# Who Gets the Blame? A Structural Theory of Scapegoating in Social Networks

*Keywords: scapegoating, social networks, opinion dynamics, game theory, network centrality*

## Extended Abstract

**Introduction.** Scapegoating, the act of shifting blame onto innocent individuals or groups, has been a persistent feature of social life across historical, political, and organizational contexts. From witch hunts and ethnic persecutions to modern-day corporate scandals and immigrant blame games during economic crises, scapegoating frequently targets marginalized individuals.

While existing literature in social science largely emphasizes *identity-based* attributes [1–4], focusing on who people are (e.g., gender, ethnicity, religion, partisanship, nationality), this study proposes a *network-based* explanation, focusing on where individuals are positioned within social networks. A leader facing systemic failure strategically decides whether to self-blame or scapegoat, while agents embedded in a network update beliefs and form public opinions through social learning. Game-theoretic analysis shows that the leader exploits structural vulnerabilities, systematically targeting peripheral nodes with lower centrality. The findings offer insights into the structural mechanism driving inequality and social injustice in interconnected communities.

**Model Setup.** The model considers a leader and $n$ agents embedded in an undirected graph $G(V,E)$. All agents are innocent, but after a systemic failure, the leader must decide whether to self-blame at cost $C$ or scapegoat some agent $k \in V$. Each agent $i$ holds prior beliefs $\pi_i^{(k)}$ about $k$'s guilt, and upon accusation, they update beliefs using the Bayesian formula as $b_i^{(k)} = \frac{\pi_i^{(k)}}{\pi_i^{(k)} + p(1-\pi_i^{(k)})}$, where $p \in [0,1]$ is the common perceived probability of the leader prosecuting an innocent agent. Thus, lower $p$ indicates stronger public trust in the leader.

Agents balance posterior fidelity and peer conformity based on a DeGroot-style utility function [5–7]: $u_i^{(k)}(x_i^{(k)}, x_{-i}^{(k)}) = -(x_i^{(k)} - b_i^{(k)})^2 - \sum_{j=1}^n A_{ij}(x_i^{(k)} - x_j^{(k)})^2$, and the opinion dynamics subgame admits a unique Nash equilibrium. Taking the first-order condition yields $x^{*(k)} = (I+L)^{-1} b^{(k)}$, where $L$ is the Laplacian of $G$.

Scapegoating succeeds if the collective opinion $\bar{x}^{(k)} = \frac{1}{n}\sum_i x_i^{*(k)} \geq \frac{1}{2}$. The leader's payoff is $-C$ if self-blaming and $-R(k)$ if scapegoating, with reputational cost $R(k) = \sum_{i=1}^n (1 - x_i^{*(k)})$. Thus, the leader scapegoats if $R(k) \leq C$ for some $k$ in perfect Bayesian equilibria (PBE).

**Network Effects and Structural Vulnerability.** The primary contribution lies in analyzing how different belief structures, tied to network position, affect scapegoating feasibility. Within a unified structural framework, I examine two canonical specifications of network effects: discrete and decay. Illustrations are provided in Figure 1.

In the discrete case, neighbors of $k$ treat $k$ as innocent while non-neighbors remain skeptical. Formally, $\pi_i^{(k)} = \begin{cases} 0, & i \in N(k), \\ \frac{1}{2}, & i \notin N(k). \end{cases}$ The reputational cost then depends on the degree of $k$, and the threshold for scapegoating (i.e., the largest possible level of public distrust) reduces to $p^{*(k)} = \frac{C - \deg(k)}{n - C}$. This result shows that scapegoats are those with the lowest degree centrality. Peripheral nodes are attractive targets because their limited connectivity prevents them from mobilizing sufficient resistance, thereby minimizing reputational penalties for the leader.

In the decay case, skepticism increases smoothly with distance. For geodesic length $l_{ik}$, priors are defined as $\pi_i^{(k)} = \frac{1}{2} - \left(\frac{1}{2}\right)^{l_{ik}}$. The derived scapegoating threshold is $p^{*(k)} = D_k^{-1}(C)$, where $D_k$ is a convex transformation of decay centrality. Hence the leader scapegoats the node with weakest global connectivity. Even moderately connected nodes may be scapegoated if they are structurally distant from the broader network. This illustrates that both local degree and global reach matter: marginalization can arise not only from few ties but also from isolation in the global structure.

**Ongoing Experiment.** To explore the model's implications, I am currently designing an individual experiment that empirically tests how blame attribution varies with network position under simulated crisis scenarios. Participants will assume the role of a leader, evaluate potential scapegoat targets in different positions, and indicate both their own choices and how they expect others to respond. The design further varies political trust as a moderator and examines perceived influence as a mediator, thereby linking formal modeling with behavioral evidence.

**Conclusion.** Overall, this interdisciplinary research integrates insights from social psychology, political science, and network economics. The findings contribute to understanding the structural roots of inequality in blame-shifting beyond identity-based biases. Prevention strategies that address structural vulnerability, such as enhancing transparency, strengthening social ties, and promoting inclusion, can effectively reduce the risk of unfair blame.

# References

[1] Gordon W Allport. "The nature of prejudice". In: *Reading/Addison-Wesley* (1954).

[2] Henri Tajfel et al. "An integrative theory of intergroup conflict". In: *Intergroup relations: Essential readings* (2001), pp. 94–109.

[3] Yann Bramoullé and Pauline Morault. "Violence against rich ethnic minorities: A theory of instrumental scapegoating". In: *Economica* 88.351 (2021), pp. 724–754.

[4] Daniel J Hopkins. "Politicized places: Explaining where and when immigrants provoke local opposition". In: *American political science review* 104.1 (2010), pp. 40–60.

[5] Morris H DeGroot. "Reaching a consensus". In: *Journal of the American Statistical association* 69.345 (1974), pp. 118–121.

[6] David Bindel, Jon Kleinberg, and Sigal Oren. "How bad is forming your own opinion?" In: *Games and Economic Behavior* 92 (2015), pp. 248–265.

[7] Javad Ghaderi and Rayadurgam Srikant. "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate". In: *Automatica* 50.12 (2014), pp. 3209–3215.

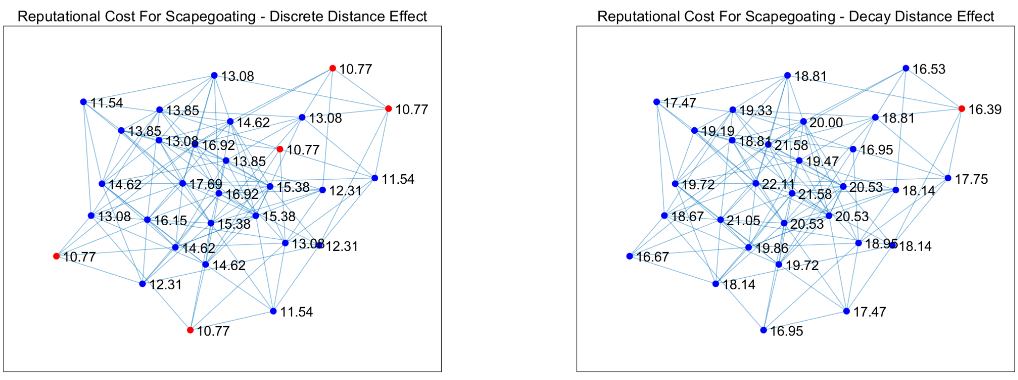


Figure 1: Scapegoat selection with $n = 30$, $p = 0.3$ (red nodes as predicted scapegoats).