# How do we learn about the value of others' advice

**Yoonseo Zoh**

**Hyeonjin Kim**

Spring 2018 : Computational Modeling Course

2018.06.08

# Introduction

## Leong & Zaki (2017)

Journal of Experimental Psychology: General



0096-3445/17/$12.00 http://dx.doi.org/10.1037/xge0000382

### Unrealistic Optimism in Advice Taking: A Computational Account

Yuan Chang Leong and Jamil Zaki
Stanford University

Expert advisors often make surprisingly inaccurate predictions about the future, yet people heed their suggestions nonetheless. Here we provide a novel, computational account of this unrealistic optimism in advice taking. Across 3 studies, participants observed as advisors predicted the performance of a stock. Advisors varied in their accuracy, performing reliably above, at, or below chance. Despite repeated feedback, participants exhibited inflated perceptions of advisors' accuracy, and reliably "bet" on advisors' predictions more than their performance warranted. Participants' decisions tightly tracked a computational model that makes 2 assumptions: (a) people hold optimistic initial expectations about advisors, and (b) people preferentially incorporate information that adheres to their expectations when learning about advisors. Consistent with model predictions, explicitly manipulating participants' initial expectations altered their optimism bias and subsequent advice-taking. With well-calibrated initial expectations, participants no longer exhibited an optimism bias. We then explored crowdsourced ratings as a strategy to curb unrealistic optimism in advisors. Star ratings for each advisor were collected from an initial group of participants, which were then shown to a second group of participants. Instead of calibrating expectations, these ratings propagated and exaggerated the unrealistic optimism. Our results provide a computational account of the cognitive processes underlying inflated perceptions of expertise, and explore the boundary conditions under which they occur. We discuss the adaptive value of this

# Task design

- Participant's Choice – Advisor's Prediction – Outcome

- 26 subjects
- 3 advisors
- 108 trials
  (36 for each advisor)


- How people learn to trust or not to trust the advisor?

# Original models (MAP method)

- Confirmation Bias model ⚠️
- Win-Stay-Lose-Switch model
- Bayesian updating model ⚠️
- Reinforcement model

- Each trial certain value is updated in Bayesian fashion.
- Each trial it calculates posterior distribution by grid approximation

Which means we need to put grid approximation in Stan code!

# Introduction

## Original models (MAP method)

- Confirmation Bias model
- Win-Stay-Lose-Switch model
- Bayesian updating model
- Reinforcement model

# WSLS model

- Propensity to stay after win and shift after lose.

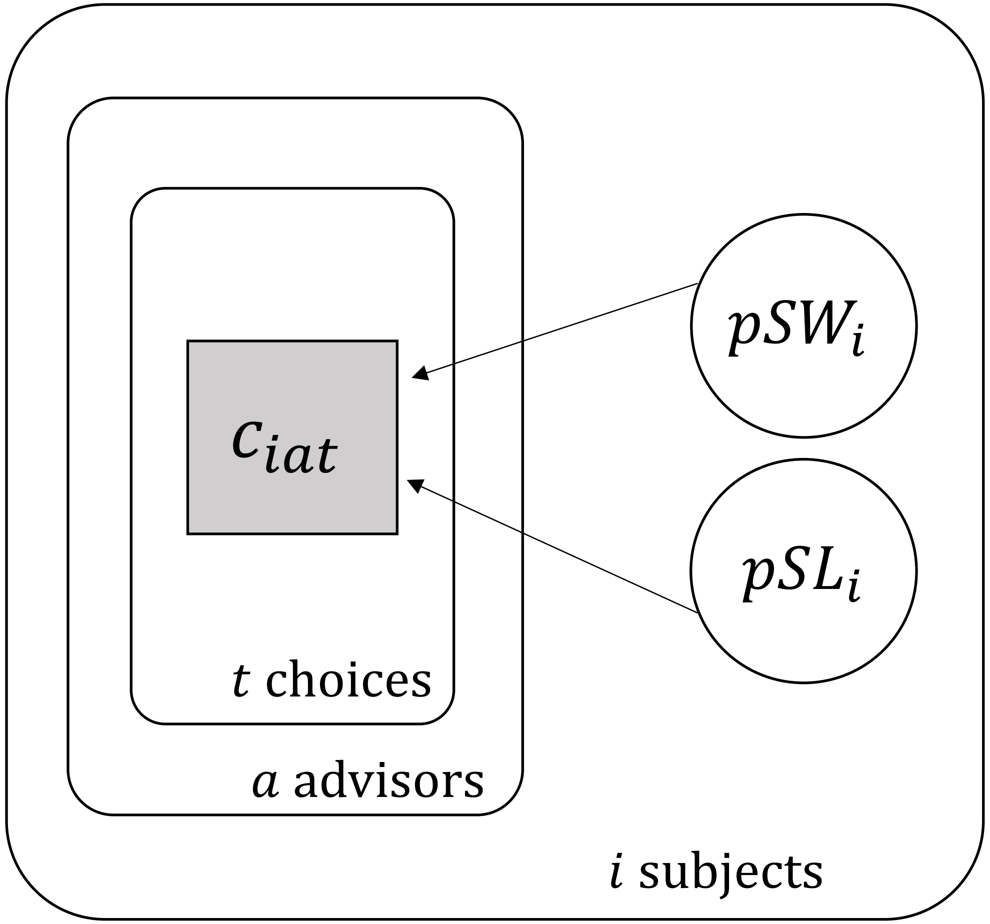
- Two parameters, each represents strength of these propensities.

$$P\left(stay|win\right)$$

$$P\left(shift|loss\right)$$

# Methods

## WSLS model



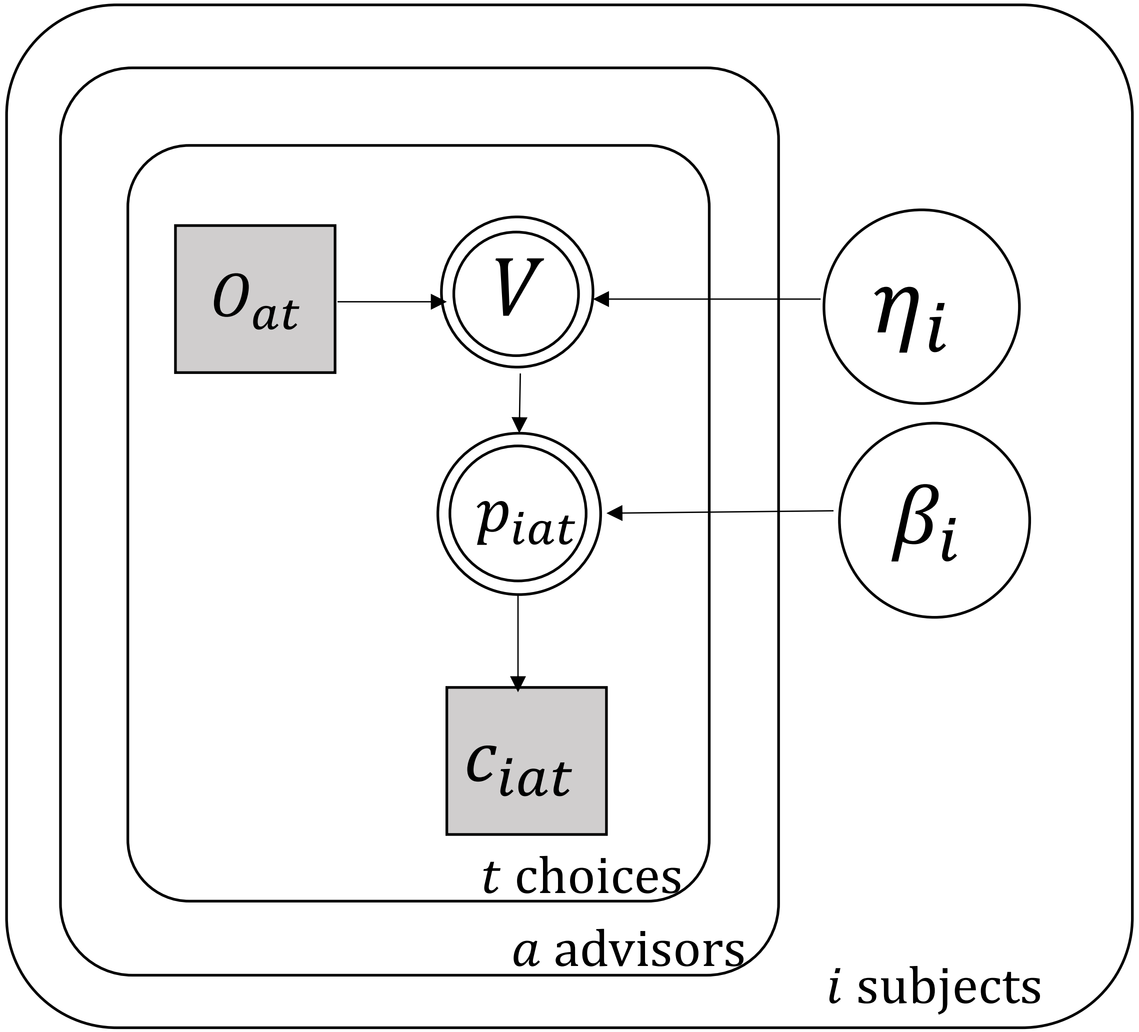
# Reinforcement Learning model

- Update the value of each advisor with prediction error
- Higher the value, higher the chance to trust the advisor

$$\delta_t = R_t - E(V_t)$$

$$E(V_{t+1}) = E(V_t) + \eta\delta_t$$

# Reinforcement Learning model



# Methods

## Proposal for better models

**A Comparison Model of Reinforcement-Learning and Win-Stay-Lose-Shift Decision-Making Processes: A Tribute to W.K. Estes**

**Darrell A. Worthy**[1] and **W. Todd Maddox**[2]
[1]*Texas A&M University*
[2]*The University of Texas at Austin*

**Abstract**

W.K. Estes often championed an approach to model development whereby an existing model was augmented by the addition of one or more free parameters, and a comparison between the simple and more complex, augmented model determined whether the additions were justified. Following this same approach we utilized Estes' (1950) own augmented learning equations to improve the fit and plausibility of a win-stay-lose-shift (WSLS) model that we have used in much of our recent work. Estes also championed models that assumed a comparison between multiple concurrent cognitive processes. In line with this, we develop a WSLS-Reinforcement Learning (RL) model that assumes that the output of a WSLS process that provides a probability of staying or switching to a different option based on the last two decision outcomes is compared with the output of an RL process that determines a probability of selecting each option based on a comparison of the expected value of each option. Fits to data from three different decision-making experiments suggest that the augmentations to the WSLS and RL models lead to a better account of decision-making behavior. Our results also support the assertion that human participants weigh the output of WSLS and RL processes during decision-making.

- WSLS learning model
- Transition RL model
  - version 1 / 2
- WSLS-RL model
  - version 1 / 2

# WSLS learning model

- Updating the strength of propensity, trial by trial
  - Successive WIN increases pSW
  - Successive LOSS increases pSL
- Trend of the outcome, not the value itself!

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

$$P(shift|loss)_{t+1} = \left(1 - \theta_{P(lose|shift)}\right) P(shift|loss)_t \quad (8)$$

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

$$P(stay|win)_{t+1} = \left(1 - \theta_{P(stay|win)}\right) P(stay|win)_t \quad (10)$$

# Transition RL model

- Learn the value of transition (Stay or Shift)
- Version 1 :  EV is updated separately on Win / Loss trial
- Version 2 :  EV is updated without distinction of Win/ Loss

$$EV_{i,t+1}=EV_{i,t}+\alpha\cdot\left[r\left(t\right)-EV_{i,t}\right]$$

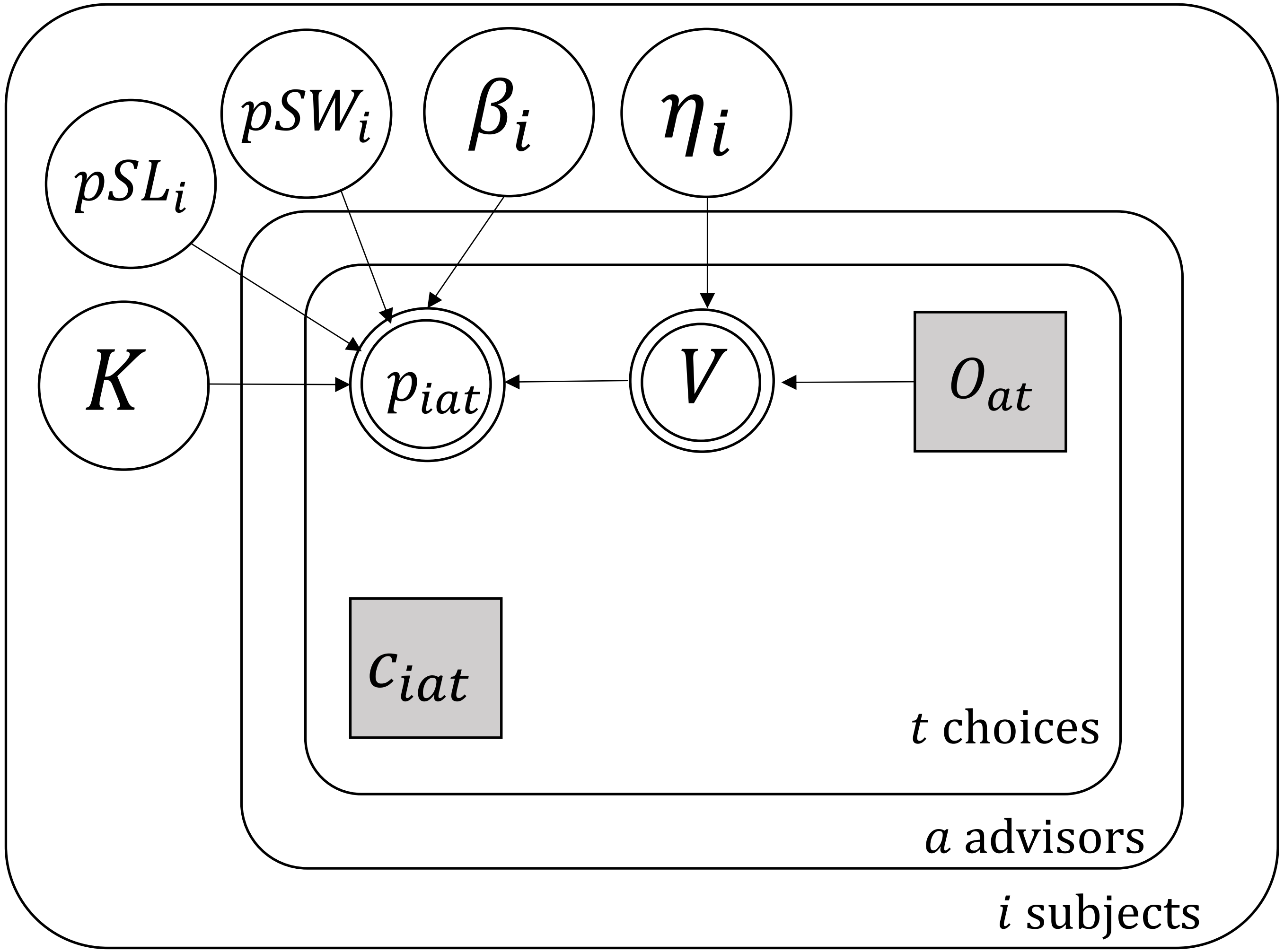
$$P\left(a_t\right)=\frac{e^{\left[\gamma\cdot EV_{a,t}\right]}}{\sum_{j=1}^{n}e^{\left[\gamma\cdot EV_{j,t}\right]}}$$

# WSLS-RL hybrid model

- WSLS + Reinforcement learning
- Version 1 : WSLS learning + Reinforcement learning
- Version 2 : WSLS (classic) + Reinforcement learning
- People consider both the trend of the outcome and the value of the advisor

$$P(a_t) = P(a_t)_{WSLS_{Learning}} \cdot W_{WSLS_{Learning}} + P(a_t)_{RL} \cdot \left(1 - W_{WSLS_{Learning}}\right)$$

# WSLS-RL hybrid model (ver 2.)



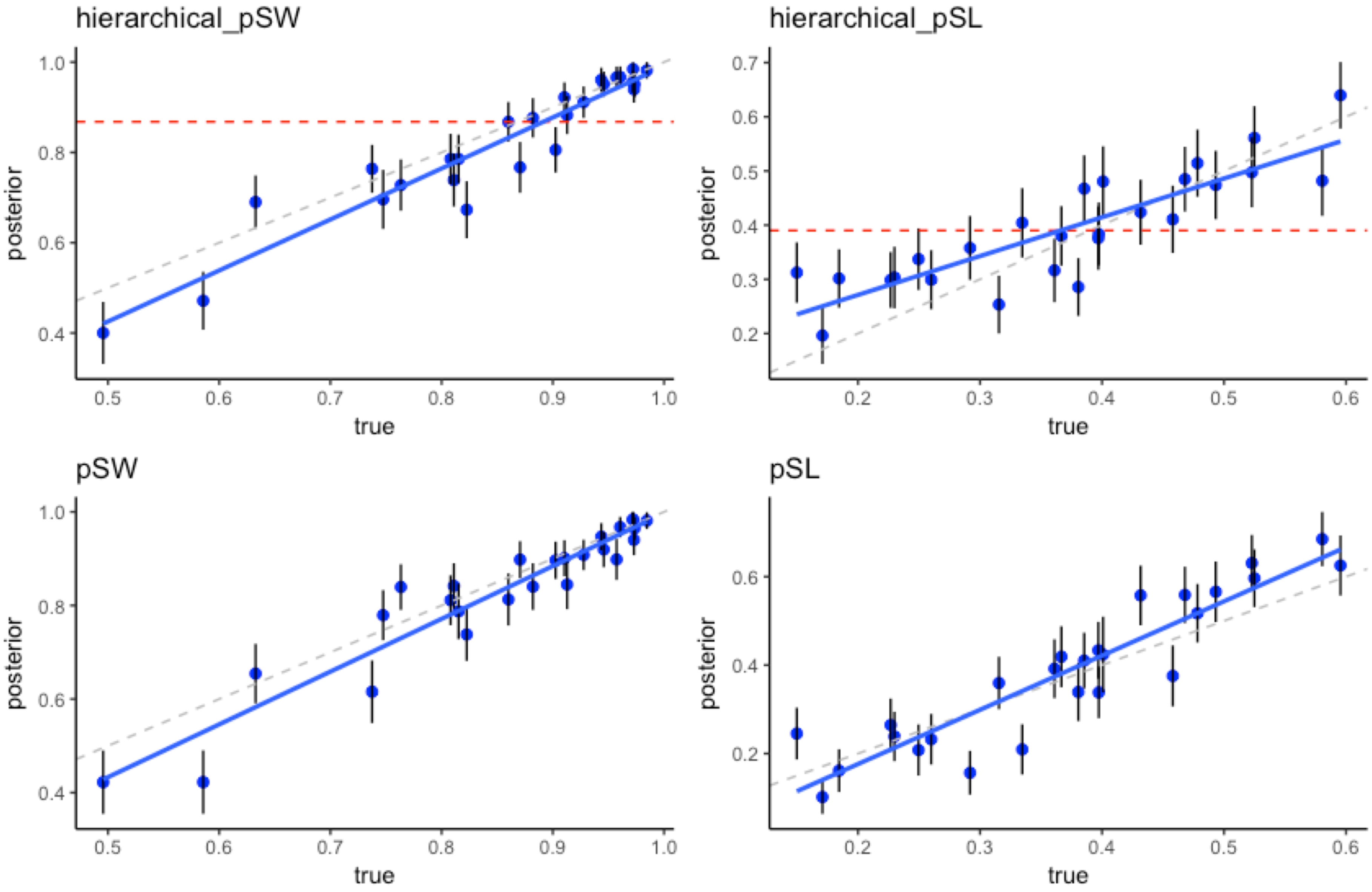
# Results

## Model Comparison

| | **LOOIC (full individual)** | **LOOIC (hierarchy)** | |
|---|---|---|---|
| WSLS model | 2644.4 | 2638.8 | ● |
| Reinforcement Learning model | 2755.8 | 2697.8 | ● |
| WSLS learning model | 2812.4 | 2662.3 | ● |
| WSLS transition model v1. | 3305.6 | 3178.4 | |
| WSLS transition model v2. | 3309.1 | 3242.8 | |
| WSLS-RL (hybrid) model | 4260.5 | 3398.0 | |
| WSLS-RL(Classic) model | 2543.5 | 2450.6 | ● |

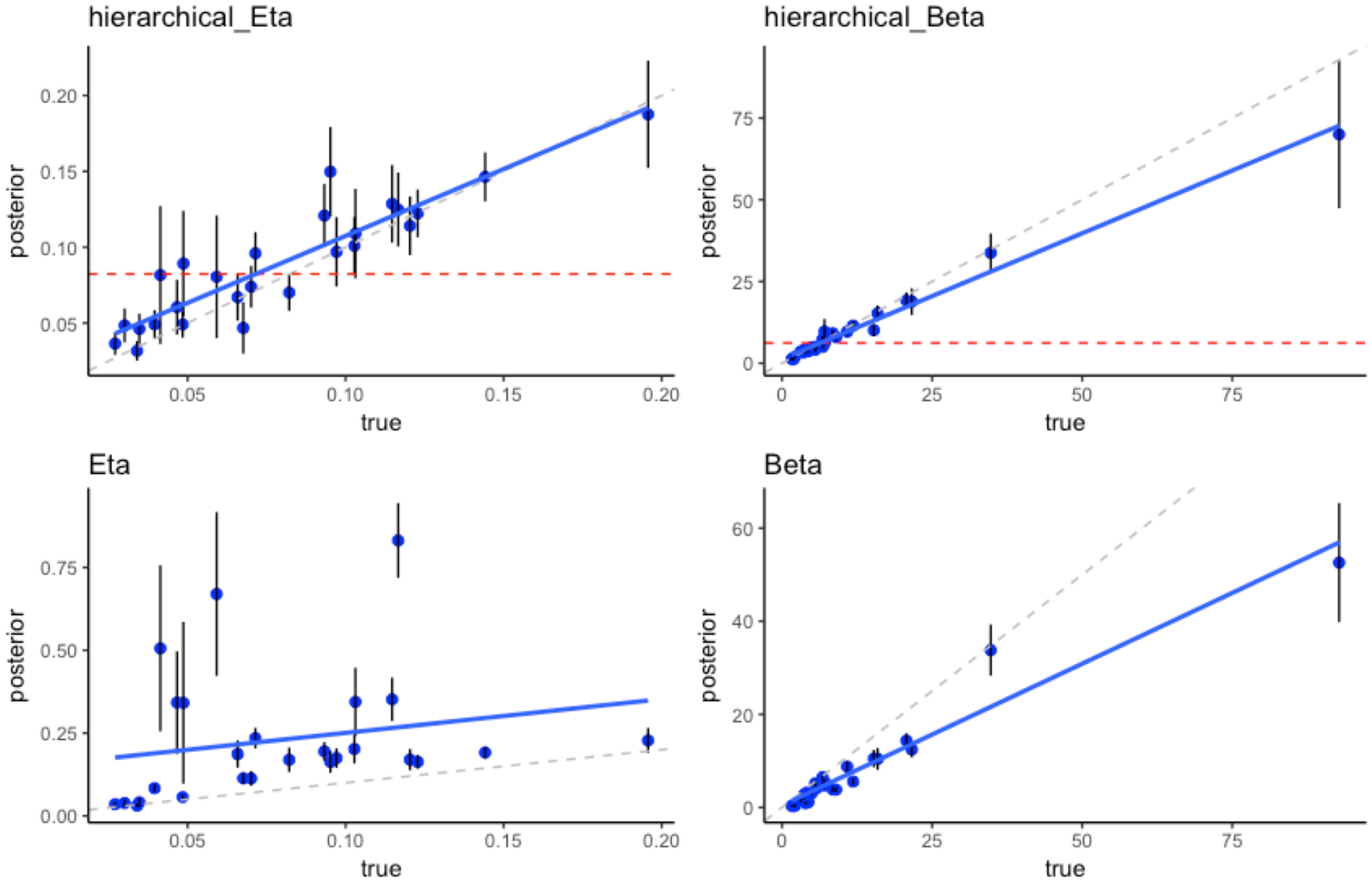
# Results

## WSLS_classic model : parameter recovery



# Results

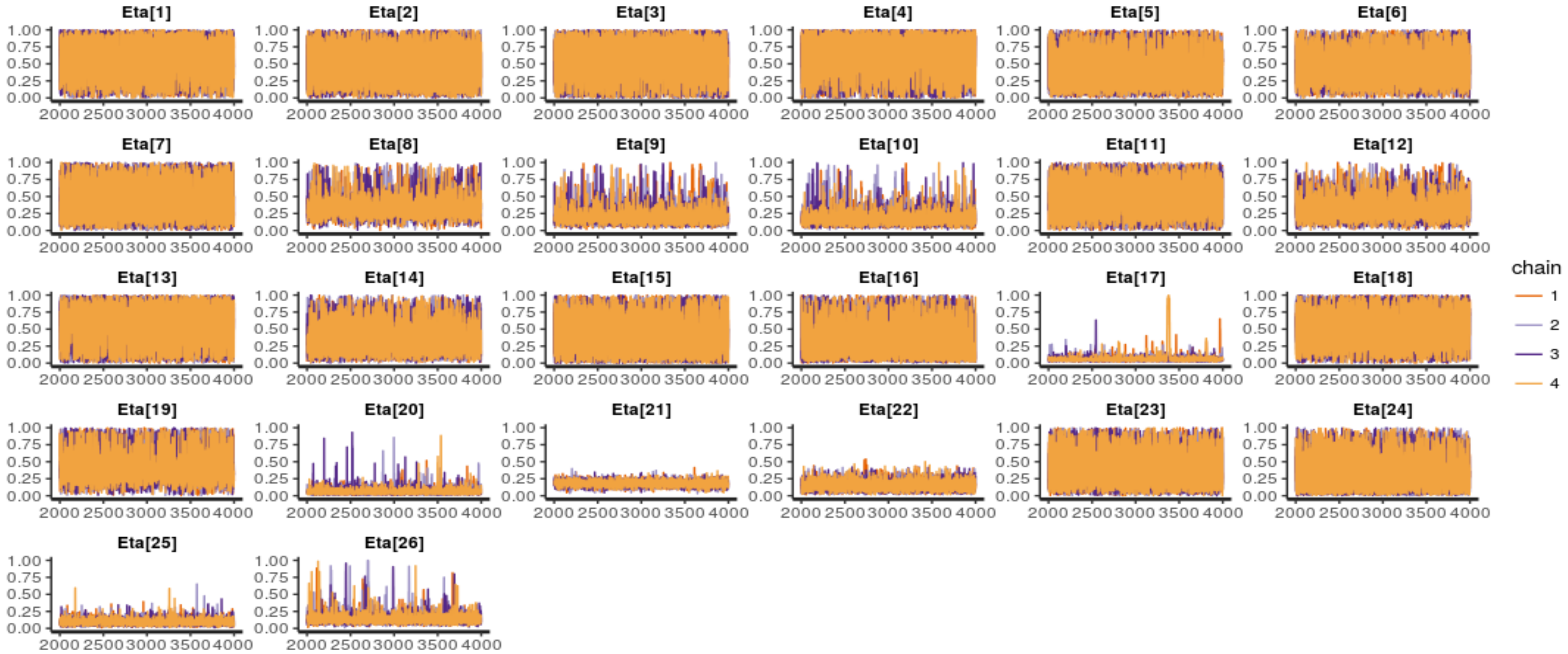
## RL_classic model : parameter recovery



# Full Individual model

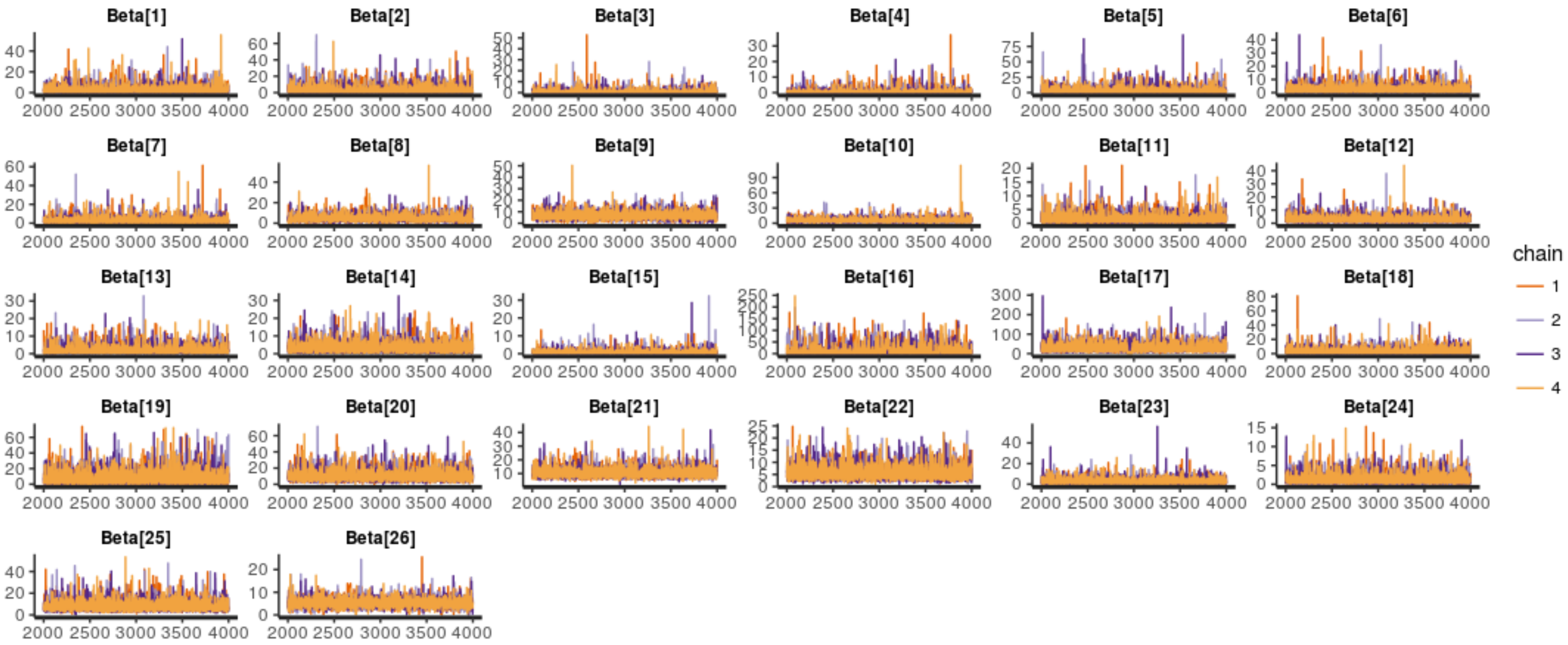
# Results

# Classic hybrid WSLS-RL model



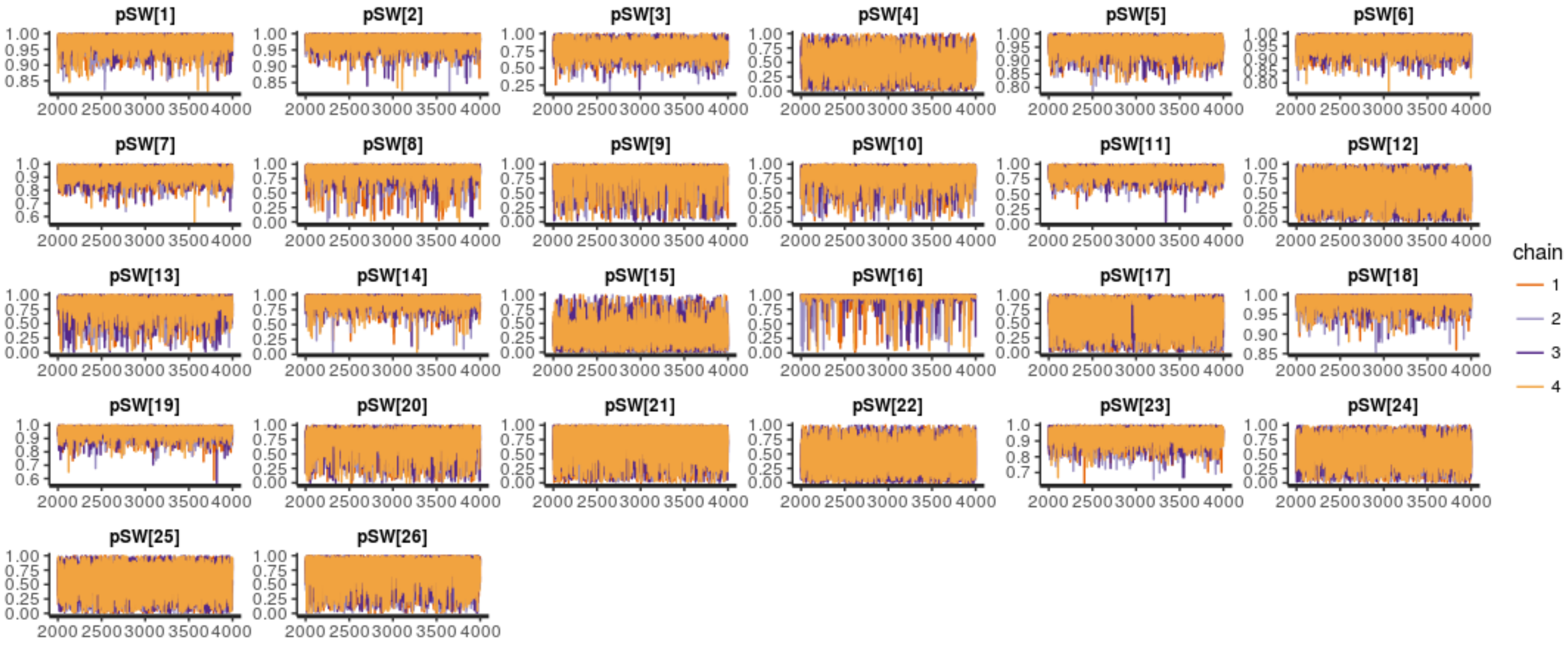
# Results

## Classic hybrid WSLS-RL model



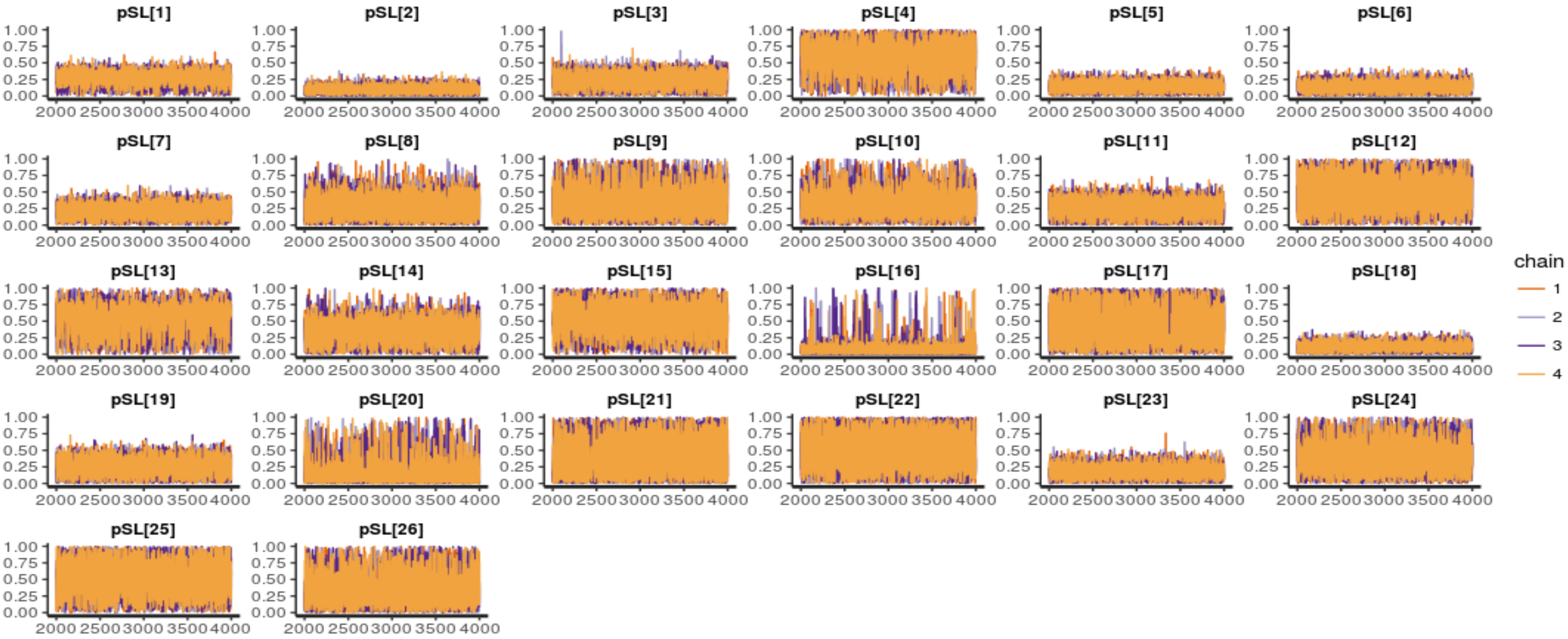
# Results

## Classic hybrid WSLS-RL model



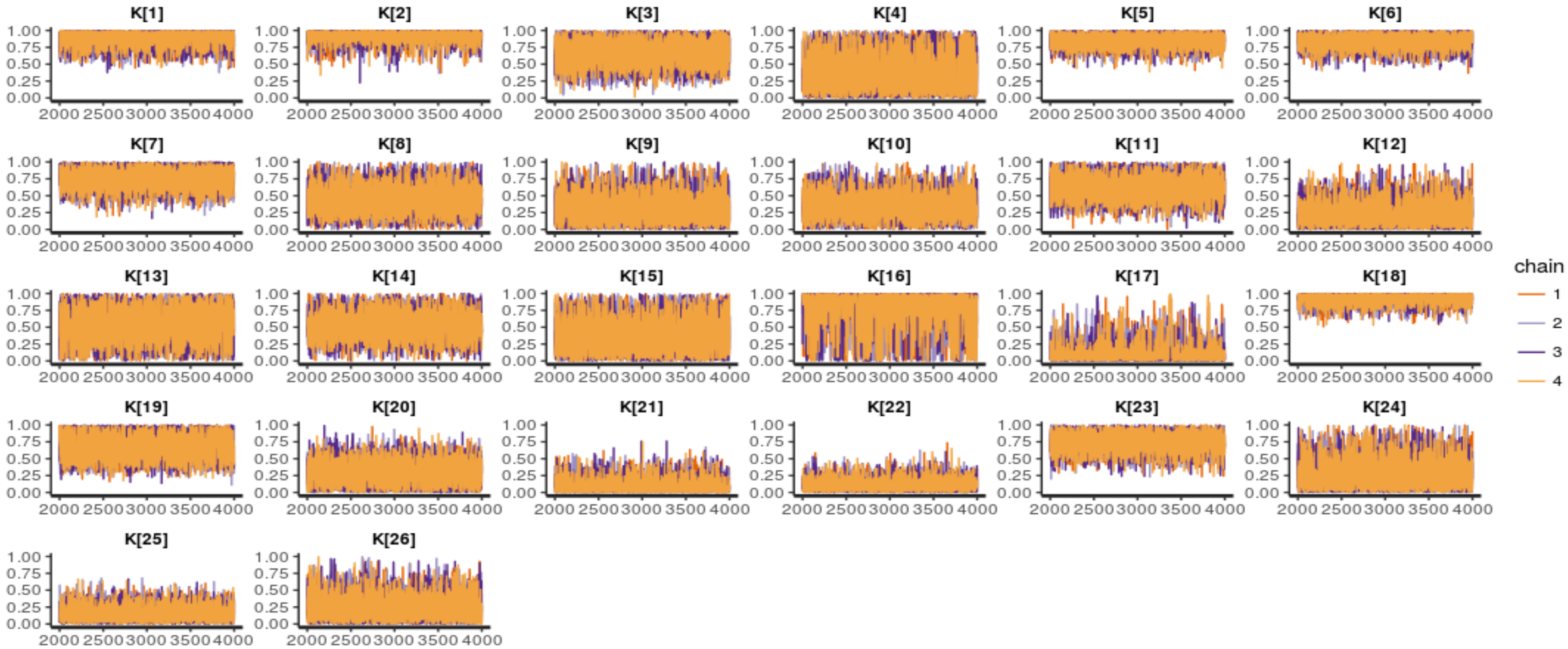
# Results

## Classic hybrid WSLS-RL model



# Results

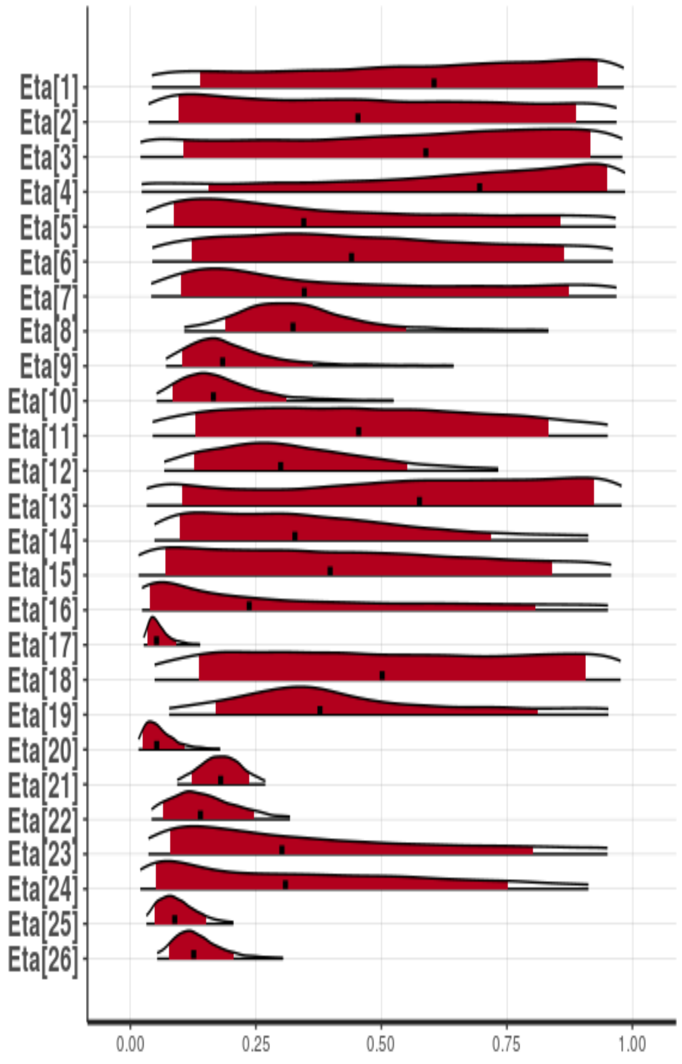
## Classic hybrid WSLS-RL model

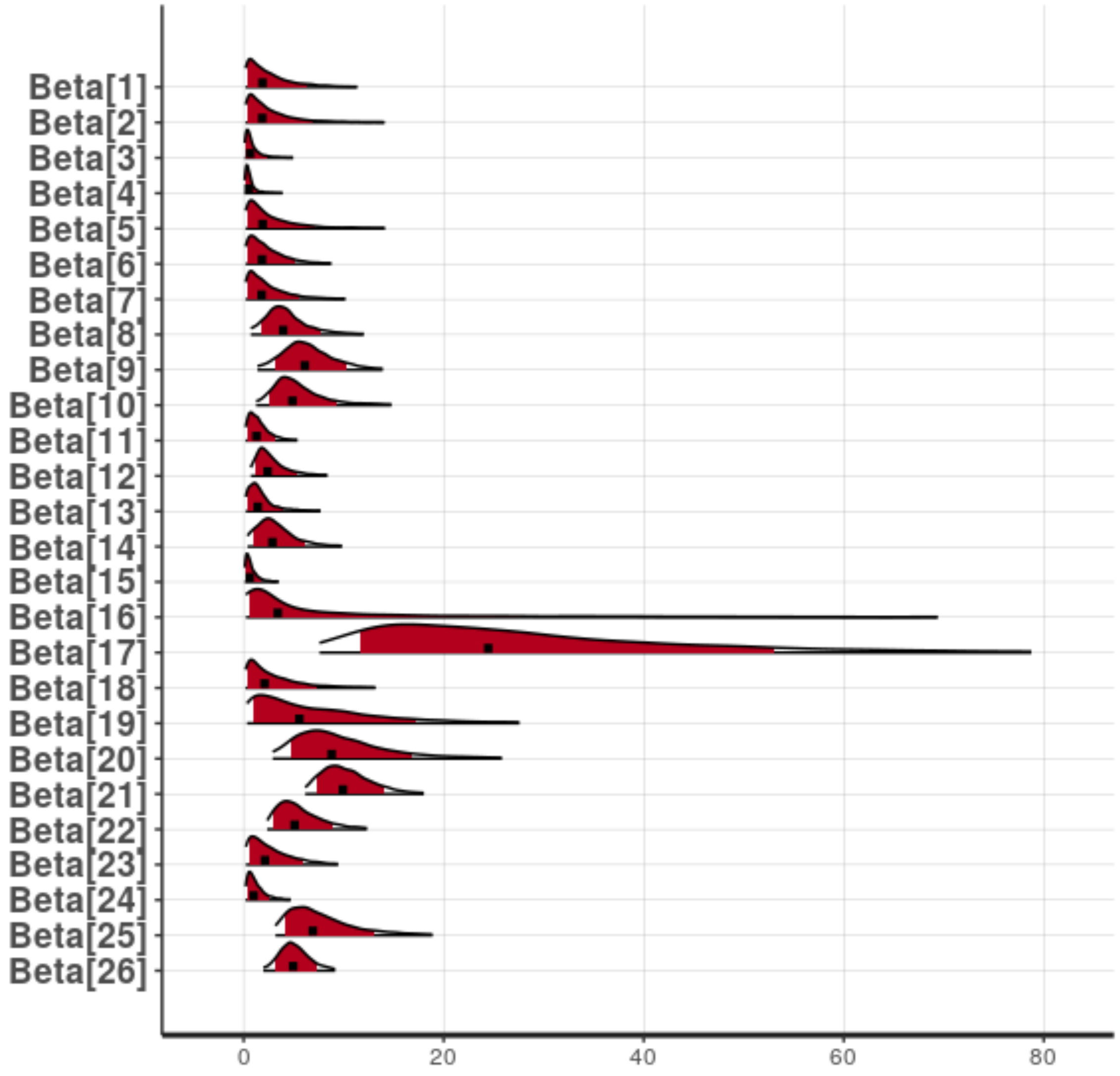


# Results

Eta

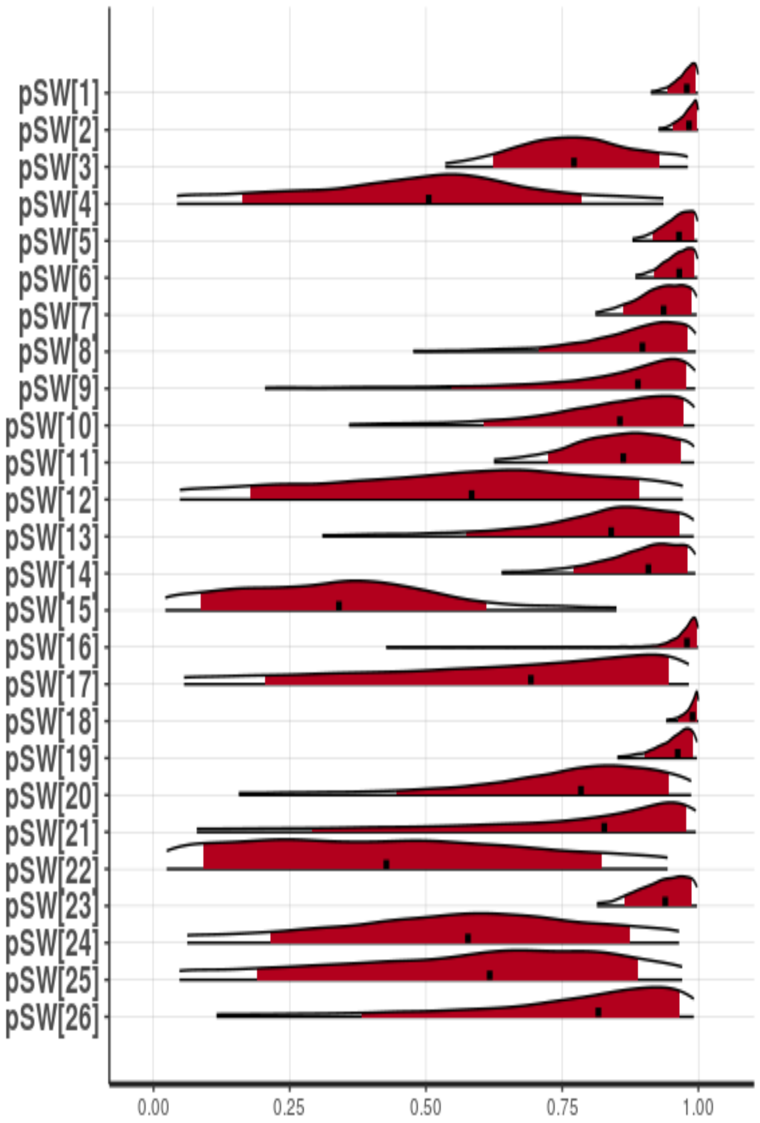
Beta



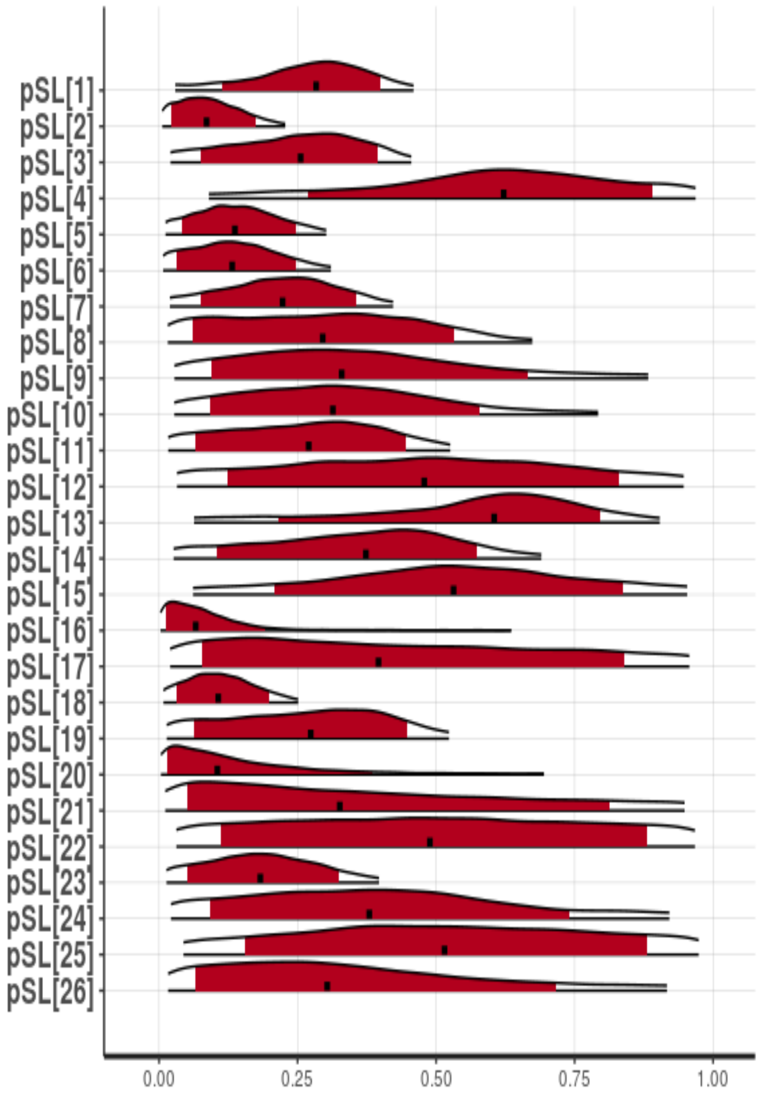



# Results

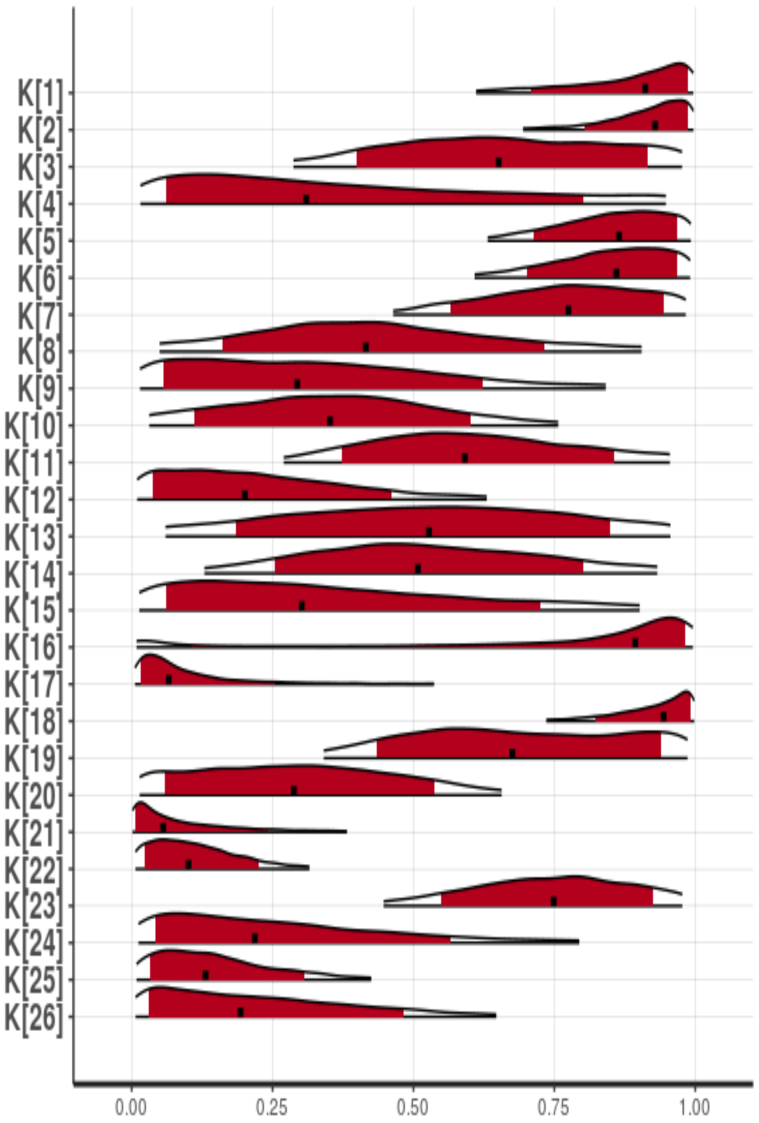
pSW




pSL



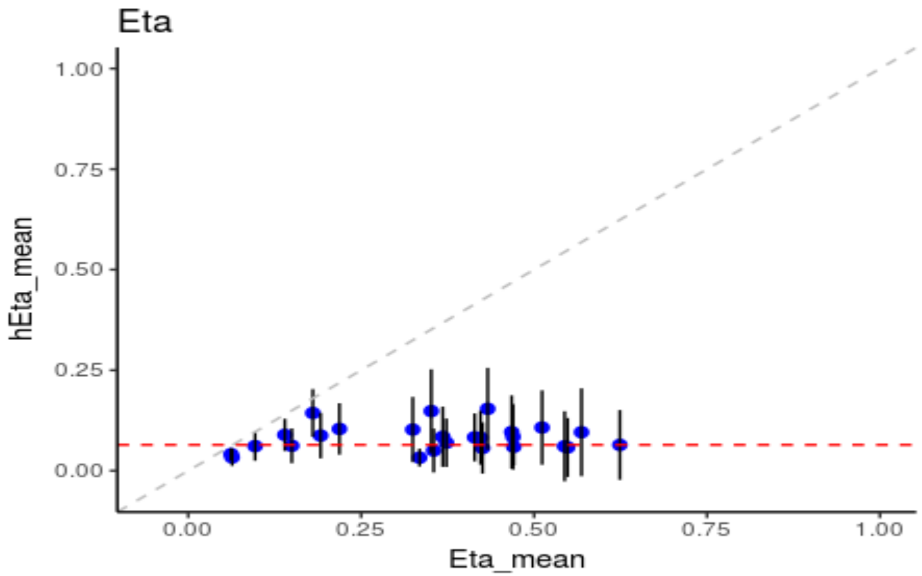

K

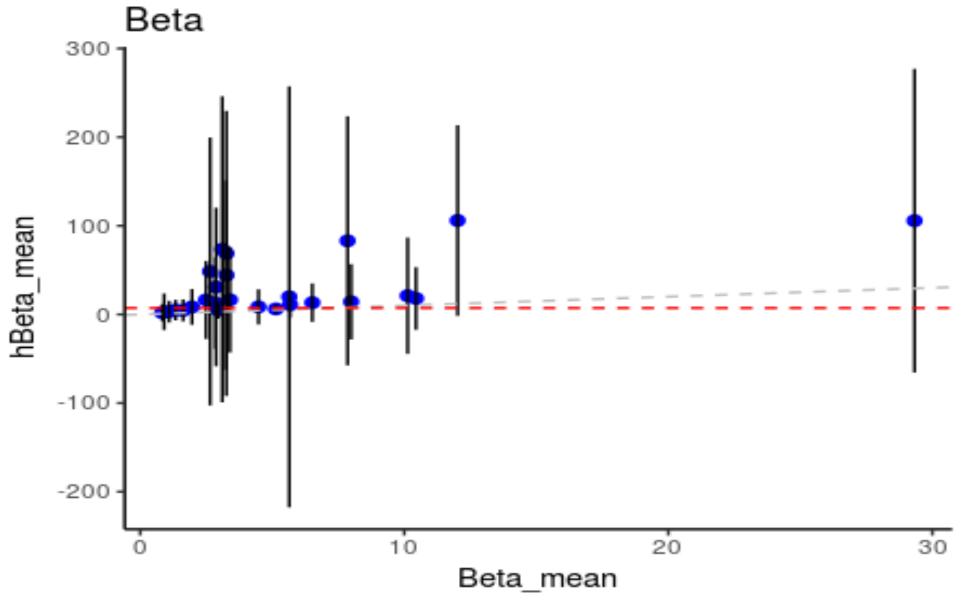


# Hierarchical model

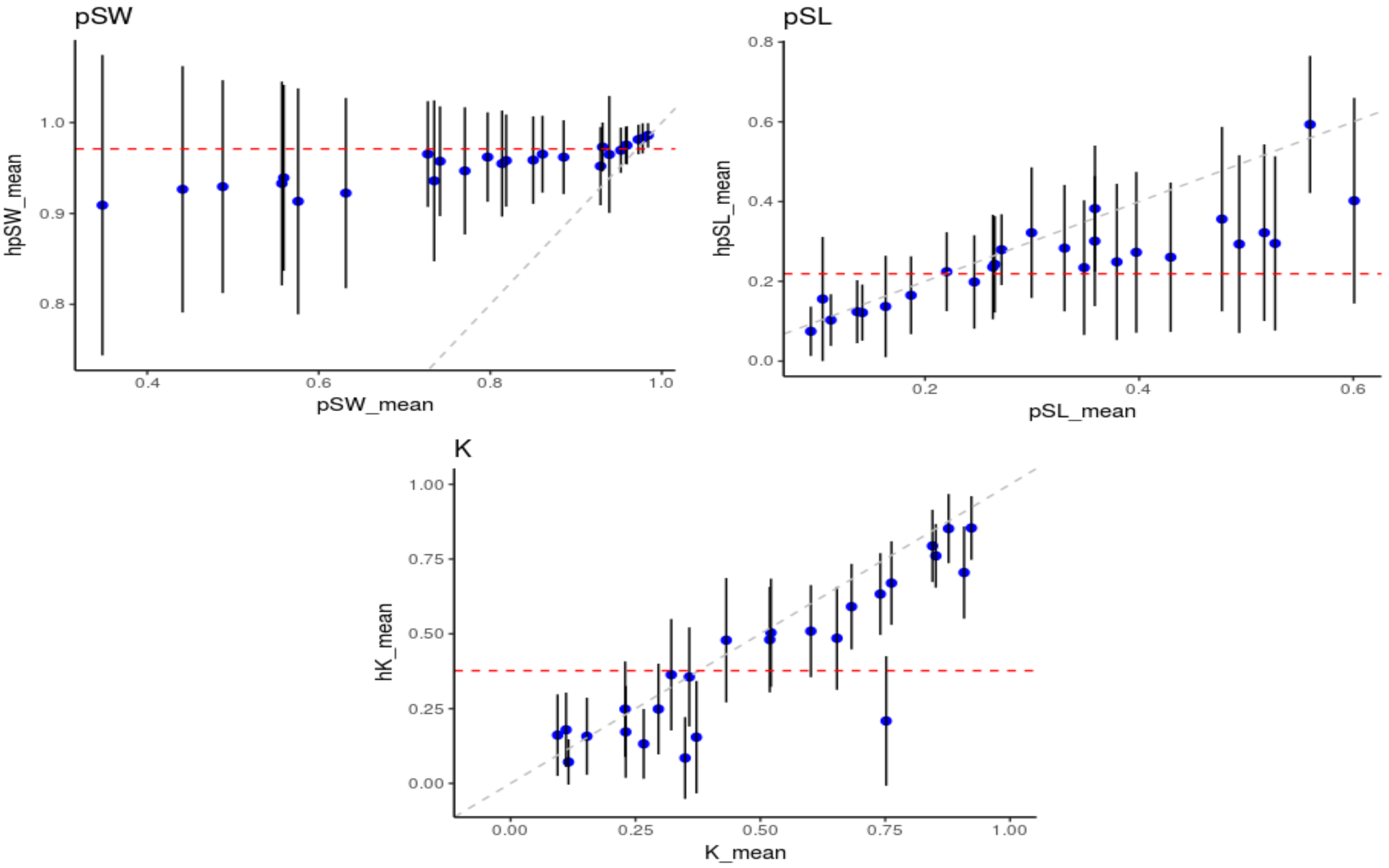
# Results

## WSLS_classic + RL model
## : Hierarchy – Nonhierarchy plot



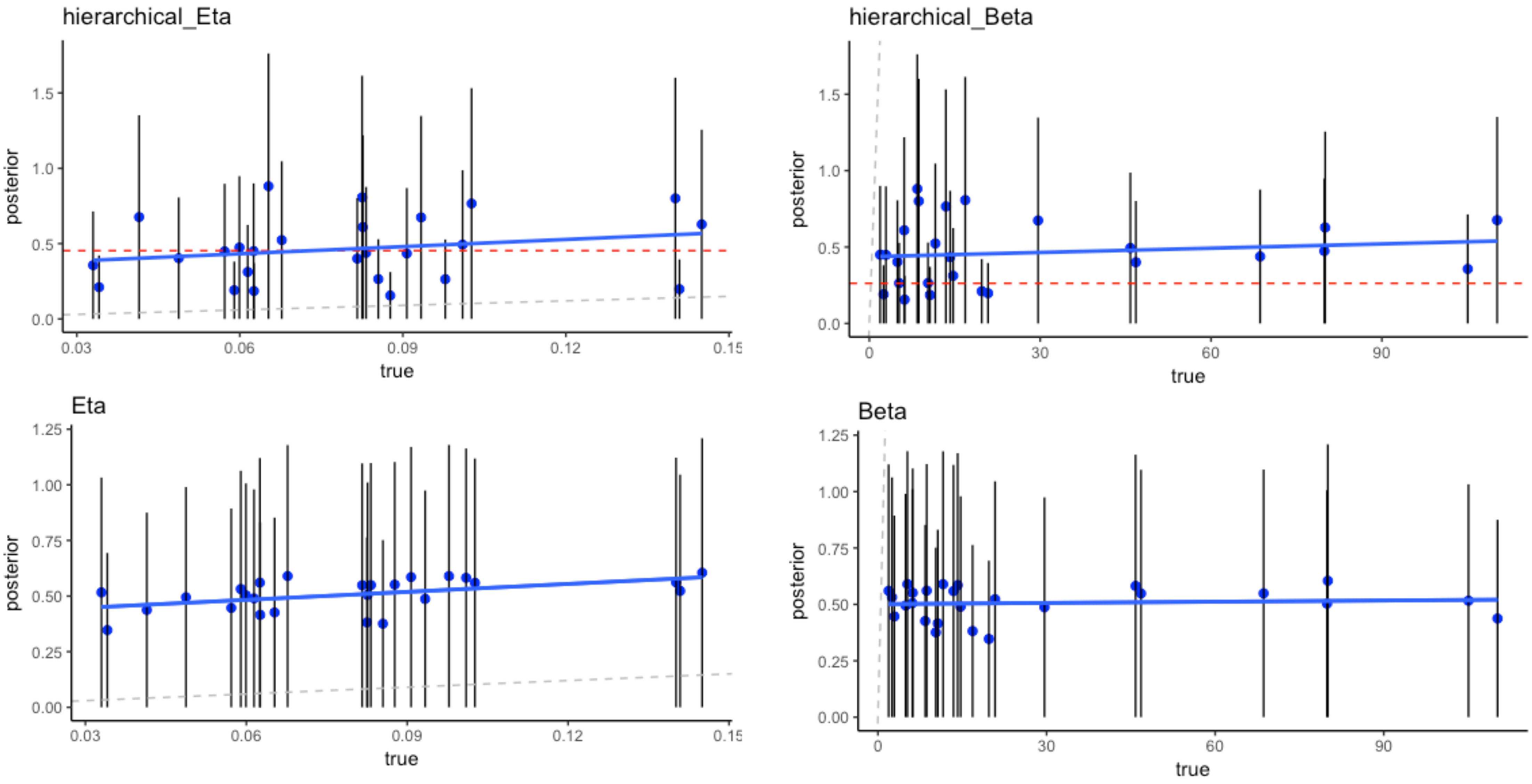



# Results


pSW
hpSW_mean
pSW_mean
1.0
0.9
0.8
0.4
0.6
0.8
1.0
pSL
hpSL_mean
pSL_mean
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
K
hK_mean
K_mean
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00

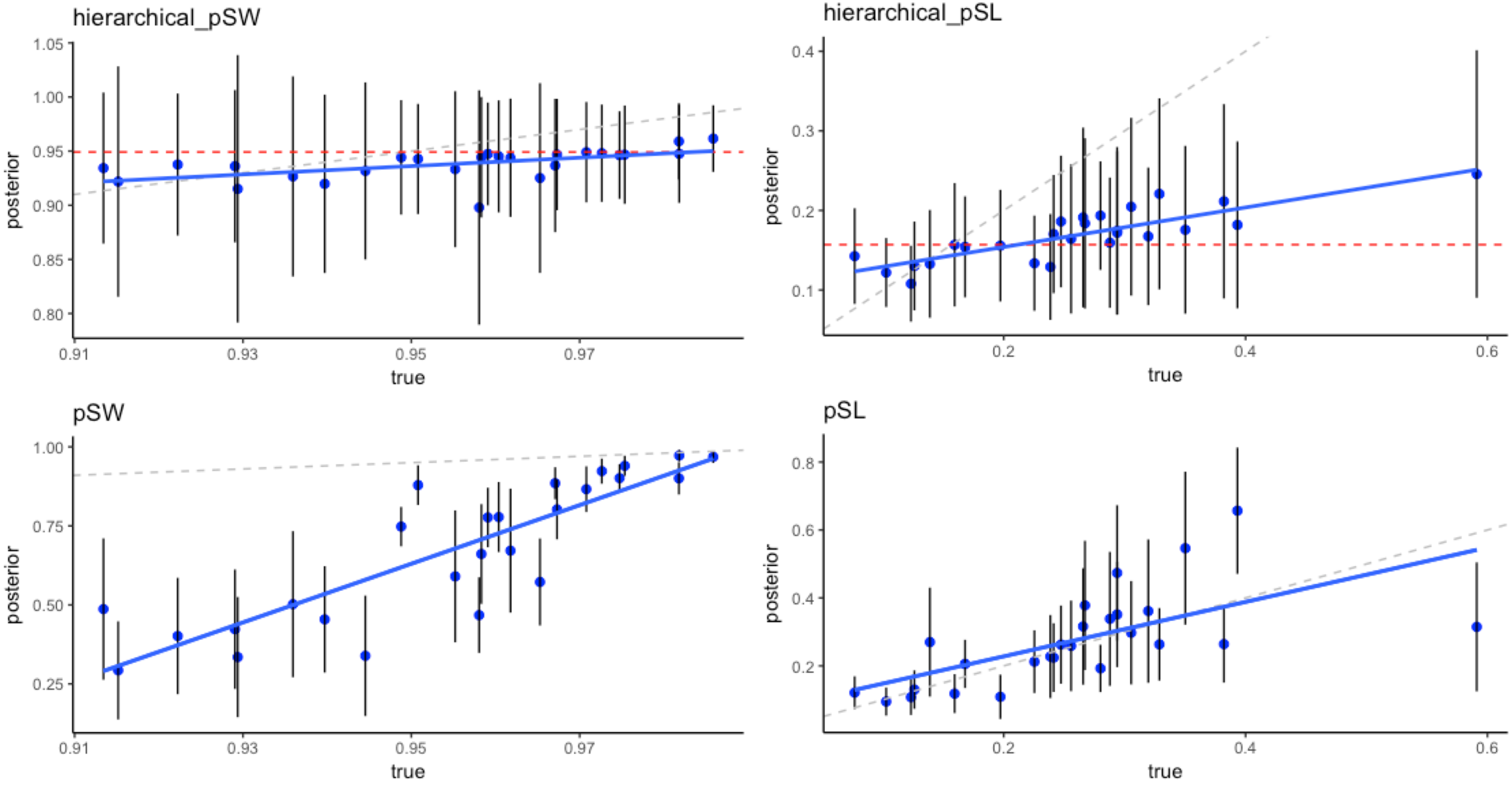
# Results

## WSLS_classic + RL model : parameter recovery



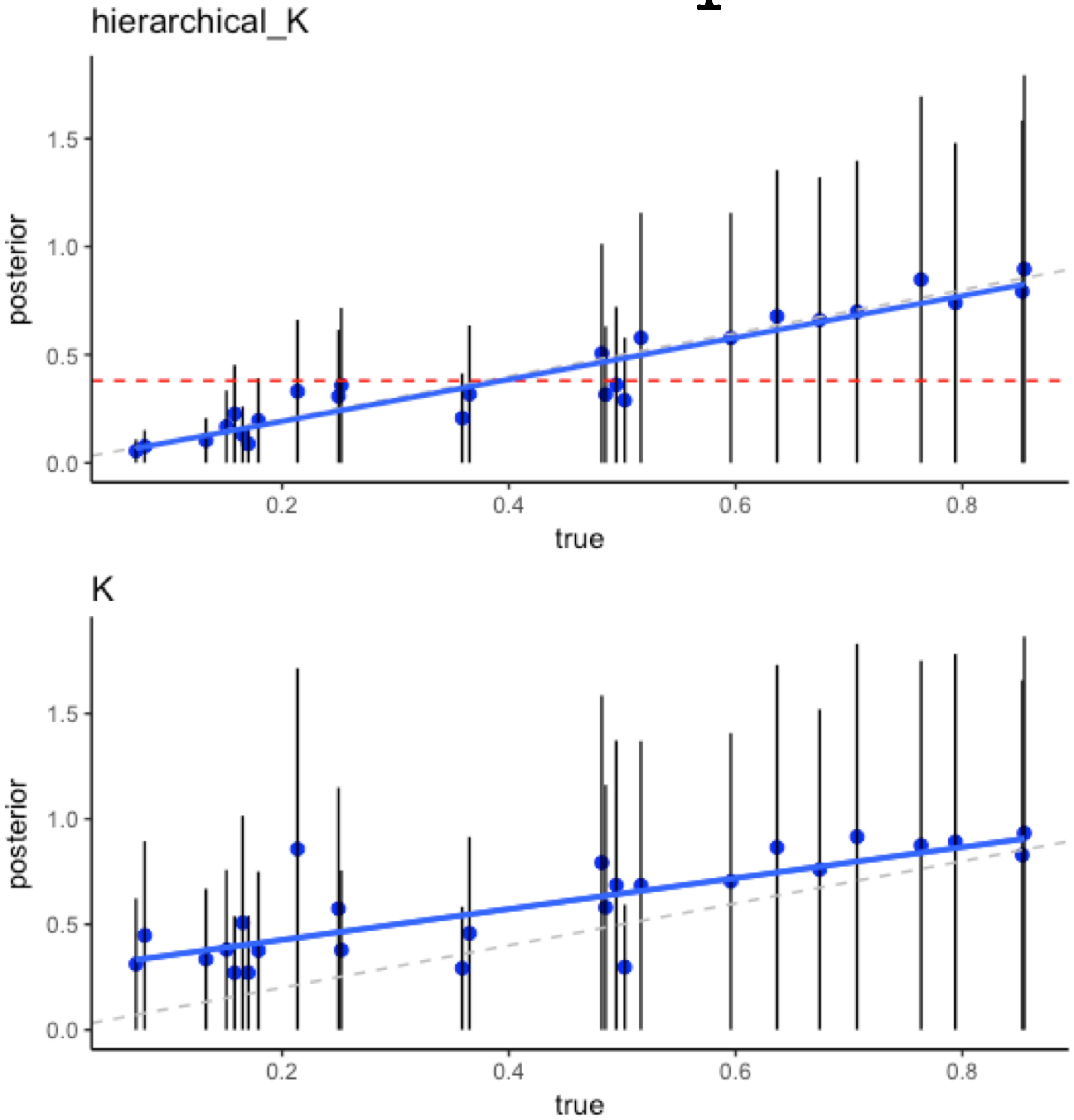
# Results

## WSLS_classic + RL model : parameter recovery



# Results

## WSLS_classic + RL model : parameter recovery



# Results

Is this model okay?

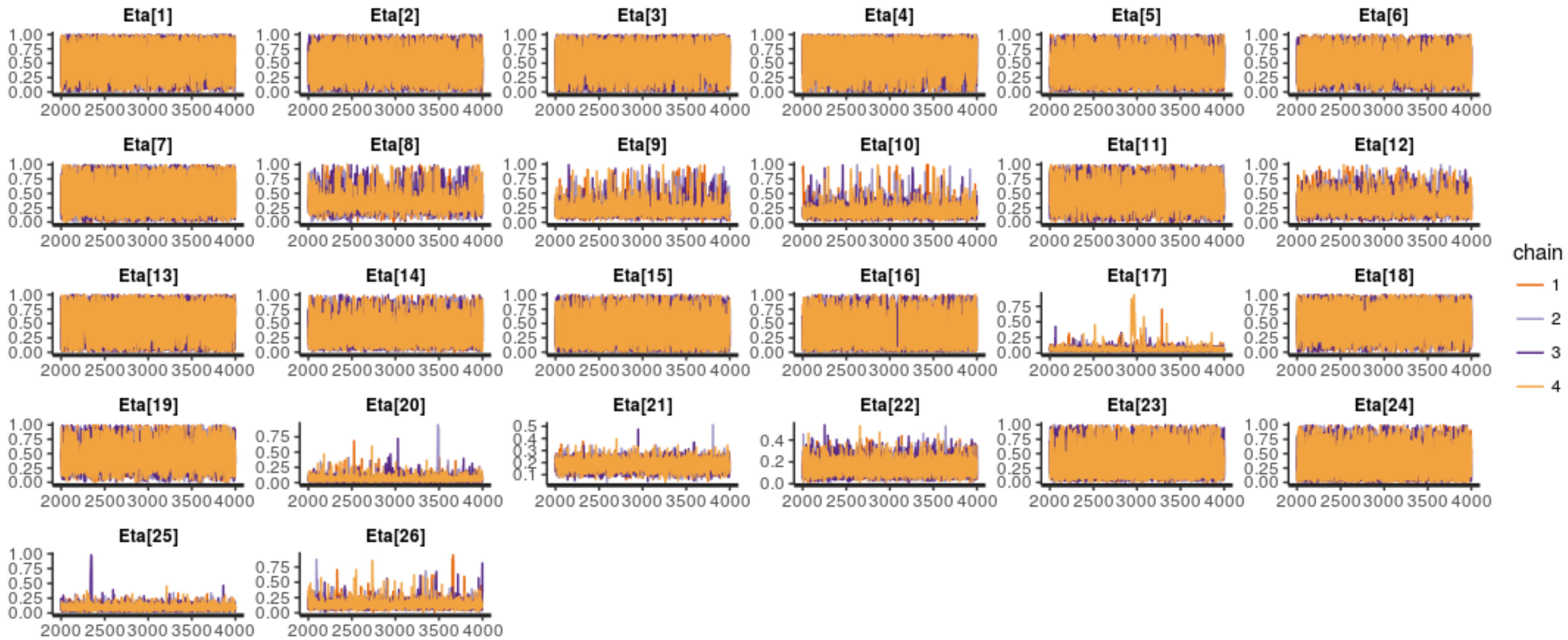
# Results

## K (full individual)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.877 | 0.908 | 0.653 | 0.372 | 0.851 | 0.845 | 0.763 | 0.432 | 0.322 | 0.358 |
| 0.601 | 0.231 | 0.521 | 0.519 | 0.349 | 0.752 | 0.111 | 0.923 | 0.683 | 0.296 |
| 0.094 | 0.116 | 0.740 | 0.266 | 0.153 | 0.229 | | | | |

WSLS : 1,2,3,5,6,7,10,11,13,14,16,18,19,23
RL : 4,8,9,12,15,17,20,21,22,24,25,26

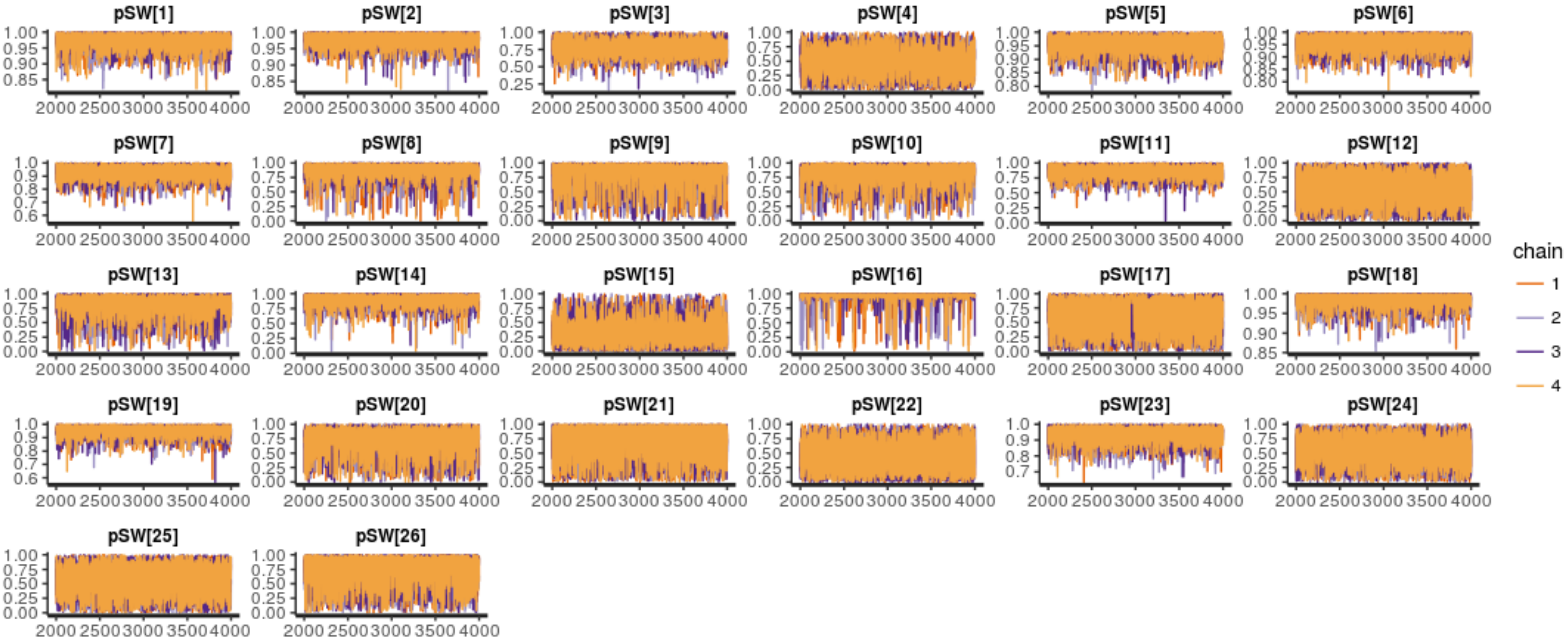
# Results




WSLS: 1,2,3,5,6,7,10,11,13,14,16,18,19,23

RL :4,8,9,12,15,17,20,21,22,24,25,26

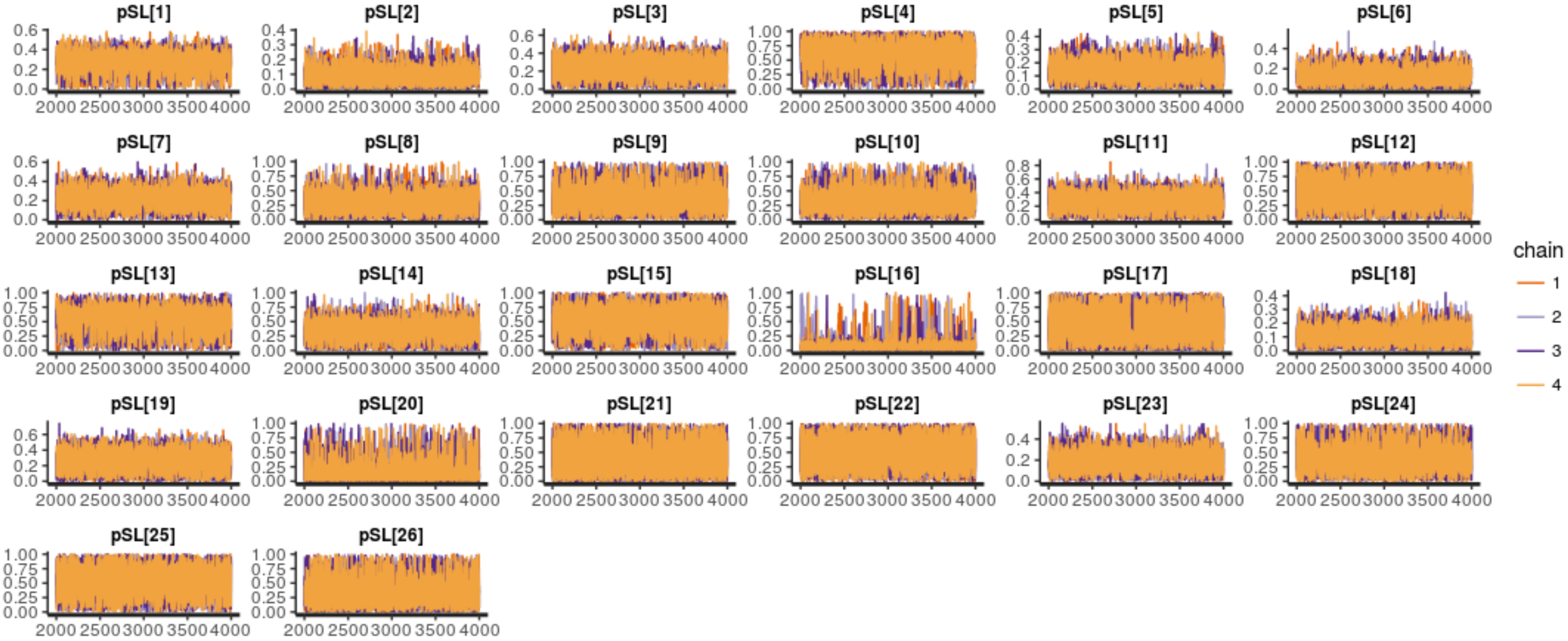
# Results




WSLS: 1,2,3,5,6,7,10,11,13,14,16,18,19,23

RL :4,8,9,12,15,17,20,21,22,24,25,26

# Results




WSLS: 1,2,3,5,6,7,10,11,13,14,16,18,19,23

RL : 4,8,9,12,15,17,20,21,22,24,25,26

# Implication

WSLS_learning

+

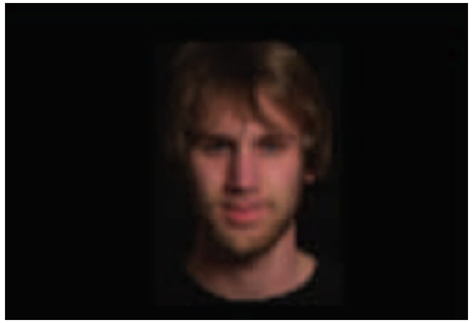
RL model

VS

WSLS

+

RL model



Thank you ☺

# Contribution

| | Yoonseo Zoh | Hyeonjin Kim |
|---|---|---|
| Codes | Transition RL, hybrid v2 | RL , hybrid v1 |
| | Parameter recovery, posterior predictive | traceplot, posterior distribution check |
| | Everything else together | |
| PPT | First section | Second section |
| Presentation | First section | Second section |

# References

Leong, Y. C., & Zaki, J. (2018). Unrealistic optimism in advice taking: A computational account. *Journal of Experimental Psychology: General*, *147*(2), 170.

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.

# Supplementary

## Posterior Predictive Check