Reverse Bayesianism: Revising Beliefs in Light of Unforeseen Events

Christoph Becker [1] Tigran Melkonyan [2] Eugenio Proto [3]
Andis Sofianos [4] Stefan Trautmann [1]

[1]University of Heidelberg [2]University of Alabama
[3]University of Glasgow, CEPR, IZA and CESIfo [4]Durham University

EEA-ESEM Annual Meeting – Barcelona 2023

Unforeseen events and Bayesian updating

- Standard Bayesian paradigm is silent about how individuals react to unforeseen events
- But the universe frequently expands - observe something that was unforeseen/unforeseeable before
- Some examples: 9/11, Fall of Berlin Wall, Global pandemics

Unforeseen events and Bayesian updating

- Standard Bayesian paradigm is silent about how individuals react to unforeseen events
- But the universe frequently expands – observe something that was unforeseen/unforeseeable before
- Some examples: 9/11, Fall of Berlin Wall, Global pandemics

Reverse Bayesianism

- Karni and Viero (2013, 2015, 2017); Karni et al. (2020):
  - The construction of the new universe maintains consistency with the old structure
  - Probability is shifted away from known outcomes proportionally $\Rightarrow$ Keep ratios of previous estimates constant
- Intuitively simple and directly amenable to testing
- But adhering to rev. Bayesianism can be cognitively demanding & hindsight bias

Main Hypotheses Tested

H1. Participants update their beliefs according to reverse Bayesianism. That is, for any $\hat{p}_i^o, \hat{p}_i^u$ and any outcomes $i, i' \in C_0^F$:

$$\frac{\hat{p}_i^o}{\hat{p}_{i'}^o} = \frac{\hat{p}_i^u}{\hat{p}_{i'}^u}$$

H2. In treatments where unforeseen consequences are ruled out, the residual estimate: $\hat{p}_x = 0$

H3. In treatments where unforeseen consequences are not ruled out, the residual estimate: $\hat{p}_x > 0$

H4. Participants will not adjust their residual belief after an unforeseen event: $\hat{p}_x^u - \hat{p}_x^o = 0$

Overview of both Experiments

Experiment 1

- Studies an *"unforeseeable"* event.
- Observe random draws from urn, then provide estimates.
- Elicits implicit residual probabilities.

Experiment 2

- Studies when individuals stop expecting new events.
- Explore urn sequentially, providing estimates after each draw.
- Elicits explicit residual probabilities.

Results Teaser:

We find evidence supporting reverse Bayesianism in both experiments.

General Design of Experiment 1

1. Observe 20 physical draws from a real urn
   - *Original* urn: 24 balls worth 80 and 36 balls worth 190

General Design of Experiment 1

1. Observe 20 physical draws from a real urn
   - *Original* urn: 24 balls worth 80 and 36 balls worth 190
2. After observing draws:
   - Report probabilities: $\hat{p}^o_{80}, \hat{p}^o_{190}$ (Karni method ▸ Details)
   - Do not need to add up to 1 $\implies \hat{p}^o_x = 1 - \hat{p}^o_{80} - \hat{p}^o_{190}$
   - Report valuation of urn through: WTA^o (BDM ▸ Details)

General Design of Experiment 1

1. Observe 20 physical draws from a real urn
   - *Original* urn: 24 balls worth 80 and 36 balls worth 190
2. After observing draws:
   - Report probabilities: $\hat{p}^o_{80}, \hat{p}^o_{190}$ (Karni method ▸ Details)
   - Do not need to add up to 1 $\implies \hat{p}^o_x = 1 - \hat{p}^o_{80} - \hat{p}^o_{190}$
   - Report valuation of urn through: WTA^o (BDM ▸ Details)
3. Previously hidden *new* urn is revealed and its content emptied into original urn → *Updated* urn
   - *New* urn: 15 balls either worth 15 or 375 (depending on condition)

General Design of Experiment 1

1. Observe 20 physical draws from a real urn
   - *Original* urn: 24 balls worth 80 and 36 balls worth 190
2. After observing draws:
   - Report probabilities: $\hat{p}^o_{80}, \hat{p}^o_{190}$ (Karni method ▸ Details)
   - Do not need to add up to 1 $\implies \hat{p}^o_x = 1 - \hat{p}^o_{80} - \hat{p}^o_{190}$
   - Report valuation of urn through: WTA^o (BDM ▸ Details)
3. Previously hidden *new* urn is revealed and its content emptied into original urn → *Updated* urn
   - *New* urn: 15 balls either worth 15 or 375 (depending on condition)
4. Participants now report: $\hat{p}^u_{80}, \hat{p}^u_{190}, \hat{p}^u_S$, and WTA^u (Karni method & BDM)

General Design of Experiment 1

1. Observe 20 physical draws from a real urn
   - *Original* urn: 24 balls worth 80 and 36 balls worth 190
2. After observing draws:
   - Report probabilities: $\hat{p}^o_{80}, \hat{p}^o_{190}$ (Karni method ▸ Details)
   - Do not need to add up to 1 $\implies \hat{p}^o_x = 1 - \hat{p}^o_{80} - \hat{p}^o_{190}$
   - Report valuation of urn through: WTA^o (BDM ▸ Details)
3. Previously hidden *new* urn is revealed and its content emptied into original urn → *Updated* urn
   - *New* urn: 15 balls either worth 15 or 375 (depending on condition)
4. Participants now report: $\hat{p}^u_{80}, \hat{p}^u_{190}, \hat{p}^u_S$, and WTA^u (Karni method & BDM)

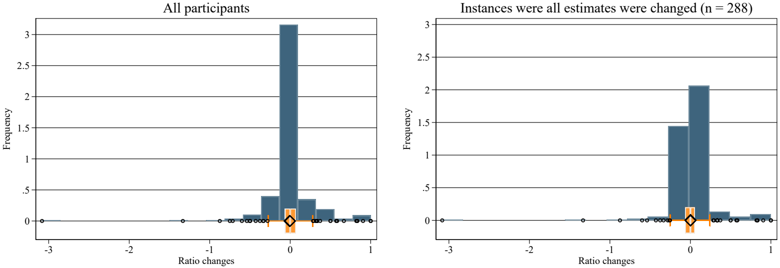
Two conditions: Information Surprise & Payment Surprise ▸ Timeline

- Students from University of Heidelberg and KIT
- 344 participants in total
- The design was pre-registered at the AEA RCT Registry

Reverse Bayesianism

Histograms of the ratio changes before vs. after the urn is updated

$$\Delta R = \frac{\hat{p}^{o}_{80}}{\hat{p}^{o}_{190}} - \frac{\hat{p}^{u}_{80}}{\hat{p}^{u}_{190}}$$



Histogram in blue, box plot in orange, outliers (circles) and mean (diamond) in black.

- Participants consistent with rev. Bayesianism. ▸ Statistical Tests
- Ratios remain constant, but individual estimates *are* updated. ▸ Evidence

Residuals

Results for H2 & H3:

- $\hat{p}_x^o = 0$ cannot be rejected in any treatment $\Rightarrow$ People do not implicitly expect the unknown when this is *reasonably* unforeseeable.
- $\hat{p}_x^u = 0$ rejected in the PS, low prize treatment. Statistical Tests
- Support for $H2$, limited support for $H3$.

Results for H4:

- Overall, $\hat{p}_x^u - \hat{p}_x^o = 0$ in most treatments.
- Some evidence of $\hat{p}_x^u \neq \hat{p}_x^o$ in (PS, *low prize*). Statistical Tests
- In line with $H4$.

Overview of both Experiments

Experiment 1

- Studies an *"unforeseeable"* event.
- Observe random draws from urn, then provide estimates.
- Elicits implicit residual probabilities.

Experiment 2

- Studies when individuals stop expecting new events.
- Explore urn sequentially, providing estimates after each draw.
- Elicits explicit residual probabilities.

Experiment 2 - Design

Participants draw 30 samples out of 4 different virtual urns containing different colours (100 marbles per urn).

- Draws and colours are randomized ▸ Example screen
- After each draw (Karni method):
  - State probability estimate for every observed outcome so far.
  - State a probability estimate for the residual, $\hat{p}_x$.

	Task 1		Task 2	Task 3	Task 4
	Two colours	Four colours			
Colour 1	55	40	53	75	48
Colour 2	45	28	35	25	28
Colour 3		20	12		12
Colour 4		12			12

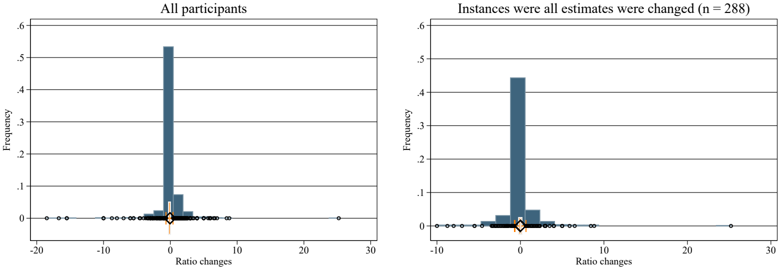
- Students from Warwick Business School
- 174 participants in total
- The design was pre-registered at the AEA RCT Registry

Reverse Bayesianism

Histograms of ratio changes before vs. after the urn is updated

Third outcome: $\Delta R^3 = \frac{\hat{p}_H^u}{\hat{p}_L^u} - \frac{\hat{p}_H^o}{\hat{p}_L^o}$

Fourth outcome: $\Delta R_1^4 = \frac{\hat{p}_H^u}{\hat{p}_M^u} - \frac{\hat{p}_H^o}{\hat{p}_M^o}$; $\Delta R_2^4 = \frac{\hat{p}_M^u}{\hat{p}_L^u} - \frac{\hat{p}_M^o}{\hat{p}_L^o}$; $\Delta R_3^4 = \frac{\hat{p}_H^u}{\hat{p}_L^u} - \frac{\hat{p}_H^o}{\hat{p}_L^o}$



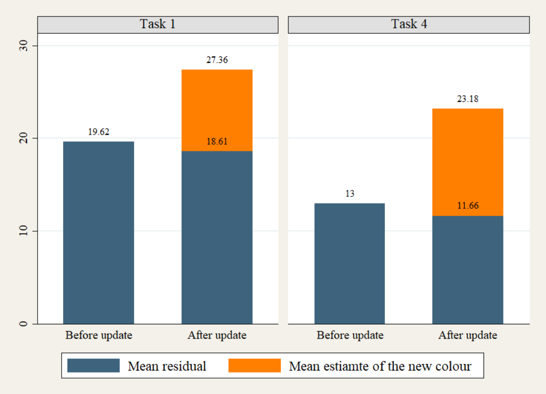
Histogram in blue, box plot in orange, outliers (circles) and mean (diamond) in black.

Again:

- Participants consistent with rev. Bayesianism. Statistical Tests
- Ratios remain constant, but individual estimates *are* updated. Evidence

To what degree are participants Bayesian updaters?

- Unpacking bias (Tversky and Koehler, 1994; Sonnemann et al., 2013)



▸ Other graphs

- Unpacked estimate is significantly larger than the original residual ($p - values < 0.001$, both before and after correction)

Concluding remarks

- Predictions of Bayesian updating are typically systematically violated in experimental studies (Charness and Levin, 2005; Charness et al., 2007; Holt, 2009).
- We find that behaviour remarkably conforms with rev. Bayesianism
  - Holds both for *foreseeable* and *unforeseeable* unknowns
  - Holds whether participants did not expect further surprises (Experiment 1) or did (Experiment 2)
  - Despite other biases in beliefs (unpacking of estimates after surprise)
- Additionally, we find that:
  - Hope dominates fear when faced with the unknown Evidence
  - Participants become complacent in their expectations of the unknown as they sample more ▸ Evidence
- Planning new experimental sessions studying situations where a paradigm shift takes place, i.e., extent by which rev. Bayesianism still adhered to

Thanks for your attention

Karni (2009) Method

- Participants are asked to express a perceived likelihood or probability for a prize – in our case, proportion of prizes equal to value X within the urn
- This declared probability is compared to a random number between 0 and 1
- **IF** the random number is greater than the declared probability, participants receive a lottery paying X according to the true proportion of prize X within the urn
- Instead, **IF** the random number is less than the declared probability, participants receive a lottery paying X according to the random number probability
- Participants were told that declaring their true perception is in their best interest, if interested in more details they could click on a button explaining the above procedure

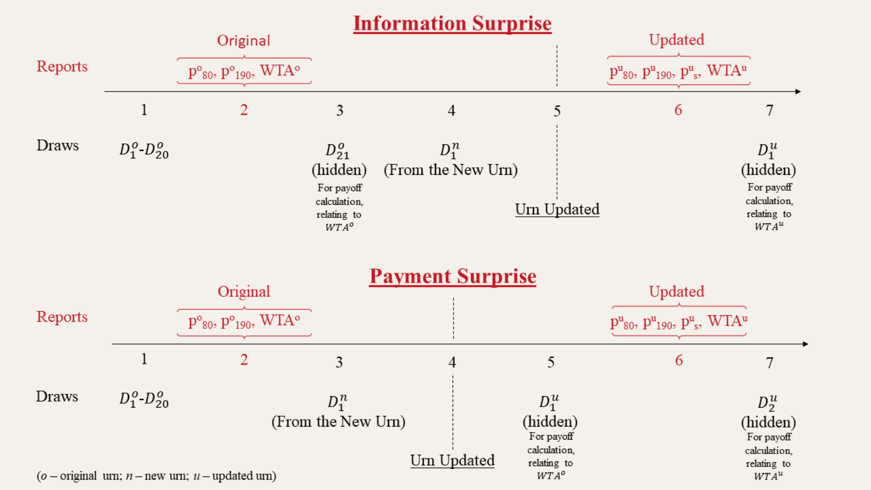
Back

Standard BDM Method

Some details

- This method asks participants to state a minimum willingness to accept (WTA) for an item – in our context a lottery
- Their stated value is then compared to a random number
- **IF** stated WTA is greater than the random number, the participant does not sell the lottery and will thus be paid according to the realisation of the lottery
- Instead, **IF** stated WTA is less than or equal to the random number, the participant gets to sell the lottery for *the value of the random number*
- BDM method is said to be incentive compatible, i.e. aligns incentives for truthful reporting

Back



▸ IS Details ▸ PS Details ▸ Treatment Reasoning ▸ Back to Design Overview

Exp. 1 Design: Information Surprise (IS) Condition

1. Original urn:
   - Participants told: *"the urn contains two and only two prizes"*.
   - Not told what these prizes or their relative proportions are.
   - **Not** alerted on possible changes to composition of urn.
2. After reports on original urn:
   - Hidden draw relating to WTA^o.
3. New ⇒ Updated urn:
   - Draw one ball from new urn and told: *"This urn contains only the prize you are (about to be) shown"*.
4. After reports on updated urn:
   - Hidden draw relating to WTA^u.

Back

Exp. 1 Design: Payment Surprise (PS) Condition

1. Original urn:
   - Participants told: *"new balls representing different tokens to what you have been observing so far may be added to this urn"*.
   - Not told about number of prizes in urn or anything about proportion of any prize.
2. New ⇒ Updated urn:
   - Draw one ball from new urn and told: *"This urn contains new prizes. One such prize is the one you see. The urn contains no prizes similar to what you have been observing as a result of random draws from the other urn"*.
3. After urn is updated:
   - Hidden draw relating to WTA^o.
4. After reports on updated urn:
   - Hidden draw relating to WTA^u.

Back

Contrasting IS with PS condition

- Our aim is to induce an unforeseeable event and study reactions to it
- For an event to be unforseeable it must:
  1. be unannounced and/or ruled out
  2. have immediate payment consequences
- Incorproprating both risks a design that would contain deception
  - either by ruling out any new event and then enforcing a payment relevant surprise
  - or by enforcing a payment relevant surprise without forewarning
- Hence, two conditions:

  IS: New event unannounced, but not instantly payment-relevant

  PS: New event instantly payment-relevant, but forewarned

Back

Reverse Bayesianism

Statistical tests

$$\Delta R = \frac{\hat{p}^o_{80}}{\hat{p}^o_{190}} - \frac{\hat{p}^u_{80}}{\hat{p}^u_{190}} = 0$$

		Obs	Avg ratio change	p-value	p-value (corr)	**95%** CI	Bayes factor
IS	low prize	75	0.007	0.375	1.000	$[-\mathbf{0.06}, \mathbf{0.05}]$	14.76
	high prize	75	-0.039	0.981	1.000	$[-\mathbf{0.06}, \mathbf{0.14}]$	6.57
PS	low prize	93	0.016	0.918	1.000	$[-\mathbf{0.06}, \mathbf{0.03}]$	9.72
	high prize	100	-0.007	0.011	0.043	$[-\mathbf{0.04}, \mathbf{0.05}]$	16.35

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure, confidence interval from one sample t-test, Bayes factor from JZS test.

		Increased	Decreased	Const ratio	p-value	p-value (corr)	Unchanged Est
IS	low prize	29	23	23	0.488	1.000	1
	high prize	31	32	12	1.000	1.000	1
PS	low prize	33	37	23	0.720	1.000	0
	high prize	29	61	10	0.001	0.004	4

Matched pairs sign test, p-values corrected by Bonferroni-Holm procedure. 'Unchanged Est.' denotes the subset of those holding their ratios constant while not changing any of their estimates.

Participants consistent with rev. Bayesianism, supporting $H1$

Back

Do estimates of known outcomes change?

Ratios remain constant, but individual estimates *are* updated

		Obs	Diff	p-value	p-value (corr)
IS, low prize	$\hat{p}^u_{80} - \hat{p}^o_{80}$	76	-0.101	0.000	0.000
	$\hat{p}^u_{190} - \hat{p}^o_{190}$	76	-0.130	0.000	0.000
IS, high prize	$\hat{p}^u_{80} - \hat{p}^o_{80}$	75	-0.102	0.000	0.000
	$\hat{p}^u_{190} - \hat{p}^o_{190}$	75	-0.125	0.000	0.000
PS, low prize	$\hat{p}^u_{80} - \hat{p}^o_{80}$	93	-0.100	0.000	0.000
	$\hat{p}^u_{190} - \hat{p}^o_{190}$	93	-0.136	0.000	0.000
PS, high prize	$\hat{p}^u_{80} - \hat{p}^o_{80}$	100	-0.075	0.000	0.000
	$\hat{p}^u_{190} - \hat{p}^o_{190}$	100	-0.108	0.000	0.000

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure.

Back

Residuals different from zero

		$\hat{p}_x^t = 0$	$\hat{p}_x^t > 0$	$\hat{p}_x^t < 0$	p-value	p-value (corr)
IS, original	low prize	74	1	1	0.993	1.000
	high prize	71	3	1	0.314	1.000
PS, original	low prize	92	0	1	0.317	1.000
	high prize	90	6	4	0.549	1.000
IS, updated	low prize	61	10	5	0.251	1.000
	high prize	65	7	3	0.228	1.000
PS, updated	low prize	74	16	3	0.004	0.028
	high prize	84	11	5	0.146	1.000

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure.

- $\hat{p}_x^o = 0$ cannot be rejected in any treatment $\Rightarrow$ People do not implicitly expect the unknown when this is *reasonably* unforeseeable
- $\hat{p}_x^u = 0$ rejected in the PS, low prize treatment
- Support for $H2$, limited support for $H3$

▸ Back

Adjusting beliefs after an unforeseen event

$$\Delta\hat{p}_x = \hat{p}_x^u - \hat{p}_x^o = 0$$

		$\Delta\hat{p}_x = 0$	$\Delta\hat{p}_x > 0$	$\Delta\hat{p}_x < 0$	p-value	p-value (corr)
IS	low prize	60	11	5	0.173	0.692
	high prize	63	6	6	0.937	1.000
PS	low prize	73	17	3	0.002	0.009
	high prize	82	11	7	0.345	1.000

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure.

- Overall, support for *H4*
- Some evidence of $\hat{p}_x^u \neq \hat{p}_x^o$ in (*PS*, *low prize*)

Back

Differences in urn valuations

Original urn: WTA^o

	IS	PS	Diff	p-value
Low prize	110.39	138.47	-28.08	0.008
High prize	110.48	134.81	-24.33	0.002

Wilcoxon signed-rank test.

Updated urn: WTA^u

	IS	PS	Diff	p-value
Low prize	86.45	96.70	-10.25	0.074
High prize	153.53	178.25	-24.72	0.160

Wilcoxon signed-rank test.

- $WTA(PS) > WTA(IS)$ in both prize conditions
- Hope seems to dominate fear
- Caveat: for more uncertain prospects, WTA leads to higher valuations (Trautmann et al., 2011; Trautmann and Schmidt, 2012)

Back

Part 1

Please draw a sample from the box.

Sample draw: 30

maroon

Please indicate in the fields below, how many marbles of a samples color you think are in this box. Remember, the box has a total of 100 marbles.

Number of **orange** marbles: [ ] Fill previous estimate + -

Number of **maroon** marbles: [ ] Fill previous estimate + -

Number of **blue** marbles: [ ] Fill previous estimate + -

Number of **salmon** marbles: [ ] Fill previous estimate + -

Number of marbles of **any other color**: [ ] Fill previous estimate + -

orange orange maroon maroon blue blue orange maroon maroon salmon orange orange maroon salmon orange

salmon orange orange maroon maroon orange orange maroon maroon salmon blue maroon maroon orange maroon

Go to part 2

Back

Exp. 2: Reverse Bayesianism

Statistical tests

Third outcome: $\Delta R^3 = \frac{\hat{p}_H^u}{\hat{p}_L^u} - \frac{\hat{p}_H^o}{\hat{p}_L^o} = 0$

Fourth outcome: $\Delta R_1^4 = \frac{\hat{p}_H^u}{\hat{p}_M^u} - \frac{\hat{p}_H^o}{\hat{p}_M^o} = 0$; $\Delta R_2^4 = \frac{\hat{p}_M^u}{\hat{p}_L^u} - \frac{\hat{p}_M^o}{\hat{p}_L^o} = 0$; $\Delta R_3^4 = \frac{\hat{p}_H^u}{\hat{p}_L^u} - \frac{\hat{p}_H^o}{\hat{p}_L^o} = 0$

		Obs	Avg ratio change	p-value	p-value (corr)	95%CI	Bayes factor
Task 1	ΔR^3	85	-1.365	0.172	1.000	$[-0.10, 0.29]$	5.32
	ΔR_1^4	84	-0.548	0.584	1.000	$[-0.79, 0.22]$	4.44
	ΔR_2^4	84	-2.134	0.033	0.362	$[-1.27, 0.04]$	1.58
	ΔR_3^4	84	-1.005	0.315	1.000	$[-0.52, 0.33]$	7.52
Pooled	ΔR_P^4	252	-2.229	0.026	0.284	$[-0.64, -0.03]$	1.49
Task 2	ΔR^3	169	-2.632	0.008	0.093	$[-0.31, 0.01]$	2.26
Task 4	ΔR^3	173	-0.648	0.517	1.000	$[-0.26, 0.06]$	5.71
	ΔR_1^4	164	-0.048	0.962	1.000	$[-0.07, 0.25]$	6.05
	ΔR_2^4	163	-0.067	0.946	1.000	$[-0.19, 0.69]$	6.09
	ΔR_3^4	163	-0.148	0.883	1.000	$[-0.14, 0.27]$	9.46
Pooled	ΔR_P^4	490	-0.203	0.839	1.000	$[-0.03, 0.30]$	5.64

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure, confidence interval from one sample t-test,

Bayes factor from JZS test.

Participants consistent with rev. Bayesianism, supporting $H1$ ▸ Back

Exp. 2: Reverse Bayesianism

Statistical tests II

		Increased	Decreased	Const ratio	p-value	p-value (corr)	Unchanged Est
Task 1	ΔR^3	16	29	40	0.072	0.797	26
	ΔR_1^4	19	21	44	0.875	1.000	31
	ΔR_2^4	16	31	37	0.040	0.440	32
	ΔR_3^4	16	23	45	0.337	1.000	35
Pooled	ΔR_P^4	51	75	126	0.040	0.440	93
Task 2	ΔR^3	35	59	75	0.017	0.189	46
Task 4	ΔR^3	45	50	78	0.682	1.000	44
	ΔR_1^4	50	57	57	0.562	1.000	36
	ΔR_2^4	54	60	49	0.640	1.000	33
	ΔR_3^4	43	47	73	0.752	1.000	37
Pooled	ΔR_P^4	147	164	179	0.364	1.000	108

Matched pairs sign test, p-values corrected by Bonferroni-Holm procedure. 'Unchanged Est' denotes the subset of those holding their ratios constant while not changing any of their estimates.

- Many keep estimates unchanged; possibly due to re-fill button
- Substantial share holds ratio constant, not trivial especially after fourth outcome

Back

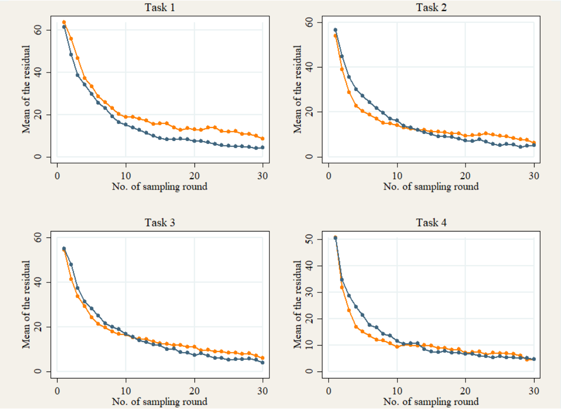
Do estimates of known outcomes change?

		Obs	Diff	p-value	p-value (corr)
Task 1, after third color	$\hat{p}_H^u - \hat{p}_H^o$	85	-0.06	0.000	0.000
	$\hat{p}_L^u - \hat{p}_L^o$	85	-0.04	0.000	0.000
Task 1, after fourth color	$\hat{p}_H^u - \hat{p}_H^o$	84	-0.04	0.000	0.000
	$\hat{p}_M^u - \hat{p}_M^o$	84	-0.02	0.000	0.005
	$\hat{p}_L^u - \hat{p}_L^o$	84	-0.02	0.000	0.000
Task 2, after third color	$\hat{p}_H^u - \hat{p}_H^o$	169	-0.07	0.000	0.000
	$\hat{p}_L^u - \hat{p}_L^o$	169	-0.05	0.000	0.000
Task 4, after third color	$\hat{p}_H^u - \hat{p}_H^o$	174	-0.07	0.000	0.000
	$\hat{p}_L^u - \hat{p}_L^o$	174	-0.07	0.000	0.000
Task 4, after fourth color	$\hat{p}_H^u - \hat{p}_H^o$	164	-0.05	0.000	0.000
	$\hat{p}_M^u - \hat{p}_M^o$	164	-0.03	0.000	0.000
	$\hat{p}_L^u - \hat{p}_L^o$	164	-0.03	0.000	0.000

Wilcoxon signed-rank test, p-values corrected by Bonferroni-Holm procedure.

Back

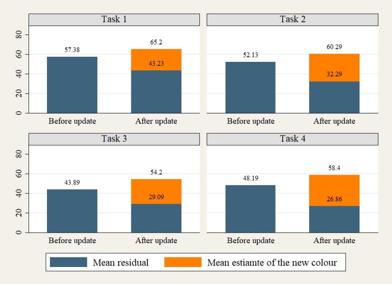
Dynamics of Residuals

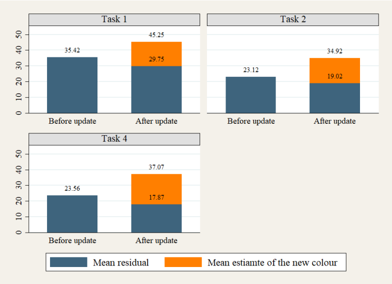


- No difference between treatments (Kolmogorov-Smirnov test, all $p - values > 0.994$).
- Pearson correlation coefficient between # of samples and $\hat{p}_x$: $\rho < -0.311$
- Spearman correlation coefficient between # of observed colours and $\hat{p}_x$: $\rho < -0.272$

Back

To what degree are participants Bayesian updaters?





Back