Network Science (2026), **14**, e11, 1–15
doi:10.1017/nws.2026.10030

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Perceived fairness in networks

Arthur Charpentier[1,2]

[1]Département de Mathématiques, Université du Québec à Montréal (UQAM), Canada and [2]Departement of Mathematics, Kyoto University, Japan
Email: charpentier.arthur@uqam.ca

Abstract

The usual definitions of algorithmic fairness focus on population-level statistics, such as demographic parity or equal opportunity. However, in many social or economic contexts, fairness is not perceived globally, but locally, through an individual's peer network and comparisons. We propose a theoretical model of perceived fairness networks, in which each individual's sense of discrimination depends on the local topology of interactions. We show that even if a decision rule satisfies standard criteria of fairness, perceived discrimination can persist or even increase in the presence of homophily or assortative mixing. We propose a formalism for the concept of fairness perception, linking network structure, local observation, and social perception. Analytical and simulation results highlight how network topology affects the divergence between objective fairness and perceived fairness, with implications for algorithmic governance and applications in finance and collaborative insurance.

Keywords: fairness; networks; homophily; stochastic block model; perceived discrimination

1. Introduction

Fairness in networked systems is often defined using population-level statistics, while individuals experience and evaluate fairness through local social comparisons. In many social, economic, and algorithmic contexts, agents observe only the outcomes of their neighbors and infer fairness from this limited information. As a result, *perceived fairness* may differ substantially from *objective fairness* defined at the population level.

Network structure plays a central role in this discrepancy. Homophily, degree heterogeneity, and clustering shape who observes whom and therefore how outcomes are locally compared. Even when a decision rule satisfies standard fairness criteria globally, network topology may distort local exposure and generate systematic differences in perceived fairness across groups.

This paper develops a concise analytical framework to study perceived fairness on networks. We model perceived fairness as a local comparison between an individual's outcome (or acceptance probability) and the average outcome observed in their network neighborhood. We then analyze how topological features affect the gap between objective and perceived fairness.

Our results show that perceived fairness converges to objective fairness only when individuals observe the entire population. At finite visibility, homophily and degree heterogeneity amplify perceived discrimination, while clustering mitigates dispersion by stabilizing local averages. These effects are structural and arise independently of any intentional bias in the decision rule.

1.1 Contributions

The paper makes four main contributions: (i) it formalizes perceived fairness as a network-dependent operator; (ii) it establishes convergence to population-level fairness as visibility grows;

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

(iii) it characterizes how homophily, degree bias, and clustering shape perceived discrimination; (iv) it provides numerical illustrations highlighting these mechanisms.

1.2 Organization of the paper

The remainder of the paper is organized as follows. Section 2 reviews the related literature and positions our contribution within the algorithmic fairness and network science strands. Sections 3 and 4 introduce the model, define our perception-based fairness operators, and establish the main convergence results as visibility increases. Section 5 analyzes how network topology (e.g., homophily, degree heterogeneity, and clustering) shapes perceived discrimination at finite visibility. Section 6 presents numerical illustrations, discusses the mechanisms highlighted by the theory, and draws policy implications for algorithmic governance in networked settings. Finally, Section 7 concludes and outlines directions for future work.

2. Related literature

2.1 Group and individual fairness

The algorithmic fairness literature traditionally distinguishes between *group fairness* and *individual fairness*. Group fairness criteria, such as demographic parity (DP) or equalized odds, impose constraints on outcome distributions across protected groups (Hardt *et al.*, 2016; Barocas *et al.*, 2023). Individual fairness, initially proposed by Dwork *et al.* (2012), requires that similar individuals be treated similarly. Recent work has extended individual fairness using counterfactual and causal frameworks, focusing on invariance to hypothetical changes in sensitive attributes (Kusner *et al.*, 2017; De Lara *et al.*, 2024; Zhou *et al.*, 2024; Fernandes Machado *et al.*, 2025). While these approaches address important normative concerns, they do not model how fairness is *perceived* in social contexts. Perceived fairness is inherently relational and group-based: individuals assess their outcomes by comparing them to those of peers, often within socially defined categories. Our framework complements individual fairness by focusing on exposure and local comparison rather than counterfactual perturbations.

2.2 Networks, exposure bias, and topology

Network structure is known to bias local observations. A classical example is the friendship paradox, whereby individuals tend to observe neighbors with higher degree or attribute values than themselves. Generalizations of this phenomenon show that any positively correlated attribute is overrepresented in local neighborhoods (Wu *et al.*, 2017; Cantwell *et al.*, 2021). These exposure biases distort inference and perception in networked systems. Recent work shows that network topology can also distort standard operations performed on networks. Charpentier and Ratz (2025) demonstrate that decentralized operations may lead to systematic biases driven purely by structure. Our results extend this insight to fairness metrics, showing how homophily, degree heterogeneity, and clustering shape perceived fairness independently of population-level parity.

2.3 Perceived discrimination and social comparison

In sociology and social psychology, perceived discrimination has long been recognized as a key determinant of behavior and well-being. Empirical studies show that perceived unfairness affects trust, motivation, and social cohesion (Pascoe & Richman, 2009; Schmitt *et al.*, 2014; Brown *et al.*, 2006; Gonzalez *et al.*, 2021). These perceptions arise through social comparison, typically based on local observations. Homophily and assortative mixing (McPherson *et al.*, 2001; Newman, 2003) create segregated local environments in which perceived fairness may diverge

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

from population-level parity. Recent work studies alternative notions of homophily based on similarity in continuous attributes, such as income, rather than group membership (Mayerhoffer & Schulz, 2022). We focus on categorical homophily relevant to discrimination across protected groups, while viewing attribute-based homophily as an important extension. Our contribution is to formalize perceived discrimination within a graph-theoretic framework, linking social comparison, exposure bias, and network topology.

3. Model of perceived fairness

3.1 Setup and notation

Let $G = (V, E, S)$ be a finite, simple, undirected graph with $|V| = n$, adjacency matrix $\boldsymbol{A} \in \{0,1\}^{n \times n}$, and degrees $d_i = \sum_j A_{ij}$. The sensitive attribute $S_i \in \{A, B\}$ induces a partition $V = V_A \cup V_B$.

A (possibly randomized) decision rule is represented by a map $h : V \to [0,1]$, where $h(i)$ denotes the *acceptance probability* of node i.

Given h, realized decisions are modeled by independent draws

$$H(i) \mid h \sim \text{Bernoulli}(h(i)), \qquad i \in V,$$

and we write $H \in \{0,1\}^V$ for the realized decision vector. This allows us to distinguish (i) fairness notions defined at the level of probabilities (h) and (ii) fairness notions defined at the level of realized outcomes (H).

For $i \in V$, denote the 1-neighborhood $N(i) = \{ j : A_{ij} = 1 \}$ and its r-hop expansion

$$N^{(r)}(i) := \left\{ j \in V : \exists k \le r \text{ with } (\boldsymbol{A}^k)_{ij} > 0 \right\}, \qquad r \in \mathbb{N}.$$

3.1.1 Running example (job-market screening)

Think of A/B as two demographic groups applying to a program (or job), and let $h(i) \in [0,1]$ denote the (possibly biased) probability that applicant i is accepted. The realized outcome is $H(i) \in \{0,1\}$ with $H(i) \mid h \sim \text{Bernoulli}(h(i))$. Individuals compare their own outcome (or acceptance probability) to what they observe in their r-hop neighborhood, so network structure shapes *perceived* fairness even when *global* acceptance rates are unchanged.

3.1.2 Objective (global) fairness

DP can be imposed either *ex ante* (on acceptance probabilities) or *ex post* (on realized outcomes):

$$\text{DP}_{\text{prob}} : \quad \mathbb{E}[h(i) \mid S_i = A] = \mathbb{E}[h(i) \mid S_i = B], \tag{1}$$

$$\text{DP}_{\text{real}} : \quad \mathbb{P}[H(i) = 1 \mid S_i = A] = \mathbb{P}[H(i) = 1 \mid S_i = B]. \tag{2}$$

When $H(i) \mid h \sim \text{Bernoulli}(h(i))$ and h is deterministic, (1) implies (2). We keep both notations because the paper studies perceived fairness both at the probability level (h) and at the realized level (H).

3.2 Local observation and perceived fairness

Individuals do not observe population averages; they observe outcomes in their r-neighborhood. Define the r-neighborhood averaging operator for any vector $x \in \mathbb{R}^V$ by

$$\mathcal{E}_i^{(r)}[x] := \frac{1}{|N^{(r)}(i)|} \sum_{j \in N^{(r)}(i)} x(j), \qquad r \in \mathbb{N}. \tag{3}$$

(For $r = 1$ we write $\mathcal{E}_i[x]$.)

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

3.2.1 Two notions of perceived fairness

We consider two closely related perception indicators.

(i) Perception based on acceptance probabilities.

$$F^{(r)}_{\text{prob}}(i;h) := \mathbf{1}\big\{\, h(i) \;\geq\; \mathcal{E}^{(r)}_i[h] \,\big\}. \tag{4}$$

This captures the idea that an individual evaluates whether their *chance* of being accepted is at least as high as the average chance observed among peers. In the running example, $F^{(r)}_{\text{prob}}(i;h)$ indicates whether individual i perceives their acceptance probability as at least the average acceptance probability in their r-hop neighborhood.

(ii) Perception based on realized outcomes.

$$F^{(r)}_{\text{real}}(i;H) := \mathbf{1}\big\{\, H(i) \;\geq\; \mathcal{E}^{(r)}_i[H] \,\big\}. \tag{5}$$

This is the literal "I was accepted vs my neighbors" comparison, and it is random even when h is fixed. In the running example, $F^{(r)}_{\text{real}}(i;H)$ captures whether the realized outcome of i (accepted or not) compares favorably to the realized outcomes observed in their r-hop neighborhood.

Equations (4)–(5) clarify which object is being compared (probabilities vs realizations); this distinction matters for convergence statements when h is randomized.

3.2.2 Fairness visibility (group-level perceived fairness)

For $s \in \{A, B\}$ define

$$\text{Vis}^{\text{prob}}_r(s;h) := \frac{1}{|V_s|}\sum_{i\in V_s} F^{(r)}_{\text{prob}}(i;h), \text{ and } \text{Vis}^{\text{real}}_r(s;H) := \frac{1}{|V_s|}\sum_{i\in V_s} F^{(r)}_{\text{real}}(i;H),$$

$$\begin{cases} \Delta^{\text{prob}}_r(h) := \text{Vis}^{\text{prob}}_r(A;h) - \text{Vis}^{\text{prob}}_r(B;h), \\ \Delta^{\text{real}}_r(H) := \text{Vis}^{\text{real}}_r(A;H) - \text{Vis}^{\text{real}}_r(B;H). \end{cases} \tag{6}$$

We say *visibility parity* holds at depth r if $\Delta_r = 0$.

3.3 Exposure operators and degree-weighting

Besides the node-average (3), it is useful to recall the edge-weighted mean

$$\bar{h}_{\text{edge}} := \frac{1}{2m}\sum_{i\in V} d_i h(i), \text{ where } m := |E|. \tag{7}$$

This identity will be the source of classical exposure biases (friendship-paradox-type effects) in our perceived fairness gap: neighborhoods may over-sample high-degree nodes, hence over-sample high-h nodes when h correlates with degree (see Proposition 5.2).

3.4 Asymptotics in neighborhood radius

We now formalize the idea that, on a connected graph, sufficiently large neighborhoods eventually recover population-level information.

Proposition 3.1 (Visibility convergence). *Assume G is connected and $|V_s| \geq 1$ for $s \in \{A, B\}$. Fix a map $h : V \to [0, 1]$.*

(a) Probability-based perception. *For each $s \in \{A, B\}$,*

$$\text{Vis}^{\text{prob}}_r(s;h) \;\longrightarrow\; \mathbb{P}\big(h(i) \geq \bar{h} \;\big|\; S_i = s\big) \qquad \text{as } r \to \infty,$$

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

where $\bar{h} := \frac{1}{|V|}\sum_{j\in V} h(j)$. *In particular, if* h *depends only on the group (i.e.,* $h(i) = h_A$ *on* V_A *and* $h(i) = h_B$ *on* V_B*), then*

$$\mathrm{Vis}_r^{\mathrm{prob}}(s; h) \longrightarrow \mathbf{1}\{h_s \geq \bar{h}\}, \qquad \text{with } \bar{h} = \frac{|V_A|h_A + |V_B|h_B}{|V|}.$$

(b) Realized-outcome perception. *Let* $H(i) \mid h \sim \mathrm{Bernoulli}(h(i))$ *independently over* $i \in V$. *Then, conditionally on* h*, for each* $s \in \{A, B\}$,

$$\mathbb{E}\big[\mathrm{Vis}_r^{\mathrm{real}}(s; H) \mid h\big] \longrightarrow \mu_s(h) + p_0(h) \qquad \text{as } r \to \infty,$$

where $\mu_s(h) := \frac{1}{|V_s|}\sum_{i\in V_s} h(i)$ *and* $p_0(h) := \mathbb{P}(\bar{H} = 0 \mid h) = \prod_{j\in V}\big(1 - h(j)\big)$. *In particular, if* $p_0(h) = 0$ *(e.g., if* $\max_{j\in V} h(j) = 1$*), then* $\mathbb{E}[\mathrm{Vis}_r^{\mathrm{real}}(s; H) \mid h] \to \mu_s(h)$.

Remark 3.1. *The limit in (a) compares* decision probabilities $h(i)$ *to the corresponding local average and is therefore deterministic once* h *is fixed. By contrast, (b) compares* realized outcomes $H(i)$ *to a realized local average, hence remains random even when* h *is fixed; it is thus most naturally stated conditionally on* h *(and, in particular, in conditional expectation). The two notions coincide for deterministic rules (when* $H = h$*), while for randomized rules they may differ due to global realization effects such as the event* $\{\bar{H} = 0\}$.

The proof is deferred to Section 4. On a connected finite graph, neighborhoods saturate: $N^{(r)}(i) = V$ for all $r \geq \mathrm{diam}(G)$, so both visibility scores stabilize beyond the diameter. In the realized case, the conditional expectation is obtained by a direct decomposition that isolates the edge event $\{\bar{H} = 0\}$ (rather than an asymptotic law-of-large-numbers argument).

4. Analytical results and proofs

This section gathers the analytical arguments underlying the results stated in Section 3. In particular, it provides the proof of the convergence of perceived fairness as the visibility radius grows.

4.1 Convergence of perceived fairness at large visibility radius

We start by recalling a simple but fundamental topological observation.

Lemma 4.1 (Neighborhood saturation). *If* G *is a finite connected graph with diameter* $\mathrm{diam}(G)$*, then for every node* $i \in V$ *and every* $r \geq \mathrm{diam}(G)$,

$$N^{(r)}(i) = V.$$

Proof. Since G is connected, any node $j \in V$ can be reached from i by a path of length at most $\mathrm{diam}(G)$. Hence $j \in N^{(r)}(i)$ for all $r \geq \mathrm{diam}(G)$. □

Lemma 4.1 implies that, beyond a finite radius, all agents observe the same population-level information. This observation underlies the proof of Proposition 3.1.

4.2 Proof of Proposition 3.1

4.2.1 Probability-based perception

Recall that

$$F_{\mathrm{prob}}^{(r)}(i; h) = \mathbf{1}\big\{h(i) \geq \mathcal{E}_i^{(r)}[h]\big\}.$$

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

By Lemma 4.1, for all $r \geq \text{diam}(G)$,

$$\mathcal{E}_i^{(r)}[h] = \frac{1}{|V|}\sum_{j\in V} h(j) =: \bar{h},\ \forall i \in V.$$

Therefore, for all sufficiently large r,

$$F_{\text{prob}}^{(r)}(i;h) = \mathbf{1}\{\bar{h} \leq h(i)\}.$$

Averaging over nodes with $S_i = s$ yields

$$\text{Vis}_r^{\text{prob}}(s) = \frac{1}{|V_s|}\sum_{i\in V_s} \mathbf{1}\{\bar{h} \leq h(i)\},$$

which coincides with $\mathbb{P}\big(h(i) \geq \bar{h} \mid S_i = s\big)$. This proves the convergence claim for probability-based perception. □

4.2.2 Realized-outcome perception

We now consider the case where realized decisions satisfy

$$H(i) \mid h \sim \text{Bernoulli}(h(i)), \qquad i \in V,$$

conditionally independent given h. By Lemma 4.1, for all $r \geq \text{diam}(G)$,

$$F_{\text{real}}^{(r)}(i;H) = \mathbf{1}\big\{H(i) \geq \bar{H}\big\}, \qquad \bar{H} := \frac{1}{|V|}\sum_{j\in V} H(j).$$

Since $H(i) \in \{0,1\}$ and $\bar{H} \in [0,1]$, we have the identity

$$\mathbf{1}\{H(i) \geq \bar{H}\} = H(i) + \mathbf{1}\{\bar{H} = 0\}.$$

Indeed, if $\bar{H} > 0$ then $\{H(i) \geq \bar{H}\} = \{H(i) = 1\}$, whereas if $\bar{H} = 0$ then $\{H(i) \geq \bar{H}\}$ holds for all i. Therefore, for all $r \geq \text{diam}(G)$,

$$F_{\text{real}}^{(r)}(i;H) = H(i) + \mathbf{1}\{\bar{H} = 0\}.$$

Taking conditional expectation given h yields

$$\mathbb{E}\big[F_{\text{real}}^{(r)}(i;H) \mid h\big] = \mathbb{E}[H(i) \mid h] + \mathbb{P}(\bar{H} = 0 \mid h) = h(i) + p_0(h),$$

where $\{\bar{H} = 0\} = \{H(j) = 0\ \forall j \in V\}$ and hence

$$p_0(h) = \mathbb{P}(H(j) = 0\ \forall j \in V \mid h) = \prod_{j\in V}(1 - h(j)).$$

Averaging over $i \in V_s$ gives, for all $r \geq \text{diam}(G)$,

$$\mathbb{E}\big[\text{Vis}_r^{\text{real}}(s;H) \mid h\big] = \frac{1}{|V_s|}\sum_{i\in V_s}\big(h(i) + p_0(h)\big) = \mu_s(h) + p_0(h),$$

which proves the claimed convergence as $r \to \infty$ (the left-hand side is in fact constant for all $r \geq \text{diam}(G)$). □

4.3 Extension and discussion

4.3.1 Asymptotic regimes

Proposition 3.1(b) makes explicit a subtle but important "edge case" of realized-outcome perception: when $\bar{H} = 0$ (i.e., when all realized outcomes are 0), every node satisfies $H(i) \geq \bar{H}$ and

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

hence $F^{(r)}_{\text{real}}(i;H)=1$ for all i. This event contributes an additive term $p_0(h)=\mathbb{P}(\bar{H}=0\mid h)=\prod_{j\in V}(1-h(j))$ to the conditional expectation of the visibility score. On a fixed finite graph, this contribution need not be negligible in general, and it is therefore natural to state the limit in terms of $\mu_s(h)+p_0(h)$.

In many network settings, however, the graph size is large and the mean approval rate is bounded away from 0 in the sense that $\frac{1}{|V|}\sum_{j\in V}h(j)\geq\varepsilon$ for some $\varepsilon>0$. In that regime, the event $\{\bar{H}=0\}$ becomes exponentially unlikely since $p_0(h)=\prod_{j\in V}(1-h(j))\leq\exp\bigl(-\sum_{j\in V}h(j)\bigr)\leq e^{-|V|\varepsilon}$. Consequently, the conditional expectation of realized visibility is well approximated by $\mu_s(h)$, which is the group-average decision probability.

The following corollary formalizes this approximation in a large-network regime.

Corollary 4.2 (Vanishing edge case). *Assume we are in a regime where $|V|\to\infty$ and there exists $\varepsilon>0$ such that $\frac{1}{|V|}\sum_{j\in V}h(j)\geq\varepsilon$ for all graphs. Then $p_0(h)\leq e^{-|V|\varepsilon}\to 0$, and thus*

$$\mathbb{E}\bigl[\mathrm{Vis}^{\text{real}}_r(s)\mid h\bigr]\longrightarrow\mu_s(h)\quad\text{as } r\to\infty \text{ and } |V|\to\infty.$$

4.3.2 Discussion

Corollary 4.2 shows that the discrepancy between probability-based and realized-outcome visibility is driven by a rare global event (all outcomes equal to 0). Thus, once the overall approval probability does not vanish, realized perception concentrates around the underlying probability model, and the two notions of visibility become asymptotically aligned. Conversely, when approvals are extremely sparse, the realized notion can be dominated by global randomness, which motivates treating the $\bar{H}=0$ event separately (or, in applications, adopting a small regularization of the threshold).

5. Topology and perceived discrimination

This section studies how network topology shapes perceived fairness at finite visibility radius. While Section 3 shows that perceived fairness converges to objective fairness as $r\to\infty$, the present section focuses on the structural mechanisms that generate systematic gaps at small or moderate visibility.

5.1 Homophily amplifies perceived discrimination

We begin with the effect of homophily. Intuitively, when individuals are more likely to connect to others from the same group, their local observations become less representative of the population, which can amplify perceived differences even under global parity.

5.1.1 Setup

We consider a K-group stochastic block model (SBM) with group proportions $(\pi_1,\dots,\pi_K)$. Conditional on group labels, edges are independent and satisfy

$$\mathbb{P}[A_{ij}=1\mid S_i=S_j]=p_{\text{in}}\text{ and }\mathbb{P}[A_{ij}=1\mid S_i\neq S_j]=p_{\text{out}},$$

with $p_{\text{in}}>p_{\text{out}}$ (assortative mixing). We define the edge-level homophily index

$$\rho:=\frac{p_{\text{in}}-p_{\text{out}}}{p_{\text{in}}+(K-1)p_{\text{out}}}\in(0,1). \tag{8}$$

In the two-group case ($K=2$), this reduces to $\rho=(p_{\text{in}}-p_{\text{out}})/(p_{\text{in}}+p_{\text{out}})$, and one may equivalently parameterize $p_{\text{in}}=p(1+\rho)$ and $p_{\text{out}}=p(1-\rho)$ for some $p>0$.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

Our homophily index (8) is a normalized within–between contrast that coincides, in the symmetric K-SBM with $p_{\text{in}} = a/n$ and $p_{\text{out}} = b/n$, with the spectral ratio $(a-b)/(a+(K-1)b)$ (equivalently λ_2/λ_1 of the SBM connectivity matrix), which is standard in the community-detection literature (Abbe, 2017, 2018). Note that assortative mixing with respect to categorical labels is classically quantified via mixing matrices and assortativity coefficients (Newman, 2002, 2003). Here we use the normalized SBM contrast (8) as a convenient edge-level homophily index.

Proposition 5.1 (Homophily and perceived fairness). *Assume DP holds at the population level,* $\mathbb{E}[h(i) \mid S_i = A] = \mathbb{E}[h(i) \mid S_i = B]$. *Then, for small* ρ, *the group-level perceived fairness gap at depth* $r = 1$, *from Equation (6), satisfies*

$$\Delta_1^{\text{prob}} = C\,\rho\,\Gamma(h) + o(\rho),$$

where $C > 0$ *depends only on* (π_A, π_B) *and* $\Gamma(h)$ *captures differences in local exposure across groups.*

5.1.2 Explicit form of $\Gamma(h)$ (neighbor-exposure contrast)

Write $\mu_s := \mathbb{E}[h(i) \mid S_i = s]$ for $s \in \{A, B\}$ and define the (depth-1) neighbor exposure

$$h_{\text{nbr}}(i) := \mathcal{E}_i[h] = \frac{1}{|N(i)|} \sum_{j \in N(i)} h(j) = \frac{1}{d_i} \sum_{j \in N(i)} h(j).$$

Let $\mu_s^{\text{nbr}} := \mathbb{E}[h_{\text{nbr}}(i) \mid S_i = s]$. The quantity appearing in Proposition 5.1 can be written as

$$\Gamma(h) := (\mu_A - \mu_B) - \left(\mu_A^{\text{nbr}} - \mu_B^{\text{nbr}}\right). \tag{9}$$

In particular, under DP_{prob} we have $\mu_A = \mu_B$, hence $\Gamma(h) = -(\mu_A^{\text{nbr}} - \mu_B^{\text{nbr}})$: the first-order amplification is entirely driven by differential neighborhood exposure.

5.1.3 Remark (DP edge cases)

When DP is imposed in a way that forces $\mu_A = \mu_B$ exactly, the mean-shift component of the linear response cancels. A nonzero perceived gap at $r = 1$ may then arise from higher-order terms and/or degree-weighted exposure effects (friendship-paradox-type sampling).

5.1.4 A general assortativity/modularity bound (depth $r = 1$)

Let $\boldsymbol{d} \in \mathbb{R}^n$ be the degree vector, $m := |E|$, and define the modularity matrix

$$B := A - \frac{\boldsymbol{d}\boldsymbol{d}^\top}{2m}.$$

Let $\boldsymbol{s} \in \{\pm 1\}^n$ encode group membership ($+1$ for $\boldsymbol{A}$, -1 for $\boldsymbol{B}$), and define the (normalized) assortativity/modularity

$$Q := \frac{1}{4m}\,\boldsymbol{s}^\top \boldsymbol{B}\boldsymbol{s}.$$

Under DP_{prob} and mild smoothness assumptions (implemented via a Lipschitz surrogate of the threshold map in $F_{\text{prob}}^{(1)}$), there exists a constant $C > 0$ such that

$$\left|\mathbb{E}\left[\Delta_1^{\text{prob}}\right]\right| \le C\,|Q|\,\text{Lip}(h). \tag{10}$$

Hence, higher assortativity/modularity amplifies perceived disparity even when population-level parity holds.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

5.1.5 Interpretation

Proposition 5.1 shows that homophily generates a *first-order amplification* of perceived unfairness, even when DP holds globally. The mechanism is purely structural: homophily changes the composition of neighborhoods, thereby distorting local comparisons without affecting population averages.

5.2 Degree bias and exposure effects

Homophily is not the only source of perceptual distortion. Even in the absence of assortative mixing, degree heterogeneity can bias local observations.

Proposition 5.2 (Degree bias at depth $r = 1$ (edge-weighted exposure)). *Assume $r = 1$ and define $\mathcal{E}_i[h] = \frac{1}{d_i} \sum_{j \in N(i)} h(j)$. Then the edge-weighted average neighborhood exposure satisfies*

$$\overline{\mathcal{E}}[h] := \frac{1}{2m} \sum_{i \in V} d_i \, \mathcal{E}_i[h] = \frac{1}{2m} \sum_{i \in V} d_i \, h(i) = \bar{h} + \frac{\mathrm{Cov}(d, h)}{\mathbb{E}[d]},$$

where $\bar{h} := \frac{1}{n} \sum_{i \in V} h(i)$, $n := |V|$, $2m := \sum_{i \in V} d_i$, and $\mathbb{E}[d] = 2m/n$ denotes the average degree (for $I \sim \mathrm{Unif}(V)$). In particular, $\overline{\mathcal{E}}[h] > \bar{h}$ iff $\mathrm{Cov}(d, h) > 0$.

5.2.1 Interpretation

This is a fairness analog of the friendship paradox: when h is positively correlated with degree, individuals are on average exposed to neighbors with higher acceptance probabilities. As a result, groups that are overrepresented among high-degree nodes may appear advantaged even under objective parity.

5.3 Clustering and variance reduction

We now turn to the role of clustering. While homophily and degree bias amplify perceived gaps, clustering has an opposite effect.

Proposition 5.3 (Clustering dampens perceived dispersion). *Fix a function $h : V \to \mathbb{R}$ and consider depth $r = 1$ neighborhood exposure*

$$\mathcal{E}_i[h] = \frac{1}{d_i} \sum_{j \in N(i)} h(j), \qquad i \in V.$$

Let G and G' be two graphs on the same vertex set V with the same degree sequence $(d_i)_{i \in V}$, and define the exposure vectors

$$e(G) := \left(\mathcal{E}_i^G[h]\right)_{i \in V}, \qquad e(G') := \left(\mathcal{E}_i^{G'}[h]\right)_{i \in V}.$$

Assume that there exists a doubly stochastic matrix $C \in \mathbb{R}^{|V| \times |V|}$ (i.e., $C\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^\top C = \mathbf{1}^\top$) such that

$$e(G') = C \, e(G).$$

Then:

(1) The mean exposure is preserved: $\frac{1}{|V|} \sum_{i \in V} e_i(G') = \frac{1}{|V|} \sum_{i \in V} e_i(G).$

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

(2) *For every convex function* $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\frac{1}{|V|}\sum_{i\in V}\varphi\big(e_i(G')\big) \;\le\; \frac{1}{|V|}\sum_{i\in V}\varphi\big(e_i(G)\big).$$

In particular, the dispersion of exposures cannot increase:

$$\mathrm{Var}\big(e(G')\big) \;\le\; \mathrm{Var}\big(e(G)\big).$$

The assumption $e(G') = Ce(G)$ formalizes the idea that increased local clustering/overlap makes neighborhood averages more similar by turning them into convex combinations of one another.

5.3.1 Interpretation

Clustering creates overlapping neighborhoods, which stabilizes local averages. As a result, individuals receive more similar signals about fairness, reducing extreme perceptions even when mean differences persist. This highlights that not all forms of segregation have the same perceptual impact: assortativity amplifies perceived gaps, whereas clustering smooths them.

5.4 Proof of Proposition 5.2

Recall that at depth $r = 1$,

$$\mathcal{E}_i[h] = \frac{1}{d_i}\sum_{j\in N(i)} h(j), \qquad d_i := |N(i)|, \qquad 2m = \sum_{i\in V} d_i.$$

Consider the *edge-weighted* average neighborhood exposure (equivalently, the average over all directed neighbor observations):

$$\overline{\mathcal{E}}[h] := \frac{1}{2m}\sum_{i\in V} d_i\,\mathcal{E}_i[h] = \frac{1}{2m}\sum_{i\in V}\sum_{j\in N(i)} h(j).$$

Swapping the order of summation,

$$\sum_{i\in V}\sum_{j\in N(i)} h(j) = \sum_{j\in V} h(j)\,|\{i\in V : j\in N(i)\}| = \sum_{j\in V} h(j)\,d_j,$$

since in an undirected graph node j appears in exactly d_j neighborhoods. Therefore,

$$\overline{\mathcal{E}}[h] = \frac{1}{2m}\sum_{j\in V} d_j\,h(j).$$

To relate this to the node-average $\bar{h} := \frac{1}{n}\sum_{j\in V} h(j)$, note that with uniformly random $I \sim \mathrm{Unif}(V)$ we have $\mathbb{E}[d_I] = \frac{2m}{n}$ and

$$\overline{\mathcal{E}}[h] = \frac{\mathbb{E}[d_I h(I)]}{\mathbb{E}[d_I]} = \mathbb{E}[h(I)] + \frac{\mathrm{Cov}(d_I, h(I))}{\mathbb{E}[d_I]} = \bar{h} + \frac{\mathrm{Cov}(d,h)}{2m/n}.$$

Hence $\overline{\mathcal{E}}[h] \neq \bar{h}$ whenever $\mathrm{Cov}(d,h) \neq 0$, which is the claimed degree-bias (friendship-paradox-type) effect.

5.5 Proof of Proposition 5.3

At depth $r = 1$, write the vector of neighborhood averages as

$$e := \big(\mathcal{E}_i[h]\big)_{i\in V} = D^{-1}A\,h,$$

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

where A is the adjacency matrix and $D = \mathrm{diag}(d_i)$. Let $\mathrm{Var}(e) := \frac{1}{n}\sum_{i\in V}(e_i - \bar{e})^2$ with $\bar{e} = \frac{1}{n}\sum_i e_i$.

We formalize the effect of increased local clustering (with degrees held fixed) through the following sufficient condition: the new neighborhood-average vector e' can be written as a *mixing* of the previous one,

$$e' = Ce,$$

for some *doubly stochastic* matrix C (i.e., $C\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^\top C = \mathbf{1}^\top$). This captures the idea that overlapping neighborhoods (stemming from triadic closure / increased clustering) make individuals' local averages more similar, as each e'_i becomes a convex combination of nearby e_j's.

Let $P := I - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ be the centering projection. Since $C\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^\top C = \mathbf{1}^\top$, we have $PC = CP$ and $Pe' = PCe = CPe$. Moreover, any doubly stochastic matrix is a convex combination of permutation matrices (Birkhoff–von Neumann theorem), hence it is a contraction in ℓ_2 on the subspace $\{\mathbf{1}\}^\perp$:

$$\|Cx\|_2 \le \|x\|_2 \quad \text{for all } x \text{ such that } \mathbf{1}^\top x = 0.$$

Applying this to $x = Pe$ gives

$$\|Pe'\|_2 = \|C(Pe)\|_2 \le \|Pe\|_2.$$

Finally, since $\mathrm{Var}(e) = \frac{1}{n}\|Pe\|_2^2$, we obtain

$$\mathrm{Var}(e') \le \mathrm{Var}(e),$$

that is, the dispersion of neighborhood averages cannot increase under such mixing. This formalizes the variance-reduction intuition: higher clustering creates overlapping neighborhoods, which stabilizes local averages and attenuates extreme perceived deviations.

5.6 Discussion

Taken together, these results show that perceived discrimination is shaped by multiple, distinct topological features. Homophily and degree heterogeneity amplify perceived unfairness by distorting local exposure, while clustering mitigates dispersion by stabilizing neighborhood-level observations. These mechanisms operate independently of population-level fairness and explain why objective parity may coexist with persistent perceived discrimination.

6. Numerical illustration and discussion

This section provides a numerical illustration of the mechanisms identified in Sections 3–5 and discusses their interpretation and implications. The simulations are intended as qualitative illustrations rather than exact tests of the theoretical results.

6.1 Numerical illustration

We simulate networks with n nodes partitioned into two groups A and B with proportions (π_A, π_B). Edges are generated according to a two-group SBM: conditional on group labels, pairs (i, j) are connected independently with probability p_{in} if $S_i = S_j$ and p_{out} if $S_i \neq S_j$. We parameterize homophily by $\rho \in [0, 1)$ by setting $p_{\text{in}} = p(1 + \rho)$ and $p_{\text{out}} = p(1 - \rho)$ (Section 5), where $p > 0$ controls the overall density. We vary ρ on a grid in $[0, \rho_{\max}]$ and generate multiple independent graph realizations for each value. Unless stated otherwise, we fix $n = 400$ and $(\pi_A, \pi_B) = (0.5, 0.5)$, and we use R independent graph realizations per value of ρ.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

6.1.1 Assigning acceptance probabilities

For each simulated graph, each node i is assigned an acceptance probability $h(i) \in [0, 1]$ according to one of the following scenarios:

- *Group-based rule:* $h(i)$ depends only on S_i (e.g., $h(i) = h_A$ on V_A and $h(i) = h_B$ on V_B).
- *Degree-based rule:* $h(i)$ depends only on the node degree d_i through a monotone mapping (e.g., a normalized degree score in $[0, 1]$).
- *Mixed rule:* $h(i)$ combines a group component and a degree component. Specifically, we set

$$h(i) = \Pi(\alpha\, h^{\mathrm{grp}}(i) + (1 - \alpha)\, h^{\mathrm{deg}}(i) + \varepsilon_i), \tag{11}$$

  where $\alpha \in (0, 1)$, $\Pi(x) = \min\{1, \max\{0, x\}\}$ clips to $[0, 1]$, and ε_i is a small mean-zero noise term. In the experiments we take $\alpha = 0.7$. We instantiate the group component by drawing $h^{\mathrm{grp}}(i) \sim \mathrm{Beta}(4, 2)$ if $S_i = A$ and $h^{\mathrm{grp}}(i) \sim \mathrm{Beta}(2, 4)$ if $S_i = B$, and we set $h^{\mathrm{deg}}(i)$ to be an increasing normalized function of d_i (e.g., rescaled ranks or $d_i / \max_j d_j$).

Unless stated otherwise, DP is *not* imposed exactly in these simulations (so the global gap may be nonzero).

6.1.2 Quantities reported

For each simulated network we compute: (i) the *global* (probability-based) fairness gap

$$\Delta_{\mathrm{global}}(h) := \frac{1}{|V_A|} \sum_{i \in V_A} h(i) - \frac{1}{|V_B|} \sum_{i \in V_B} h(i),$$

and (ii) the *perceived* fairness gap at visibility radius $r = 1$, $\Delta_1^{\mathrm{prob}}(h) = \mathrm{Vis}_1^{\mathrm{prob}}(A; h) - \mathrm{Vis}_1^{\mathrm{prob}}(B; h)$ (Equation (6)). Results are displayed across realizations; in Figure 1, each point corresponds to one network realization and the solid line is a LOESS smoother.

6.1.3 Main observation

Figure 1 plots the perceived and global fairness gaps as functions of the homophily index ρ. For small to moderate values of ρ, perceived unfairness increases approximately linearly with homophily, consistent with the first-order expansion in Proposition 5.1. At larger values of ρ, the perceived gap may decrease as neighborhoods become nearly homogeneous, reducing cross-group comparisons. This non-monotonic behavior does not contradict the theoretical results, which characterize local behavior around $\rho = 0$.

In summary, the simulations are consistent with a positive local dependence on ρ (as captured by Proposition 5.1), but also show that at high homophily the perceived gap may saturate or decrease.

6.2 Interpretation and policy implications

We now discuss how the proposed framework should be interpreted and how it may inform policy analysis.

6.2.1 Perceived versus objective fairness

Perceived unfairness is not a substitute for objective fairness criteria such as DP or equality of outcomes. Rather, it captures how individuals experience fairness through local social comparisons. As shown in Sections 4 and 5, objective parity may coexist with substantial perceived unfairness

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

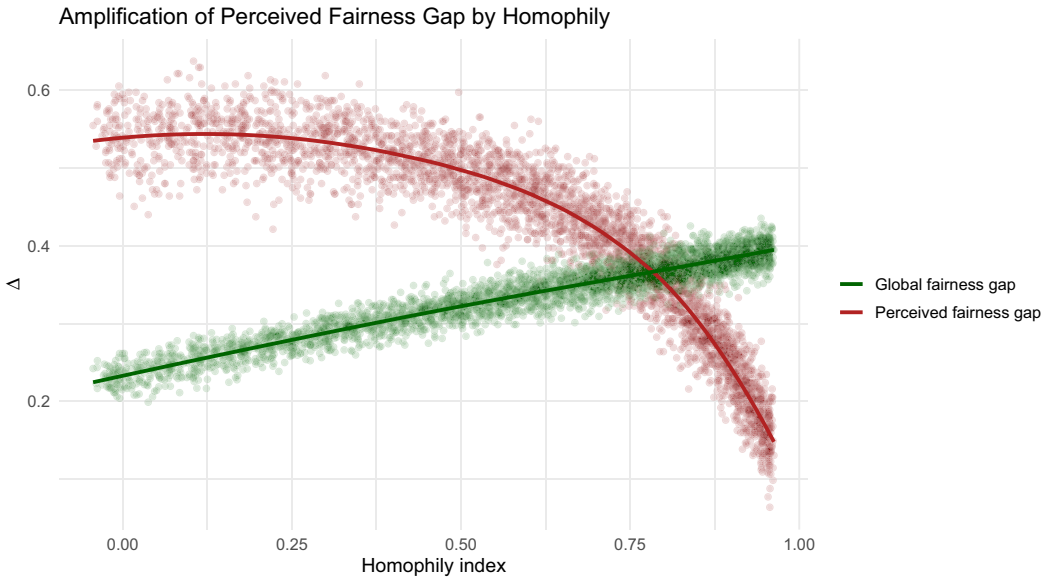


Figure 1. Perceived and global fairness gaps as functions of the homophily index ρ. Each point corresponds to one network realization; solid lines show LOESS smoothing. For small to moderate homophily, perceived unfairness increases approximately linearly, while non-monotonic behavior may arise at high homophily due to neighborhood homogenization.

when network structure distorts local exposure, and conversely low perceived unfairness may arise in highly segregated settings despite persistent demographic disparities.

6.2.2 Why perceptions matter

Perceptions of fairness influence behavior, trust, and participation in social and economic systems. Even when decisions satisfy formal fairness constraints, persistent perceived unfairness may undermine legitimacy or compliance. Our results provide a structural explanation for such mismatches by showing how network topology shapes local observations.

6.2.3 Policy objectives

Policy objectives concerning outcomes and perceptions are conceptually distinct. For realized outcomes, the goal is to reduce disparities in expected values while improving outcomes for all groups. For perceptions, the goal is to reduce systematic differences in how fairness is experienced locally. Our framework is intended to inform the latter objective by identifying how network features such as homophily, degree heterogeneity, and clustering affect perceived fairness. It does not advocate relaxing outcome-based fairness constraints.

6.2.4 Limits and extensions

Finally, we emphasize that reducing perceived unfairness through segregation or information restriction is not a desirable policy solution. While such mechanisms may mechanically reduce perceived gaps, they do so by limiting exposure rather than addressing underlying inequalities. Future work could extend the present framework to attribute-based homophily, dynamic networks, or learning processes, and study how perception and behavior co-evolve over time.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

7. Discussion and extensions

This paper studies perceived fairness as a network-dependent phenomenon. While objective fairness criteria are defined at the population level, individuals evaluate fairness through local comparisons shaped by network structure. Our results show that these two perspectives may diverge systematically.

7.1 Interpretation

Perceived unfairness is not a substitute for objective fairness. Low perceived unfairness may coexist with substantial demographic disparities in segregated networks, while perceived unfairness may persist even under DP. The framework highlights this potential misalignment without advocating any normative trade-off.

7.2 Policy implications

Policy objectives concerning outcomes and perceptions are distinct. Reducing disparities in expected outcomes remains essential, but addressing persistent perceived unfairness may require complementary interventions that modify exposure or information aggregation. Network structure thus plays a critical role in how fairness is experienced.

7.3 Extensions

Several extensions are natural. Future work could study attribute-based homophily, dynamic or adaptive networks, and learning processes in which perceptions feed back into behavior. Another direction is to couple perceived fairness with endogenous network formation.

7.4 Conclusion

Perceived fairness is a structural property of networked systems. By formalizing how topology shapes local comparisons, this paper provides a unified framework to study the gap between objective and perceived fairness in social and algorithmic networks.

8. Conclusion

We proposed a mathematical framework linking network structure and perceived fairness. Our analysis highlights how local perception can deviate from global fairness even when algorithms are unbiased in aggregate. Future research will connect these theoretical insights with empirical data from collaborative or decentralized systems.

Acknowledgements. AC thanks the audience of the lecture at the at the "decentralized insurance and risk sharing" workshop at DePaul University, Chicago, in 2024. AC acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [RGPIN-2019-07077], and the SCOR Foundation for Science.

Competing interests. None.

References

Abbe, E. (2017). Community detection and stochastic block models. arXiv:1703.10146.

Abbe, E. (2018). Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, *18*, 1–86.

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press

Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and machine learning: Limitations and opportunities*. MIT Press.
Brown, C., Matthews, K. A., Bromberger, J. T., & Chang, Y. (2006). The relation between perceived unfair treatment and blood pressure in a racially/ethnically diverse sample of women. *American Journal of Epidemiology*, *164*, 257–262.
Cantwell, G. T., Kirkley, A., Newman, M. E. J., & Estrada, E. (2021). The friendship paradox in real and model networks. *Journal of Complex Networks*, *9*, cnab011.
Charpentier, A., & Ratz, P. (2025). Linear risk sharing on networks. arXiv preprint arXiv:2509.21411.
De Lara, L., González-Sanz, A., Asher, N., Risser, L., & Loubes, J. M. (2024). Transport-based counterfactual models. *Journal of Machine Learning Research*, *25*, 1–59.
Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference* (pp. 214–226).
Fernandes Machado, A., Charpentier, A., & Gallic, E. (2025). Sequential conditional transport on probabilistic graphs for interpretable counterfactual fairness. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 19358–19366).
Gonzalez, D., McDaniel, M., Kenney, G. M., & Skopec, L. (2021). *Perceptions of Unfair Treatment or Judgment Due to Race or Ethnicity in Five Settings*. Washington, DC: Urban Institute.
Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. *Advances in Neural Information Processing Systems*, *29*, 3315–3323.
Kusner, M., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4069–4079). Curran Associates Inc.
Mayerhoffer, D. M., & Schulz, J. (2022). Perception and privilege. *Applied Network Science*, *7*, 32. https://doi.org/10.1007/s41109-022-00467-x.
McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444.
Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, *67*, 026126.
Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*, 208701.
Pascoe, E. A., & Richman, L. S. (2009). Perceived discrimination and health: a meta-analytic review. *Psychological Bulletin*, *135*, 531.
Schmitt, M. T., Branscombe, N. R., Postmes, T., & Garcia, A. (2014). The consequences of perceived discrimination for psychological well-being: A meta-analytic review. *Psychological Bulletin*, *140*, 921.
Wu, X. Z., Percus, A. G., & Lerman, K. (2017). Neighbor-neighbor correlations explain measurement bias in networks. *Scientific Reports*, *7*, 5576.
Zhou, Z., Liu, T., Bai, R., Gao, J., Kocaoglu, M., & Inouye, D. I. (2024). Counterfactual fairness by combining factual and counterfactual predictions. In *Proceedings of the 38th International Conference on Neural Information Processing Systems* (pp. 47876–47907). Curran Associates Inc.

Cite this article: Charpentier A. (2026). Perceived fairness in networks. *Network Science* **14**, 1–15. https://doi.org/10.1017/nws.2026.10030

https://doi.org/10.1017/nws.2026.10030 Published online by Cambridge University Press