Detecting Zealots in Social Networks

A Spatial Autoregressive (SAR) Model Approach

Introductory Econometrics Final Presentation

陳柏安 (B10202029)
陳立凡 (B10601001)

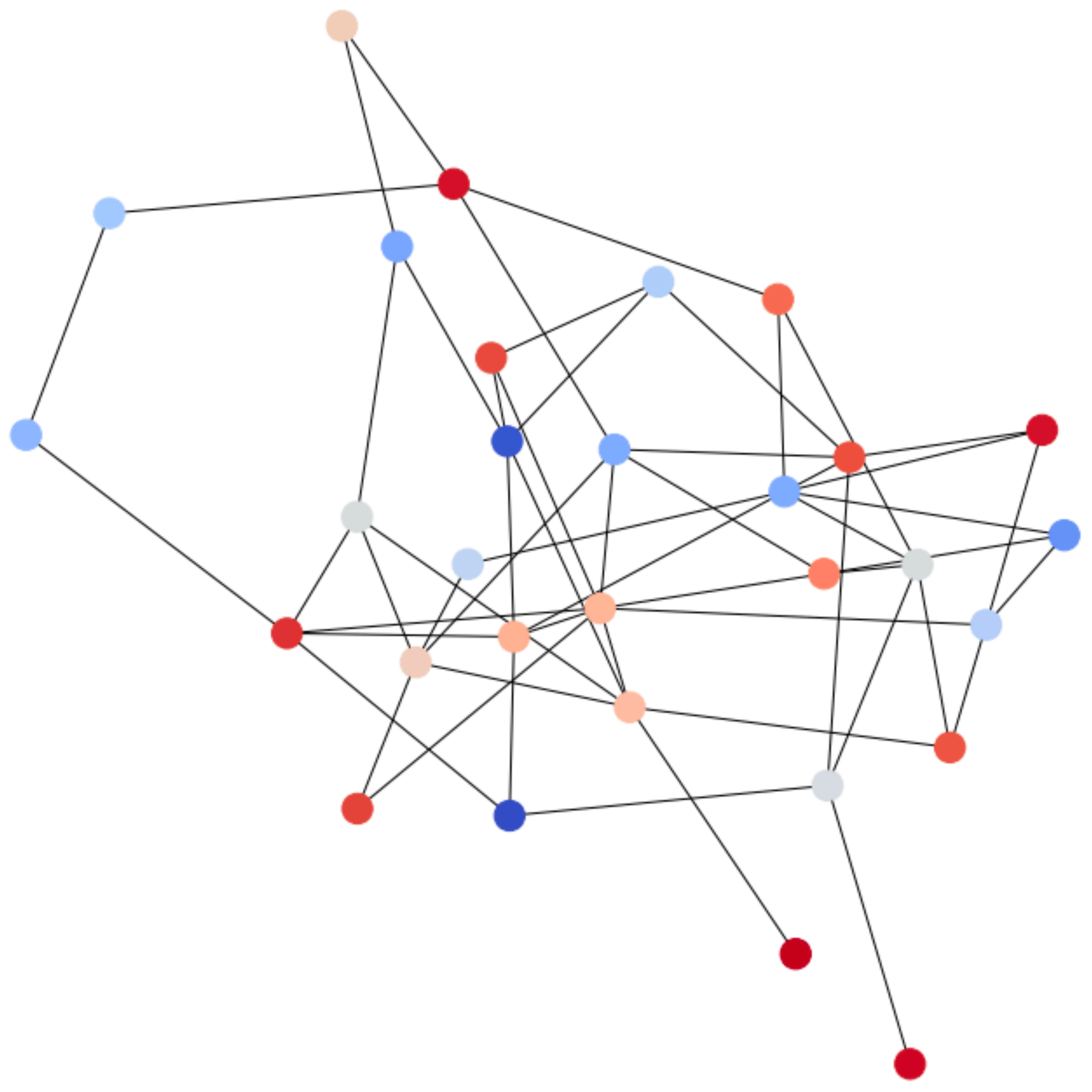
Outlines

- Opinion Dynamics and Social Network
- The Interference of Zealots
- Conventional Deffuant model: Test Statistics
- Adaptive Network: Spatial Autoregressive model
- Toward Applied Microeconomic Theory Studies

Opinion Dynamics

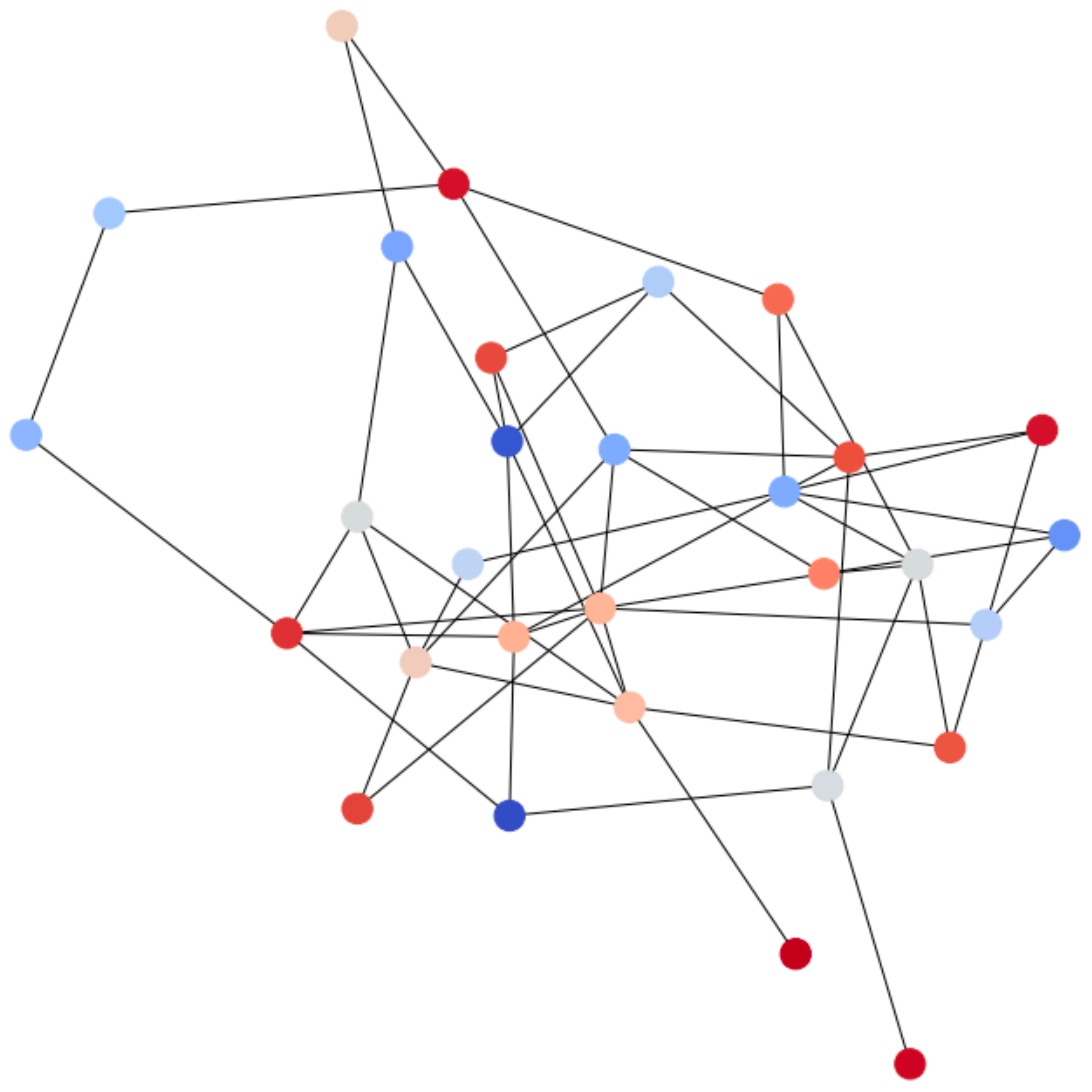
Social Network

- Network = Nodes + Edges
- Nodes: social actors
- Edges: connections



Opinion Dynamics

- Node attributes:
  - opinion: -1~1
- Nodes will exchange their opinion



Opinion Dynamics

Deffuant model

- For each time, an edge is chosen randomly
- If the difference of their opinion $< c$:
  - $o_1 \leftarrow o_1 + \alpha(o_2 - o_1)$
  - $o_2 \leftarrow o_2 + \alpha(o_1 - o_2)$

Opinion Dynamics

Deffuant model with zealot

- For each time, an edge is chosen randomly
- If the difference of their opinion $< c$:
    - $o_1 \leftarrow o_1 + \alpha(o_2 - o_1) \cdot z_1$
    - $o_2 \leftarrow o_2 + \alpha(o_1 - o_2) \cdot z_2$
- where $z_i = 0$ if i is zealot, otherwise 1

The Interference of Zealots

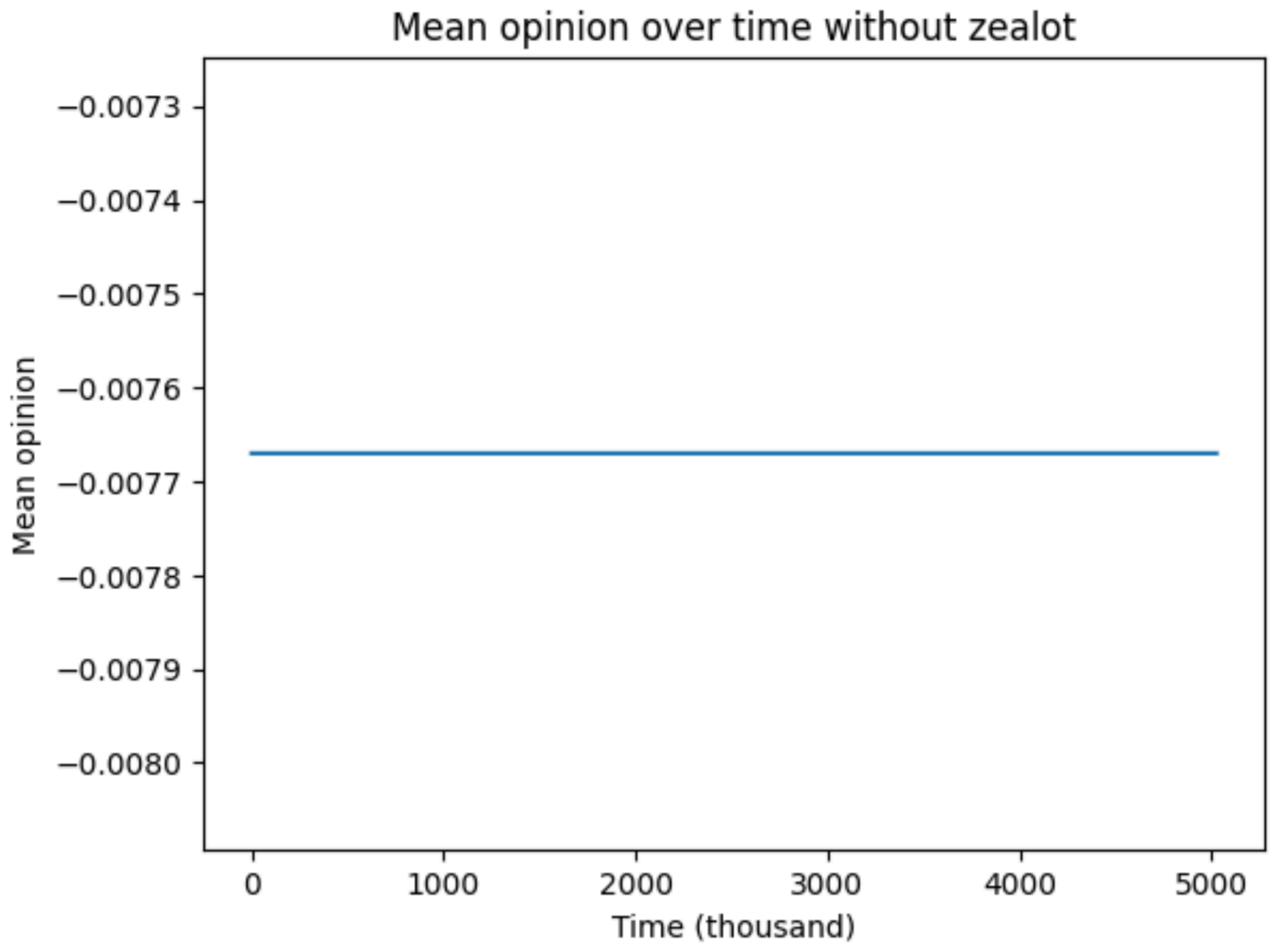
Conservation without zealots

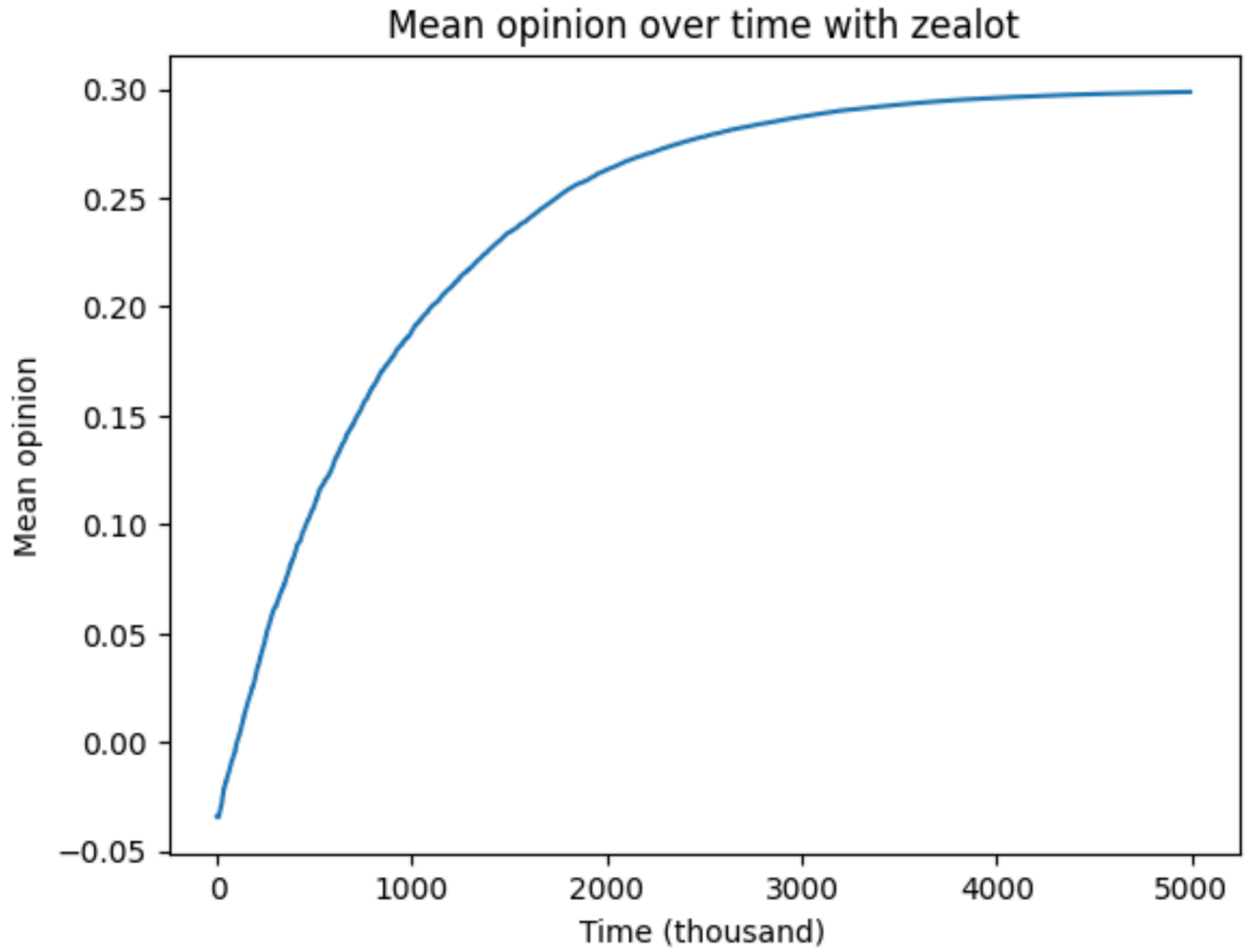
- $o_1^{t+1} = o_1^t + \alpha(o_2^t - o_1^t)$
- $o_2^{t+1} = o_2^t + \alpha(o_1^t - o_2^t)$
- $o_1^{t+1} + o_2^{t+1} = o_1^t + o_2^t$
- $\sum o_i^{t+1} = \sum o_i^t$

The Interference of Zealots

Conservation broke by zealots

- $o_1^{t+1} = o_1^t$ (zealot)
- $o_2^{t+1} = o_2^t + \alpha(o_1^t - o_2^t)$
- $o_1^{t+1} + o_2^{t+1} = o_1^t + o_2^t + \alpha(o_1^t - o_2^t)$
- $\sum o_i^{t+1} = \sum o_i^t + \alpha(o_1^t - o_2^t)$
- Mean opinion moves towards o_1


Mean opinion over time without zealot
−0.0073
−0.0074
−0.0075
−0.0076
−0.0077
−0.0078
−0.0079
−0.0080
Mean opinion
0
1000
2000
3000
4000
5000
Time (thousand)


Mean opinion over time with zealot
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
Mean opinion
0
1000
2000
3000
4000
5000
Time (thousand)

Test Statistics

Motivation

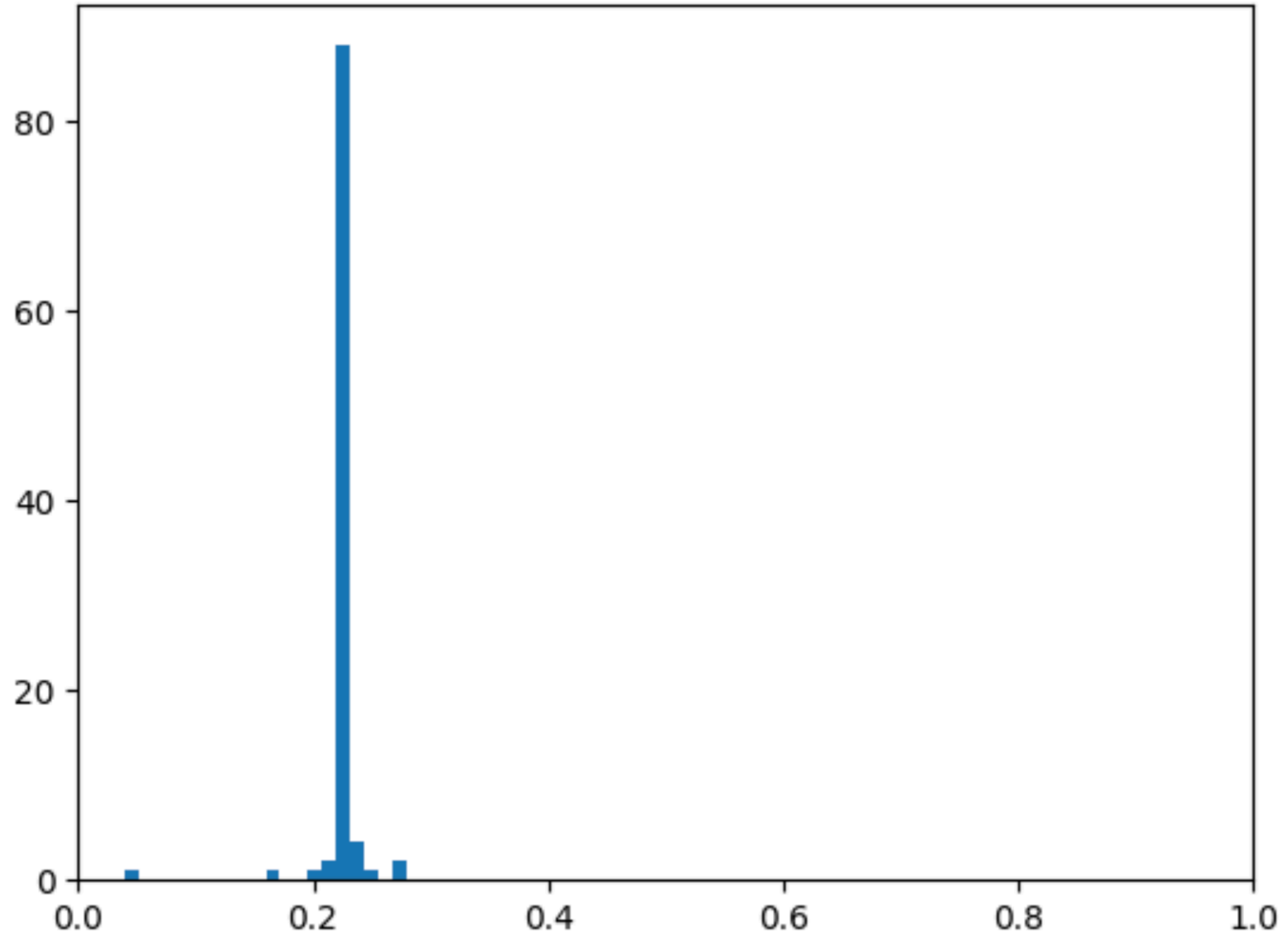
- Population mean will remain the same if there is no zealot
- Population mean will move toward a value if there is zealots
- Suppose we
  - Sample periodically
  - Calculate sample mean
  - Run linear regression $\bar{o}_t \sim t$ and get R^2
- We get to distinguish the two cases

Test Statistics
Monte Carlo

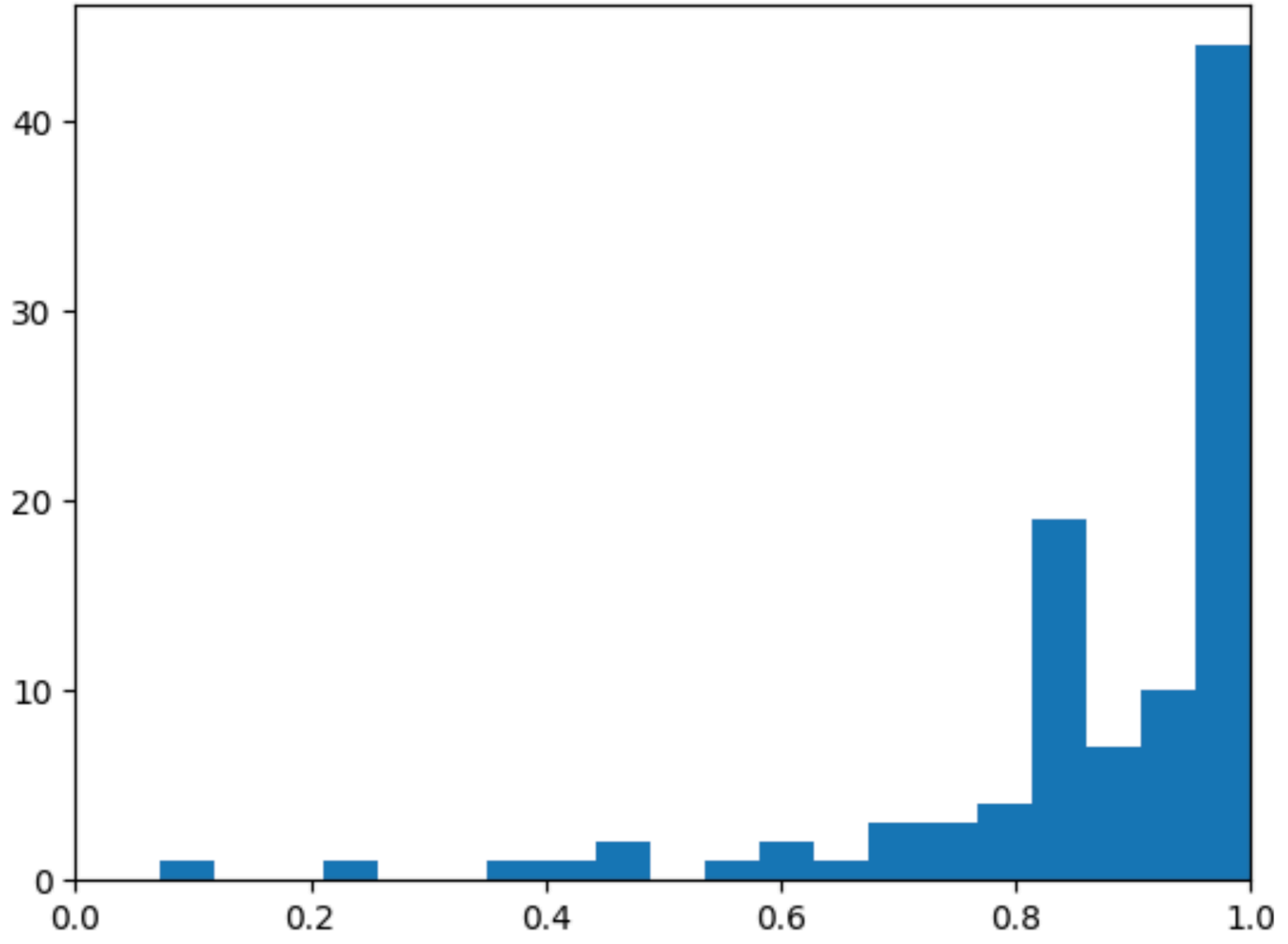
- Repeat 100 times:
    - Simulate opinion dynamics (with (10*edge number) iterations)
    - Sample opinions with same / different people for 25 panel waves
    - Calculate the R^2 of mean opinions — time evolution

Conventional Deffuant model: Test Statistics

Result: same people



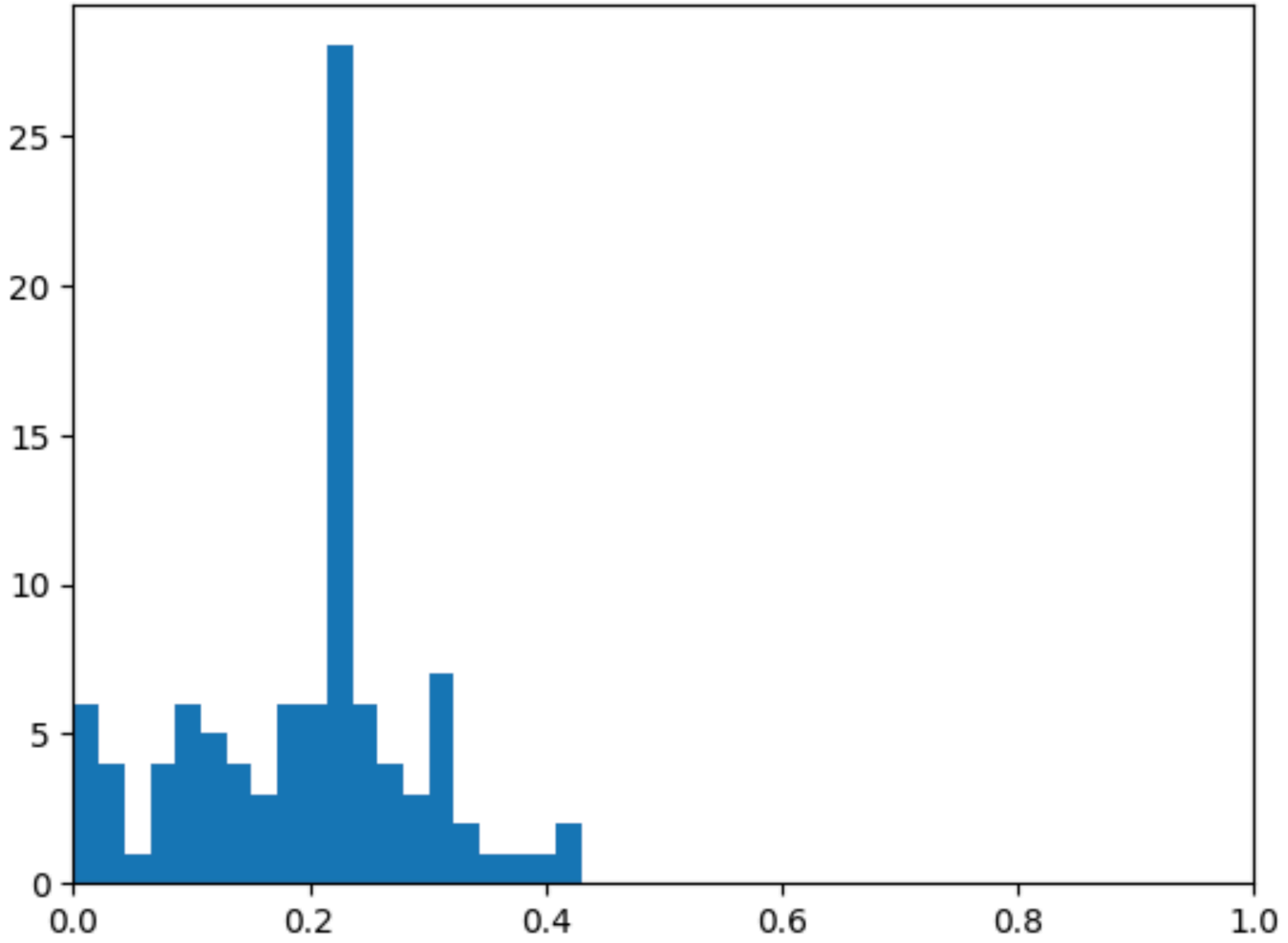
R^2 Without Zealot



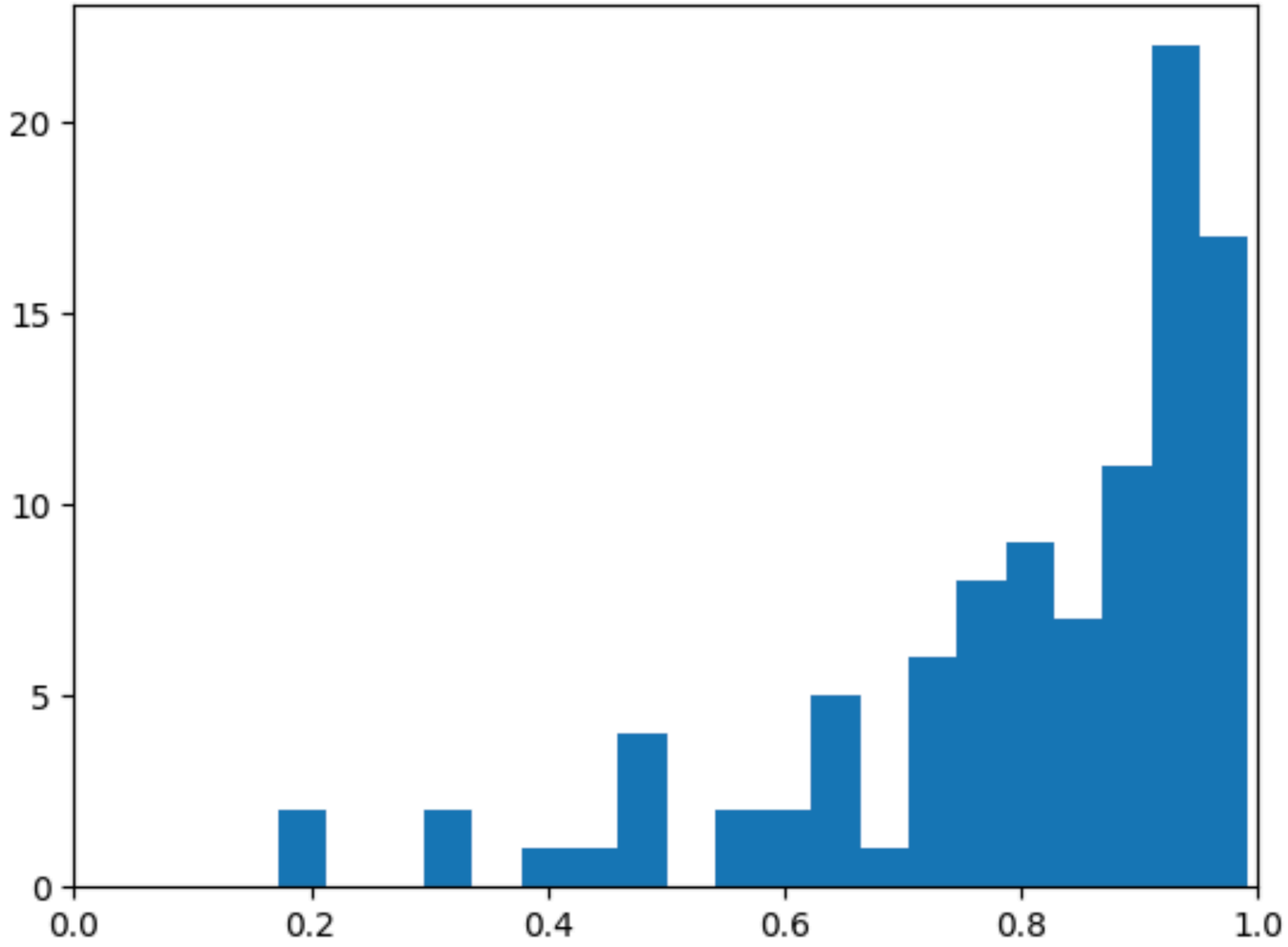
R^2 With Zealot (o = 0.3)

Conventional Deffuant model: Test Statistics

Result: different people



R^2 Without Zealot



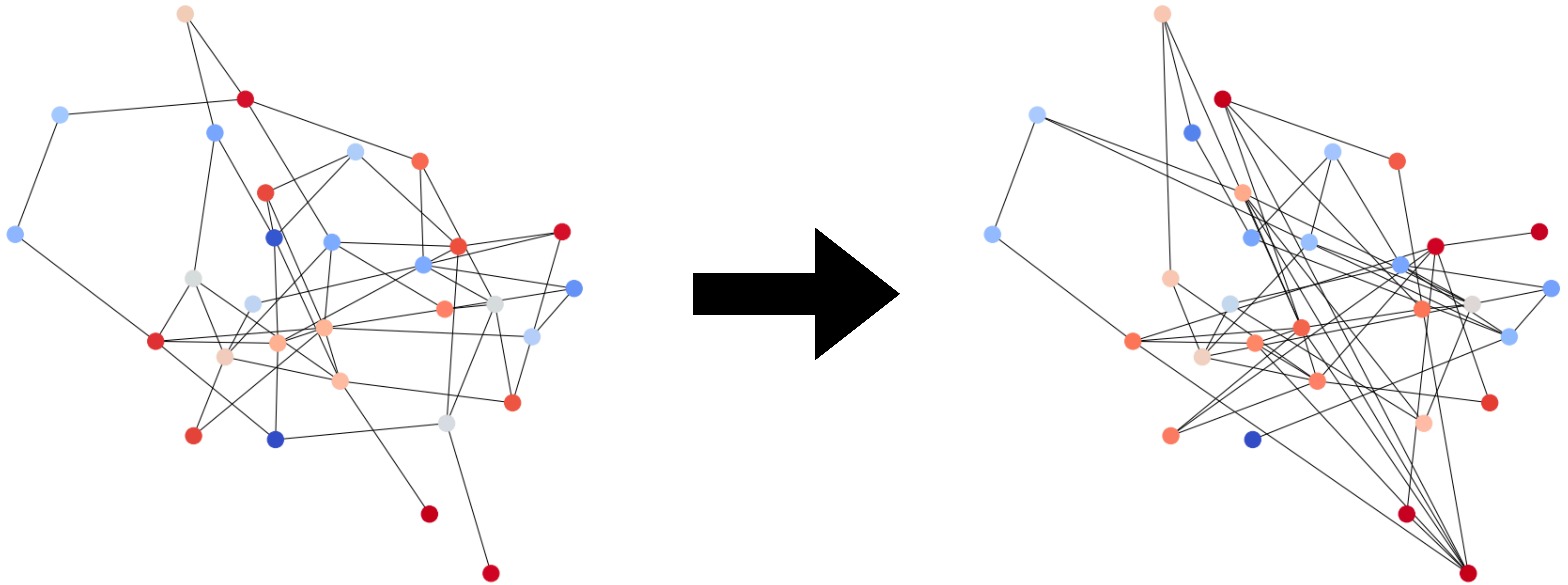
R^2 With Zealot (o = 0.3)

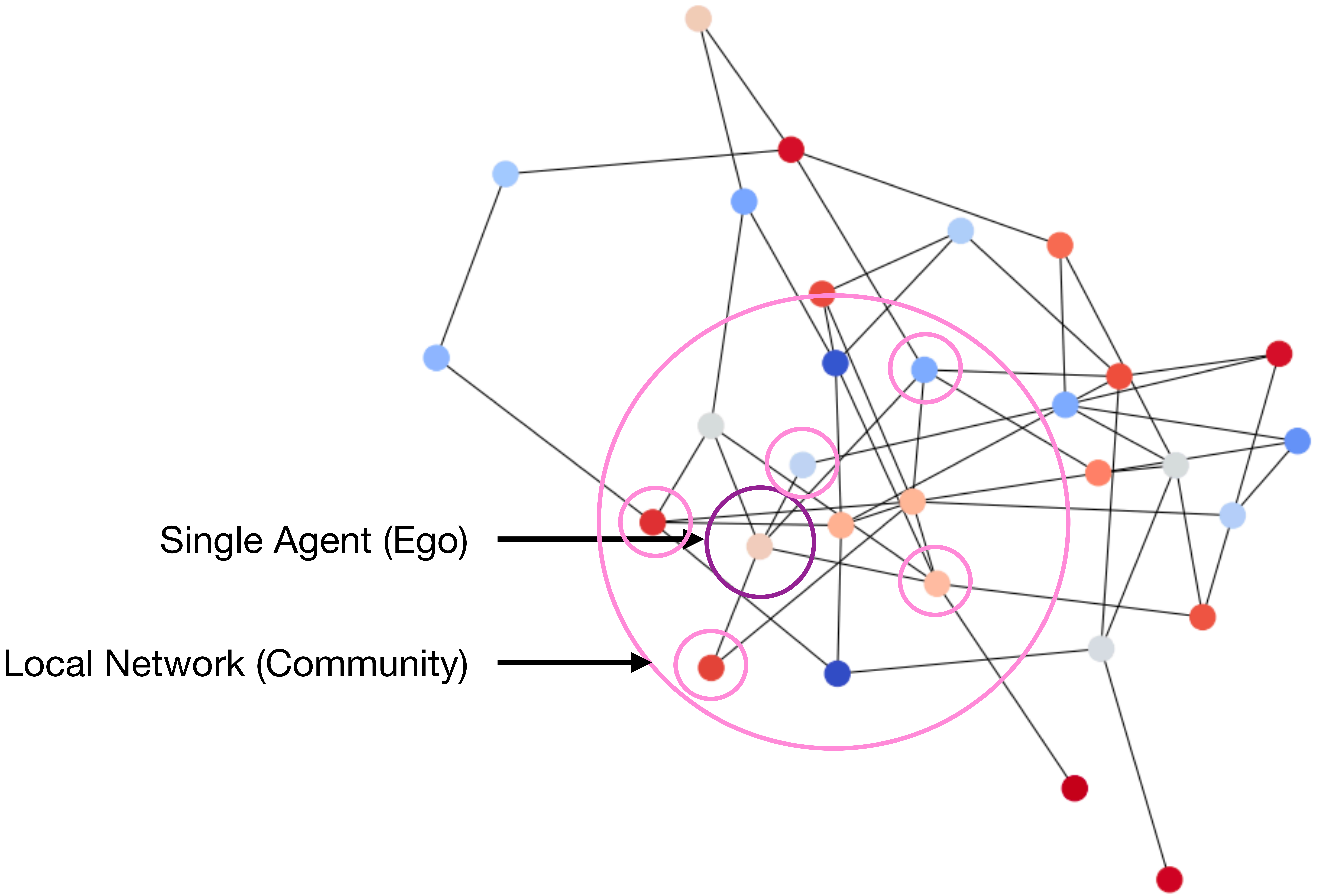
Adaptive Network

- People are smart (maybe not), and they choose whom to be friends
- In this study, they make decision based on ***Similarity***
- We consider opinion similarity, which is also known as the social learning (DeGroot, 1974)

✳ Generalization is possible (See Appendix)

Credit: iStock




Single Agent (Ego)
Local Network (Community)

Adaptive Network

Spatial Autoregressive Model

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \beta + \epsilon$$

- $\mathbf{y}$ is the response variable
- ρ is the spatial lag coefficient
- $\mathbf{W}$ is the adjacency matrix of the network
- $\mathbf{X}$ is the exogenous variables
- β is the coefficient of exogenous effects
- ϵ is the error term

Adaptive Network

Spatial Autoregressive Model

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \beta + \epsilon$$

- In the past, linear-in-means models are used to study the peer effects (Manski, 1993)
- However, it can not deal with endogenous effects
- Hence, we use MLE methods proposed by Lee et al (2004). See Appendix

Adaptive Network

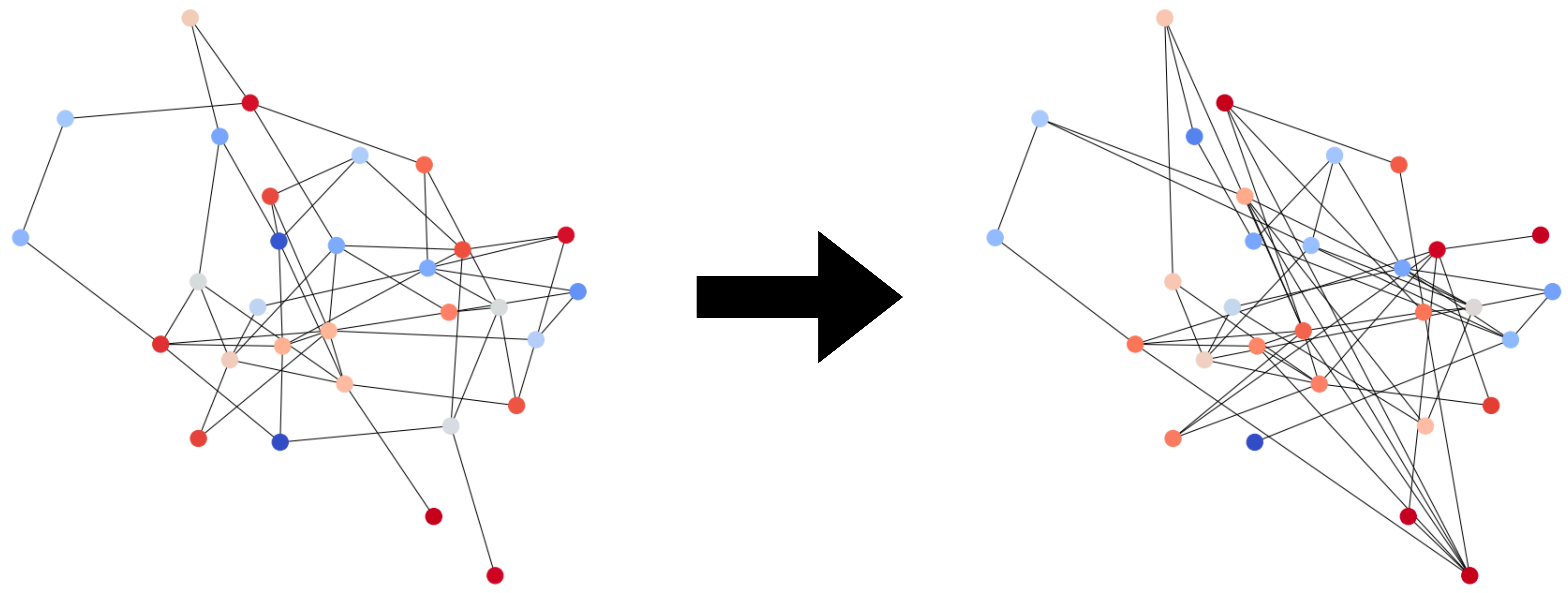
SAR Model on Opinion Dynamics

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \beta + \epsilon$$

- $\mathbf{y}$ is the opinions of agents
- $\mathbf{x}_1$ is the degree of nodes (baseline)
- $\mathbf{x}_2$ is the clustering coefficients
- We collect multiple panel waves from MCMC

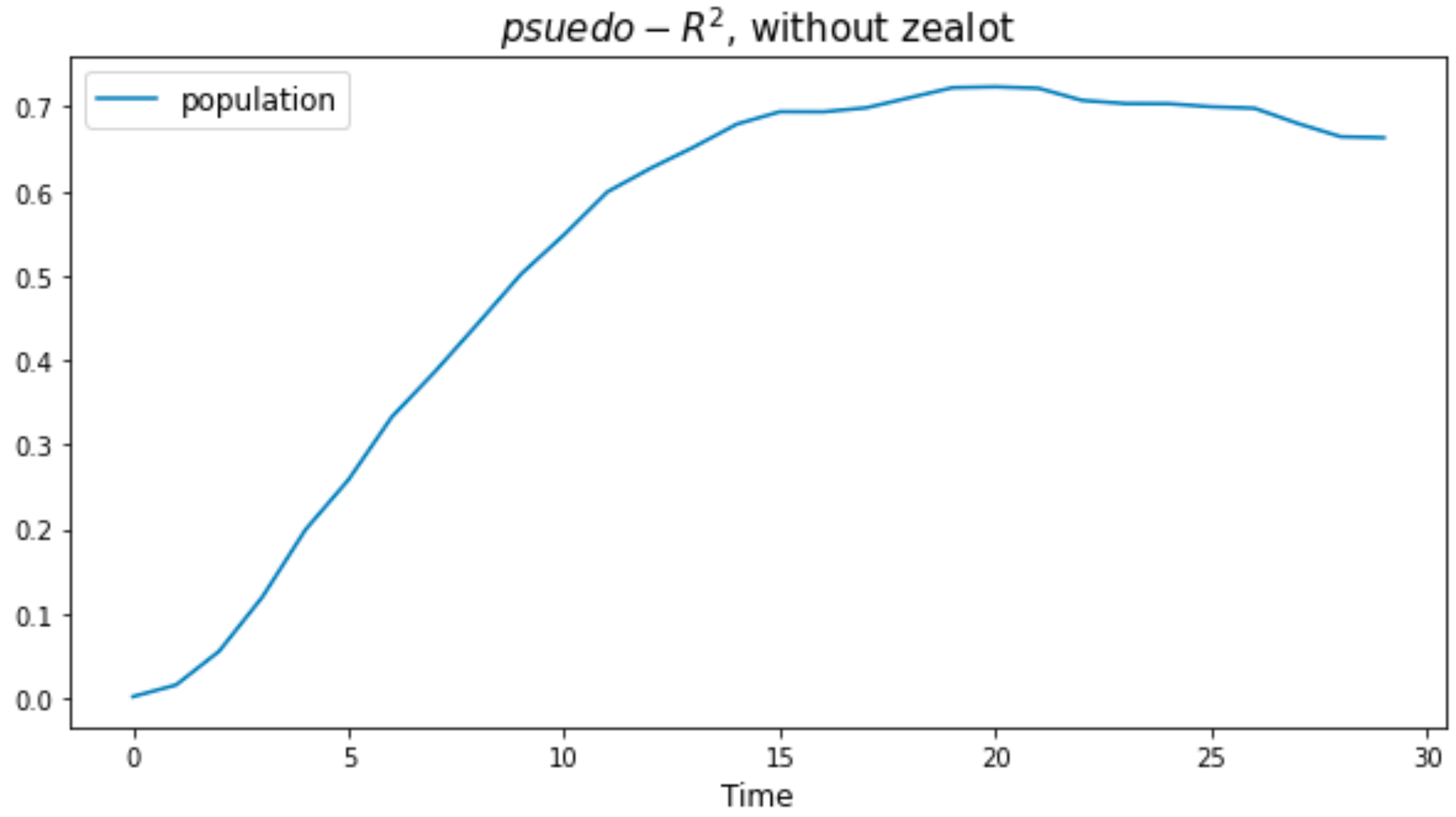
✳ We mostly care about ρ and β_2

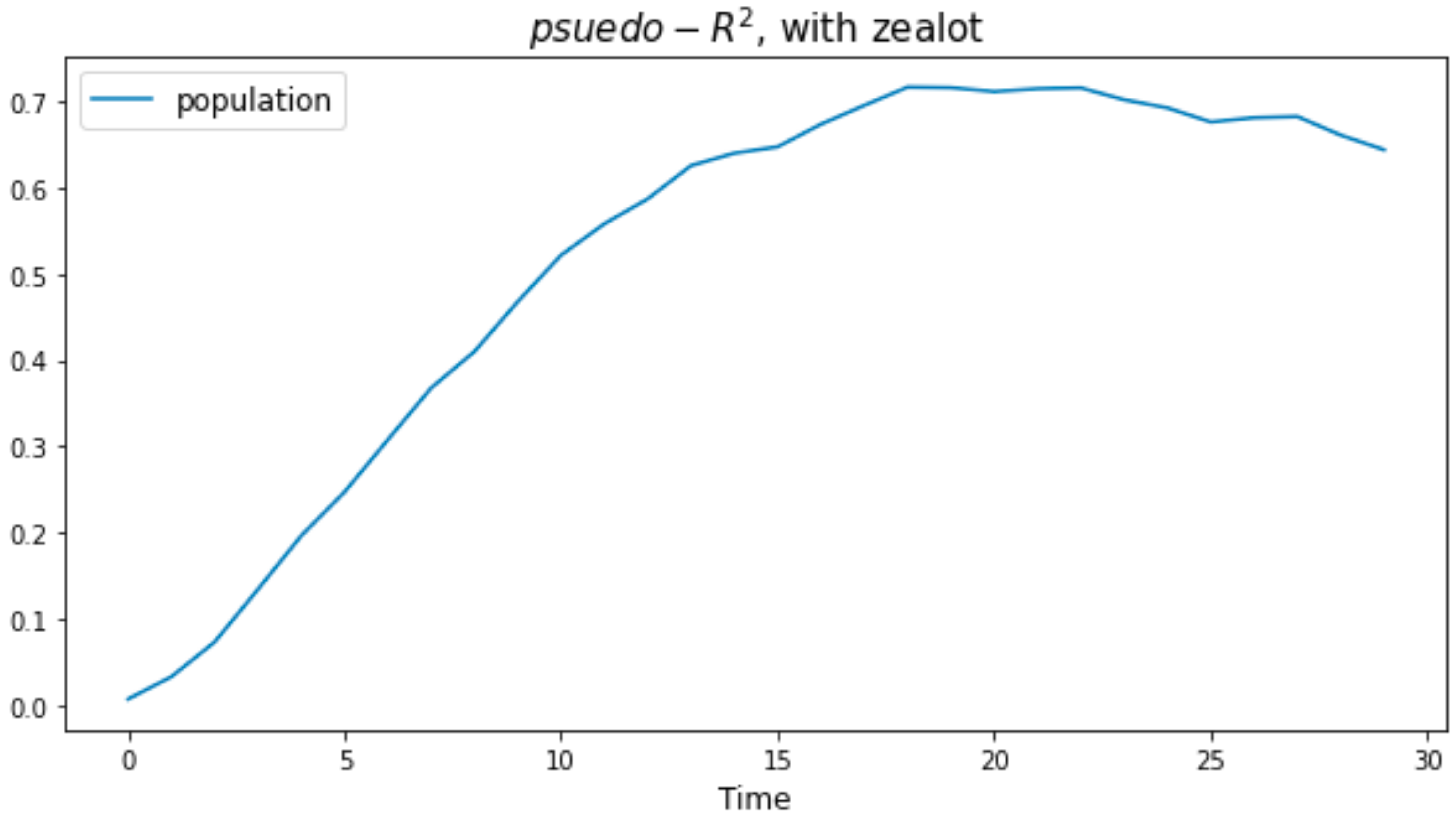
$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X}\beta + \epsilon$$



Adaptive Network

Population Analysis — Goodness of Model Fitting

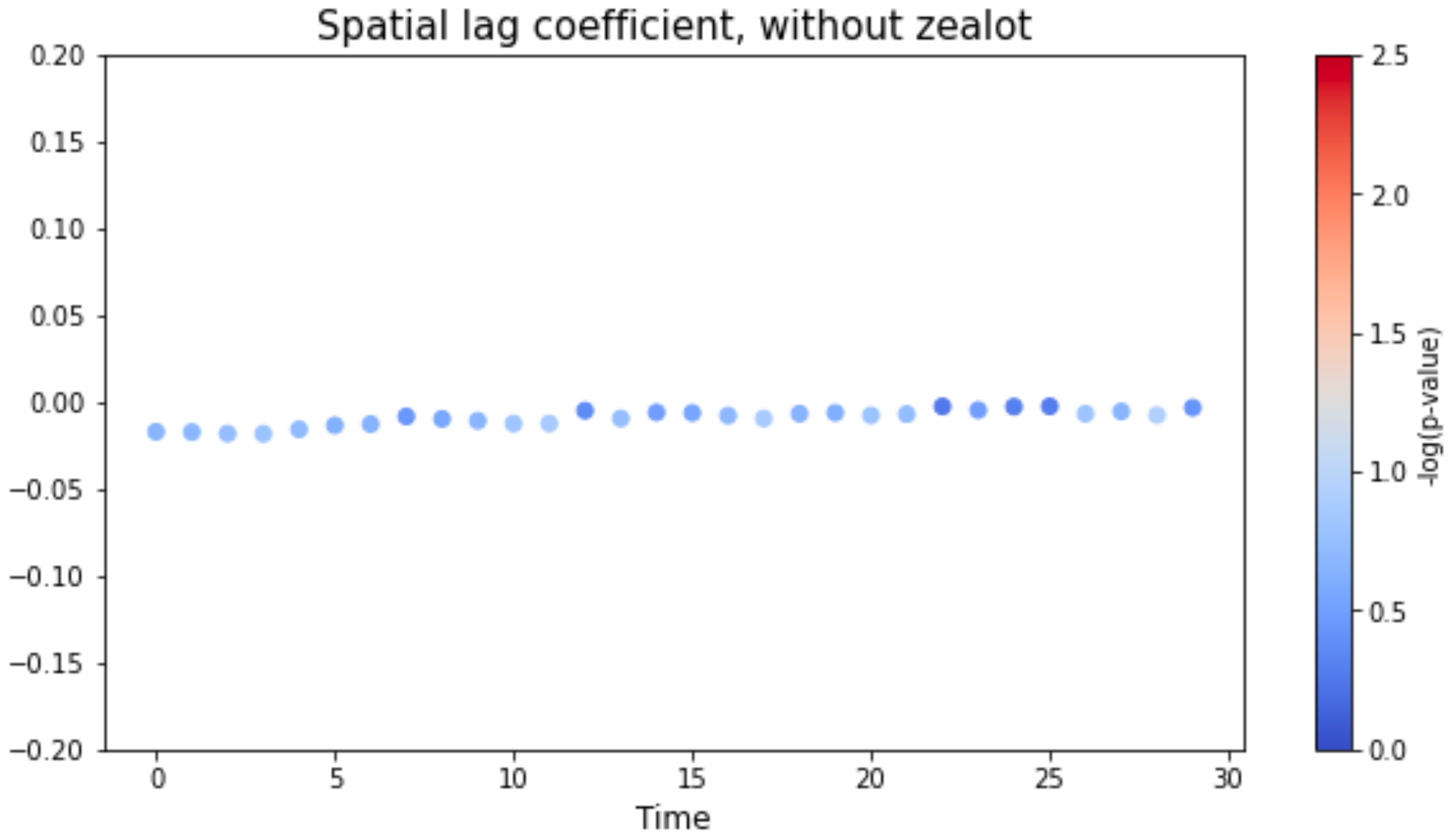


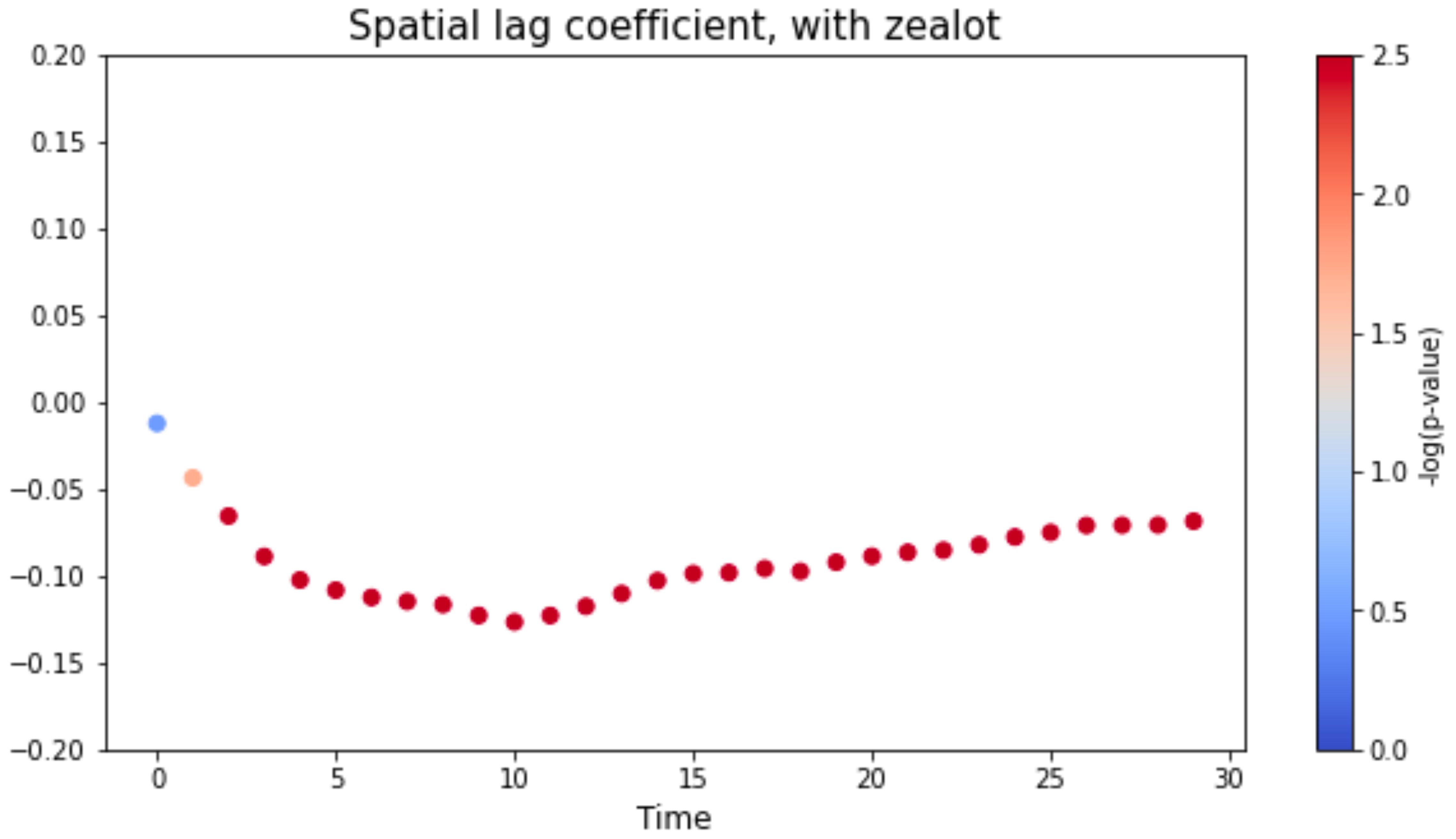


10 zealot with $o(z_i) = 0.9, \forall i \in Z$

Adaptive Network

Population Analysis — Peer Effects

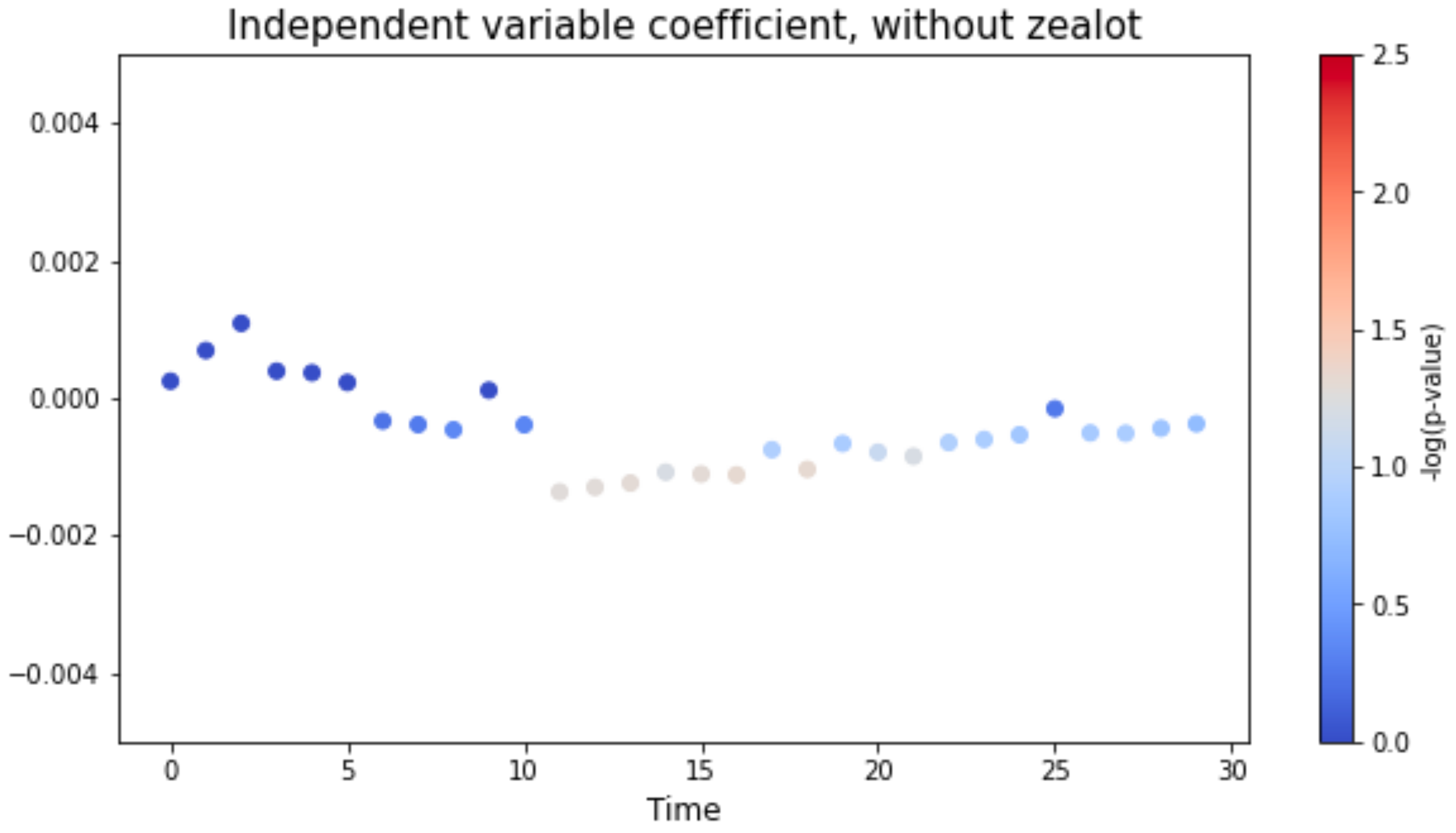


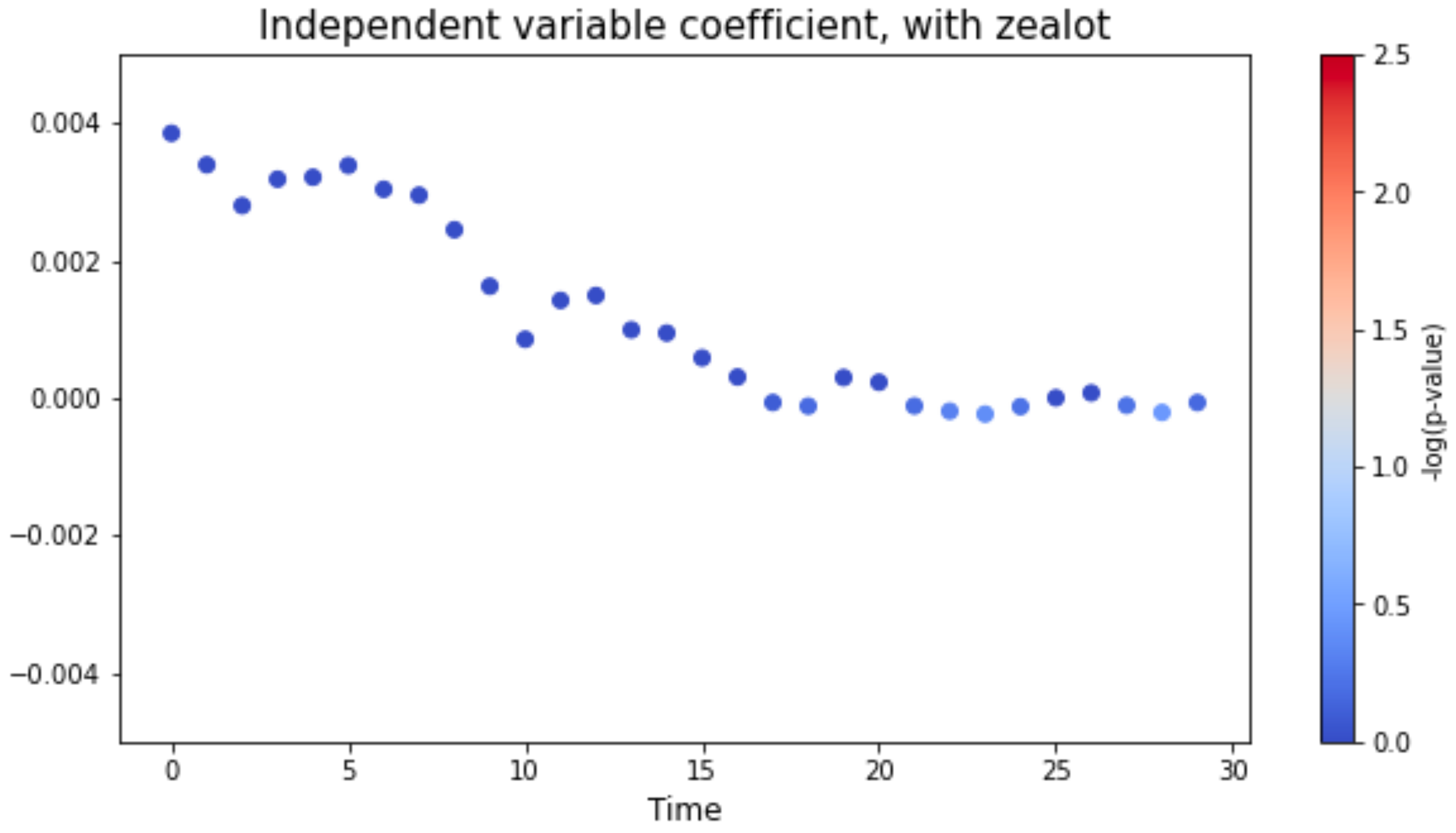


10 zealot with $o(z_i) = 0.9, \forall i \in Z$

Adaptive Network

Population Analysis — Clustering Effects

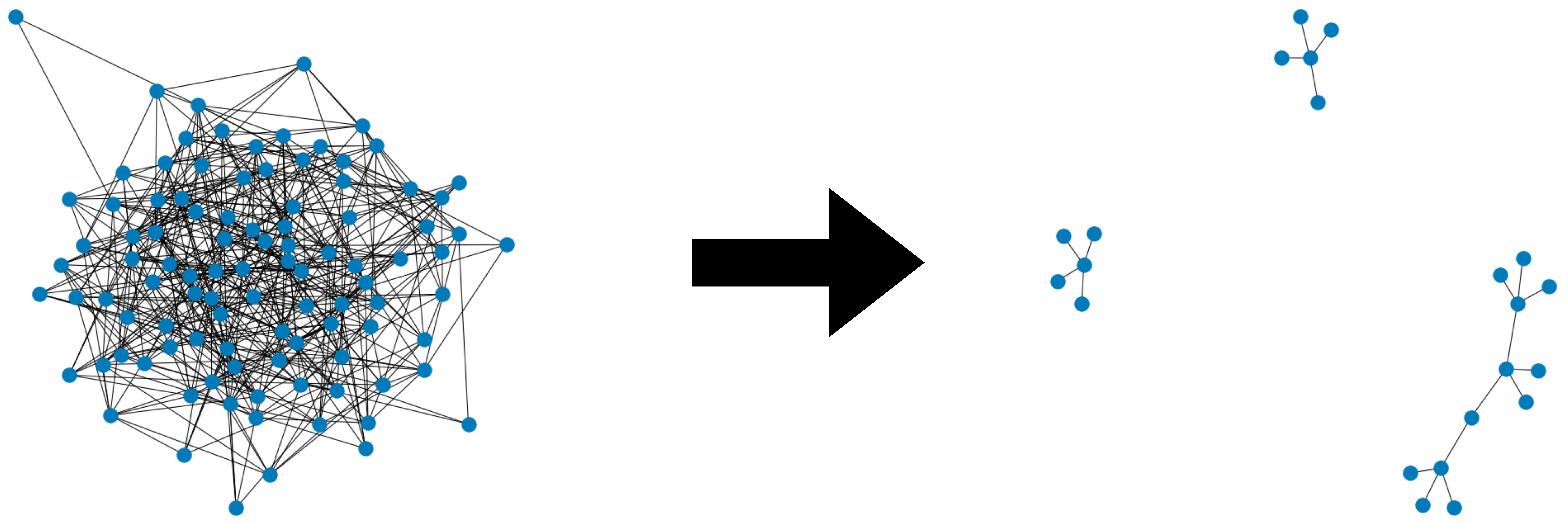




10 zealot with $o(z_i) = 0.9, \forall i \in Z$

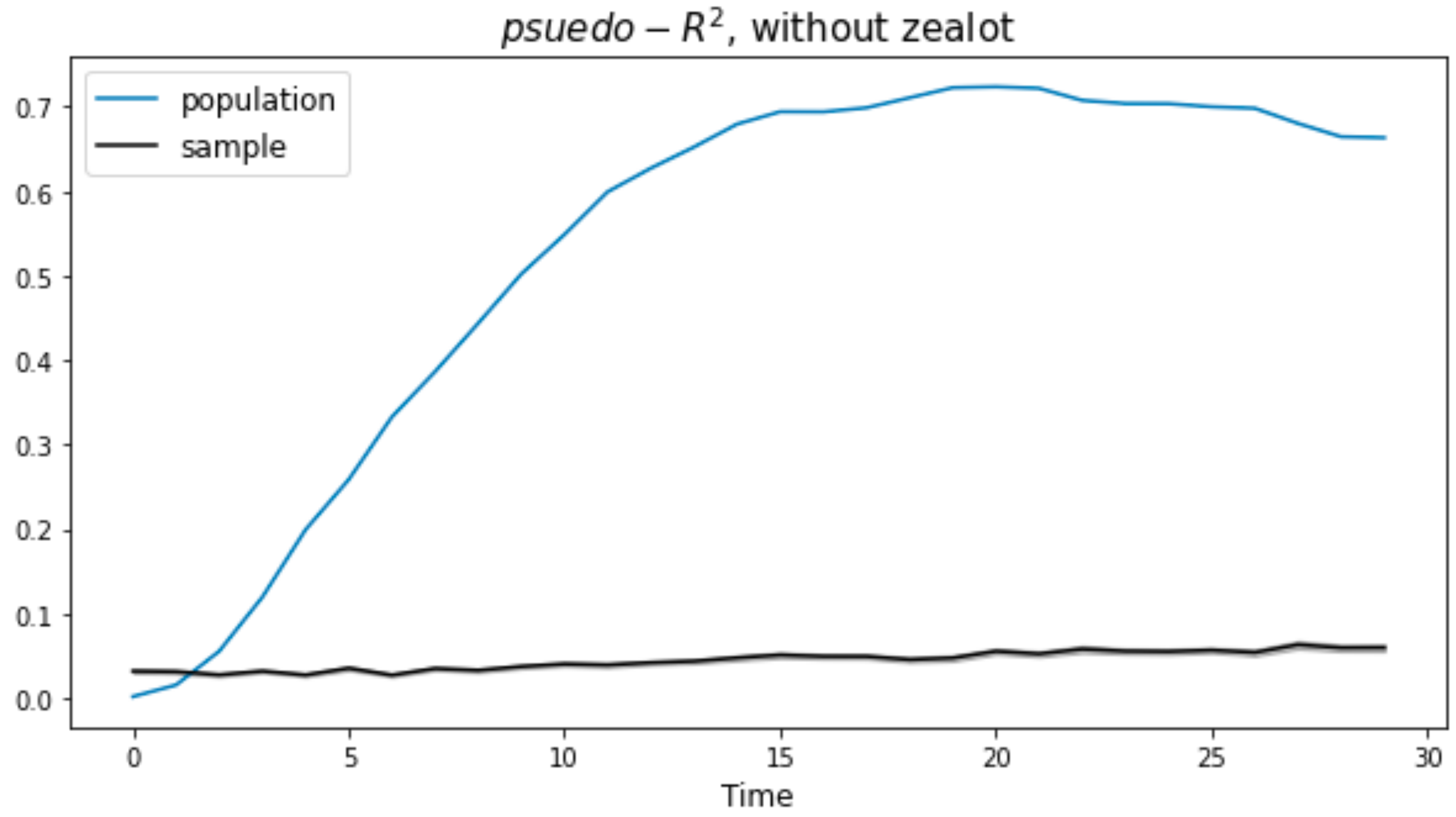
Adaptive Network

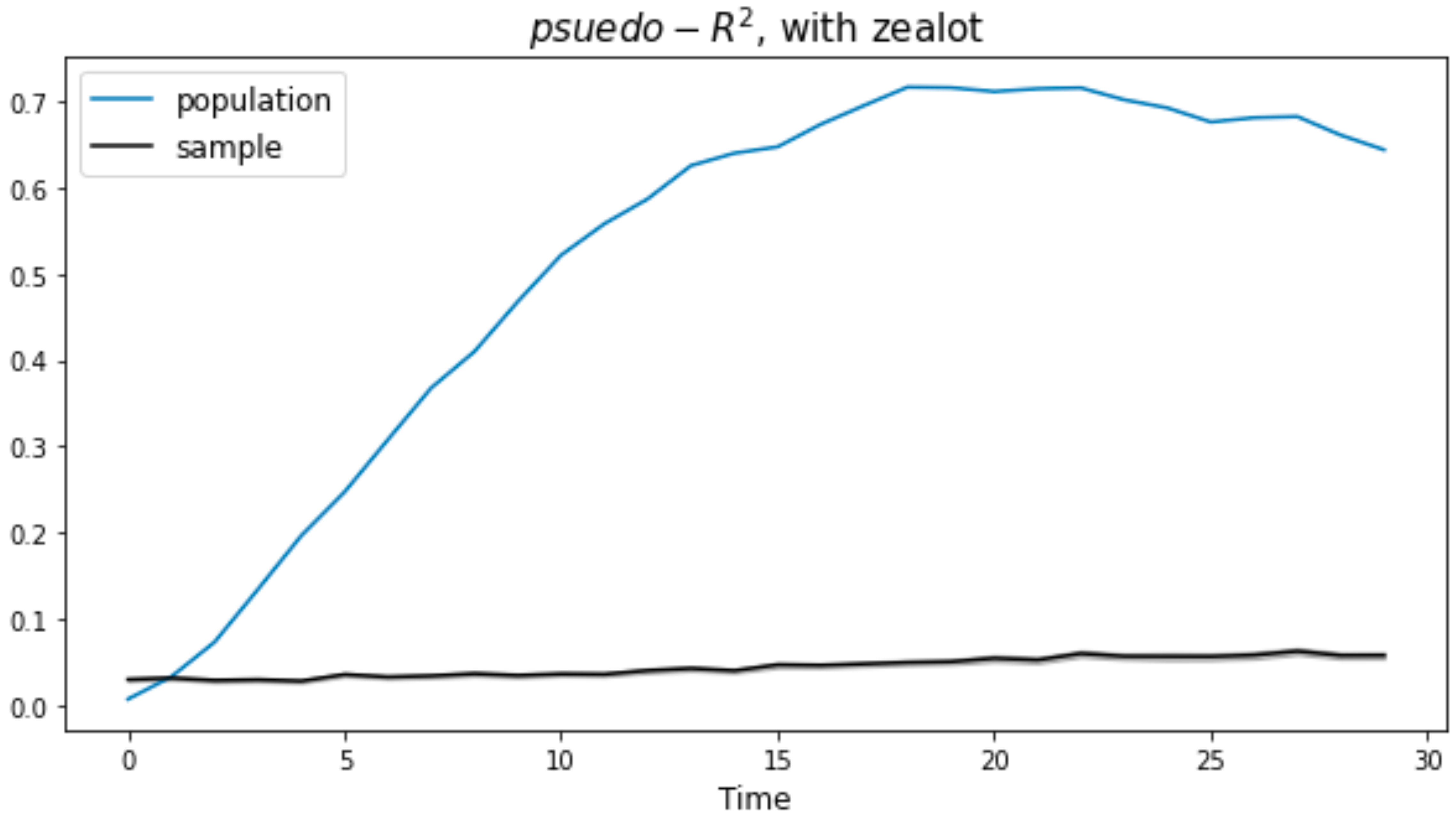
Snowball Sampling



Adaptive Network

Snowball Sampling Analysis — Goodness of Model Fitting

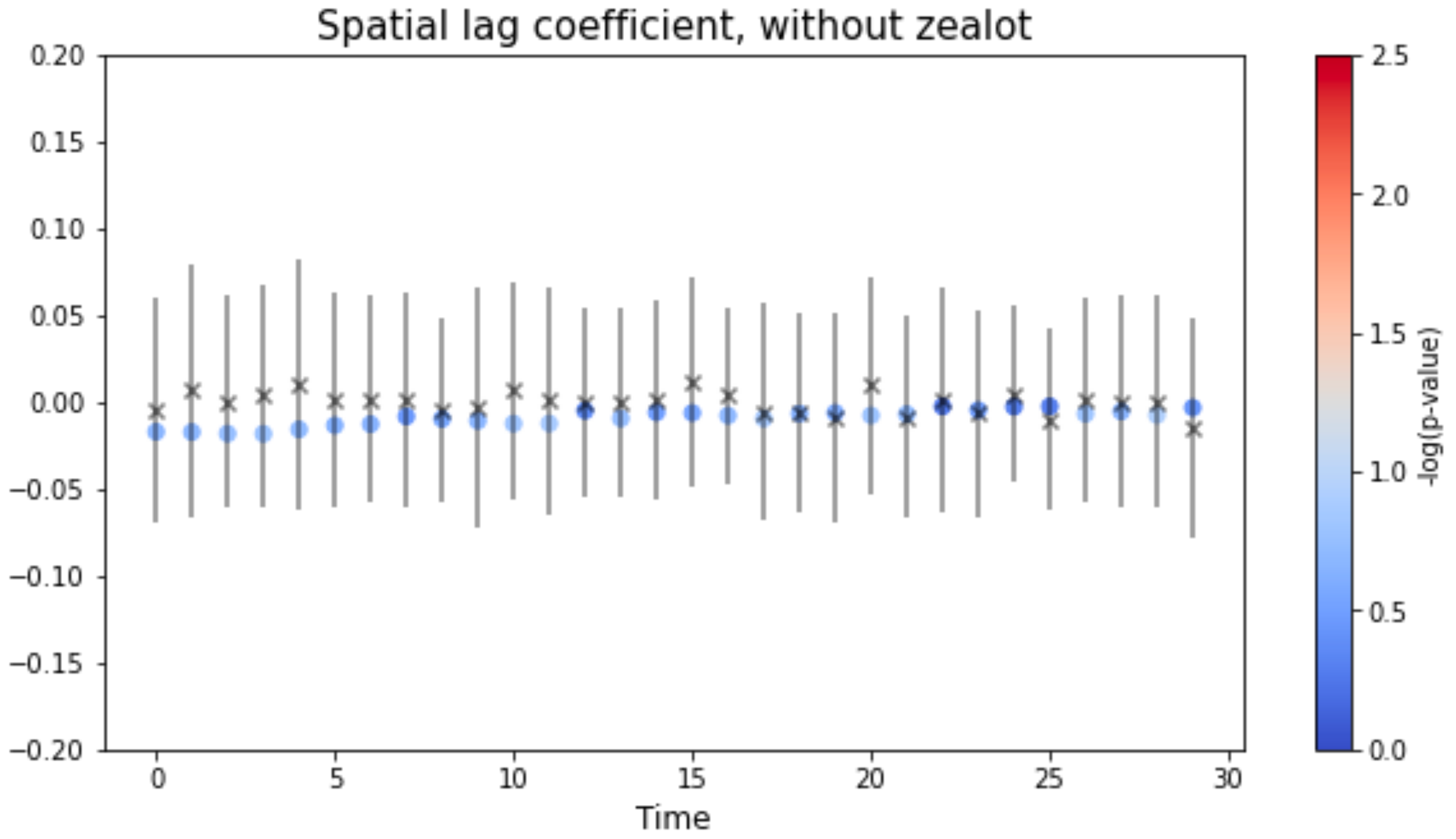


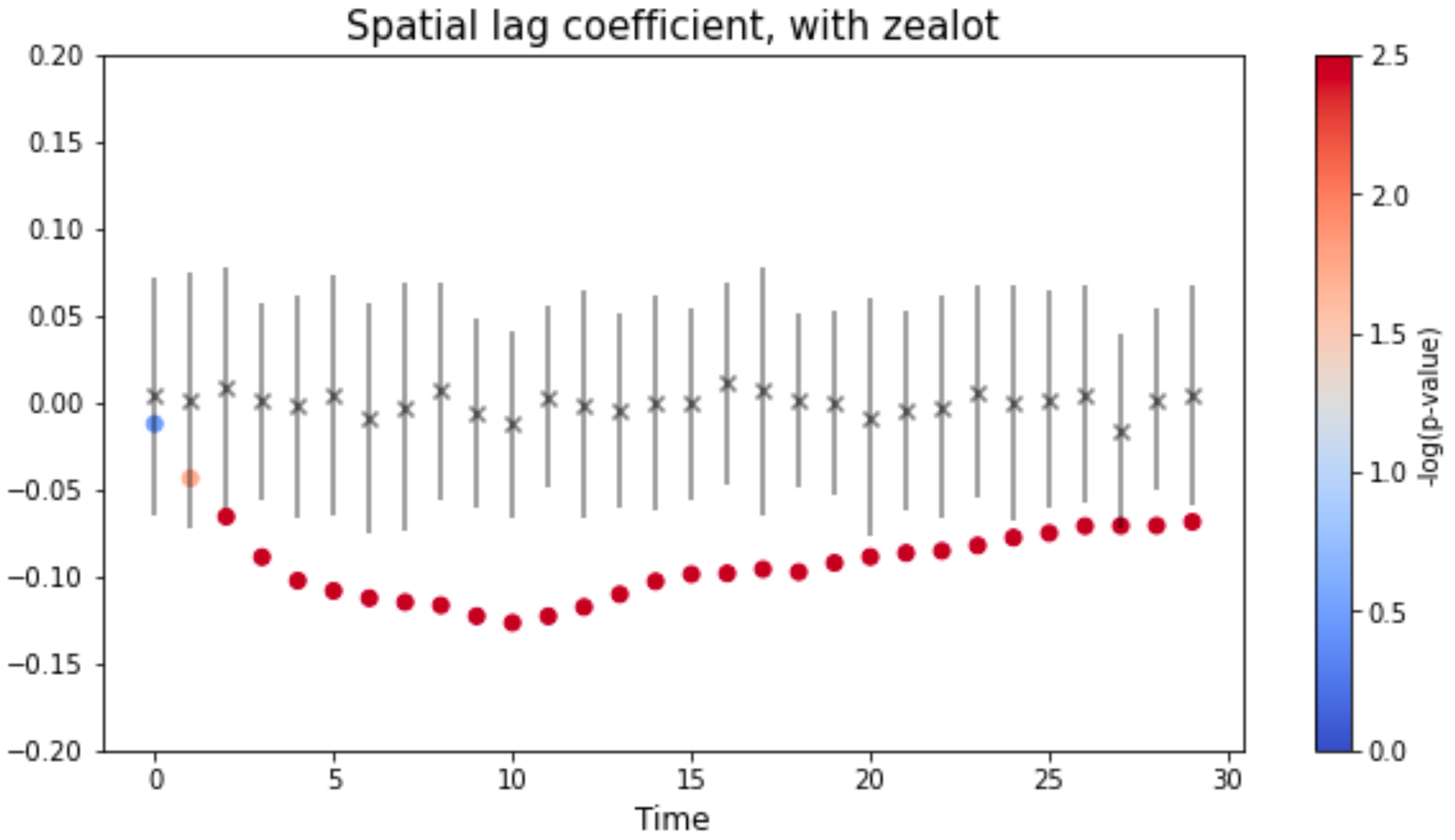


10 zealot with $o(z_i) = 0.9, \forall i \in Z$

Adaptive Network

Snowball Sampling Analysis — Peer Effects





10 zealot with $o(z_i) = 0.9, \forall i \in Z$

Adaptive Network

Population Analysis — Inference

- SAR model could infer peer effects in opinion dynamics
- Comparison between **spatial lag coefficients** suggests that **zealots induce rewiring and social learning**
- There is no significance on clustering effects
- Social learning is observed on **the network's global topology**

Don’t be fooled by your ego and belief.

Hold your faith with a good personality if you want to change something.

Toward Applied Microeconomic Theory Studies

Extensions in Theoretical & Empirical Studies

- **Microeconomic initiated concepts**
    1. Mechanism design (Renou, 2012)
    2. Bounded rationality (Mueller-Frank, 2013)
    3. Collective behavior (Acemoglu, 2014)
- **Empirical Studies**
    1. Social preferences (Hsieh, 2018)
    2. Altruistic TU game strategy (Leider, 2009)
- **Statistical/Econometric Approaches**
    1. Stochastic Actor Based model (Snijder, 2011)
    2. Unbiased Network Sampling (Hsieh, 2022)

Toward Applied Microeconomic Theory Studies

My work — An Advertisement

- Testing altruism in the core of social movements — On the egalitarian coalition-strategies
- On the biasness of the estimation for social behavior in dynamic networks

Share your ideas and suggestions with us!

陳柏安 (Bo-An Chen): seer2239@gmail.com

陳立凡 (Li-Fan Chen): koru1130@gmail.com

References

- DeGroot, (1974), *Journal of the American Statistical Association*, *69(345), 118-121.*
- Kan et al., (2023), *Journal of Complex Networks*, *Volume 11, Issue 1*
- Manski, (1993), *The Review of Economic Studies, Vol. 60, pp. 531–542.*
- Lee, (2004), *Econometrica, Vol. 72, pp. 1899–1925.*
- Renou & Tomala, T. (2012), *Theoretical Economics*, *7: 489-533.*
- Mueller-Frank, M. (2013), *Theoretical Economics*, *8: 1-40.*
- Acemoglu et al, (2014), *Theoretical Economics*, *9: 41-97.*
- Hsieh, C.-S. and van Kippersluis, H. (2018), *Quantitative Economics, 9: 825-863.*
- Leider et al, (2009), *The Quarterly Journal of Economics*, *Volume 124, Issue 4, November 2009, Pages 1815–1851*
- Snijder et al, (2011), *Social Networks*, *Volume 32, Issue 1, January 2010, Pages 44-60*
- Hsieh et al, (2022), ”Non-randomly Sample Networks: Biases and Corrections” (Revision requested by *Journal of Econometrics*)

Appendix: MLE for SAR

$$log(ML)_{SAR} = -\frac{N}{2}log(2\pi) - \frac{1}{2}log(det(\mathbf{V})) + log(det(\mathbf{S})) - \frac{1}{2}\epsilon^{\mathbf{T}}\mathbf{V}^{-\mathbf{1}}\epsilon$$

- N is the number of nodes (agents)
- $\mathbf{V}$ is the covariance matrix
- $\mathbf{S}$ is spatial correlation matrix $\mathbf{I}_{N\times N} - \rho\mathbf{W}$

Appendix: Clustering Coefficient

$$c_u = \frac{2T(u)}{deg(u)(deg(u) - 1)}$$

- $T(u)$ is the number of triangles through node u
- $deg(u)$ is the degree of u

NetworkX Documentation

Appendix: Pseudo-code for Adaptive DW model

Algorithm 1 Pseudocode for our Adaptive DW Model

parameters: $N, p, M, K, \alpha, \beta, C$

$t \leftarrow 0$; $G \leftarrow G(N, p)$
for $i \in G$.nodes() **do**
    $x_i(0) \leftarrow \text{Unif}[0, 1]$
end for
while (time $t <$ bail-out time) and (sum of the magnitudes of the opinion changes $<$ tol for fewer than 100 consecutive steps) **do**
    $E_d^\beta(t) \leftarrow \emptyset$ [initialize set of discordant edges]
    for (i, j) in the set of edges $E(t)$ **do**
        if $|x_i(t) - x_j(t)| < \beta$ **then** $E_d^\beta(t) \leftarrow E_d^\beta(t) \cup \{(i, j)\}$
        end if
    end for
    if $|E_d^\beta(t)| > M$ **then** select M edges uniformly at random from $E_d^\beta(t)$
    else select all edges from $E_d^\beta(t)$
    end if
    for each discordant edge (i, j) **do**
        dissolve and remove the edge from the edge set $E(t)$
        select node i or j with equal probability
        compute probabilities to rewire to other nodes using (3.1)
        randomly pick another node using the computed rewiring probabilities
        connect the node to the previously selected node with an edge; add the new edge to $E(t)$
    end for
    select K edges uniformly at random from $E(t)$
    for each selected edge (i, j) **do**
        if $|x_i(t) - x_j(t)| \leqslant C$ **then** update the opinions of the nodes using (3.2)
            [see the main text for further discussion]
        end if
    end for
    compute the sum of the magnitudes of the opinion changes
    $t \leftarrow t + \Delta t$
end while

Parameters:
$x_i(o) \in (-1,1)$
G: Erdős–Rényi Graph ~ G(1000, 0.1)
N: 1000
p: 0.1
M: 50
K: 30
α: 0.6
β: 0.6
C: 0.6

Kan+23