Online Platforms and the Fair Exposure Problem under Homophily

Jakob Schoeffer*[1], Alexander Ritchie*[2], Keziah Naggita*[3], Faidra Monachou*[4], Jessie Finocchiaro*[5], Marc Juarez[6]

1 Karlsruhe Institute of Technology 2 University of Michigan 3 Toyota Technical Institute of Chicago 4 Yale School of Management
5 Harvard University CRCS 6 University of Edinburgh

INFORMS 2023 MD18 Fairness in Platforms and Recommendations

* denotes equal contribution; listed in reverse alphabetical order

A brief introduction

A brief introduction

Algorithmic decision-making

Design of optimization problems

A brief introduction

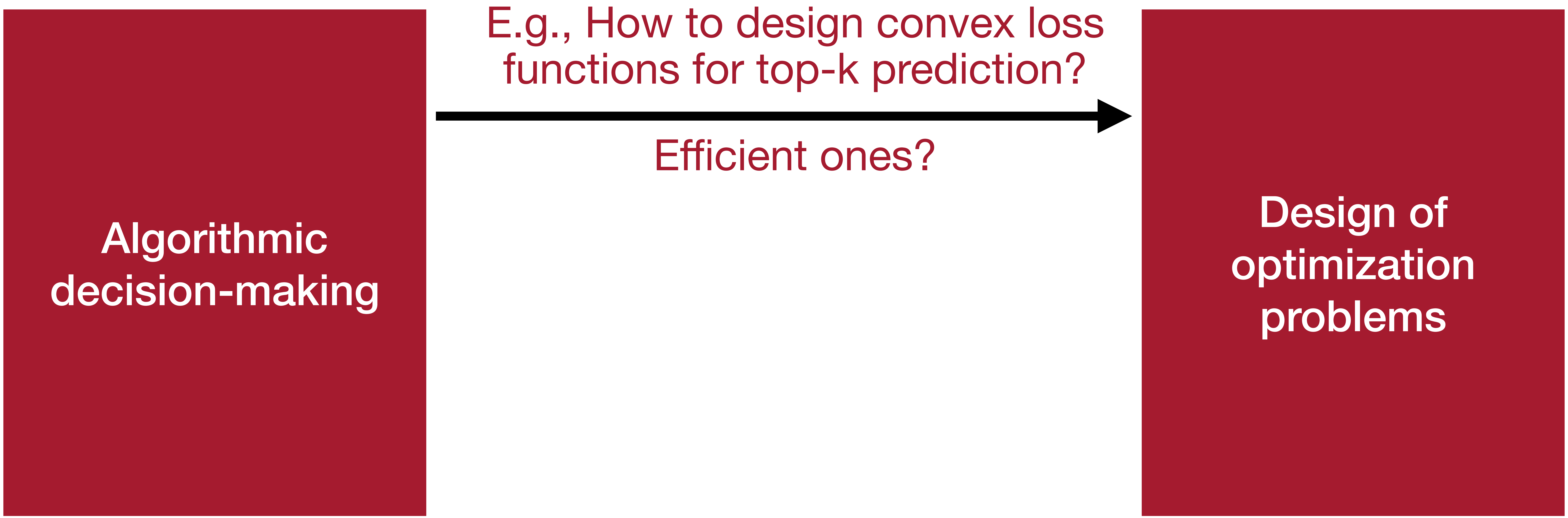
A brief introduction

Algorithmic decision-making

E.g., How to design convex loss functions for top-k prediction?

Design of optimization problems

A brief introduction



A brief introduction

Algorithmic decision-making

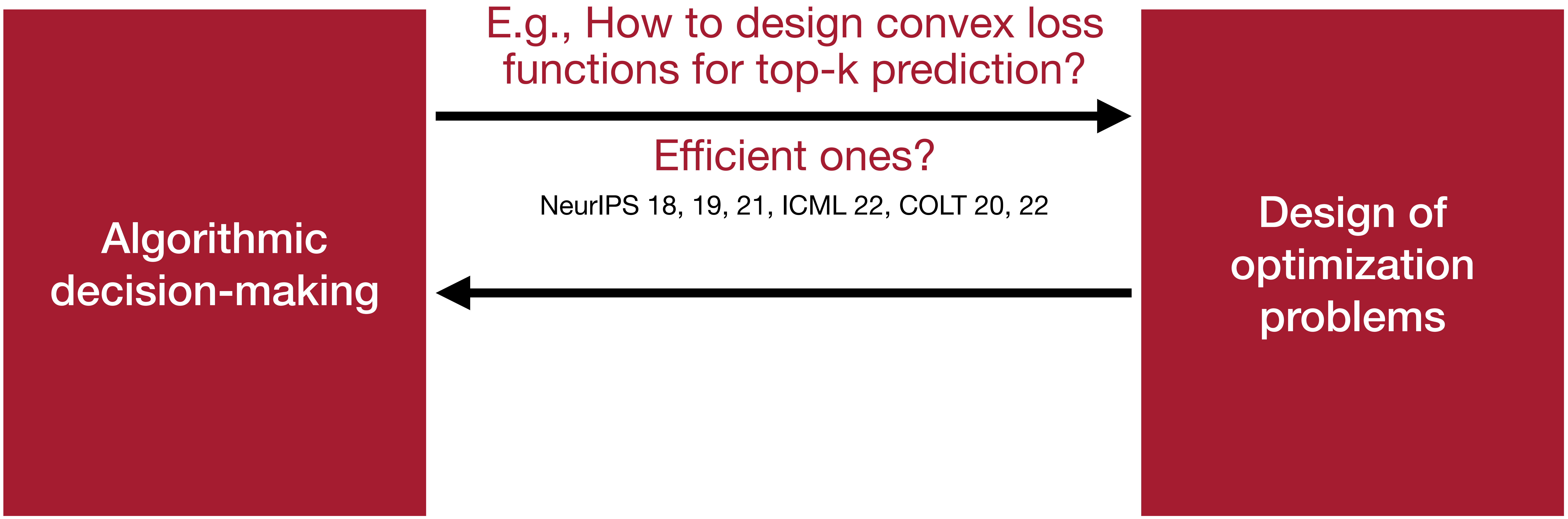
E.g., How to design convex loss functions for top-k prediction?

Efficient ones?

NeurIPS 18, 19, 21, ICML 22, COLT 20, 22

Design of optimization problems

A brief introduction



A brief introduction

Algorithmic decision-making

E.g., How to design convex loss functions for top-k prediction?

Efficient ones?

NeurIPS 18, 19, 21, ICML 22, COLT 20, 22

Design of optimization problems

E.g., How do fairness constraints change algorithmic decision-making?

A brief introduction

Algorithmic decision-making

E.g., How to design convex loss functions for top-k prediction?

Efficient ones?

NeurIPS 18, 19, 21, ICML 22, COLT 20, 22

Design of optimization problems

E.g., How do fairness constraints change algorithmic decision-making?

IEEE Transactions on Games 20, FAccT 21,
ITCS 24 (in submission), ALT 24 (in submission)

A brief introduction

Algorithmic decision-making

E.g., How to design convex loss functions for top-k prediction?

Efficient ones?

NeurIPS 18, 19, 21, ICML 22, COLT 20, 22

Design of optimization problems

E.g., How do fairness constraints change algorithmic decision-making?

IEEE Transactions on Games 20, FAccT 21,
ITCS 24 (in submission), ALT 24 (in submission)

Today: in online platforms (AAAI 23)

Increasing polarization in online platforms

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

Increasing polarization in online platforms

Yarchi et al. (2021) discuss three types of polarization on social media

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

Increasing polarization in online platforms

Yarchi et al. (2021) discuss three types of polarization on social media

Interactional: "increasing interaction with like-minded individuals"

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

Increasing polarization in online platforms

Yarchi et al. (2021) discuss three types of polarization on social media

Interactional: “increasing interaction with like-minded individuals”

Positional: “increases in antagonistic and extreme political preferences”

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

Increasing polarization in online platforms

Yarchi et al. (2021) discuss three types of polarization on social media

Interactional: “increasing interaction with like-minded individuals”

Positional: “increases in antagonistic and extreme political preferences”

“Interactional polarization aggravates positional polarization”

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

Increasing polarization in online platforms

Yarchi et al. (2021) discuss three types of polarization on social media

Interactional: "increasing interaction with like-minded individuals"

Positional: "increases in antagonistic and extreme political preferences"

"Interactional polarization aggravates positional polarization"

Promote informational diversity to mitigate positional polarization

Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

Accusations that broadcast stations (radio, TV news) were being used to air one-sided viewpoints and less mainstream views had no platform

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

Accusations that broadcast stations (radio, TV news) were being used to air one-sided viewpoints and less mainstream views had no platform

FCC fairness doctrine asked broadcasters…

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

Accusations that broadcast stations (radio, TV news) were being used to air one-sided viewpoints and less mainstream views had no platform

FCC fairness doctrine asked broadcasters…

  Provide adequate coverage of public issues

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

Accusations that broadcast stations (radio, TV news) were being used to air one-sided viewpoints and less mainstream views had no platform

FCC fairness doctrine asked broadcasters…

- Provide adequate coverage of public issues
- Ensure coverage fairly represented opposing views

A historical attempt to promote informational diversity

FCC Fairness Doctrine (1949)

Accusations that broadcast stations (radio, TV news) were being used to air one-sided viewpoints and less mainstream views had no platform

FCC fairness doctrine asked broadcasters…

Provide adequate coverage of public issues

Ensure coverage fairly represented opposing views

Informational diversity: Could something in this spirit translate to online platforms?

The fair exposure problem

The fair exposure problem:
Given limited intervention power of a platform, add constraints to enforce a balance in the exposure of content (e.g., impressions of news articles) among groups of users

Broader contexts

Broader contexts

How polarization arises (advertisers' strategy)

Hązła et al. 2023, Gaitonde et al. 2021

How user opinions change

Dean et al. 2022

How users enter/leave platform

Full computational resources: Celis et al. 2019

Limited intervention capacity: Ali et al. 2019, Fish et al. 2019, Stoica et al. 2020

Empirical examinations of polarization

Bakshy et al. 2015, Garimella et al. 2017, Yarchi et al. 2021

Broader contexts

How polarization arises (advertisers' strategy)

Hązła et al. 2023, Gaitonde et al. 2021

How user opinions change

Dean et al. 2022

How users enter/leave platform

Full computational resources: Celis et al. 2019

Limited intervention capacity: Ali et al. 2019, Fish et al. 2019, Stoica et al. 2020

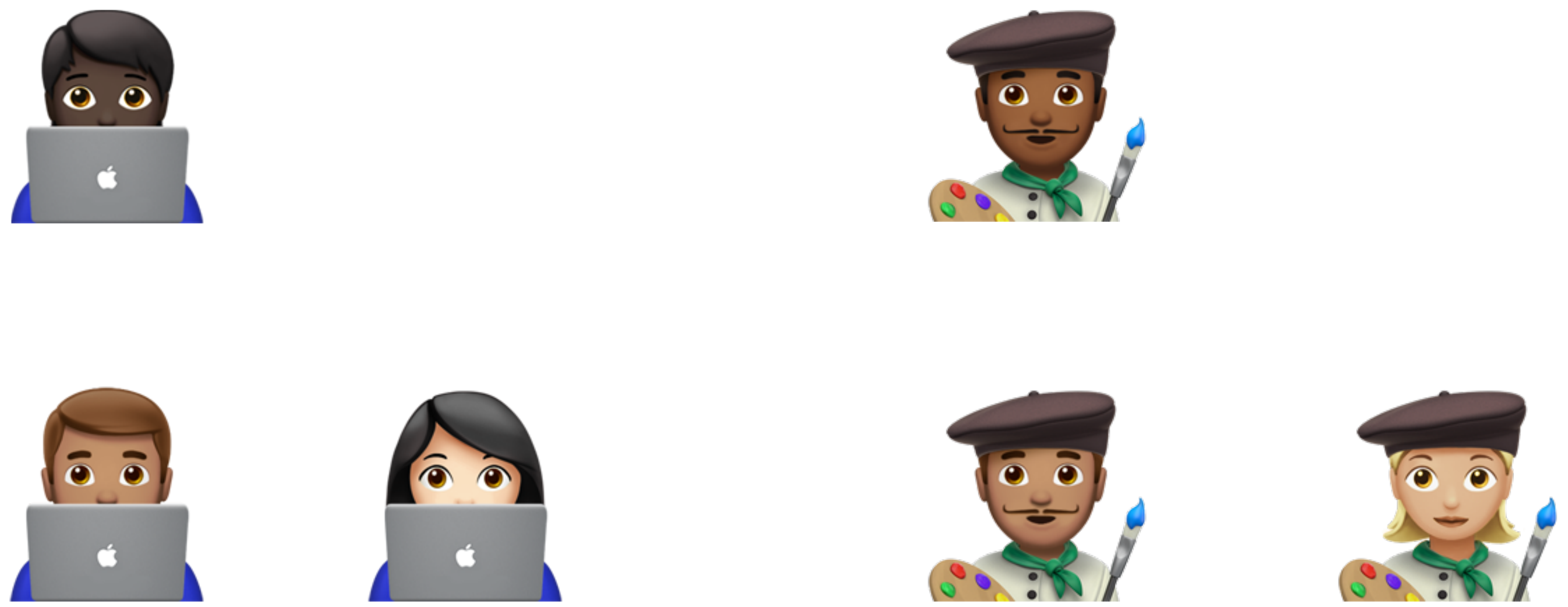
Empirical examinations of polarization

Bakshy et al. 2015, Garimella et al. 2017, Yarchi et al. 2021

We look at how policy constraints change *platform* behavior with limited intervention capacity

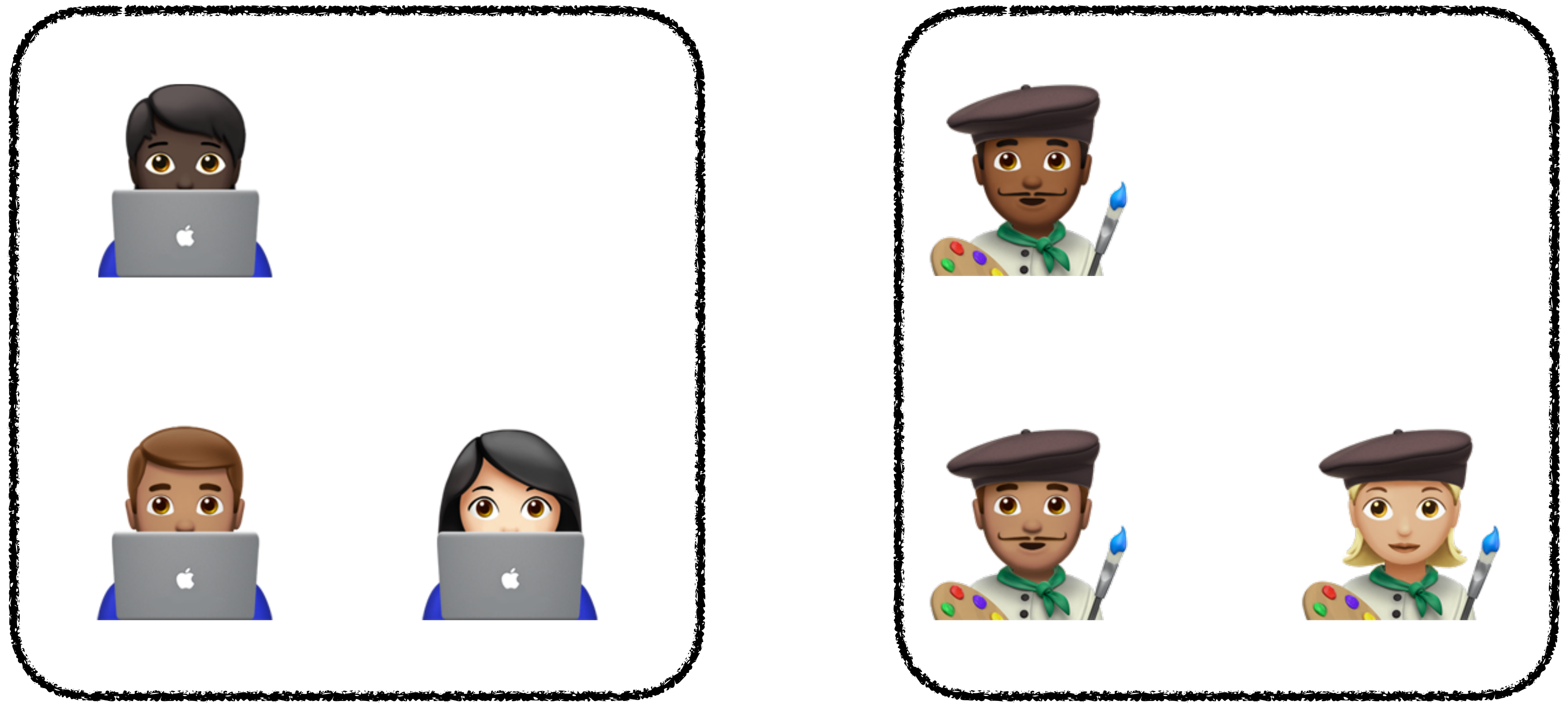
Model participants

Model participants

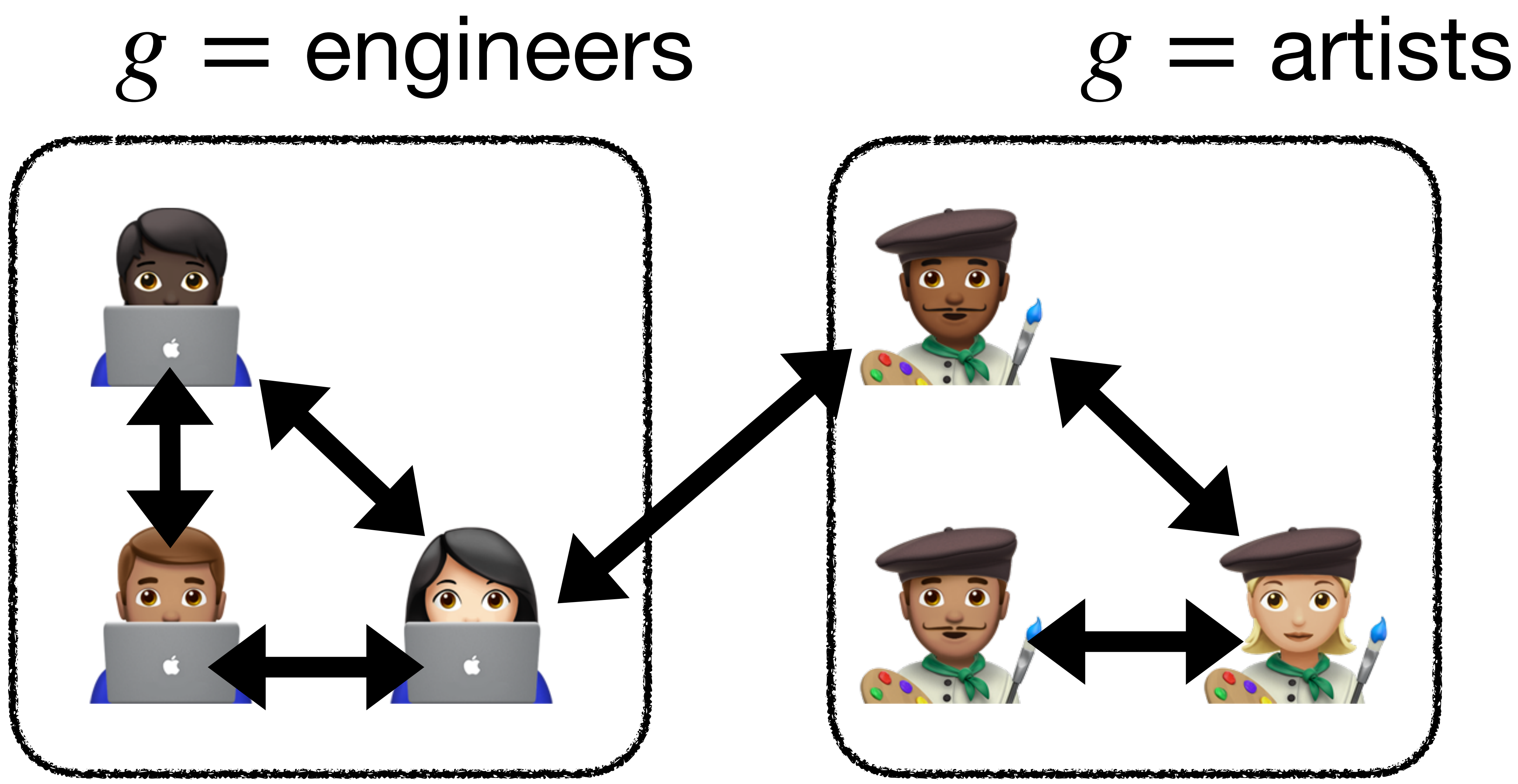


Model participants

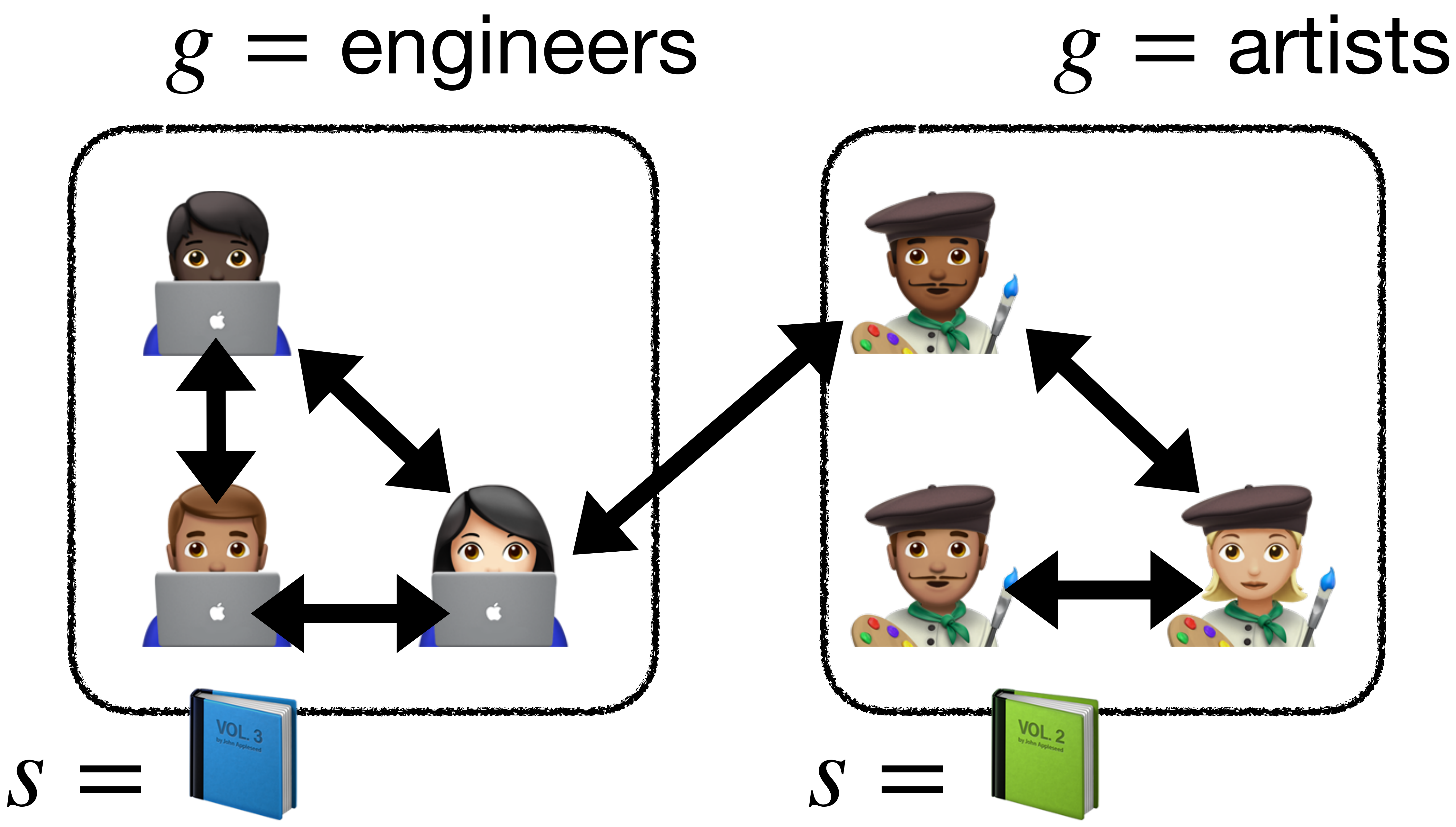
$g =$ engineers $\quad$ $g =$ artists



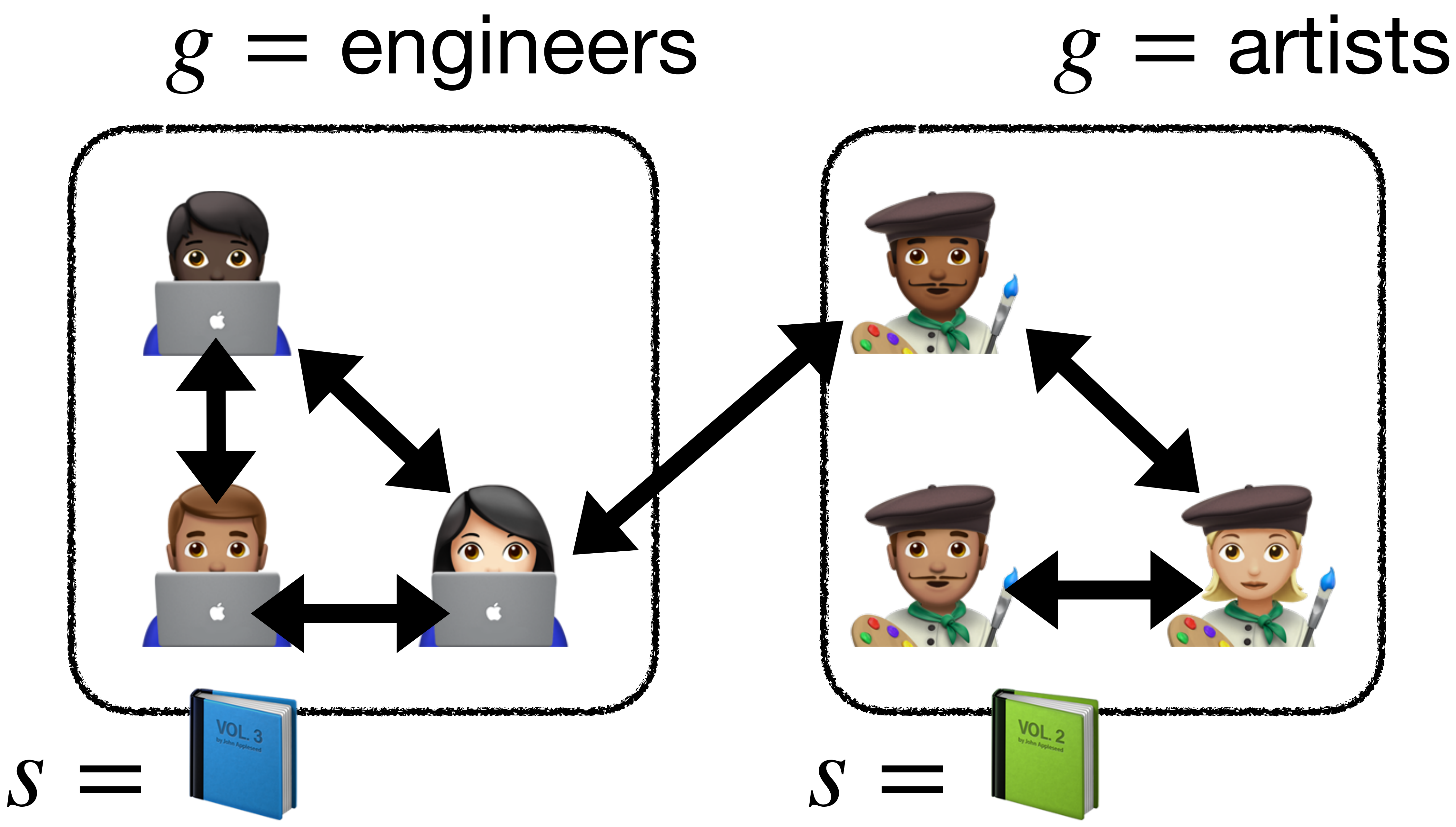
Model participants



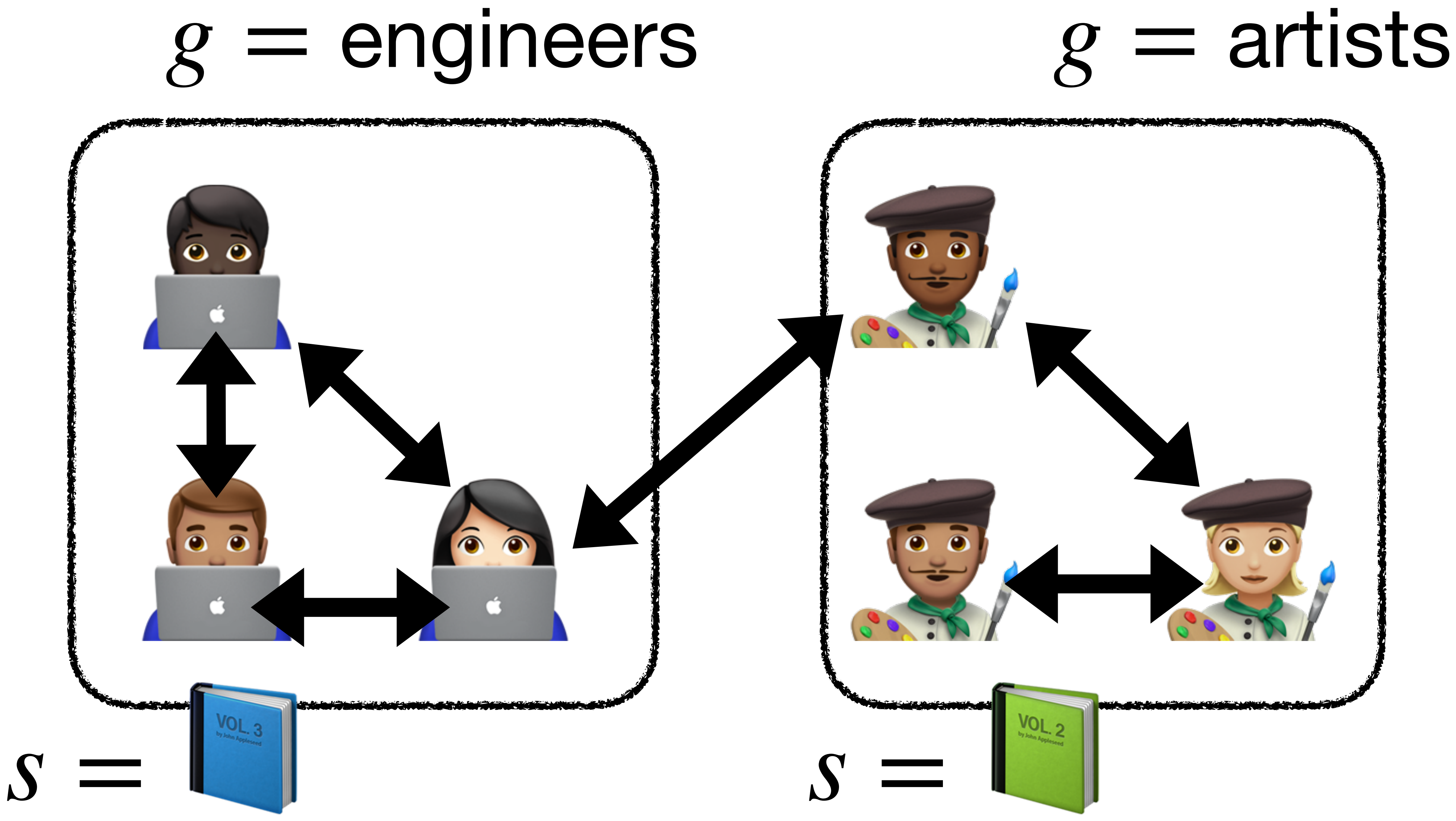
Model participants



Model participants



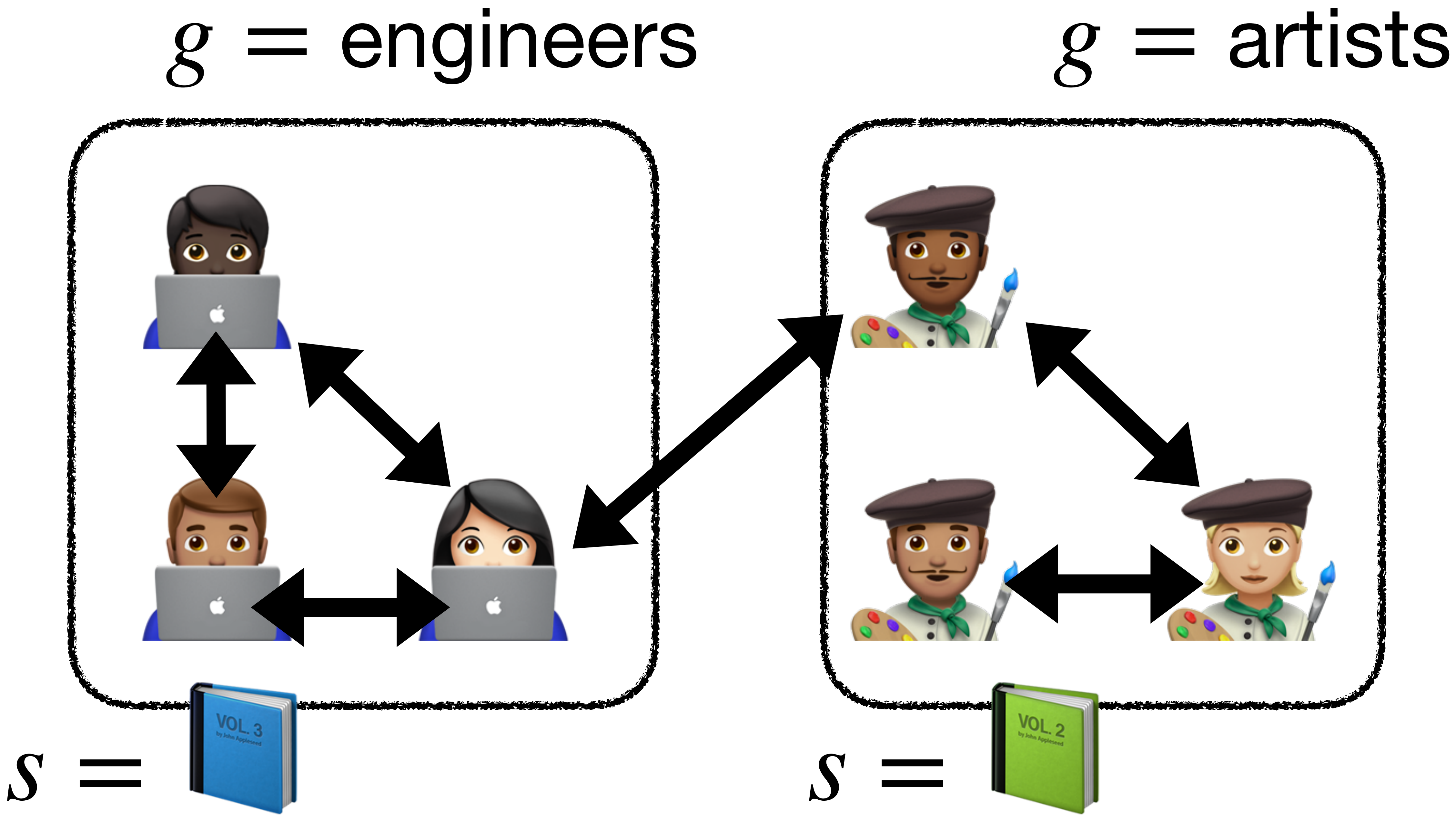
Model participants



$\Theta_{g,s} = \Pr[\text{platform shows content } s \text{ to user in group } g]$

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

Model participants



$\Theta_{g,s} = \Pr[$platform shows content s to user in group $g]$

$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$

$\Theta_{b,B} = 1 \implies$ Artist gets shown green content with probability 1.

👩🏾‍💻:📘, 👩🏼‍🎨:📗

Model dynamics

👩🏾‍💻:📘, 👩🏼‍🎨:📗

Model dynamics

$t = 1$

👩🏾‍💻:📘, 👩🏼‍🎨:📗

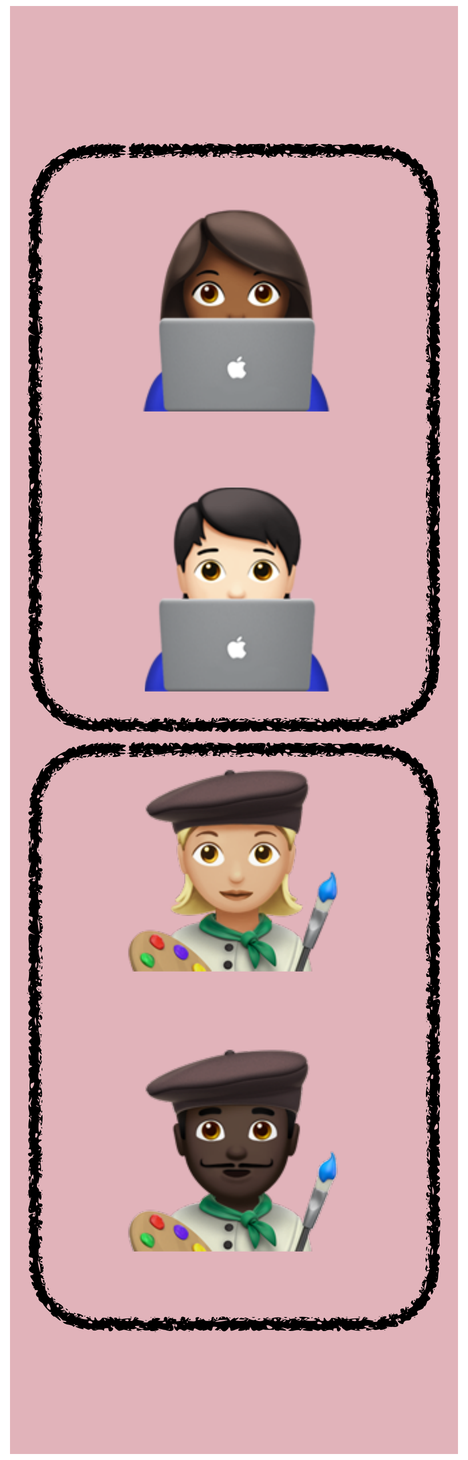
Model dynamics

$t = 1$

$t = \{2, 3, \ldots, T\}$

Model dynamics

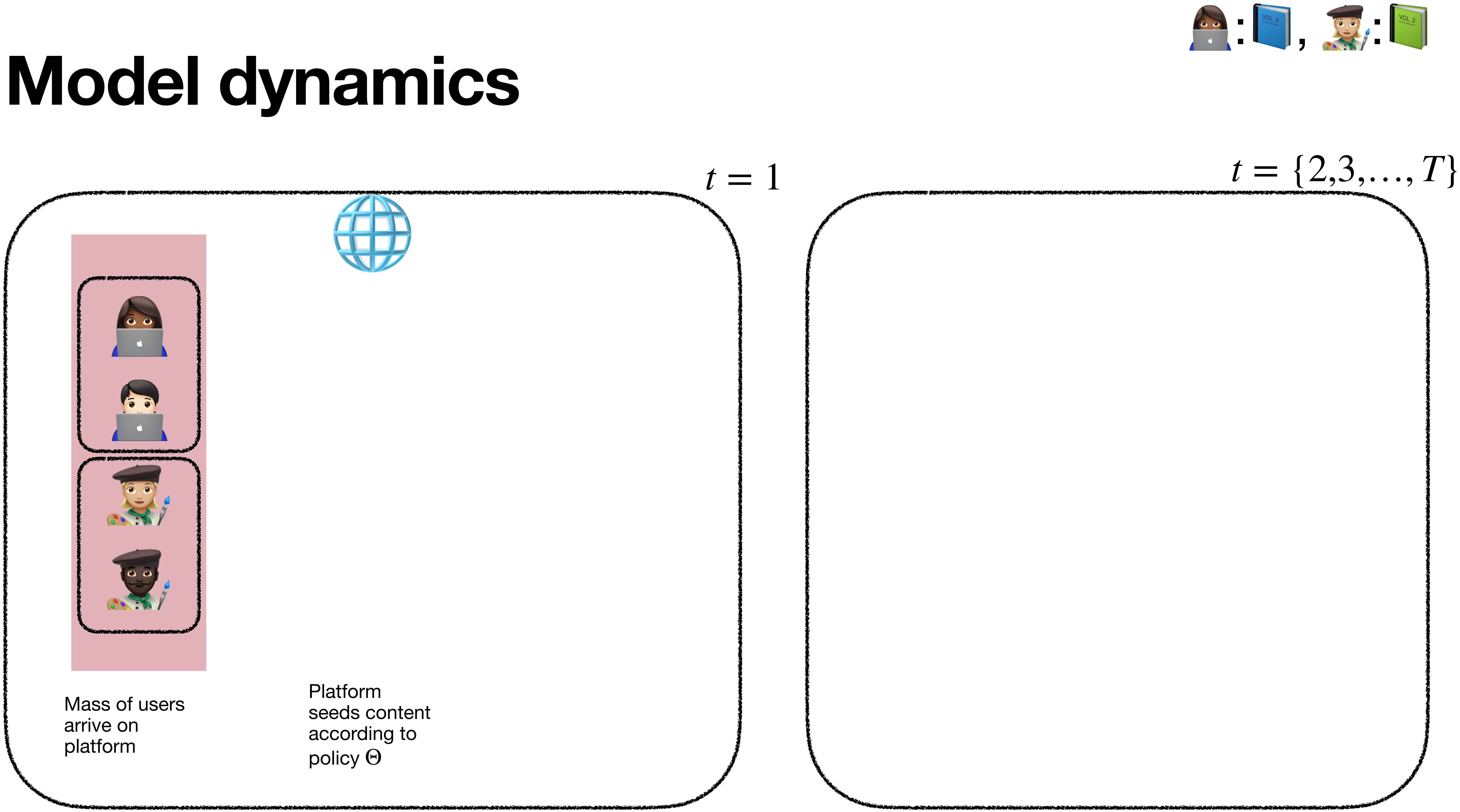
$t = 1$



Mass of users
arrive on
platform

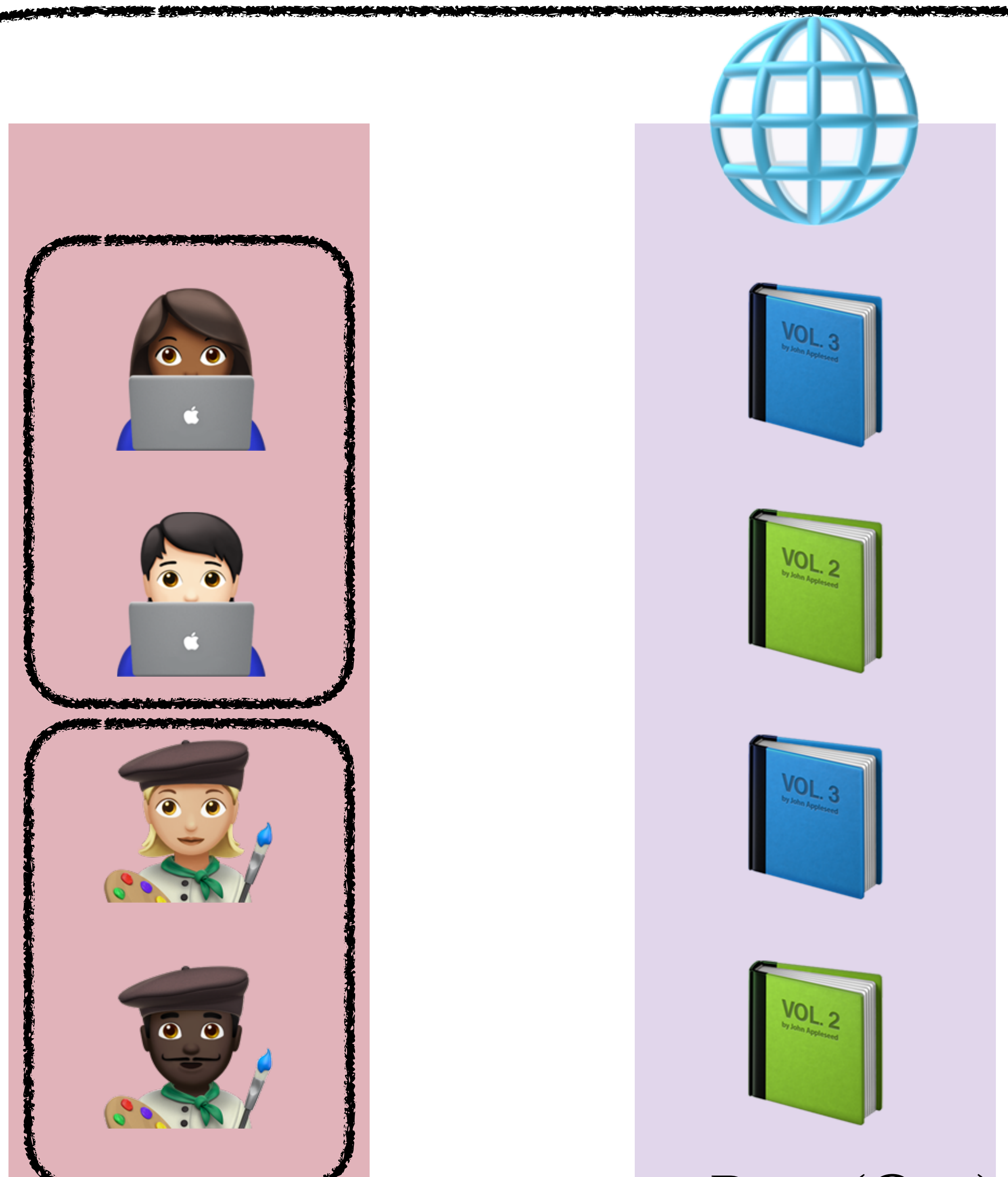
$t = \{2,3,\ldots,T\}$

Model dynamics



Model dynamics

$t = 1$



$\sim Bern(\Theta_{g,s})$

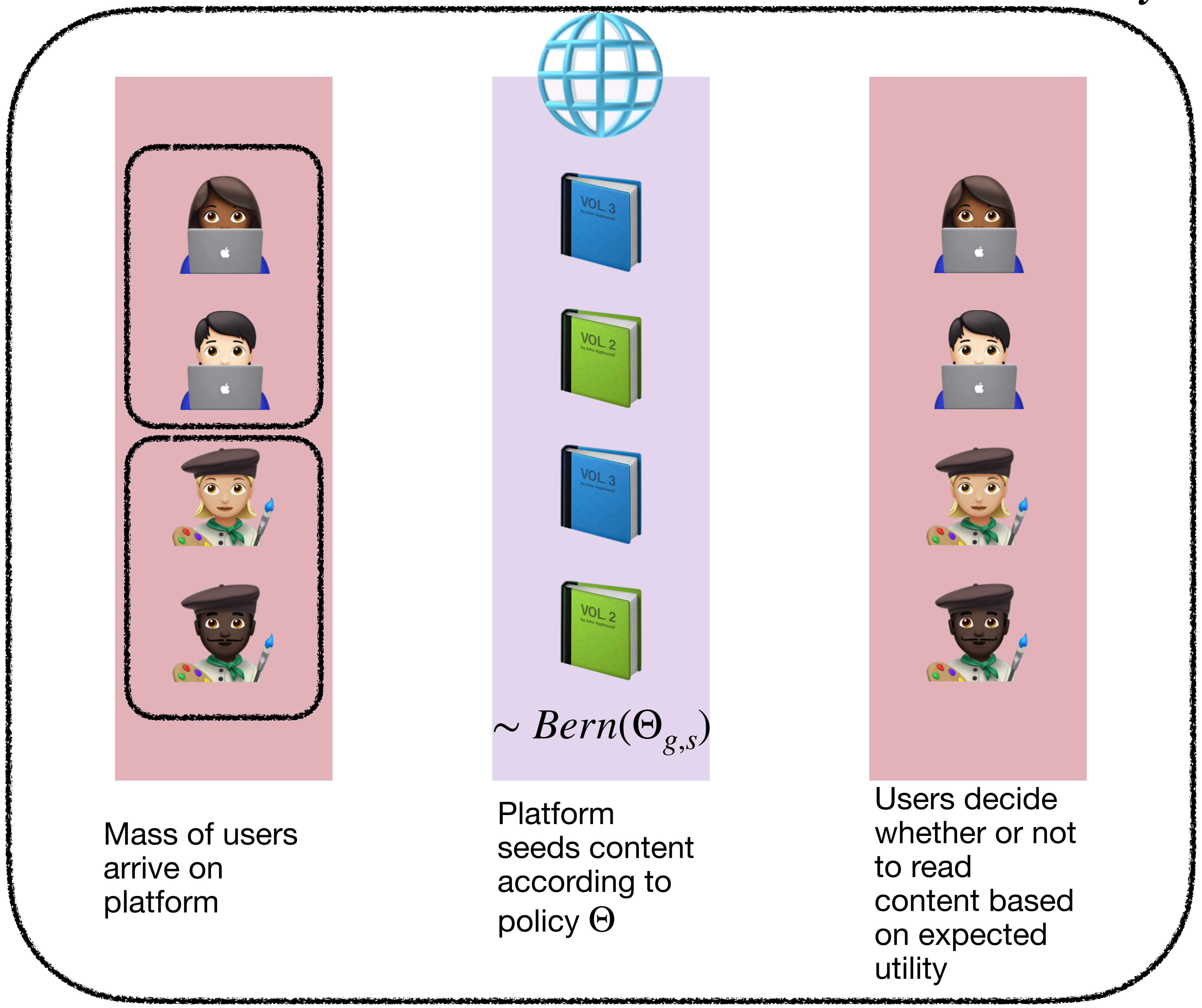
Mass of users arrive on platform

Platform seeds content according to policy Θ

$t = \{2,3,\ldots,T\}$

Model dynamics

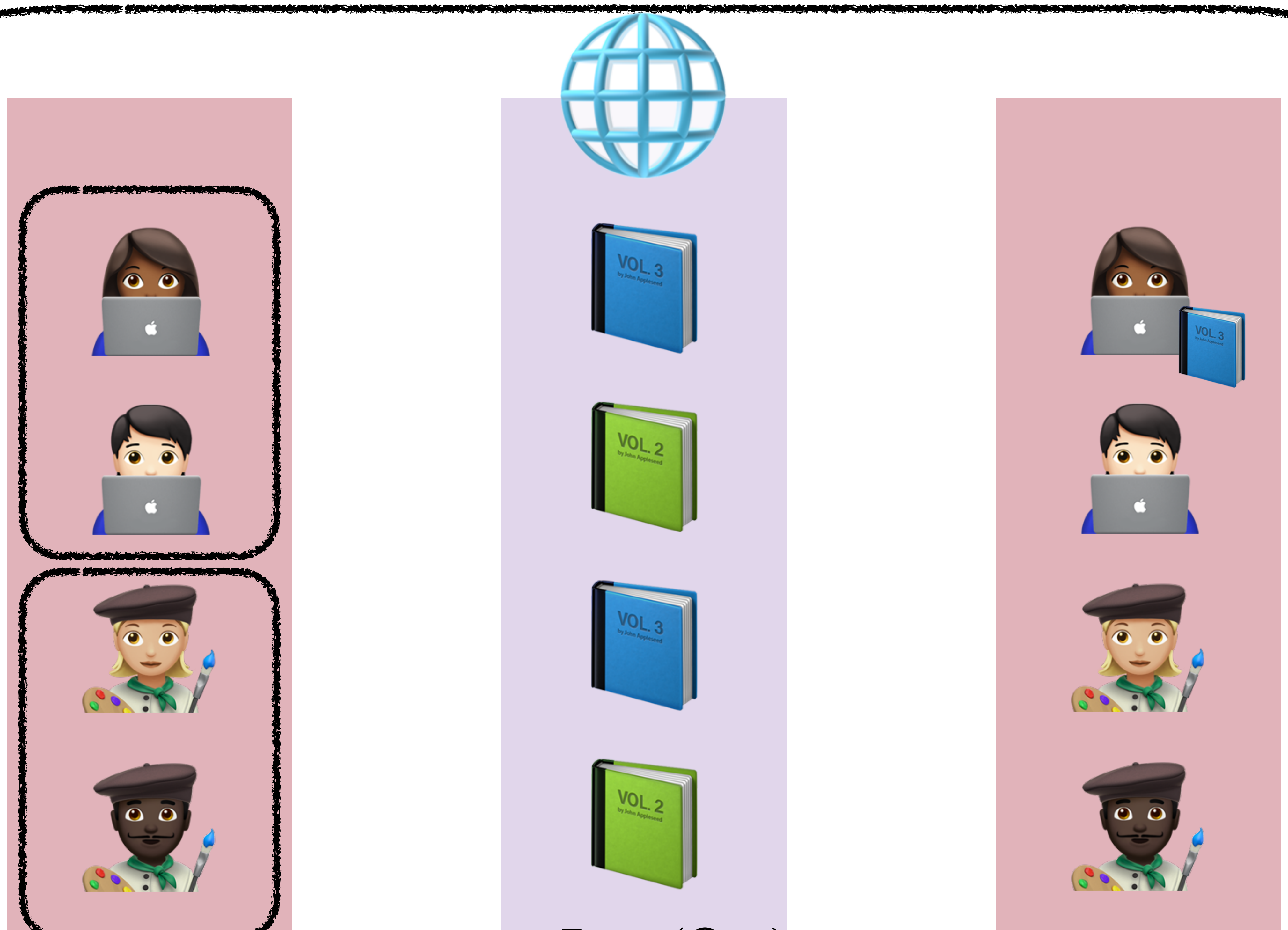
$t = 1$



$t = \{2, 3, \ldots, T\}$

Model dynamics

$t = 1$



$\sim Bern(\Theta_{g,s})$

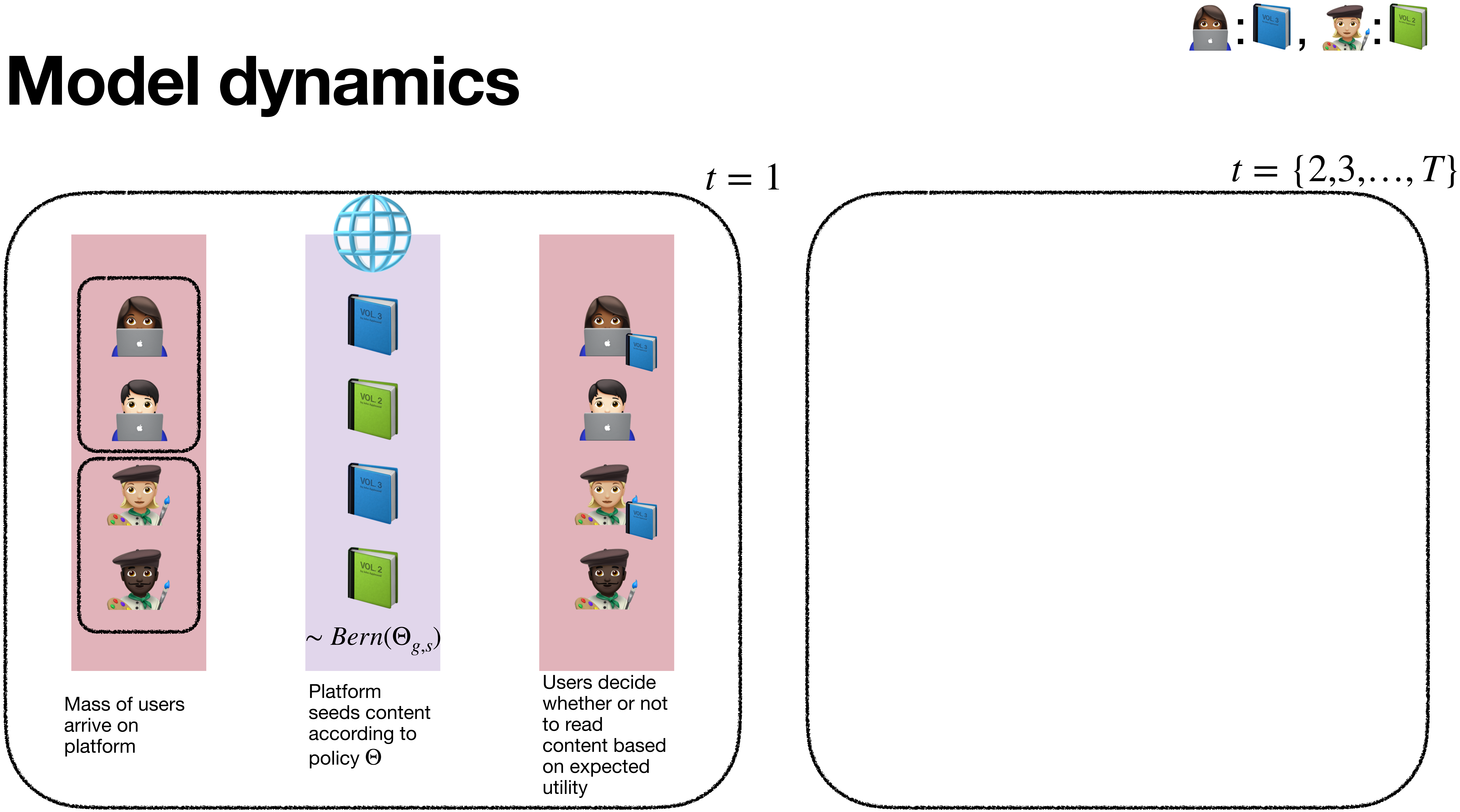
Mass of users arrive on platform

Platform seeds content according to policy Θ

Users decide whether or not to read content based on expected utility

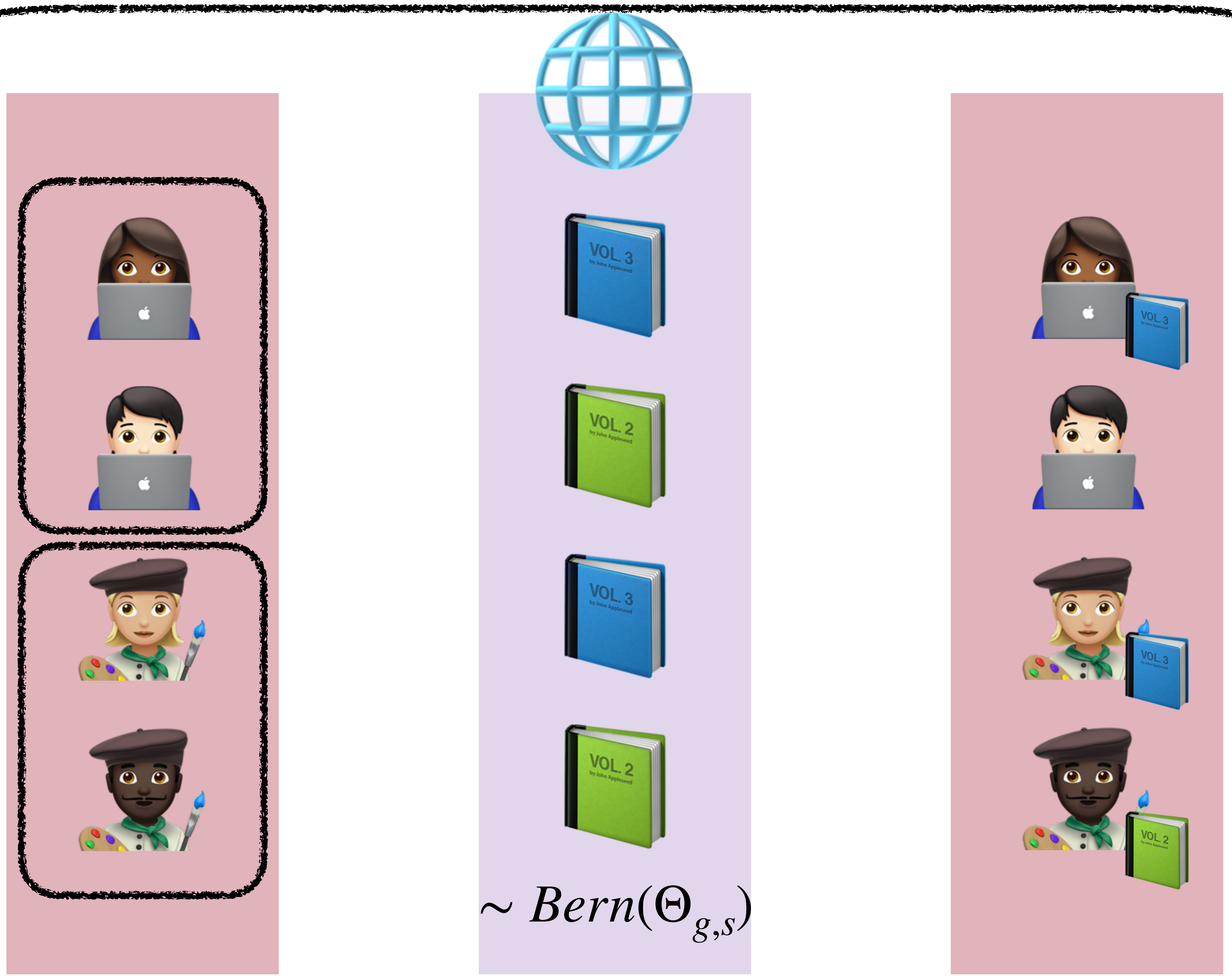
$t = \{2,3,\ldots,T\}$

Model dynamics



Model dynamics

$t = 1$



$\sim Bern(\Theta_{g,s})$

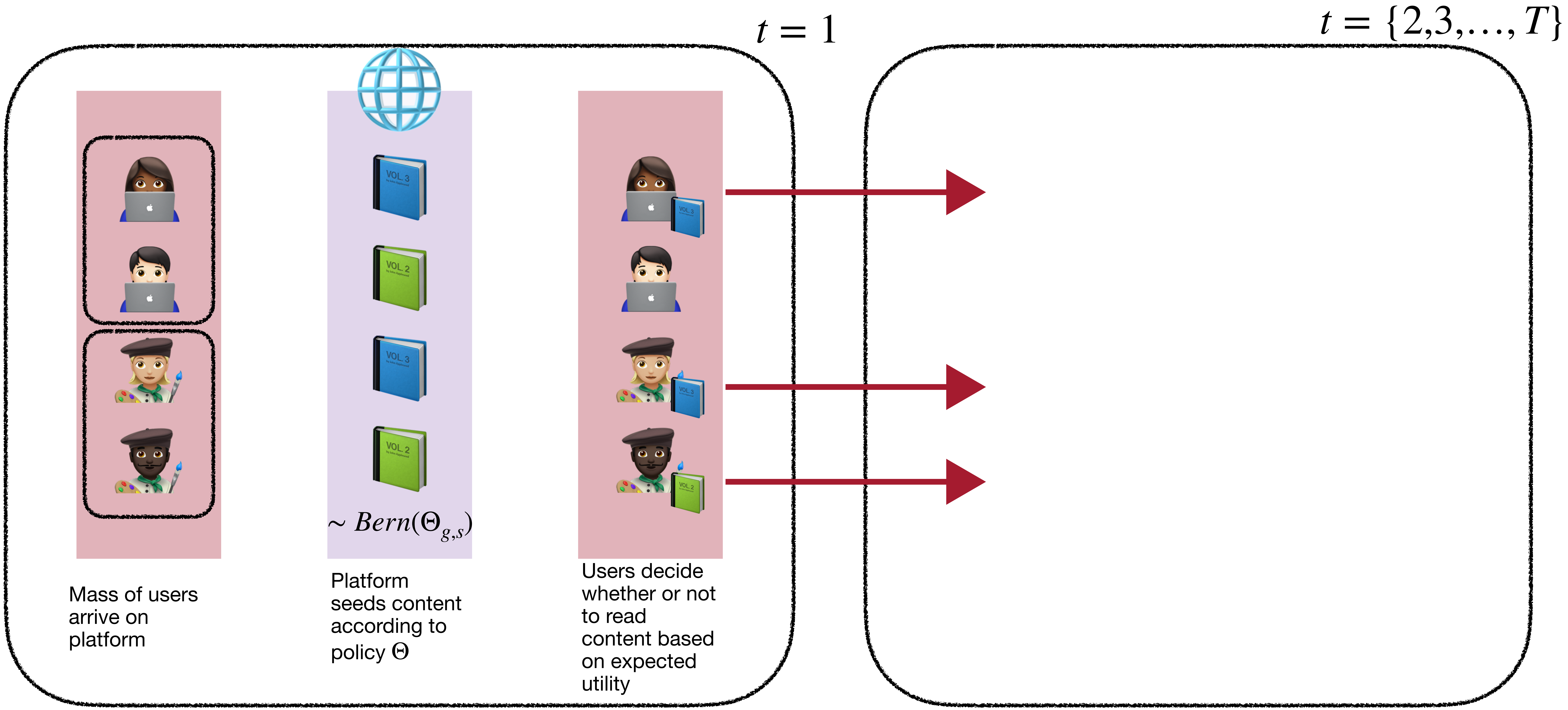
Mass of users arrive on platform

Platform seeds content according to policy Θ

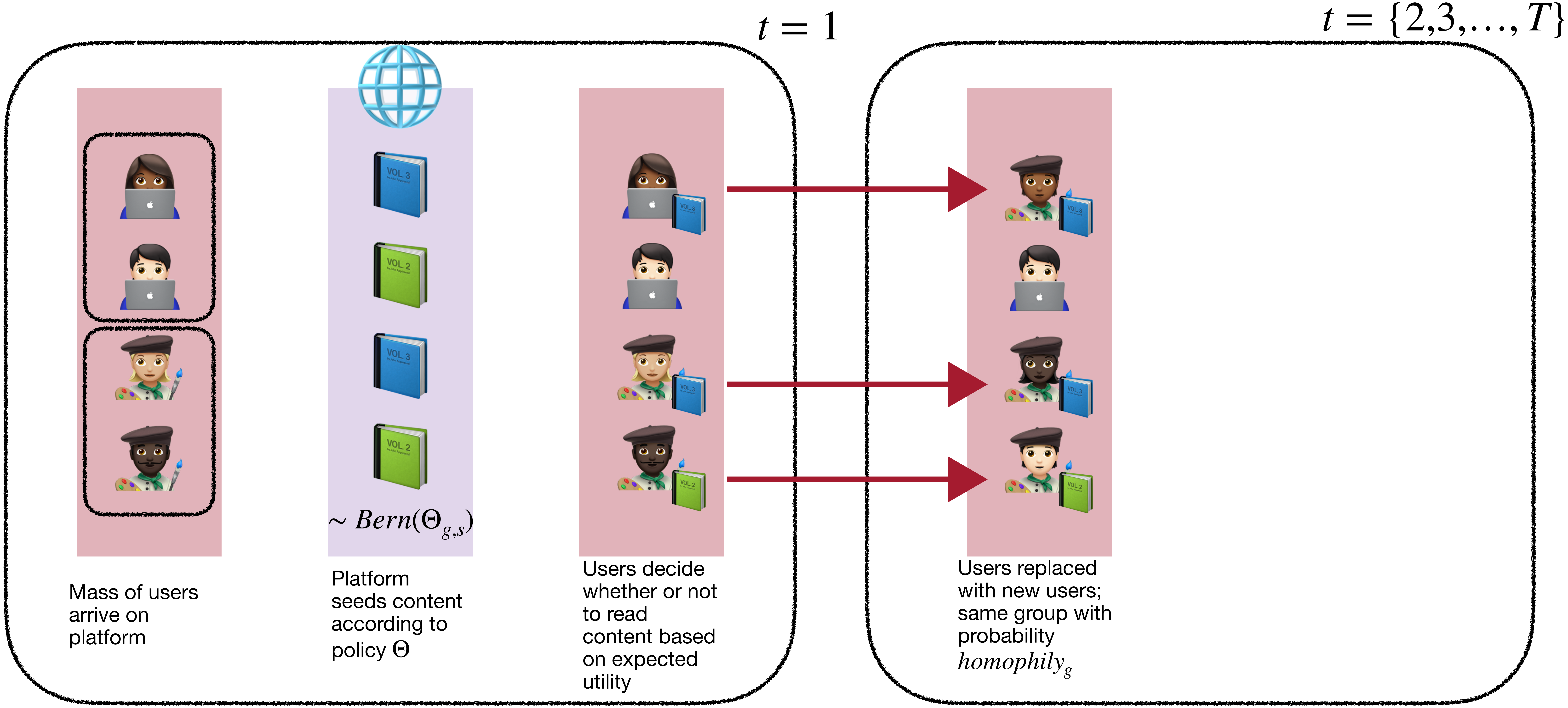
Users decide whether or not to read content based on expected utility

$t = \{2,3,\ldots,T\}$

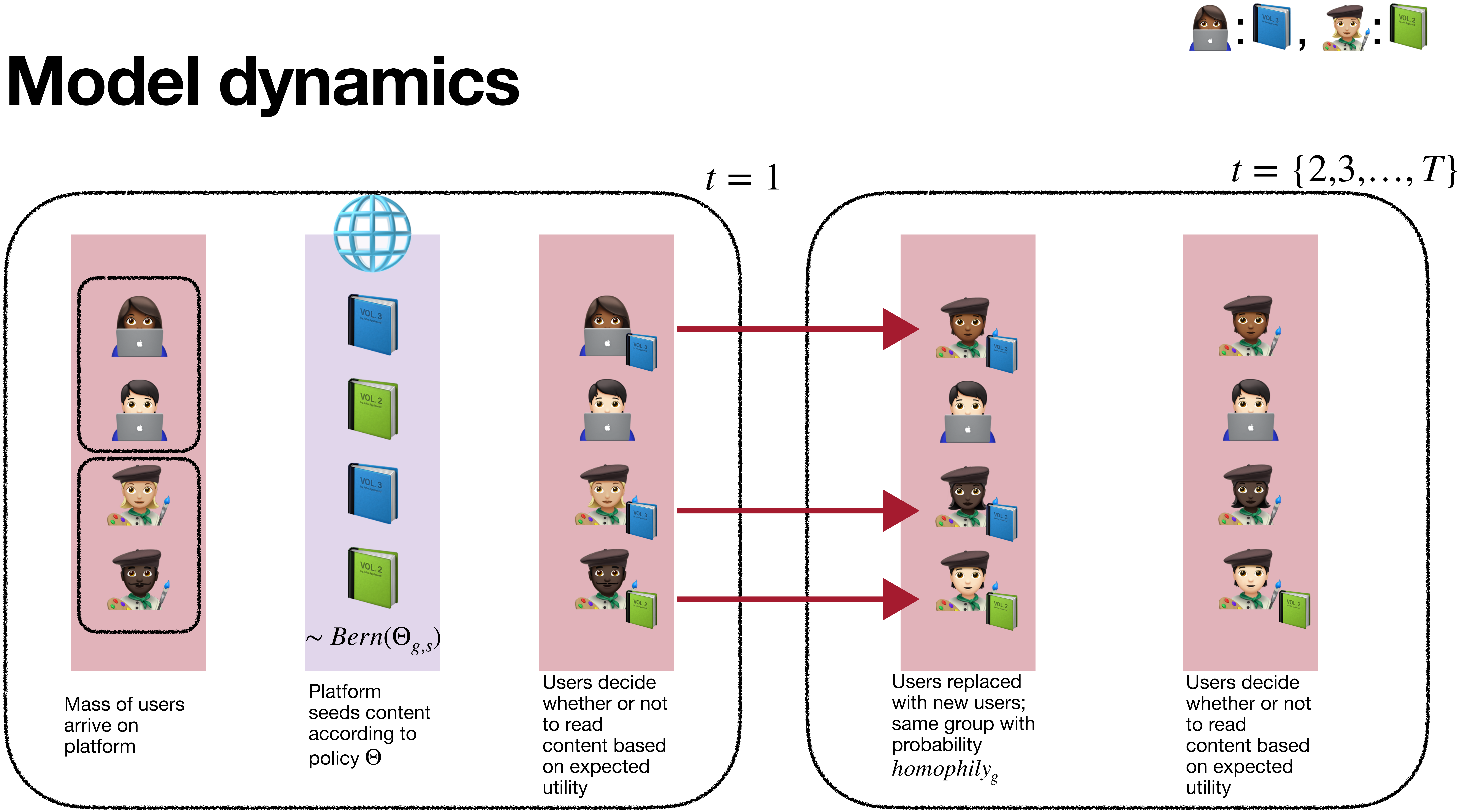
Model dynamics



Model dynamics



Model dynamics

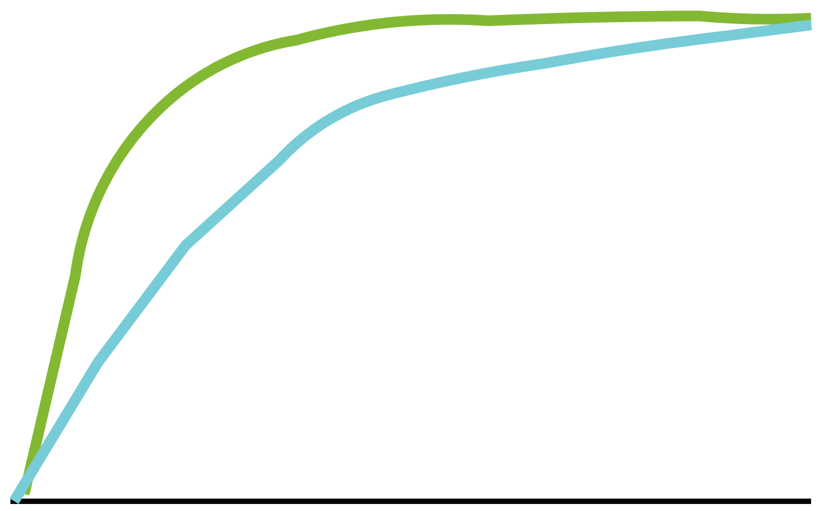


👩🏾‍💻:📘, 👩🏼‍🎨:📗

Player's decision process

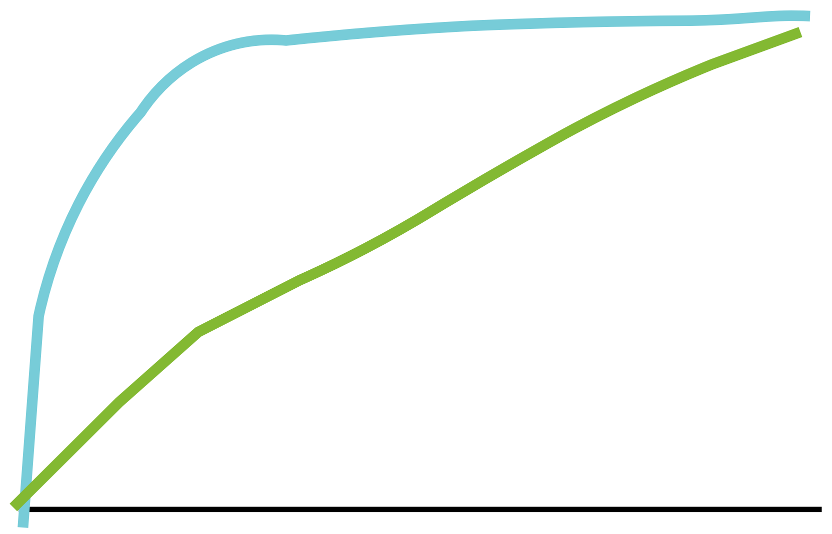
Player’s decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



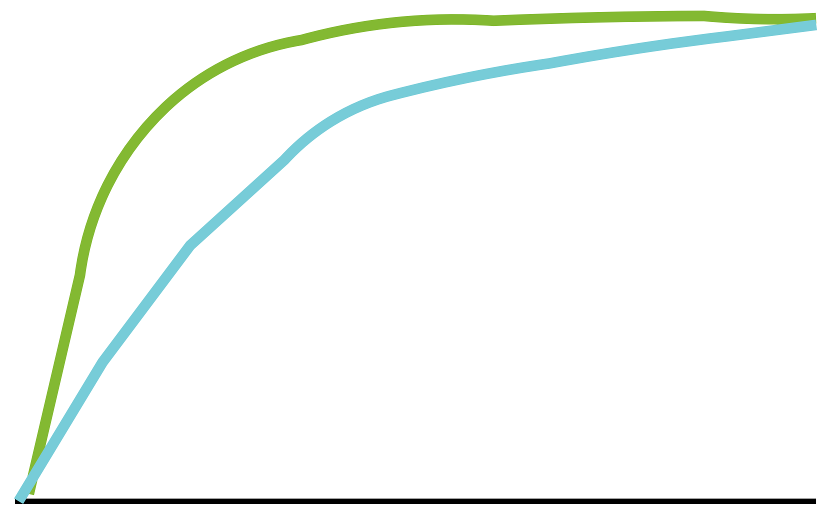
$p_{art,s} \sim F_{art,s}$

👩🏾‍💻:📘, 👩🏼‍🎨:📗

Player’s decision process

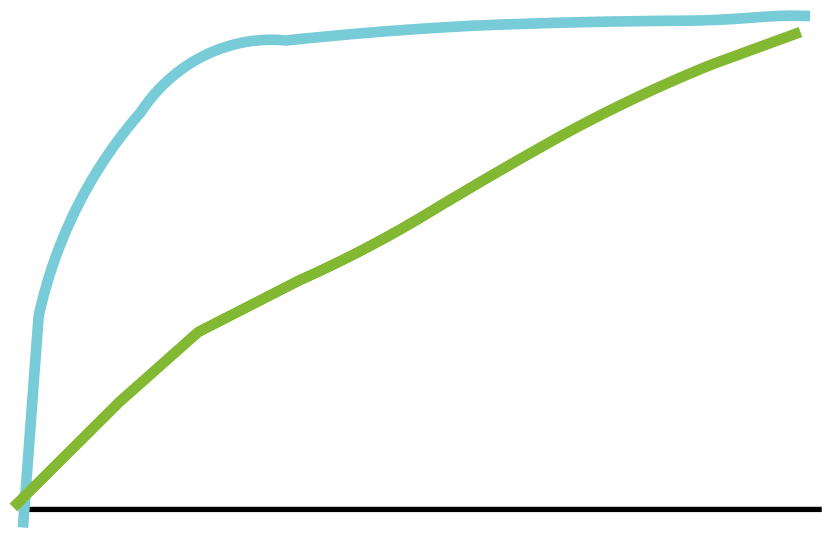
CDF[probability liking|engineer] $F_{eng,s}$

🧑🏿‍💻



$p_{eng,s} \sim F_{eng,s}$

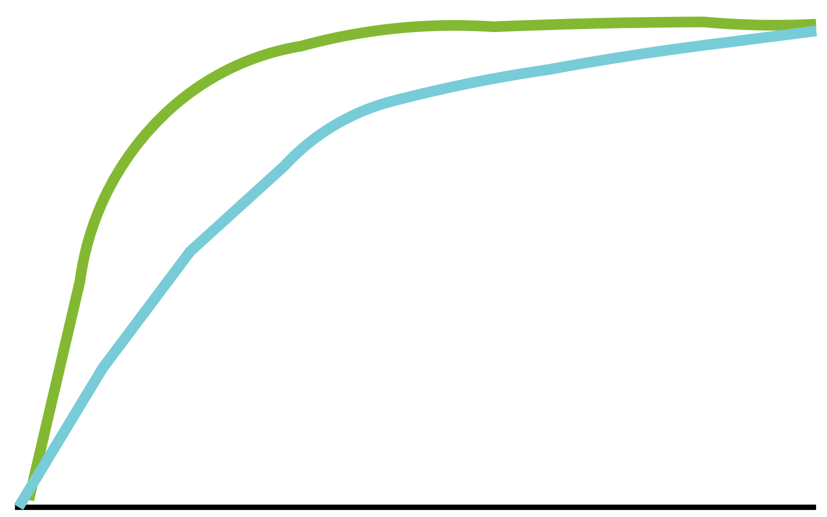
CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$

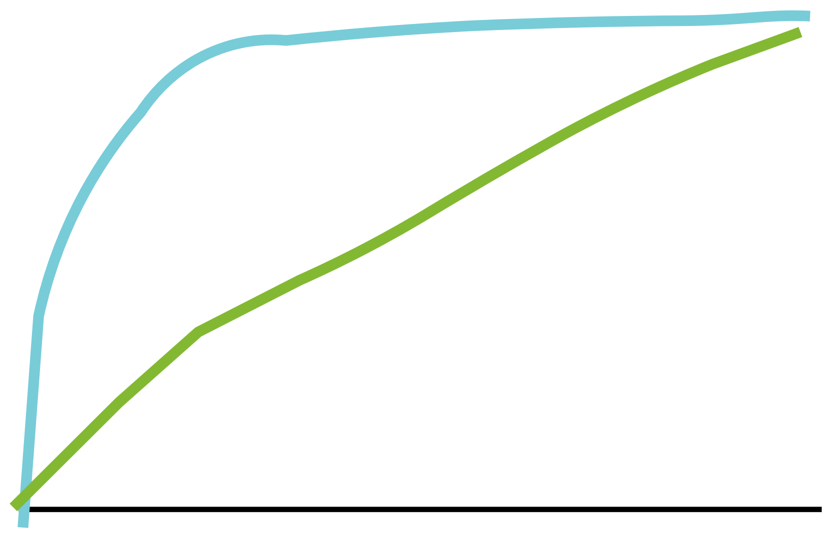


$p_{eng,s} \sim F_{eng,s}$

$$p = \frac{2}{3}$$

$$p = \frac{1}{2}$$

CDF[probability liking|artist] $F_{art,s}$

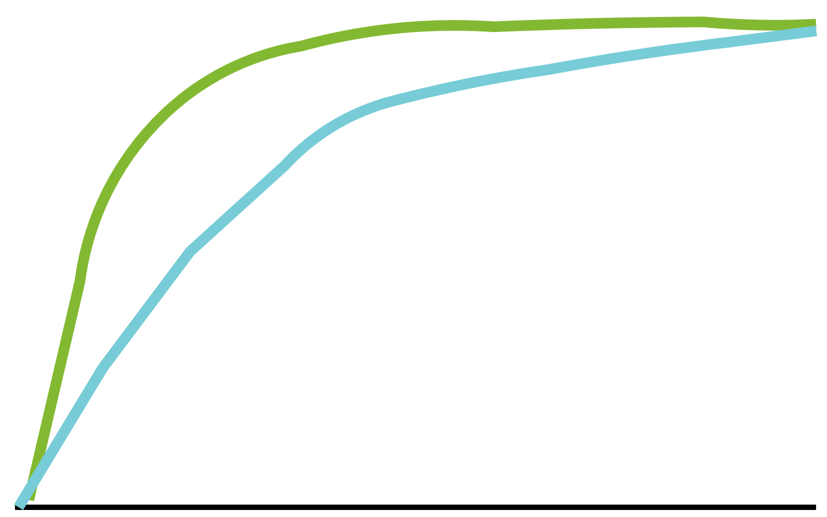


$p_{art,s} \sim F_{art,s}$

👩🏾‍💻:📘, 🧑🏼‍🎨:📗

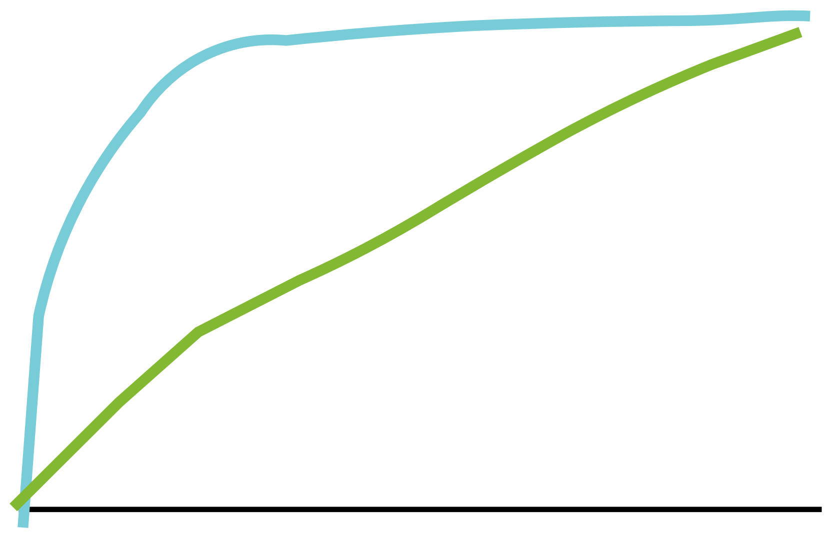
Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

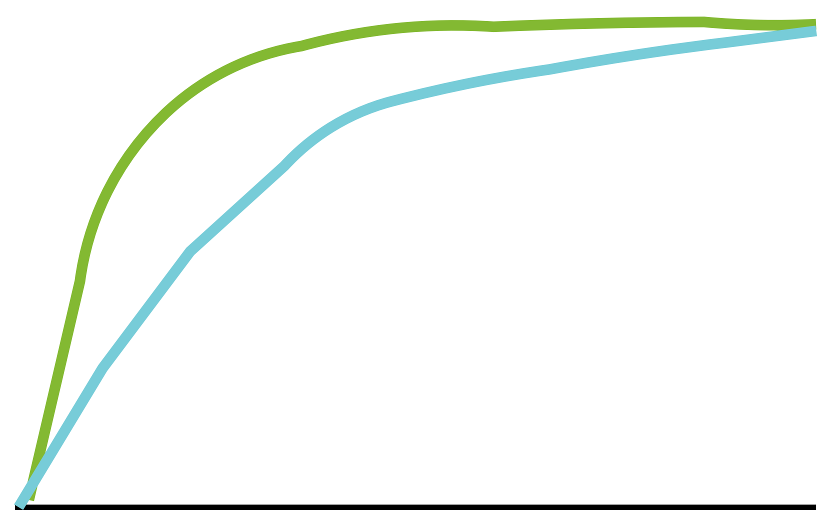
🧑🏿‍💻

	Value	Cost
$p_{\text{📘}} = \frac{2}{3}$	$v_{\text{📘}} = \frac{1}{10}$	$c_{\text{📘}} = \frac{1}{20}$
$p_{\text{📗}} = \frac{1}{2}$	$v_{\text{📗}} = \frac{1}{10}$	$c_{\text{📗}} = \frac{1}{10}$

👩🏾‍💻:📘, 👩🏼‍🎨:📗

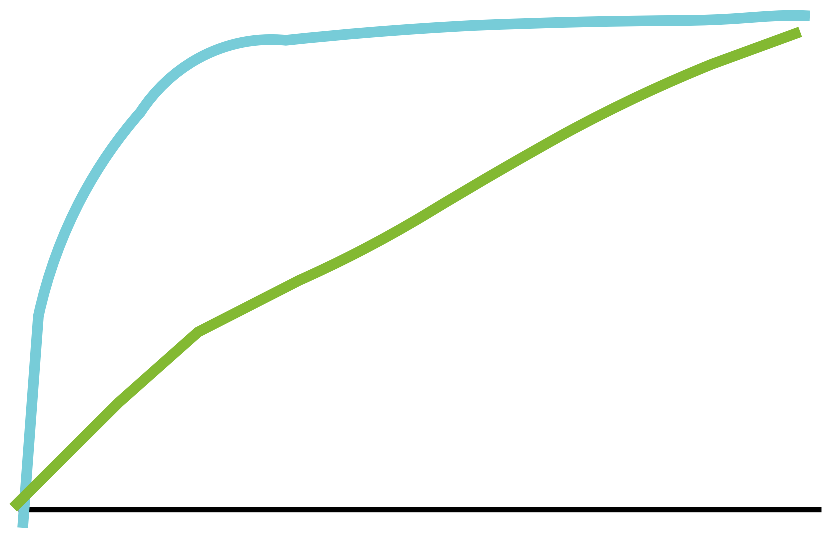
Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

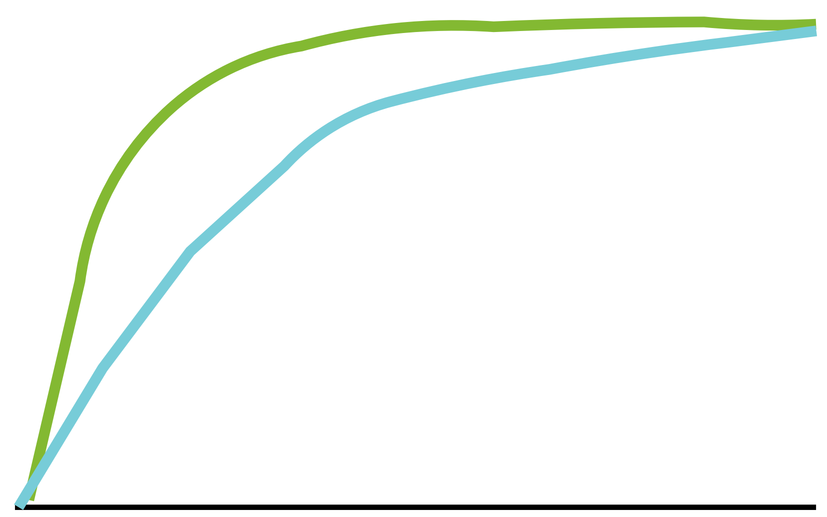
🧑🏿‍💻

	Value	Cost	Expected utility
$p_{\text{📘}} = \frac{2}{3}$	$v_{\text{📘}} = \frac{1}{10}$	$c_{\text{📘}} = \frac{1}{20}$	
$p_{\text{📗}} = \frac{1}{2}$	$v_{\text{📗}} = \frac{1}{10}$	$c_{\text{📗}} = \frac{1}{10}$	

👩🏾‍💻:📘, 🧑🏼‍🎨:📗

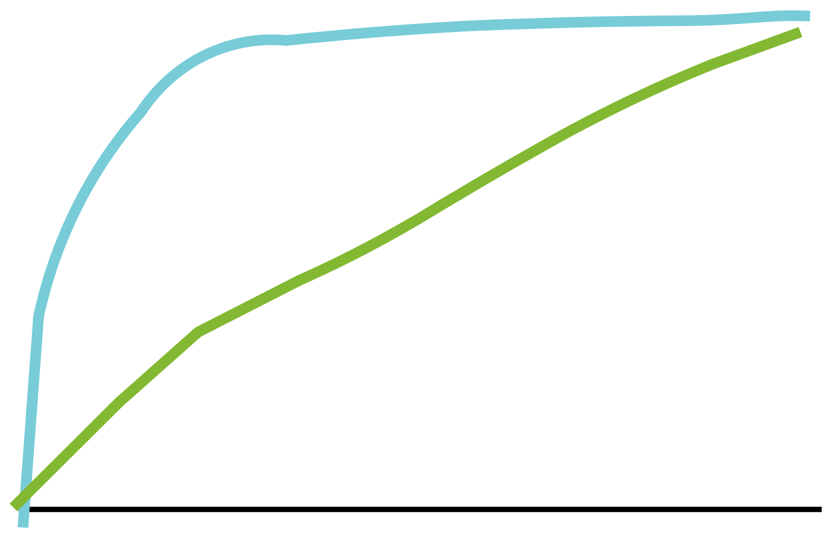
Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

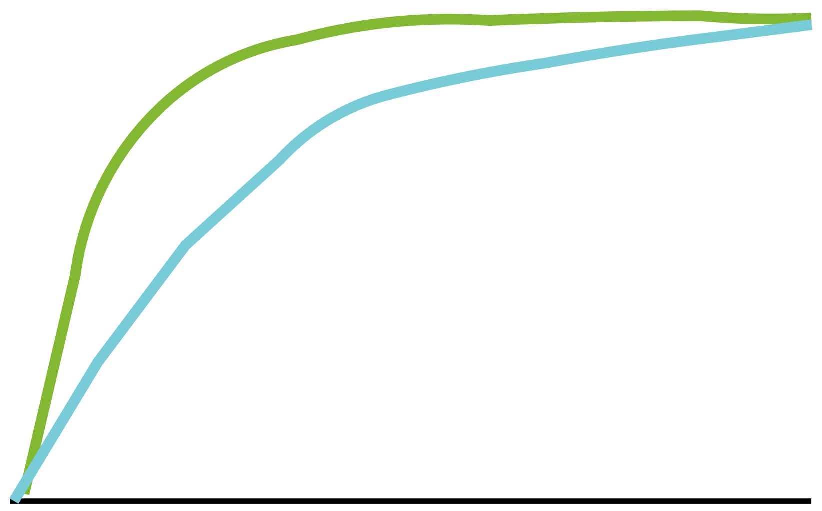
🧑🏿‍💻

	Value	Cost	Expected utility
$p_{📘} = \frac{2}{3}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$\frac{1}{60}$
$p_{📗} = \frac{1}{2}$	$v_{📗} = \frac{1}{10}$	$c_{📗} = \frac{1}{10}$	$-\frac{1}{20}$

👩🏾‍💻:📘, 👩🏼‍🎨:📗

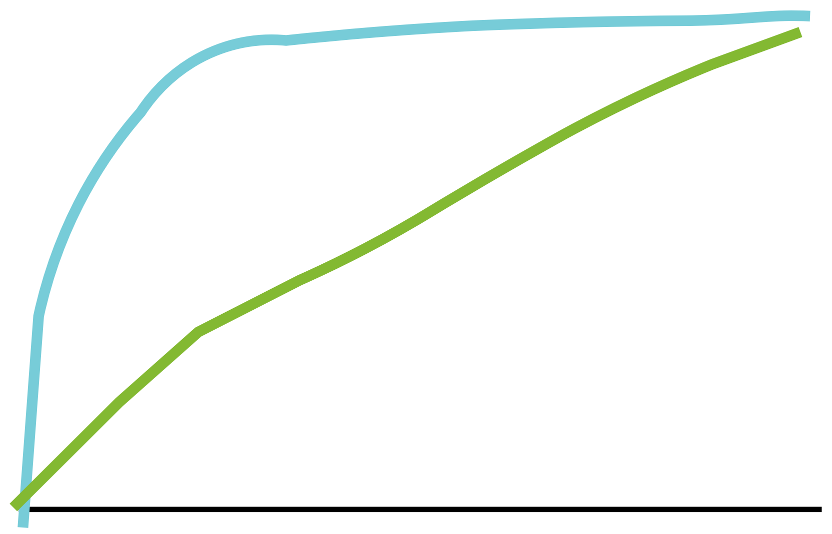
Player’s decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

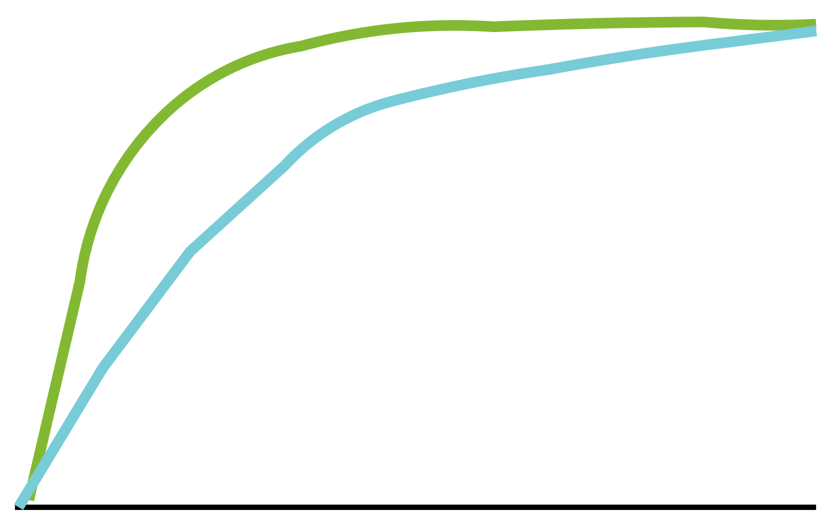
🧑🏿‍💻

	Value	Cost	Expected utility	
$p_{📘} = \frac{2}{3}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$\frac{1}{60}$	✅
$p_{📗} = \frac{1}{2}$	$v_{📗} = \frac{1}{10}$	$c_{📗} = \frac{1}{10}$	$-\frac{1}{20}$	

👩🏾‍💻:📘, 👩🏼‍🎨:📗

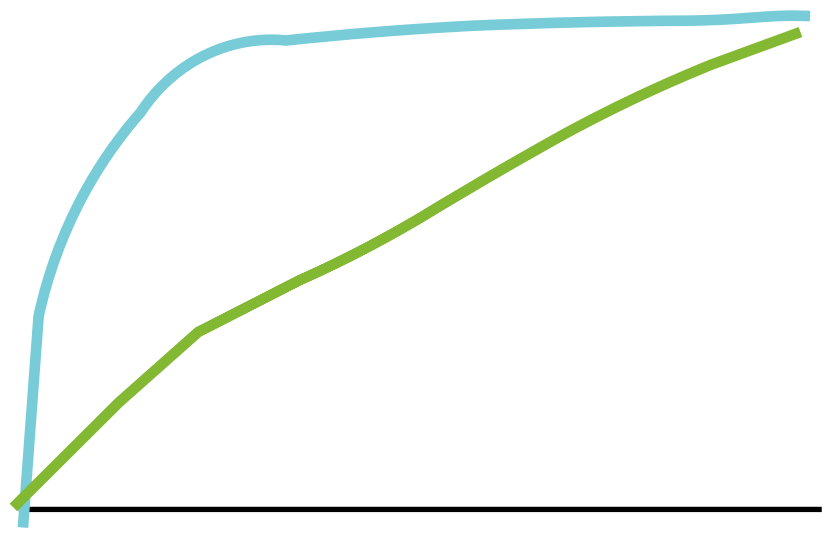
Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



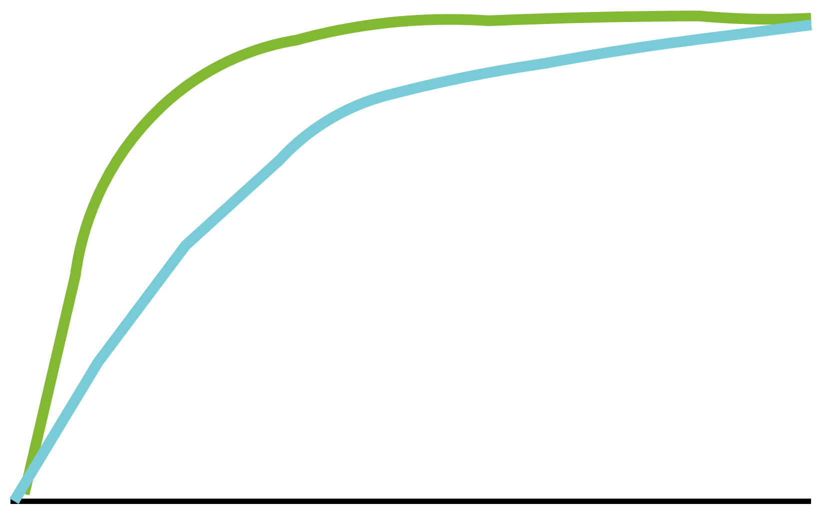
$p_{art,s} \sim F_{art,s}$

		Value	Cost	Expected utility	
🧑🏿‍💻	$p_{\text{📘}} = \frac{2}{3}$	$v_{\text{📘}} = \frac{1}{10}$	$c_{\text{📘}} = \frac{1}{20}$	$\frac{1}{60}$	✅
	$p_{\text{📗}} = \frac{1}{2}$	$v_{\text{📗}} = \frac{1}{10}$	$c_{\text{📗}} = \frac{1}{10}$	$-\frac{1}{20}$	❌

👩🏾‍💻:📘, 🧑🏼‍🎨:📗

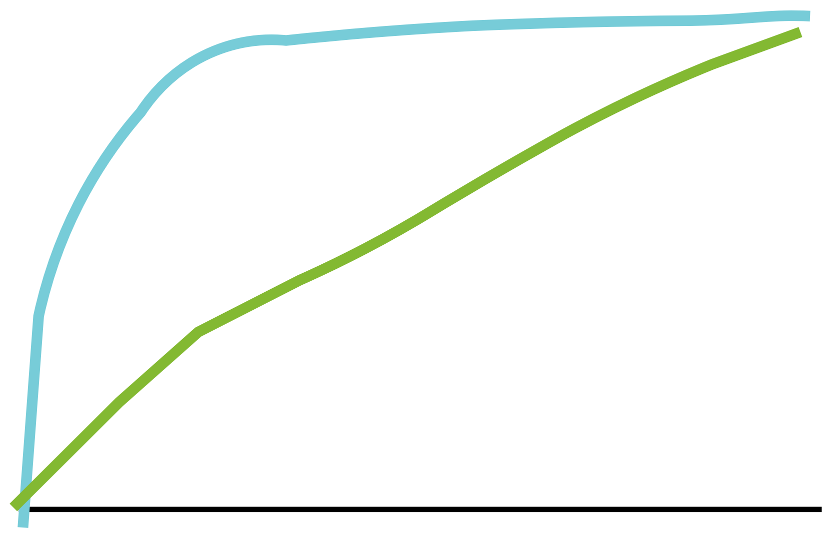
Player's decision process

CDF[probability liking|engineer] $F_{eng,s}$



$p_{eng,s} \sim F_{eng,s}$

CDF[probability liking|artist] $F_{art,s}$



$p_{art,s} \sim F_{art,s}$

		Value	Cost	Expected utility	
🧑🏿‍💻	$p_{📘} = \frac{2}{3}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$\frac{1}{60}$	✅
	$p_{📗} = \frac{1}{2}$	$v_{📗} = \frac{1}{10}$	$c_{📗} = \frac{1}{10}$	$-\frac{1}{20}$	❌
👩🏻‍💻	$p_{📘} = \frac{3}{4}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$\frac{1}{40}$	✅
	$p_{📗} = \frac{1}{3}$	$v_{📗} = \frac{1}{10}$	$c_{📗} = \frac{1}{10}$	$-\frac{1}{15}$	❌
👨🏾‍🎨	$p_{📘} = \frac{2}{3}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$\frac{1}{60}$	✅
	$p_{📗} = \frac{3}{4}$	$v_{📗} = \frac{1}{3}$	$c_{📗} = \frac{1}{40}$	$\frac{9}{40}$	✅
👨🏽‍🎨	$p_{📘} = \frac{1}{4}$	$v_{📘} = \frac{1}{10}$	$c_{📘} = \frac{1}{20}$	$-\frac{1}{40}$	❌
	$p_{📗} = \frac{5}{6}$	$v_{📗} = \frac{1}{3}$	$c_{📗} = \frac{1}{40}$	$\frac{91}{360}$	✅

Platform’s Optimization Problem

Platform’s Optimization Problem

Maximize expected number of clicks at time t

Platform's Optimization Problem

Maximize expected number of clicks at time t

$$\max_{\Theta} \qquad\qquad\qquad \mathit{total\_clicks}\;(\;\; ,\Theta)$$

Platform's Optimization Problem

Maximize expected number of clicks at time t

over time, and across groups, and sources

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{a,b\}} \sum_{s \in \{A,B\}} total\_clicks_{gs}(t, \Theta)$$

Platform's Optimization Problem

Maximize expected number of clicks at time t

over time, and across groups, and sources

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{a,b\}} \sum_{s \in \{A,B\}} total\_clicks_{gs}(t,\Theta)$$

$total\_clicks_{g,s}(t+1,\Theta) = ind\_clicks_{g,s} \times$ expected number of users in group g shown content s at time t

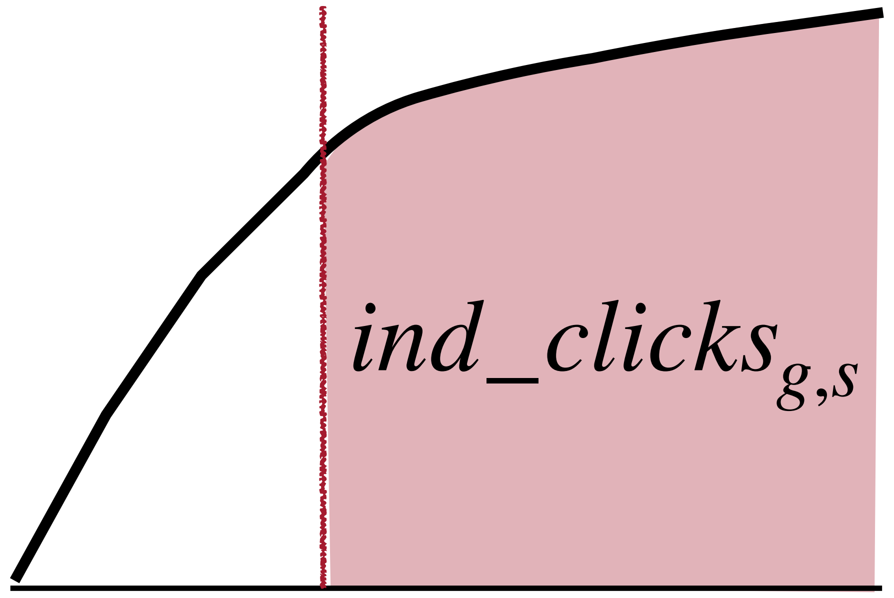
Platform's Optimization Problem

Maximize expected number of clicks at time t

over time, and across groups, and sources

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{a,b\}} \sum_{s \in \{A,B\}} total\_clicks_{gs}(t, \Theta)$$

$cost_{g,s}/value_{g,s}$



Prob. Liking content

$total\_clicks_{g,s}(t+1, \Theta) = ind\_clicks_{g,s} \times$ expected number of users in group g shown content s at time t

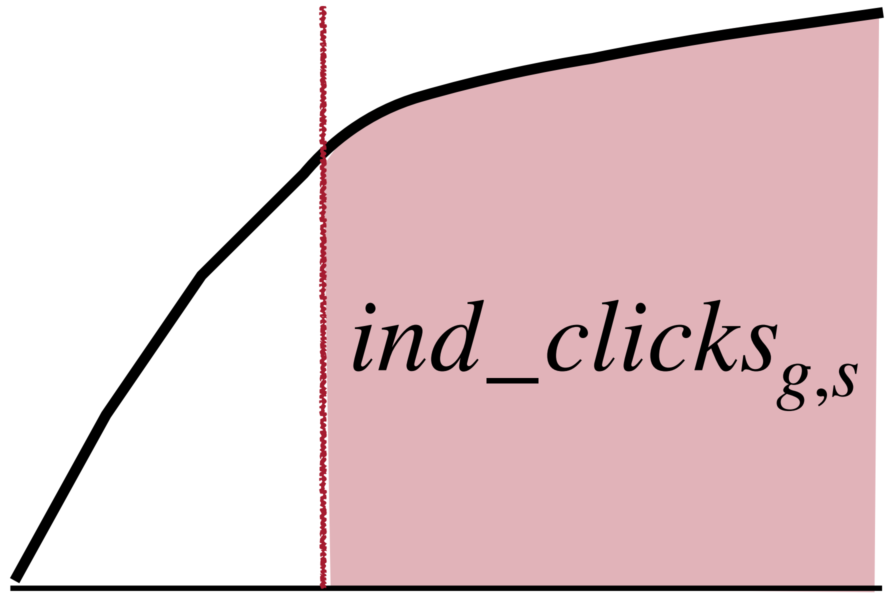
Platform's Optimization Problem

Maximize expected number of clicks at time t

over time, and across groups, and sources

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{a,b\}} \sum_{s \in \{A,B\}} total\_clicks_{gs}(t, \Theta)$$

$cost_{g,s}/value_{g,s}$



Prob. Liking content

$total\_clicks_{g,s}(t+1,\Theta) = ind\_clicks_{g,s} \times$ expected number of users in group g shown content s at time t

Proposition (informal): The exclusion of any fairness constraints in the platform's optimization problem ➡️ group homogeneous targeting

Possible notions of fair exposure

Possible notions of fair exposure

Recall $total\_clicks_{g,s}(t, \Theta)$: expected number of users in g at time t who click content s

Possible notions of fair exposure

Recall $total\_clicks_{g,s}(t, \Theta)$: expected number of users in g at time t who click content s

Constant fair exposure: the rate of exposure of users who preferred content s is constant and at level $e \in [0,1]$ at each t and across groups

Possible notions of fair exposure

Recall $total\_clicks_{g,s}(t, \Theta)$: expected number of users in g at time t who click content s

Constant fair exposure: the rate of exposure of users who preferred content s is constant and at level $e \in [0,1]$ at each t and across groups

$$\frac{total\_clicks_{b,s}(t, \Theta)}{\text{frac. population in group } a} \approx \frac{total\_clicks_{b,s'}(t, \Theta)}{\text{frac. population in group } b} \approx e \qquad \forall t \leq T, \forall s \neq s'$$

Possible notions of fair exposure

Recall $total\_clicks_{g,s}(t, \Theta)$: expected number of users in g at time t who click content s

Constant fair exposure: the rate of exposure of users who preferred content s is constant and at level $e \in [0,1]$ at each t and across groups

$$\frac{total\_clicks_{b,s}(t, \Theta)}{\text{frac. population in group } a} \approx \frac{total\_clicks_{b,s'}(t, \Theta)}{\text{frac. population in group } b} \approx e \qquad \forall t \leq T, \forall s \neq s'$$

Proposition (informal): Constant fair exposure is generally not feasible

Relaxing fair exposure

Relaxing fair exposure

Approximate fair average exposure: total exposure of users to their preferred and non preferred content is approximately equal across groups. Given $\underline{\delta} < 1 < \overline{\delta}$

Relaxing fair exposure

Approximate fair average exposure: total exposure of users to their preferred and non preferred content is approximately equal across groups. Given $\underline{\delta} < 1 < \overline{\delta}$

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,a}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,b}(t, \Theta)} \leq \overline{\delta}$$

Clicks on preferred content
approximately equal

Relaxing fair exposure

Approximate fair average exposure: total exposure of users to their preferred and non preferred content is approximately equal across groups. Given $\underline{\delta} < 1 < \overline{\delta}$

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,a}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,b}(t, \Theta)} \leq \overline{\delta}$$

Clicks on preferred content approximately equal

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,b}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,a}(t, \Theta)} \leq \overline{\delta}$$

Clicks on non-preferred content approximately equal

Platform's optimization problem

With approximately fair average exposure constraints

Platform’s optimization problem

With approximately fair average exposure constraints

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g\in\{A,B\}} \sum_{s\in\{a,b\}} total\_clicks_{g,s}(t, \Theta)$$

Maximize expected number of clicks on the time horizon

Platform's optimization problem

With approximately fair average exposure constraints

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{A,B\}} \sum_{s \in \{a,b\}} total\_clicks_{g,s}(t, \Theta)$$

Maximize expected number of clicks on the time horizon

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,a}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,b}(t, \Theta)} \leq \overline{\delta}$$

Clicks on preferred content approximately equal

Platform's optimization problem

With approximately fair average exposure constraints

$$\max_{\Theta} \sum_{t=1}^{T} \sum_{g \in \{A,B\}} \sum_{s \in \{a,b\}} total\_clicks_{g,s}(t, \Theta)$$

Maximize expected number of clicks on the time horizon

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,a}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,b}(t, \Theta)} \leq \bar{\delta}$$

Clicks on preferred content approximately equal

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} total\_clicks_{A,b}(t, \Theta)}{\sum_{t=1}^{T} total\_clicks_{B,a}(t, \Theta)} \leq \bar{\delta}$$

Clicks on non-preferred content approximately equal

How does the platform optimally recommend?

How does the platform optimally recommend?

Analyzing the optimal solutions to platform's optimization problem, we see that

How does the platform optimally recommend?

Analyzing the optimal solutions to platform’s optimization problem, we see that

1. Introducing fairness constraints does not necessarily imply that the outcome is truly fair or balanced

How does the platform optimally recommend?

Analyzing the optimal solutions to platform's optimization problem, we see that

1. Introducing fairness constraints does not necessarily imply that the outcome is truly fair or balanced
2. Constraints forbid group homogeneous targeting for both groups, but there exist optimal solutions that target members of one group deterministically

“Lowering one group to the other” in practice

“Lowering one group to the other” in practice

Estimate population parameters via MLE with data available from Bakshy et al. (2015) and Garimella et al. (2017)

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*

Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *NeurIPS*

“Lowering one group to the other” in practice

Estimate population parameters via MLE with data available from Bakshy et al. (2015) and Garimella et al. (2017)

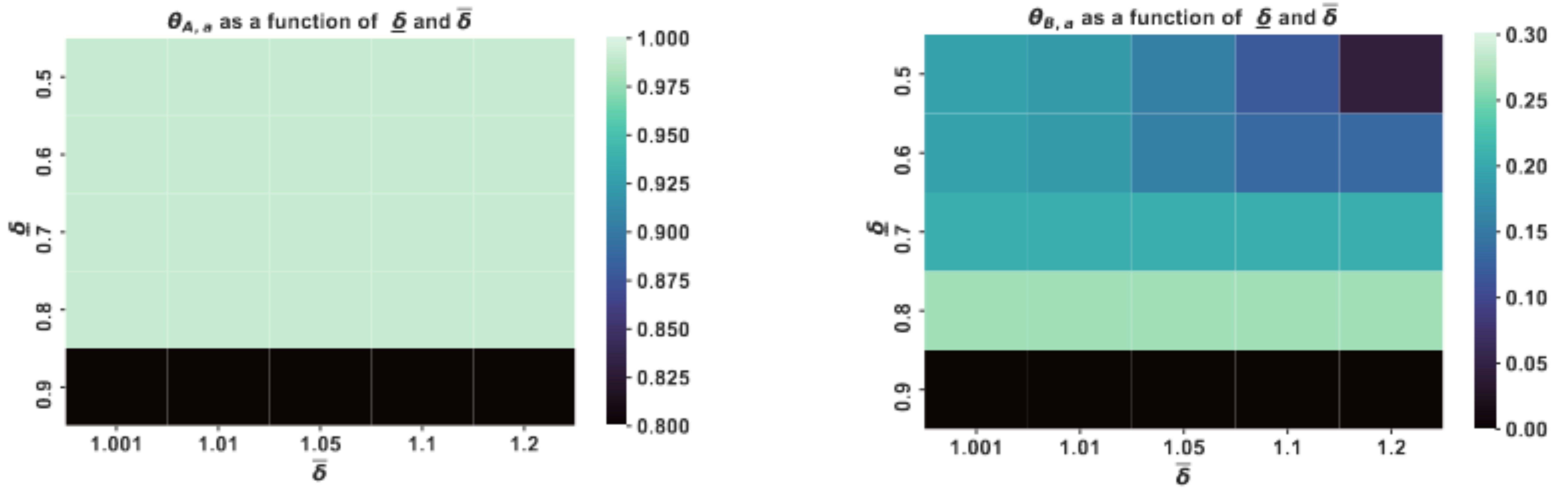


Figure 3: Calculating $\theta_{A,a}$ (left) and $\theta_{B,a}$ (right) as a function of $\underline{\delta}$ and $\overline{\delta}$ with parameters estimated from Bakshy et al. [2015]. Black cells in the bottom row indicate no feasible solution to the fairness-constrained problem.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*

Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *NeurIPS*

“Lowering one group to the other” in practice

Estimate population parameters via MLE with data available from Bakshy et al. (2015) and Garimella et al. (2017)

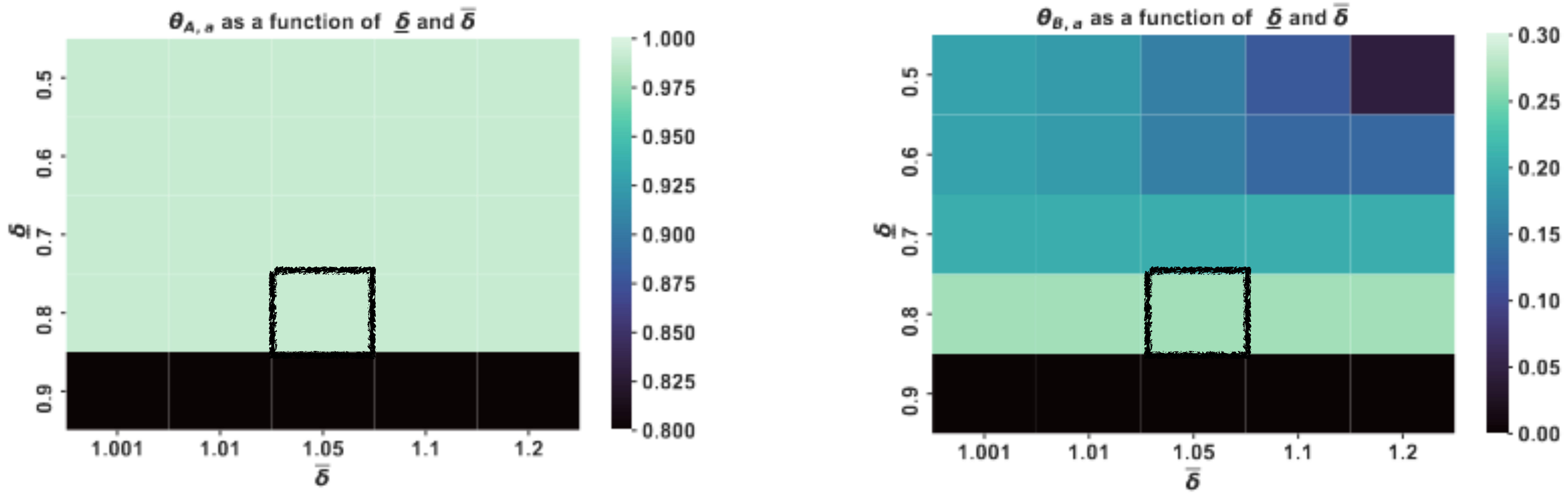


Figure 3: Calculating $\theta_{A,a}$ (left) and $\theta_{B,a}$ (right) as a function of $\underline{\delta}$ and $\overline{\delta}$ with parameters estimated from Bakshy et al. [2015]. Black cells in the bottom row indicate no feasible solution to the fairness-constrained problem.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*

Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *NeurIPS*

“Lowering one group to the other” in practice

Estimate population parameters via MLE with data available from Bakshy et al. (2015) and Garimella et al. (2017)

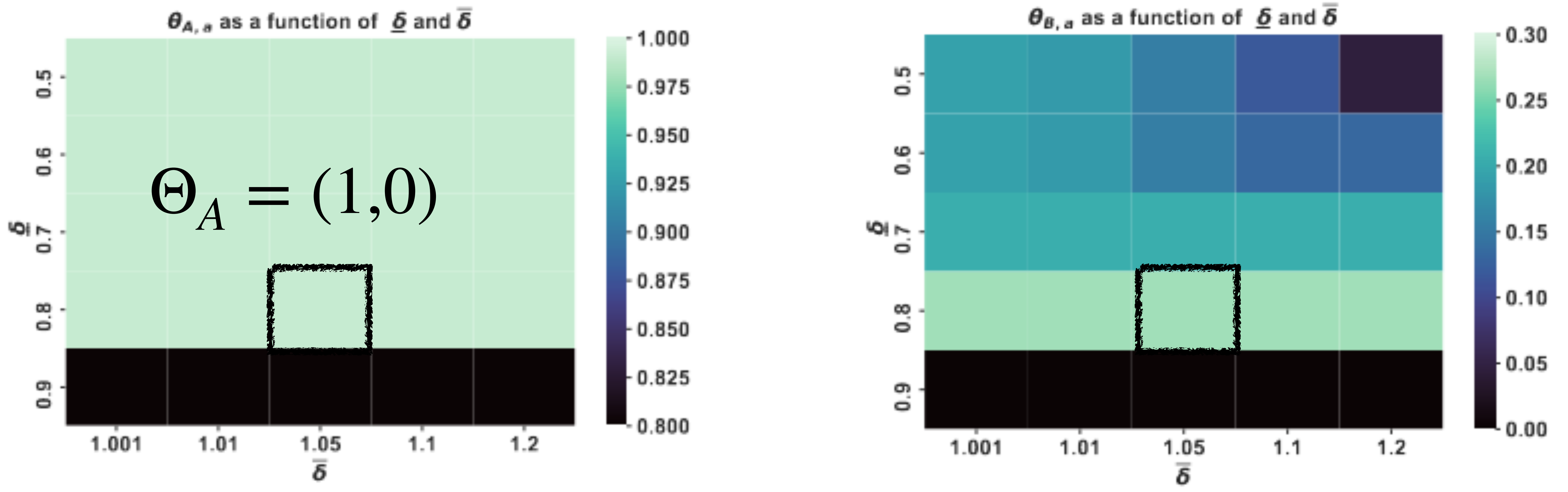


Figure 3: Calculating $\theta_{A,a}$ (left) and $\theta_{B,a}$ (right) as a function of $\underline{\delta}$ and $\overline{\delta}$ with parameters estimated from Bakshy et al. [2015]. Black cells in the bottom row indicate no feasible solution to the fairness-constrained problem.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*

Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *NeurIPS*

"Lowering one group to the other" in practice

Estimate population parameters via MLE with data available from Bakshy et al. (2015) and Garimella et al. (2017)

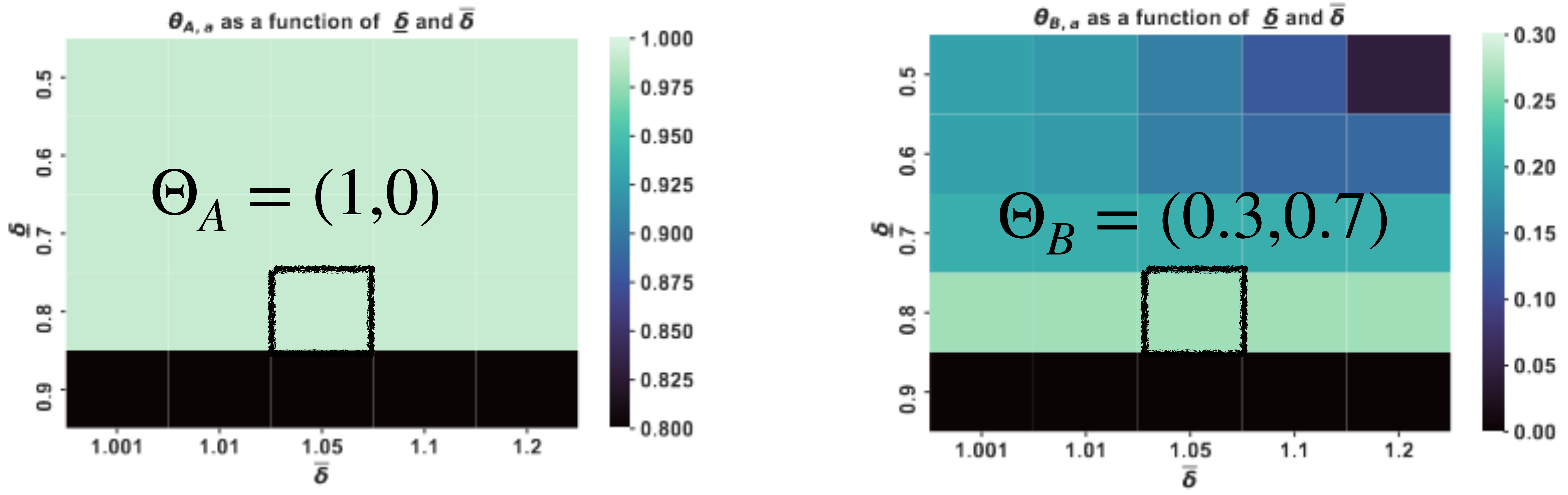


Figure 3: Calculating $\theta_{A,a}$ (left) and $\theta_{B,a}$ (right) as a function of $\underline{\delta}$ and $\overline{\delta}$ with parameters estimated from Bakshy et al. [2015]. Black cells in the bottom row indicate no feasible solution to the fairness-constrained problem.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*

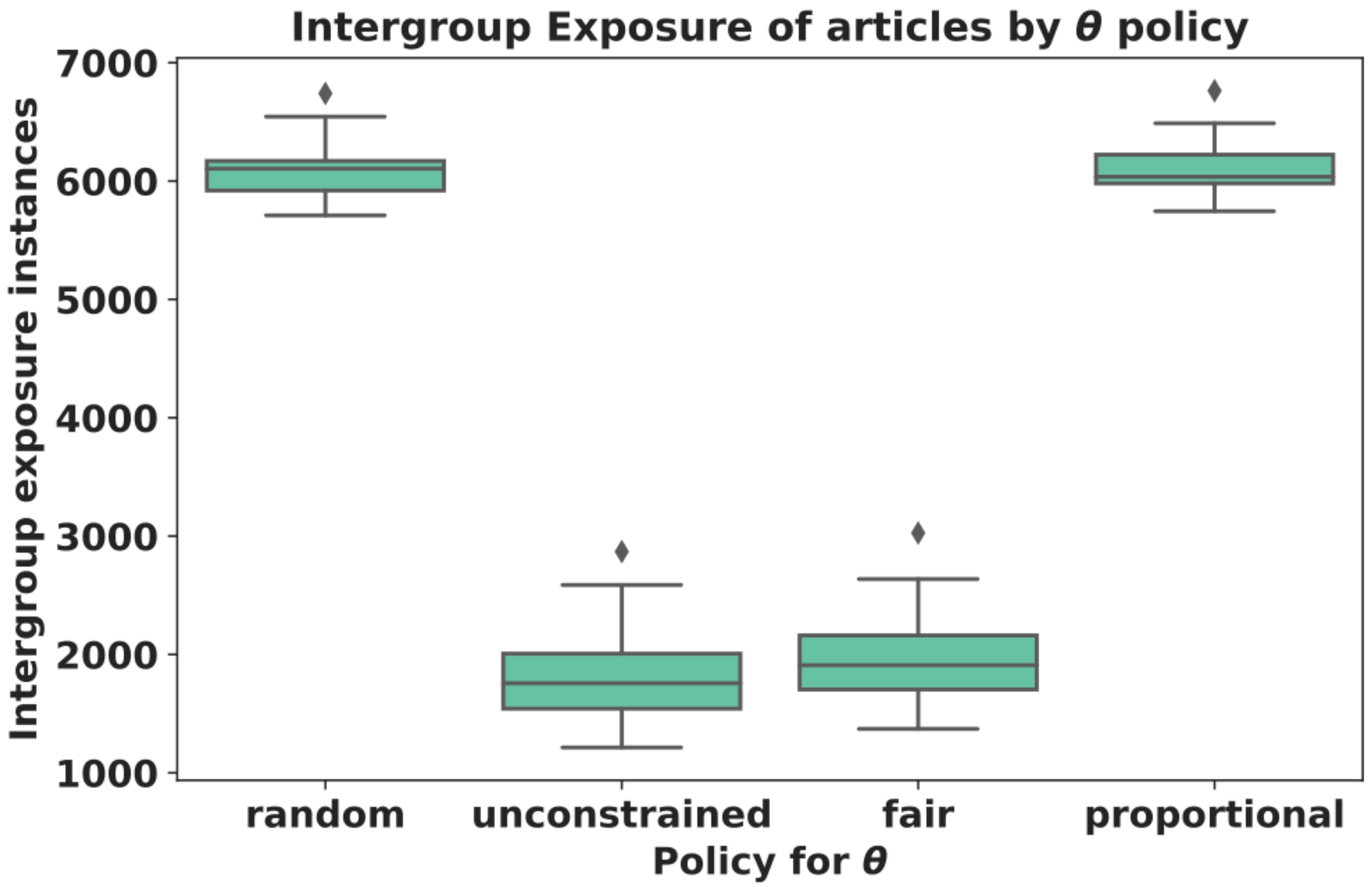
Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N. (2017). Balancing information exposure in social networks. *NeurIPS*

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

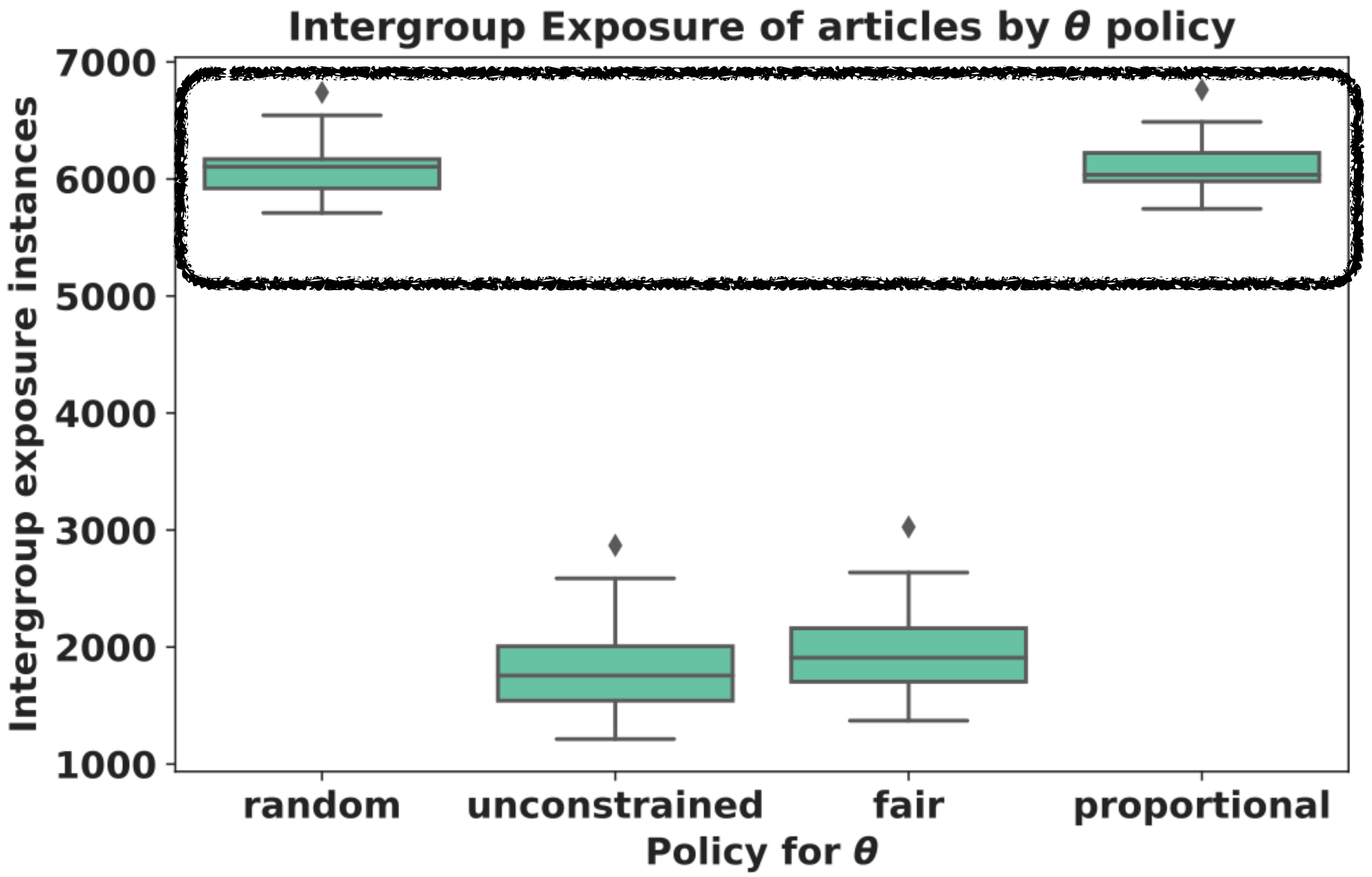


$$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

$\Theta = \big($prop group a, prop group b, prop group a, prop group $b\big)$

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

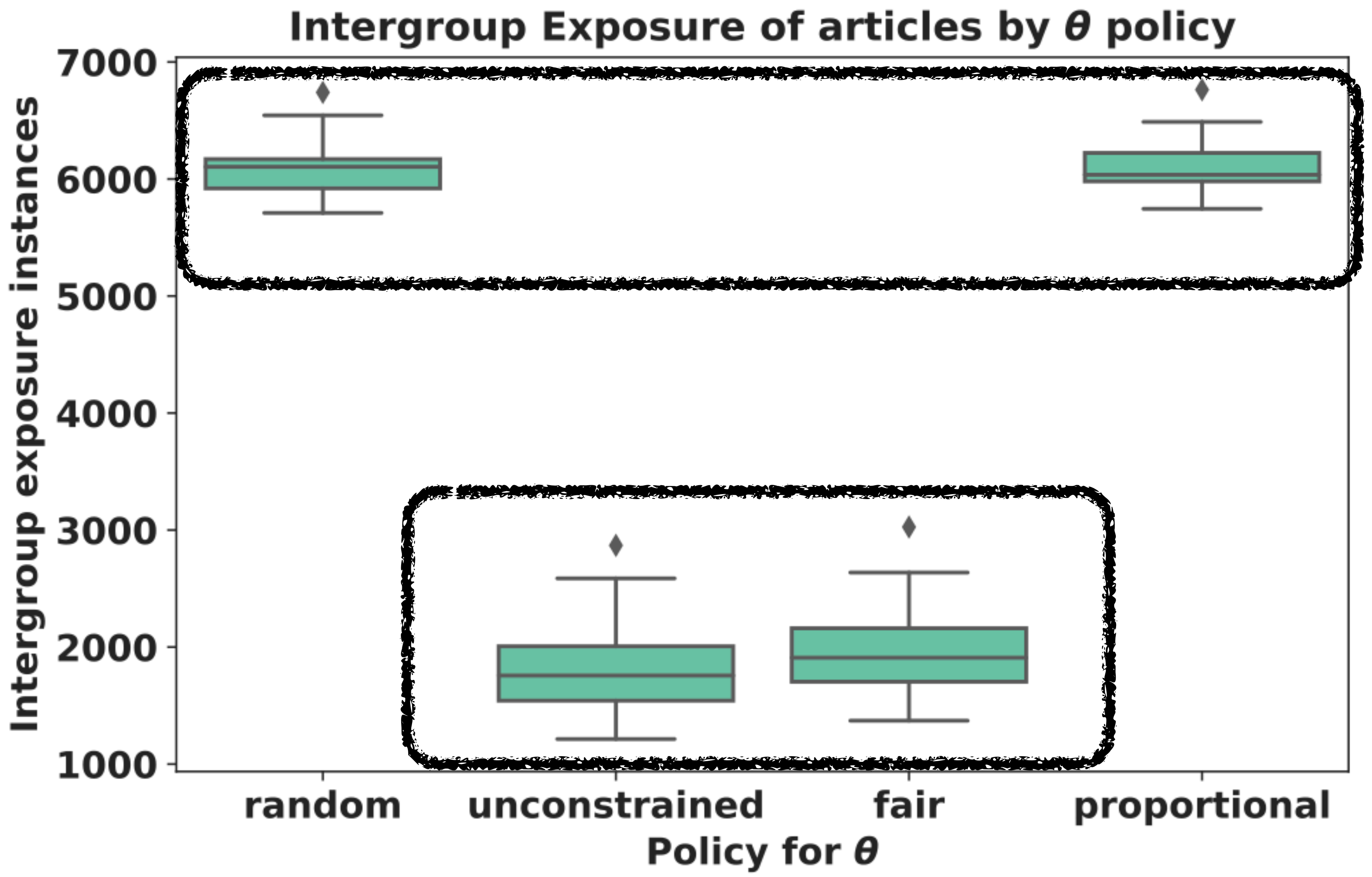


$$\Theta = \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$$

$$\Theta = \left(\text{prop group } a, \text{prop group } b, \text{prop group } a, \text{prop group } b\right)$$

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

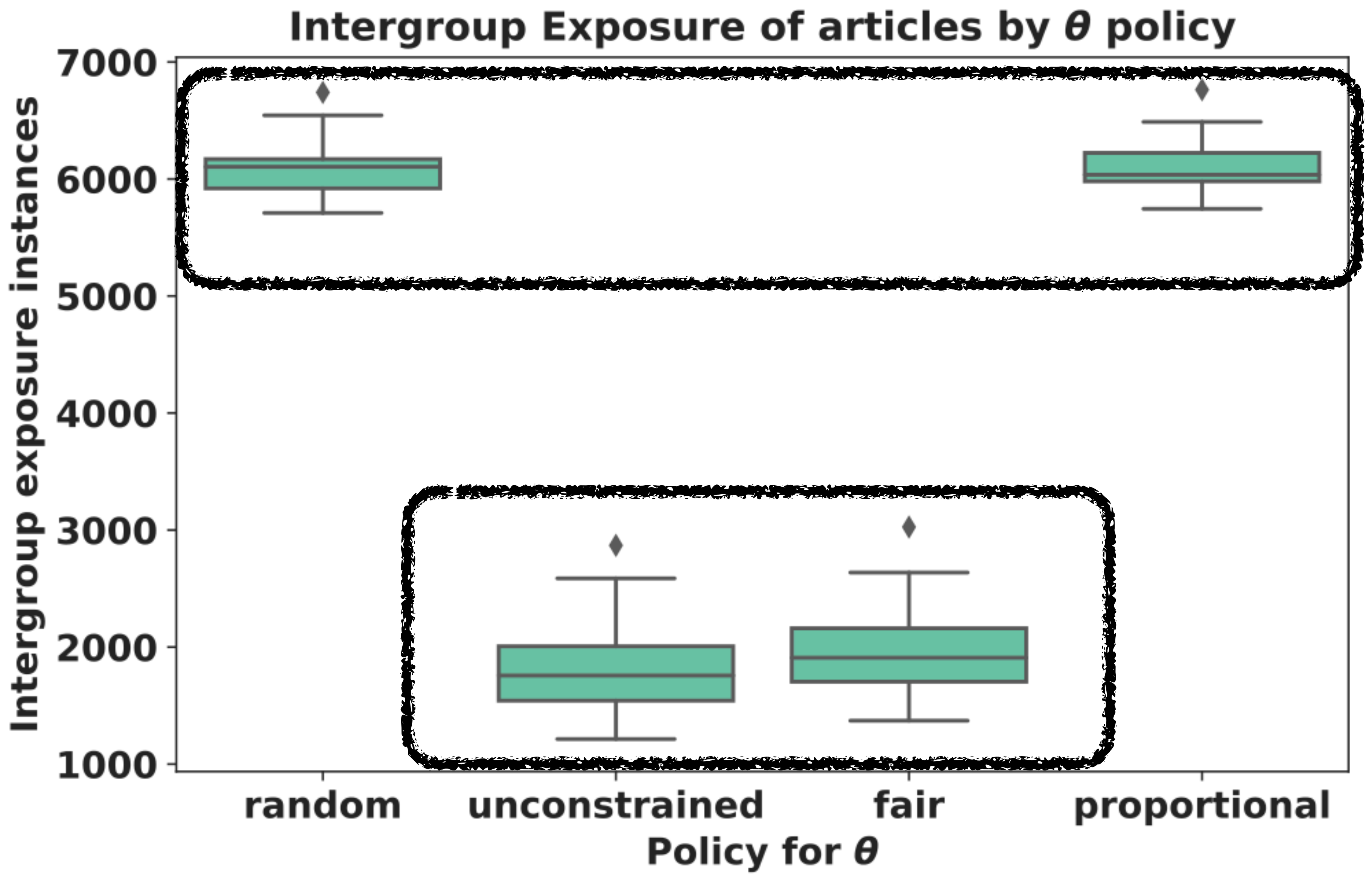


$$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

$$\Theta = \left(\text{prop group } a, \text{prop group } b, \text{prop group } a, \text{prop group } b\right)$$

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$

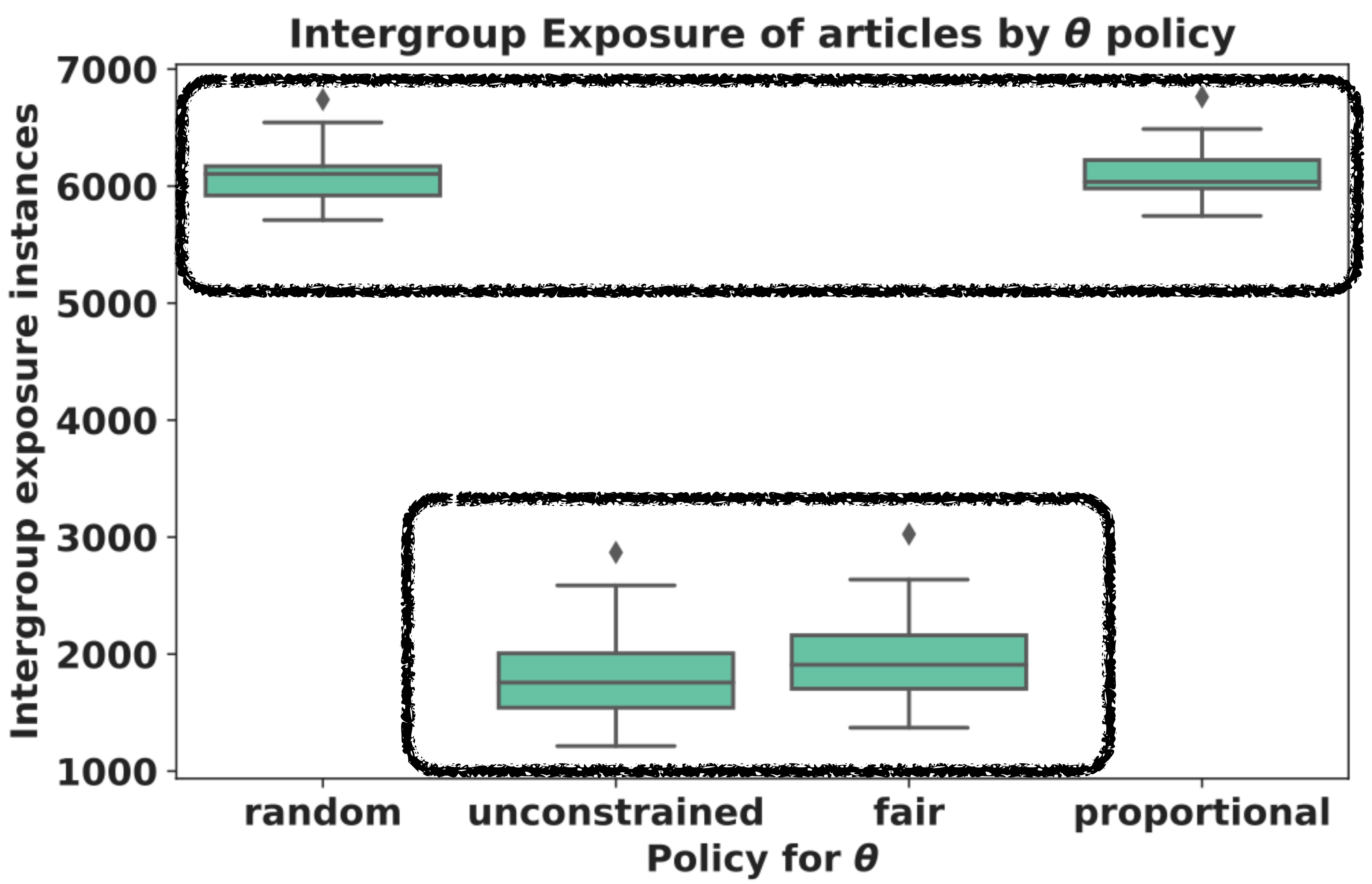


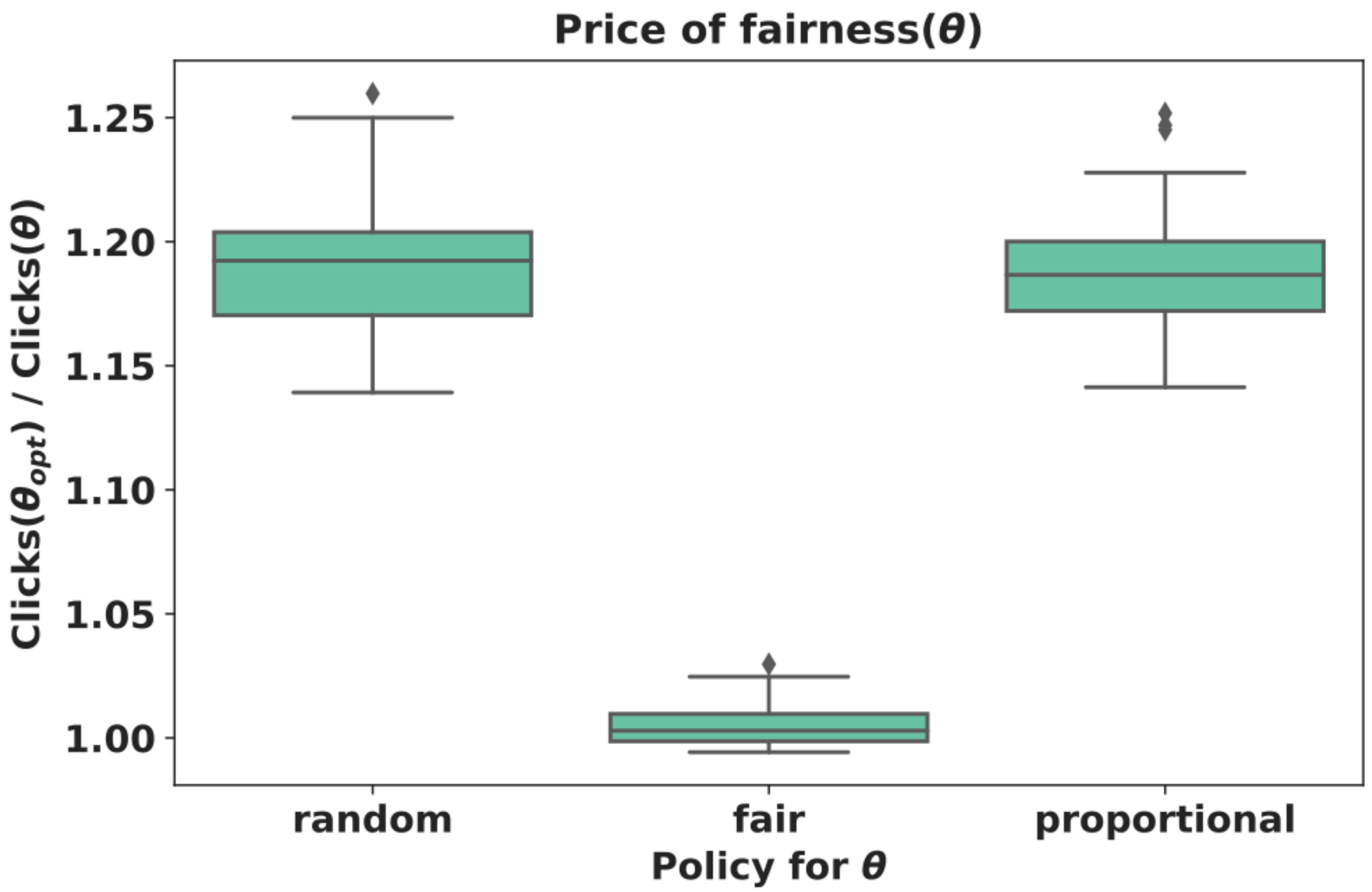
$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ $\Theta = \left(\text{prop group } a, \text{prop group } b, \text{prop group } a, \text{prop group } b\right)$

1. Modest improvements to intergroup exposures with fairness constraints, but significantly reduced from “naive” baseline policies

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$





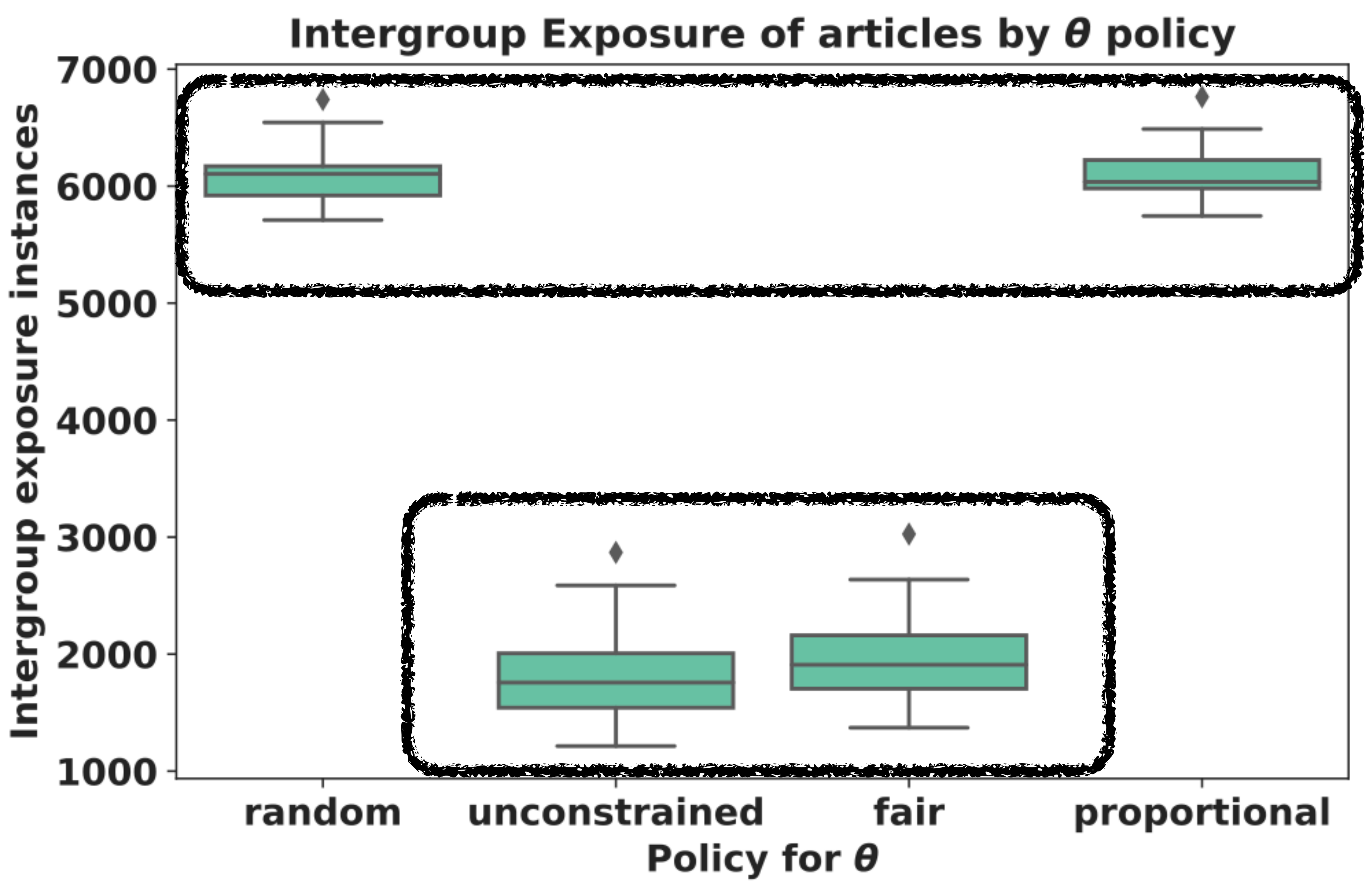
$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$

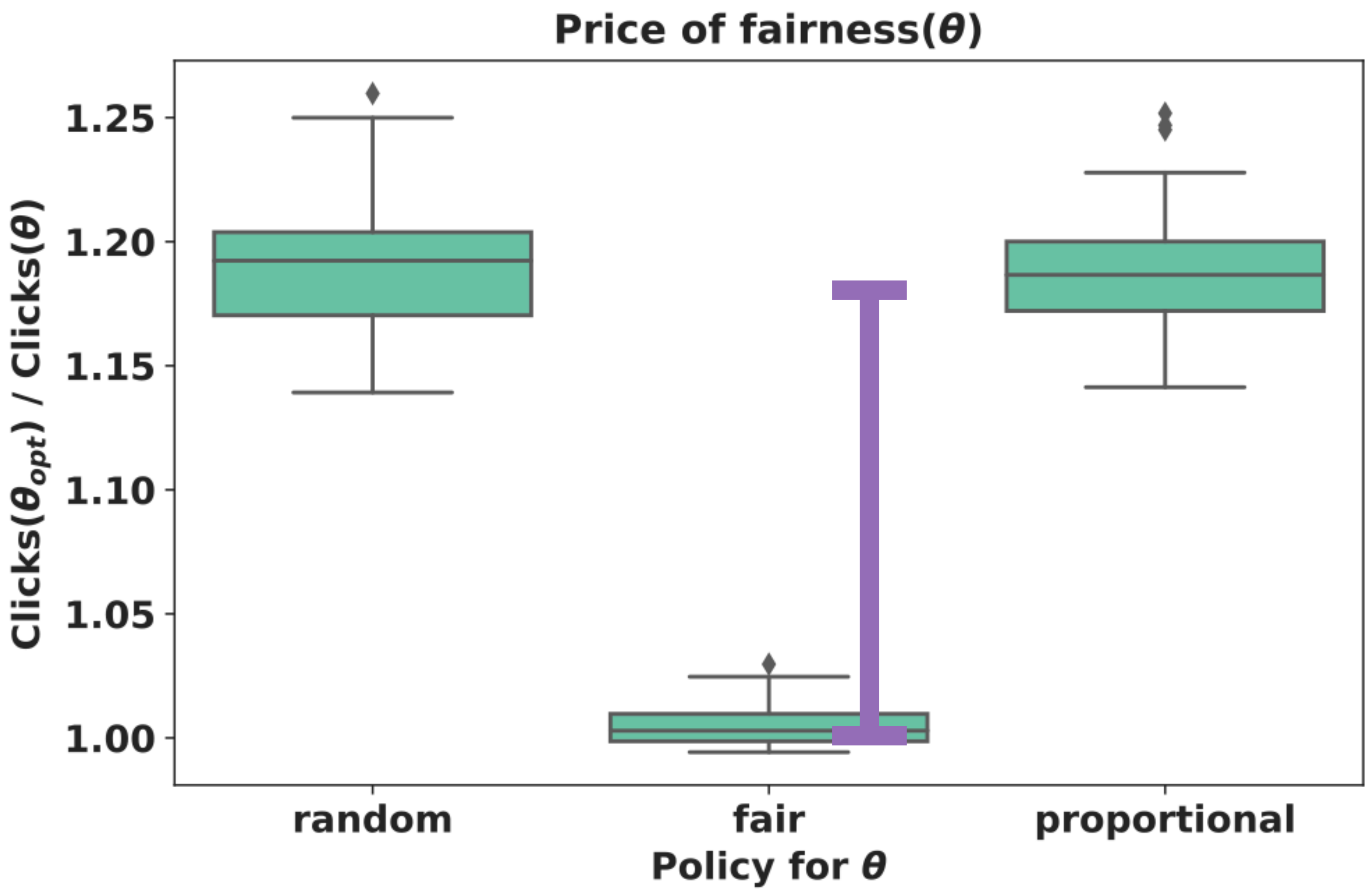
$\Theta = \left(\text{prop group } a, \text{prop group } b, \text{prop group } a, \text{prop group } b\right)$

1. Modest improvements to intergroup exposures with fairness constraints, but significantly reduced from "naive" baseline policies

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$





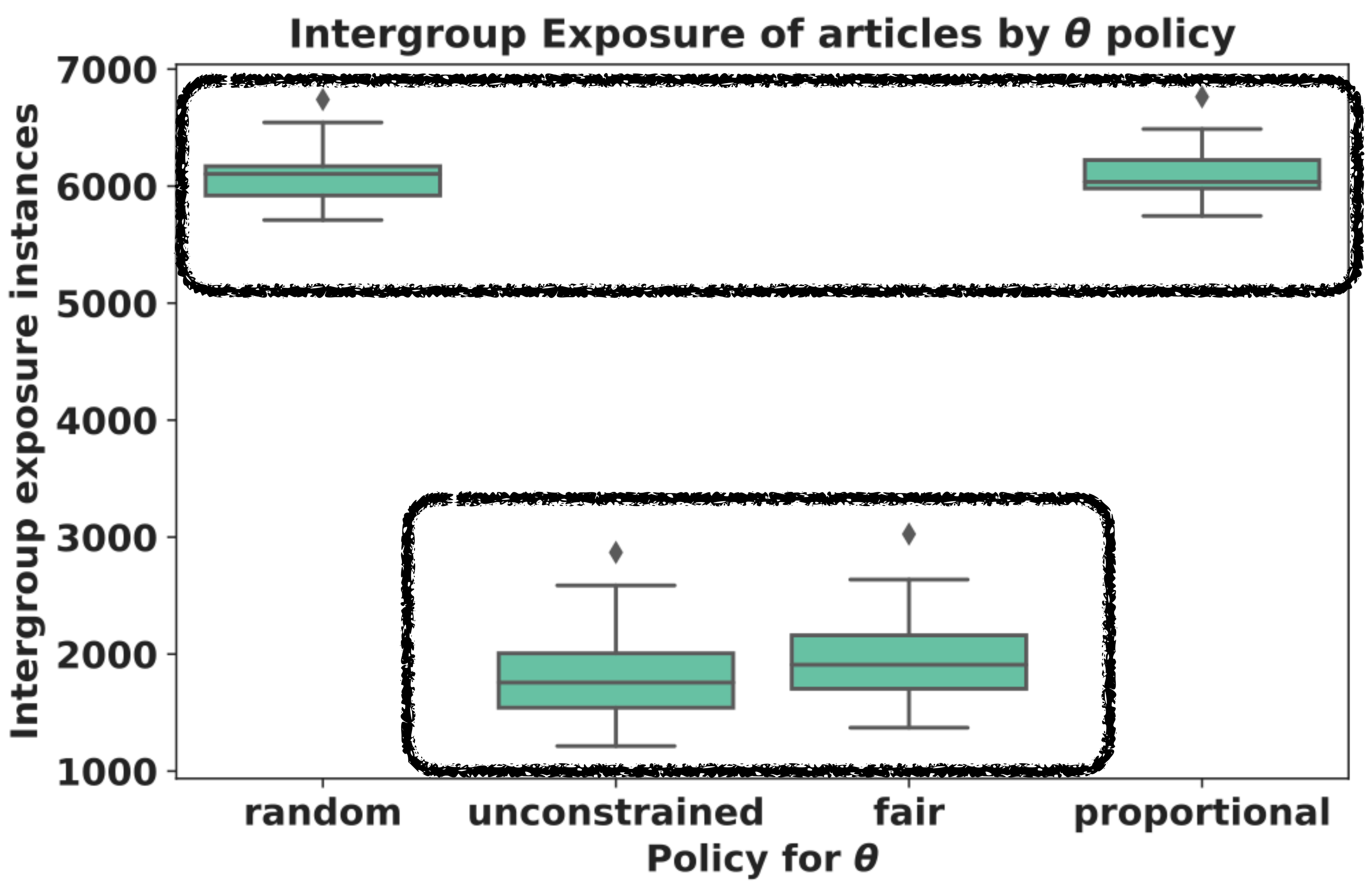
$$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

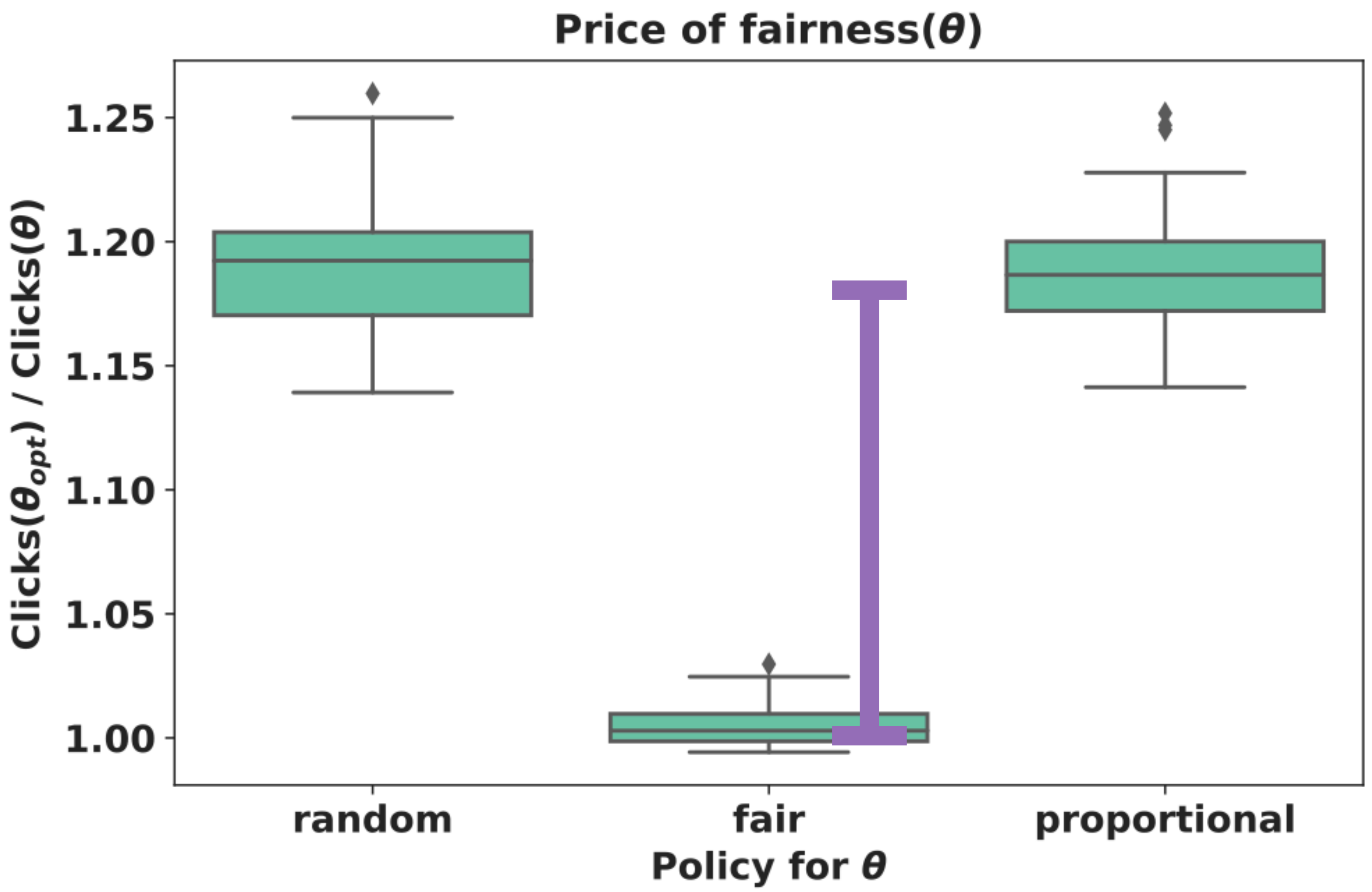
$$\Theta = \left(\text{prop group } a, \text{prop group } b, \text{prop group } a, \text{prop group } b\right)$$

1. Modest improvements to intergroup exposures with fairness constraints, but significantly reduced from "naive" baseline policies

Do constraints improve intergroup interactions?

$$\Theta = (\Theta_{a,A}, \Theta_{a,B}, \Theta_{b,A}, \Theta_{b,B})$$





$$\Theta = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

$\Theta = \big($prop group a, prop group b, prop group a, prop group $b\big)$

1. Modest improvements to intergroup exposures with fairness constraints, but significantly reduced from “naive” baseline policies
2. High benefit in clicks to optimize policy, fair exposure constraints come at marginal cost

In summary…

In summary…

Present a stylized model to understand if the essence of the Fairness Doctrine can be applied to online platforms

In summary…

Present a stylized model to understand if the essence of the Fairness Doctrine can be applied to online platforms

A policy enforcing fair exposure is insufficient on its own

In summary…

Present a stylized model to understand if the essence of the Fairness Doctrine can be applied to online platforms

A policy enforcing fair exposure is insufficient on its own

Empirically, can lead to modest improvements in cross-group content exposure

In summary…

Present a stylized model to understand if the essence of the Fairness Doctrine can be applied to online platforms

A policy enforcing fair exposure is insufficient on its own

Empirically, can lead to modest improvements in cross-group content exposure

Implications: Integrating principles from the Fairness Doctrine into online platforms seems to require more than imposing constraints.

In summary…

Present a stylized model to understand if the essence of the Fairness Doctrine can be applied to online platforms

A policy enforcing fair exposure is insufficient on its own

Empirically, can lead to modest improvements in cross-group content exposure

Implications: Integrating principles from the Fairness Doctrine into online platforms seems to require more than imposing constraints.

Approaches to mitigating interactional polarization may be more effective.

Thank you!

arXiv

Questions? jessie@seas.harvard.edu





SIMONS INSTITUTE for the Theory of Computing

MD4SG Mechanism Design for Social Good

Omitted citations

Hązła, Jan, et al. "A geometric model of opinion polarization." *Mathematics of Operations Research* (2023).
Jason Gaitonde, Jon Kleinberg, and Éva Tardos. "Polarization in geometric opinion dynamics." *EC* 2021.
Dean et al. "Preference dynamics under personalized recommendations." *EC* 2022.
Celis et al. "Controlling polarization in personalization: An algorithmic framework." *FAccT* 2019.
Ali et al. "On the fairness of time-critical influence maximization in social networks." *IEEE Transactions on Knowledge and Data Engineering* (2021).
Fish et al. "Gaps in information access in social networks?." *TheWebConf* 2019.
Stoica al. "Seeding network influence in biased networks and the benefits of diversity." *TheWebConf* 2020.
Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 2015.
Garimella, Venkata Rama Kiran, and Ingmar Weber. "A long-term analysis of polarization on Twitter." *ICWSM* 2017.
Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. "Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media." *Political Communication* (2021).