Generation and Prevention
in continuous time causal induction

Tia Gong
MRes student @Bramleylab
2 March 2020

Background

- Paradigm Demo
    - https://www.bramleylab.ppls.ed.ac.uk/experiments/tia/diamond_demo/
- Causal structure learning
- From micro-temporal dynamics

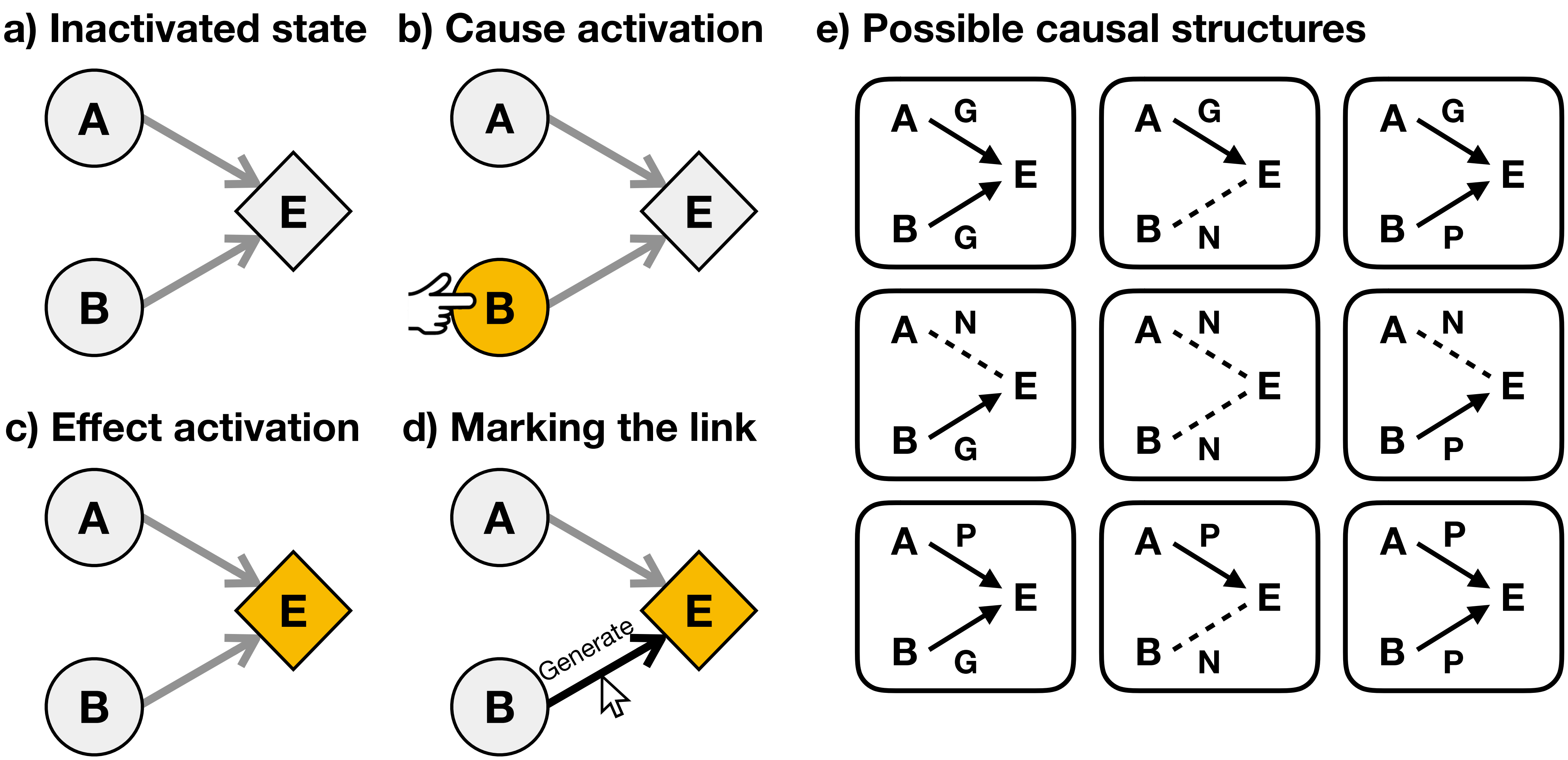
Background

- Including generative and preventative causal learning in the same causal system
- Building both computational-level and algorithmic-level models to predict human inference
- Examining previous findings under the contingency/contiguity issues

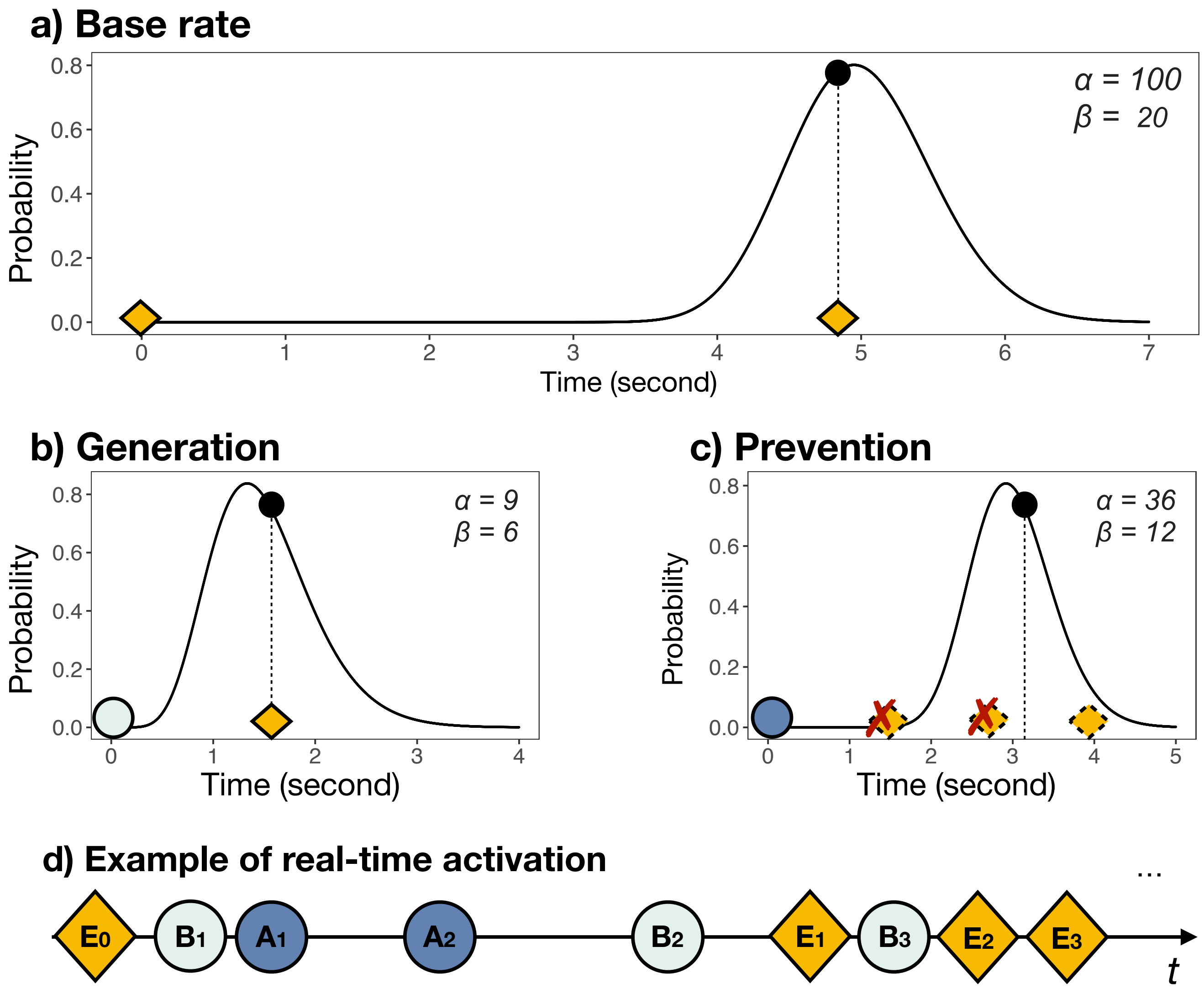
Background

- Including generative and preventative causal learning in the same causal system
- Building both computational-level and algorithmic-level models to predict human inference
- Examining previous findings under the contingency/contiguity issues

Experimental design



Experimental design

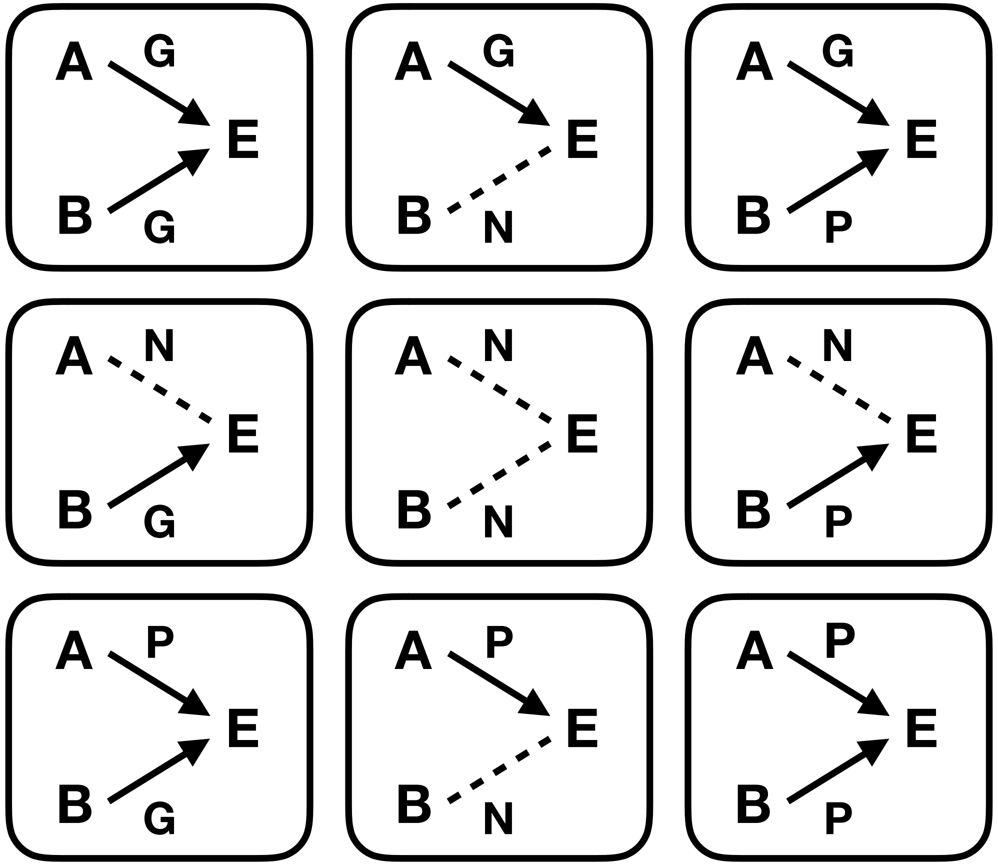


Background

- Including generative and preventative causal learning in the same causal system
- Building both computational-level and algorithmic-level models to predict human inference
- Examining previous findings under the contingency/contiguity issues

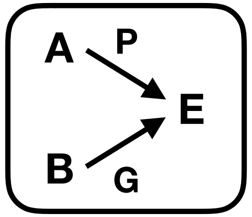
1.Actual Causal Attribution (Normative)

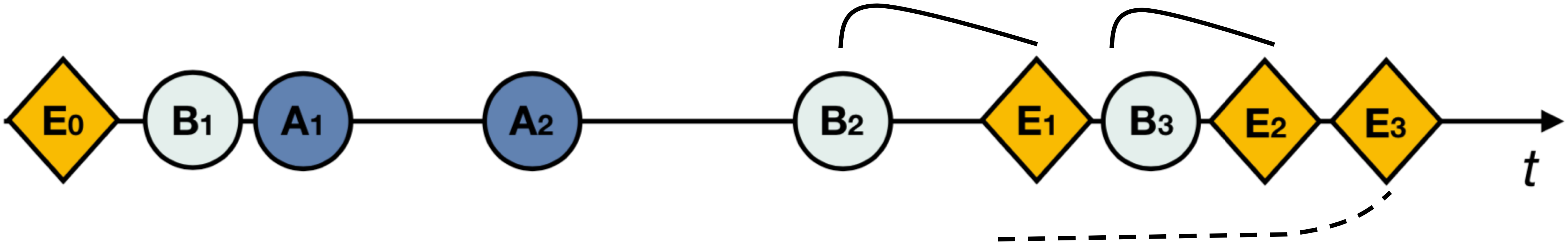
$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$

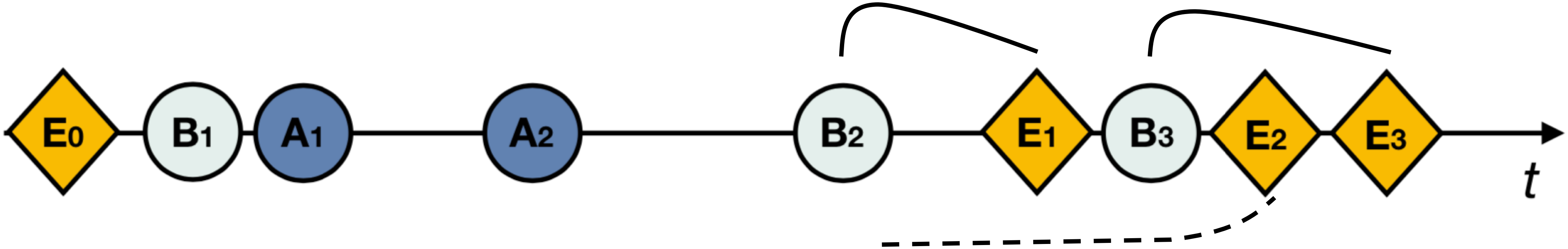


1.Actual Causal Attribution (Normative)

$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$

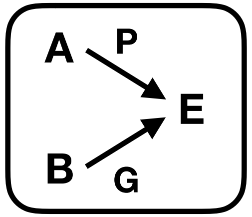




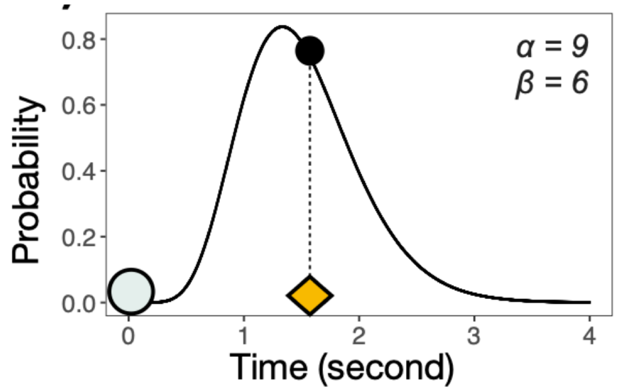


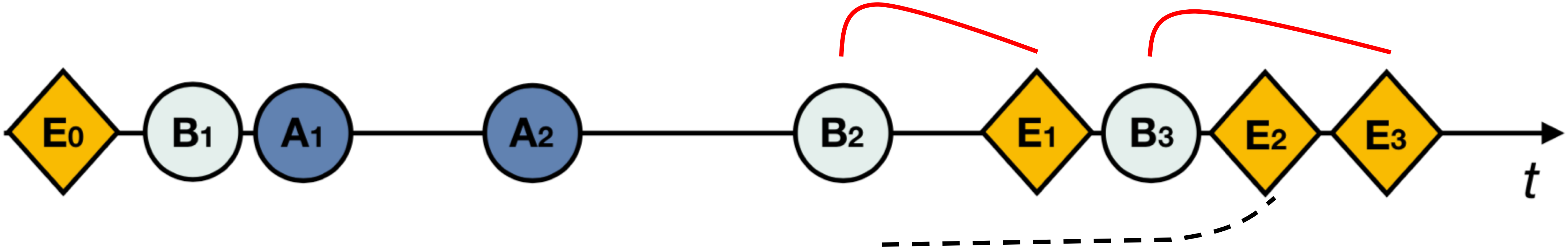
1.Actual Causal Attribution (Normative)

$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$



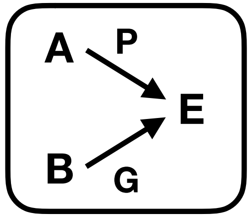
Generative path: the delay between the effect and its supposed actual cause.



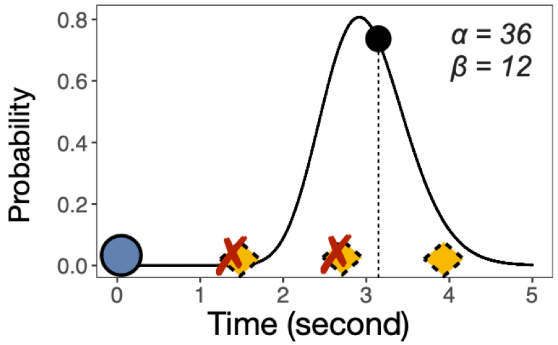


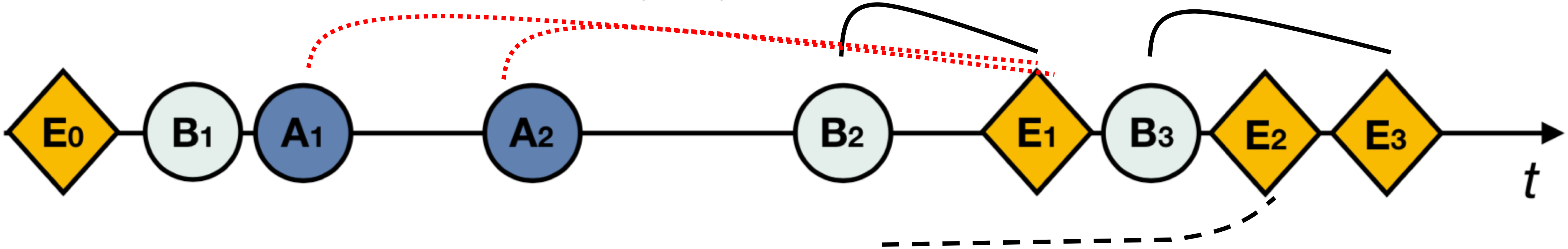
1.Actual Causal Attribution (Normative)

$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$



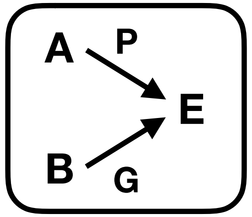
Preventative path: the delay between the supposed preventative control component's activation and its following subsequent target component activation.



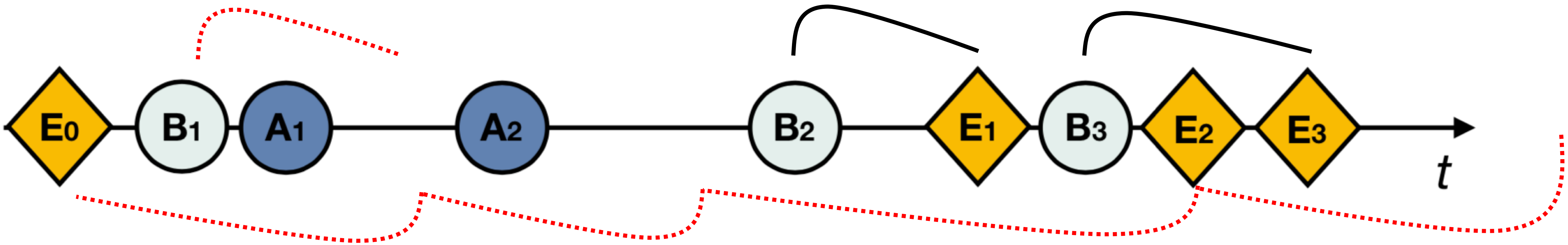


1.Actual Causal Attribution (Normative)

$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$



Reflective path: for supposed causes that did not have their corresponding effect *e* in the path, we must attribute them as either occurring after the end of the clip or as prevented.

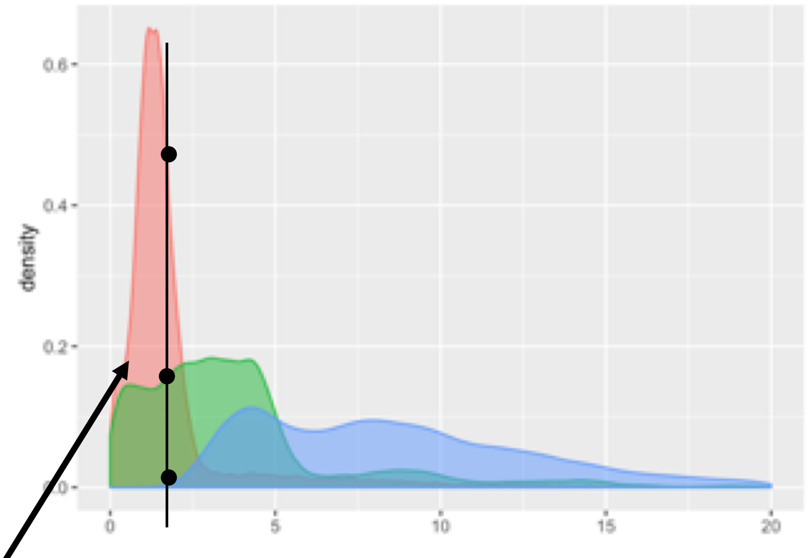


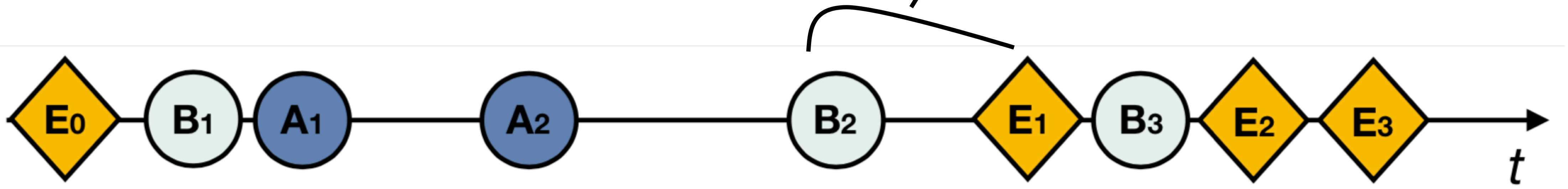
2. Feature-based Inference

$$P(S|\mathbf{d}_\tau, \mathbf{w}) \propto p(\mathbf{d}_\tau|S, \mathbf{w}) \cdot P(S)$$

"combine imagination with the reality to draw causal conclusions"

Feature 1: the ***delay*** between control component's activation and its nearest target component's activation.
Feature 2: following effect ***number*** before any other activation of any control component.





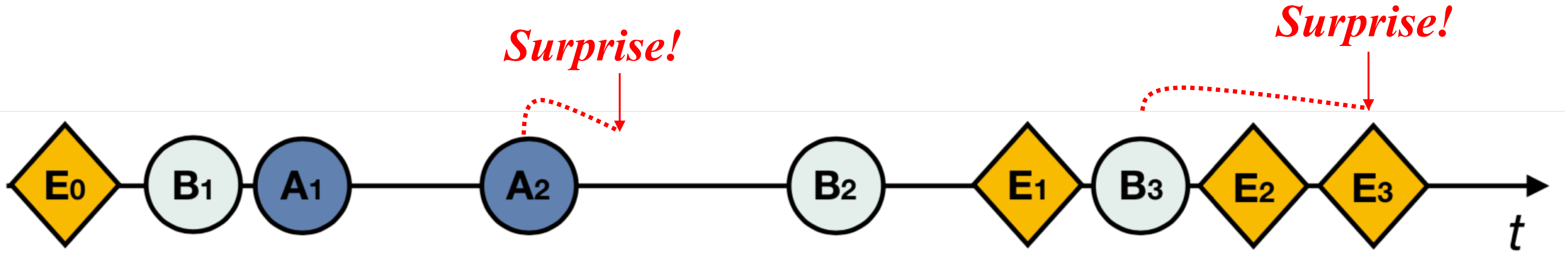
3. Expectation-violation local inference

"update their belief about one connection (local updating) every time the expectation of a target component's activation given the current belief is violated."

Expectation accumulation: $\mathbb{E}(e) = -\sum_{g' \in \mathbf{G_b}} \ln(1 - p(t_{g' \to e} > t_{g'e} | \mathbf{w}, b) \cdot c)$

Penalization: $c = \prod_{p' \in \mathbf{P_b}} p(t_{p' \to e} < t_{p'e} | \mathbf{w}, b))$

Initialization: both non-causal
Upper boundary (for preventative attribution): ***η***
Lower boundary (for generative attribution): ***exp(1−η)***



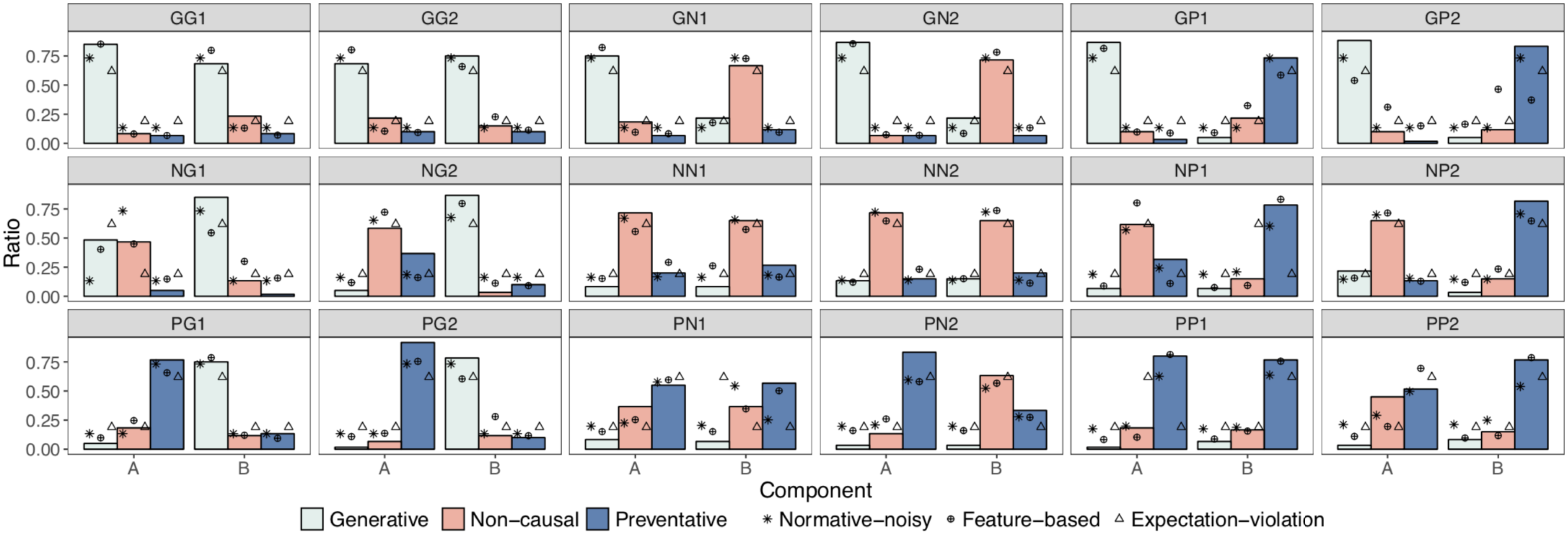
Model comparison

Table 1: Feature comparison between models.

Model	Forward thinking	Actual attribution	Full evidence	Multi-links updating
NN	✔	✔	✔	✔
FB	✔	✗	✗	✔
EV	✔	✔	✗	✗

Note: NN:Normative & Normative-noisy; FB:Feature-based; EV:Expectation-violation.

Human results



Human results

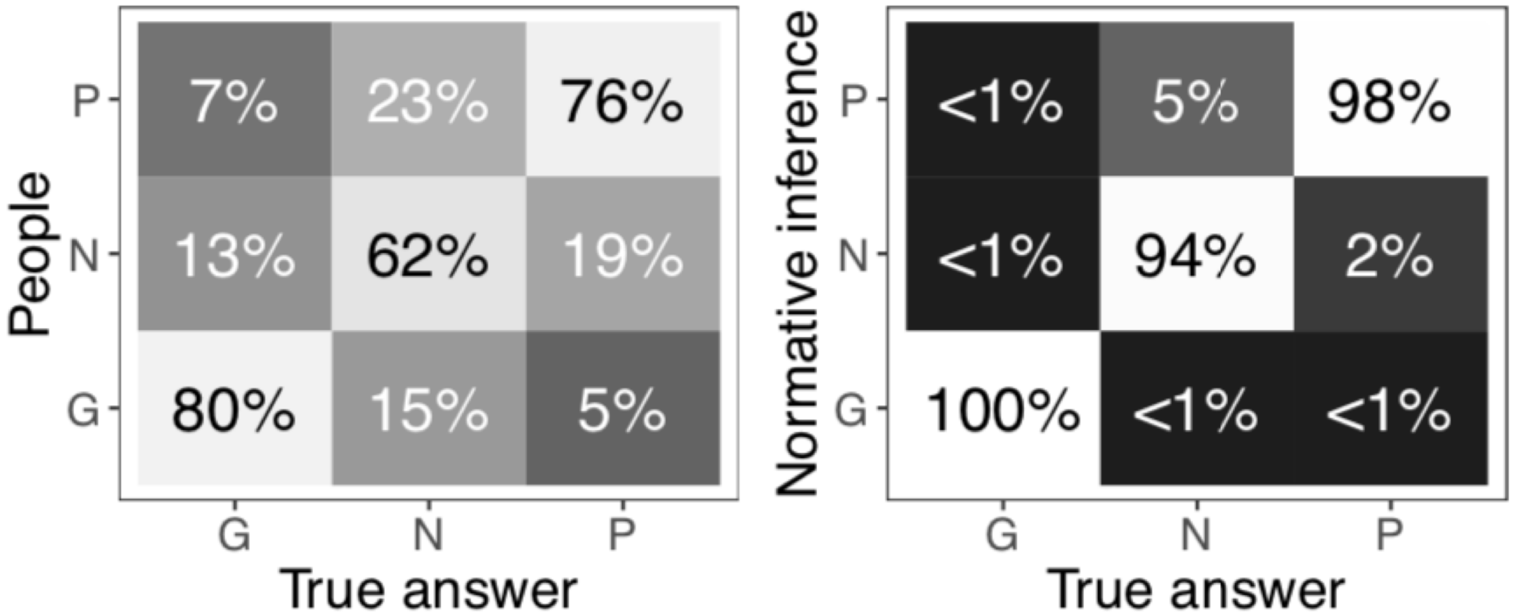


Figure 4: Overall choice patterns between different kinds of causal connections (participants vs. normative model).

Model comparison

$$p(n) = \frac{exp(\lambda \cdot v_n)}{\sum_{n' \in N} exp(\lambda \cdot v_{n'})}$$

Table 2: Model accuracy and fitting results at device level.

Model	Accuracy	Parameters	BIC	N Best
NN:	82-95%	λ:2.67; θ:3	3382	14/60
FB:				**35/60**
delay	60%	λ:3.11	3621	(12)
number	44%	λ:5.20	3559	(10)
combine	44-60%	λ_d:1.92; λ_n:3.52	**3278**	(13)
EV:	11-78%	λ:2.05; η:3.86	3782	6/60
RD:	11%		4768	5/60

Note: NN:Normative & Normative-noisy; FB:Feature-based; EV:Expectation-violation; RD:Random. Model accuracy was calculated prior to the fitting of human data and under consideration of noise (θ) and threshold (η) parameters.

Background

- Including generative and preventative causal learning in the same causal system
- Building both computational-level and algorithmic-level models to predict human inference
- Examining previous findings under the contingency/contiguity issues

Regularity

- Regular
  - Base rate: 5 ± 0.5
  - Generative: 1.5 ± 0.5
  - Preventative: 3 ± 0.5

- Semi-regular
  - Base rate: 5 ± 5 (memoryless)
  - Generative: 1.5 ± 0.5
  - Preventative: 3 ± 0.5

- Irregular
  - Base rate: 5 ± 5 (memoryless)
  - Generative: 1.5 ± 1
  - Preventative: 3 ± 1

Regularity-normative

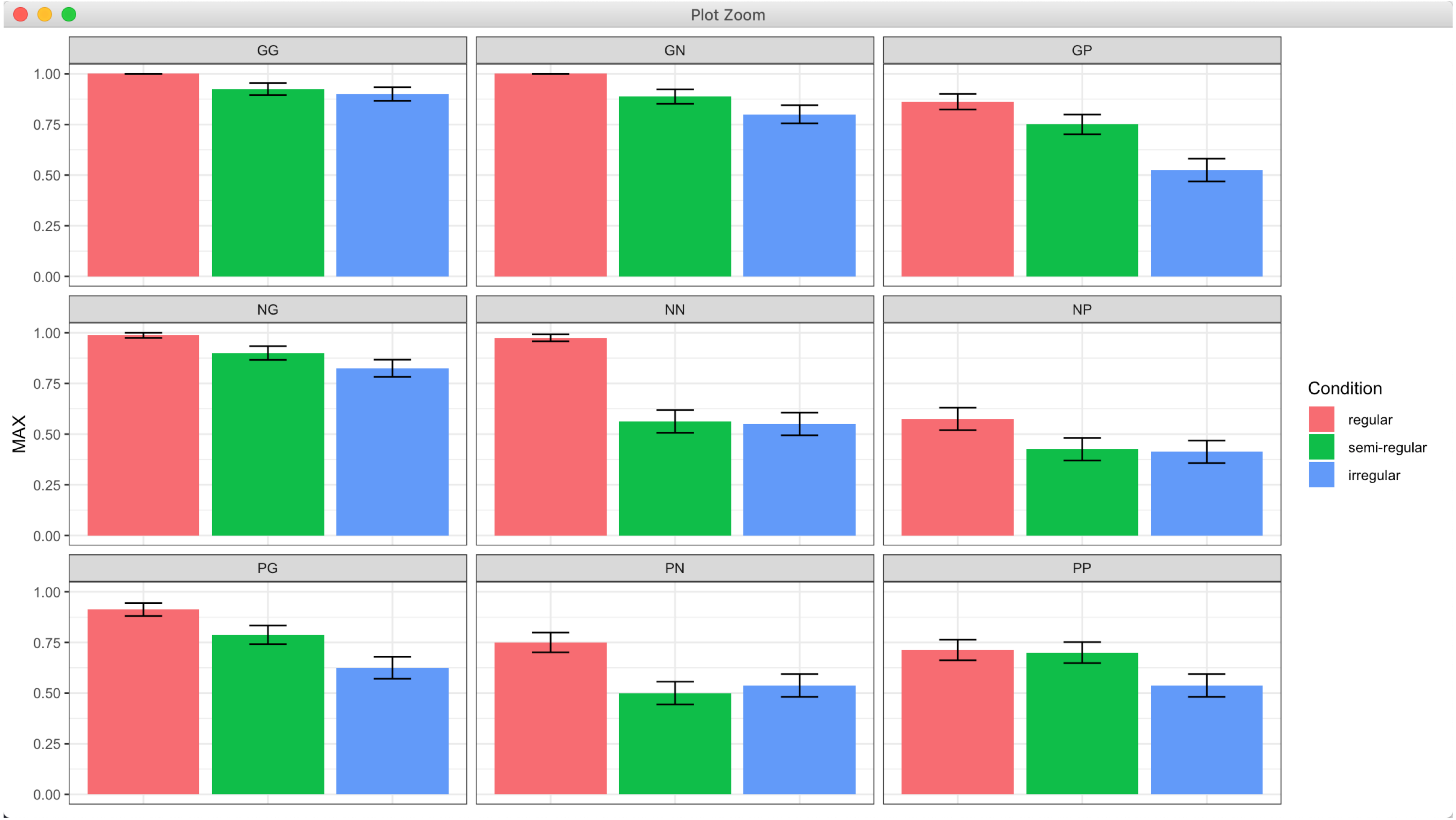
*Clip number in each regularity*structure = 80*

ACC

Regularity-normative

*Clip number in each regularity*structure = 80*

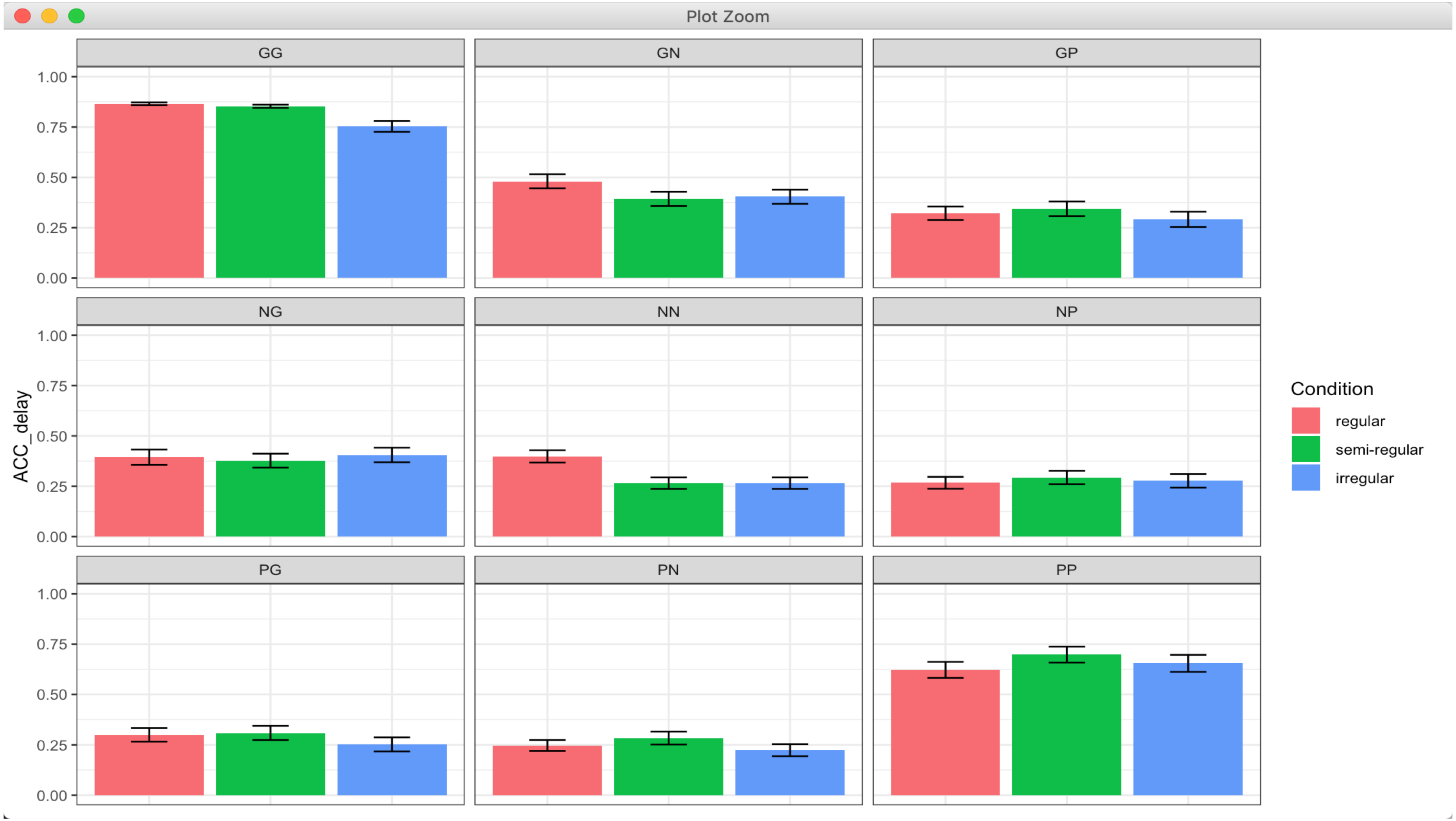
MAP



Regularity- F: delay

*Clip number in each regularity*structure = 60 (expectation simulation=1000)*

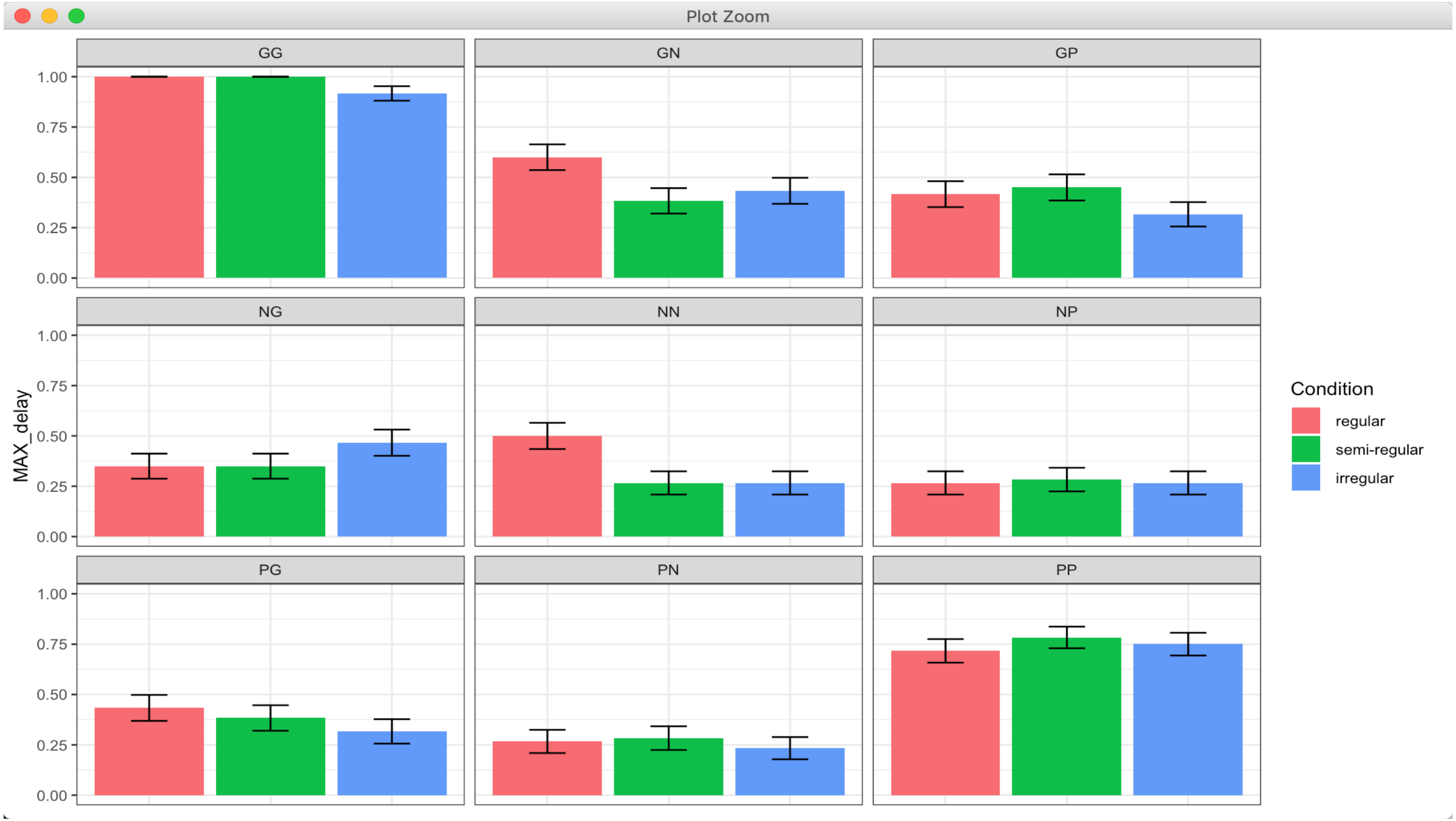
ACC



Regularity- F: delay

*Clip number in each regularity*structure = 60 (expectation simulation=1000)*

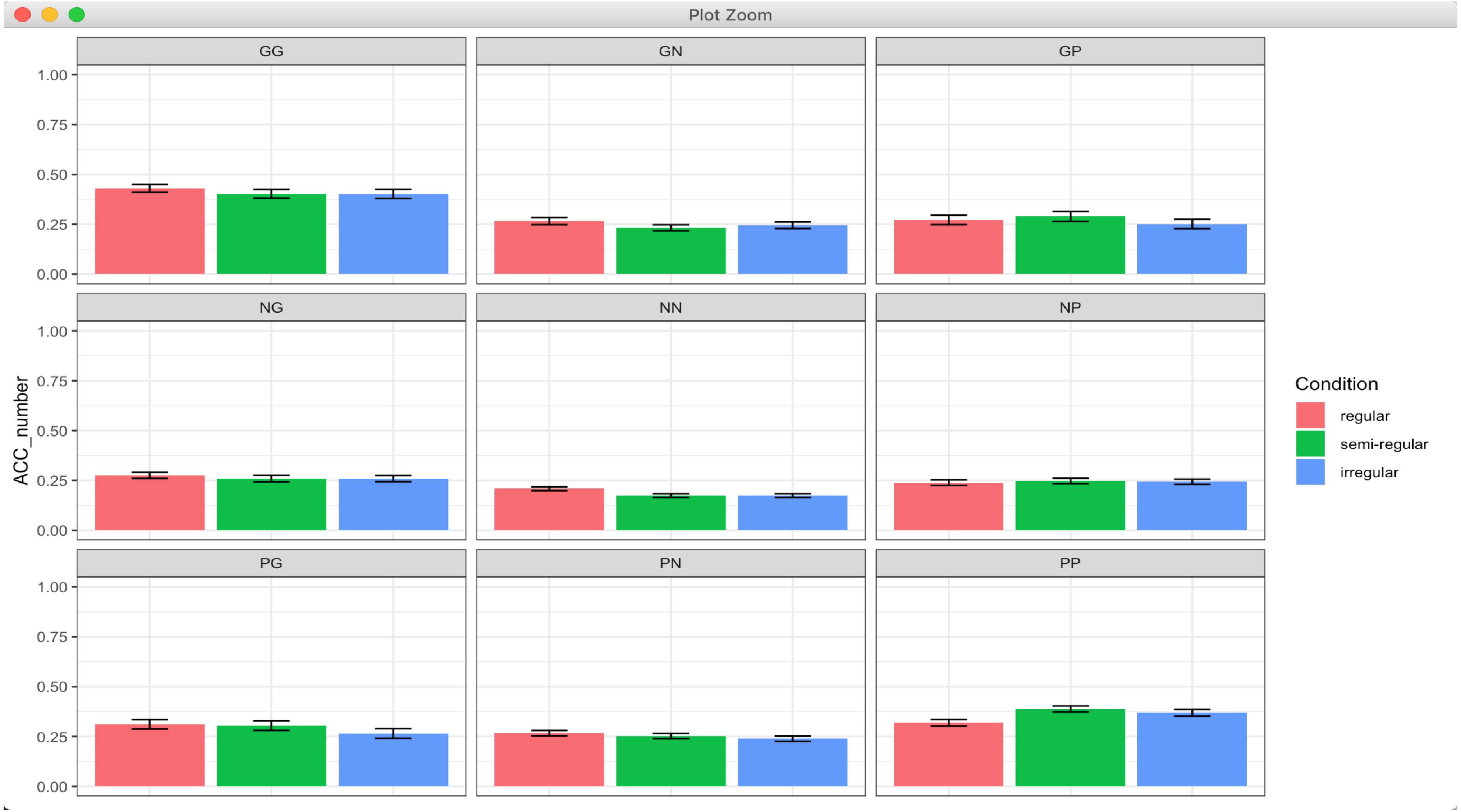
MAP



Regularity- F: num

*Clip number in each regularity*structure = 60 (expectation simulation=1000)*

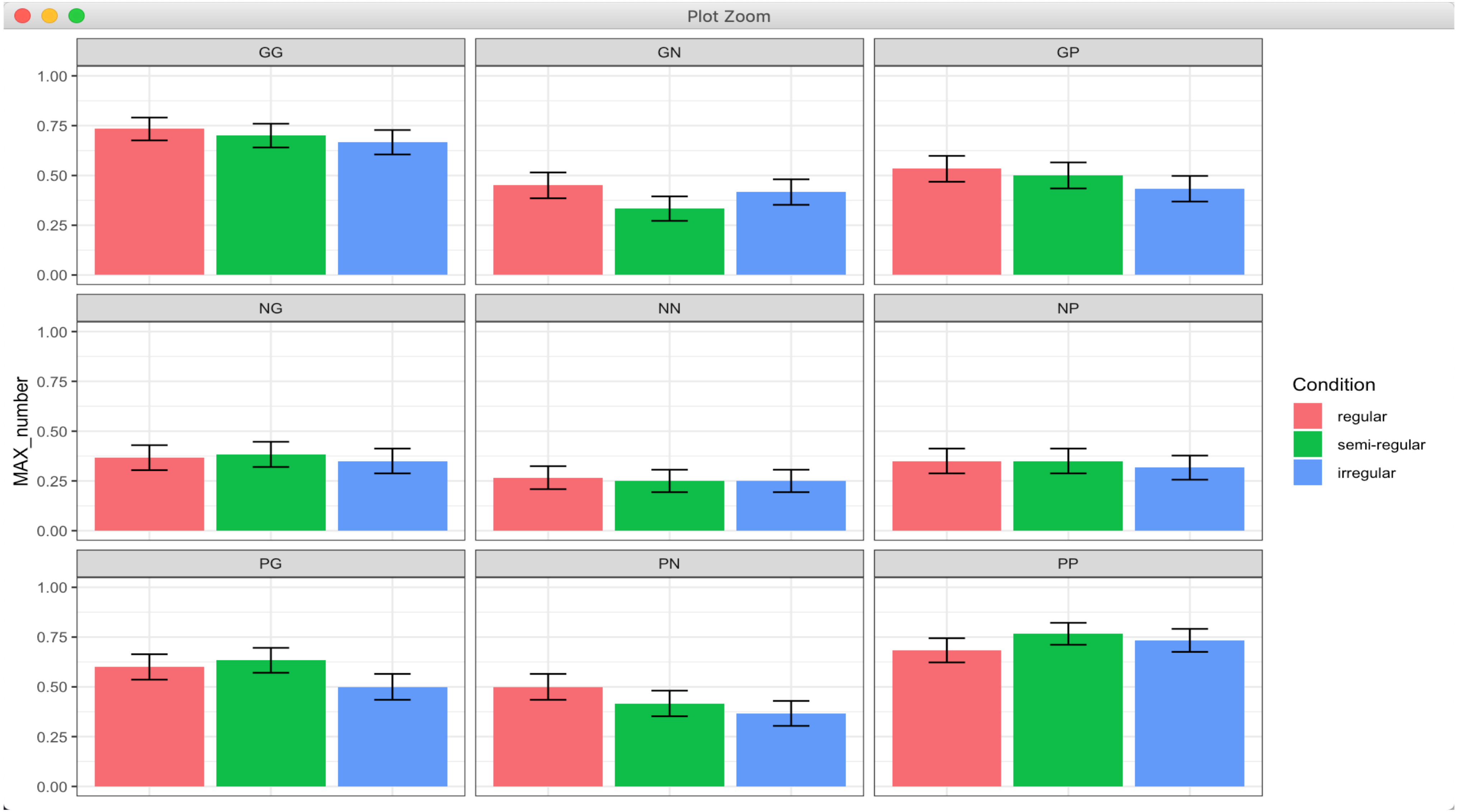
ACC



Regularity- F: num

*Clip number in each regularity*structure = 60 (expectation simulation=1000)*

MAP



Base Rate

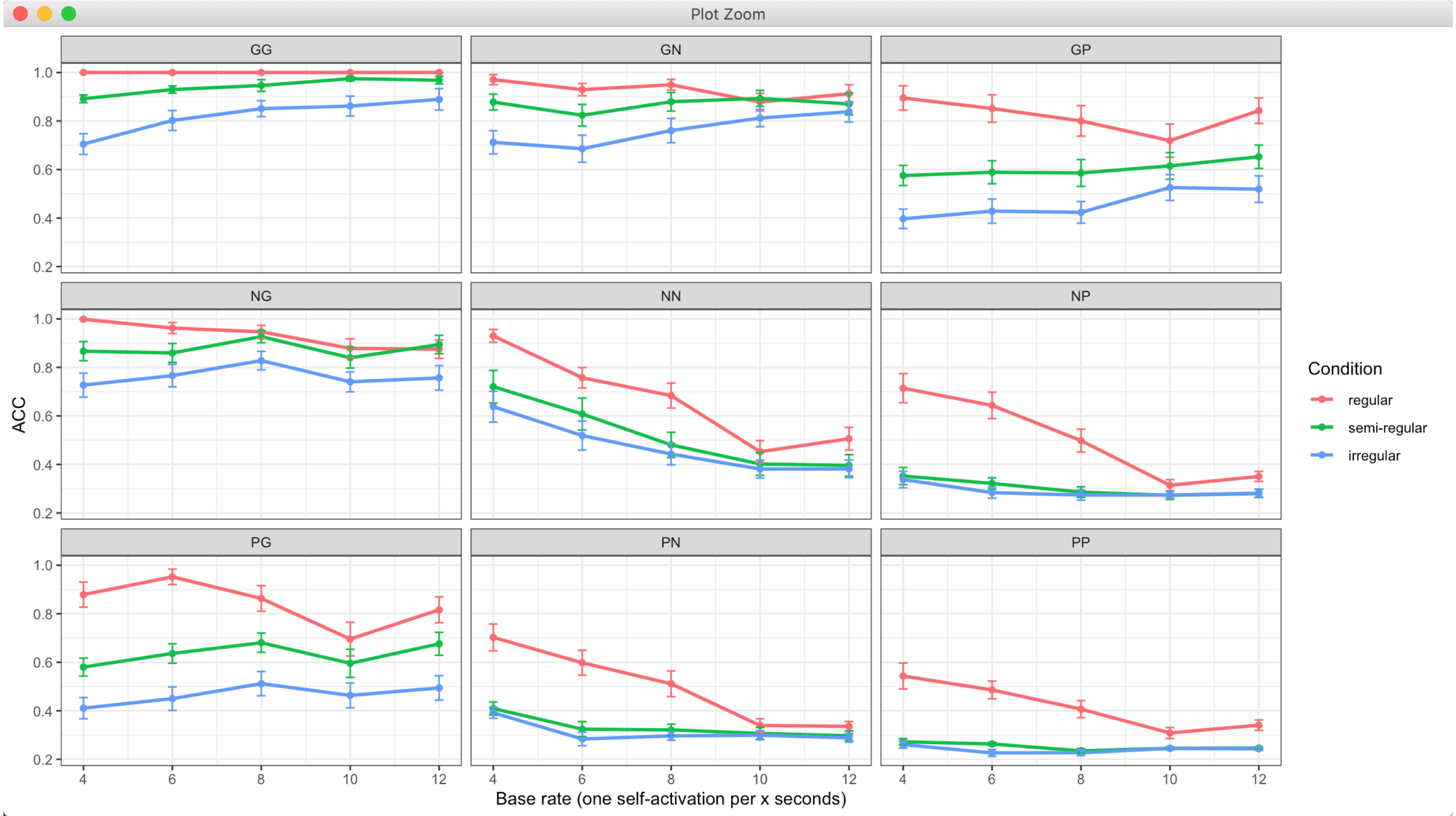
- Regular
  - Base rate: 5 ± 0.5 — *{4±0.5, 6±0.5, 8±0.5, 10±0.5, 12±0.5}*
  - Generative: 1.5 ± 0.5
  - Preventative: 3 ± 0.5

- Semi-regular
  - Base rate: 5 ± 5 (memoryless) — *{4±4, 6±6, 8±8, 10±10, 12±12}*
  - Generative: 1.5 ± 0.5
  - Preventative: 3 ± 0.5

- Irregular
  - Base rate: 5 ± 5 (memoryless) — *{4±4, 6±6, 8±8, 10±10, 12±12}*
  - Generative: 1.5 ± 1
  - Preventative: 3 ± 1

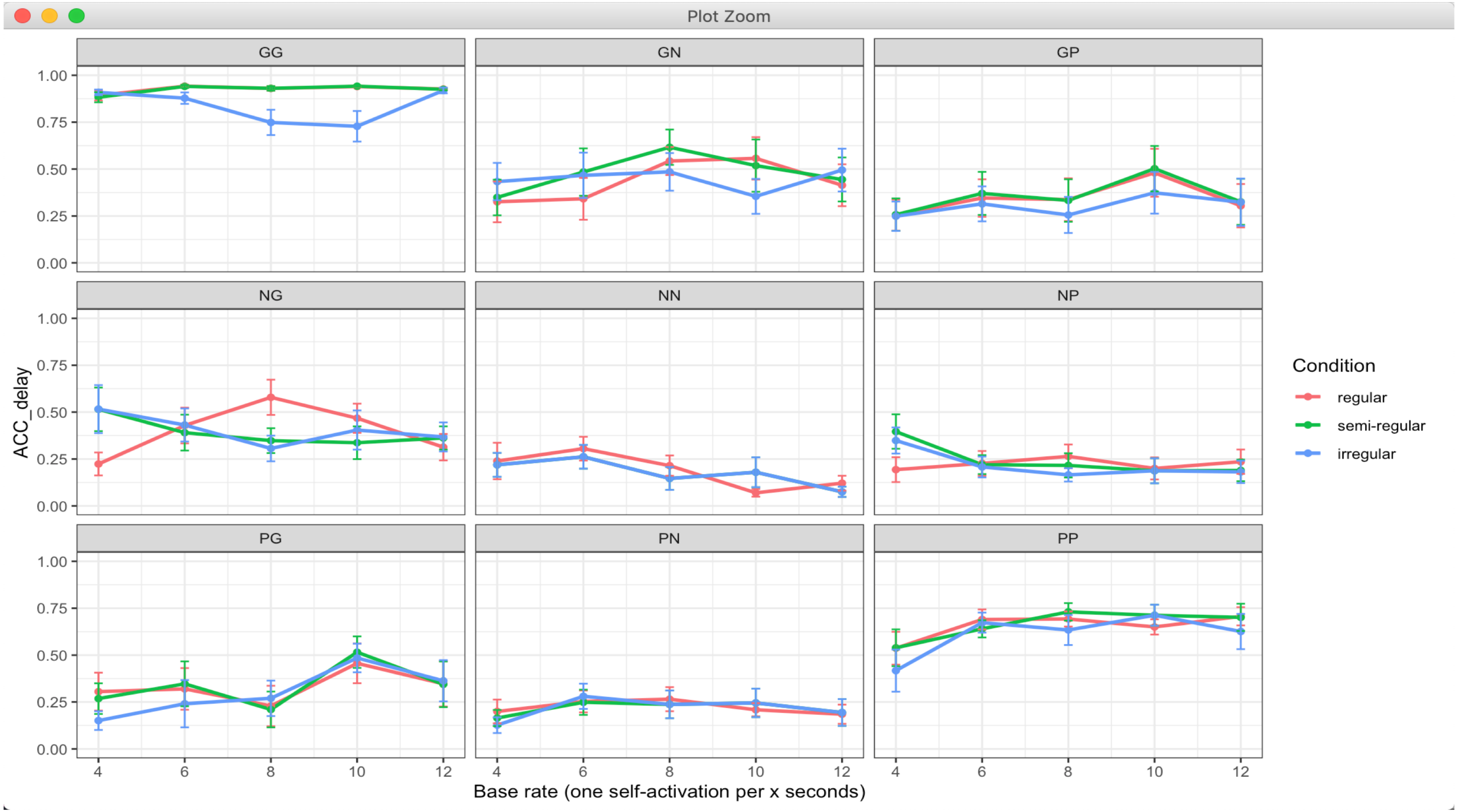
Base Rate-normative

Clip number in each rate
**structure = 30*



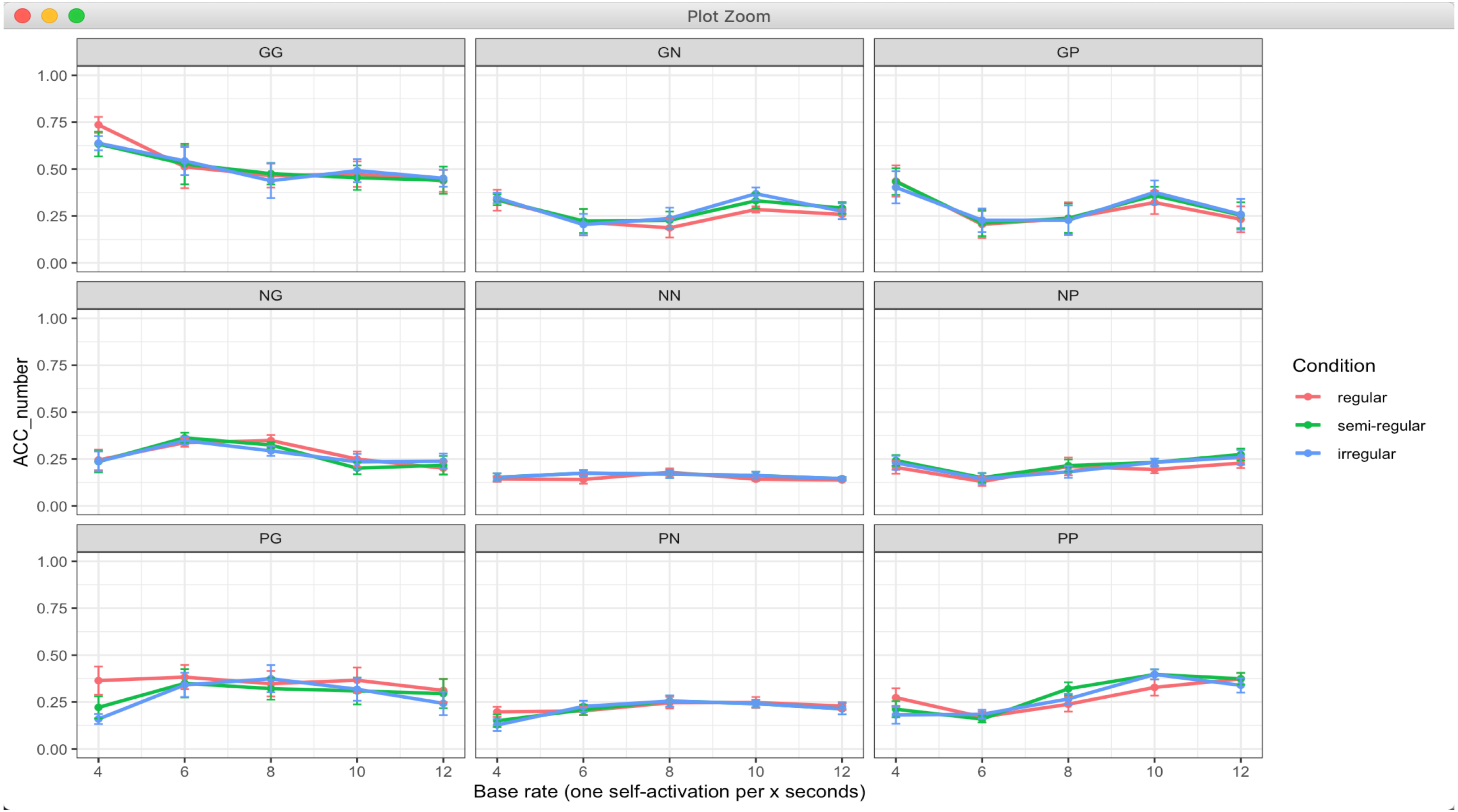
Base Rate-F: delay

*Clip number in each regularity*structure = 8 (expectation simulation=720)*



Base Rate-F: num

*Clip number in each regularity*structure = 8 (expectation simulation=720)*



Causal Power (noisy)

Noisy-OR & Noisy-AND-NOT theory (causal structure learning)

$$P_1(e^+|b,c;w_0,w_1) = w_0^b(1-w_1)^c$$

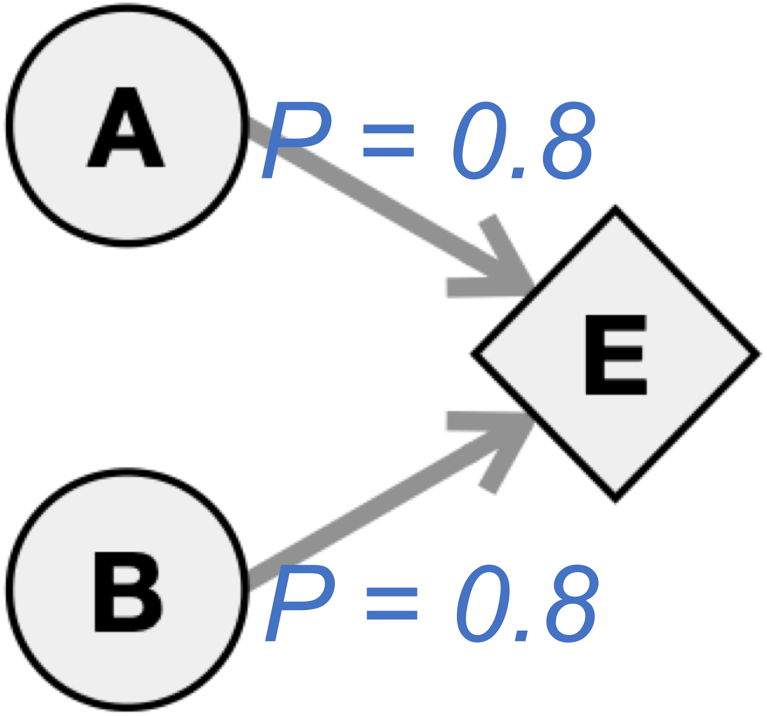
$$P_1(e^+|b,c;w_0,w_1) = 1-(1-w_0)^b(1-w_1)^c.$$

Causal Power / delta-P (causal strength learning)

$$\Delta P = P(E|C) - P(E|\neg C)$$

$$w_C = \frac{\Delta P}{1-P(E|\neg C)}$$

$$w_C = \frac{-\Delta P}{P(E|\neg C)}$$

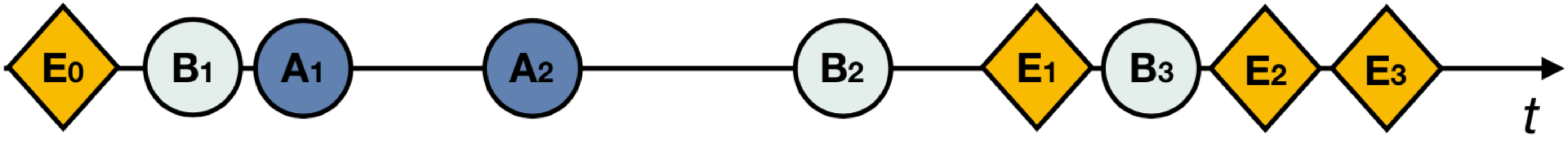


Causal Power (noisy)

P = 0.8 ?

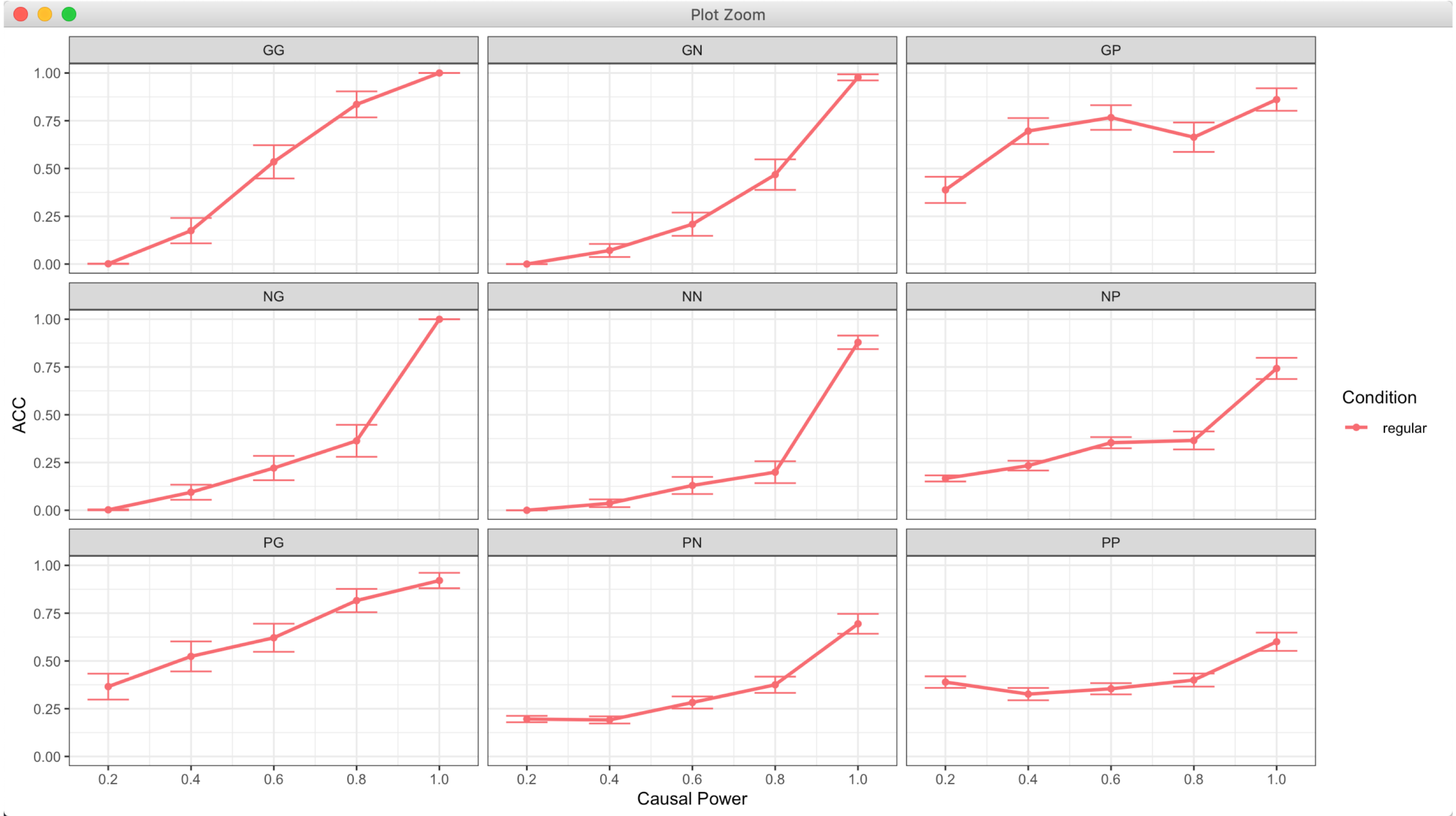
- Effect's self-activation (E)
  - P(self-activating once) = 0.8
- Generative effect (e.g., B1)
  - P(producing an effect) = 0.8
- Preventative effect (e.g., A1)
  - P(blocking all effect) = 0.8
  - P(not blocking any effect) = 0.2

- Effect's self-activation (E)
  - P(self-activating once) = 0.8
- Generative effect (e.g., B1)
  - P(producing an effect) = 0.8
- Preventative effect (e.g., A1)
  - P(blocking one effect) = 0.8
  - P(blocking two effect) = 0.8*0.8



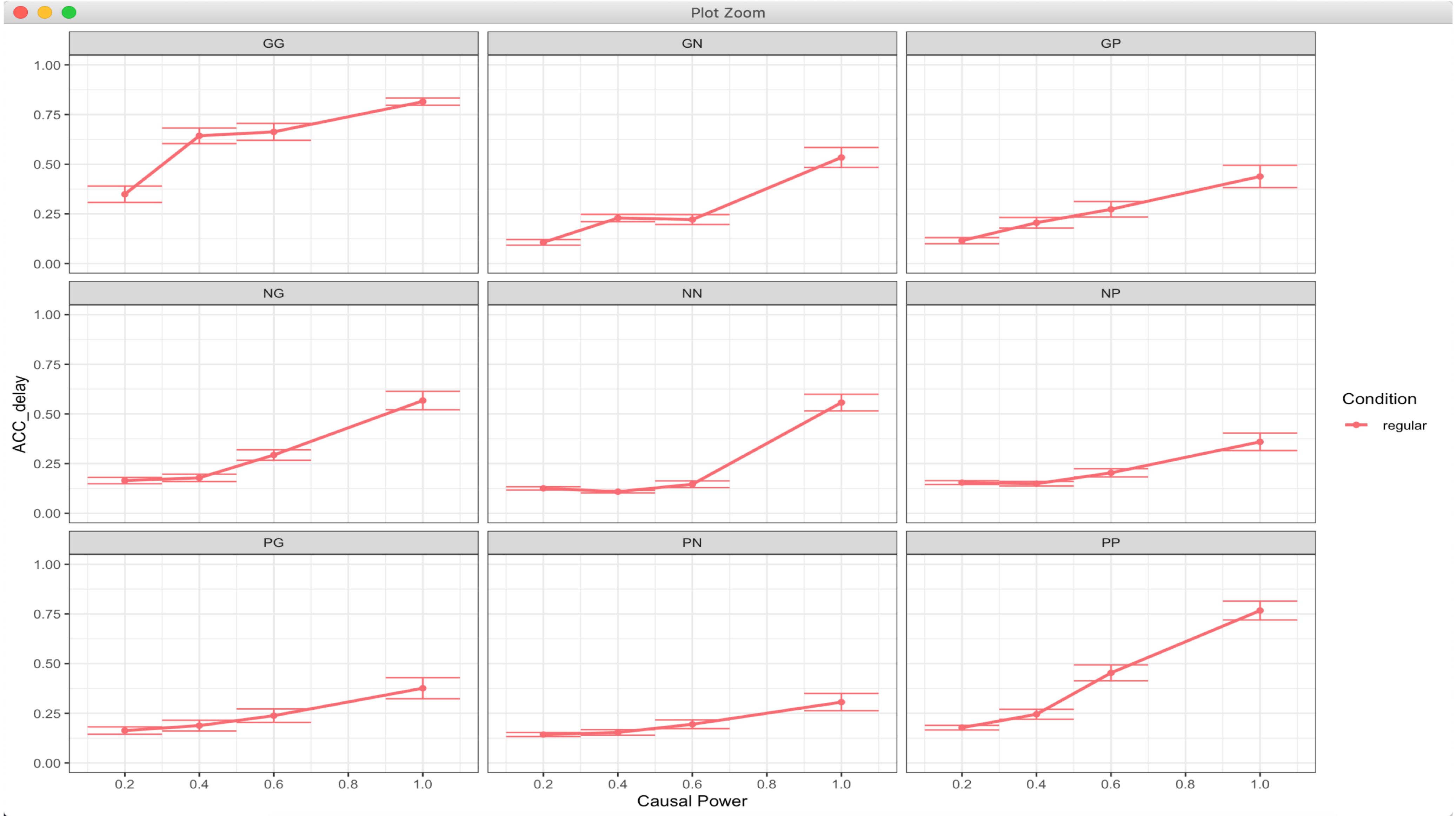
Cpower-normative

*Clip number in each power*structure = 30*



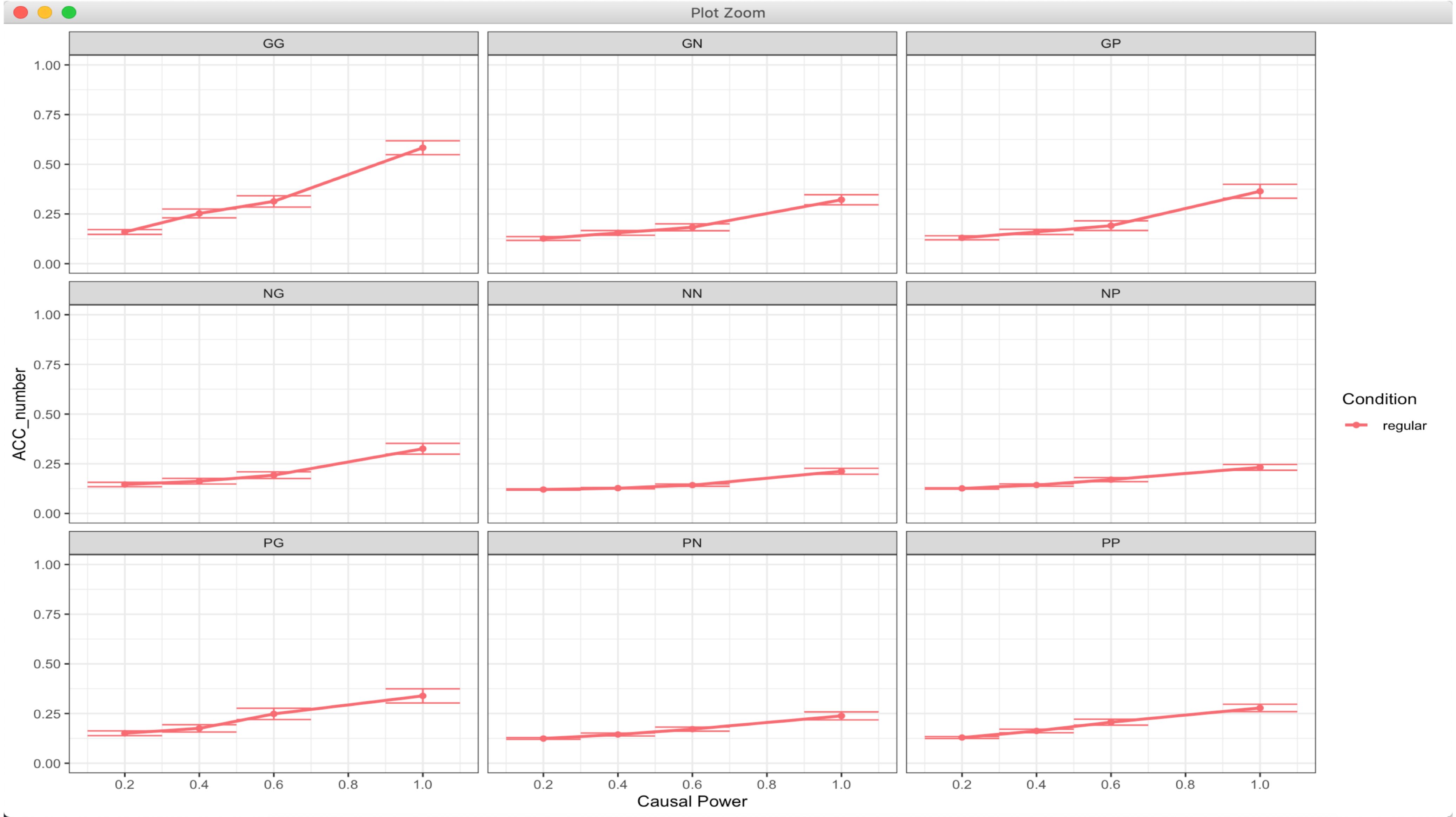
Cpower- F: delay

*Clip number in each regularity*structure =40 (expectation simulation=570)*



Cpower- F: num

*Clip number in each regularity*structure =40*
(expectation simulation=570)



Prevention Scope

Hypothesis space = 3