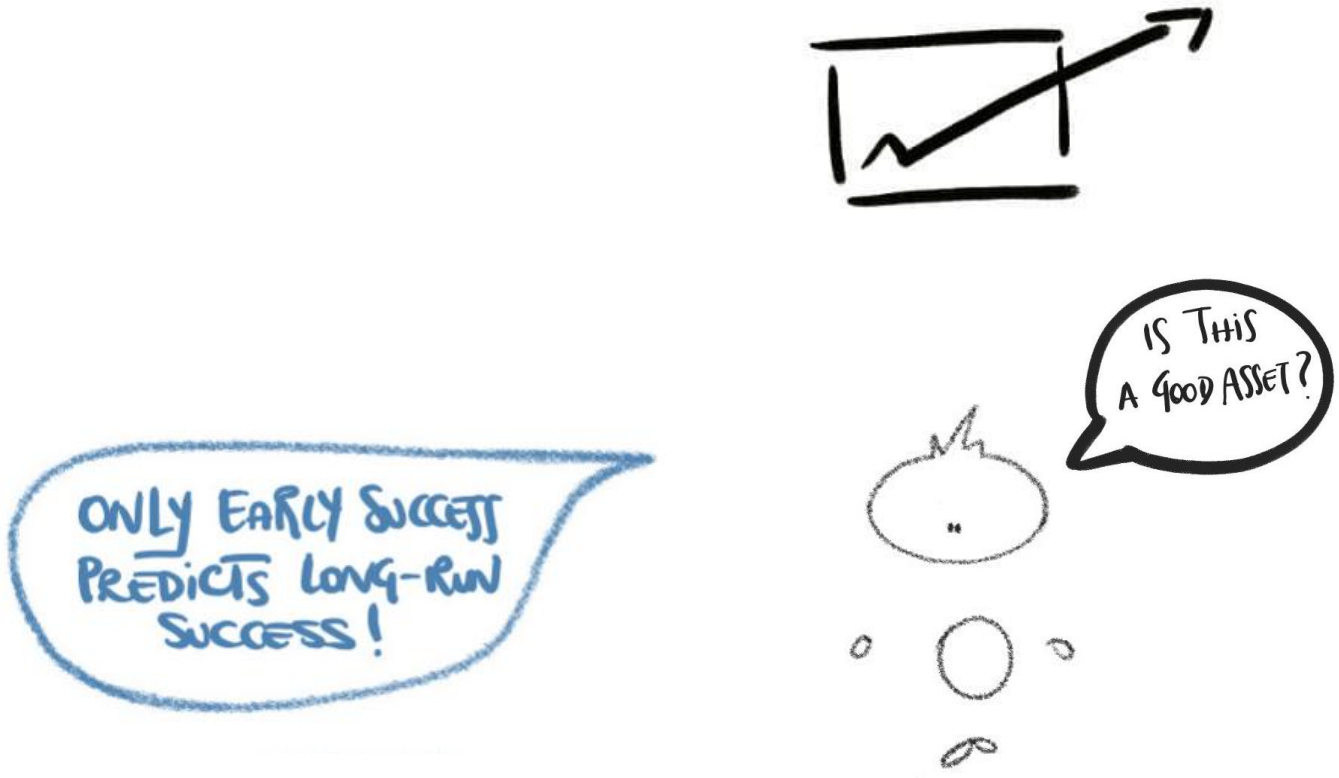
Florian Schneider
University of Copenhagen

June 2025

Motivation



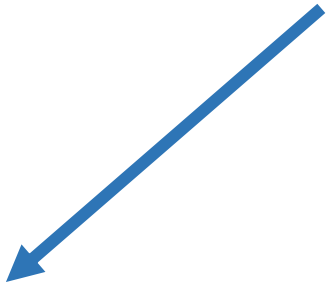
Motivation


ONLY EARLY SUCCESS PREDICTS LONG-RUN SUCCESS!
IS THIS A GOOD ASSET?

Motivation

New piece of information

past financial performance

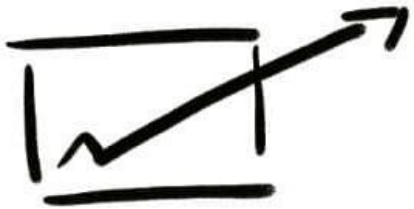


Learn about unknown state

quality of the asset (good vs bad)

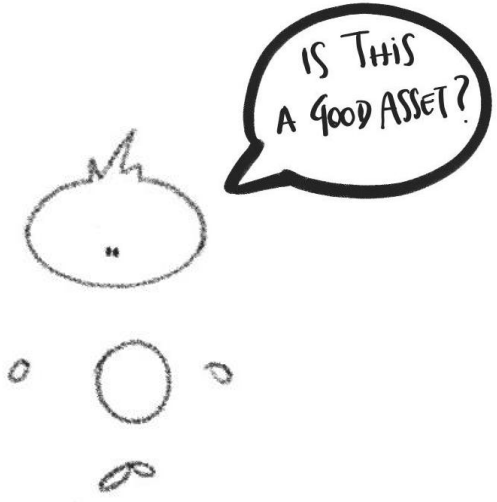
Motivation


ONLY LOSSES ARE INFORMATIVE!



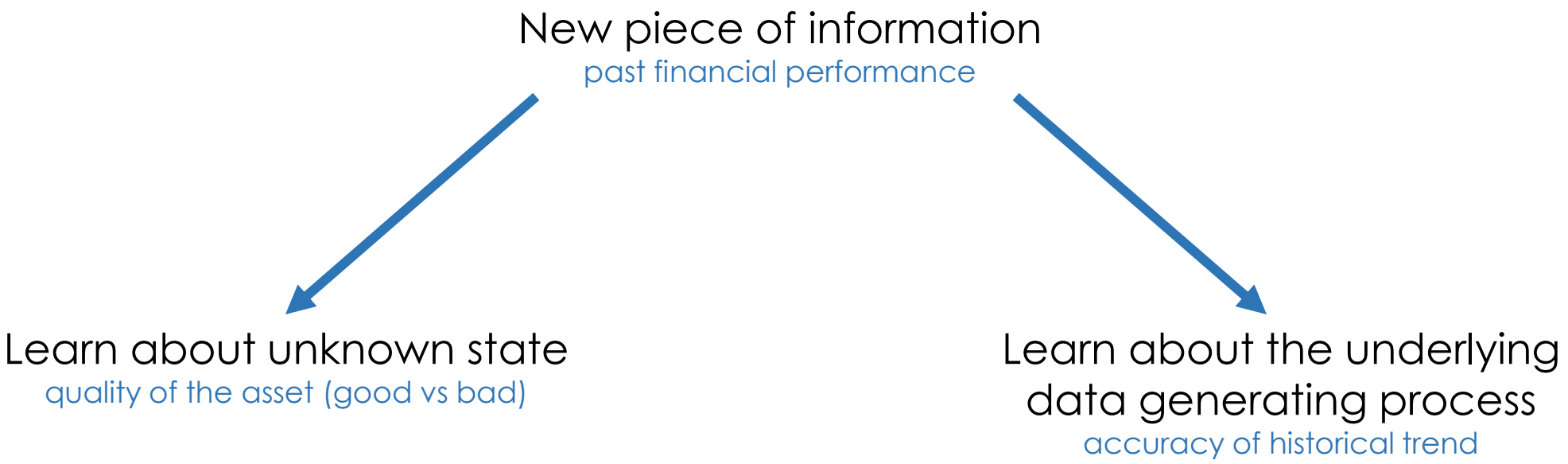

WHAT GOES DOWN GOES UP!


ONLY EARLY SUCCESS PREDICTS LONG-RUN SUCCESS!

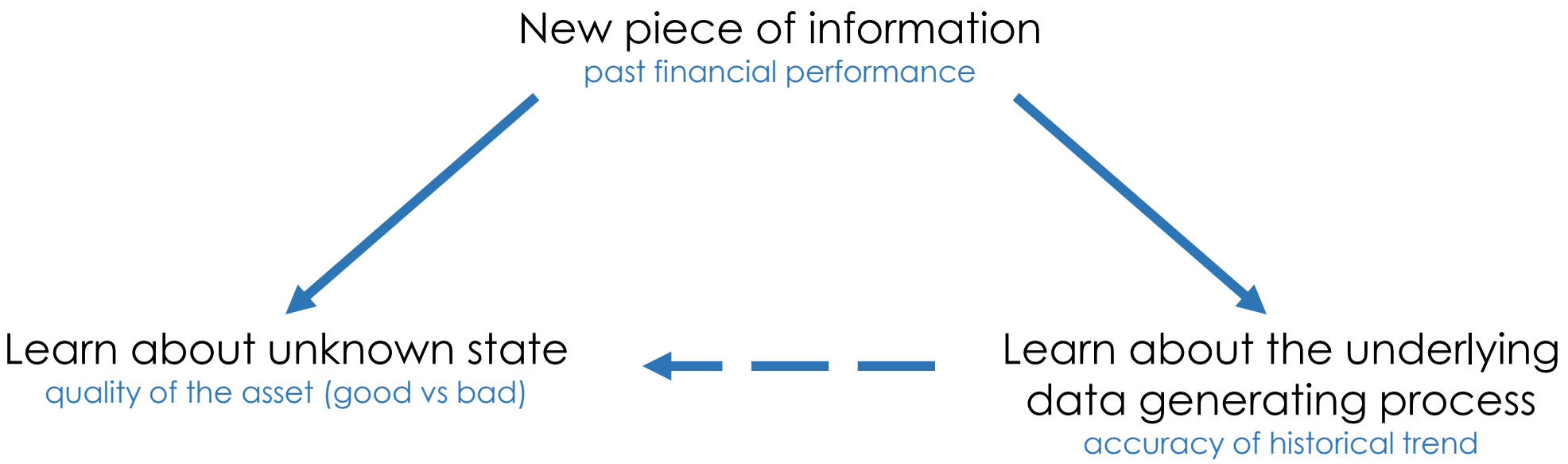

IS THIS A GOOD ASSET?


PAST & FUTURE PERFORMANCES ARE NOT LINKED!

Motivation


New piece of information
past financial performance
Learn about unknown state
quality of the asset (good vs bad)
Learn about the underlying data generating process
accuracy of historical trend

Motivation


New piece of information
past financial performance
Learn about unknown state
quality of the asset (good vs bad)
Learn about the underlying data generating process
accuracy of historical trend

This Paper

With competing models, how do we update beliefs?

Study biases in model weighting using an online experiment

- One-to-one with the literature on narratives as models (model persuasion)
  e.g., Schwartzstein & Sunderam (2021, 2024), Aina (2024), Ichihashi & Meng (2021), Ispano (2024), Yang (2023), Jain (2024), Barron and Fries (2024), …

- Little is known about how people update in these settings
  e.g., Liang (2023)

  - Do people select one model or rather put positive weights on multiple models? According to which criteria?

  - Are these updating rules used consistently?

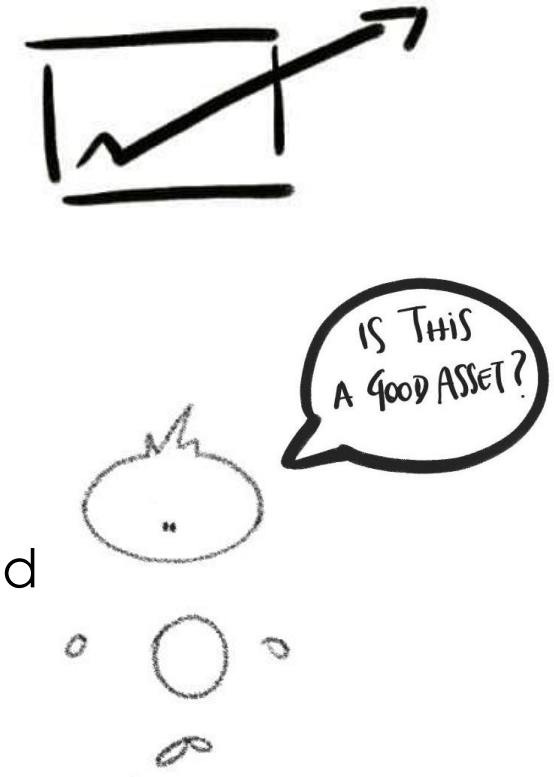
more

Framework

Example

Past data can be either low or high

Observe high!



Asset can be either good or bad

Pr(G)= 50%

G= good asset, B = bad asset
H = high signal, L = low signal

Models

Model = a map that assigns conditional probabilities of signals given states

- How the signals are generated given the state (data-generating process)
- How to interpret the signals to learn about the state using Bayes rule

$$\Pr^{model}(H \mid G)$$

$$\Pr^{model}(H \mid B)$$

G= good asset, B = bad asset
H = high signal, L = low signal

Updating about States

models

$Pr^{m1}(H \mid G) = 100\%$

$Pr^{m1}(H \mid B) = 0\%$

$Pr^{m2}(H \mid G) = 17\%$

$Pr^{m2}(H \mid B) = 67\%$

model predictions

$Pr^{m1}(G \mid H) = 100\%$

$P^{m2}(G \mid H) = 20\%$

G= good asset, B = bad asset
H = high signal, L = low signal

Updating about States and Models

models

$Pr^{m1}(H \mid G) = 100\%$ $\quad$ $Pr^{m2}(H \mid G) = 17\%$

$Pr^{m1}(H \mid B) = 0\%$ $\quad$ $Pr^{m2}(H \mid B) = 67\%$

model predictions

$Pr^{m1}(G \mid H) = 100\%$ $\quad$ $P^{m2}(G \mid H) = 20\%$

reported belief

$$Pr(G \mid H) = \rho_H^{m1}\ 100\% + \rho_H^{m2}\ 20\%$$

We study biases in model weighting!

G= good asset, B = bad asset
H = high signal, L = low signal

Updating Rules

Bayesian Updating

$\rho_s^m = \Pr(m \mid s)$

model predictions weighted by the model posterior given the signal & prior over models

→ learning about the models

Updating Rules

Bayesian Updating

$\rho_s^m = \Pr(m \mid s)$

model predictions weighted by the model posterior given the signal & prior over models

→ learning about the models

One-stage Updating

$\rho_s^m = \Pr(m)$

model predictions weighted by prior over models

→ extreme form of underinfence about the models

Updating Rules

Bayesian Updating

$\rho_s^m = \Pr(m \mid s)$

model predictions weighted by the model posterior given the signal & prior over models

→ learning about the models

One-stage Updating

$\rho_s^m = \Pr(m)$

model predictions weighted by prior over models

→ extreme form of underinfence about the models

Model Selection via Maximum Likelihood

$\rho_s^{m^*} = 100\%$ w/ $m^* \in \text{argmax } \Pr(s \mid m)$

select the best-fitting model & only form posteriors using that model

→ extreme form of overinference about the models

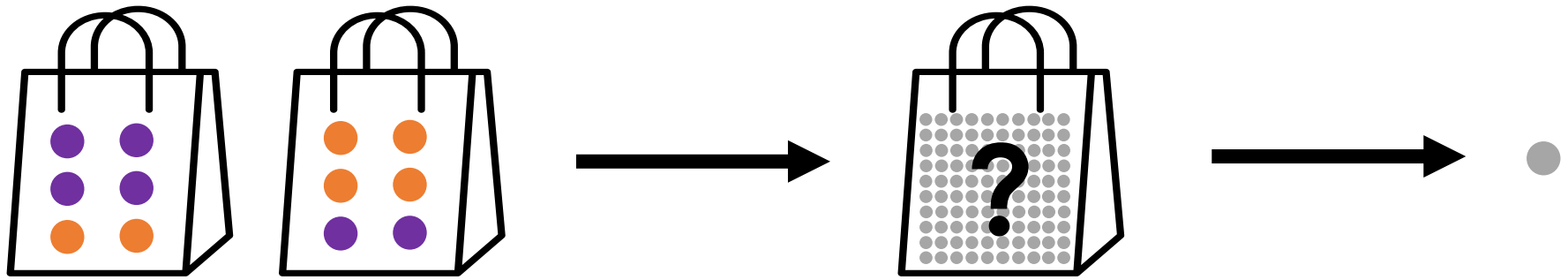
- Stochastic Model Selection
- Model Selection based on Informativeness
- Dogmatic Model Selection

Experimental Design

Task

Classic "balls and urns" updating exercise

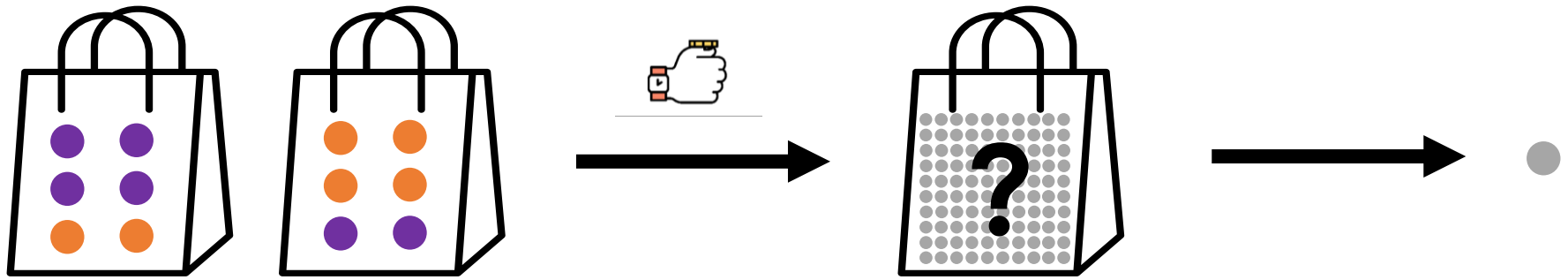
- Binary state (uniform priors) & binary signal
- Subject in control of each step (animated)



Task

Classic "balls and urns" updating exercise

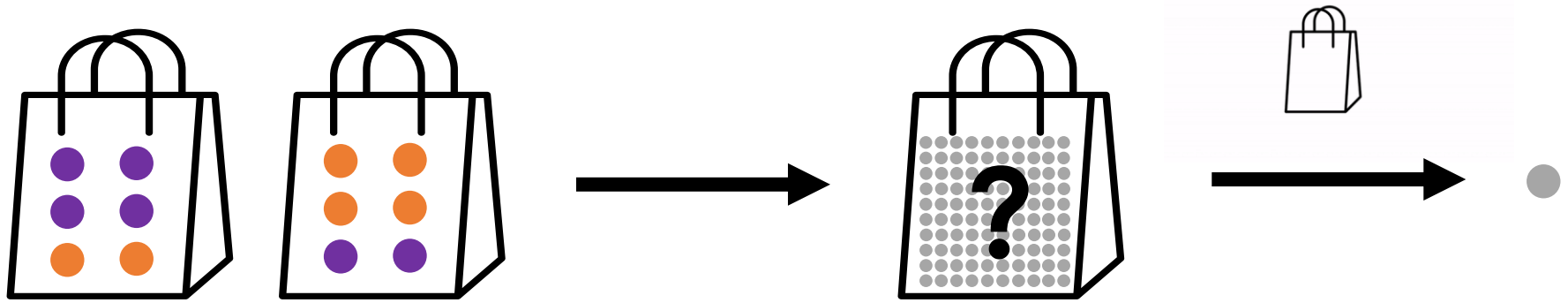
- Binary state (uniform priors) & binary signal
- Subject in control of each step (animated)



Task

Classic "balls and urns" updating exercise

- Binary state (uniform priors) & binary signal
- Subject in control of each step (animated)

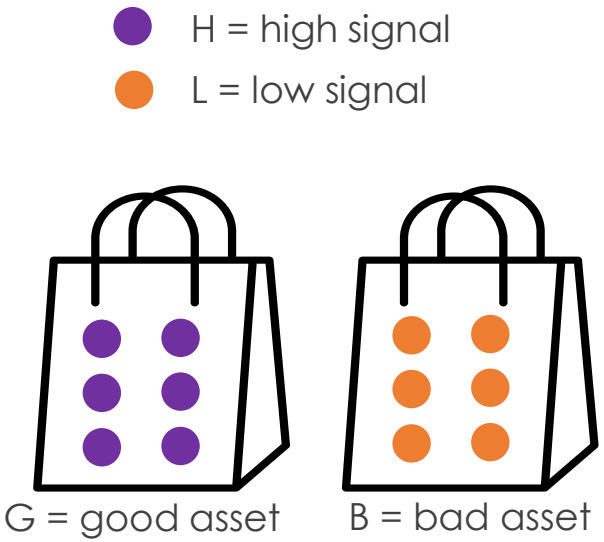


Why This Task?

model = distribution of balls in the bags

$Pr^{m1}(H \mid G) = 100\%$

$Pr^{m1}(H \mid B) = 0\%$

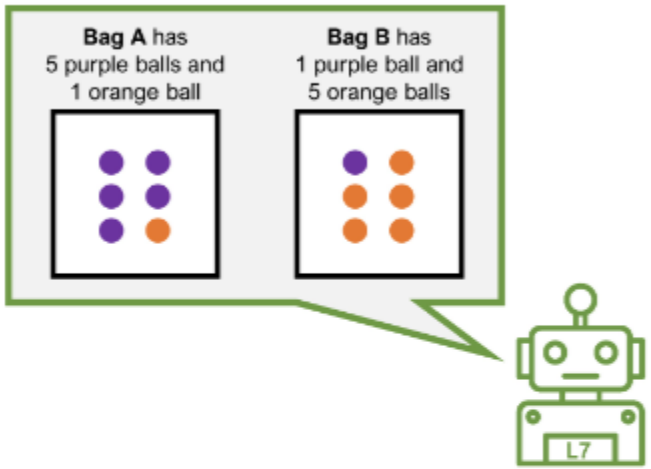


$Pr^{m2}(H \mid G) = 17\%$

$Pr^{m2}(H \mid B) = 67\%$

One Model

Remember that for this task we consider robot L7.

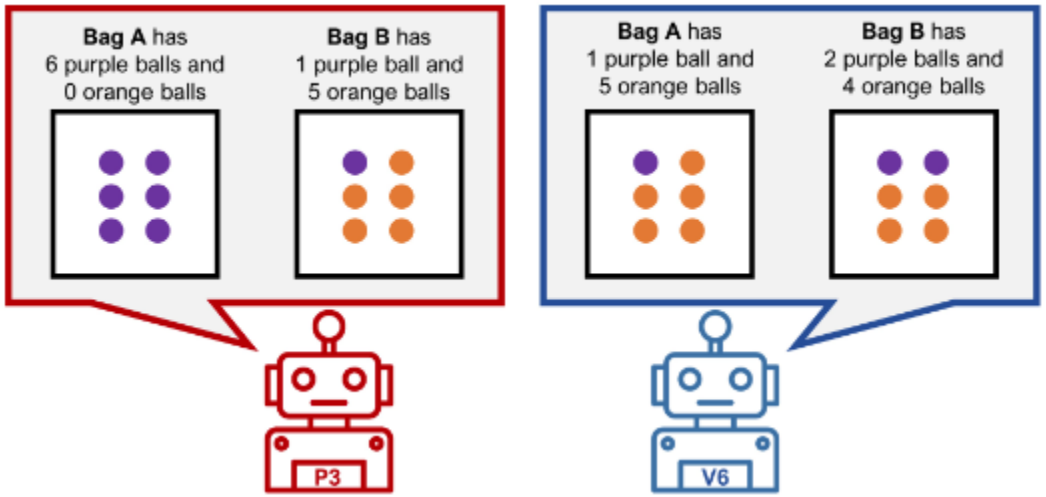


Remember, the computer has drawn an orange ball

What do you think is the chance that bag A was selected given that the color of the drawn ball is orange? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

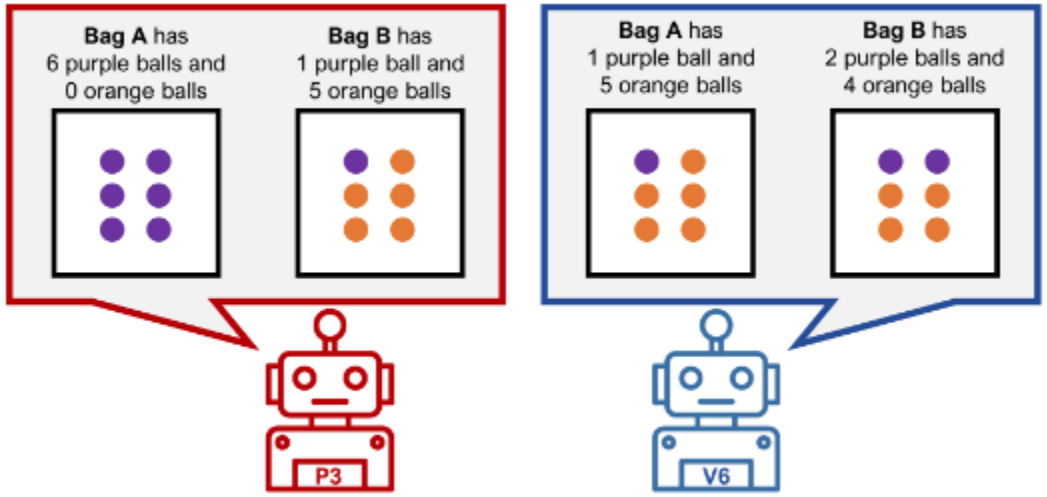
Competing Models



When you click "**START**", the following steps take place:

- **Step 1:** the computer selects one of the two robots by rolling a fair 6-sided die
  - If the number 1, 2 or 3 is rolled: the computer selects robot P3
  - If the number 4, 5 or 6 is rolled: the computer selects robot V6
- **Step 2:** the computer fills the bags with purple and orange balls according to the robot selected in Step 1
- **Step 3:** the computer selects one bag by flipping a fair coin
- **Step 4:** the computer draws a ball from the bag that was selected in Step 3

Competing Models



Remember, the computer has drawn a purple ball

The chance that bag A was selected given that the color of the drawn ball is purple would be…

86%
if I, robot P3, were selected

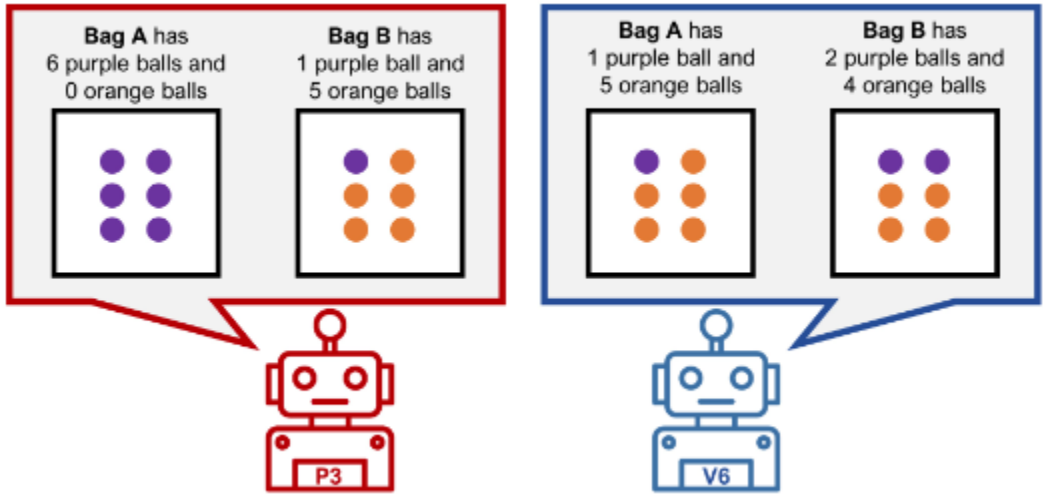
33%
if I, robot V6, were selected

What do you think is the chance that bag A was selected given that the color of the drawn ball is purple? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

Competing Models

Model selection



Remember, the computer has drawn a purple ball

The chance that bag A was selected given that the color of the drawn ball is purple would be…

86%
if I, robot P3, were selected

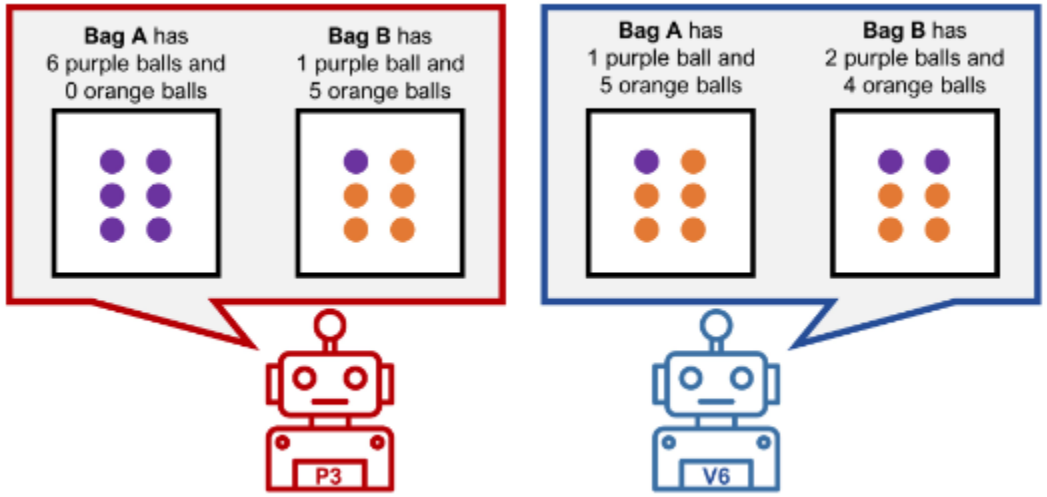
33%
if I, robot V6, were selected

What do you think is the chance that bag A was selected given that the color of the drawn ball is purple? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

Competing Models

Best-fitting model?
model w/ more purple balls
P3: 7 > 3: V6



Remember, the computer has drawn a purple ball

The chance that bag A was selected given that the color of the drawn ball is purple would be...

Selecting best-fitting model →

86%
if I, robot P3, were selected

33%
if I, robot V6, were selected

← Selecting worst-fitting model

What do you think is the chance that bag A was selected given that the color of the drawn ball is purple? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

Competing Models

Pr(P3) = Pr(V6) = 50%

One-stage updating

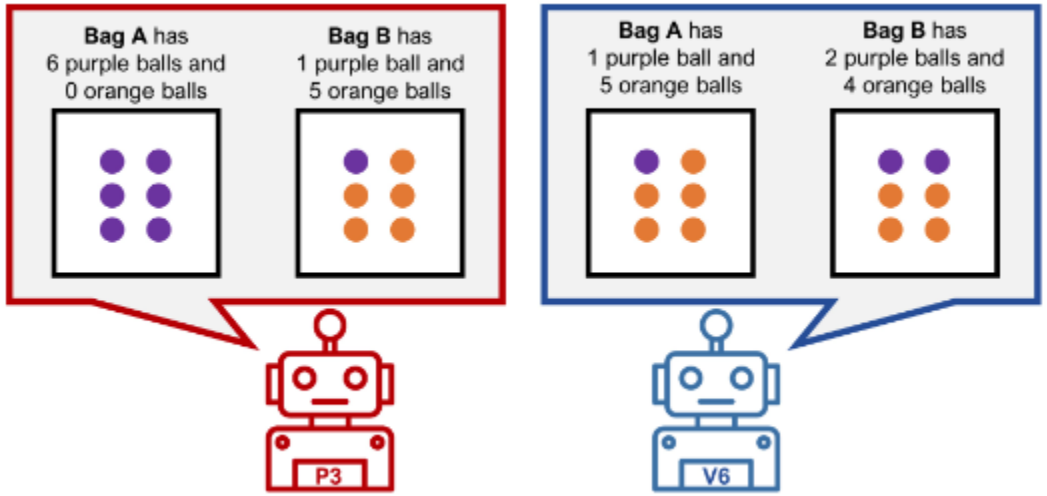
0.5 x 86% + 0.5 x 33% = 60%

Bayesian updating

0.7 x 86% + 0.3 x 33% = 70%

or

$$\frac{\#\text{ purple balls in A}}{\#\text{ purple balls}} = \frac{7}{10}$$



Remember, the computer has drawn a purple ball

The chance that bag A was selected given that the color of the drawn ball is purple would be...

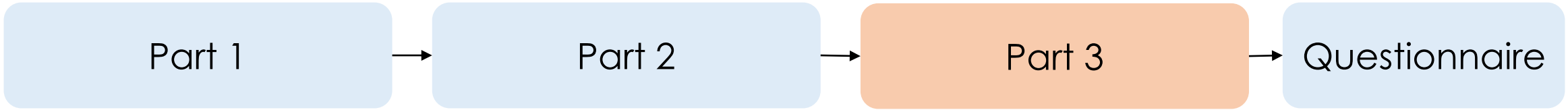
86%
if I, robot P3, were selected

33%
if I, robot V6, were selected

What do you think is the chance that bag A was selected given that the color of the drawn ball is purple? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

Design



one model	one model w/prediction	two models w/predictions
2 tasks	2 tasks	5+2 tasks

- Starting w/ classing paradigm then progressively introducing new elements
- Simple setting: 50-50 priors over bags/models & total of 6 balls in each bag
- 300 Prolific participants (April 2024), incentivization of accurate beliefs

logistics more part 4

Results

Results

- Quality of Data

1. Analysis of Individual Guesses
2. Consistency of Updating Rules
3. Analysis of Posteriors across Signal Realizations

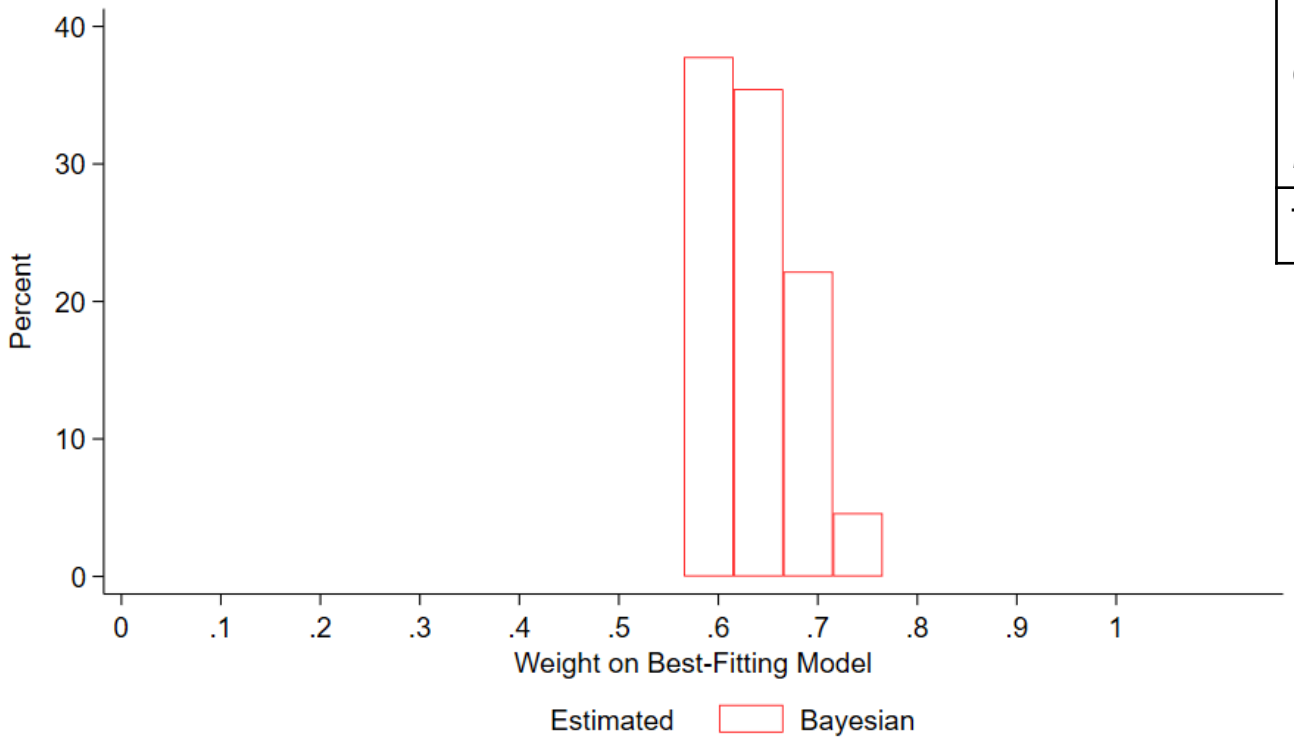
Results

- Quality of Data

1. Analysis of Individual Guesses
2. Consistency of Updating Rules
3. Analysis of Posteriors across Signal Realizations

Reported Guesses

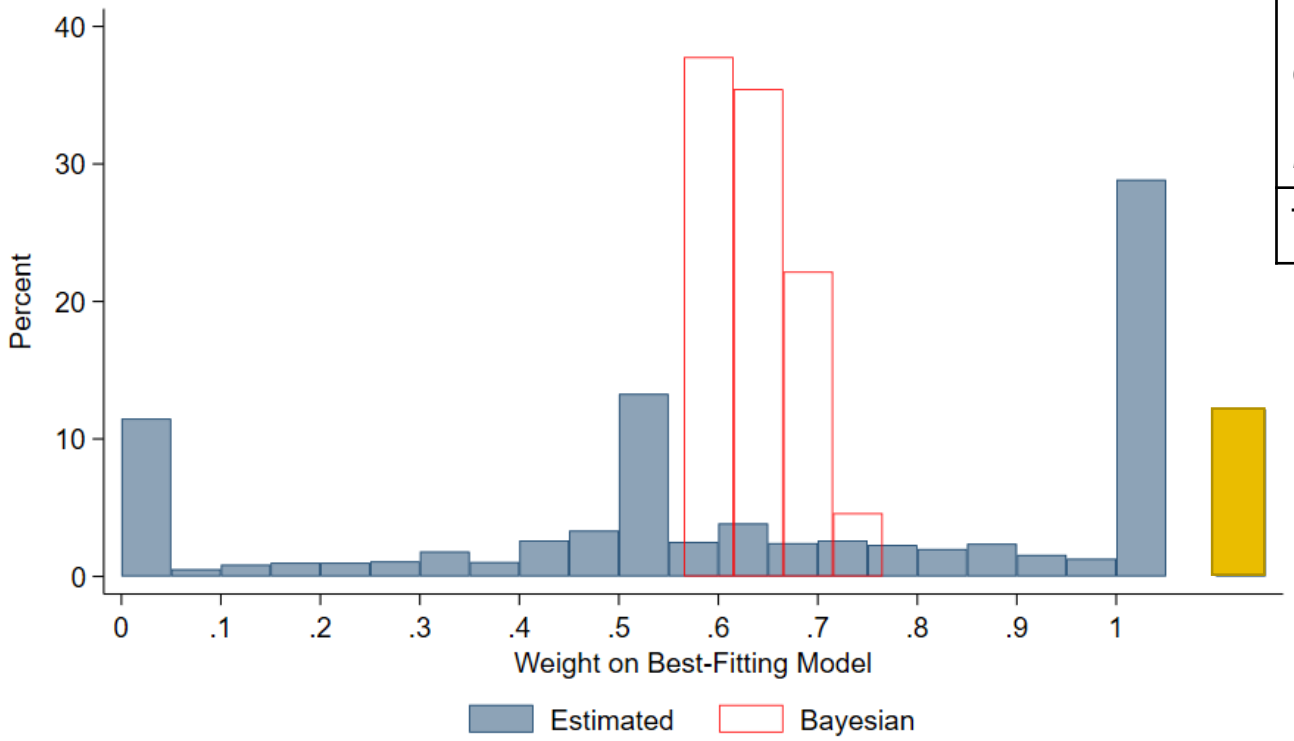
$$\text{Reported } \Pr(A \mid s) = \rho \Pr^{\text{best-fitting}}(A \mid s) + (1-\rho) \Pr^{\text{worst-fitting}}(A \mid s)$$



Predictions	% exact	% within 2 p.p.
Bayesian One-stage Model Selection		
Total		

Reported Guesses

$$\text{Reported } Pr(A \mid s) = \rho\, Pr^{\text{best-fitting}}(A \mid s) + (1-\rho)\, Pr^{\text{worst-fitting}}(A \mid s)$$



Predictions	% exact	% within 2 p.p.
Bayesian	3.05	8.14
One-stage	10.95	18.00
Model Selection	39.81	45.53
Total	53.81	71.67

- Most frequent updating rule: selecting the best-fitting model!

Results

- Quality of Data

1. Analysis of Individual Guesses
2. Consistency of Updating Rules
3. Analysis of Posteriors across Signal Realizations

Consistency Within Participant

# tasks	Bayesian	One-stage	Model Selection	Prior
0	67.67	56.33	33.00	49.00
1	21.00	18.33	13.00	39.00
2	7.00	8.67	6.00	8.67
3	1.67	4.67	4.00	1.67
4	0.33	1.33	4.33	0.33
5	0.33	0.67	3.67	0.33
6	0.00	2.33	6.00	0.33
7	2.00	7.67	30.00	1.67
Total	100.00	100.00	100.00	100.00

more

Consistency Within Participant

# tasks	Bayesian	One-stage	Model Selection	Prior
0	67.67	56.33	33.00	49.00
1	21.00	18.33	13.00	39.00
2	7.00	8.67	6.00	8.67
3	1.67	4.67	4.00	1.67
4	0.33	1.33	4.33	0.33
5	0.33	0.67	3.67	0.33
6	0.00	2.33	6.00	0.33
7	2.00	7.67	30.00	0.67
Total	100.00	100.00	100.00	100.00

Consistency Within Participant

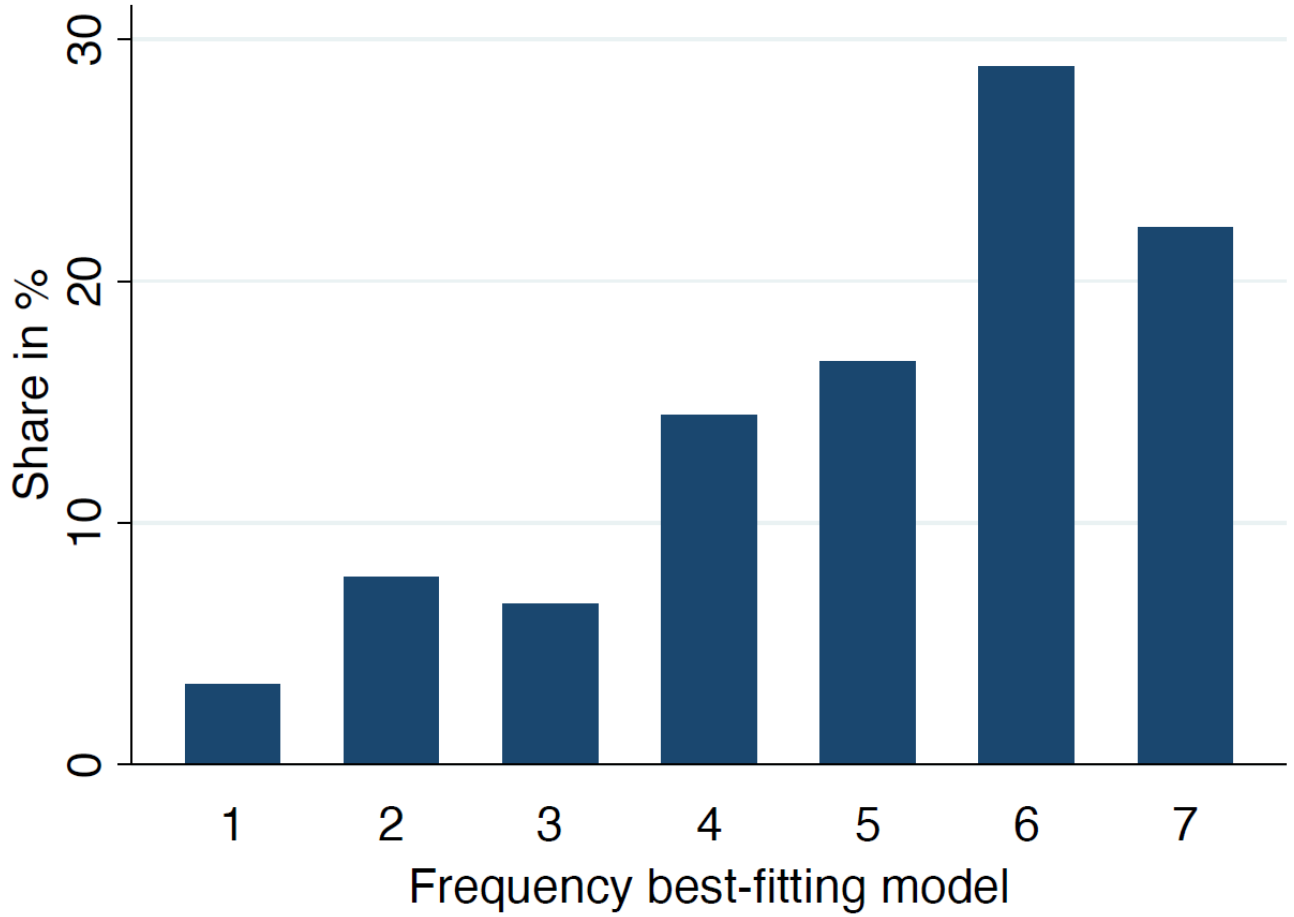
# tasks	Bayesian	One-stage	Model Selection	Prior
0	67.67	56.33	33.00	49.00
1	21.00	18.33	13.00	39.00
2	7.00	8.67	6.00	8.67
3	1.67	4.67	4.00	1.67
4	0.33	1.33	4.33	0.33
5	0.33	0.67	3.67	0.33
6	0.00	2.33	6.00	0.33
7	2.00	7.67	30.00	0.67
Total	100.00	100.00	100.00	100.00

Consistency Within Participant

# tasks	Bayesian	One-stage	Model Selection	Prior
0	67.67	56.33	33.00	49.00
1	21.00	18.33	13.00	39.00
2	7.00	8.67	6.00	8.67
3	1.67	4.67	4.00	1.67
4	0.33	1.33	4.33	0.33
5	0.33	0.67	3.67	0.33
6	0.00	2.33	6.00	0.33
7	2.00	7.67	30.00	0.67
Total	100.00	100.00	100.00	100.00

factors | first-last

Mistakes in Model Selection via ML



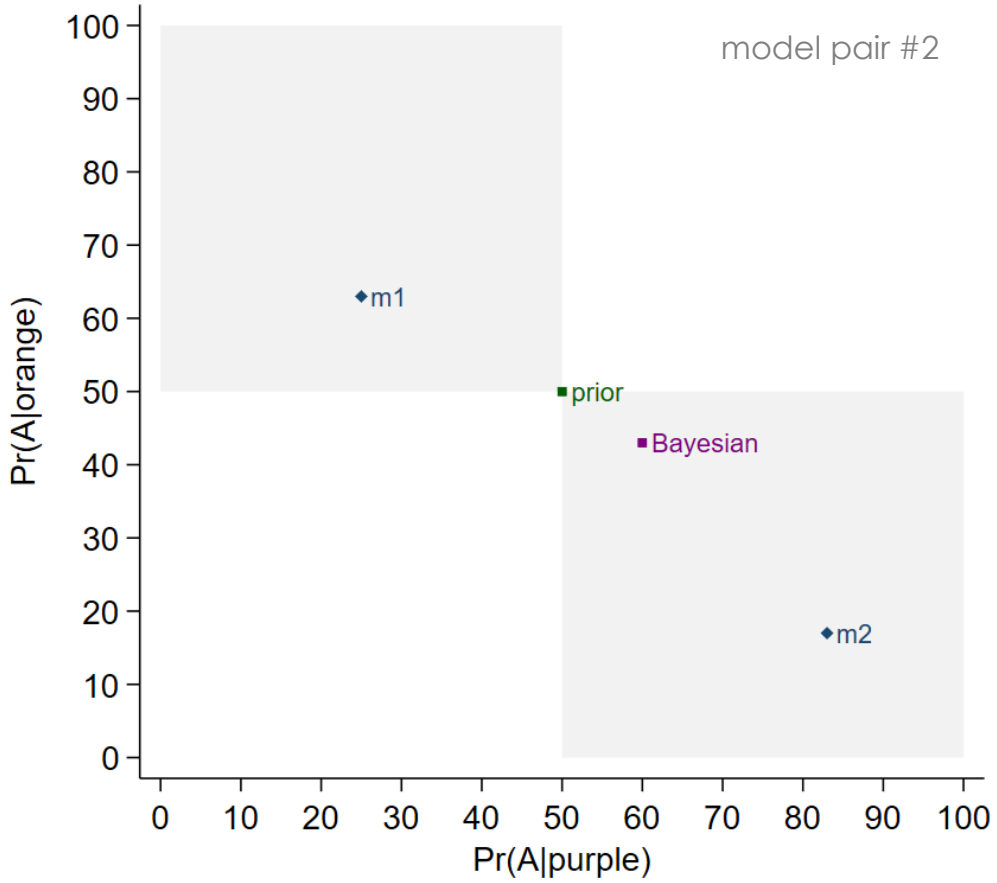
more

Results

- Quality of Data

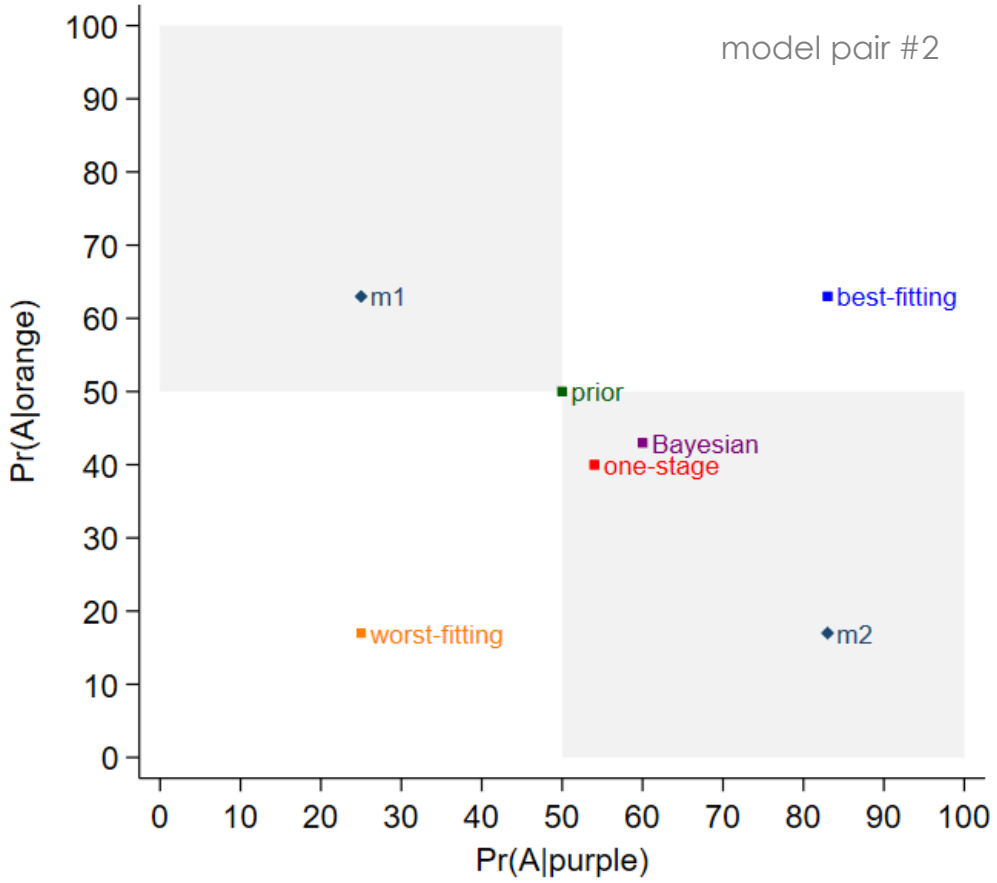
1. Analysis of Individual Guesses
2. Consistency of Updating Rules
3. Analysis of Posteriors across Signal Realizations

Vectors of Posteriors



- Vectors of posteriors from repeated model pairs
- Each model corresponds to a vector of posteriors (posteriors across signal realizations)
- Bayesian reasoning → Bayes-consistent areas (gray)
- Bayes-inconsistent = becoming more confident about one state regardless of what it is observed

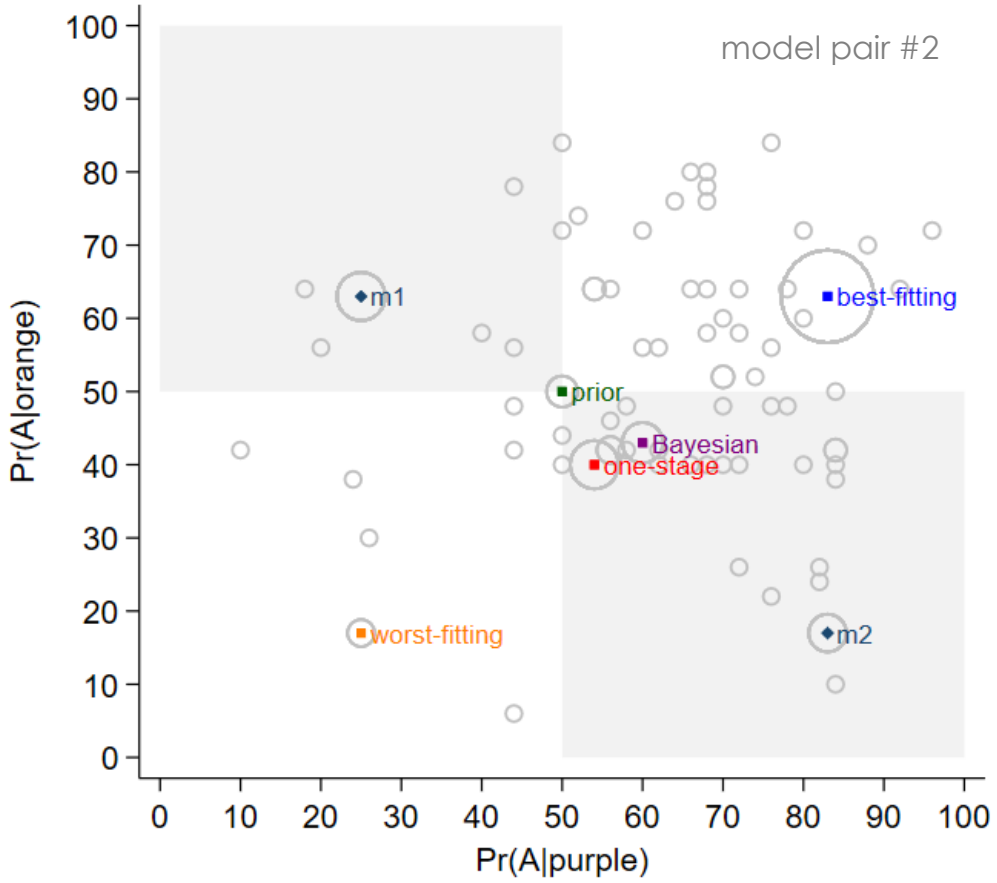
Vectors of Posteriors



- Vectors of posteriors from repeated model pairs
- Each model corresponds to a vector of posteriors (posteriors across signal realizations)
- Bayesian reasoning → Bayes-consistent areas (gray)
- Bayes-inconsistent = becoming more confident about one state regardless of what it is observed

  → Always for selecting the best-fitting model!

Vectors of Posteriors



- Similar distribution as before
  - Bayesian: 3%
  - One-stage:10%
  - Model selection: 37%
- Bayes-inconsistent vectors: 50%

Additional Results

- Explore other criteria for model selection
- No evidence of difference in updating rules with experience
- No evidence of other updating rules consistently used in the data
  - Use of constant model weights across signals
  - Biases in updating about the state given a model
  - Combination of rules across tasks
- Less extreme form of over- and underinference about the models
- Determinants of Model Selection vs Model Weighting
- Robust to other updating environments

selection
first-last
factors
more

Conclusion

With competing models, how do we update beliefs?

- Support for theoretical assumptions, but need to account for heterogeneity
  - Bayes-inconsistencies highlights the power of persuasion using models
- Many systematically select a model
  - Most select the best-fitting model (overinference about models)
  - …but sometimes make mistakes
- Other weights models, mostly in a biased way (underinference about models)
- Individuals consistently use these different updating rules

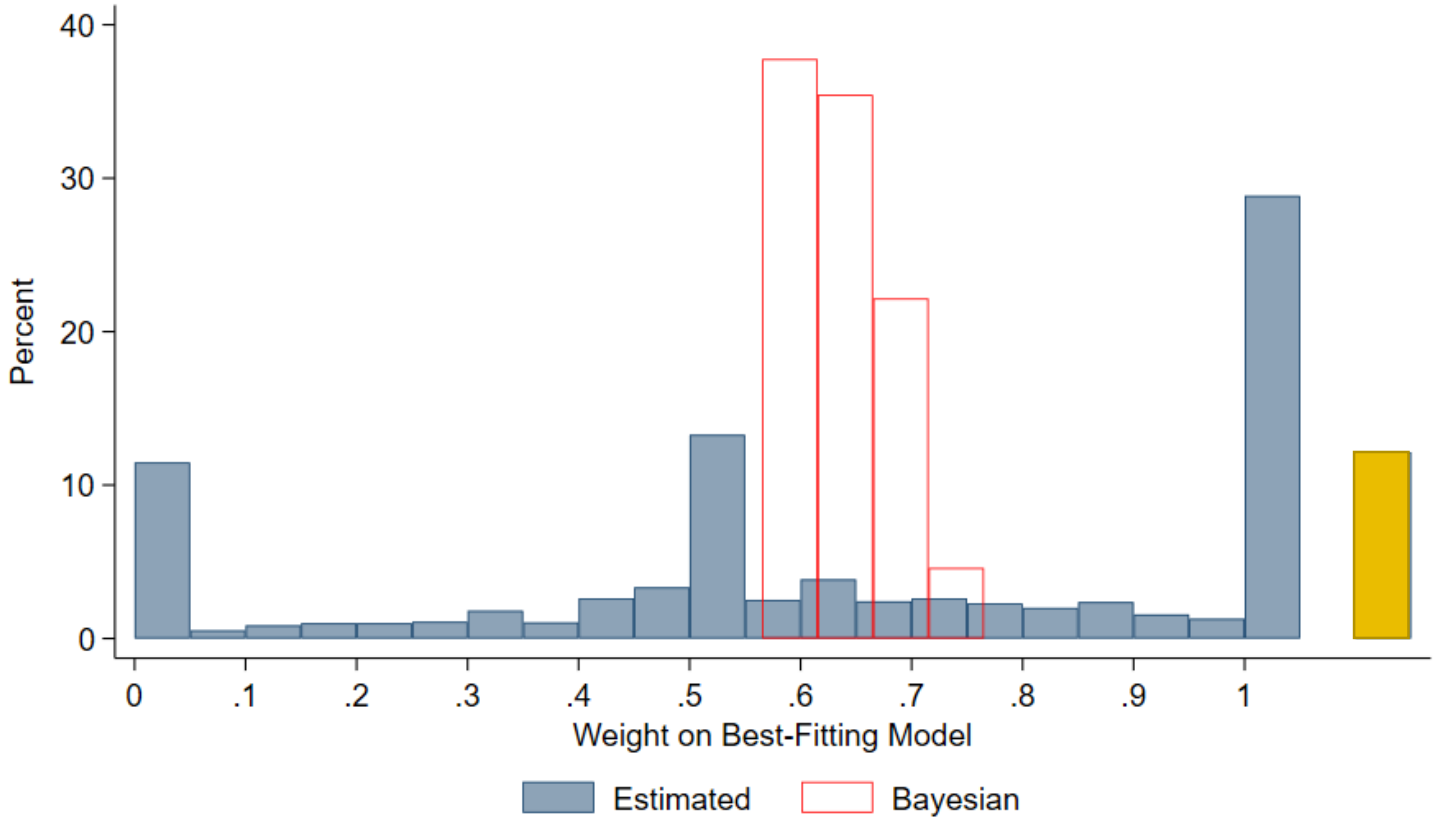
Thank you!

chiara.aina@upf.edu

Appendix

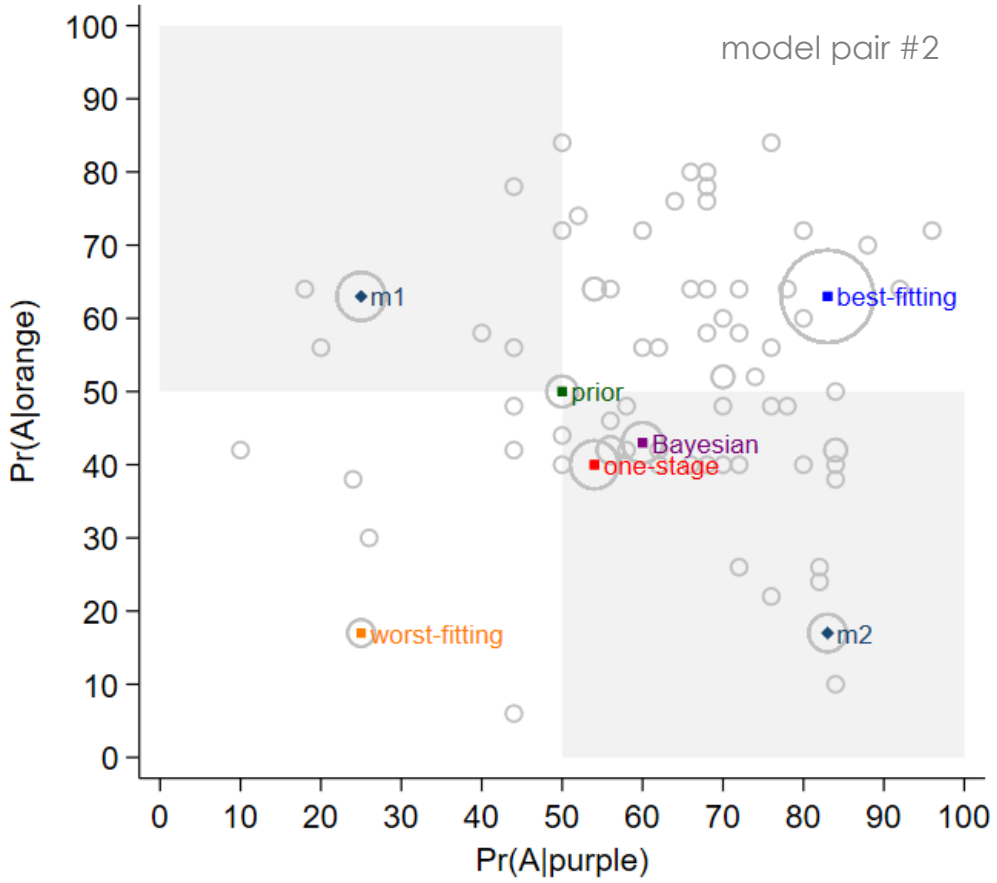
Reported Guesses

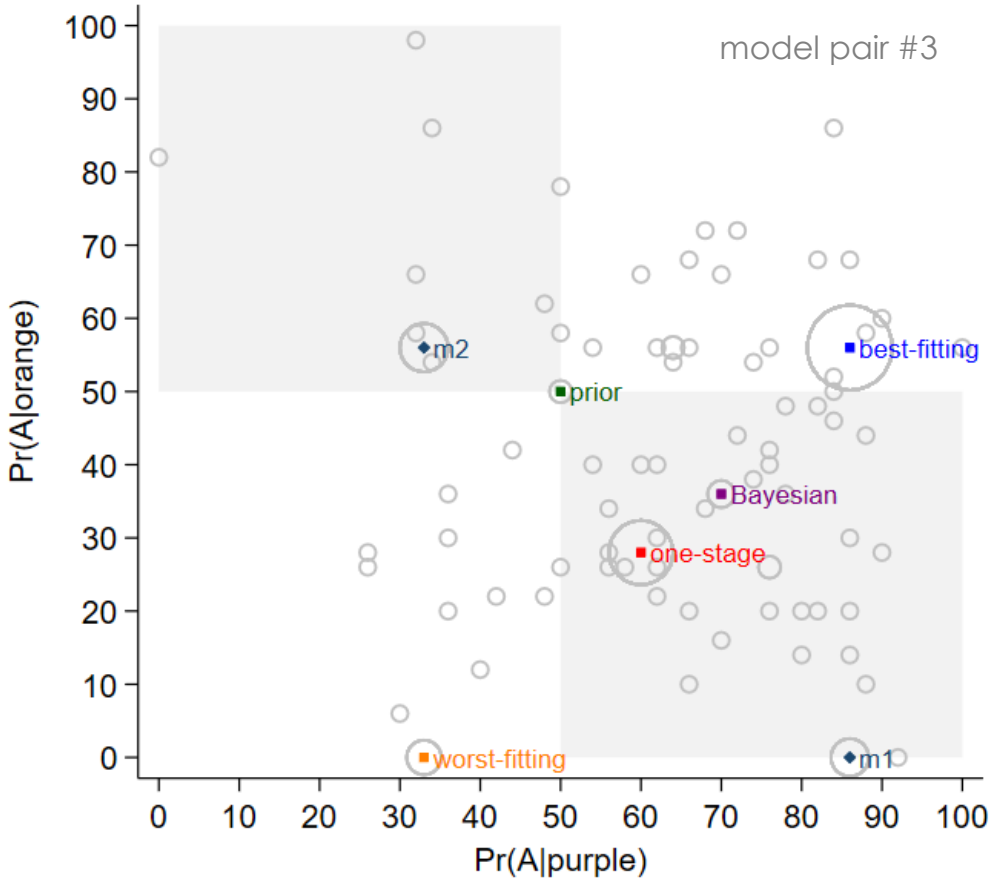
$$\text{Reported } \Pr(A \mid s) = \rho \Pr^{\text{best-fitting}}(A \mid s) + (1-\rho) \Pr^{\text{worst-fitting}}(A \mid s)$$



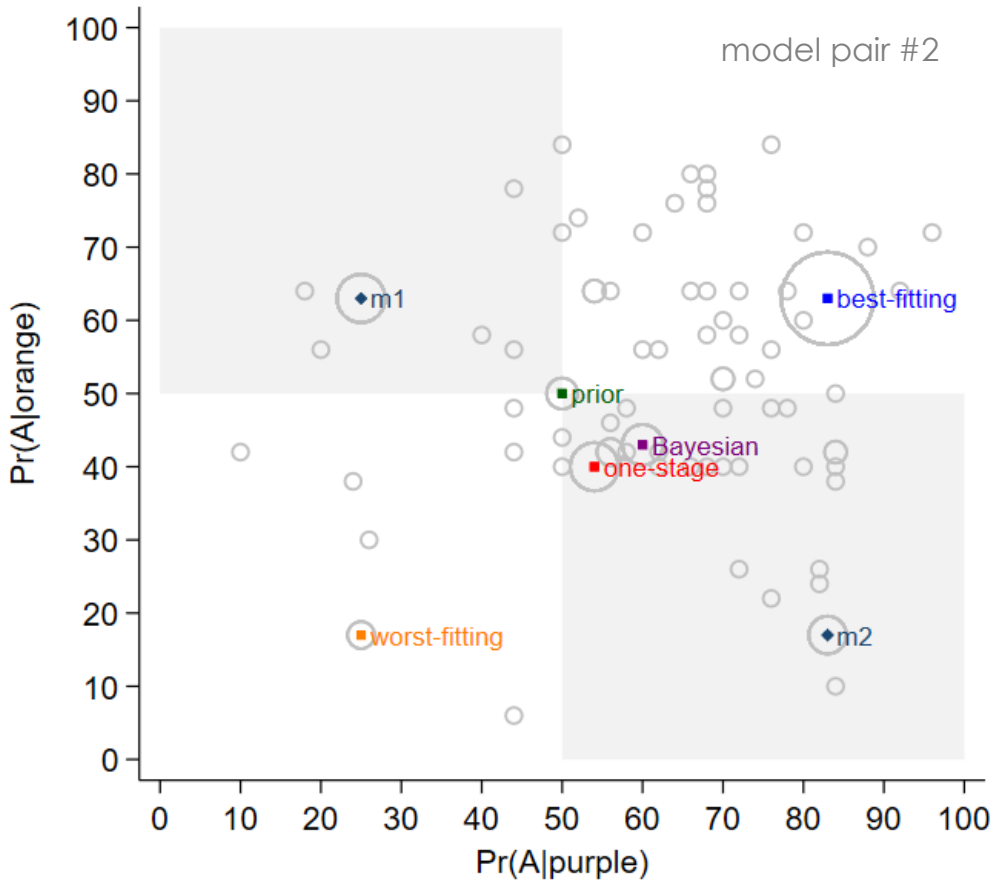
Predictions	% exact	% within 2 p.p.
Bayesian	3.05	8.14
One-stage	10.95	18.00
Best-fitting	28.86	33.10
Worst-fitting	10.95	12.43
Within	35.90	
Outside	12.14	
Total	100.00	71.67

Vectors of Posteriors





Vectors of Posteriors



- Use repeated tasks (pooled data)

Predictions	% exact	% within 2 p.p.
Bayesian	2.67	3.33
One-stage	6.00	9.67
Best-fitting	19.67	23.33
Worst-fitting	2.33	2.67
Model 1	4.33	5.33
Model 2	4.67	6.00
Prior	2.00	2.00
Total	41.67	52.33

more alpha

Related Literature

1. Narratives, formalized in multiple ways

- Models: Schwartzstein & Sunderam (2021, 2024), Aina (2024), Ichihashi & Meng (2021), Ispano (2024), Yang (2023), Jain (2024), Barron and Fries (2024) → use this approach
- DAGs: Eliaz & Spiegler (2020), Eliaz, Spiegler, & Weiss (2021), Eliaz, Galperti, & Spiegler (2022), Horz & Kocak (2022), Charles and Kendall (2023), Ambuehl & Thysen (2024), Frechette et al. (2024)
  - These papers assume subjects select one narrative
  - Persuasion via models is effective(Barron and Fries, 2023) → not about persuasion
  - Complement papers on DAGs (Charles and Kendall, 2023; Ambuehl & Thysen, 2024)
  - Others: e.g., Levy & Razin (2021), Izzo, Martin, & Callander (2021), Olea et al. (2022), Szeidl & Szucs (2024), Graeber et al. (2022), Bursztyn et al. (2022), Andre et al. (2022, 2021), …

2. Belief Updating, e.g., Benjamin (2019), Augenblick et al., (2024), Ba et al. (2024), Bordalo et al. (2024)

- Learning from uncertain information: Liang (2023) → not about weights

back

ML Selection

- Assumed by Schwartzstein & Sunderam (2021, 2023), Aina (2024), …
- Similar to "Inference to the Best Explanation" (Harman, 1965)
  only the best hypothesis is used to make inference
  - Hypotheses are supported by the same observations they are supposed to explain (Lipton, 2003; Keil, 2006; Douven et al., 2015)
  - The better a hypothesis explains the data, the more confidence in it (Koehler, 1991; Pennington and Hastie, 1992; Lombrozo and Carey, 2006)
- Ambiguity-aversion: under ambiguity aversion, learning efficiency is maximal with Maximum Likelihood updating (Frick et al., 2022)
- Bayesian statistics: model comparison based on the model (marginal) likelihood, also known as model evidence

back

Bayesian Updating

Two ways to calculate the Bayesian prediction:

1. Weighting models by their posterior given the signal

$$\Pr(G \mid s) = \Pr(m1 \mid s)\, \Pr^{m1}(G \mid s) + \Pr(m2 \mid s)\, \Pr^{m2}(G \mid s)$$

2. Bayes' rule with reduced model, derived by combining the models using the prior

$$\Pr(G \mid s) = \Pr^{m'}(G \mid s)$$

$$\Pr^{m'}(s \mid G) = \Pr(m1)\, \Pr^{m1}(s \mid G) + \Pr(m2)\, \Pr^{m2}(s \mid G)$$

$$\Pr^{m'}(s \mid B) = \Pr(m1)\, \Pr^{m1}(s \mid B) + \Pr(m2)\, \Pr^{m2}(s \mid B)$$

back

Bayesian Updating

Model predictions weighted by the models' posterior given the signal & models' prior

$$\rho_s^m = \Pr(m \mid s)$$

For each model m, use observed signal s to learn about states and models:

a. Posterior of the state given the signal via Bayes rule: $\Pr^m(G \mid s)$

b. Posterior of the model given the signal via Bayes rule given models' prior: $\Pr(m \mid s)$

- Equivalent to update using a "reduced model"

Bayesian Updating

models

$Pr^{m1}(H \mid G) = 100\%$

$Pr^{m1}(H \mid B) = 0\%$

$Pr^{m2}(H \mid G) = 17\%$

$Pr^{m2}(H \mid B) = 67\%$

model predictions

$Pr^{m1}(G \mid H) = 100\%$

$P^{m2}(G \mid H) = 20\%$

Ex-ante equally likely models
Pr(m1)=Pr(m2)=50%

model posteriors

$Pr(m1 \mid H) = 55\%$

$Pr(m2 \mid H) = 45\%$

$$Pr^{Bayes}(G \mid H) = 55\% \cdot 100\% + 45\% \cdot 20\% = 64\%$$

G= good asset, B = bad asset
H = high signal, L = low signal

back

One-stage Updating

Model predictions weighted by the models' prior

$$\rho_s^m = \Pr(m)$$

- Extreme form of underinfence about models:
  recognize that the signal is informative about the states,
  but fail to update beliefs about models

back

One-stage Updating

models

$Pr^{m1}(H \mid G) = 100\%$

$Pr^{m1}(H \mid B) = 0\%$

$Pr^{m2}(H \mid G) = 17\%$

$Pr^{m2}(H \mid B) = 67\%$

model predictions

$Pr^{m1}(G \mid H) = 100\%$

$P^{m2}(G \mid H) = 20\%$

Ex-ante equally likely models
Pr(m1)=Pr(m2)=50%

model posteriors

$Pr(m1 \mid H) = 55\%$

$Pr(m2 \mid H) = 45\%$

$$Pr^{\text{one-stage}}(G \mid H) = 50\%\ 100\% + 50\%\ 20\% = 60\%$$

G= good asset, B = bad asset
H = high signal, L = low signal

back

Model Selection

Model Selection via Maximum Likelihood

Select the best-fitting model & only form posteriors using that model

$$m^* \in \text{argmax } Pr^m(s) = Pr(G)\ Pr^m(s \mid G) + Pr(B)\ Pr^m(s \mid B)$$

$$\rho_s^{m^*} = 100\%$$

- Extreme form of overinference about models!
- Other form of model selection:
  - Stochastic Maximum Likelihood Selection: sometimes making errors by selecting the worst-fitting model
  - Model Selection based on Informativeness: select the most informative model, inspired by Yang (2024)
  - Dogmatic Model Selection: tendency to be always drawn to the same models, inspired by Ba (2024)

back

ML Selection

models

$Pr^{m1}(H \mid G) = 100\%$

$Pr^{m1}(H \mid B) = 0\%$

$Pr^{m2}(H \mid G) = 17\%$

$Pr^{m2}(H \mid B) = 67\%$

model predictions

$Pr^{m1}(G \mid H) = 100\%$

$P^{m2}(G \mid H) = 20\%$

model fit levels

$Pr^{m1}(H) = 50\%$

$\rightarrow \rho_H^{m1} = 100\%$

$Pr^{m2}(H) = 42\%$

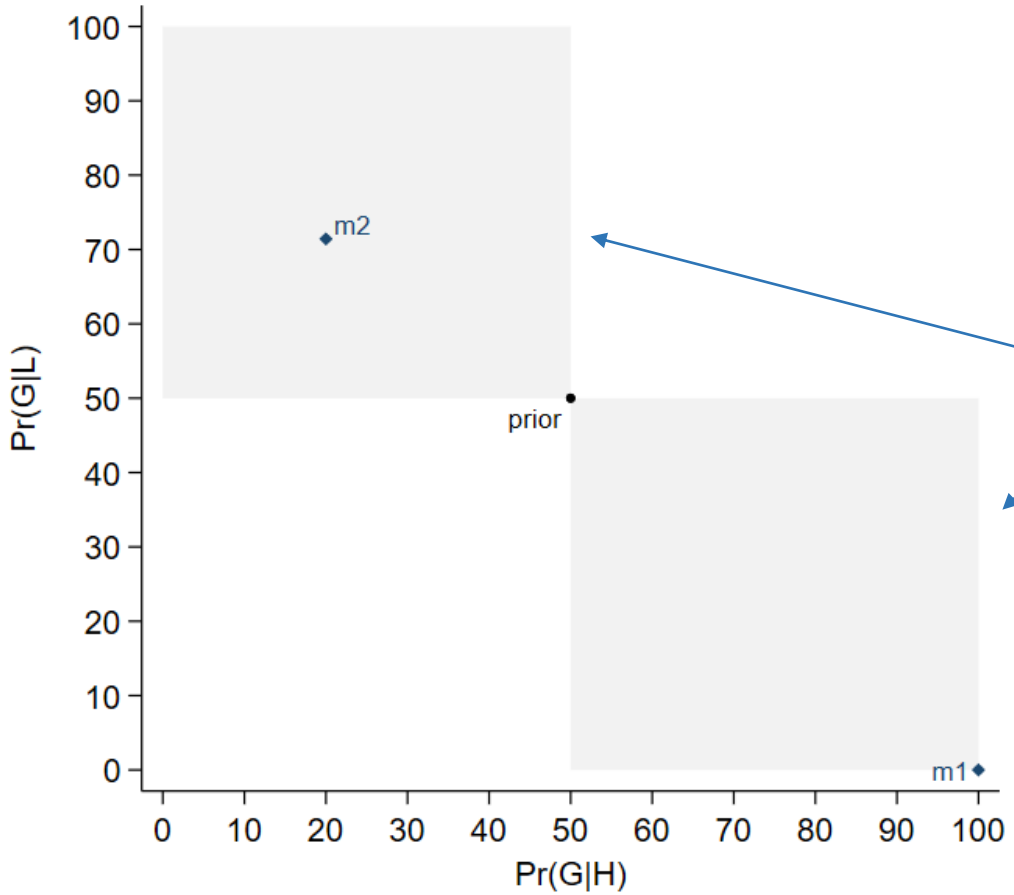
$\rightarrow \rho_H^{m2} = 0\%$

$$Pr^{ML}(G \mid H) = 100\%\ 100\% + 0\%\ 20\% = 100\%$$

G= good asset, B = bad asset
H = high signal, L = low signal

back

Vectors of Posterior Beliefs

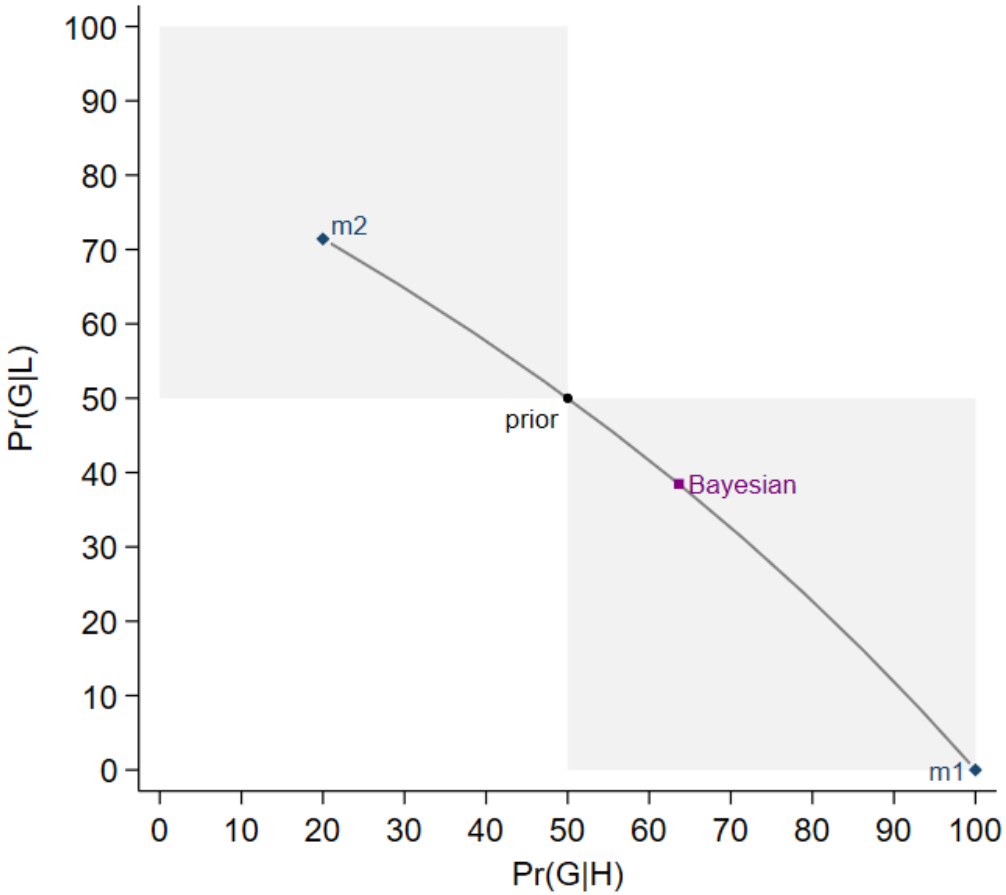


Bayes-Consistent Vectors

G= good asset, B = bad asset
H = high signal, L = low signal

back

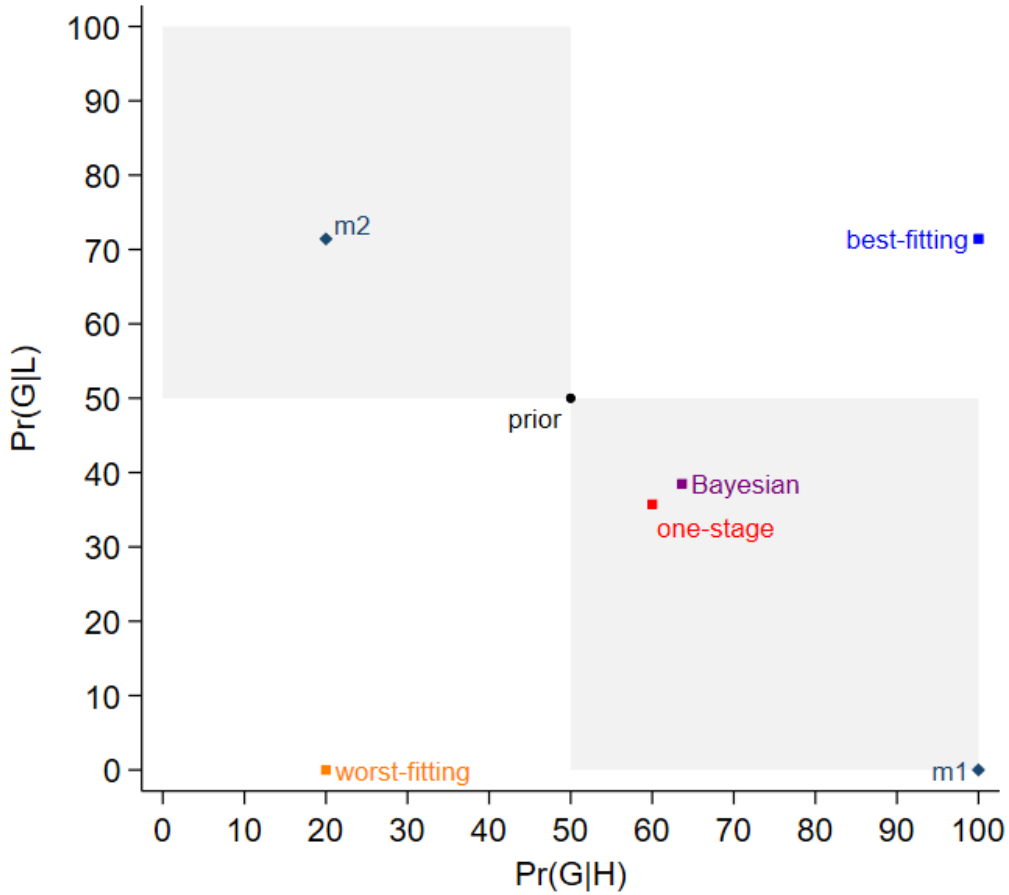
Vectors of Posterior Beliefs



G= good asset, B = bad asset
H = high signal, L = low signal

back

Vectors of Posterior Beliefs

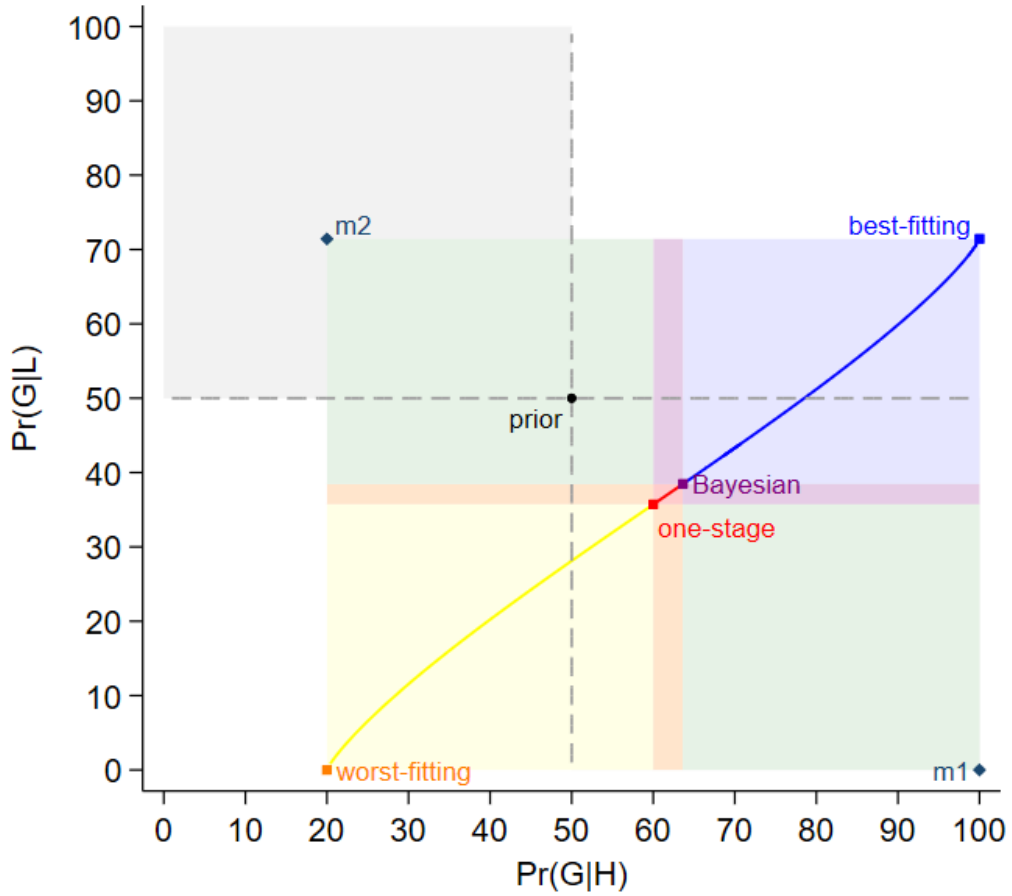


Pr(m1) = Pr(m2) = 50%

G= good asset, B = bad asset
H = high signal, L = low signal

back alpha

Biases in Model Weighting



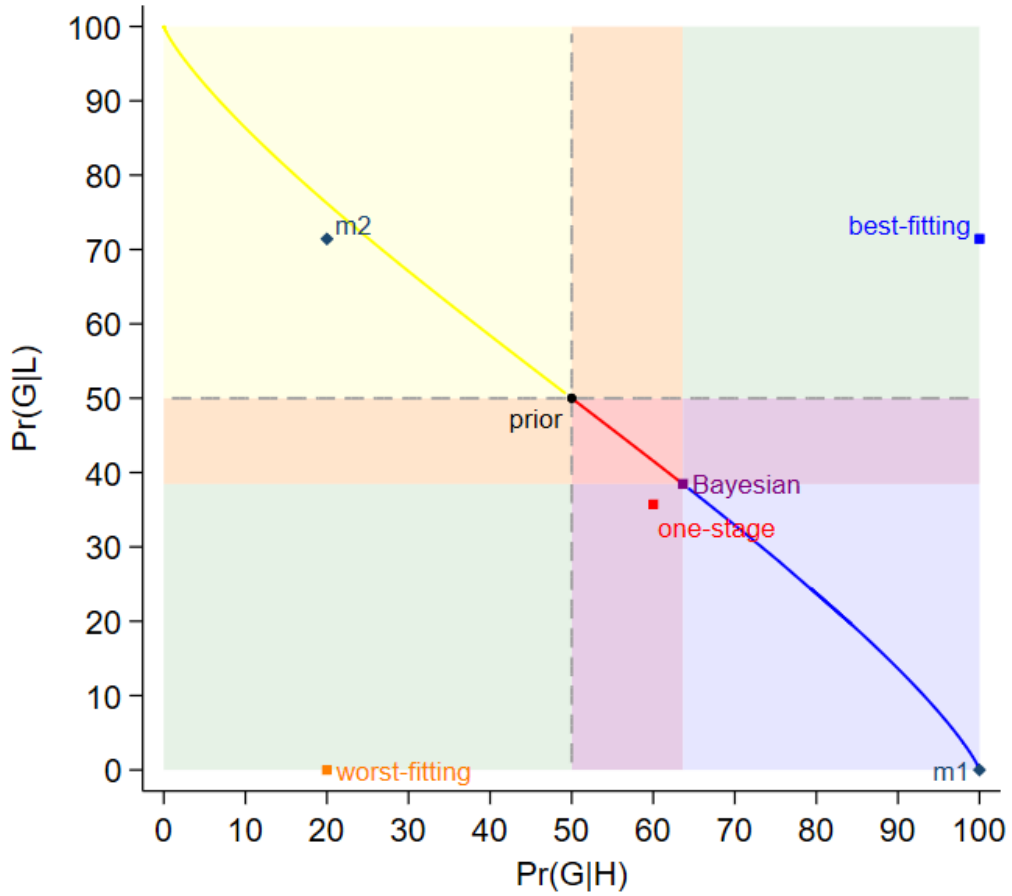
Grether (1980)'s approach to learning about the models:

$$\rho_s^m = \frac{1}{1 + \left(\frac{\Pr(s|m')}{\Pr(s|m)}\right)^{\alpha_m}}$$

- $\alpha_m = 1$ Bayesian updating
- $\alpha_m \in [0,1]$ Underinference about models
  w/ $\alpha_m = 0$ One-stage updating
- $\alpha_m > 1$ Overinference about models
  w/ $\alpha_m \to \infty$ ML selection
- $\alpha_m < 0$ Inference in the wrong direction
  w/ $\alpha_m \to -\infty$ worst-fitting

alpha back

Biases in Updating about the States



Grether (1980)'s approach to learning about the states:

$$\hat{\Pr}(A|s) = \frac{1}{1 + \left(\frac{\Pr(A|s,m')}{\Pr(A|s,m)}\right)^{\alpha_s}}$$

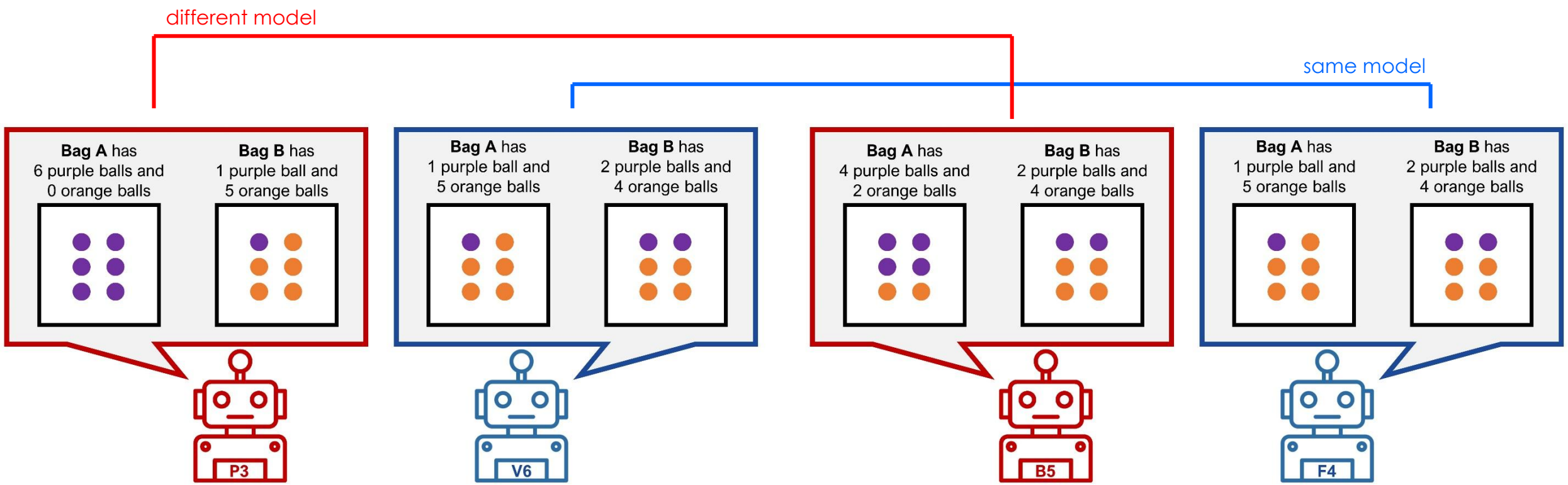
- $\alpha_s = 1$ Bayesian updating
- $\alpha_s \in [0,1]$ Underinference about states

  w/ $\alpha_m = 0$ Prior over states
- $\alpha_s > 1$ Overinference about states

- $\alpha_m < 0$ Inference in the wrong direction

back

Comparison across Model Pairs



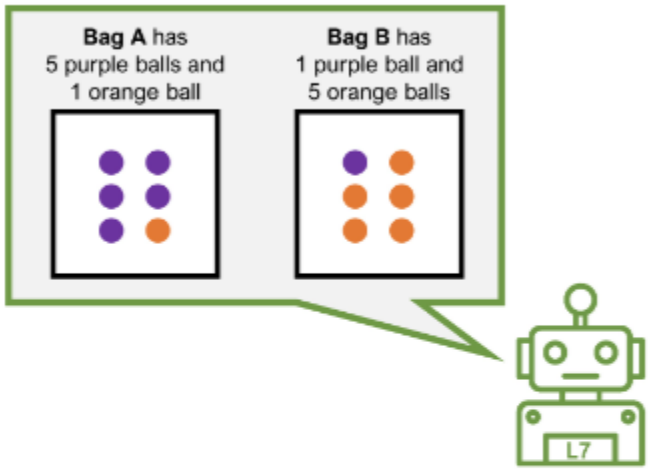
model pair #3

model pair #5

back

Part 1

Remember that for this task we consider robot L7.



Remember, the computer has drawn an orange ball

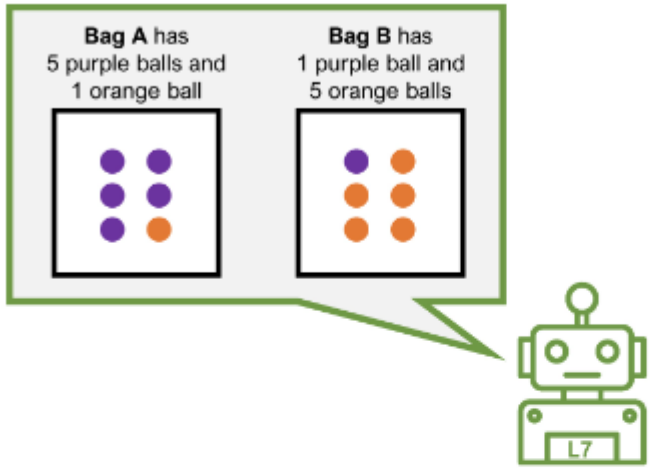
What do you think is the chance that bag A was selected given that the color of the drawn ball is orange? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

back

Part 2

Remember that for this task we consider robot L7.



Remember, the computer has drawn an orange ball

The chance that bag A was selected given that the color of the drawn ball is orange would be **17%**

What do you think is the chance that bag A was selected given that the color of the drawn ball is orange? Make your guess using the slider

Chance of **bag A (in %): Please click on the slider**

back

Logistics

- Data collection on Prolific (April 2024)
- Pre-registered
- 300 participants
  - Part 3: 7 x 300 = 2,100 obs
- Avg payment 7.5$ for an avg duration of 38 min
  - Completion fee 6$
  - Bonus 2$ for a randomly chosen task
- Belief elicitation using intuitive description (Danz et al., 2022)
  - Binarizing scoring rule (Hossain & Okui, 2013)
    the higher the reported guess to the realized state, the higher prob of the bonus

back

Pairs of Models

# model pair	$Pr^1(p \mid A)$	$Pr^1(p \mid B)$	$Pr^2(p \mid A)$	$Pr^2(p \mid B)$	repeated	notes
1	1/6	2/6	5/6	2/6		
2	1/6	3/6	5/6	1/6	yes	
3	6/6	1/6	1/6	2/6	yes	share model w/ #5
4	1/6	1/6	0/6	6/6		fully info vs uninformative
5	4/6	2/6	1/6	2/6		share model w/ #3

back design back

Part 4 – Questionnaire

- 3 unincentivized but contextualized tasks similar to Part 3
- Selection of items to determine thinking style (Pacini and Epstein, 1999)
- Cognitive reflection (Frederick, 2005), modified version
- Self-reported choice rule to make guesses
- Self-reported level of challenge
- Demographics: gender, age, and political affiliation

back

Mistakes in Part 2

# Mistakes for subject in Part 2	%
0	79.33%
1	11.67%
2	9.00%

- Mistake if reported guess not in 2p.p. window of given Bayesian prediction
- Correlation btw #mistakes in Part 2 & absolute bias (= asb distance btw reported guess and Bayesian prediction) in Part 3 (p-value = 0.002)

back design back

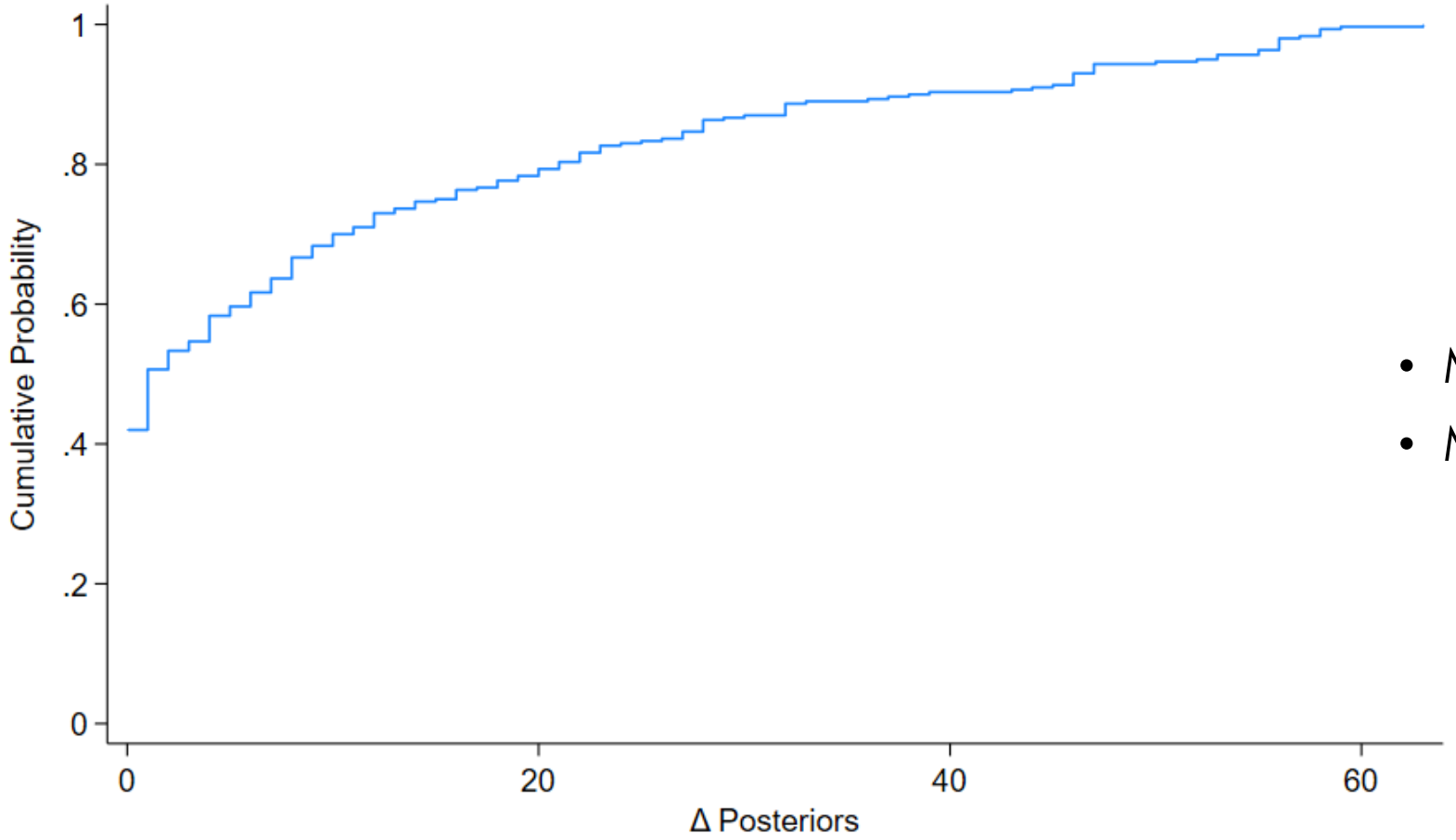
Quality of Data

1. Consistency – in the same situation, do participants report the same guesses?
   - Use repeated model pairs and observe the same signal (N=300)
   - 42% same guess, 53% within 2p.p.

2. Across model pairs – does aggregate data react to making a model more extreme?
   - Comparison across model pairs w/ one model in common
   - Requires 2 tests: one for each signal
   - Avg belief is more extreme in the model pair w/ the more extreme mode

back
more more part 2

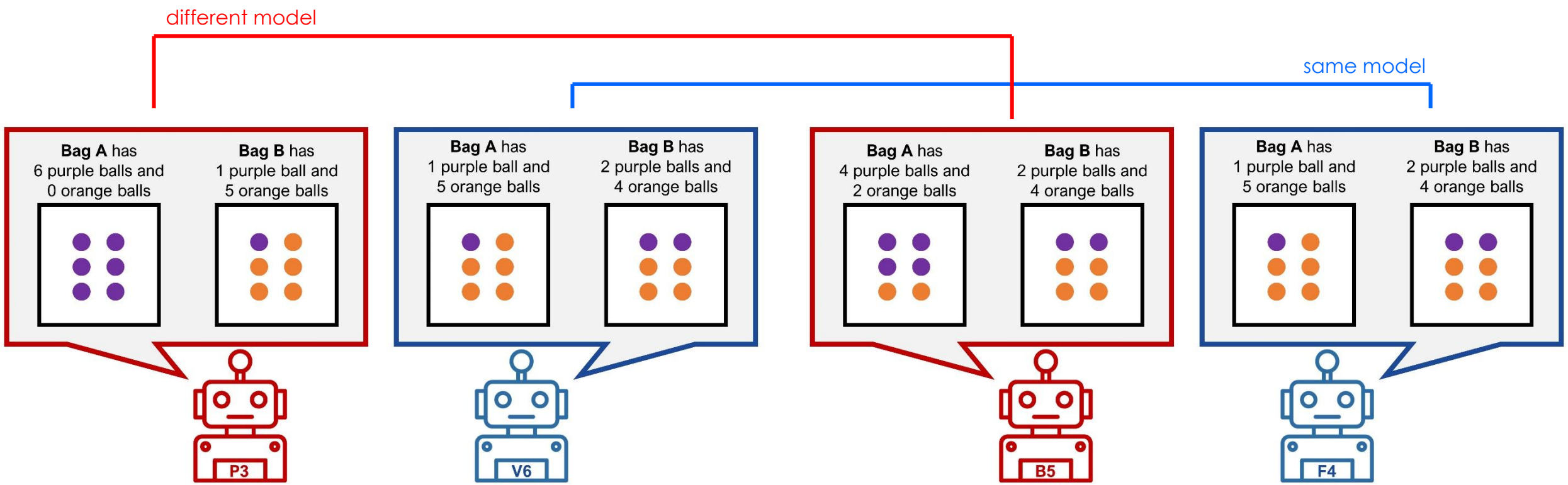
Consistency – Δ Posteriors



- Mean Δ Posterior: 11%
- Median Δ Posterior: 1%

back

Comparison across Model Pairs

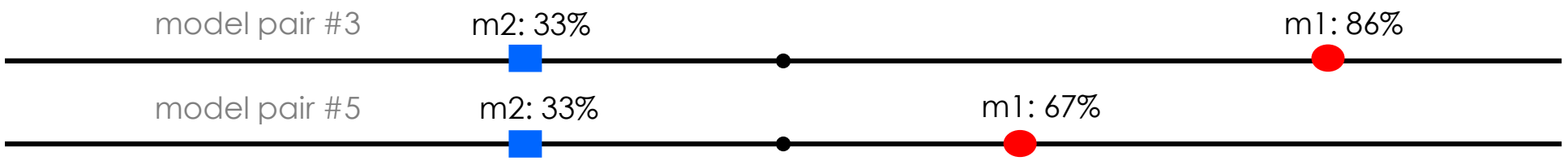


model pair #3

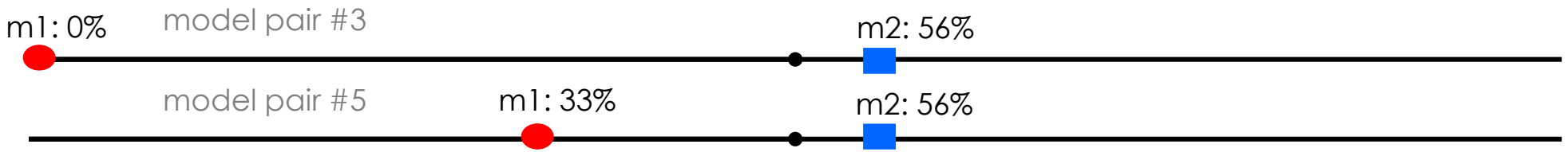
model pair #5

back

Comparison across Model Pairs



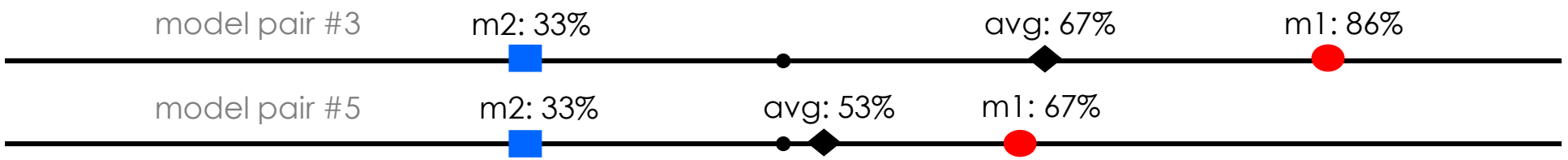
- Given a **purple ball**, expect a higher guess in #3 than #5



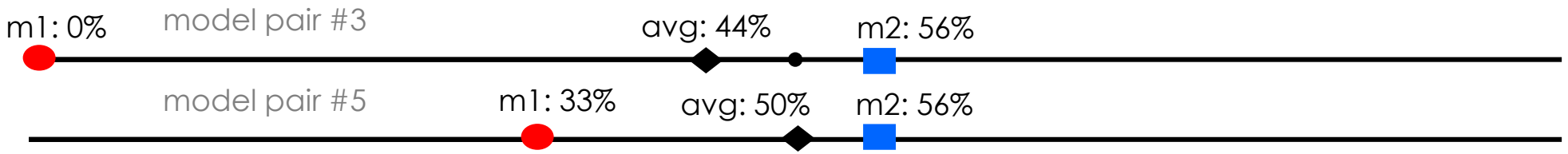
- Given an **orange ball**, expect a lower guess in #3 than #5

back

Comparison across Model Pairs



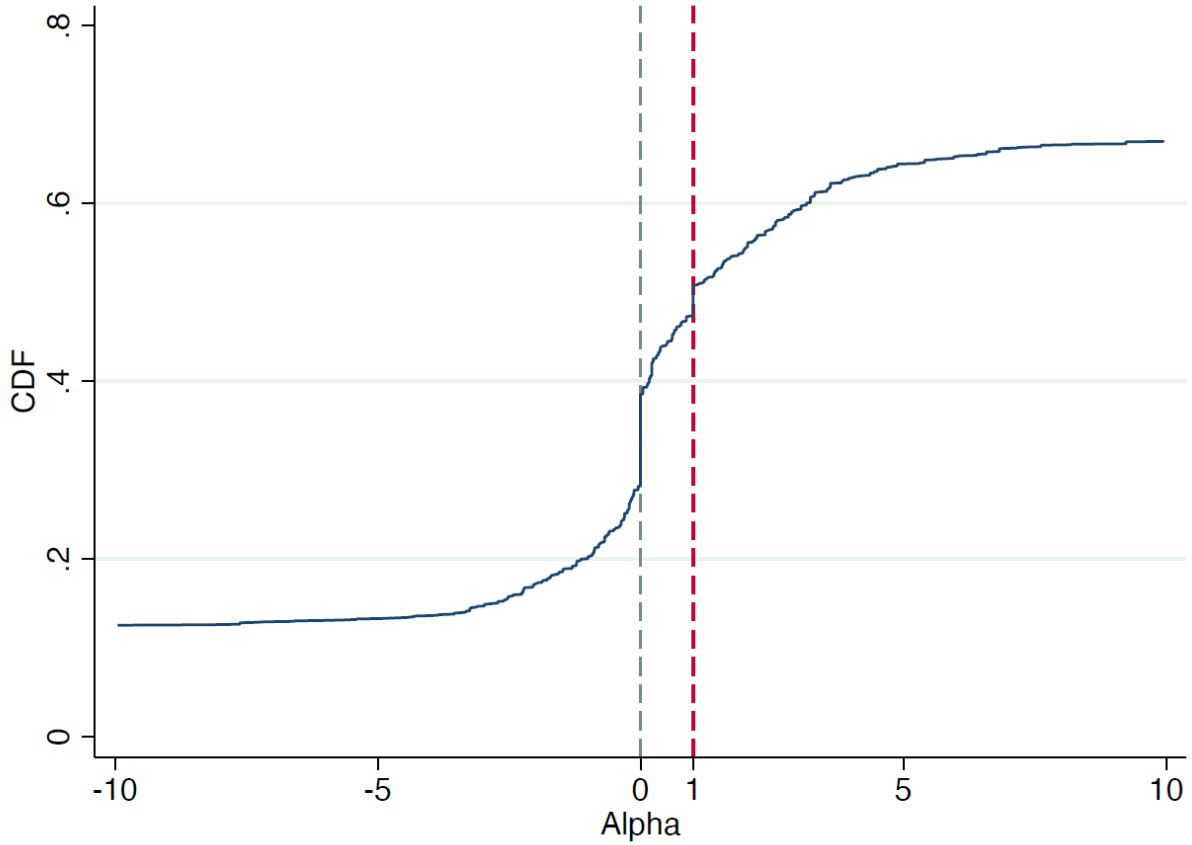
- Given a **purple ball**, expect a higher guess in #3 than #5 (p-value<0.001)
  - Avg #3: 67% (N=117)
  - Avg #5: 53% (N=105)



- Given an **orange ball**, expect a lower guess in #3 than #5 (p-value=0.001)
  - Avg #3: 44% (N=183)
  - Avg #5: 50% (N=195)

back

Cumulative Distribution of α_m



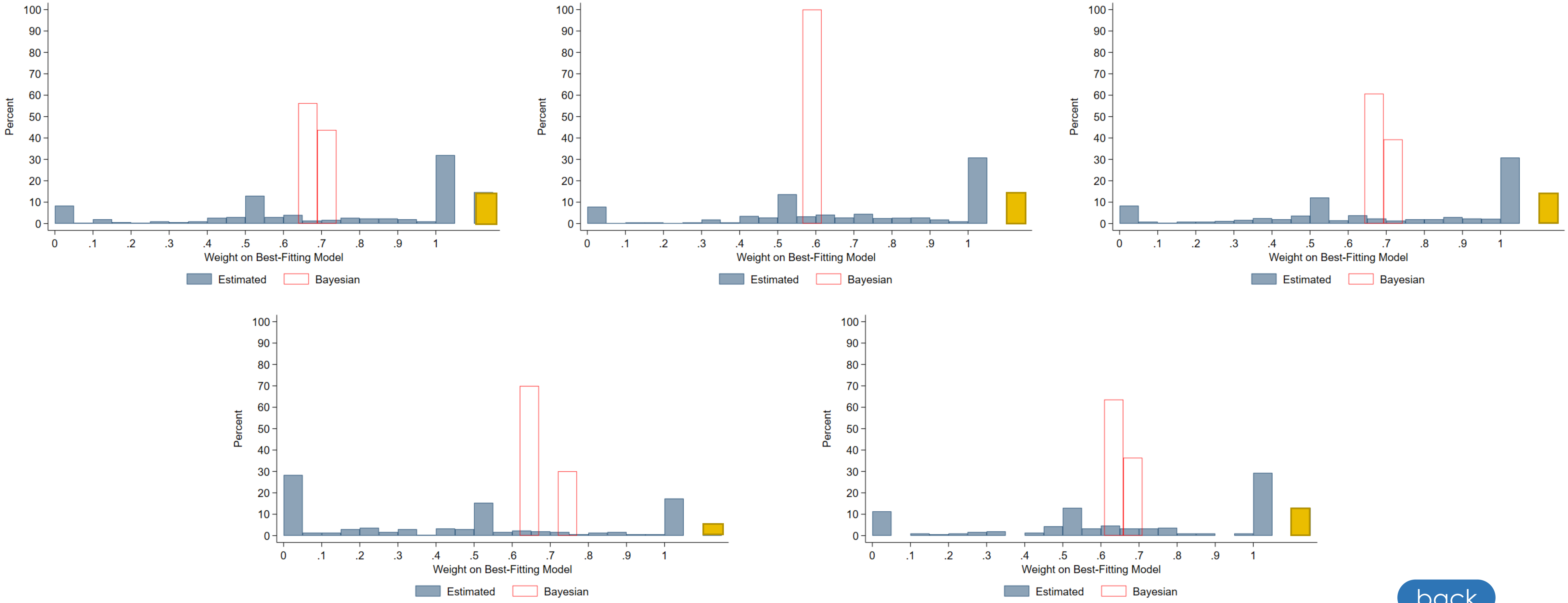
Grether (1980)'s approach to learning about the models:

$$\rho_s^m = \frac{1}{1 + \left(\frac{\Pr(s|m')}{\Pr(s|m)}\right)^{\alpha_m}}$$

- $\alpha_m = 1$ Bayesian updating
- $\alpha_m \in [0,1]$ Underinference about models
  w/ $\alpha_m = 0$ One-stage updating
- $\alpha_m > 1$ Overinference about models
  w/ $\alpha_m \to \infty$ ML selection
- $\alpha_m < 0$ Inference in the wrong direction
  w/ $\alpha_m \to -\infty$ worst-fitting

from vectors
framework
back

Weights by Model Pair



back

Type of Guesses: Consistent Participants

Predictions	% exact	% within 2.p.p.
Bayesian	3.15	5.90
One-stage	14.42	22.80
Best-fitting	42.86	47.84
Worst-fitting	12.32	13.37
Within	19.66	
Outside	7.60	
Total	100.00	88.21

- 2 p.p. 1.7% overlap between Bayesian and One-stage

back

Drivers of Model Selection

	I	II	III	IV	V
Best-fitting	Maximum Likelihood criterion				
Most Informative	Informativeness criterion				
Model Pair 2	Dogmatic Model Selection				
Model Pair 3					
Model Pair 4					
Model Pair 5					
Constant					
Observations	836	836	836	836	836
R2					

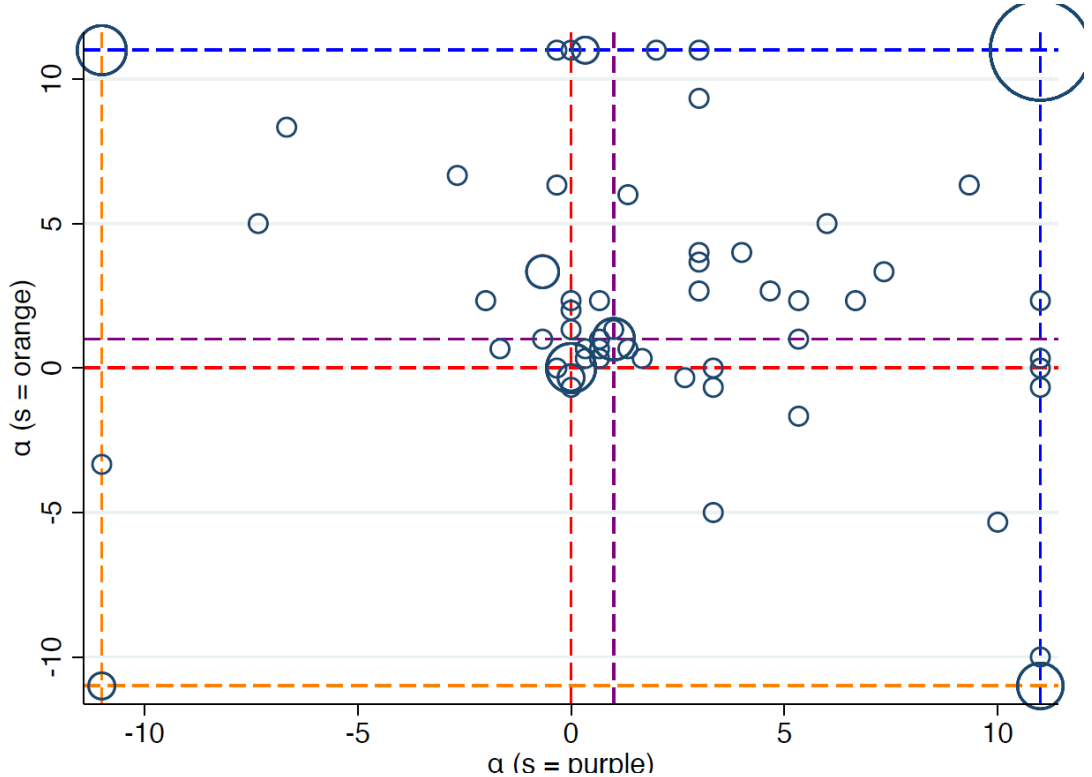
end back

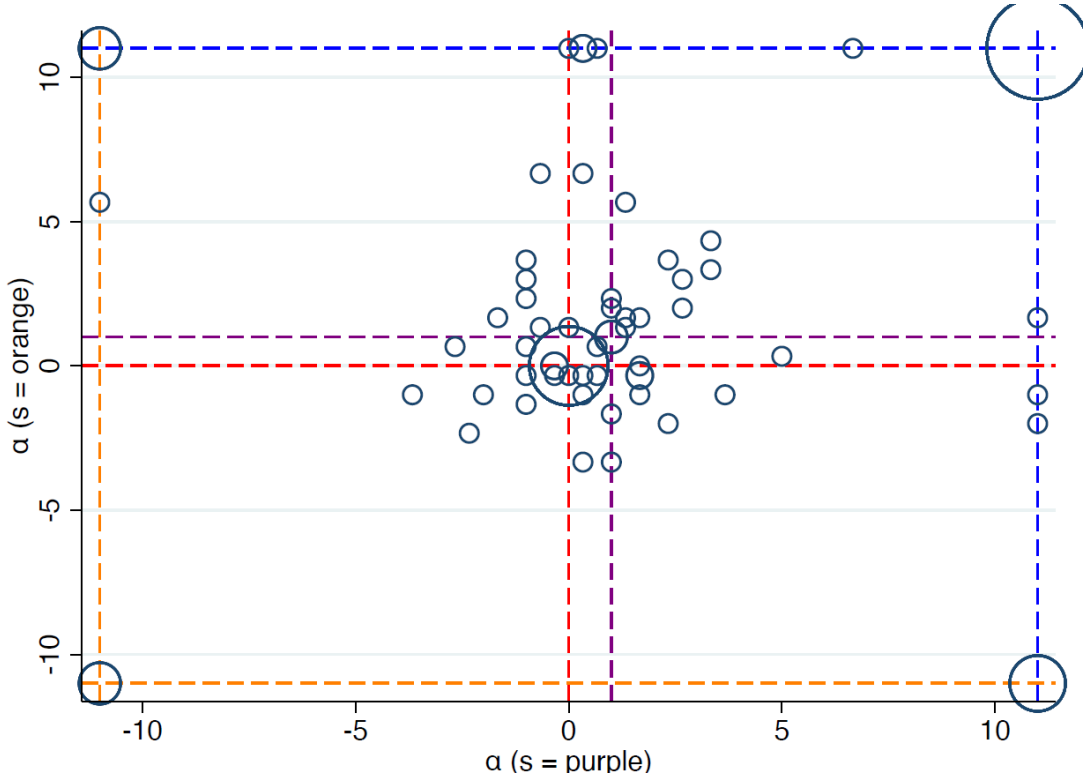
Drivers of Model Selection

	I	II	III	IV	V
Best-fitting	0.450***		0.449***		0.446***
	(0.041)		(0.043)		(0.043)
Most Informative		-0.151***	-0.005		-0.002
		(0.039)	(0.035)		(0.080)
Model Pair 2				0.148***	0.111***
				(0.048)	(0.041)
Model Pair 3				-0.009	0.049
				(0.060)	(0.091)
Model Pair 4				0.035	-0.011
				(0.066)	(0.064)
Model Pair 5				-0.038	0.021
				(0.065)	(0.077)
Constant	0.257***	0.540***	0.259***	0.446***	0.214***
	(0.027)	(0.023)	(0.034)	(0.045)	(0.041)
Observations	836	836	836	836	836
R2	0.203	0.021	0.203	0.020	0.211

end back

Estimated α_m from Vectors of Posteriors





from guesses

back

Stochastic Model Selection

- Using the repeated model pairs where the same signal is observed to estimate the probability of making mistakes in applying a specific model selection criterion
  - For model selection in both (N=118), 16.1% in which two different models where selected
  - HP: constant probability of mistakes across participants and model selection criteria

Probability of mistakes = 8.33%

- Use this estimate and reported vectors of posteriors to determine the shares of participants that follow the different model selection criteria making mistakes 8.33% of the times
  - 73.10% selects the best-fitting model (maximum likelihood criterion)
  - 21.02% use either dogmatic model selection or model selection based on informativeness
  - 5.88% consistently select the worst-fitting model

back

Consistency Within Participant

	Bayesian	One-stage	Best-fitting	Worst-fitting	Other
Bayesian	2.33	0.00	0.00	0.00	1.33
One-stage	1.33	11.67	0.00	0.00	2.67
Best-fitting	0.67	1.00	28.67	2.67	7.33
Worst-fitting	0.00	0.33	3.67	4.33	1.33
Other	2.67	3.33	3.00	1.00	20.67

- Pooled repeated model pairs where the same signal is observed
- Use within 2 p.p. classification
- Very similar to other classification

back

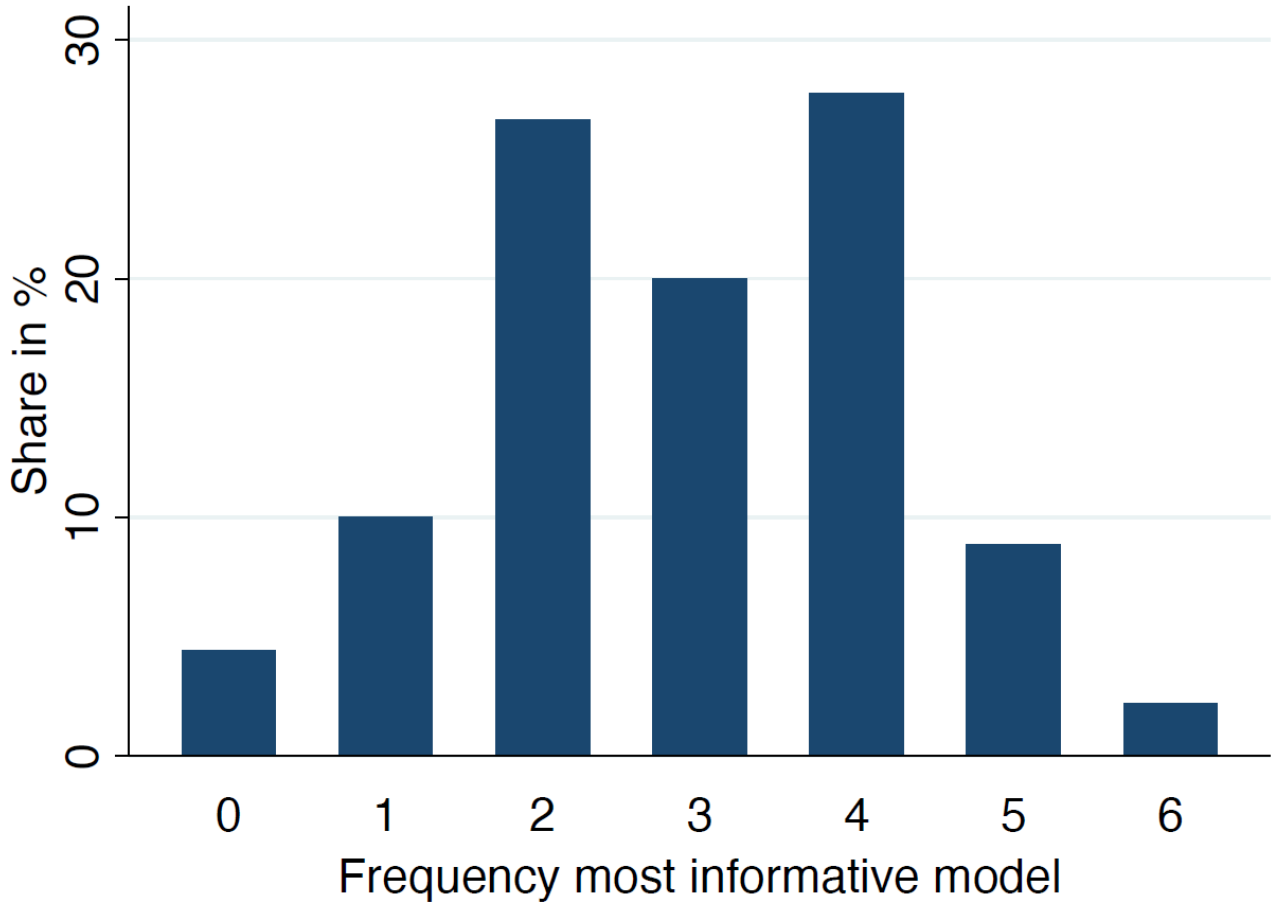
Updating Rule: First-Last Task

	Bayesian	One-stage	Best-fitting	Worst-fitting	Other	Total
Bayesian	2.67	0.00	0.00	0.00	3.33	6.00
One-stage	1.00	10.00	1.00	0.67	4.33	17.00
Best-fitting	0.67	3.33	19.33	6.67	5.00	35.00
Worst-fitting	0.00	0.67	6.67	4.67	2.33	14.33
Other	0.67	5.67	3.67	2.67	15.00	27.67
Total	5.00	19.67	30.67	14.67	30.00	100.00

- Row: first updating task w/ multiple models
- Column: last updating task w/ multiple models before repeated model pairs
- Use 2 p.p. classification

end back

Mistakes in Model Selection



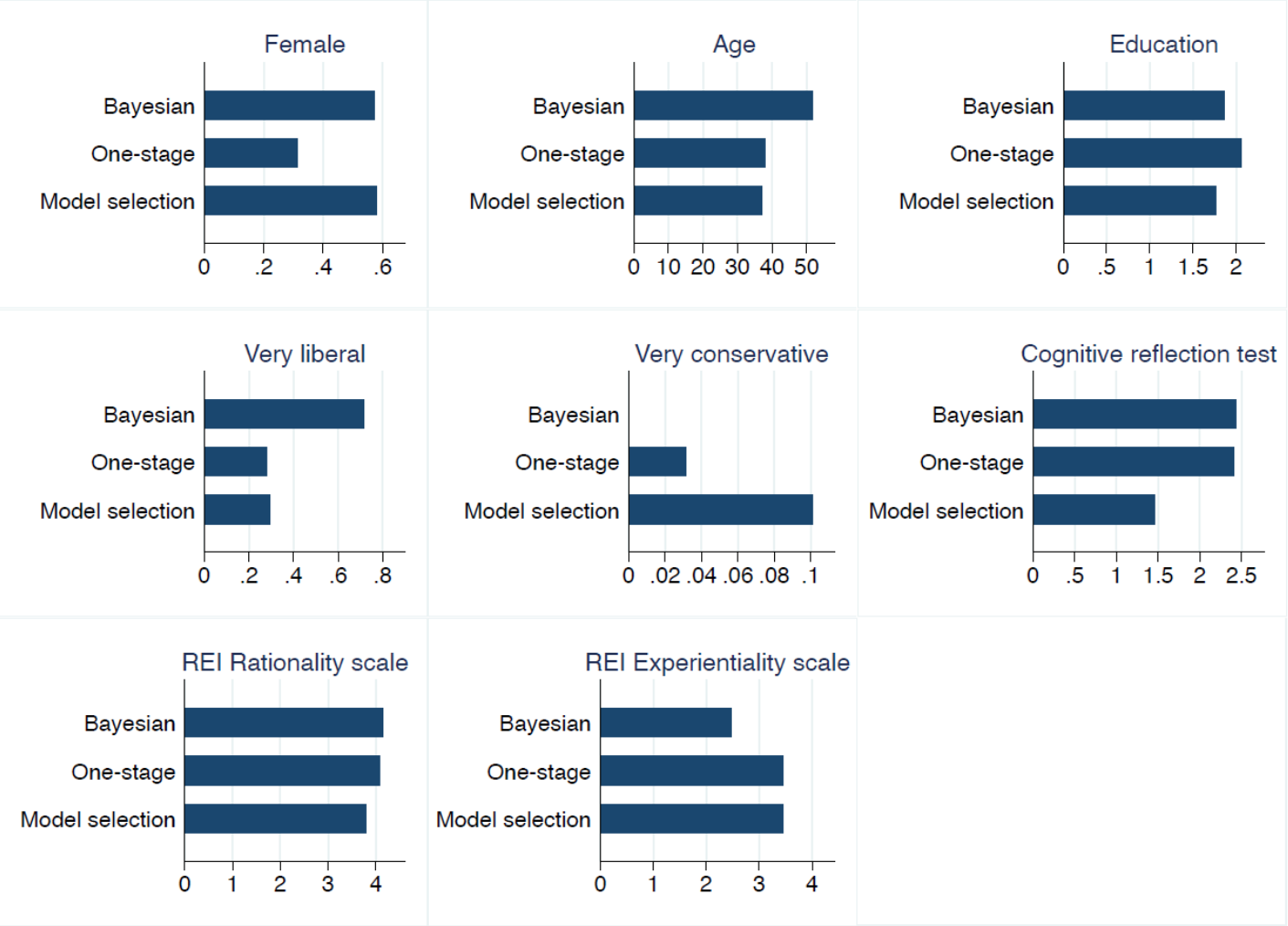
back

Model Selection vs Model Weighting

	I	II	III	IV	V	VI	VII	VIII	IX
Female	0.180***								0.147**
	(0.069)								(0.069)
Age		-0.005							-0.005*
		(0.003)							(0.003)
Education			-0.063						-0.009
			(0.047)						(0.042)
Very Liberal				-0.050					-0.067
				(0.077)					(0.071)
Very Conservative					0.205**				0.091
					(0.083)				(0.088)
CRT						-0.132***			-0.109***
						(0.025)			(0.0206)
REI Rationality							-0.161***		-0.107*
							(0.056)		(0.056)
REI Experientiality								0.054	0.002
								(0.045)	(0.041)
Observations	162	162	162	162	162	162	162	162	162
R2	0.042	0.015	0.012	0.003	0.016	0.114	0.041	0.011	0.185

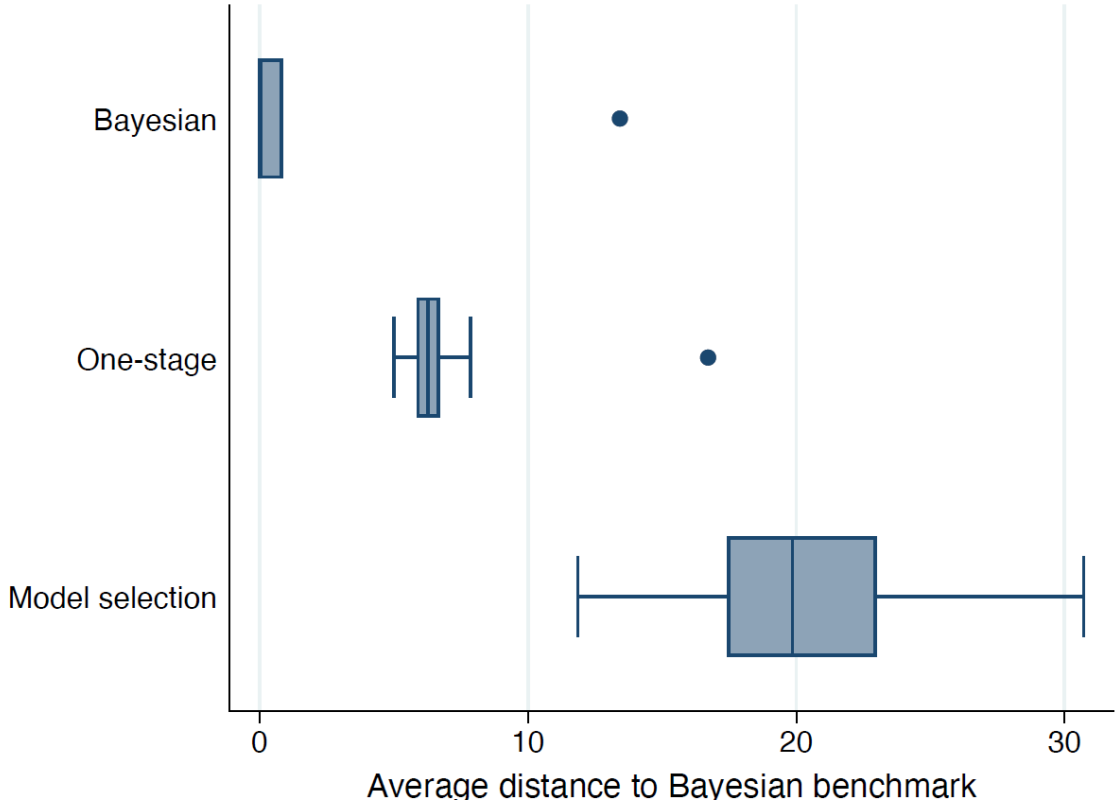
back end trade-off

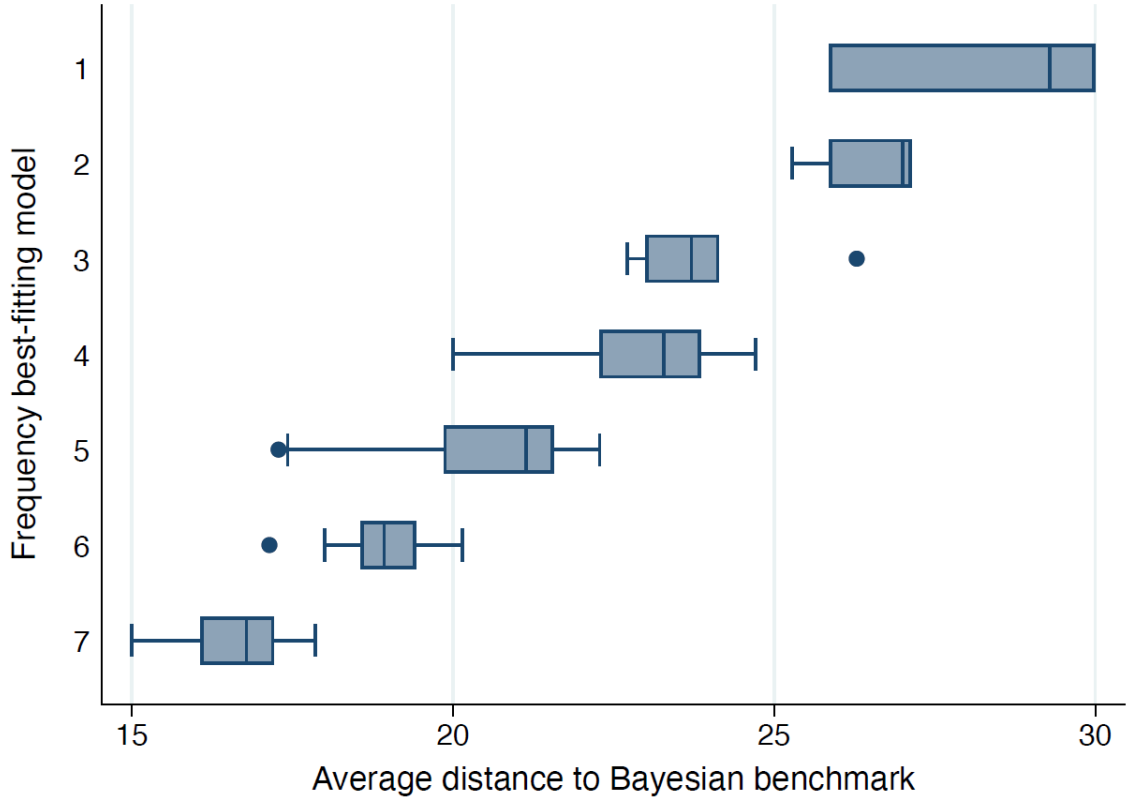
Model Selection vs Model Weighting



back

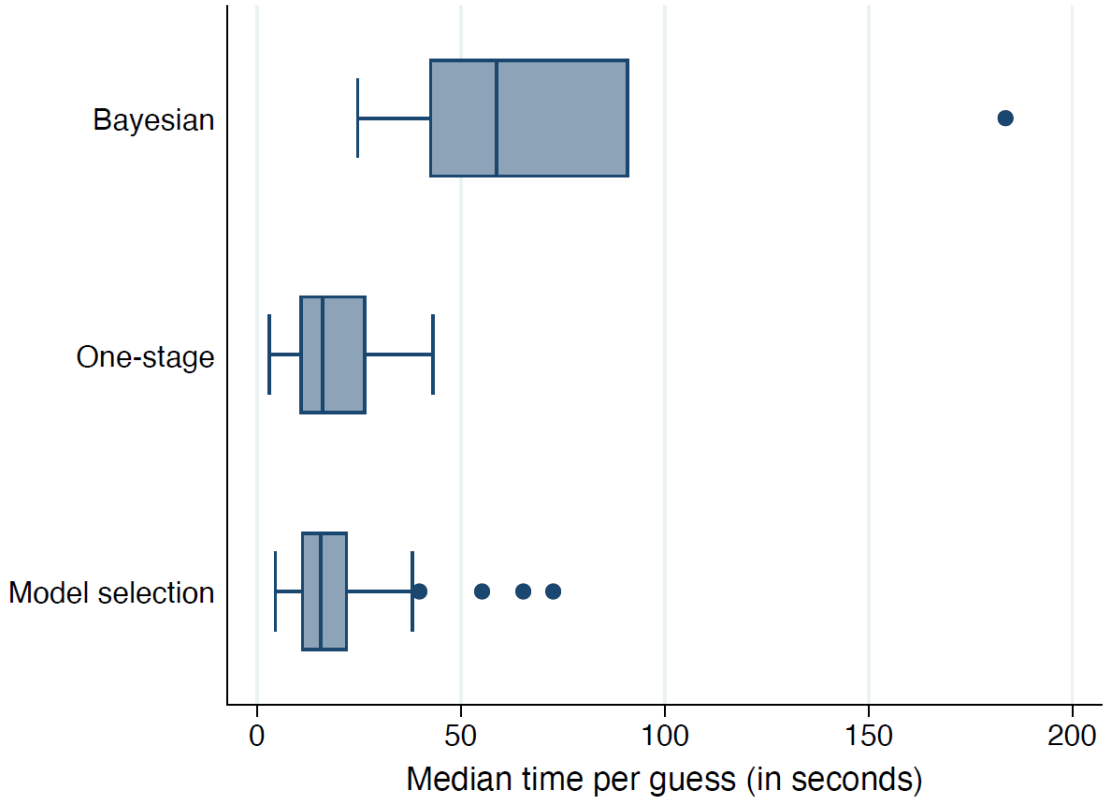
Model Selection vs Model Weighting





back

Model Selection vs Model Weighting



back

Additional Data Collection

Baseline study

- Objective prior over the models → No-prior treatment
  link to literature on ambiguity

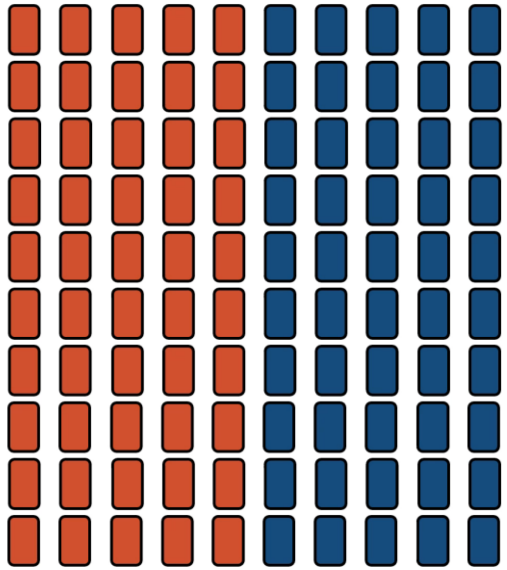
- Model predictions readily available → Click treatment
  deliberate action to access them

- Data collection on Prolific (February 2025), pre-registered
- 592 participants
  - 3 treatments x 2 choices in Part 3 (Model Pair 2 & Model Pair 3)

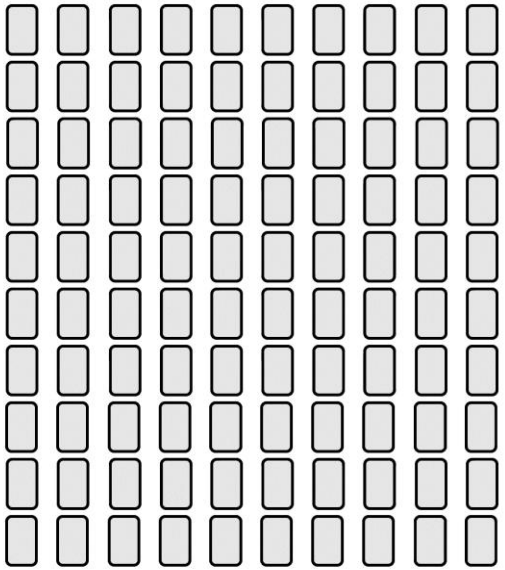
back

Additional Data Collection

Baseline



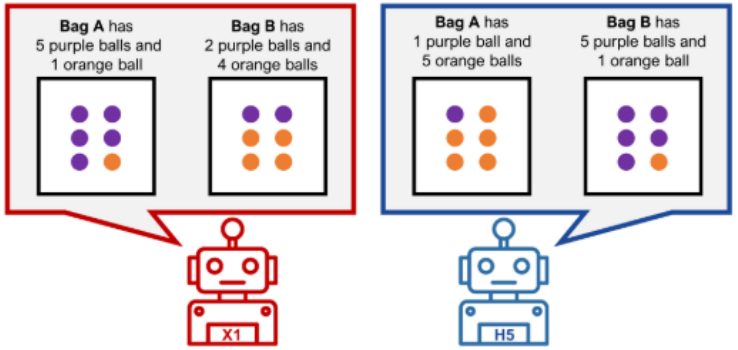
No-prior



back

Additional Data Collection

Baseline

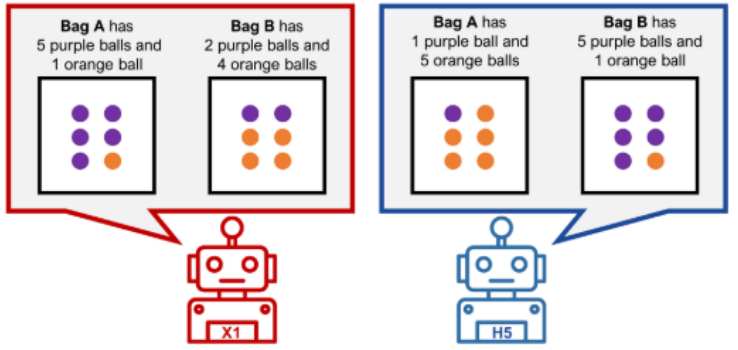


Remember, the computer has drawn a purple ball.

The chance that bag A was selected given that the color of the drawn ball is purple would be…

71% if I, robot X1, were selected	**17%** if I, robot H5, were selected

Click



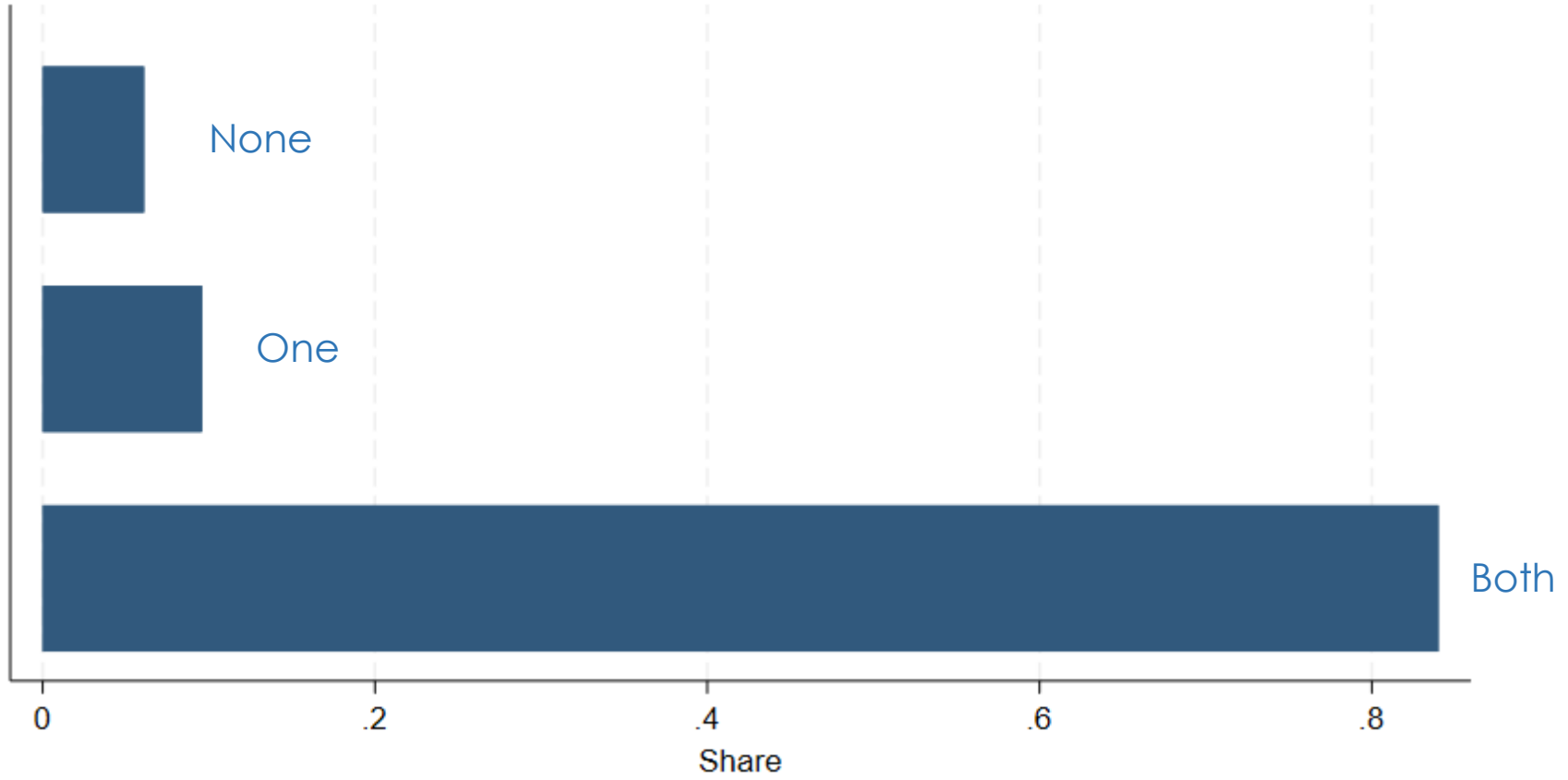
Remember, the computer has drawn a purple ball.

The chance that bag A was selected given that the color of the drawn ball is purple would be…

Press 'R' five times to see the chance if robot X1 were selected	Press 'B' five times to see the chance if robot H5 were selected

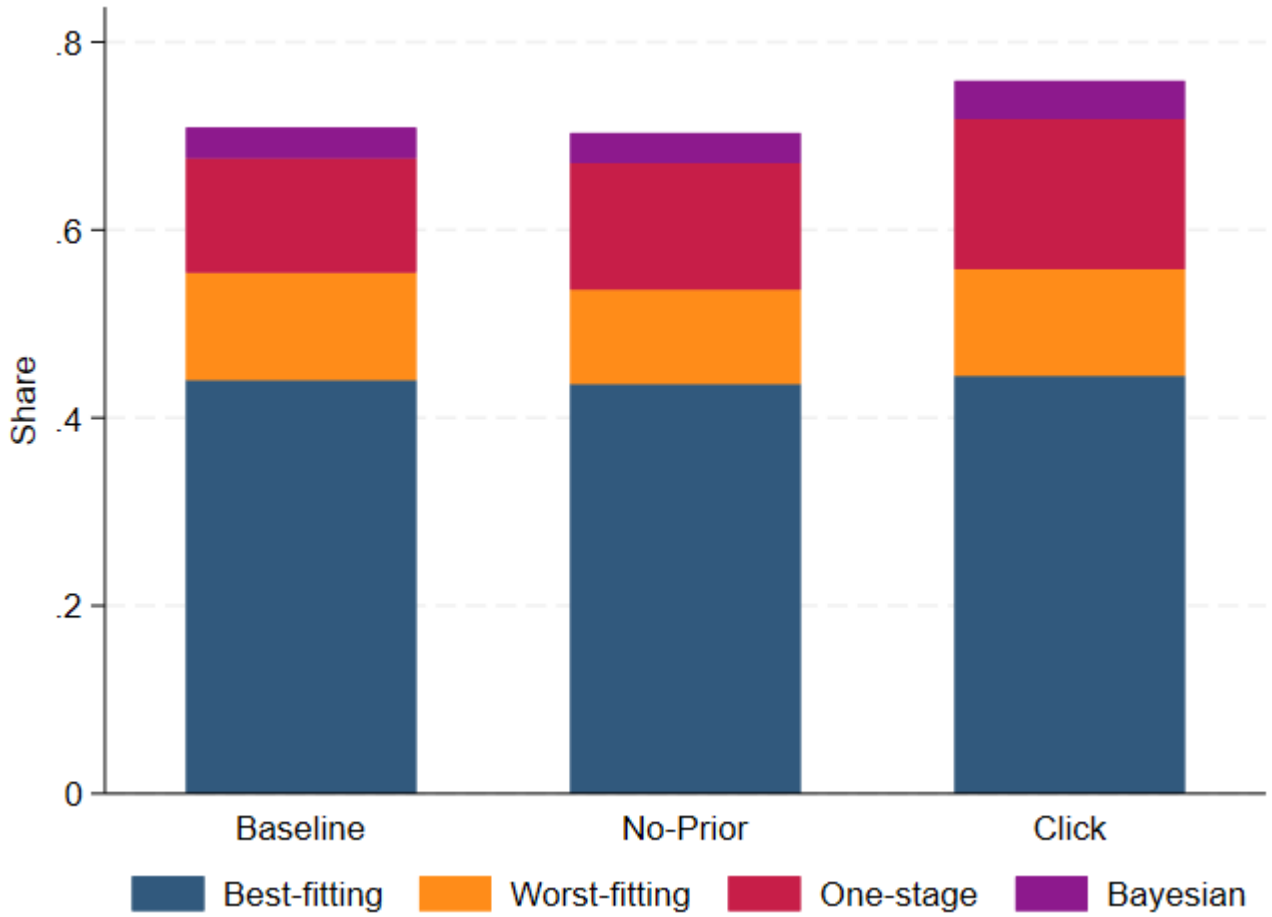
back

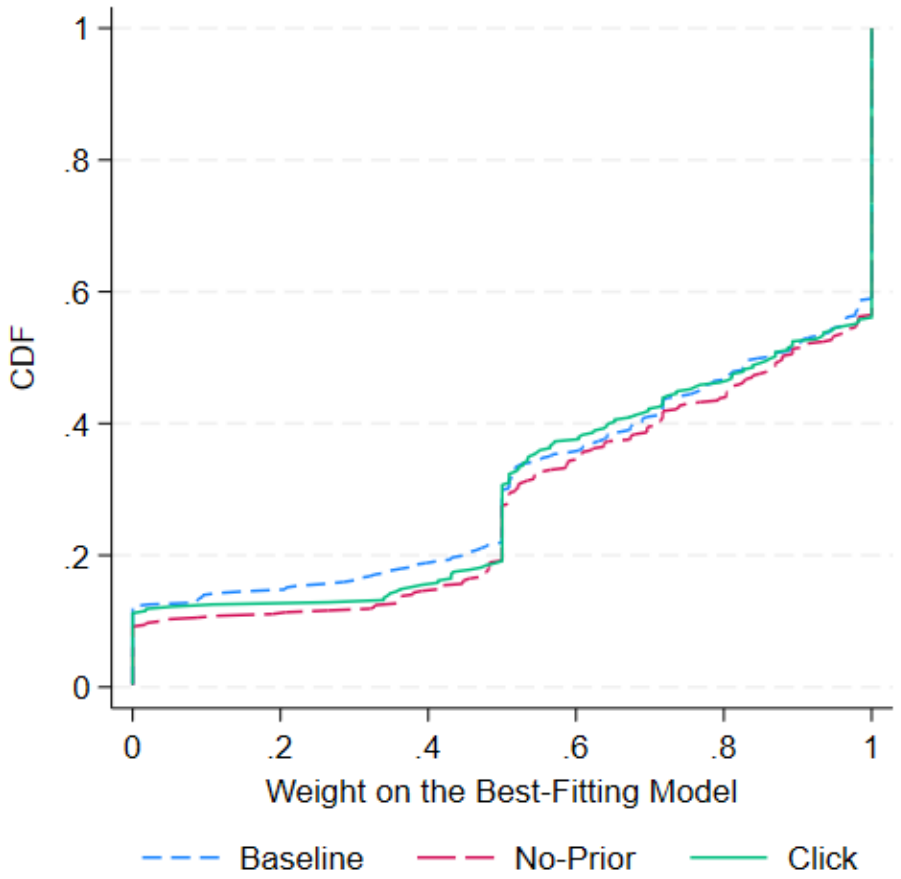
Clicking



back

Treatment Comparison





back