Bayesian modeling of the Iowa Gambling Task for substance dependent individuals: With variants on perseverance models

Seminar in Experimental Psychology: Computational Modeling

Spring 2018

Bokyung Son (sbk1652@snu.ac.kr)

Computational Linguistics Lab, SNU

Abstract

- Iowa Gambling Task data from substance dependent individuals were analyzed with Bayesian modeling techniques. Three variants of perseverance models were tested, which differ in complexity and conceptual framework: two were Win-Stay-Lose-Shift (WSLS-2) models, each with 2 (WSLS-2) and 6 (WSLS-6) parameters; the other one was a dual strategy model, the Value-Plus-Perseverance(VPP). Amphetamine and heroin user groups were compared against healthy controls based on parameter estimates and posterior predictives of the models. The results indicated heavy reliance on RL strategy than perseverance in all three groups, however, heroin users displayed a lower loss aversion and higher perseverance to gains. Amphetamine users showed a more moderate but still noticeable preference of big gains as well. Parameter correlation checks, posterior predictive checks, *post-hoc* fit and parameter recovery tests were used to compare performances of three models. Model comparison results showed that VPP was the winning model, although with possibility of parameter misspecification.

Introduction

- **Iowa Gambling Task (IGT)**
  - A popular decision making task in the medical/addiction literature
  - “Simulates real-life decision making in the way it factors uncertainty of premises and outcomes, as well as reward and punishment” (Bechara et al., 1994)
  - “Poor behavioral performance could be the result of deficits in various distinct neurocognitive processes, such as hypersensitivity to reward and/or hyposensitivity to losses, failure to learn from past outcomes and losses, and/or erratic and impulsive response style” (Ahn et al., 2014)

- Decision making of substance dependent individuals
  - SDIs often exhibit selection pattern that yields high immediate gains but even higher future losses (Bechara et al., 2001)

Methods: Dataset

- Studied and provided by Ahn et al. (2014)
- 3 groups
  - 48 healthy controls **(HC)**
  - 38 amphetamine users **(AM)** : stimulant
  - 43 heroine users **(HR)** : opiate
- Contains information on gain, loss, chosen deck at each trial
- Used a modified IGT design
  - Original design by Bechara et al.(1994), with a fixed amount of gain/loss throughout all trials.
  - Modified design by Bechara et al.(2001), with changing amounts of gain/loss per trial. Used a block design of 6 blocks (each with 10 trials); each deck's net outcome throughout a block increases/decreases compared to the previous block.

- Available at:
https://figshare.com/articles/IGT_raw_data_Ahn_et_al_2014_Frontiers_in_Psychology/1101324

Methods: Task

	Disadvantageous ("bad")		Advantageous ("good")	
Each deck has 60 cards (6 blocks x 10 cards) Each subject plays 100 trials Net outcome changes over blocks	A	B	C	D
Initial block net	-250		+250	
Block-wise net increase by	-150		+25	
Gain policy	Always bigger than C, D		Always smaller than A, B	
Loss occurrence frequency	50%	10%	50%	10%

Methods: Models

- Models for IGT under RL framework
  - Expected reward values are developed for each option; choice is made by comparing the value of each option against others.
  - Expectancy Valence (EV) (Busemeyer & Stout, 2002)
  - Prospect Valence Learning (PVL) (Ahn et al., 2008)

- Here, I focus on alternative models that take **Perseverance** into consideration.

- Win-Stay-Lose-Shift (WSLS) policy
  - People have tendencies to stay with the same option after "win" trials, and shift to a different option after "lose" trials.
  - **2-parameter WSLS (WSLS-2)** (Novak & Sigmund, 1993)
  - **6-parameter WSLS (WSLS-6)** (Worthy & Maddox, 2014)

- Dual strategy: RL + WSLS
  - **Value-Plus-Perseveration (VPP)** (Worthy & Byrne, 2013)
  - WSLS-RL (Worthy & Maddox, 2014)*

*I implemented this model, but could not make analysis due to convergence issues.

Methods: Models [WSLS-2]

ws	P(stay|win)	[0, 1]
ls	P(shift|loss)	[0, 1]

If $r\,(t) \geq r\,(t-1)$, then the trial is considered a "win" trial

$$P\,(a_{t+1}|choice_t = a \text{ and } r\,(t) \geq r\,(t-1)) = P\,(stay|win)$$

If $r\,(t) < r\,(t-1)$, then the trial is considered a "loss" trial

$$P\,(b_{t+1}|choice_t = a \text{ and } r\,(t) < r\,(t-1)) = P\,(shift|loss)$$

r(t): outcome (reward) at trial t

Methods: Models [WSLS-6]

- Unlike WSLS-2, this variant allows P(stay|win) and P(shift|loss) to change over trials.
- This model was not officially published as an independent model. It was used inside the WSLS-RL, but I tested it detached from the nested model.

ws_{init}	Initial value of P(stay|win)	[0, 1]
ls_{init}	Initial value of P(shift|loss)	[0, 1]
θ_{ws}	Change rate of P(stay|win) on each trial	[0, 1]
θ_{ls}	Change rate of P(shift|loss) on each trial	[0, 1]
ws_{fin}	Asymptotic ending value of P(stay|win)	[0, 1]
ls_{fin}	Asymptotic ending value of P(shift|loss)	[0, 1]

If $r(t) \geq r(t-1)$, then the trial is considered a "win" trial

$$P(stay|win)_{t+1} = P(stay|win)_t + \theta_{P(stay|win)} \times (P(stay|win)_{final} - P(stay|win)_t)$$

If $r(t) < r(t-1)$, then the trial is considered a "loss" trial

$$P(shift|loss)_{t+1} = P(shift|loss)_t + \theta_{P(lose|shift)} \times (P(shift|loss)_{final} - P(shift|loss)_t)$$

Methods: Models [VPP]

- A dual strategy model with 8 parameters.

k	Perseverance decay	[0, 1]
ε_{pos}	Perseverance bias after positive outcome	[-∞, ∞]*
ε_{neg}	Perseverance bias after negative outcome	[-∞, ∞]*

- Decay rule for Perseverance to the *chosen option i* at trial t

$$P_i(t) = \begin{cases} k \cdot P_i(t-1) + \varepsilon_{\text{pos}} & \textit{if } x(t) \geq 0 \\ k \cdot P_i(t-1) + \varepsilon_{\text{neg}} & \textit{if } x(t) < 0 \end{cases}$$

- Unlike in the WSLS models, the outcome of current trial is evaluated by simply whether it is positive or negative, without comparison to the previous trials.

* The range is [-1, +1] in the original model. The code I used for the VPP model is from the *hbayesdm* package (Ahn et al., 2017). This configuration follows Ahn et al.(2014) and is reflected in *hbayesdm*.

Methods: Models [VPP] (cont'd)

α	Shape of utility	(0, 1)*
λ	Loss aversion If $\boldsymbol{\lambda} > 1$, more sensitive to losses than gains	(0, 5)*

- Utility function

$$u(t) = \begin{cases} x(t)^{\alpha} & \text{if } x(t) \geq 0 \\ -\lambda|x(t)|^{\alpha} & \text{if } x(t) < 0 \end{cases}$$

ϕ	Recency If large, a greater weight to recent outcomes	[0, 1]

- Value updating
  - Utility is used to update expectancy for the *chosen option i* on trial t.
  - Delta rule: expectancies are recency-weighted averages of the rewards received.

$$E_i(t) = E_i(t-1) + \phi \cdot [u(t) - E_i(t-1)]$$

* I made (minimal) modification of the *hbayesdm* code regarding the range of these parameters: *hbayesdm* puts $0<\alpha<2$ and $0<\lambda<10$ (configuration by Ahn et al.(2014)), but I follow the original model.

Methods: Models [VPP] (cont'd)

w	Interpolation weight between EV and perseverance strength If > 0.5, greater weight on EV	[0, 1]

- Value is determined through dual strategy, considering both EV (RL term) and perseverance.

$$V_j(t) = w_{E_j} \cdot E_j(t) + (1 - w_{E_j}) \cdot P_j(t)$$

c	Exploitation; Response consistency If large, likely to select options with higher expected values If small, likely to explore other options	[0, 5]

- Action (choice) selection
  - Softmax probability of each deck j to be chosen on trial t $$Pr(G_j(t)) = \frac{e^{[\theta(t) \cdot V_j(t)]}}{\sum_{j=1}^{4} e^{[\theta(t) \cdot V_j(t)]}}$$
  - With trial-independent Θ determined by c $\theta(t) = 3^c - 1$

Methods: Settings

- Hierarchical model
  - Each individual-level parameters come from a group distribution.
- Prior
  - $\mu \sim$ Normal(0, 1) for bounded parameters, Normal(0, 10) for unbounded parameters
  - $\sigma \sim$ Cauchy(0, 5)
- For initial choice, all decks have the same probability (0.25) of being chosen

- Implemented with Stan + PyStan
- NUTS sampler
- For each group, 2,000 samples x 3 chains = 6,000 samples (excluding warm-ups)
  - 5,000 warm-ups per chain for WSLS-RL
  - 1,000 warm-ups per chain for all other models
- Fed in hand-set initial parameter values only for WSLS-RL.

Methods: Analysis

* Marked diagnostics are omitted to conserve space.

<table>
<tr><td rowspan="2">Convergence check</td><td>Traceplot*</td><td>./traceplot</td></tr>
<tr><td>Gelman-Rubin diagnositcs (Rhat), ESS*</td><td>./summary</td></tr>
<tr><td rowspan="4">Posterior distribution analysis</td><td>Mean, SD of samples</td><td>(code print)</td></tr>
<tr><td>Plot by models*</td><td>./posterior_plot/{model}_{group}.png</td></tr>
<tr><td>Plot by parameters (for group comparison)</td><td>./posterior_plot/paramwise_{model}.png</td></tr>
<tr><td>Joint / Correlation plot</td><td>./posterior_plot/
{joint|correlation}_{model}_{group}.png</td></tr>
<tr><td rowspan="3">Posterior predictive check</td><td>Of each trial (with/without error bands)*</td><td>./posterior_predictive/
{model}_{group}(_errband).png</td></tr>
<tr><td>Block-wise trend (each block with 10 trials)</td><td>./posterior_predictive/
10block_{model}_{group}.png</td></tr>
<tr><td>Real behavior (for reference)</td><td>./posterior_predictive/gt_{group}.png</td></tr>
<tr><td>Post-hoc fit</td><td>PSIS-LOO</td><td>(code print)</td></tr>
<tr><td colspan="2">Parameter recovery test</td><td>./recovery</td></tr>
</table>

Results: Mean, SD of samples

WSLS-2

```
Group parameters of: HC
mu_ws              0.342   0.033
mu_ls              0.768   0.030
----------------------------
Group parameters of: AM
mu_ws              0.312   0.041
mu_ls              0.776   0.032
----------------------------
Group parameters of: HR
mu_ws              0.367   0.043
mu_ls              0.729   0.035
```

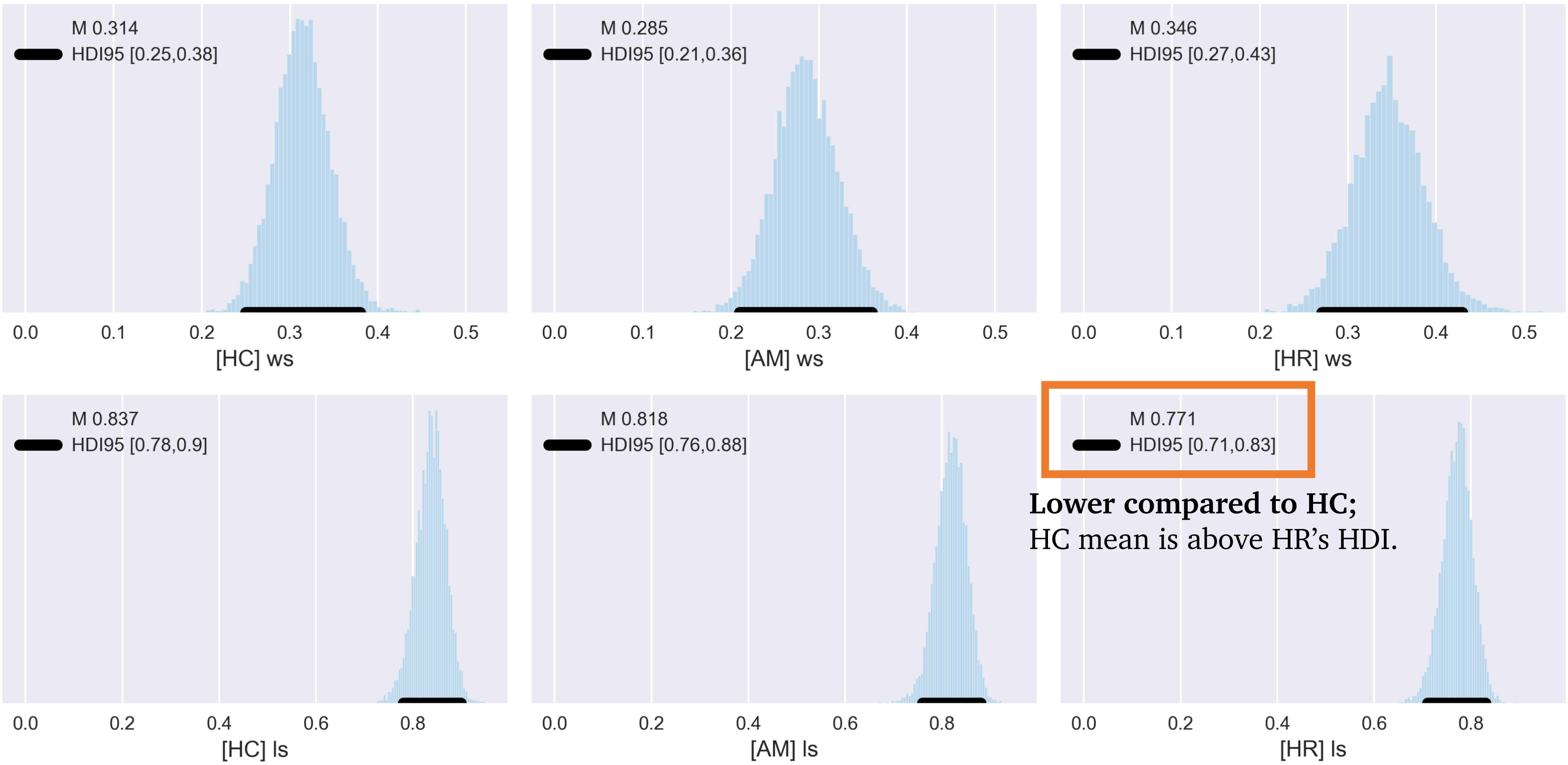
WSLS-6

```
Group parameters of: HC
mu_ws_init     0.267   0.163
mu_ls_init     0.751   0.039
mu_theta_ws    0.434   0.301
mu_theta_ls    0.006   0.006
mu_ws_fin      0.367   0.150
mu_ls_fin      0.660   0.214
----------------------------
Group parameters of: AM
mu_ws_init     0.331   0.206
mu_ls_init     0.913   0.053
mu_theta_ws    0.552   0.293
mu_theta_ls    0.019   0.010
mu_ws_fin      0.287   0.105
mu_ls_fin      0.563   0.109
----------------------------
Group parameters of: HR
mu_ws_init     0.183   0.133
mu_ls_init     0.773   0.042
mu_theta_ws    0.433   0.269
mu_theta_ls    0.007   0.002
mu_ws_fin      0.499   0.128
mu_ls_fin      0.480   0.260
----------------------------
```

VPP

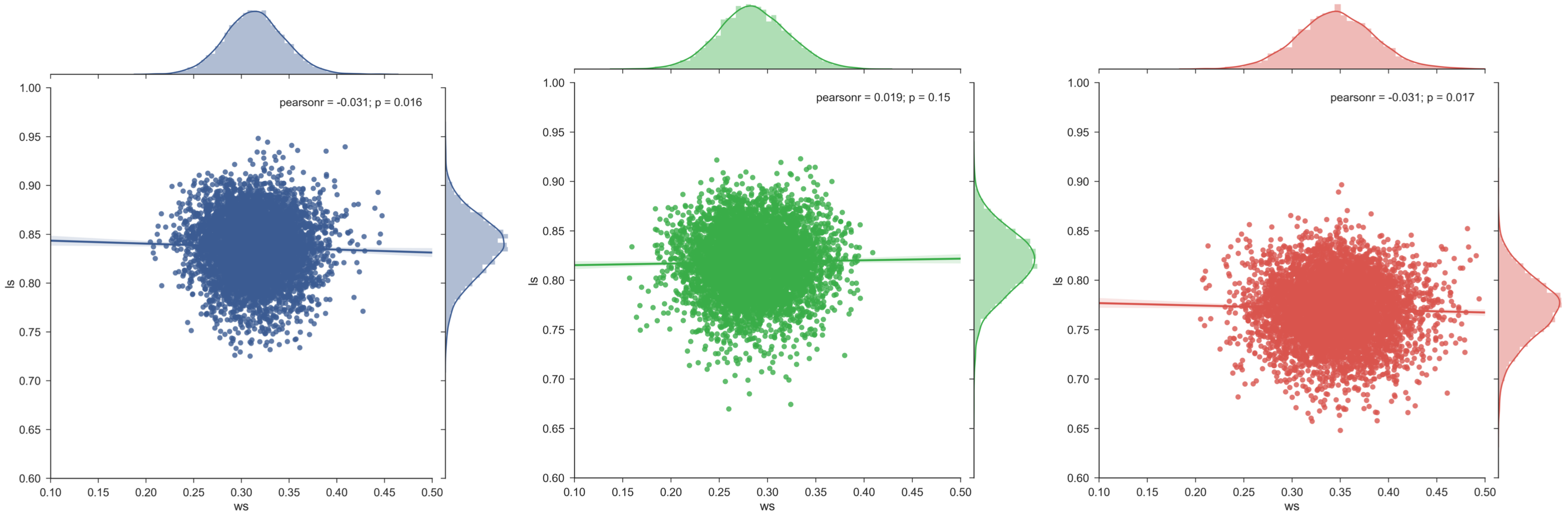
```
Group parameters of: HC
mu_alpha        0.545   0.224
mu_lambda       0.777   0.423
mu_phi          0.008   0.007
mu_k            0.489   0.058
mu_ep_pos       0.030   1.479
mu_ep_neg      -5.353   3.890
mu_w            0.960   0.043
mu_c            1.893   0.466
----------------------------
Group parameters of: AM
mu_alpha        0.390   0.221
mu_lambda       0.608   0.494
mu_phi          0.013   0.007
mu_k            0.408   0.068
mu_ep_pos      -1.644   2.106
mu_ep_neg      -4.701   3.512
mu_w            0.960   0.041
mu_c            1.779   0.321
----------------------------
Group parameters of: HR
mu_alpha        0.564   0.189
mu_lambda       0.052   0.178
mu_phi          0.070   0.047
mu_k            0.336   0.074
mu_ep_pos       0.626   2.628
mu_ep_neg      -2.421   2.789
mu_w            0.932   0.052
mu_c            1.109   0.121
```

Results: Posterior Distribution [WSLS-2]



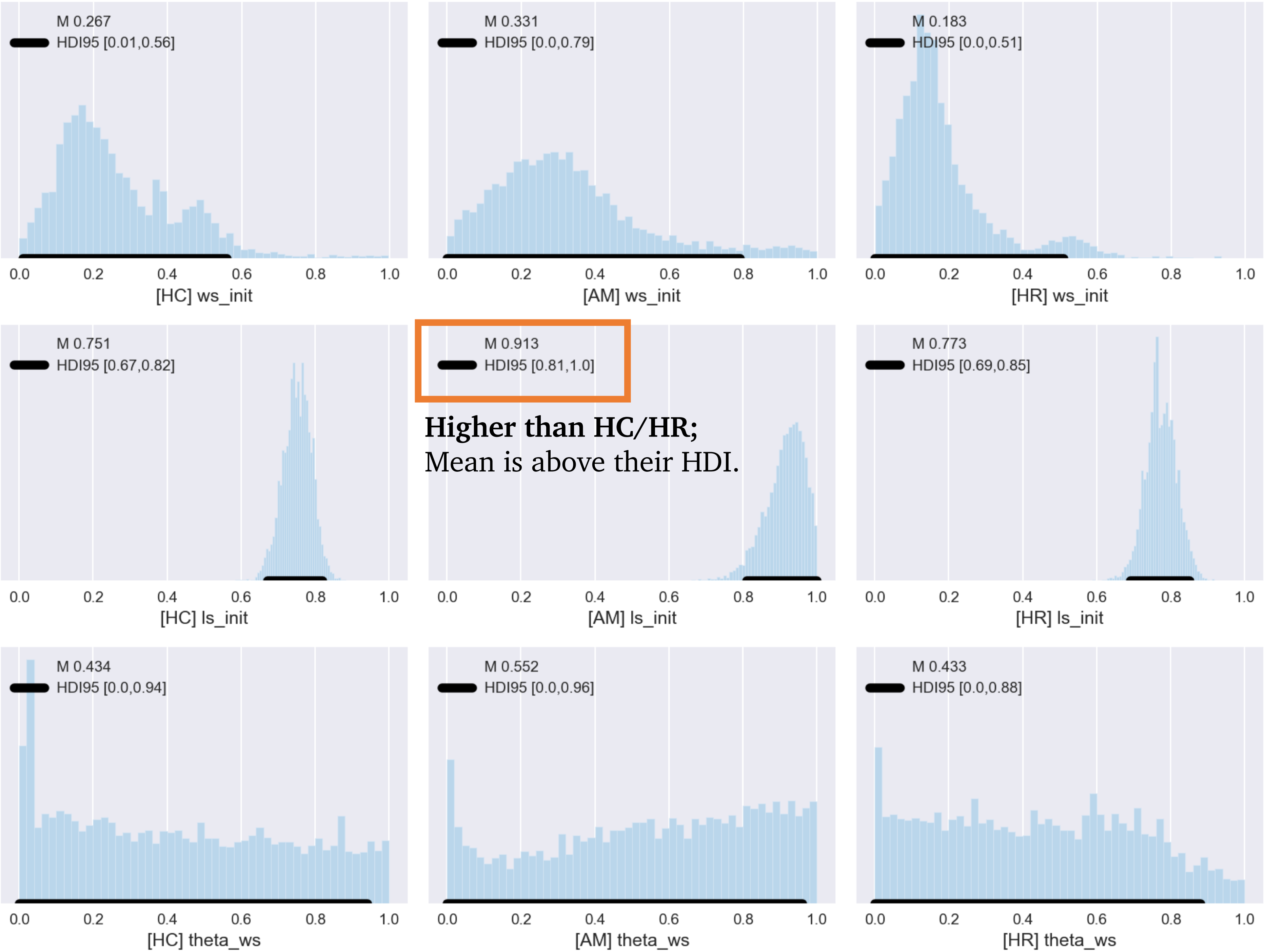
Lower compared to HC;
HC mean is above HR's HDI.

Results: Posterior Distribution [WSLS-2] (cont'd)



- Difference between HR and HC/AM is readily shown through visual inspection of the joint plot.
- HR has higher *ws* (yet in others' HDIs) and significantly lower *ls* compared to other groups.
- Parameters do not suffer from high correlation.

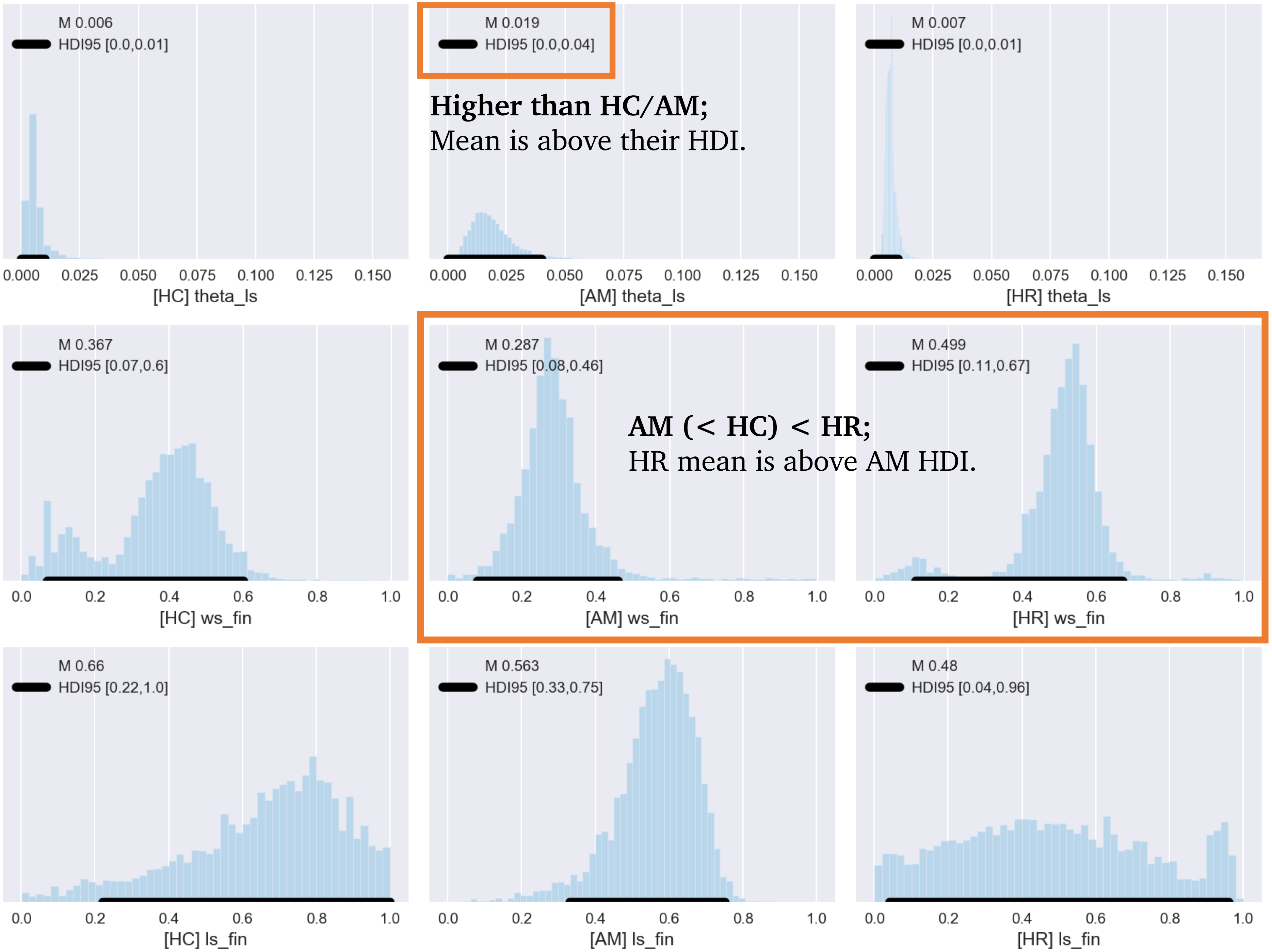
Results: Posterior Distribution [WSLS-6]



Higher than HC/HR;
Mean is above their HDI.

θ shows very wide HDI.

Results: Posterior Distribution [WSLS-6] (cont'd)

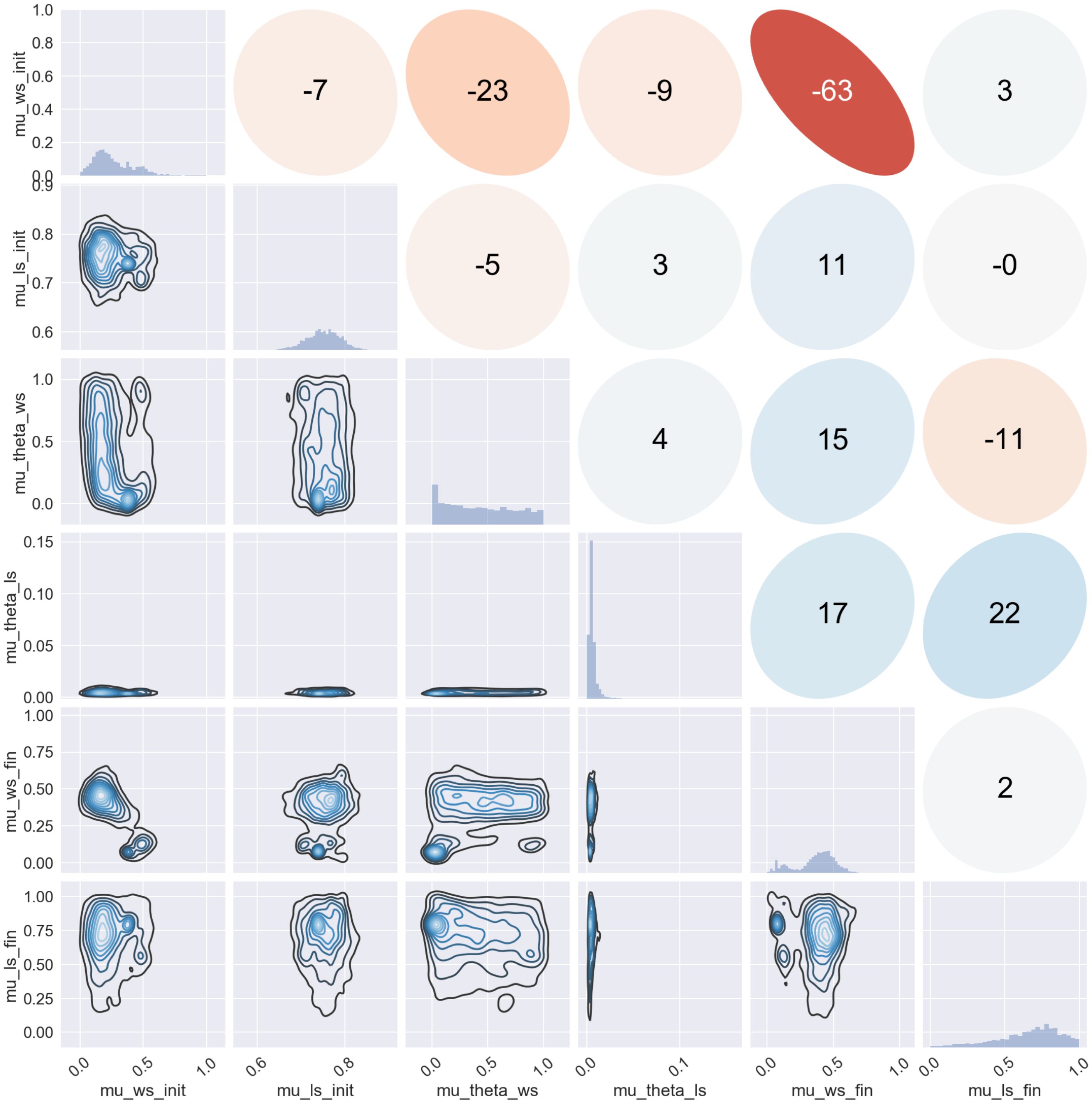


Higher than HC/AM;
Mean is above their HDI.

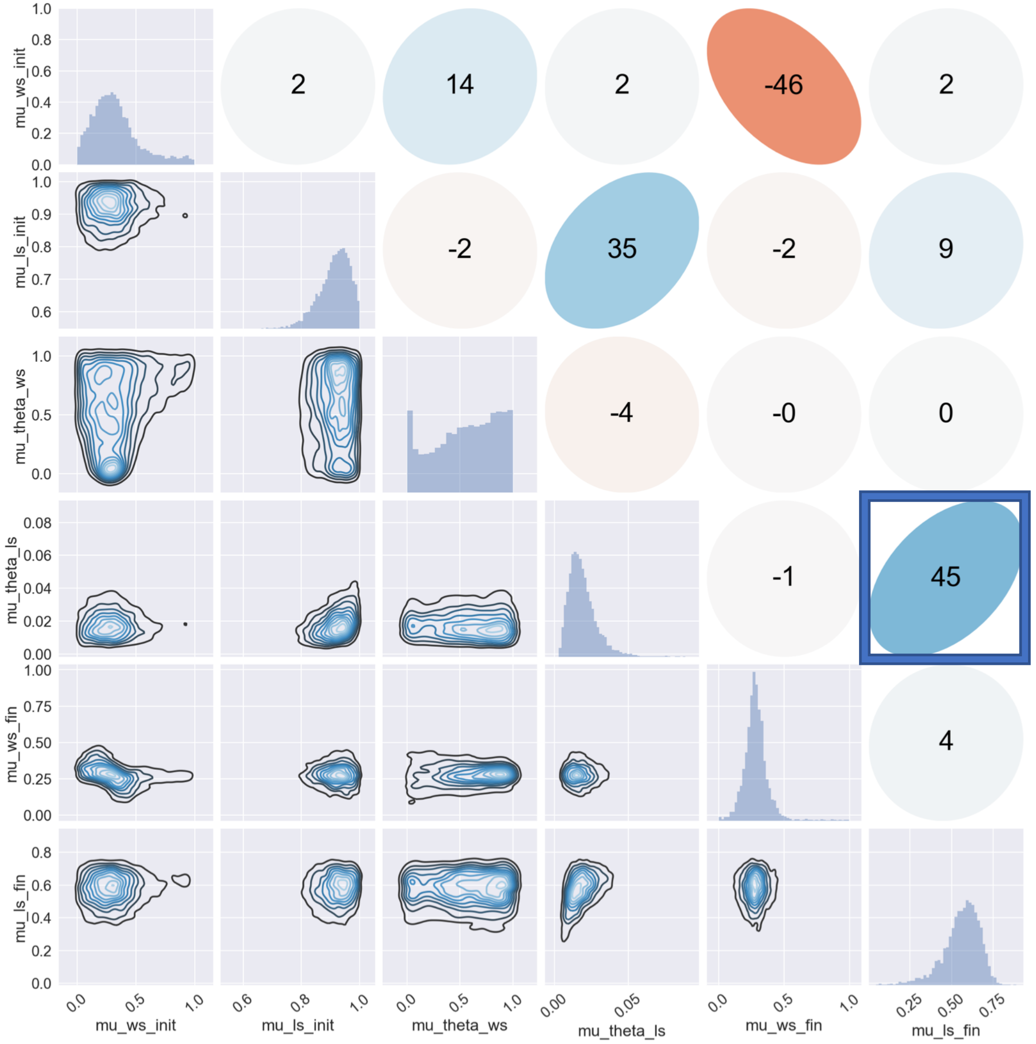
AM (< HC) < HR;
HR mean is above AM HDI.

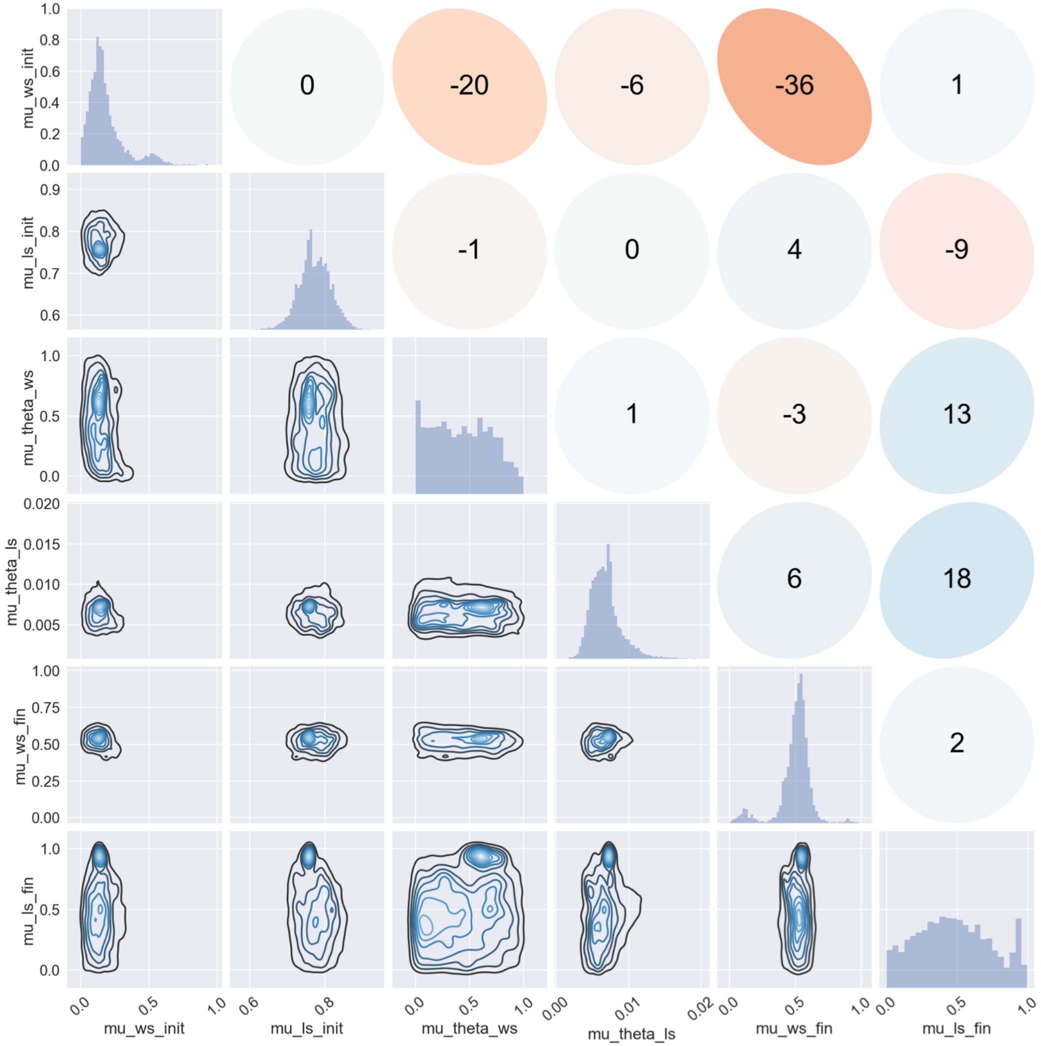
Results: Parameter Correlation [WSLS-6]

- ws_init and ws_fin show moderate negative correlation (-.63).
- However, since the conceptually parallel ls_init and ls_fin pair exhibit zero correlation, it seems fair to maintain initial and final values as parameters.
- Overall, this model has a good parameter scheme in terms of low correlation.



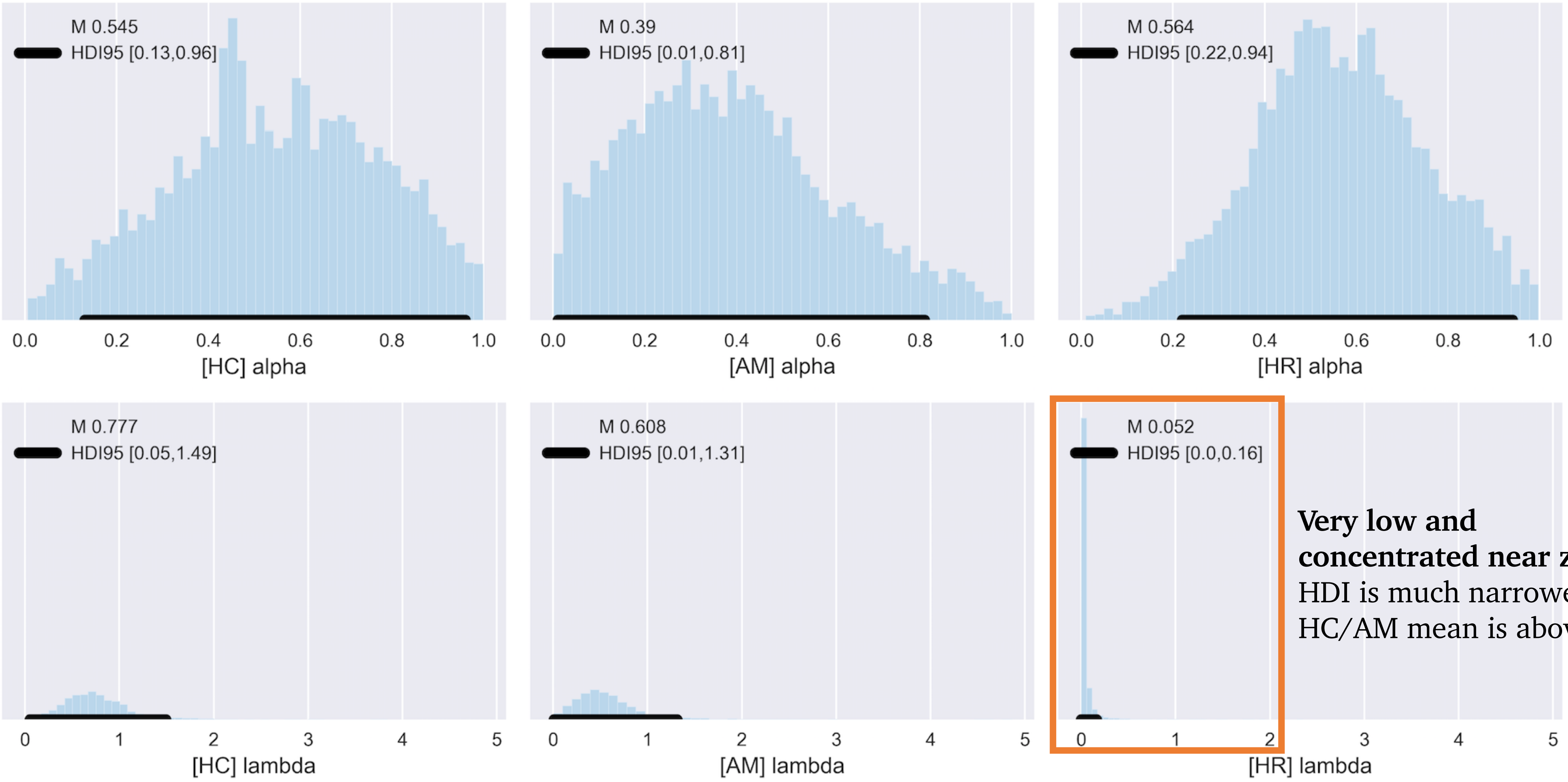
Results: Parameter Correlation [WSLS-6] (cont'd)





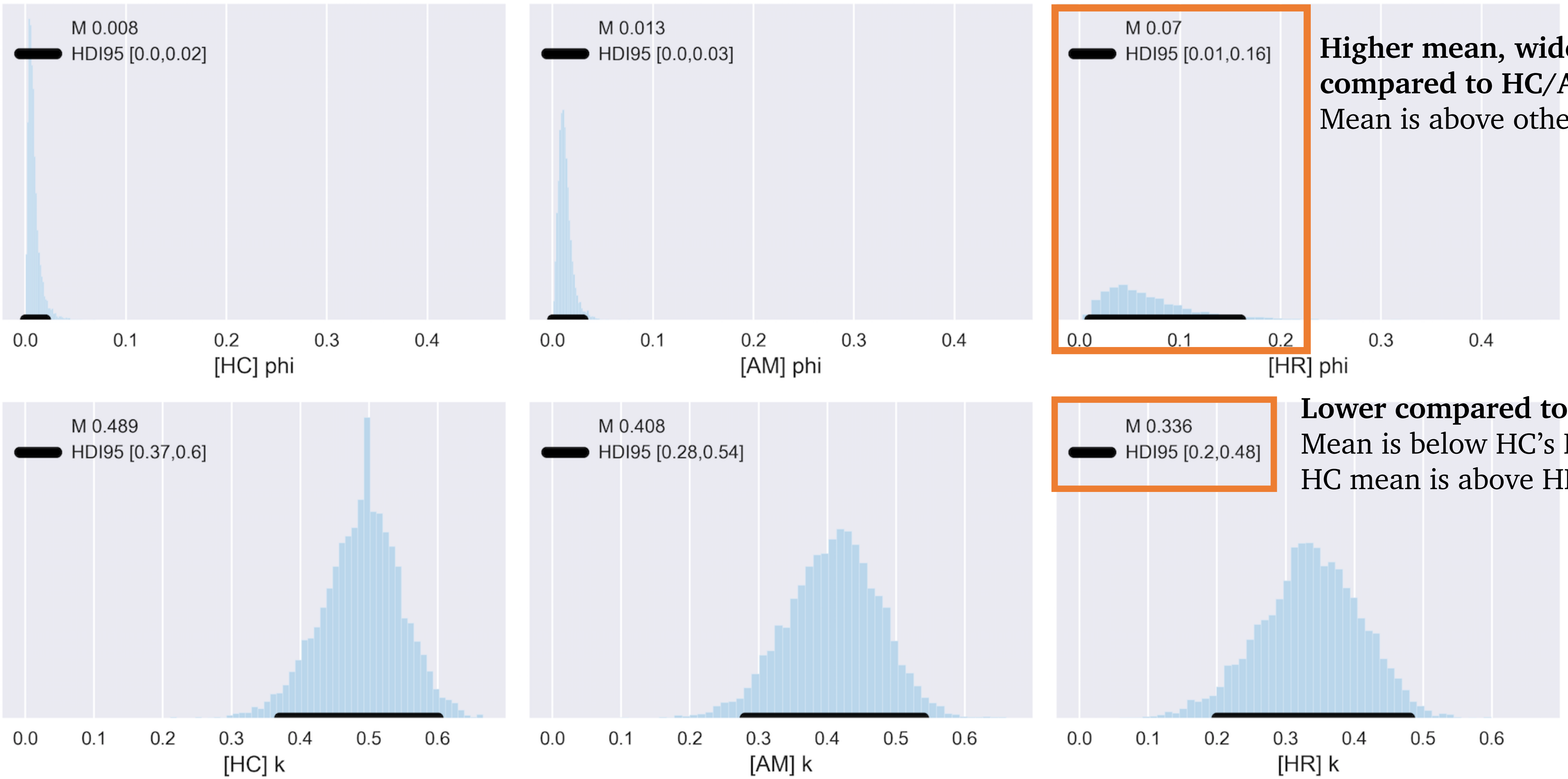
- AM(left) shows new moderate correlation between change rate and final value and change rate of ls.
- Still, both groups show similar pattern to HC.

Results: Posterior Distribution [VPP]



Very low and concentrated near zero;
HDI is much narrower.
HC/AM mean is above HR's HDI.

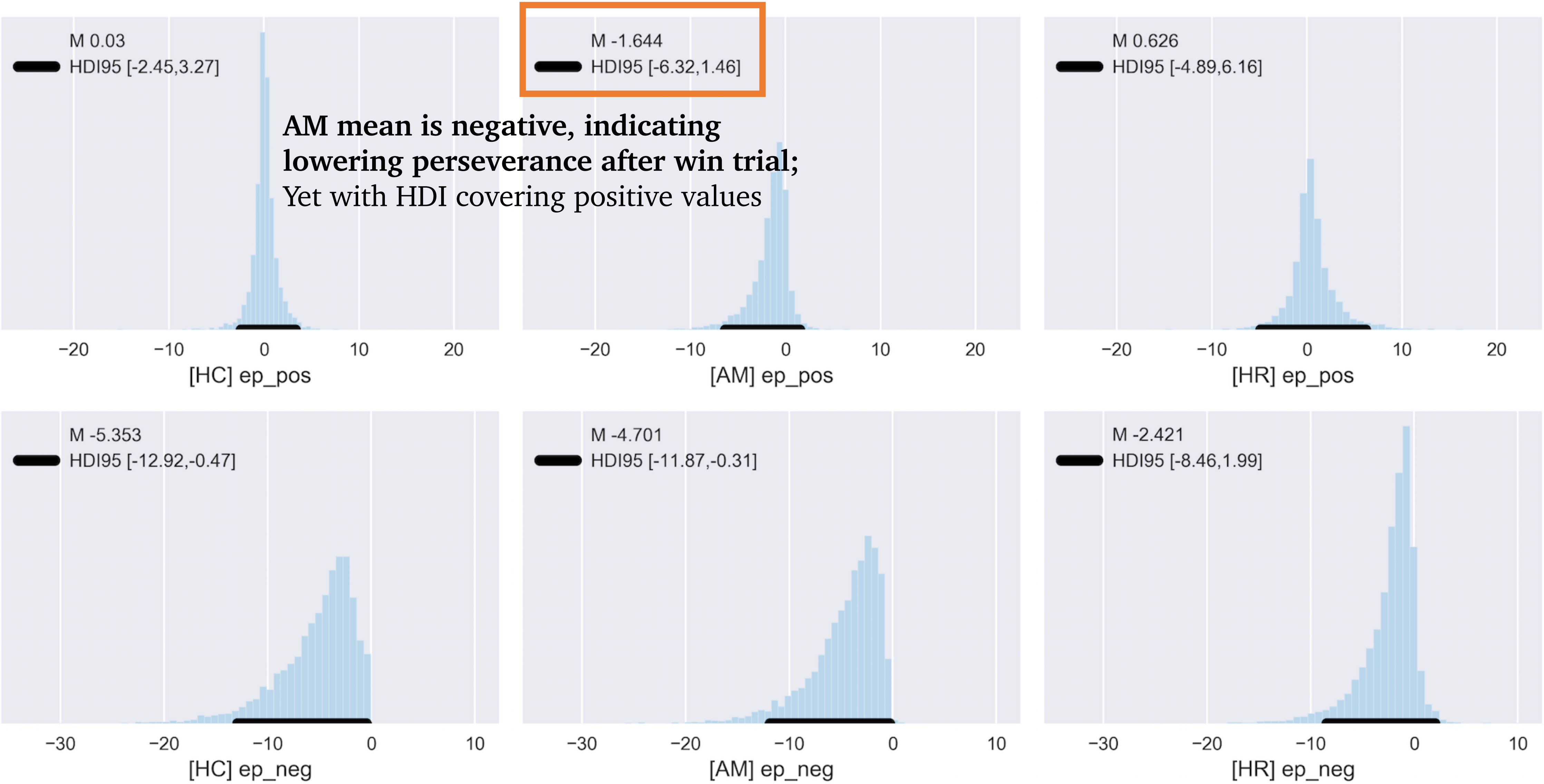
Results: Posterior Distribution [VPP] (cont'd)



Higher mean, wider HDI compared to HC/AM;
Mean is above others' HDI

Lower compared to HC;
Mean is below HC's HDI,
HC mean is above HR's HDI

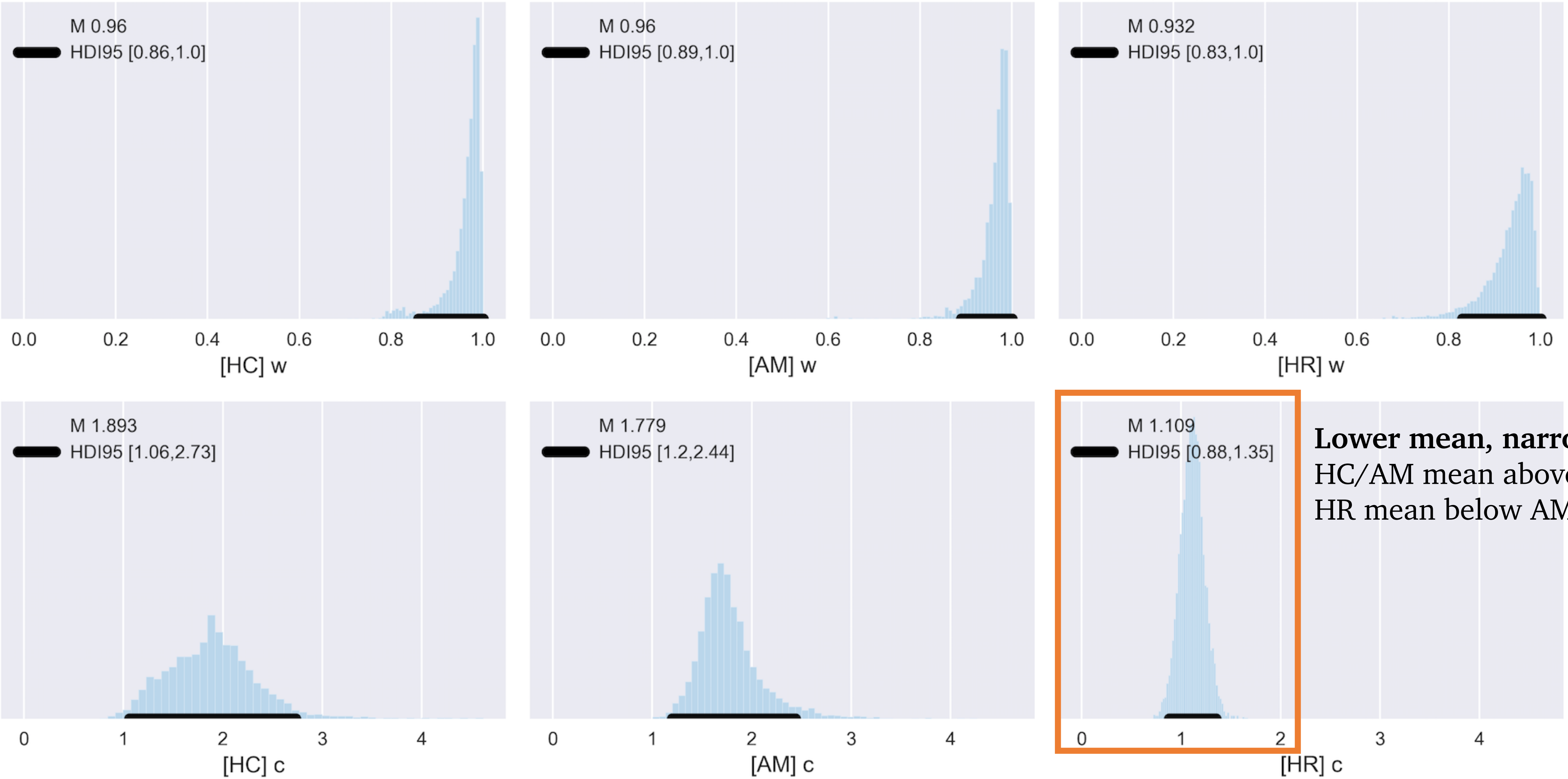
Results: Posterior Distribution [VPP] (cont'd)



AM mean is negative, indicating lowering perseverance after win trial;
Yet with HDI covering positive values

Results: Posterior Distribution [VPP] (cont'd)

All groups give most of the weight to EV rather than perseverance ($w \gg 0.5$)

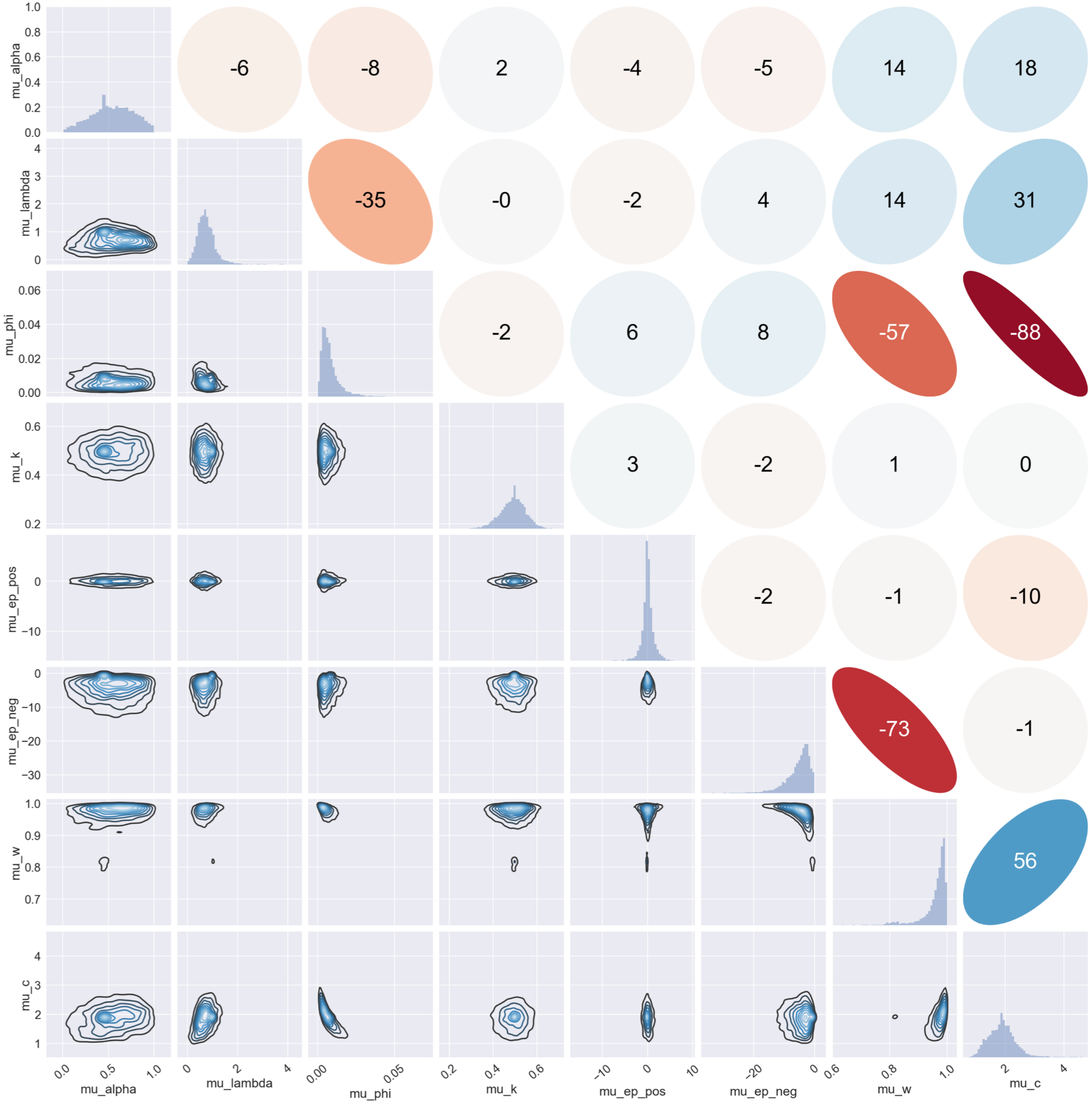


Lower mean, narrower HDI;
HC/AM mean above HR's HDI,
HR mean below AM' HDI.

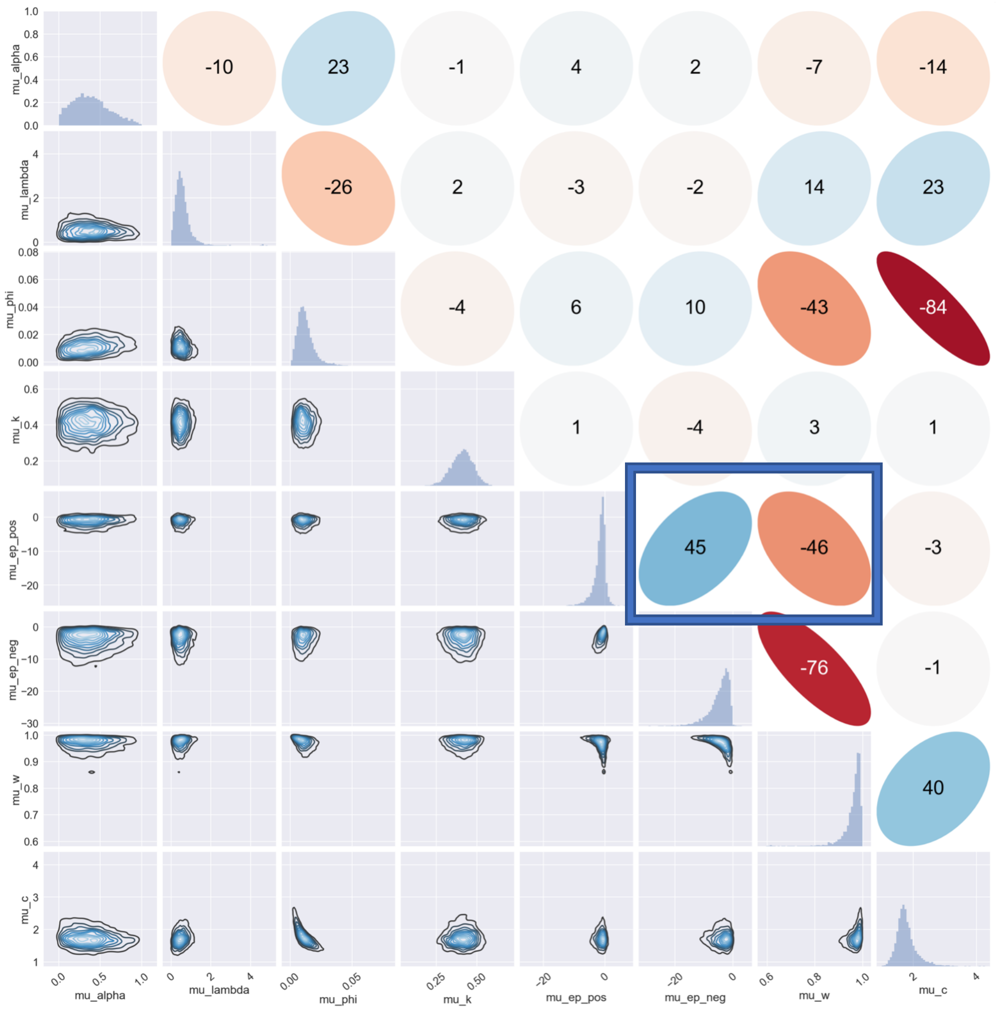
Results: Parameter Correlation [VPP]

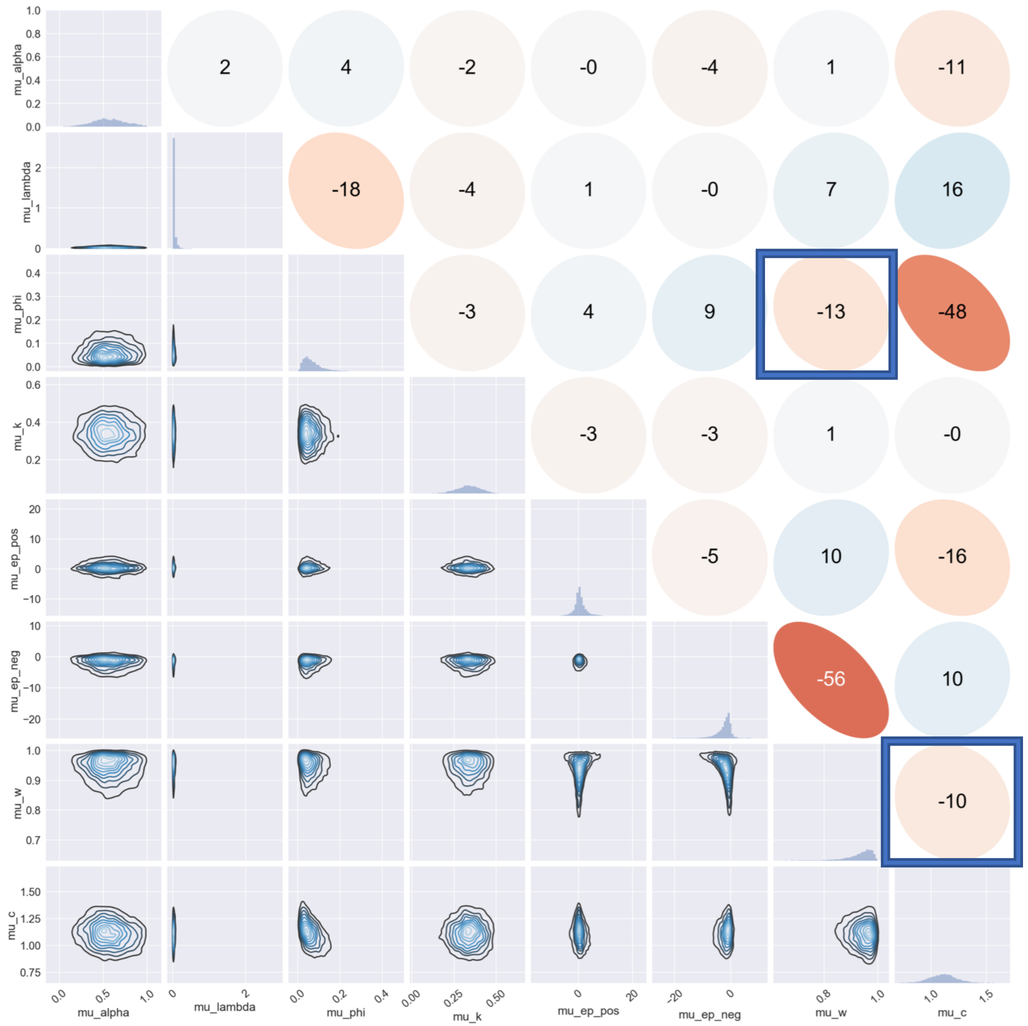
- Some parameters **overlap** in HC.
- Conceptually predictable

- Exploitation(c) - Recency(ϕ) (-.88)
  - A more consistent EV-based choice implies being less affected by recent outcomes
- Exploitation(c) - Perseverance bias after negative outcome($\boldsymbol{\varepsilon}_{neg}$) (-.73)
  - A more consistent EV-based choice implies less staying after loss
- Exploitation(c) - EV weight(w) (+.56)
  - A more consistent EV-based choice indicates relying more on EV
- Recency(ϕ) - EV weight(w) (-.57)
  - Being more affected by recent outcomes may indicate relying less on EV



Results: Parameter Correlation [VPP] (cont'd)

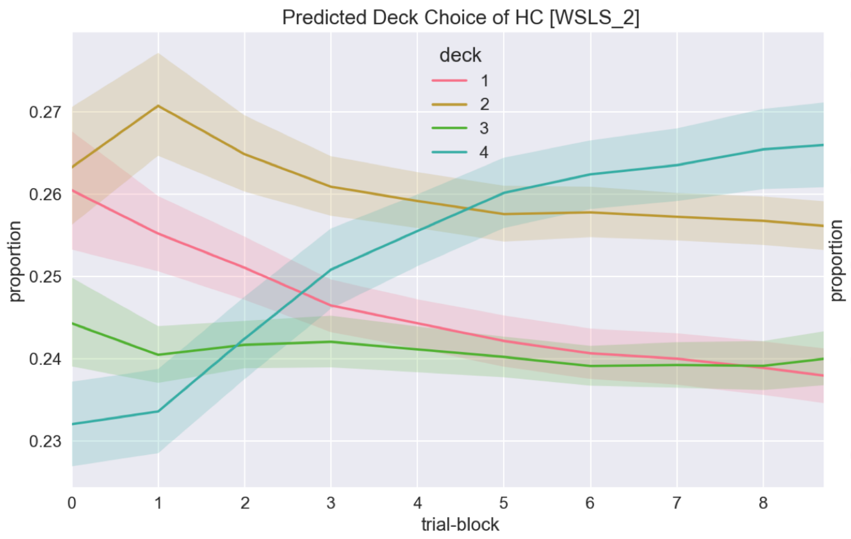


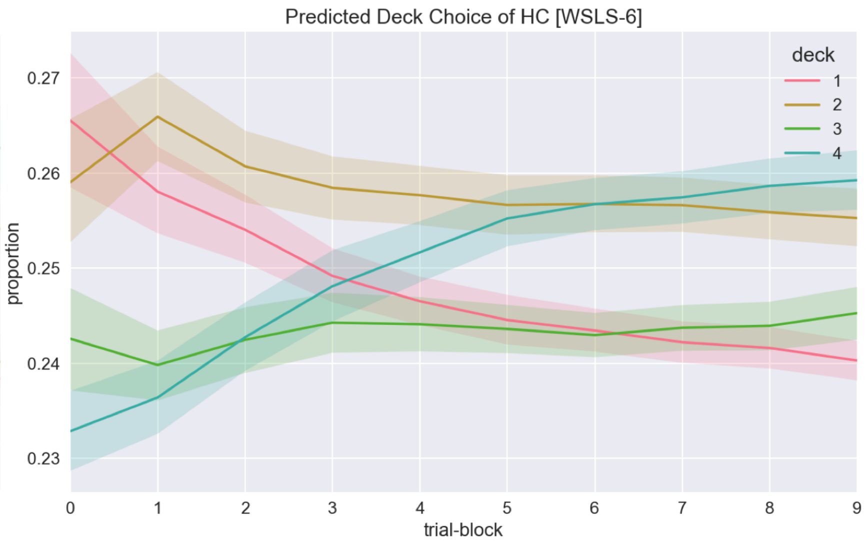


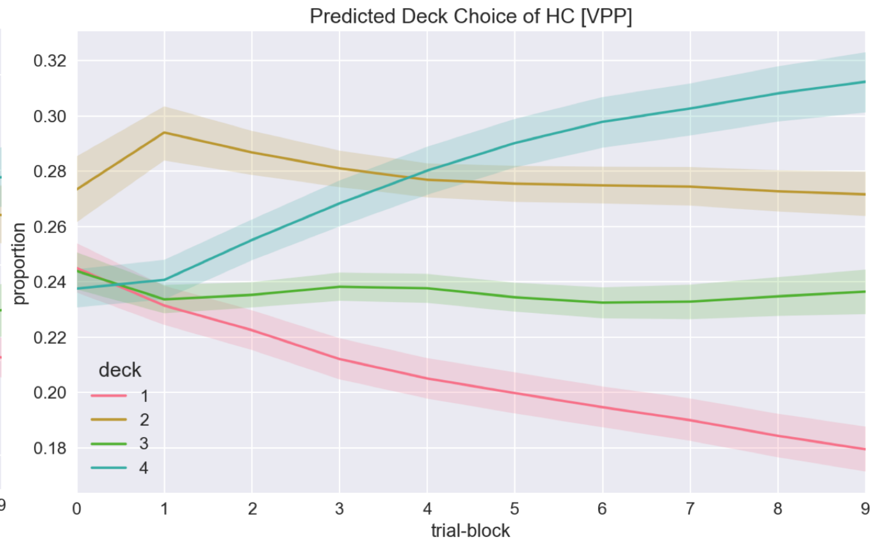
- AM(left) shows new correlations that are absent in HC/HR.
- Both involves $\boldsymbol{\varepsilon}_{pos}$, which was distinctively negative in AM.

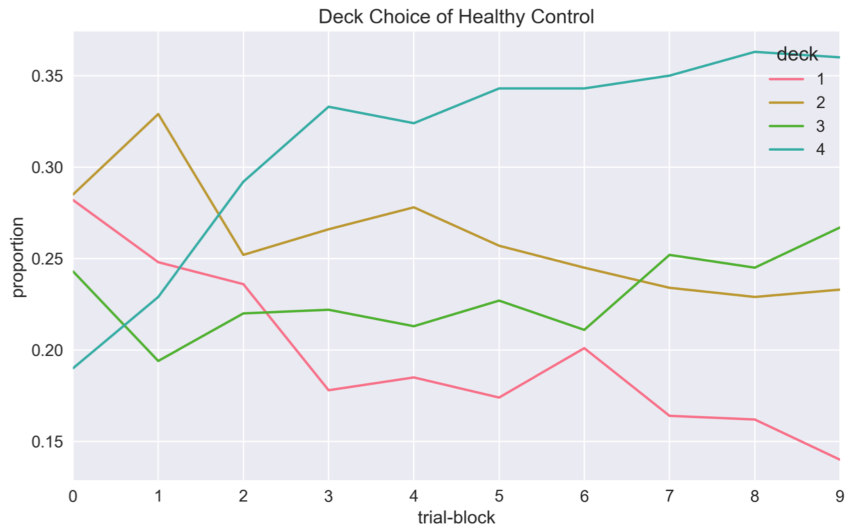
- *c* exhibits a smaller, and *w* exhibits (almost) no correlation with others in HR(right), which were present in HC.

Results: Posterior Predictive Check [HC]



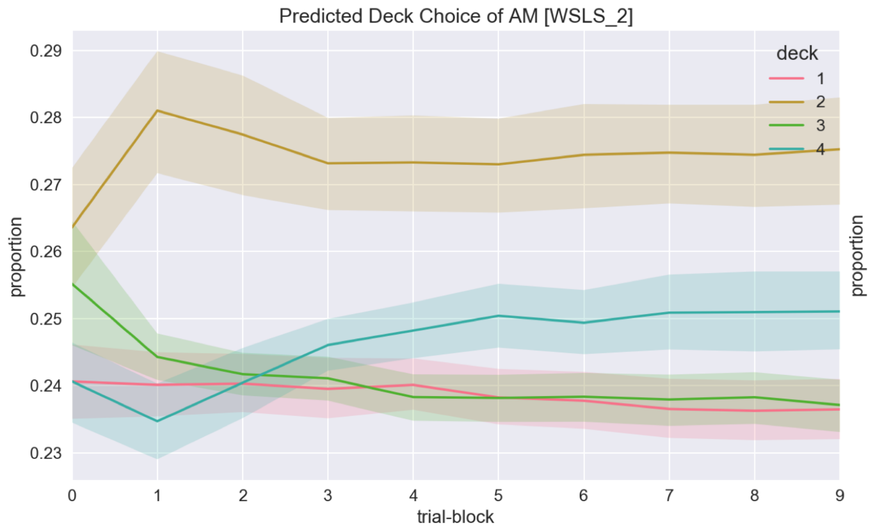


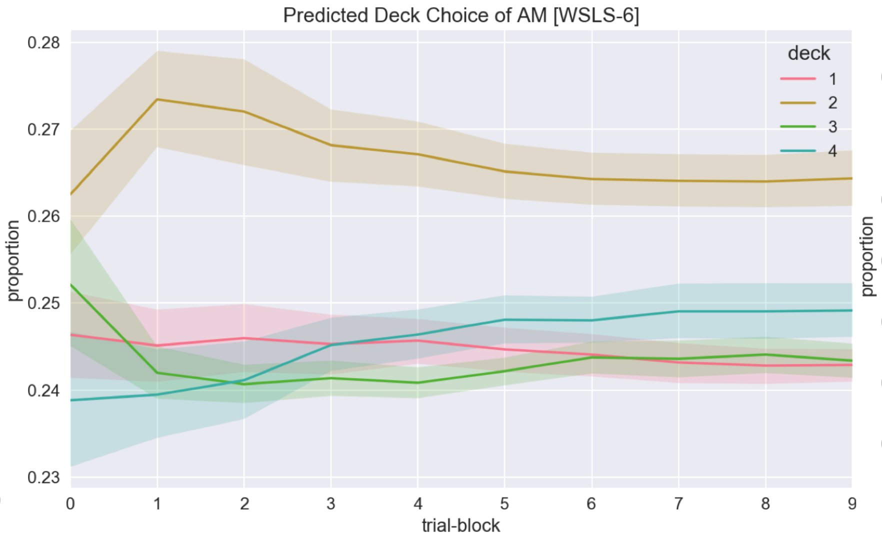


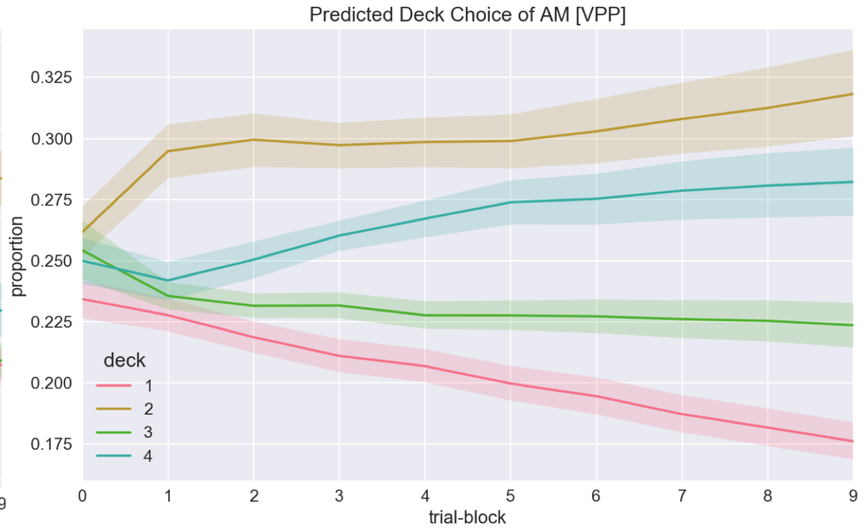


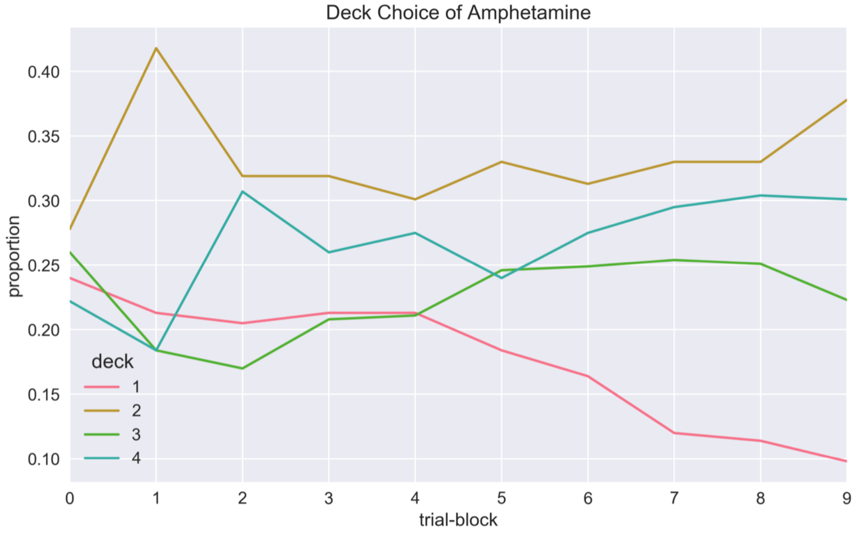
- Real choice data (bottom left) shows a **relatively strong preference for deck 4** (with positive outcome overall and less frequent loss), and avoidance of deck 1 (with negative outcome overall and more frequent loss). This pattern quickly settles after the first block. Decks with less frequent loss (2 and 4) are preferred to their counterparts (1 and 3, respectively).
- Three models also predict winning of deck 4, but at later trials than the actual behavior. **VPP** predicts the fastest transition to deck 4, but its asymptotic ending proportion is not as high as the actual data. Other models predict a even lower ending value for deck 4, and the differences between each deck is not as high as the actual data or VPP's prediction.

Results: Posterior Predictive Check [AM]





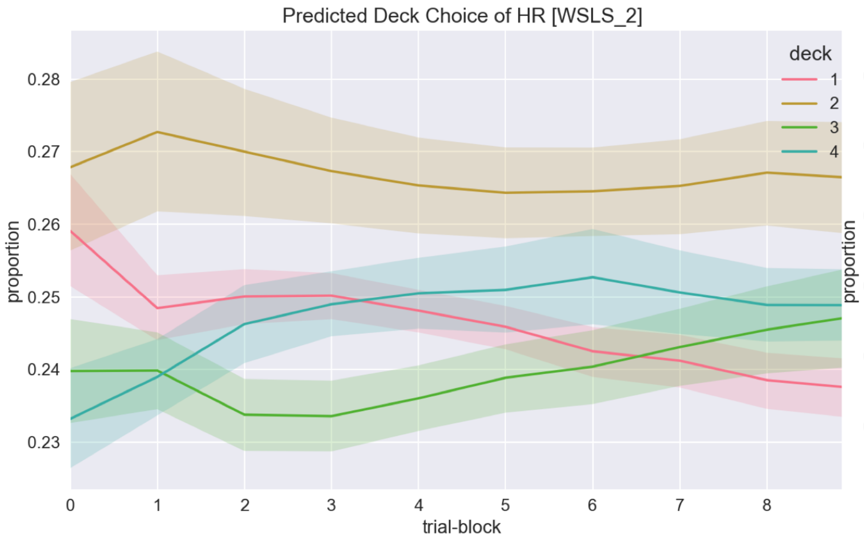


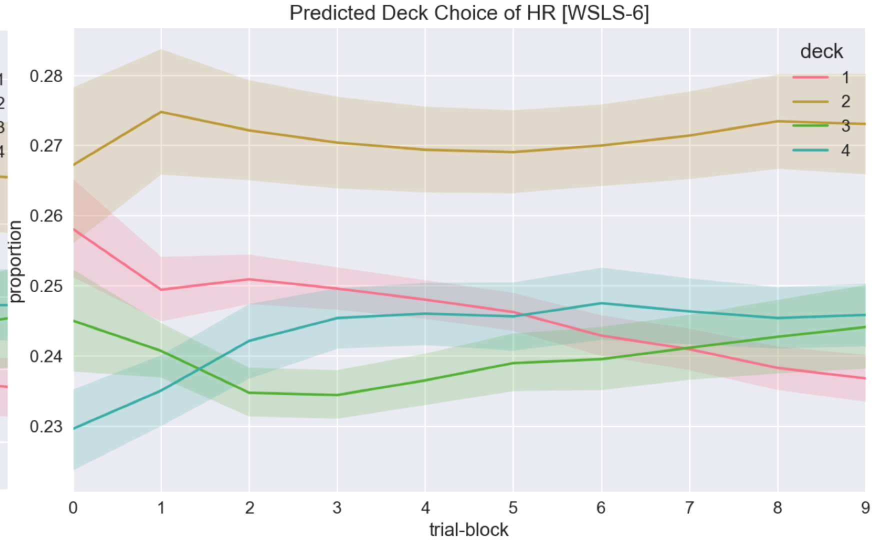


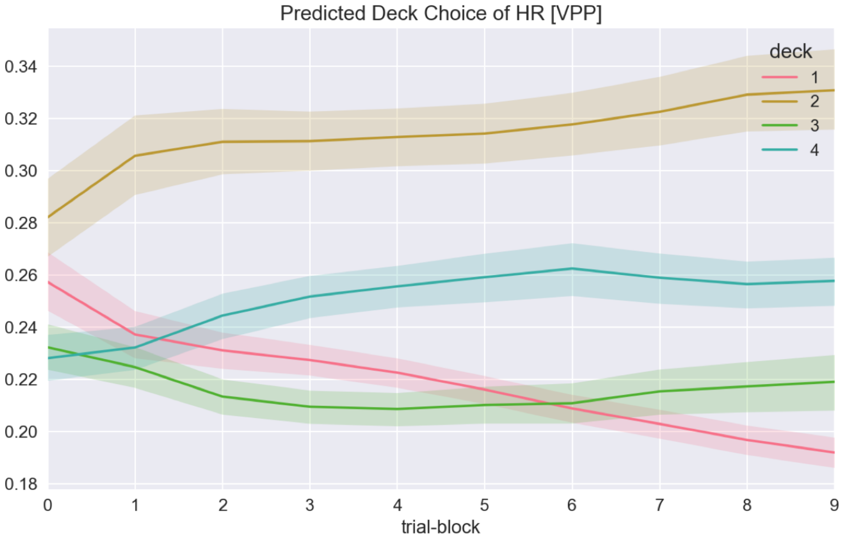
- Unlike HC, AM subjects show **preference for deck 2**, which gives bigger gains although disadvantageous overall. However, the difference between deck 2 and 4 is not big, especially from the second block. Strong avoidance of deck 1 as well as preference for less frequent loss decks still hold.

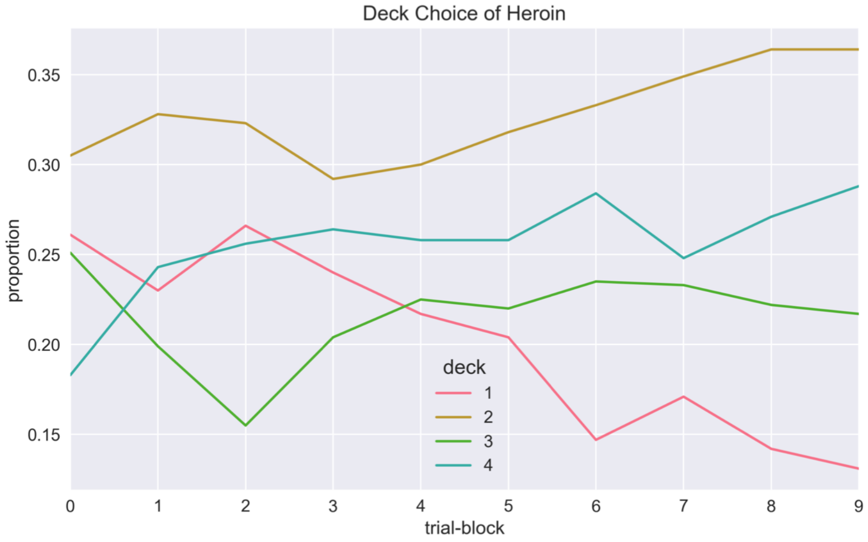
- Among the three models, **VPP** shows the best prediction. Other two models predict winning of deck 2, however, the proportion is not as high. Avoidance of deck 1 is not strongly predicted.

Results: Posterior Predictive Check [HR]









- HR subjects shows a similar pattern to AM, but with a higher preference of deck 1 and 2 giving big gains. In the earlier blocks, deck 1 is explored more often compared to other groups. Difference between deck 2 and 3 is also clearer than in AM.

- A strong preference for deck 2 is clearly predicted by **VPP**. Compared to actual behavior, VPP predicts a faster departure from deck 1. Other models fail to capture the losing of deck 1. They also predict a smaller win for deck 2.

Results: Post-hoc Fit

- PSIS-LOO (Vehtari et al., 2017)
  - Model comparison by computing the standard errors of expected log pointwise predictive density for a new dataset (elpd).
- VPP > WSLS-2 ≫ WSLS-6 (for HC and AM).
  - However, with many unreliable estimates (k > 0.7)
  - For HR, log likelihood estimate for a single subject kept being NaN throughout multiple samplings for some reason I could not figure out, hence the NaN elpd here.

```
PSIS leave-one-out cross validation.
Estimates are unreliable if k > 0.7

VPP
-     ELPD
HC   -5870.621  Note: there are 46 unreliable points.
AM   -4585.050  Note: there are 37 unreliable points.
HR   -5116.428  Note: there are 42 unreliable points.

WSLS-2
-     ELPD
HC   -6271.139  Note: there are 46 unreliable points.
AM   -4930.923  Note: there are 36 unreliable points.
HR   -5469.343  Note: there are 42 unreliable points.

WSLS-6
-     ELPD
HC   -11053.387  Note: there are 48 unreliable points.
AM   -8543.726  Note: there are 38 unreliable points.
HR     nan
---------------------------
Model Comparisons: Estimate of ELPD difference

VPP      - WSLS-2
    estimate   SE
HC   400.518   53.152
AM   345.873   48.068
HR   352.915   52.926

VPP      - WSLS-6
    estimate   SE
HC   5182.766   1013.147
AM   3958.676   953.221
HR     nan      nan

WSLS-2   - WSLS-6
    estimate   SE
HC   4782.249   1007.699
AM   3612.803   954.201
HR     nan      nan
```

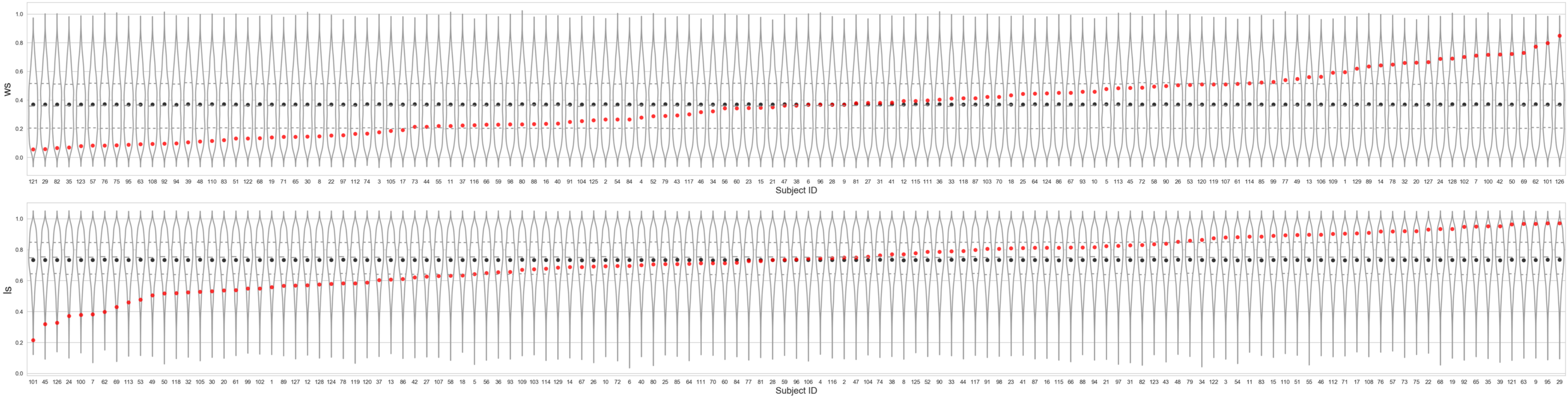
Results: Parameter Recovery Test

- Gain/loss generation
  - Since the exact gain/loss data for the current dataset was not known, I generated each deck's outcome following the prescribed gain/loss policy: deck A, B always yield bigger gains; deck A, C has loss probability of 0.5 and deck B, D has 0.1; each block yields certain amount of net outcome.
  - Partially referring to the gain/loss instances in the original data, I set the gain range to [40, 80] for C, D and [80, 140] for A, B. Gain range does not change over block, and each gain is a factor of 5. For implementation ease, loss is neither constrained in range nor it is a factor of 10 (unlike in the original data). Still, each block's net outcome strictly sums up to the assigned amount.
  - A single set of gain/loss data is used for testing all models.
- Generated IGT result for all models are in ./simulData

- Subjects from 3 groups are combined and sorted by "true" parameters.
  - Red dots mark the "true" parameters (estimated parameter mean of each individual).
  - Black dots mark the means of recovered parameter values. Tick marks in violins indicate quantiles.

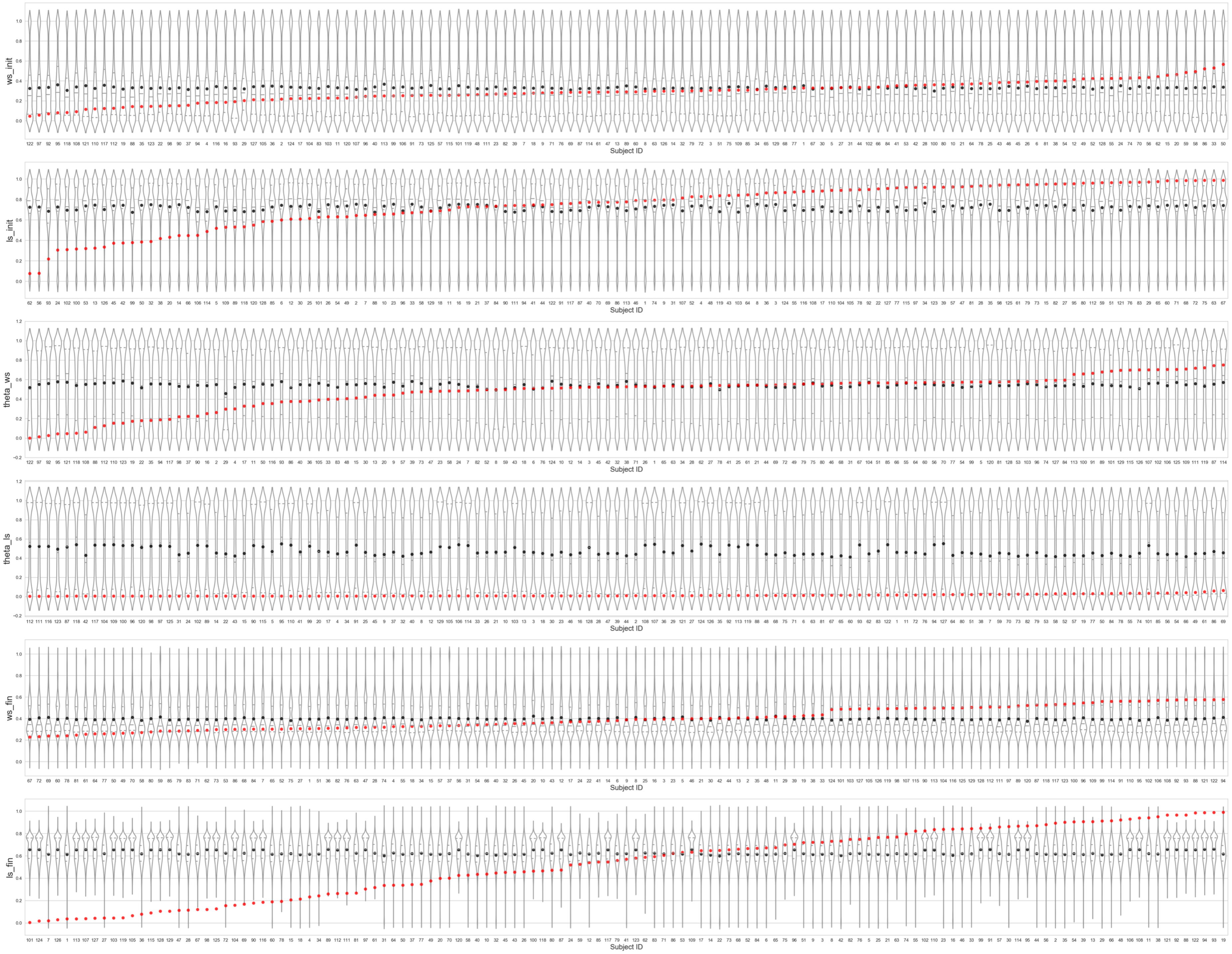
Results: Parameter Recovery Test [WSLS-2]

- Both parameters are poorly recovered.
- Recovered values are almost uniform for all subjects.



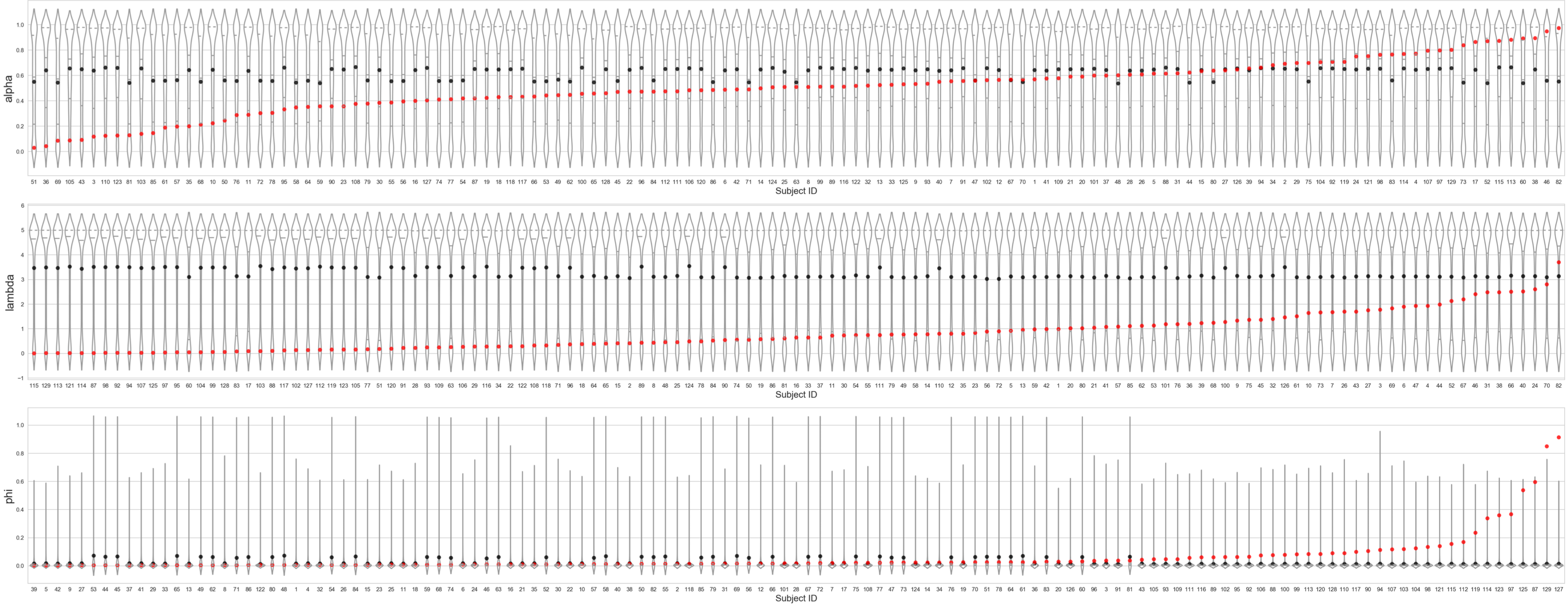
Results: Parameter Recovery Test [WSLS-6]

- Probably aside ws_fin, parameters are poorly recovered.
- θ_ls is systematically overestimated.
- Recovered parameter means do not show drastic inter-subject variance.



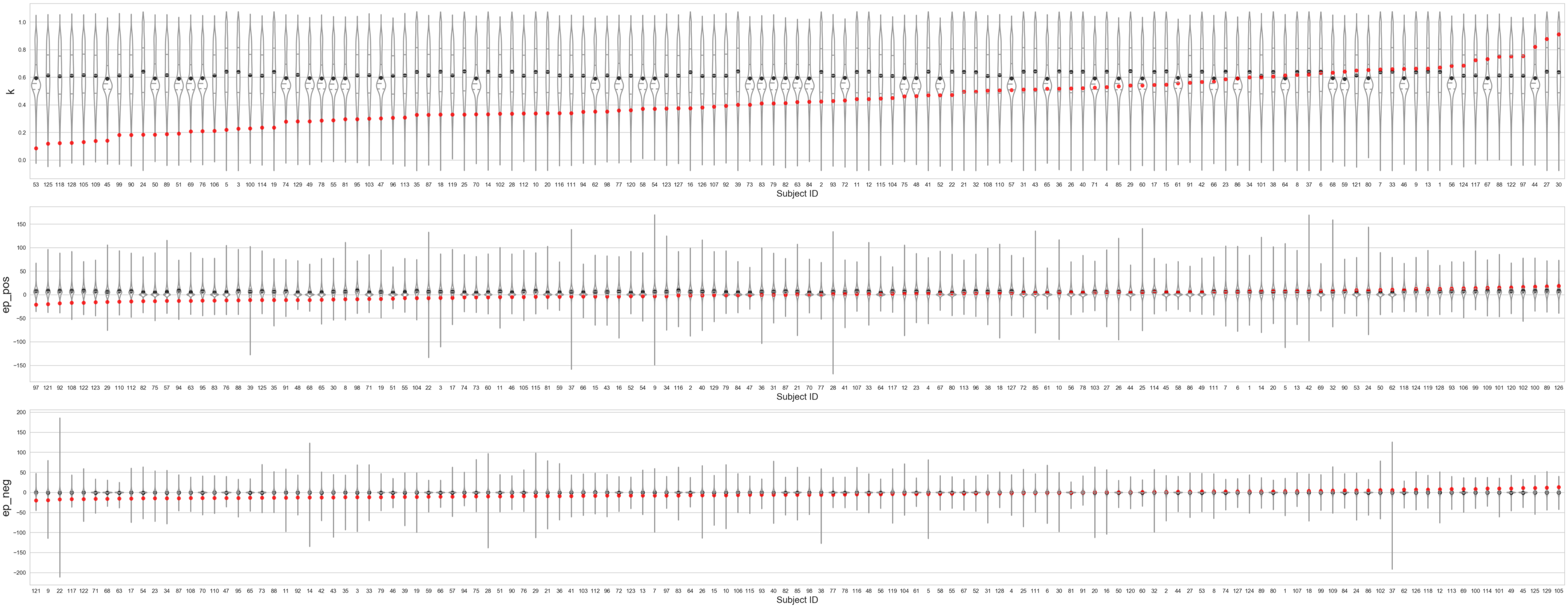
Results: Parameter Recovery Test [VPP]

- λ is systematically overestimated.
- ϕ is estimated near zero for all subjects.



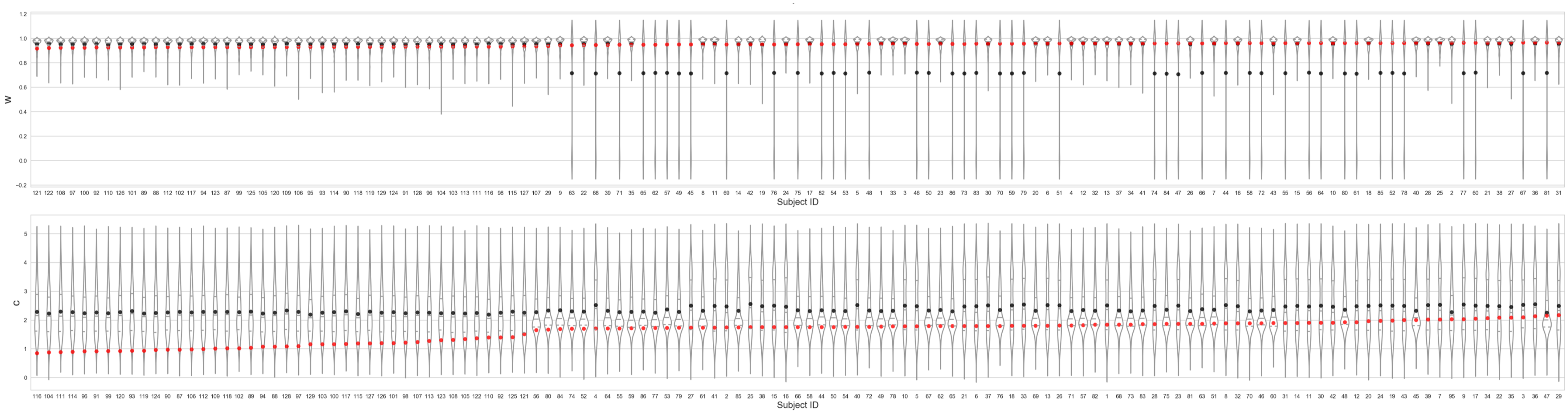
Results: Parameter Recovery Test [VPP] (cont'd)

- ε_pos and ε_neg are well recovered; k is not.



Results: Parameter Recovery Test [VPP] (cont'd)

- While "true" values of w are near 1.0, recovered values for 1/3 of the subjects are around 0.7
- c is systematically overestimated.



- Overall, there seems to be no solid "winning" model (among WSLS-2, WSLS-6 and VPP) by qualitative evaluation of parameter recovery.

Discussion & Conclusion: Model Selection

- VPP model shows that all three groups weight EV much more than perseverance (mean w > 0.9 for all groups). This explains VPP's superior result in posterior predictive test and post-hoc fit compared to other two models, which are only based on perseverance.

- When modeling IGT or, more generally, bandit problems/decision making, dual strategy or solely RL-based models seems to be more promising than simple perseverance models.

- Although VPP shows the best post-hoc fit and posterior predictive results among the three models, it has high correlation between certain parameters, exploitation(c) in particular. This suggests that VPP can take a more condensed form, with c reduced to other parameters in the model. However, considering that VPP also suffers from poor parameter recovery, it might be the case that current parameter combinations of the model are largely misspecified.

Discussion & Conclusion: Group Comparison

- Although EV is a dominant strategy, choice by heroin users were less EV-based and showed higher perseverance to recent gains, compared to healthy controls or amphetamine users. This is shown by parameter estimates in the three models, and is in line with findings in Ahn et al. (2014).
  - Lower probability of shifting after loss trial (*ls*) in WSLS-2
  - Higher asymptotic final probability of staying after win trial (*ws_fin*) in WSLS-6
  - Lower loss aversion (λ), perseverance decay (k) and exploitation of EV (c), as well as higher weight on recent outcomes (ϕ) in VPP.

- Amphetamine users were estimated to have a higher initial value (*ls_init)* and change rate (θ_*ls*) for probability of shifting after loss trial. This might indicate that they start with high loss aversion, however, become less sensitive to loss as they go through trials, and finally end at a value (*ls_fin*) which is not higher (actually a lower mean than HC) compared to other groups. The posterior predictive results show that AM prefer a deck with bigger gains, although it is disadvantageous overall.

References

- Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science*, *32*(8), 1376-1402.
- Ahn, W. Y., Vasilev, G., Lee, S. H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., & Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Frontiers in Psychology*, *5*, 849.
- Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing neuro-computational mechanisms of reinforcement learning and decision-making with the hBayesDM package. Computational Psychiatry, 1, 24-57. https://doi.org/10.1162/CPSY_a_00002
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7-15.
- Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, *39*(4), 376-389.
- Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychological assessment*, *14*(3), 253.
- Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413-1432.
- Worthy, D. A., Pang, B., & Byrne, K. A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. *Frontiers in psychology*, *4*, 640.
- Worthy, D. A., & Maddox, W. T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to WK Estes. *Journal of mathematical psychology*, *59*, 41-49.