Supplementary Material

Breaking the Echo-Chamber Effect via Consent

Naomi Korem Tammar Shrot Hadassa Daltrophe

Overview

This supplementary document provides additional material supporting the results presented in the paper *Breaking the Echo-Chamber Effect via Consent*. The contents include simulation details, supplementary figures and tables, and further experimental analyses.

Contents

1 Simulation Methodology **2**
1.1 The *Blogger* dataset . . . 2
1.2 Synthetic Networks Generation . . . 2

2 Additional Experimental Results **3**
2.1 Consent-Driven Regulation intervention outcome . . . 3
2.2 Examining the selection of the Opt-in set . . . 3

3 Robustness Analysis **4**
3.1 Alternative Networks Structure . . . 4
3.2 Alternative Parameter Values . . . 4
3.2.1 Consent-Driven Regulation Function parameters . . . 4
3.3 User Response Function parameters . . . 5

4 Comparison Between Avin et al. and Our Model **7**

5 Data and Code Availability **7**

1 Simulation Methodology

1.1 The *Blogger* dataset

A visualization of the *Bloggers* network [1], as described in the Experimental Setup section (Section 4.1), is provided in Figure 1.

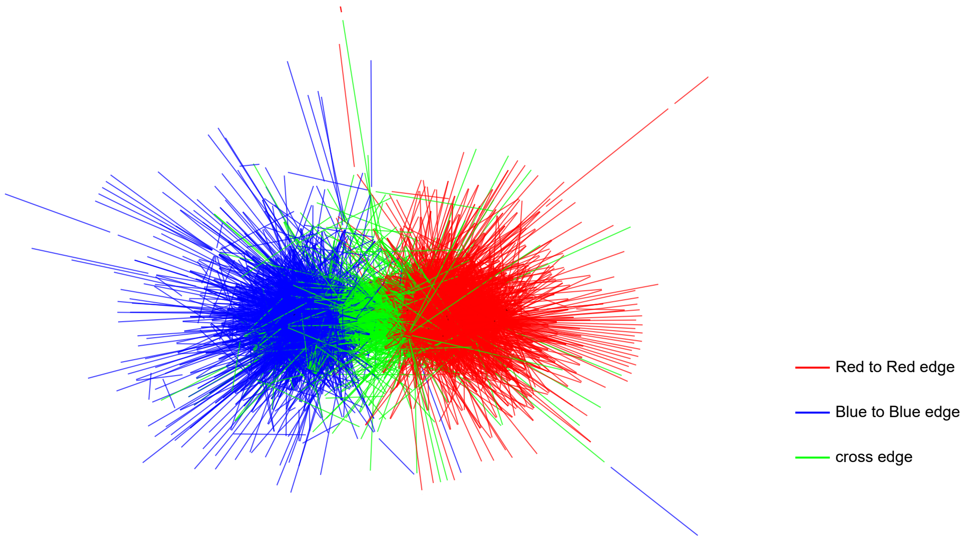

Figure 1: The *Bloggers* graph is a real-world, bi-populated social network derived from a 2005 study of U.S. political weblogs [1]. It consists of two distinct user groups: 52% labeled as `red` and 48% as `blue`. This network serves as the structural foundation for simulating the dynamics of information spreading in a social media platform.

1.2 Synthetic Networks Generation

In the Results section (Section 4), we described the construction of two bi-population networks—a heterophilic network and a homophilic network—using the Mixed Preferential Attachment (MPA) model [2]. The parameter settings used to generate each network are detailed below:

Heterophilic network $G(n, r, \delta, \pi)$ such that:

- G_0 is a complete 6 nodes, bipartite $\texttt{red} - \texttt{blue}$ graph.
- Total number of nodes: $n = 1222$
- Rate of red nodes $r = 0.5$
- Edges per step: $\delta = 14$
- Mixing matrix: $\pi = \begin{pmatrix} \pi_{\texttt{red}} \\ \pi_{\texttt{blue}} \end{pmatrix} = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$

Homophilic network $G(n, r, \delta, \pi)$ such that:

- G_0 is a complete 6 nodes, bipartite $\texttt{red} - \texttt{blue}$ graph.
- Total number of nodes: $n = 1222$
- Rate of red nodes $r = 0.5$
- Edges per step: $\delta = 14$
- Mixing matrix: $\pi = \begin{pmatrix} \pi_{\texttt{red}} \\ \pi_{\texttt{blue}} \end{pmatrix} = \begin{pmatrix} 0.999 & 0.001 \\ 0.001 & 0.999 \end{pmatrix}$

2 Additional Experimental Results

2.1 Consent-Driven Regulation intervention outcome

We intentionally refrain from reporting traditional symmetric uncertainty measures (such as standard confidence intervals) or standard null-hypothesis significance tests across our experimental configurations. Because the underlying user-response distributions are highly bimodal, with runs converging to either very low or very high engagement states, as clearly visible in the violin plots. Traditional interval metrics would misleadingly center around low-density intermediate values. Furthermore, because the sample size (the number of independent simulation runs) is programmatically defined, standard significance tests (p-values) can be rendered arbitrarily significant by scaling up run counts without reflecting meaningful physical effect sizes. We therefore present the full empirical distributions using violin plots and medians to faithfully preserve the dual-state characteristics of the system.

The distribution charts of the aggregated outcomes under the pure homophily regime $\mathcal{F}_P(p = 1, q = 0)$ are shown in Figure 2. The distribution charts of the aggregated outcomes under the pure homophily regime $\mathcal{F}_P(p = 0.999, q = 0.001)$ are shown in Figure 3.

2.2 Examining the selection of the Opt-in set

In the main paper, the opt-in set was selected uniformly at random. Interestingly, selecting the opt-in set based on user centrality measures does not substantially affect the results. Figure 4 presents the outcomes when the opt-in set consists of the users with the highest degree centrality, whereas Figure 5 shows the results when the opt-in set is composed of the users with the highest PageRank scores. Overall, the observed behavior remains qualitatively similar to that obtained under random selection.

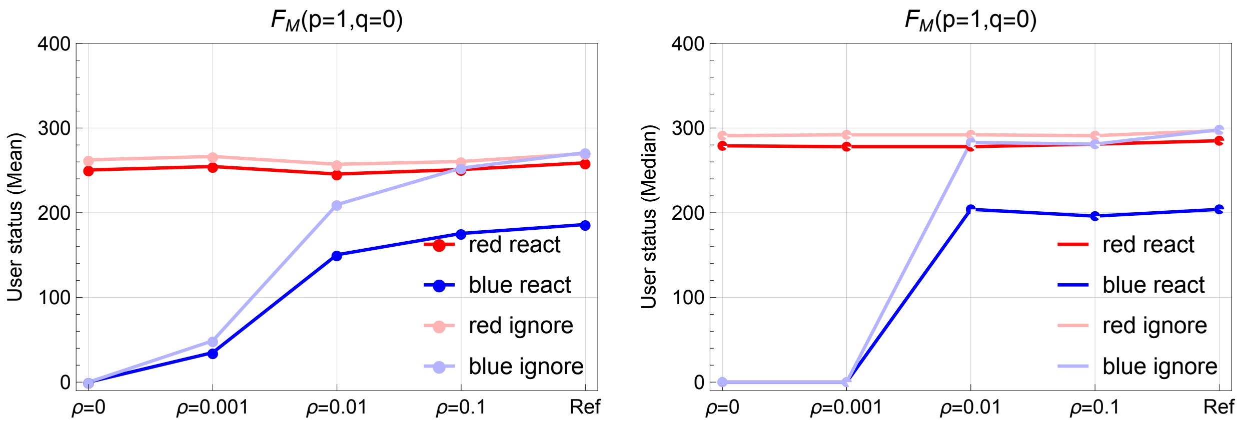

Figure 4: Opt-in set with the highest degree centrality. Consent-Driven Regulation intervention outcome expressed by the mean (left) and median (right) of the users' responses.

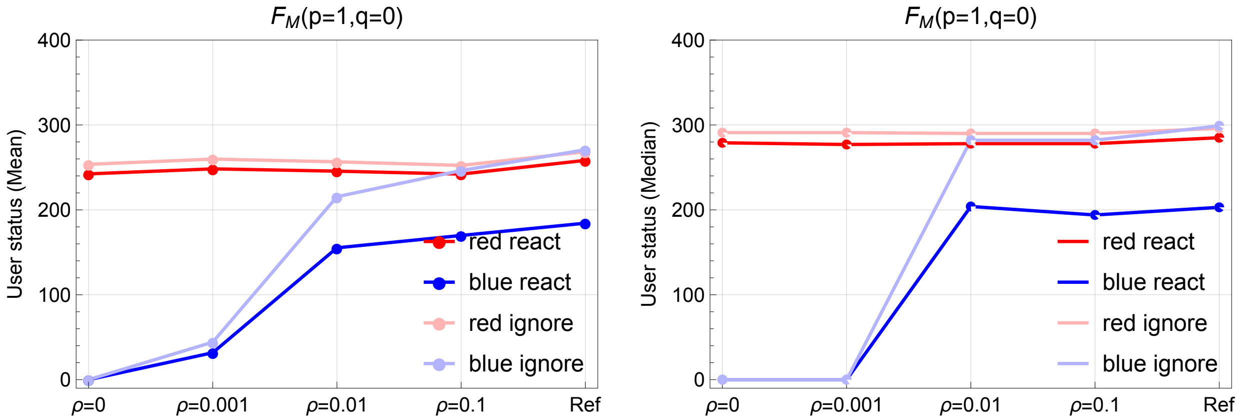

Figure 5: Opt-in set with the highest PageRank centrality. Consent-Driven Regulation intervention outcome expressed by the mean (left) and median (right) of the users' responses.

3 Robustness Analysis

3.1 Alternative Networks Structure

To ensure that our findings are not artifacts of the specific network topology used in the main experiments, we replicated the analysis on two additional networks topologies: one highly homophilic and one more heterophilic. We constructed both bi-population networks to have a similar number of users and links as the Bloggers network as described in Section 1.2. The distribution charts of the aggregated outcomes of the *hetrophilic network* under the *pure homophily* regime $\mathcal{F}_P(p = 1, q = 0)$ are shown in Figure 6. The distribution charts of the aggregated outcomes of the *homophilic* network under the *relaxed* homophily regime $\mathcal{F}_P(p = 0.999, q = 0.001)$ are shown in Figure 7.

3.2 Alternative Parameter Values

3.2.1 Consent-Driven Regulation Function parameters

Recall that the Consent-Driven Regulation is defined as:

$$\mathcal{F}_R(N, \mathcal{S}_{t-1}, P_t, Q, \rho) = P_t \cup Q_t,$$

where $Q_t \subseteq Q$ is a set of additional users selected uniformly at random from the set Q, and $|Q_t| = \lceil \rho \cdot |P_t| \rceil$, such that $\rho \in [0, 1]$ is the regulation parameter. Figure 8 shows the Consent-Driven Regulation intervention outcome for additional ρ values (related to the paper).

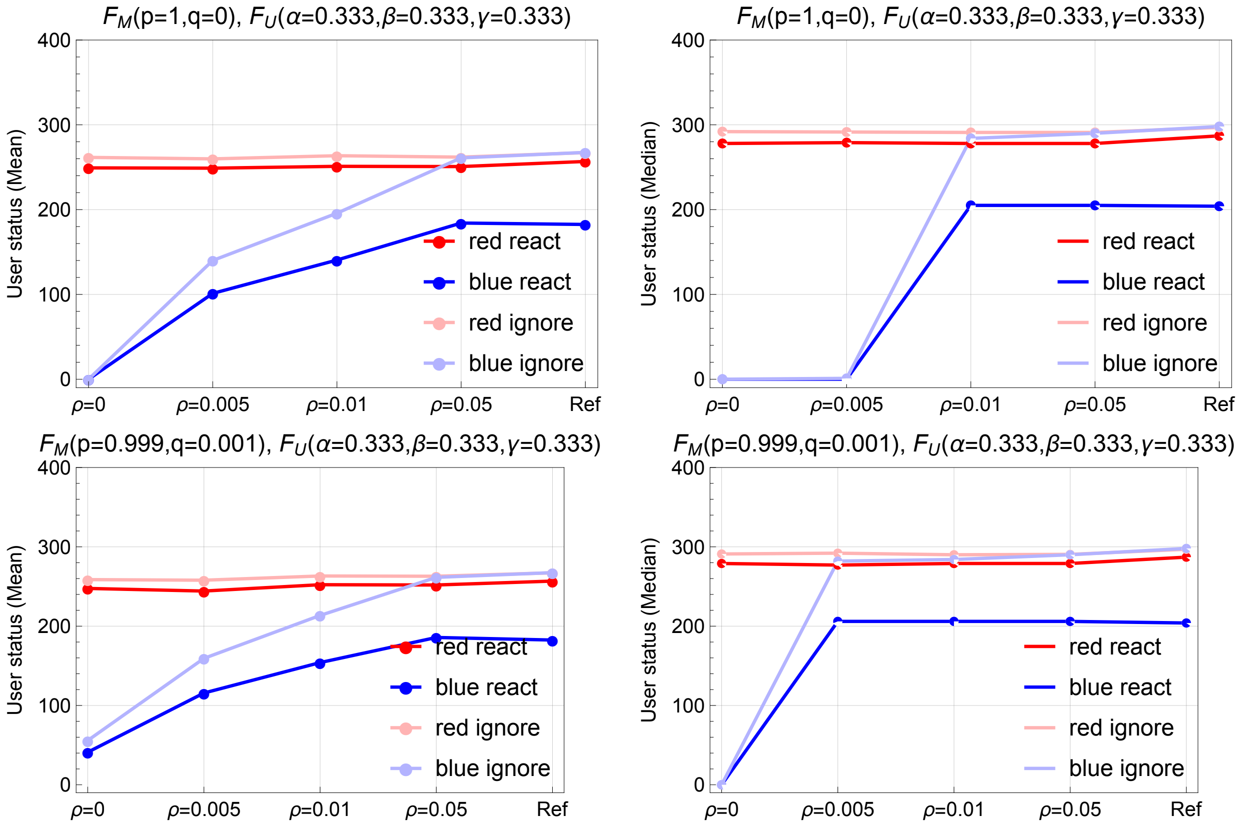

Figure 8: Consent-Driven Regulation outcome expressed by the mean (left) and median (right) of the users' responses for $\rho = 0, .0.005, 0.01$ and 0.05.

3.3 User Response Function parameters

Recall that the User Response Function $\mathcal{F}_U$ is defined as:

$$s(u) = \mathcal{F}_U(N, \mathcal{S}_{t-1}, v, u) = \begin{cases} \texttt{react} & \text{with probability } \sigma\Big(\alpha \cdot g(u) + \beta \cdot f(v,u) + \gamma \cdot h\Big) \\ \texttt{ignore} & \text{otherwise} \end{cases}$$

Here, α, β, and γ are weight parameters that determine the relative influence of the color-dependent function g, the friendship function f, and the virality function h, respectively. These parameters satisfy $\alpha + \beta + \gamma \leq 1$, thereby defining a valid probability distribution.

Weight parameters

We investigate the effect of different weight configurations on the model's behavior. Figure 9 presents the Consent-Driven Regulation outcomes when the color (opinion) component is omitted from the user response function, namely, $\alpha = 0$, $\beta = 0.5$, and $\gamma = 0.5$. Figure 10 shows the results obtained when the friendship component is omitted ($\alpha = 0.5$, $\beta = 0$, and $\gamma = 0.5$), while Figure 11 presents the outcomes when the virality component is omitted ($\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 0$).

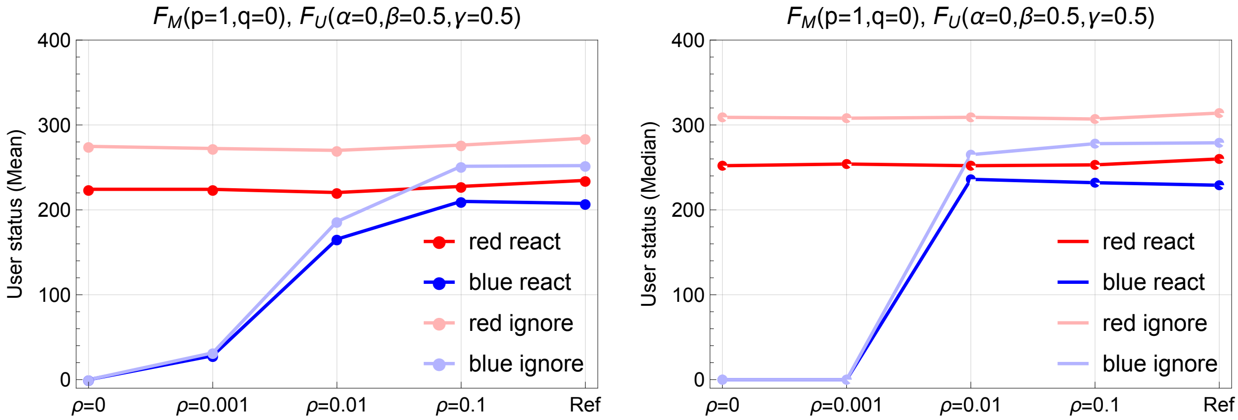

Figure 9: User Response Function based on friendship and virality only (excluding the color component).

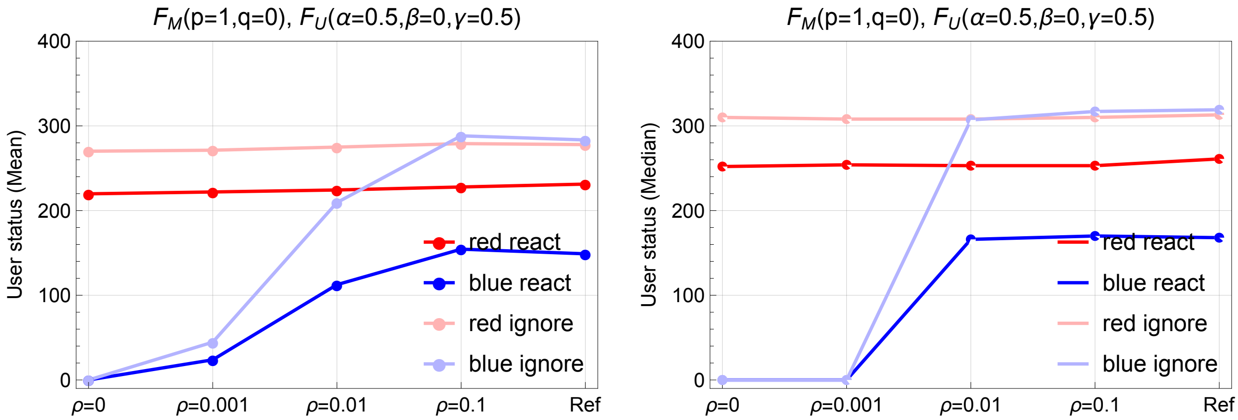

Figure 10: User Response Function based on color (opinion) and virality only (excluding the friendship component).

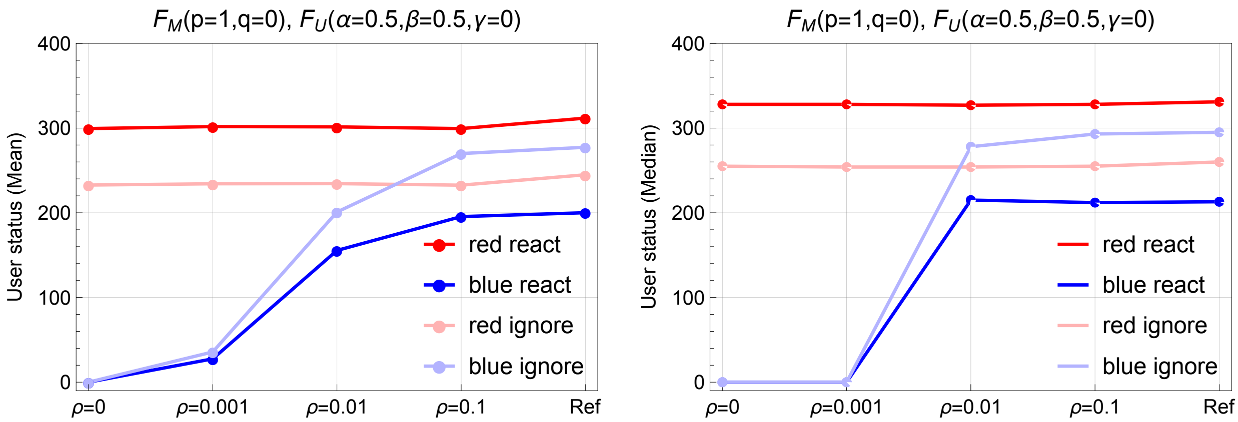

Figure 11: User Response Function based on color (opinion) and friendship only (excluding the virality component).

Sigmoid bias.

In the User Response Function $\mathcal{F}_U$, the function $\sigma(x) = \frac{1}{1+e^{-(x-3/4)}}$ is the Sigmoid function, modeling the stochastic nature of the response. The bias term in the main paper is set to $-3/4$ to shift the baseline likelihood of a `react` response to reflect the low probability of `react` related to `ignore`. Figure 12 shows the median of the Consent-Driven Regulation intervention outcome of the users' responses for bias term$=-1/2, -3/4$ and -1.

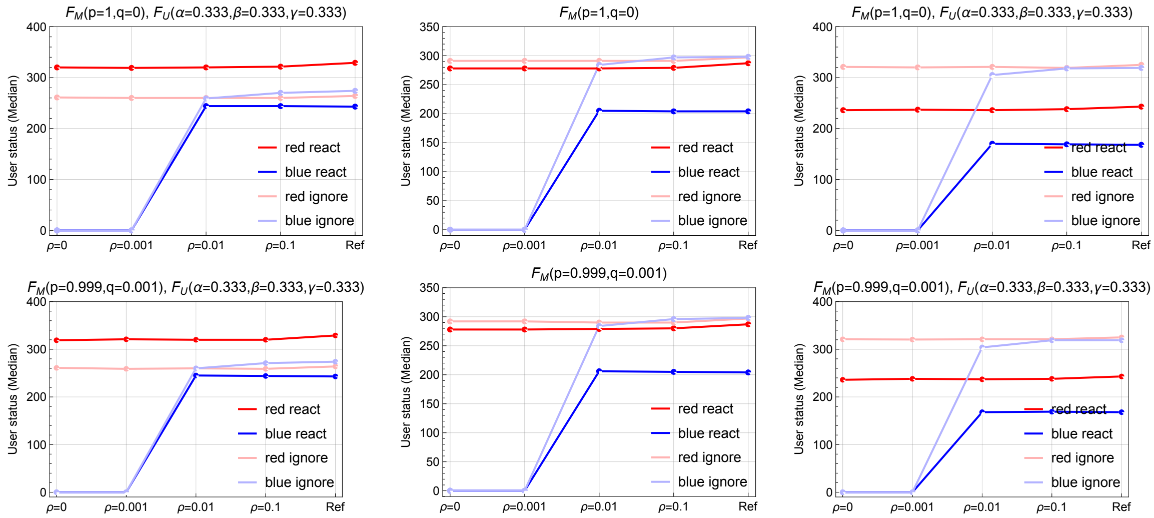

Figure 12: User Response Function using bias term of $-1/2, -3/4$, and -1 (left to right), for Pure (top) and relaxed (bottom) homophily.

4 Comparison Between Avin et al. and Our Model

Table 1: Comparison of the spreading processes models.

	Avin et al. Model [3]	**Our Model**
Social network definition	$N = (G, c)$	$N = (G, c, s)$
	$G = (V, E)$	$G = (V, E)$
	$c : V \to \{$`red`, `blue`$\}$	$c : V \to \{$`red`, `blue`$\}$
	$\{$`inactive`, `agree`, `disagree`, `ignore`$\}$	$s : V \to \{$`inactive`, `react`, `ignore`$\}$
$\mathcal{F}_U$	$\mathcal{F}_U(g)$ (color-based only)	$\mathcal{F}_U(g, f, h)$ (based on color, friendship and virality)
$\mathcal{F}_R$	$\mathcal{F}_R(\cdot)$	$\mathcal{F}_R(\cdot, Q)$, where Q is the opt-in set
Reference point	Ratio of `red` and `blue` users in the network.	Uniform platform spreading mode in the network.

5 Data and Code Availability

The code and data supporting the findings of this study are available from the corresponding author, H. Daltrophe, upon reasonable request at `hadasda1@sce.ac.il`.

References

[1] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43, 2005.

[2] Chen Avin, Hadassa Daltrophe, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg, and Yvonne-Anne Pignolet. Mixed preferential attachment model: Homophily and minorities in social networks. *Physica A: Statistical Mechanics and its Applications*, 555:124723, 2020.

[3] Chen Avin, Hadassa Daltrophe, and Zvi Lotker. On the impossibility of breaking the echo chamber effect in social media using regulation. *Scientific Reports*, 14(1):1107, 2024.

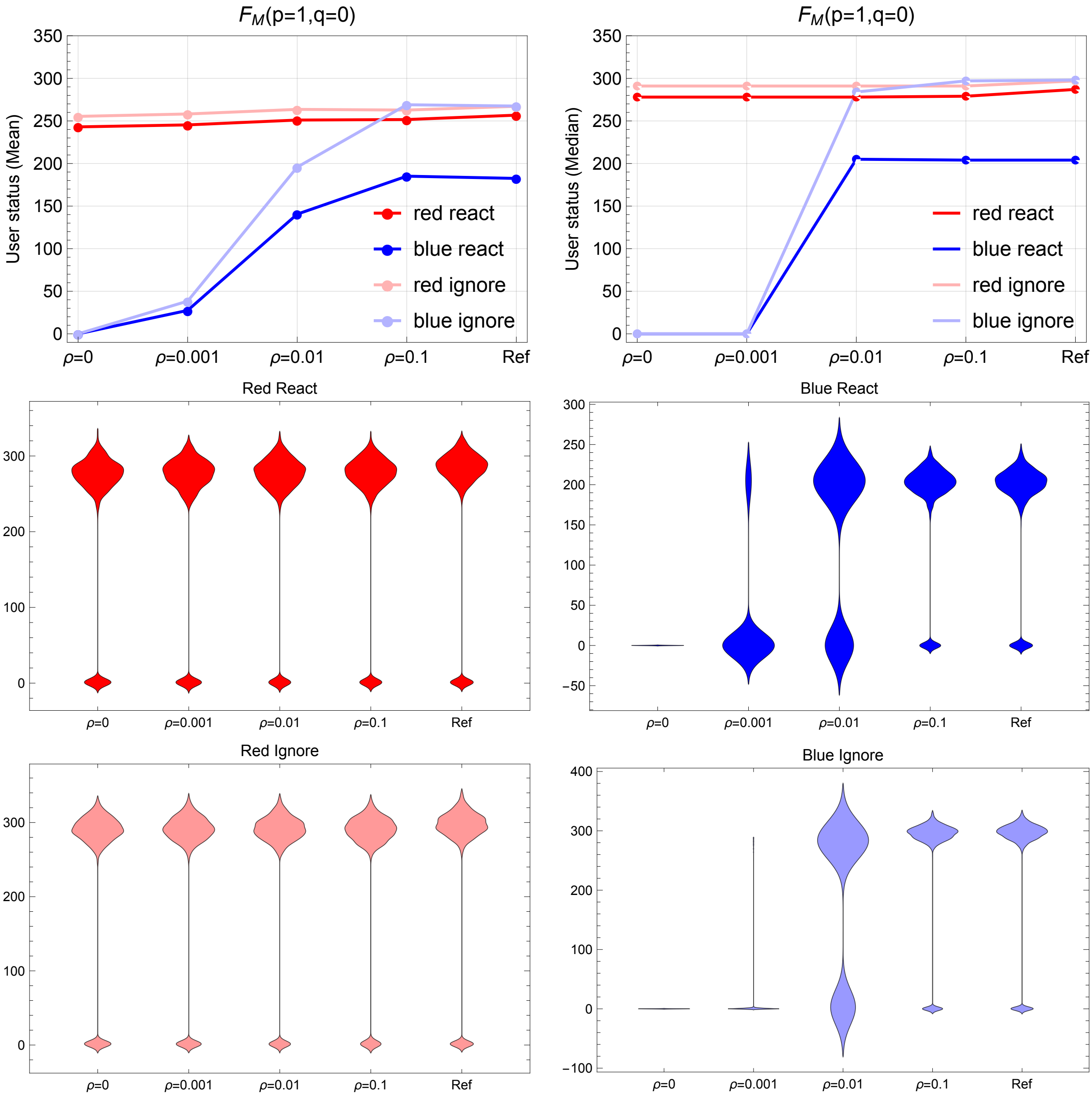

Figure 2: Consent-Driven Regulation intervention outcome of the users' responses (`red react`, `blue react`, `red ignore` and `blue ignore`) for different ρ values, under pure homophily $\mathcal{F}_P(p = 1, q = 0)$, where messages are transmitted to opposite-color neighbors with a small probability. The underlying user-response distribution is bimodal, exhibiting either very small or very large values.

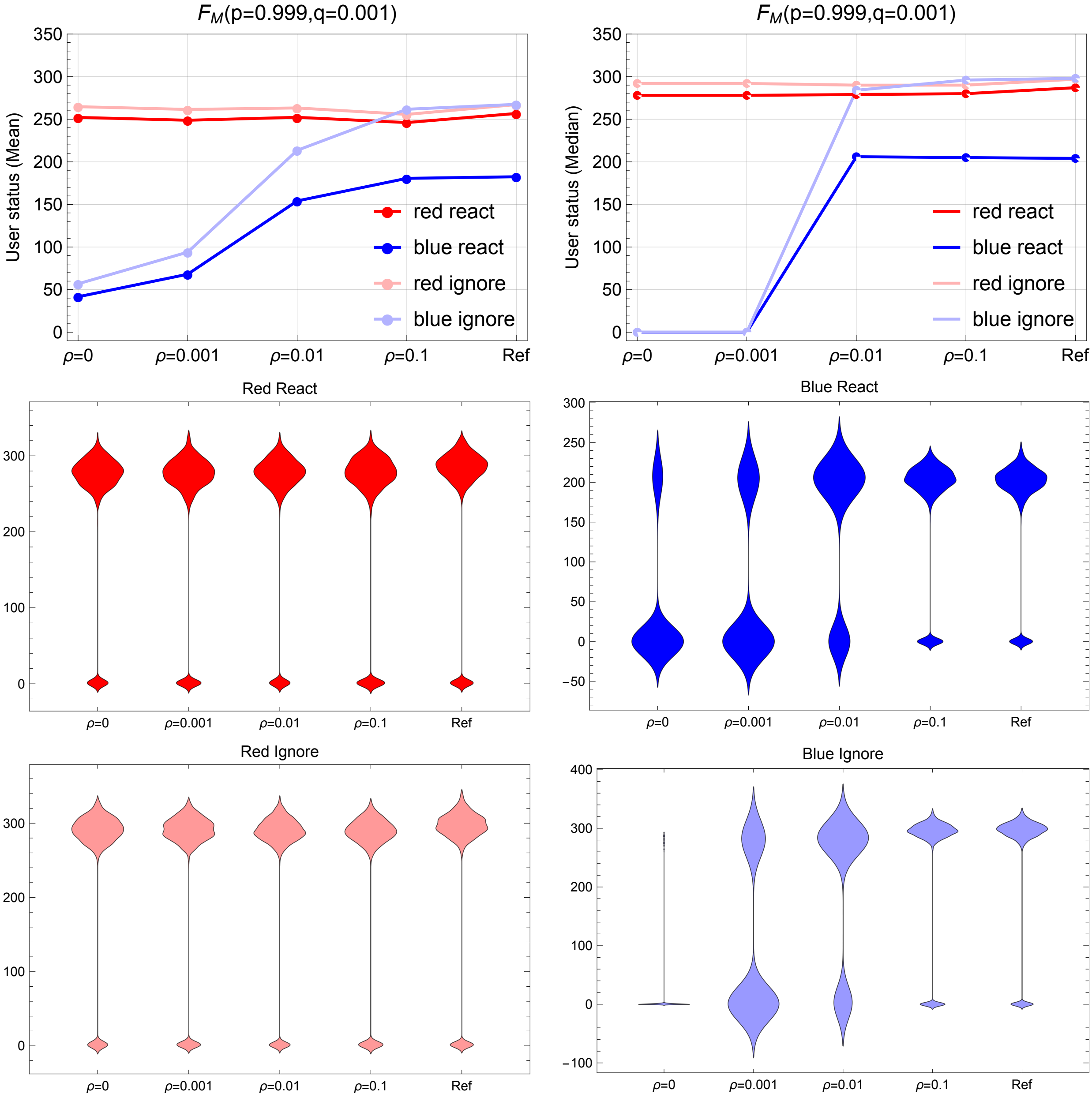

Figure 3: Consent-Driven Regulation intervention outcome of the users' responses (`red react`, `blue react`, `red ignore` and `blue ignore`) for different ρ values, under relaxed homophily $\mathcal{F}_P(p = 0.999, q = 0.001)$, where messages are transmitted to opposite-color neighbors with a small probability. The underlying user-response distribution is bimodal, exhibiting either very small or very large values.

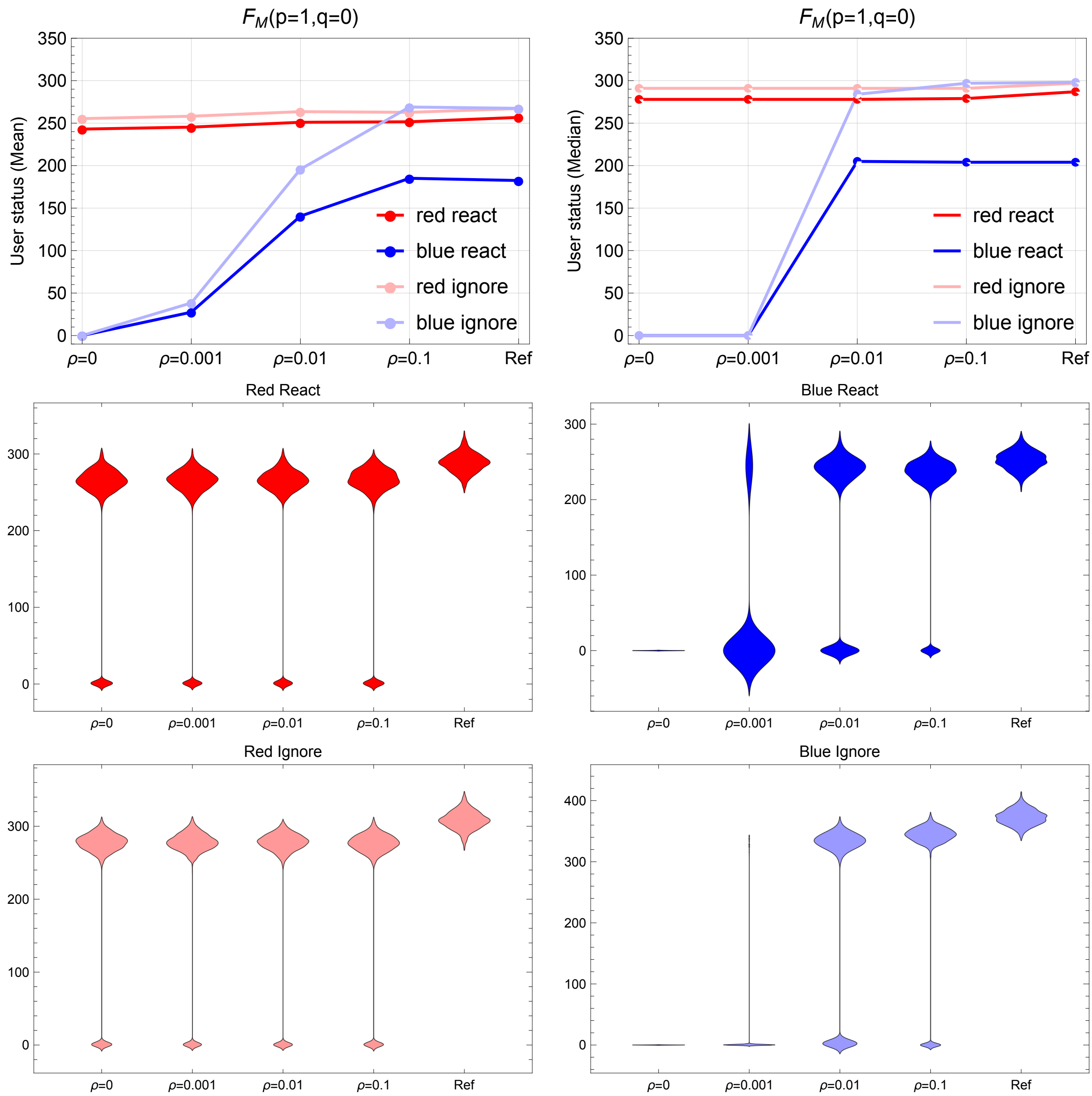

Figure 6: Hetrophilic network, under pure homophily. The underlying user-response distribution is bimodal, exhibiting either very small or very large values.

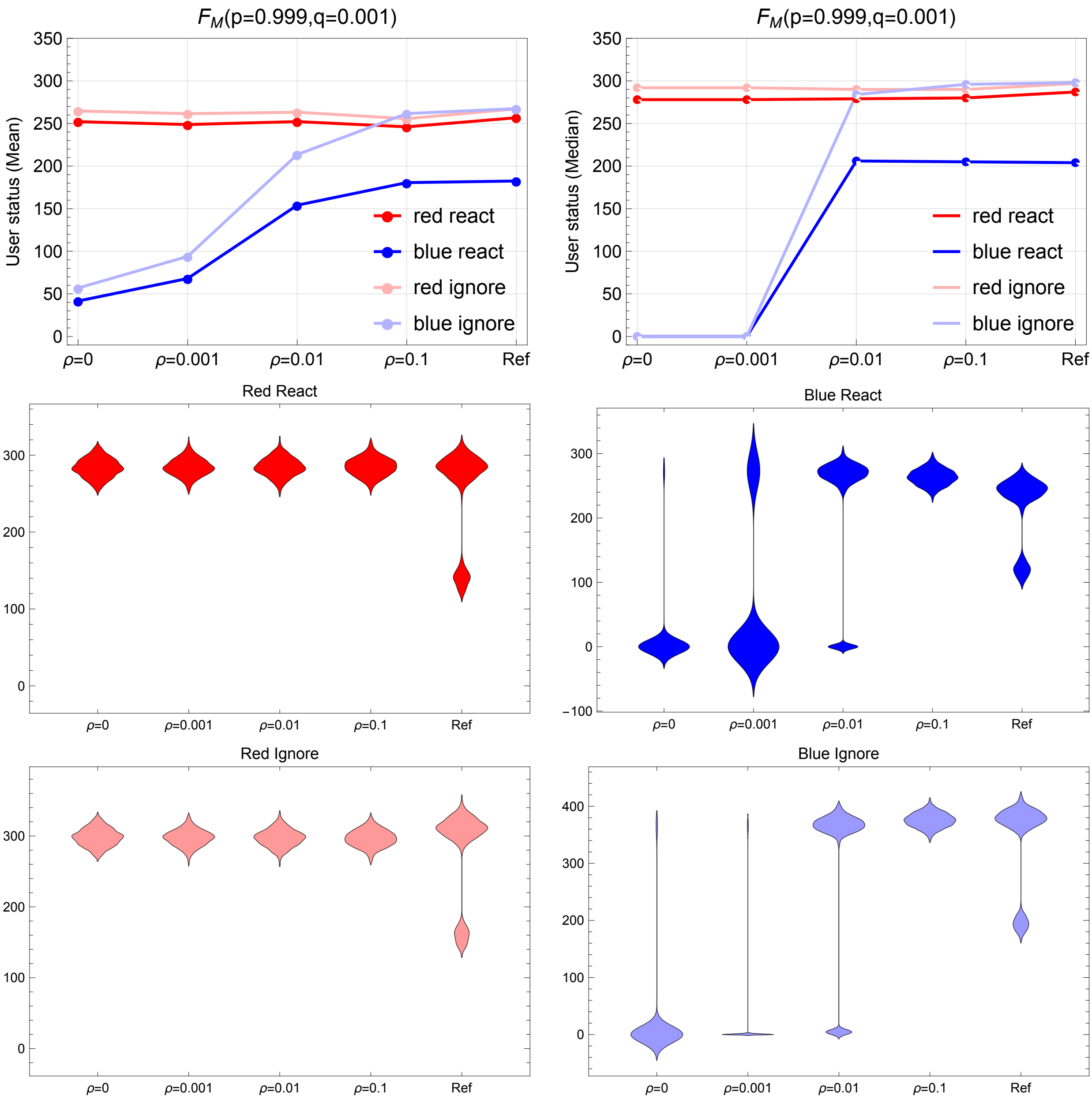

Figure 7: Homophilic network under relaxed homophily. The underlying user-response distribution is bimodal, exhibiting either very small or very large values.