“Ding-Dong! The Wicked Witch Is Dead!”: A Model of Scapegoating in Social Networks

Qinyang Yu (MSEC’ 26)

ECON 605 Term Paper

December 12, 2024

“The best way to bring folks together is to give them a real good enemy.”

— The Wonderful Wizard of Oz, in *Wicked*

Abstract

This paper studies scapegoating, the phenomenon of blame-shifting to innocent individuals or groups in a community. We consider a networked game with strategic political control and opinion dynamics. The model explores how network structure influences scapegoat selection, suggests that a leader with higher political trust identifies scapegoats easier, and explains why marginalized agents are more vulnerable to scapegoating through graph centrality measures. Ultimately, this research offers insights into the mechanisms driving inequality and social injustice in interconnected communities.

Keywords: Scapegoating, Social Networks, Game Theory, Political And Social Trust, Opinion Dynamics, Marginalization, Social Justice

1 Introduction

The phenomenon of scapegoating, where individuals or groups are unfairly blamed for problems they did not cause, has existed throughout human history and remains pervasive in modern society. The act of identifying a scapegoat serves a variety of purposes, such as restoring social order, transferring internal conflict, or shifting blame away from those in positions of authority. Despite its long-standing presence in human behavior, scapegoating remains a complex and multifaceted topic, influencing both interpersonal interactions and the larger dynamics of social trust and political power.

The universality of scapegoating is evident in a range of historical events, from religious persecutions and ethnic scapegoating in nationalist policies to more contemporary examples, such as political blame games, corporate scapegoating, economic scapegoating during financial crises (e.g., immigrants being blamed for unemployment), and even within smaller social networks. In each of these contexts, the allocation of blame often reflects underlying social structures, such as hierarchies, group dynamics, and the distribution of information. Whether in small communities or large institutions, scapegoating is closely related to the ways in which individuals form beliefs about each other, interact within the social networks, and contribute to social opinion or reputation.

In the modern world, the growth of complex networks has made scapegoating even more pertinent. With the rise of digital platforms and the interconnectedness of global society, the ways in which individuals and groups communicate and form beliefs about each other have been transformed. The speed and breadth of information flow, combined with the increasingly networked nature of society, raise questions about how people make decisions, assign blame, and manage social relationships in a world where trust and information are often dispersed and fragmented. Understanding how scapegoating behavior emerges and spreads within these networks has significant implications for social justice, conflict resolution, and the prevention of unfair blame attribution.

Existing economic studies have incorporated political and psychological evidence, while limited attention has been paid to the opinion and group dynamics of scapegoating. This paper contributes to understanding this aspect in a network framework. We design a networked game with strategic political control of the leader and the opinion dynamics of the agents in a community. We examine three different belief distributions across network—with constant, discrete, and decay network effects—and their influences on the process of belief formation and scapegoat selection. The model suggests that the leader scapegoats more easily with a higher political trust and tends to scapegoat marginalized agent with lower local and global graph centrality.

The research significance lies in its exploration of the intersection between social blame and network economics. By incorporating network-dependent heterogeneity, we offer a more nuanced understanding of how scapegoating behaviors emerge not only from individual psychology but also from the structural properties of social networks. This approach highlights the broader social implications of scapegoating, providing insights into how such dynamics contribute to inequality and marginalization.

The structure of the paper proceeds as follows. Section 2 reviews the existing literature on scapegoating; Section 3 presents the theoretical framework for scapegoating within a networked game and conducts a preliminary analysis; Section 4 incorporates network measures for heterogeneous prior beliefs and performs numerical simulations for scapegoat selection; and Section 5 concludes with implications and future directions for research.

2 Literature Review

The metaphor of scapegoats, or "the goat that departs", originates from ancient religious and cultural practices, with its earliest documented reference in the Hebrew Bible. Specifically, it appears in the Book of Leviticus (16:21–22) as part of the Yom Kippur ritual (Day of Atonement).

> *Then Aaron shall lay both his hands on the head of the live goat, and confess over it all the iniquities of the people of Israel, and all their transgressions, all their sins, putting them on the head of the goat, and sending it away into the wilderness by means of someone designated for the task. The goat shall bear on itself all their iniquities to a barren region; and the goat shall be set free in the wilderness* (Omanson, 1991).

The concept symbolized the removal of sin and guilt from the community, which is absorbed by a sacrificial person or an animal (Crossman, 2019).

Anthropologists have explored scapegoating as a universal cultural practice in human society. Frazer (1900) examines rituals of transferring guilt or misfortune to animals or people, exploring the role of myth, magic, and religion in global human practices. Burkert (1983) discusses the functional role of sacrifice in ancient Greek religion, which is tied to myth and social cohesion. As pointed out by Turner et al. (2017), Douglas (2003), and Durkheim (1965), religious rituals transfer collective emotions and resolve communal tensions, unify the group through shared beliefs and practices, and preserve social order, which aligns with the function of scapegoating as a tool for social cohesion and control (Girard, 1977).

During the Medieval and Renaissance periods, scapegoating expanded beyond its religious and ritualistic roots and evolved into systematic social and political persecution. The idea broadened to describe a person or group that is unfairly blamed for the faults or problems of others in general (Girard, 1989). After the religious nature of scapegoating has been overlaid with political overtones and narratives, scapegoating is widely discussed in historical analyses of societal blame mechanisms to deflect responsibility or maintain power structures. History events where the society assigned unfair blames to marginalized groups for societal crises include European witch hunts (Arsal and Yavuz, 2014; Campbell, 2012; Levack, 2013) and antisemitism towards Jewish communities (Brown, 1992; Gibson and Howard, 2007).

Theories in political science are closely related to the explanation of scapegoating mechanisms in perspective of political strategies for power consolidation, distract from leadership failures, or enforcement of social hierarchies. As summarized by Cavanaugh (2011), both religious and secular ideologies have the potential for violence when used to justify power and control, and the myth of religious violence has been instrumental in legitimizing the secular state and discrediting non-secular political actors or societies, which can explain how scapegoating phenomenon (e.g., witchcraft) leads to collective and political violence. Arendt and May (1958) examine how scapegoating Jews was central to the Nazi regime's strategy to unify the masses and distract from internal failings. Herman and Chomsky (2021) develop the propaganda model, which can explain why dissenting groups or enemy states are portrayed as scapegoats by modern mass media to manufacture public support.

Except anthropology, sociology, and political science, scapegoating has also been studied as a psychological mechanism to alleviate collective anxiety. Allport (1954) explains scapegoating as a response to frustration and societal stress, where marginalized groups become "safe" targets for collective aggression. Newman and Caldwell (2005) and Glick (2005) argue that scapegoating happens when people are unable to confront or punish the actual source of their anger and instead shift their aggression towards other individuals.

Beyond focus on collective emotions, scapegoating behaviors have two notable features based on social psychologists' views. The first feature is that the formation of scapegoating is characterized by group dynamics, peer influence, and social learning. According to Garfinkel (2023), Bandura (2024), Zimbardo (2007), and Allport (1954), scapegoating is a communal process, where shared violence and blame are transferred to a victim, whose guilt is affirmed and instigated by the group's opinion and consensus. It is often the result of social contagion, where individuals adopt the group's hostility toward a perceived enemy. Moreover, it thrives in environments where aggressive behaviors are normalized and modeled by influential figures or groups. Girard (1989) proposes mimetic theory, which stated that human desire is merely an imitation of the other's desire towards objects endowed with values, resulting in scapegoating.

The second feature of scapegoating is that the group often targets weak, disadvantaged, socially vulnerable, and deprecated minority individuals or subgroups due to their lack of power, preexisting social marginalization, and the psychological or political advantages gained by blaming them (Bettelheim and Janowitz, 1950; Dollard et al., 2013; Tajfel, 1979). Concepts such as attribution theory (Heider, 2013; Kelley, 1980) and in-group/out-group dynamics (Brewer, 1979; Tajfel, 1979) also explained how blame is shifted to marginalized groups to maintain internal group cohesion.

Although scapegoating receives relatively limited attention in current economic literature, there has been a growing body of research in this area recently. Researchers start to deem scapegoating as an economic phenomenon and delve into the economic incentives behind such a behavior, encompassing both theoretical frameworks and empirical evidence.

Most studies applies an empirical or experimental approach. Miguel (2005) studies witch murder in rural Africa and India and proposes income shocks as an economic channel beyond religious and cultural mechanism. Zussman (2021) investigates how scapegoating affects evaluation decisions made by Israeli driving testers and analyzed behavioral biases in ethnicity and gender. Luca et al. (2022) extend the analysis to market dynamics, showing how discrimination intensifies during economic downturns such as the COVID-19 pandemic, as evidenced in online platforms like Airbnb. Bauer et al. (2023) design an incentivized task of punishing bystanders, showing how punishment is disproportionately directed toward innocent ethnic minorities compared to the dominant group, transforming individualized tensions into a group conflict. Relatedly, using an experiment, Bursztyn et al. (2022) demonstrate that crises exacerbate scapegoating, with leaders strategically using it to consolidate political power, often to the detriment of marginalized populations.

For theories, most studies can be classified into two categories. One category adopts an organizational perspective. Winter (2001) explores the optimal investment-inducing mechanisms with allocation of responsibility in case of failure across the levels of the hierarchy, highlighting how blame-shifting can be strategically utilized to maintain group cohesion or minimize personal costs. Dezso (2009) proposes a rational model in which firms sometimes fire senior executives as scapegoats in response to performance failures to leverage their reputation, even if failure is caused by exogenous factors rather than incompetence.

Another category lies at the intersection of economics and political science and provides game-theoretic analyses. Gent (2009) links scapegoating to the diversionary theory of war, where leaders strategically initiate conflict by pursuing aggressive foreign policies to distract from domestic issues and deflect blame. Bramoullé and Morault (2021) propose a theory of instrumental scapegoating to explain violence against wealthy ethnic minorities, illustrating how elites exploit existing social biases for political gain.

However, most economic research in scapegoating focuses on specific applications or case study, without providing a general theory. Based on the literature review from above, a general model should incorporate both political-scientific and social psychological views. It should:

1. contextualize authoritarian strategies as a tool of political control;

2. tie into discussions of groupthink, peer influence, opinion dynamics, and social learning;

3. reflect why marginalized individuals or minority groups are more prone to become the targeted innocent scapegoats.

Specifically, the second point is rarely considered in existing literature. This field, named Opinion Dynamics, which was firstly developed by DeGroot (1974), has been widely studied using graph theory and network tools in social learning research. Therefore, a network framework can be used to extend the model of scapegoating.

This study aims to fill this research gap by establishing a two-phase networked game to explain the general scapegoating phenomenon. The network nature of this model captures group interaction dynamics and opinion spread of scapegoating behaviors. It also explains why political leaders strategically persecute marginalized individuals who are essentially innocent despite being widely considered guilty, using network measures such as centrality. Section 3 presents the model in detail.

3 The Model

3.1 Model Setup

Suppose a community consists of a leader, as a representative of any political elite or ruling party, and a population of $|N| = n$ agents. All the agents are connected through a social network, represented as an undirected graph $G(V, E)$, where each node $v \in V$ corresponds to an individual and each edge $e \in E$ represents a social connection between two agents.

In this study, G can either be weighted or unweighted. In the unweighted case, edges binarily indicate the existence of a relationship between two agents, while in the weighted case, edges reflect the strength of their connection. For generality, let A denote the adjacency matrix of G which captures the connectivity of G, whether it is weighted or unweighted. The graph G may also be connected or disconnected, allowing for the presence of isolated vertices within G.

The community faces a systematic problem or failure that affects everyone, which can be caused by the leader, any agent, or other external factors. The true state of each agent is binary, indicating whether the agent is innocent or guilty in relation to the problem. Formally, the true state of agent i is denoted as $\theta_i \in \{0, 1\}$, where 0 represents innocence and 1 represents guilt.

Assumption 1. *For all $i \in V$, $\theta_i = 0$.*

By Assumption 1,we assume that all agents are innocent, meaning the problem is either exogenous or caused by the leader. The rationale behind this assumption is to focus on the scenario where an innocent individual is strategically framed as guilty by the leader, and this prosecution is collectively accepted at the social level. This setup is consistent with the definition of scapegoating.

3.1.1 Beliefs

Suppose both the leader and the agents share complete information about the network structure. For each agent k in the network, agent i holds a prior belief $\pi_i^{(k)}$ about the probability that agent k is guilty.

Assumption 2. *When $i = k$, $\pi_k^{(k)} = 0$ for all $k \in V$.*

Assumption 2 states that agents have complete information about their own states and are certain about their innocence.

At the same time, the leader has complete knowledge of the states of all agents. Every agent shares a common belief that the leader will certainly prosecute a guilty agent with probability 1 in the interest of social justice. However, the leader may also prosecute an innocent agent with probability $p \in [0, 1]$ driven by personal interest. As explained earlier, this study focuses on the second case, where the leader may unjustly prosecute an innocent agent. The parameter p quantifies the degree of political trust people have in their leader: specifically, $p = 0$ indicates full trust in the leader, while $p = 1$ suggests complete distrust.

3.1.2 Strategies

The game is divided into three phases: (i) scapegoat selection, (ii) opinion dynamics, and (iii) collective voting.

The first phase models the leader's decision-making process and draws upon frameworks such as the "cheap talk" information game (Crawford and Sobel, 1982) and the public broadcast model (Bloch et al., 2018). During this phase, the leader either selects an agent $i \in V$ to scapegoat (i.e., publicly declares agent i as guilty) or broadcasts nothing and takes the blame himself. The leader's strategy can thus be defined as:

$$s_L : G \to V \cup \{\emptyset\} \tag{1}$$

If $s_L(G) = \emptyset$, the game ends immediately. Otherwise, the game proceeds to the second phase.

The second phase models the opinion dynamics of all agents through the social network. Suppose the leader selects agent k in the first phase. In this case, each agent i updates their belief about agent k using Bayes' rule. The posterior belief is given by:

$$b_i^{(k)} = \frac{\pi_i^{(k)}}{\pi_i^{(k)} + p(1 - \pi_i^{(k)})}. \tag{2}$$

To avoid the undefined case when both $\pi_i^{(k)}$ and p is 0, we assume $p > 0$.

After updating their beliefs, each agent i forms an opinion $x_i^{(k)} \in [0, 1]$ regarding the probability that agent k is guilty. Unlike the initial belief $b_i^{(k)}$, which is formed before any social interaction, the opinion $x_i^{(k)}$ is dynamically adjusted through network-based social learning, reflecting the collective influence of the network. Thus, while $b_i^{(k)}$ represents an agent's internal opinion before social interactions, $x_i^{(k)}$ reflects the public opinion that emerges and evolves after agents exchange information and influence one another within the network.

The third phase models the verdict for the selected scapegoat. This phase follows directly from the opinion dynamics of the second phase and assumes no further strategic behavior. A scapegoat is punished if and only if the scapegoating decision is publicly acknowledged. To formalize this, we first define the social opinion. Let $x_i^{*(k)}$ represent the equilibrium opinion of agent i in the second phase.

Definition 1. *The social opinion $\bar{x}^{(k)}$ towards agent k is defined as the average of the public opinions across all agents. Formally, we have $\bar{x}^{(k)} = \frac{1}{n} \sum_{i=1}^{n} x_i^{*(k)}$.*

Next, we assume that the scapegoating decision is considered successful—and the scapegoat is punished—if the social opinion indicates that agent k is more likely guilty than not. This automatic collective voting mechanism is formalized as follows:

Definition 2. *Scapegoating towards k is successful if $\bar{x}^{(k)} \geq \frac{1}{2}$.*

3.1.3 Utilities

To better fit in a game-theoretic setting, we adopt a version of the opinion dynamics model proposed by Bindel et al. (2015) and Ghaderi and Srikant (2014). The utility function of agent i consists of two components: (i) the penalty for the discrepancy between their public opinion and internal belief, and (ii) the penalty for the disagreement between their public opinion and the opinions of others, representing peer influence. The utility function for agent i is therefore given by:

$$u_i^{(k)}(x_i^{(k)}, x_{-i}^{(k)}) = -(x_i^{(k)} - b_i^{(k)})^2 - \sum_{j=1}^{n} A_{ij}(x_i^{(k)} - x_j^{(k)})^2 \tag{3}$$

Here, $x_i^{(k)}$ denotes agent i's public opinion about agent k, $b_i^{(k)}$ denotes agent i's internal belief about agent k, and A_{ij} is the entry of the adjacency matrix representing the connection between agent i and agent j. Note that $A_{ij} \neq 0$ if and only if agents i and j are directly connected in the network. This means that an agent is influenced only by the opinions of their neighbors when averaging their own opinion.

For the leader's utility function, if he chooses to self-blame, he assumes responsibility for the entire community and bears a fixed cost $C > 0$ for the problem. For later purposes, we assume the network size is large enough such that $n \geq 2C$. On the other hand, if he chooses to scapegoat agent k, he successfully shifts the blame to agent k and avoids the constant cost. However, this comes with a trade-off: the leader suffers a reputational cost if the social opinion deviates from 1, indicating that the community does not recognize the legitimacy of the leader's decision to prosecute. We define the the leader's reputational cost as follows.

Definition 3. *The leader's reputational cost $R^{(k)}$ for scapegoating agent k is given by $R^{(k)} = n(1 - \bar{x}^{(k)})$.*

The reputational cost is defined as the discrepancy between the social opinion and 1, scaled by the network size. Equivalently, it can be written as:

$$R^{(k)} = \sum_{i=1}^{n}(1 - x_i^{*(k)}). \tag{4}$$

Based on this definition, the leader's utility function takes the following form:

$$v_L(s_L) = \begin{cases} -R^{(k)} & \text{if } s_L(G) = k \text{ for each } k \in V, \\ -C & \text{if } s_L(G) = \emptyset. \end{cases} \tag{5}$$

Here, $v_L(s_L)$ represents the leader's payoff. If the leader scapegoats agent k, his utility is negatively affected by the reputational cost $R^{(k)}$. If the leader chooses to self-blame (i.e., not scapegoat anyone), he incurs a fixed cost C.

3.2 Preliminary Analysis

3.2.1 Equilibrium Solution

In this section, we begin by assuming that the distribution of each agent i's prior belief regarding agent k is homogeneous across the network, i.e., $\pi_i^{(k)} = \frac{1}{2}$ for all i and k with constant network effects. This assumption represents a scenario where all individuals in the community are unknown strangers and hold completely skeptical beliefs about each other's guilt or innocence. Consequently, the network structure has no influence on their initial prior beliefs. This simplification provides a useful starting point for a preliminary analysis and derives benchmark results for more complex scenarios later.

Under this assumption, we solve for the Perfect Bayesian Equilibrium (PBE) of the game using backward induction. Since the collective voting phase does not involve strategic behaviors, we start from the opinion game in the second phase.

Suppose the equilibrium opinions formed by all agents in the second phase are represented by the vector $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})^{\mathrm{T}}$. Similarly, the posterior beliefs of all agents are captured by the vector $b^{(k)} = (b_1^{(k)}, b_2^{(k)}, \ldots, b_n^{(k)})^{\mathrm{T}}$.

Let L denote the Laplacian matrix of the graph G, defined as $L = D - A$, where D is the diagonal degree matrix of G and A is the weighted or unweighted adjacency matrix. We show that the stage game admits a unique Nash equilibrium.

Claim 1. *The opinion game in the second phase has a unique Nash equilibrium* $x^{*(k)} = (I + L)^{-1} b^{(k)}$.

Proof. In the Nash equilibrium, each agent maximizes their utility function, which is concave in $[0, 1]$. The optimal solution is achieved by setting the first-order condition $\frac{\partial u_i^{(k)}}{\partial x_i^{(k)}} = 0$, resulting in the following system of equations:

$$(x_i^{(k)} - b_i^{(k)}) + \sum_{j=1}^{n} A_{ij}(x_i^{(k)} - x_j^{(k)}) = 0, \forall i \in V.$$

Rearranging the terms:

$$(1 + \sum_{j=1}^{n} A_{ij})x_i^{(k)} - \sum_{j=1}^{n} A_{ij}x_j^{(k)} = b_i^{(k)}, \forall i \in V.$$

Writing this in matrix form, we get $(I+D)x^{(k)} - Ax^{(k)} = b^{(k)}$. This simplifies to $(I+L)x^{(k)} = b^{(k)}$. Since the Laplacian matrix L is positive semi-definite, we know that the matrix $I + L$ is positive definite, thus invertible. Therefore, the Nash equilibrium is uniquely represented by $x^{*(k)} = (I + L)^{-1}b^{(k)}$.

The proof is also provided in Bindel et al. (2015). □

Since L represents the algebraic connectivity of the graph G, the formation of the public opinion vector $x^{*(k)}$ can be interpreted as the outcome of social interactions, where agents average their beliefs based on their initial internal opinions $b^{(k)}$ and the network structure, which is captured by the Laplacian matrix L.

Next, we substitute this equilibrium solution into the leader's utility function from the first phase to derive the condition for scapegoating. To simplify the results and obtain a more elegant solution, we present the following lemma.

Lemma 1 (Conservation Principle)**.** $\sum_{i=1}^{n} x_i^{*(k)} = \sum_{i=1}^{n} b_i^{(k)}$ *for all* $k \in V$*. Equivalently,* $\bar{x}^{(k)} = \bar{b}^{(k)}$*, where* $\bar{b}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} b_i^{(k)}$*.*

Proof. We start with the equilibrium condition for the opinion dynamics game, which is given by $(I + L)x^{*(k)} = b^{(k)}$. Multiply both sides of this equation by the vector of ones, denoted by $\mathbf{1} = (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$, we get $\mathbf{1}^{\mathrm{T}}(I + L)x^{*(k)} = \mathbf{1}^{\mathrm{T}}b^{(k)}$. Since the Laplacian matrix L is row-stochastic, we have $\mathbf{1}^{\mathrm{T}}L = 0$. This simplifies to $\mathbf{1}^{\mathrm{T}}(I + L)x^{*(k)} = \mathbf{1}^{\mathrm{T}}x^{*(k)}$. We also have $\mathbf{1}^{\mathrm{T}}x^{*(k)} = \sum_{i=1}^{n} x_i^{*(k)}$ and $\mathbf{1}^{\mathrm{T}}b^{(k)} = \sum_{i=1}^{n} b_i^{(k)}$. Thus, the equation becomes

$\sum_{i=1}^{n} x_i^{*(k)} = \sum_{i=1}^{n} b_i^{(k)}$. This proves the first part of the lemma.

Based on the definition, $\bar{x}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} x_i^{*(k)}$ and $\bar{b}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} b_i^{(k)}$. From the previous result, we have $\bar{x}^{(k)} = \bar{b}^{(k)}$. This completes the proof. □

Lemma 1 reflects a crucial and intuitive property for opinion dynamics. It states that the total sum of opinions after social learning remains equal to the total sum of initial opinions before social learning. This preservation implies that, while social learning redistributes and adjusts opinions through the averaging process, it does not create or destroy the overall "mass" of opinions.

The conservation principle of opinions is easily understood in real-life scenarios. Consider a group of people, each holding strong and differing beliefs about an event. After communication and exchange of ideas, it is natural to expect that the group's collective opinion will remain strong on average, even though individual opinions may have shifted.

From this lemma, we see that the social opinion is essentially predetermined by the initial internal beliefs of the agents, along with the outcome of collective punishment in the third phase.

Corollary 1 (Scapegoating Success). *Scapegoating towards k is successful if and only if $\bar{b}^{(k)} \geq \frac{1}{2}$.*

Proof. The result follows directly from Lemma 1. □

The conservation principle also ensures that the information transmission process within the network and the expected reputational cost is predictable and stable, which, in turn, makes the equilibrium solution in the first phase tractable.

The equilibrium condition in the first phase is straightforward given the comparable utility structures. If $R^{(k)} > C$ for any $k \in V$, scapegoating becomes costly and the leader's best response is to take responsibility upon himself. On the other hand, if $R^{(k)} \leq C$, the leader strategically selects scapegoats within the network to minimize his utility loss. Therefore, the equilibrium for the scapegoat selection game in the first phase is:

$$s_L^*(G) = \begin{cases} \emptyset & \text{if } R^{(k)} > C \text{ for any } k \in V, \\ \arg\min_{k \in V} R^{(k)} & \text{if } R^{(k)} \leq C \text{ for some } k \in V. \end{cases} \tag{6}$$

If there exist multiple agents $k_1 \neq \ldots \neq k_l$ such that $R^{(k_1)} = \cdots = R^{(k_l)} = \arg\min_k R^{(k)}$,then in the case of a tie, we assume that the leader randomly selects one agent from the set $\{1, \ldots, l\}$ to scapegoat.

The above framework can be generalized to an l-scapegoat selection problem with $l \geq 2$, where the leader selects l agents corresponding to the l smallest values of $R^{(k)}$. This generalization extends the model to scenarios involving scapegoating of minority groups. For simplicity, we assume $l = 1$ and focus on the scapegoating of a single marginalized individual in this study.

3.2.2 Benchmark Results

We now present the following proposition, which establishes the feasibility condition for the leader's decision to scapegoat versus self-blame, based on the trust level p in the community and the properties of the social network.

Proposition 1 (Scapegoating Feasibility). *Given a social network $G(V, E)$, there always exists a threshold trust value $p_G^{*(k)}$ such that for all $k \in V$, the leader is willing to scapegoat agent k if $p \leq p_G^{*(k)}$ and self-blame if $p > p_G^{*(k)}$. Moreover, let $p_G^* = \max_{k \in V} p_G^{*(k)}$. The leader will choose to scapegoat if and only if $p \leq p_G^*$.*

Proof. By the condition in Equation 6, the leader is willing to scapegoat agent k if and only if $R^{(k)} \leq C$. By Equation 4 and Lemma 1, $R^{(k)}$ is equivalent to $\sum_{i=1}^{n}(1 - b_i^{(k)})$. Substituting from Equation 2, this simplifies further to $\sum_{i=1}^{n}(1 - \frac{\pi_i^{(k)}}{\pi_i^{(k)} + p(1-\pi_i^{(k)})})$. Thus, $R^{(k)}$ can be viewed as a function of p, denoted as $f^{(k)}(p)$. Therefore, the leader will choose to scapegoat if and only if $f^{(k)}(p) \leq C$.

Consider the following cases: (i) If $\pi_i^{(k)} = 1$ for all $i \in V$, then $f^{(k)}(p) = 0 < C$. Set $p_G^{*(k)} = 1$; (ii) If $\pi_i^{(k)} = 0$ for all $i \in V$, then $f^{(k)}(p) = n \geq 2C > C$. Set $p_G^{*(k)} = 0$; (iii) If there exists some $i \in V$ such that $\pi_i^{(k)} \in (0, 1)$, observe that $f^{(k)}(p)$ is strictly increasing with p. This follows from the derivative $\frac{\partial f^{(k)}(p)}{\partial p} = \sum_{i=1}^{n} \frac{\pi_i^{(k)}(1-\pi_i^{(k)})}{(\pi_i^{(k)}+p(1-\pi_i^{(k)}))^2} > 0$. Since $f^{(k)}(p)$ is strictly increasing in p and continuous, there exists a unique $p' = f^{(k)-1}(C)$ that satisfies $f^{(k)}(p^*) = C$. Set $p_G^{*(k)} = p' = f^{(k)-1}(C)$.

In all three cases, we are able to determine the threshold trust value $p_G^{*(k)}$ for all $k \in V$ based on the conditions. This completes the first part of the proof.

Finally, the leader will choose to scapegoat if $p \leq p_G^{*(k)}$ for at least one $k \in V$, which is equivalent to $p \leq \max_{k \in V} p_G^{*(k)} = p_G^*$. This completes the proof. $\square$

Proposition 1 captures how varying trust levels affect the leader's choice to scapegoat or not. The feasibility is derived from the condition of reputational cost and reflected by an agent-specific threshold trust value. For each agent, the specific trust threshold indicates when scapegoating that agent becomes just worthwhile compared to self-blame. If the network's trust level p is higher than this threshold, scapegoating that agent is not appealing; if it is lower, scapegoating is preferred. This is intuitive since a lower p represents a higher political trust in the leader by the community, thus the agents in the network are more likely to believe the leader's claim about the scapegoat's guilt, even when there is limited evidence.

Among all agents, the highest threshold $p_G^* = \max_{k \in V} p_G^{*(k)}$ represents the most favorable conditions for scapegoating across the network. It specifies the worst-case scenario of the leader's collective disbelief required for the leader to successfully scapegoat.

By Proposition 1, we establish $p_G^{*(k)}$ and p_G^*, which serve as the local and global boundaries for the leader's scapegoating decision.

When the feasibility condition is satisfied for multiple agents, the leader chooses the one that minimizes the reputational cost. We link this uniqueness condition to $p_G^{*(k)}$ that we have just established.

Corollary 2 (Scapegoating Uniqueness)**.** *Given a social network $G(V, E)$, if there exists $k \in V$ such that $p \leq p_G^{*(k)}$, $s_L^*(G) = \arg\max_{k \in V} p_G^{*(k)}$.*

Proof. Suppose, for the sake of contradiction, that there exists $j \in V$ such that $s_L^*(G) = j$ with $p_G^{*(k)} > p_G^{*(j)}$. By definition, $f^{(j)}(p_G^{*(k)}) > f^{(j)}(p_G^{*(j)}) = C$, which violates the feasibility condition for $p \in (p_G^{*(j)}, p_G^{*(k)})$. Thus, we must have $s_L^*(G) = \arg\max_{k \in V} p_G^{*(k)}$. $\square$

Corollary 2 proves that the final selected scapegoat has a maximal threshold trust value. In short, scapegoating an agent becomes the most favorable if and only if the agent allows the loosest limits on political trust.

Note that the above conclusions apply to any network G and prior belief distribution. Given a specific prior distribution, i.e., the prior belief with constant network effects, we can further determine the specific value of $p_G^{*(k)}$ and draw more detailed conclusions in this benchmark section.

Theorem 1. *Given a social network $G(V, E)$ and a constant prior $\pi_i^{(k)} = \frac{1}{2}$ for all $k \in V$, the following results hold:*

(a) $p_G^{*(k)} = \frac{C}{n-C}$;

(b) $P(s_L^*(G) = i) = \frac{1}{n}, \ \forall i \in V$;

(c) Given $p \leq p_G^{*(k)}$, scapegoating towards k is always successful.

Proof. For (a), since $\pi_i^{(k)} = \frac{1}{2}$ for all $k \in V$, we have $b_i^{(k)} = \frac{1}{1+p}$ for all $k \in V$. The scapegoating condition is thus $n(1 - \frac{1}{1+p}) \leq C$, which simplifies to $p_G^{*(k)} = \frac{C}{n-C}$. By the assumption that $n > 2C$, $\frac{C}{n-C} \in [0, 1]$ is a valid probability.

For (b), since $p_G^{*(k)}$ is constant for all $k \in V$, by Corollary 2, the leader will randomly select one agent from the entire vertex set V.

For (c), consider $\bar{b}^{(k)} = \frac{1}{1+p} \geq \frac{1}{2}$, which holds for all $p \leq p_G^{*(k)}$. By Corollary 1, scapegoating towards k is always successful. □

Theorem 1 states the scapegoating feasibility, uniqueness, and success conditions with a constant prior belief across the network.

With the homogeneous belief distribution, the threshold trust value is irrelevant to the network structure and depends only on two factors: the size of the community n and the fixed cost of self-blame C. When the fixed cost C increases, the threshold becomes larger, indicating that the leader has stronger incentives to avoid C and chooses to scapegoat. The threshold also decreases with n, which reflects the idea that in larger communities, the leader's reputational cost is diffused by agents, thus scaling with the number of them and making scapegoating riskier.

Since the prior belief is constant across the network, the reputational cost and threshold trust value is identical for all agents. Consequently, the leader's scapegoating decision is indifferent among agents, and each has an equal probability $\frac{1}{n}$ of being scapegoated. In this case, scapegoating is driven purely by symmetry in the prior beliefs, with no network structure influencing the leader's decision.

Moreover, the condition $p \leq p_G^{*(k)}$ ensures that the leader's decision to scapegoat is not only feasible but also successfully acknowledged by the community. As long as the scapegoating behavior is beneficial and chosen by the leader, it will always be perceived as legitimate by the community.

While the constant assumption $\pi_i^{(k)} = \frac{1}{2}$ simplifies the analysis and is mathematically robust under the given assumptions, it may not fully capture real-world networks, and its practical relevance may vary based on the specific network characteristics and prior belief distributions in these scenarios. We consider these more complex scenarios in Section 4.

4 Network Heterogeneity

In this section, we vary each agent's prior beliefs about others' types assuming network-dependent heterogeneity. Specifically, agents hold heterogeneous prior beliefs based on their positions in the network. Naturally, this prior belief should be negatively affected by the distance between two agents.

Let $N(i)$ denote the set of the neighbors of i, and let l_{ik} denote the geodesic or shortest path length between i and k. If k is an isolated point, define $l_{ik} = \infty$ for any $i \neq k$.

Assumption 3. *If $i \in N(k)$, then $\pi_i^{(k)} = 0$; If $l_{ik} \to \infty$, then $\pi_i^{(k)} \to \frac{1}{2}$.*

This assumption models the influence of network topology and local information on belief formation, incorporating two key elements: neighbor-based trust and distance-based skepticism.

We assume each agent has only local information about the network, i.e., each agent has complete information only about their neighbors. The first statement in the assumption reflects how social relationships and proximity shape initial opinions about guilt. This aligns with the idea that stronger social ties (e.g., through direct connections) foster greater trust and mutual understanding (Coleman, 1988; Granovetter, 1973).

The second statement reflects increasing unfamiliarity and skepticism as the distance between agents grows. This also aligns with the idea that distant agents have less information or interaction with each other, making them more likely to default to skeptic beliefs.

By restricting agents' knowledge to their local neighborhoods, the model mirrors real-world scenarios where individuals are most informed about their immediate connections but rely on broader social dynamics (e.g., rumors, leader broadcasts) for opinions about distant agents, making the opinion formation process more dependent on the network's structure.

For the rate of change of the prior belief towards non-neighbors, we consider two network effects: (i) discrete network effects; and (ii) decay network effects. We discuss how the two belief distributions across the network lead to different scapegoat selection conditions in the next two subsections 4.1 and 4.2.

4.1 Scapegoat Selection with Discrete Network Effects

We first consider the simpler case where the skepticism towards non-neighbors changes discretely. Formally, we define such a prior belief as follows.

Definition 4. *A prior belief $\pi_i^{(k)}$ with discrete network effects is defined as*

$$\pi_i^{(k)} = \begin{cases} 0 & \text{if } i \in N(k), \\ \frac{1}{2} & \text{if } i \notin N(k). \end{cases}$$

The discrete network effects assume step-function beliefs based on clear thresholds for distance, which simplifies the analysis. It reflects the idea that agents only have complete knowledge about and full trust in their neighbors, but view all the others in the community as strangers and hold complete skeptic beliefs.

Analogous to Theorem 1, we derive the scapegoating feasibility, uniqueness, and success conditions with a discrete prior belief distribution across the network.

Theorem 2. *Given a social network $G(V, E)$ and a prior belief $\pi_i^{(k)}$ with discrete network effects, the following results hold:*

(a) $p_G^{*(k)} = \frac{C - \deg(k)}{n - C}$;

(b) $s_L^*(G) = \arg\min_{k \in V} \deg(k)$;

(c) Given $p \leq p_G^{*(k)}$, scapegoating towards k is always successful.

Proof. For (a), by Definition 4, we have $b_i^{(k)} = \frac{1}{1+p}$ for $i \notin N(k)$ and $b_i^{(k)} = 0$ for $i \in N(k)$. The scapegoating condition is thus $\sum_{i \notin N(k)}(1 - \frac{1}{1+p}) + \sum_{i \in N(k)}(1-0) \leq C$, which simplifies to $(n - \deg(k))(1 - \frac{1}{1+p}) + \deg(k) \leq C$. Rearranging gives $p \leq \frac{C-\deg(k)}{n-C}$, which is set as p_G^*.

For (b), the proof is immediate by Corollary 2.

For (c), consider the condition $\bar{b}^{(k)} \geq \frac{1}{2}$, which implies $(1 - \frac{1}{1+p}) + \frac{\deg(k)}{n(1+p)} \leq \frac{1}{2}$. Simplifying this inequality shows that the success of scapegoating towards k is guaranteed if and only if $p \leq 1 - \frac{2\deg(k)}{n}$ by Corollary 1. To prove the statement, it suffices to show that $p_G^{*(k)} = \frac{C-\deg(k)}{n-C} \leq 1 - \frac{2\deg(k)}{n}$. Simplifying this inequality yields $\deg(k) \leq n$, which clearly holds for all $k \in V$. Therefore, if $p \leq p_G^{*(k)}$, it follows that $p \leq 1 - \frac{2\deg(k)}{n}$, ensuring that scapegoating towards k is always successful. □

We further link the scapegoat selection to the graph centrality measure.

Proposition 2. *Given a social network $G(V,E)$ and a prior belief $\pi_i^{(k)}$ with discrete network effects, the leader will either self-blame or scapegoat the agent with the lowest degree centrality.*

Proof. The result follows directly from the statement (b) in Theorem 2. □

Degree centrality reflects the number of direct connections an agent has in the network. Agents with low degree centrality have fewer neighbors and therefore weaker local influence, thus can be viewed as marginalized individuals in the community.

Unlike the feasibility condition in Theorem 1, the threshold trust value for each agent k depends on more than n and C. It now also relies on the network structure, specifically, the agent k's degree centrality. The threshold decreases as the degree of agent k increases. This is because agents with higher degree have more direct connections to other agents and broader local influence, making it harder for the leader to scapegoat them successfully without incurring significant reputational costs. The feasibility threshold in Theorem 1 represents a best-case scenario, as it assumes all agents are strangers to each other. In this case, the network can be considered a trivial graph where all nodes are isolated, resulting in $\deg(k) = 0$ for all $k \in V$. Consequently, the threshold $p_G^{*(k)}$ derived in Theorem 1 serves as a degenerate case of the more general thresholds that account for network connectivity.

The leader optimally chooses the agent with the lowest degree centrality as the scapegoat. This is because such agents are less connected and less influential in the network. Their neighbors, who hold $\pi_i^{(k)} = 0$ and do not contribute to public disbelief, have a limited number and social influence, making the average belief about their guilt higher. Therefore, the low degree centrality limits their ability to mobilize neighbors to challenge the scapegoating decision and minimizes the reputational cost of scapegoating. As an implication, agents with fewer connections are more socially vulnerable. Their limited ability to influence opinions makes them ideal targets for blame without risking significant public opposition, which aligned with the previous literature.

The relationship between the leader's trust threshold and the community's opinion dynamics ensures that once scapegoating is feasible, it will also be successful. This creates a strong strategic alignment between the leader's choices and the network's response. Overall, the leader's choice to scapegoat the least connected agent aligns with a "divide-and-conquer" strategy, leveraging the structural isolation of low-degree nodes to ensure a favorable social opinion. It may lead to localized dissent among direct neighbors but prevents widespread opposition due to the agent's limited influence. This strategy is particularly effective in sparse networks or networks with a high degree of heterogeneity, where low-degree or isolated agents are more prominent and easier to find.

Figure 1 provides four numerical simulations of randomly generated social networks with varying sizes and an assumed p of 0.3. The red nodes indicate the selected scapegoats, and the values on the nodes represent the reputational costs incurred by scapegoating each agent. The reputational cost increases as the network size n grows, reflecting the scaling effect of the network size. The leader's strategy is consistent: scapegoating low-degree agents regardless of network size. As a result, selected nodes are locally less connected nodes compared to other nodes.

However, these nodes, though with a high probability, are not necessarily the most distant node in the network, and the scapegoat candidate may not be unique, as displayed in the third figure.

Figure 2 illustrates another example where degree centrality presents a poor prediction of the scapegoat selection. In this example, all nodes except the central node are chosen as

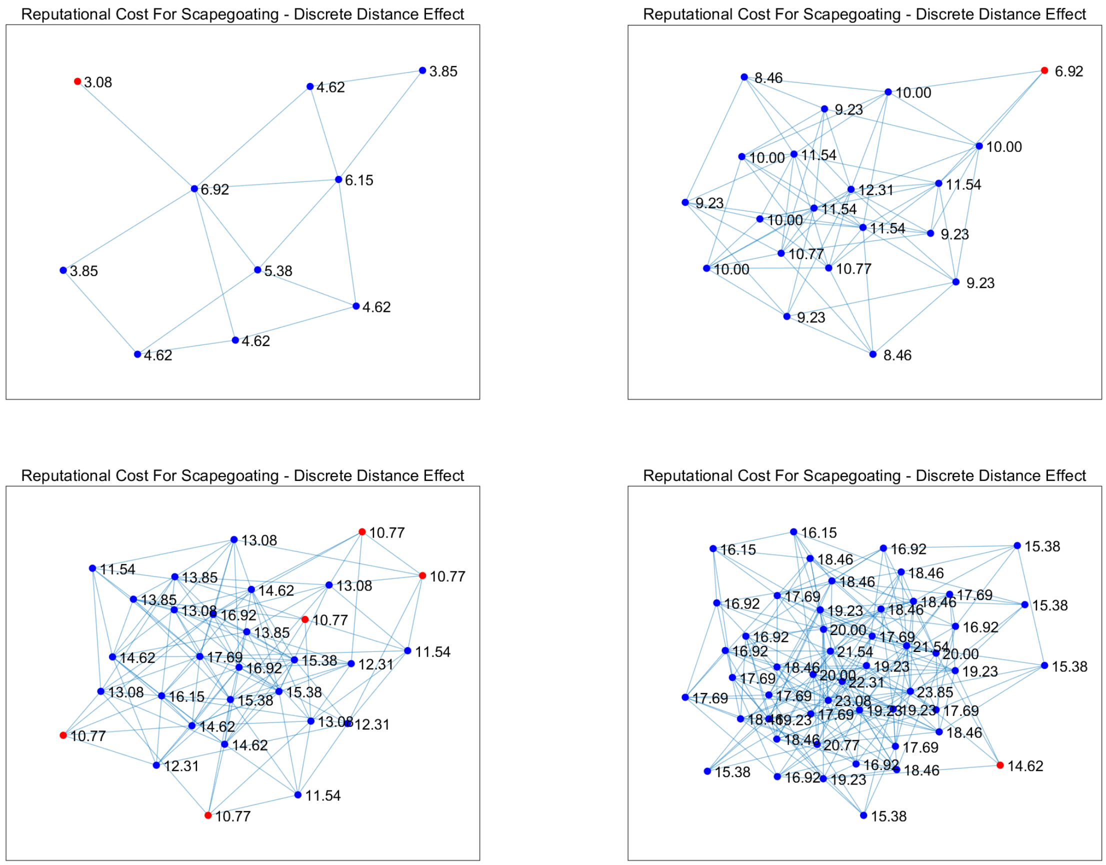

Figure 1: Social networks with $n = 10, 20, 30, 50$, $p = 0.3$, discrete network effects, and scapegoat selection (red nodes as selected scapegoats)

scapegoat candidates with degree 2. However, the far northeast node seems more distant and harder to reach the other nodes, justifying it to be a more proper scapegoat. The degree centrality fails to capture this remoteness since it focuses only on local connectivity and ignores the global network structure. Section 4.2 improves this problem.

4.2 Scapegoat Selection with Decay Network Effects

We now consider a more complex but realistic representation where the skepticism towards non-neighbors decreases exponentially. Formally, we define such a prior belief as follows.

Definition 5. *A prior belief* $\pi_i^{(k)}$ *with decay network effects is defined as* $\pi_i^{(k)} = \frac{1}{2} - (\frac{1}{2})^{l_{ik}}$.

The decay network effects serve as a refinement to the previous discrete network effects.

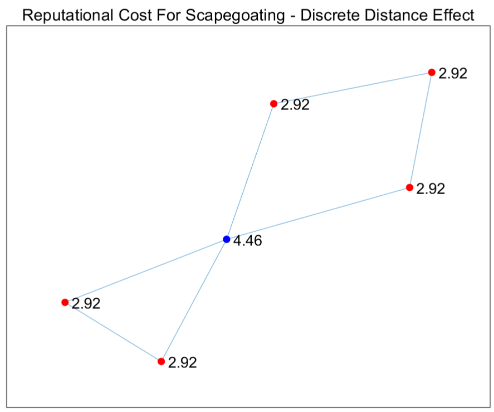

Figure 2: A poor scapegoat prediction with $n = 6$, $p = 0.3$, discrete network effects, and scapegoat selection (red nodes as selected scapegoats)

This prior distribution allows beliefs to degrade continuously with distance and provides a more gradual transition in beliefs. Specifically, the marginal decay effect is larger for closer nodes and smaller for distant nodes. Intuitively, the trust level towards a friend's friend may decrease sharply, but will be similar between a l-away person and a $(l+1)$-away person when l is sufficiently large. This aligns more closely with real-world scenarios where trust and skepticism do not change abruptly but decay smoothly and marginally with distance.

We derive the scapegoating feasibility, uniqueness, and success conditions with a decay prior belief distribution across the network.

Theorem 3. *Given a social network $G(V, E)$ and a prior belief $\pi_i^{(k)}$ with decay network effects, the following results hold:*

(a) $p_G^{*(k)} = D_k^{-1}(C)$, where D_k is a metric determined by algorithm 1;

(b) $s_L^*(G) = \arg\min_{k \in V} D_k$;

(c) Given $p \leq p_G^{*(k)}$, scapegoating towards k is always successful.

Proof. For (a), by Definition 5, $b_i^{(k)}$ is determined by the decay prior $\pi_i^{(k)} = 0.5 - 0.5^{l_{ik}}$. Using the decay prior, we calculate $b_i^{(k)}$ as $b_i^{(k)} = \frac{\pi_i^{(k)}}{\pi_i^{(k)} + p(1-\pi_i^{(k)})}$ for all $i \in V$. The scapegoating condition requires that $\sum_{i=1}^{n}(1 - b_i^{(k)}) \leq C$, which defines the metric $D_k = \sum_{i=1}^{n}(1 - b_i^{(k)})$ as an non-decreasing function of p. Solving for the threshold trust value p_G^* by the similar

Algorithm 1 Scapegoat Selection Algorithm with Decay Distance Effect

Require: Graph $G(V,E)$, parameter p

Ensure: Selected nodes with the lowest D_k

```
for each node k ∈ V do
    Initialize vector π_k ∈ ℝ^{|V|}
    for each node i ∈ V, i ≠ k do
        Calculate the shortest path length l_{ki} between k and i using Breadth-First Search (BFS) for unweighted G and Dijkstra's or Bellman-Ford Algorithm for weighted G
        π_i^{(k)} ← 0.5 − 0.5^{l_{ki}}
    end for
    π_k^{(k)} ← 0
    Generate vector b_k by b_i^{(k)} = π_i^{(k)} / (π_i^{(k)} + p·(1 − π_i^{(k)})) where ∀i ∈ V.
    Calculate D_k = Σ_{i∈V} (1 − b_i^{(k)})
end for
Select the nodes with the lowest D_k values
Output the selected nodes
```

procedure in Proposition 1, we find $p_G^{*(k)} = D_k^{-1}(C)$, where D_k is computed as described in Algorithm 1.

For (b), the proof is immediate by the construction of D_k, which equals $R^{(k)}$.

For (c), consider the condition $\bar{b}^{(k)} \geq \frac{1}{2}$, which implies $\frac{D_k}{n} \leq \frac{1}{2}$. Thus, by Corollary 1, the success of scapegoating towards k is guaranteed if and only if $p \leq D_k^{-1}(\frac{n}{2})$. To prove the statement, it suffices to show that $p_G^{*(k)} = D_k^{-1}(C) \leq D_k^{-1}(\frac{n}{2})$. Since D_k is a non-decreasing and bijective function of p, its inverse D_k^{-1} is also non-decreasing. Simplifying this inequality yields $C \leq \frac{n}{2}$, which always holds by the assumption. Therefore, if $p \leq p_G^{*(k)}$, it follows that $p \leq D_k^{-1}(\frac{n}{2})$, ensuring that scapegoating towards k is always successful. □

Similarly, we again link the scapegoat selection into another graph centrality measure—decay centrality.

Proposition 3. *The metric D_k is a component-wise convex transformation of decay centrality $C_k = \sum_{i=1}^{n} \alpha^{l_{ik}}$ with $\alpha = \frac{1}{2}$. Specifically, when $p = 1$, D_k is an equivalent centrality measure to C_k.*

Proof. By definition, $D_k = \sum_{i\in V}(1 - b_i^{(k)})$, where $b_i^{(k)}$ is derived from the decay prior $\pi_i^{(k)} = \frac{1}{2} - (\frac{1}{2})^{l_{ik}}$. Thus, D_k can be written as $D_k = \sum_{i\in V} f((\frac{1}{2})^{l_{ik}})$, where $f(x) = 1 - \frac{\frac{1}{2}-x}{\frac{1}{2}-x+p(\frac{1}{2}+x)}$.

Observe that f is strictly convex and increasing if $p \in (0, 1]$ and linear if $p = 1$. Therefore, D_k is a component-wise convex transformation of C_k^d, and the two metrics are equivalent when $p = 1$ with $\alpha = \frac{1}{2}$. $\square$

Algorithm 1 provides a computational method for determining the optimal scapegoat based on the decay prior belief distribution, which integrates the influence of both direct and indirect connections. By incorporating the geodesic measure, the decay model makes scapegoating more dependent on the global centrality in the network rather than solely on local connectivity. Closer nodes with smaller l_{ki} are weighted more heavily in the prior belief and affect further belief updates, suggesting the belief about agent k is influenced more strongly by agents close to k in the network. The metric D_k then aggregates the total reputational cost of scapegoating agent k. Based on this metric, peripheral or isolated nodes with weaker global connectivity are more likely to have lower D_k, making them prime candidates for scapegoating.

By Proposition 3, the disbelief aggregation D_k serves as a stricter measure than decay centrality, since it penalizes the longer l_{ik} and extremes their values by its convexity, making scapegoat selection even more sensitive to the agent's position in the network.

Similarly, this weakly stricter decay centrality D_k determines the specific trust threshold for each agent k and the optimal scapegoat solution. The leader selects the agent with the smallest D_k who incurs the lowest reputational cost. Since D_k not only considers local connectivity but also how well-connected agent k is to distant nodes, its value depends on more nuanced network distances and connectivity patterns. Therefore, the added dimension for comparison provides finer distinctions between nodes and makes its minimum for scapegoat selection more likely to be unique. The guaranteed success of scapegoating is also preserved in this case.

Figure 3 demonstrates the improved version of Figure 2 by using decay network effects. This modification ensures that the selected scapegoat is unique, highlighting the agent who is the most disconnected both in terms of local and global network influence.

Figure 4 provides the same network simulations as Figure 1 using decay network effects. The first two simulations generate the same results as the previous graph. The third simulation now selects a unique scapegoat compared to the five before. The fourth simulation

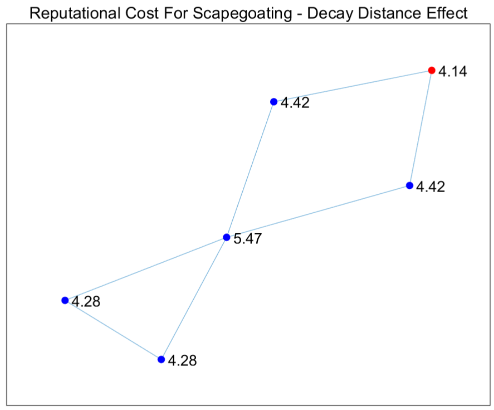

Figure 3: An improved scapegoat prediction with $n = 6$, $p = 0.3$, decay network effects, and scapegoat selection (red nodes as selected scapegoats)

selects a unique scapegoat which is different from the previous selection. This inconsistency reflects the difference that incorporating global connectivity can make. Even though the new scapegoat has a higher degree and denser local connections, it is overall more distant from all other nodes and thus is more unfamiliar to more agents, making its reputational cost higher when decay effects are added to this model.

In sum, peripheral nodes with lower global influence and centrality, farther from the network's core, less easier to reach other nodes are more likely to be selected. As a result, the leader's strategy is still consistent with previous literature: marginalized individuals are more likely to become scapegoats.

However, a disadvantage of the decay measure is that the algorithm requires calculating shortest paths for every pair of nodes, which is computationally expensive, especially for large-scale networks. The time complexity of Algorithm 1 for unweighted graph is $O(nm + n^2)$ and for weighted graph is $O(nm + n^2 \log(n))$. In large-scale networks, the number of iterations explodes, making the computational step inefficient and unrealistic. Simplified approximations or heuristic methods might be necessary for real-world implementations.

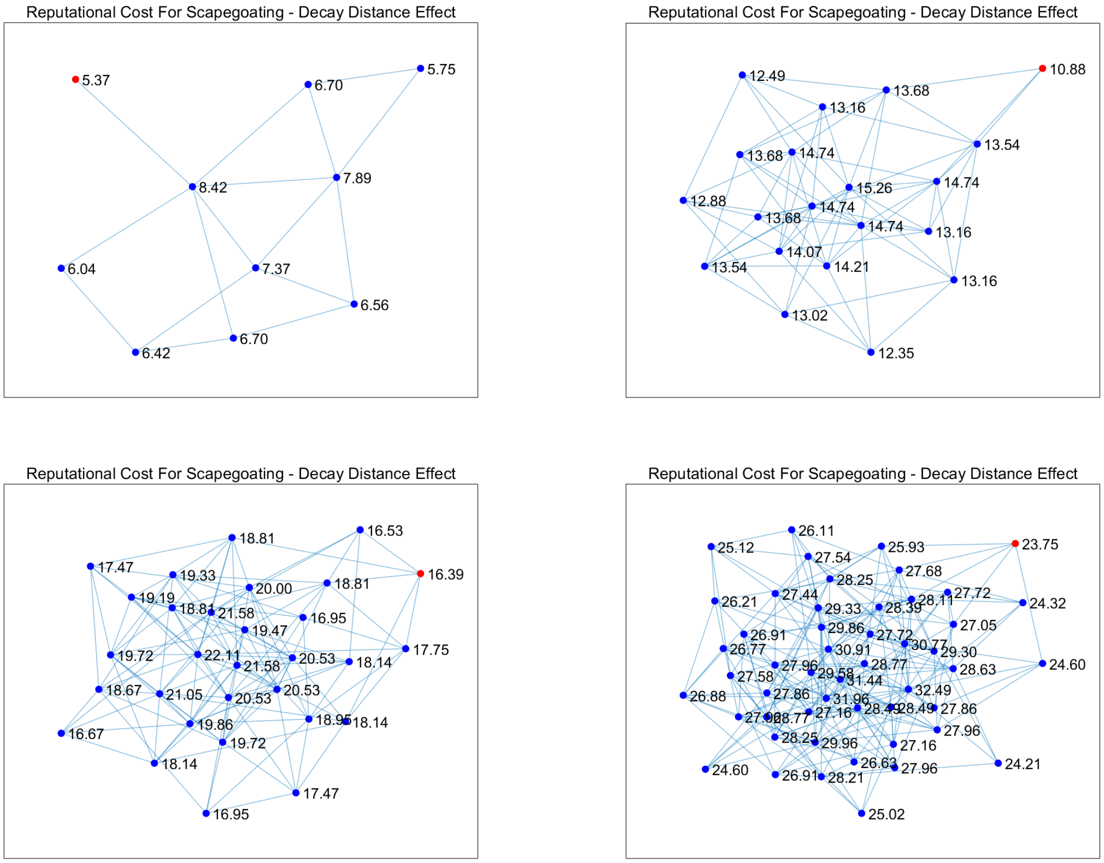

Figure 4: Social networks with $n = 10, 20, 30, 50$, $p = 0.3$, decay network effects, and scapegoat selection (red nodes as selected scapegoats)

5 Discussion and Conclusion

This paper provides a novel analysis of scapegoating behavior within a networked game, emphasizing how the network structure influences this social dynamics. By comparing the benchmark with the two heterogeneous belief models with discrete and decay network effects, we show how proximity and network position shape scapegoat selection, with agents who are more isolated or peripheral more likely to be blamed. Through this, we highlight the role of network topology in perpetuating social inequalities and marginalization.

Our findings are particularly relevant in the context of modern social networks, where the dynamics of blame are often hidden beneath layers of anonymity and distance. The study suggests that enhancing social trust, ensuring information transparency, and reinforcing community bonding can help alleviate social injustices and inequalities, as well as promote anti-collective violence and humanitarianism in interconnected environments, since it underscores the importance of fairer and more inclusive decision-making processes that minimize harm to marginalized individuals. With the above insights, we can foster more just and equitable social systems that resist the tendency to blame the vulnerable.

In future work, it would be valuable to extend the model to include incomplete information, uncertainty, or additional social and psychological factors such as norms and moral concerns. Other graph centrality can also be incorporated to study scapegoating, such as betweenness centrality focusing on importance as an intermediary connector, and eigenvector or Bonacich centrality focusing on influence, prestige, power in social networks.

Acknowledgement

I offer my sincerest gratitude to Professor Charles Becker and all the TAs in ECON 605. The course has greatly broadened my perspective and ignited my interest in network economics.

I also wish to thank my dear friends Chenzhe Zhang, Ling Lei, Runxin Zhi for their valuable inspirations, suggestions, and support that contributed to this term paper.

References

Allport, G. W. (1954). The nature of prejudice. *Reading/Addison-Wesley*.

Arendt, H. and May, N. (1958). The origins of totalitarianism.

Arsal, S. Y. and Yavuz, S. (2014). The portrait of the'scapegoat woman'as witch. *Synergies Turquie*, (7).

Bandura, A. (2024). Social learning analysis of aggression. In *Analysis of delinquency and aggression*, pages 203–232. Routledge.

Bauer, M., Cahlíková, J., Chytilová, J., Roland, G., and Želinskỳ, T. (2023). Shifting punishment onto minorities: Experimental evidence of scapegoating. *The Economic Journal*, 133(652):1626–1640.

Bettelheim, B. and Janowitz, M. (1950). Dynamics of prejudice.

Bindel, D., Kleinberg, J., and Oren, S. (2015). How bad is forming your own opinion? *Games and Economic Behavior*, 92:248–265.

Bloch, F., Demange, G., and Kranton, R. (2018). Rumors and social networks. *International Economic Review*, 59(2):421–448.

Bramoullé, Y. and Morault, P. (2021). Violence against rich ethnic minorities: A theory of instrumental scapegoating. *Economica*, 88(351):724–754.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological bulletin*, 86(2):307.

Brown, M. (1992). From stereotype to scapegoat: Anti-jewish sentiment in french canada from confederation to world war i. *Antisemitism in Canada: History and Interpretation*, pages 39–66.

Burkert, W. (1983). *Homo necans: The anthropology of ancient Greek sacrificial ritual and myth*. Univ of California Press.

Bursztyn, L., Egorov, G., Haaland, I., Rao, A., and Roth, C. (2022). Scapegoating during crises. In *AEA Papers and Proceedings*, volume 112, pages 151–155. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Campbell, C. (2012). *Scapegoat: A history of blaming other people*. Abrams.

Cavanaugh, W. T. (2011). The myth of religious violence. *The Blackwell companion to religion and violence*, pages 23–33.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, 94:S95–S120.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Crossman, A. (2019). Definition of scapegoat, scapegoating, and scapegoat theory. ThoughtCo.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical association*, 69(345):118–121.

Dezso, C. L. (2009). Scapegoating and firm reputation. *Robert H. Smith School Research Paper No. RHS*, pages 06–105.

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., Sears, R. R., Ford, C. S., Hovland, C. I., and Sollenberger, R. T. (2013). *Frustration and aggression*. Routledge.

Douglas, M. (2003). *Purity and danger: An analysis of concepts of pollution and taboo*. Routledge.

Durkheim, E. (1965). *The elementary forms of the religious life [1912]*. na.

Frazer, J. G. (1900). The golden bough: A study in magic and religion (vol. 10 of 12). *Revue des Traditions Populaires*, 15:471.

Garfinkel, H. (2023). Studies in ethnomethodology. In *Social Theory Re-Wired*, pages 58–66. Routledge.

Gent, S. E. (2009). Scapegoating strategically: Reselection, strategic interaction, and the diversionary theory of war. *International Interactions*, 35(1):1–29.

Ghaderi, J. and Srikant, R. (2014). Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50(12):3209–3215.

Gibson, J. L. and Howard, M. M. (2007). Russian anti-semitism and the scapegoating of jews. *British Journal of Political Science*, 37(2):193–223.

Girard, R. (1977). *Violence and the Sacred*. Johns Hopkins University Press, Baltimore.

Girard, R. (1989). *The scapegoat*. JHU Press.

Glick, P. (2005). Choice of scapegoats. *On the nature of prejudice: Fifty years after Allport*, pages 244–261.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1360–1380.

Heider, F. (2013). *The psychology of interpersonal relations*. Psychology Press.

Herman, E. S. and Chomsky, N. (2021). Manufacturing consent. In *Power and Inequality*, pages 198–206. Routledge.

Kelley, H. (1980). Attribution theory and research. *Annual Review of Psychology/Annual Reviews*.

Levack, B. P. (2013). *The witch-hunt in early modern Europe*. Routledge.

Luca, M., Pronkina, E., and Rossi, M. (2022). Scapegoating and discrimination in times of crisis: Evidence from airbnb. Technical report, National Bureau of Economic Research.

Miguel, E. (2005). Poverty and witch killing. *The Review of Economic Studies*, 72(4):1153–1172.

Newman, L. S. and Caldwell, T. L. (2005). Allport's "living inkblots": The role of defensive projection in stereotyping and prejudice. *On the nature of prejudice: Fifty years after Allport*, pages 377–392.

Omanson, R. L. (1991). The new revised standard version with apocrypha.

Tajfel, H. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations/Brooks/Cole.*

Turner, V., Abrahams, R., and Harris, A. (2017). *The ritual process: Structure and anti-structure*. Routledge.

Winter, E. (2001). Scapegoats and optimal allocation of responsibility. *American Economic Review, forthcoming.*

Zimbardo, P. G. (2007). *The Lucifer Effect: Understanding How Good People Turn Evil.* Random House, New York.

Zussman, A. (2021). Scapegoating in evaluation decisions. *Journal of Economic Behavior & Organization*, 186:152–163.