Synergistic barriers to algorithmic recourse in healthcare and administrative systems

A.C Demidont ORCID[1*]

[1]Nyx Dynamics LLC, Philadelphia, Pennsylvania, United States of America.

Abstract

Algorithmic decision systems mediate access to healthcare, credit, employment and housing, yet individuals who experience adverse decisions face multi-stage barriers when seeking recourse. We formalize these barriers as a series-structured system with 11 empirically parameterized stages across three layers (data integration, data accuracy and institutional access) and prove that single-barrier interventions are bounded by baseline system success. Under baseline parameterization derived from federal datasets and peer-reviewed algorithmic audit studies, end-to-end recourse probability is 0.0018%. Removing any single barrier yields negligible improvement (<0.02%). Factorial decomposition reveals that the three-way cross-layer interaction accounts for 87.6% of achievable improvement, confirmed by Shapley attribution, Sobol sensitivity analysis and bootstrap resampling ($\boldsymbol{n = 1{,}000}$). These results provide a structural explanation for the limited impact of incremental reforms and support coordinated multi-layer intervention approaches for clinical AI governance and algorithmic fairness.

Keywords: algorithmic fairness, clinical AI governance, series reliability systems, interaction effects, healthcare algorithms

1 Introduction

Healthcare algorithms increasingly determine which patients receive timely interventions, with documented instances where cost-based proxy variables reduced identification of high-risk Black patients by approximately 62%[1]. In clinical settings, algorithmic decisions include risk stratification scores, automated triage, prior authorization determinations, resource allocation and eligibility assessments; when these produce adverse outcomes, patients must navigate error detection, data correction and

institutional appeals processes that we term *clinical AI recourse*. Analogous barrier structures arise in credit, employment and housing: approximately 75% of US employers use recruitment management systems to filter candidates before human review[2], and credit scoring algorithms shape financial access for virtually all Americans[3]. When these systems encode discrimination, individuals face multi-stage procedural frictions when seeking recourse. We focus on healthcare AI as the primary context and use credit, employment and housing domains as supporting analogs where empirical barrier data are available[4, 5] that span institutional boundaries regulated by separate agencies, creating regulatory fragmentation[6].

The sociotechnical nature of algorithmic systems means discrimination arises not from any single component but from interactions among design choices, data practices, institutional contexts and deployment environments[7]. This parallels well-characterized phenomena in reliability engineering, where series systems exhibit failure dynamics dominated by component interactions rather than individual failure rates[8, 9], and in patient safety, where alignment of multiple barrier failures produces system-level adverse events[10]. Despite sustained regulatory effort targeting individual barriers—improving data accuracy, providing legal resources, mandating algorithmic audits—discrimination persists, suggesting current approaches may misunderstand the problem's structure.

We hypothesized that barriers to algorithmic recourse function synergistically such that incremental single-layer interventions produce near-zero absolute gains unless coordinated across layers. Specifically, we ask: in a multi-stage barrier system governing recourse from adverse algorithmic decisions, when do single-barrier reforms have negligible absolute impact, and how much of achievable improvement is attributable to cross-layer interactions? We formalize this as a series-structured barrier system, prove closed-form bounds on single-barrier intervention effects and quantify the interaction structure using factorial decomposition, Shapley attribution and global sensitivity analysis. This framework provides a quantitative basis for coordinated intervention design in clinical AI governance and beyond.

2 Results

2.1 Baseline system characteristics

The 11-barrier model (Table 2) yielded a baseline success probability of 0.0018%, indicating that fewer than 2 in 100,000 individuals successfully navigate all barriers to resolve adverse algorithmic decisions.

2.2 Analytical bounds on single-barrier interventions

Proposition 1 establishes that removing barrier j yields absolute improvement $\Delta_j = P \cdot (1/p_j - 1)$, which is linear in baseline success P. Counterfactual analysis confirmed this prediction: all individual barrier removal effects were <0.02%, with the maximum observed for Legal Knowledge Gap ($\Delta = 0.0055\%$). This value is consistent with the analytical bound $P \cdot (1/0.25 - 1) \approx 0.0054\%$ for the barrier with lowest pass probability (Fig. 1).

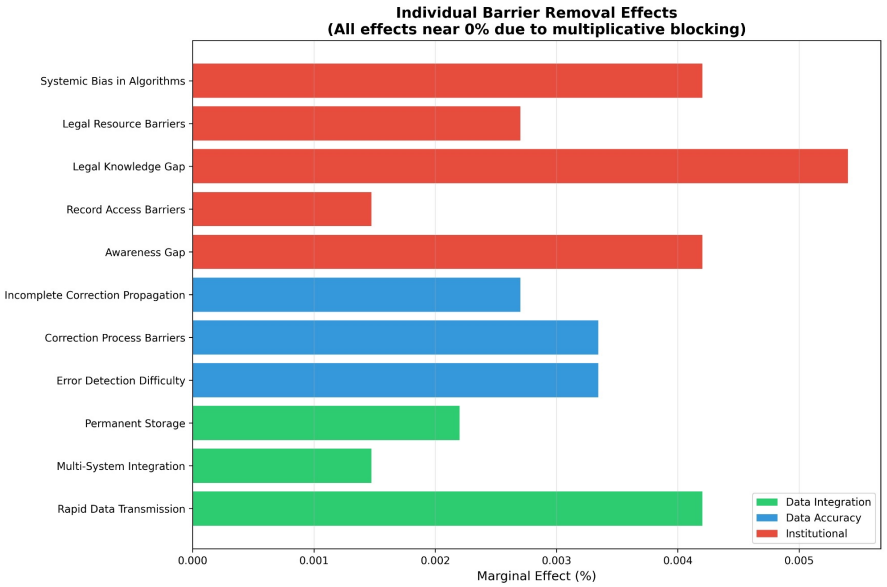


Fig. 1 Individual barrier removal effects. Counterfactual analysis showing marginal effect on success probability of removing each barrier while all others remain. All effects approach zero (<0.02%), consistent with Proposition 1. Error bars: 95% bootstrap CI (n = 1,000). Parameter sources in Tables 2–3.

2.3 Strategy equivalence

Comparison of barrier removal strategies—forward, backward, greedy by marginal impact and random ordering—revealed characteristic threshold trajectories with near-zero improvement until removal of final barriers (Fig. 2). Strategy equivalence (ANOVA: $F = 0.07$, $P = 0.98$) confirms that removal ordering is irrelevant; only completeness determines outcome.

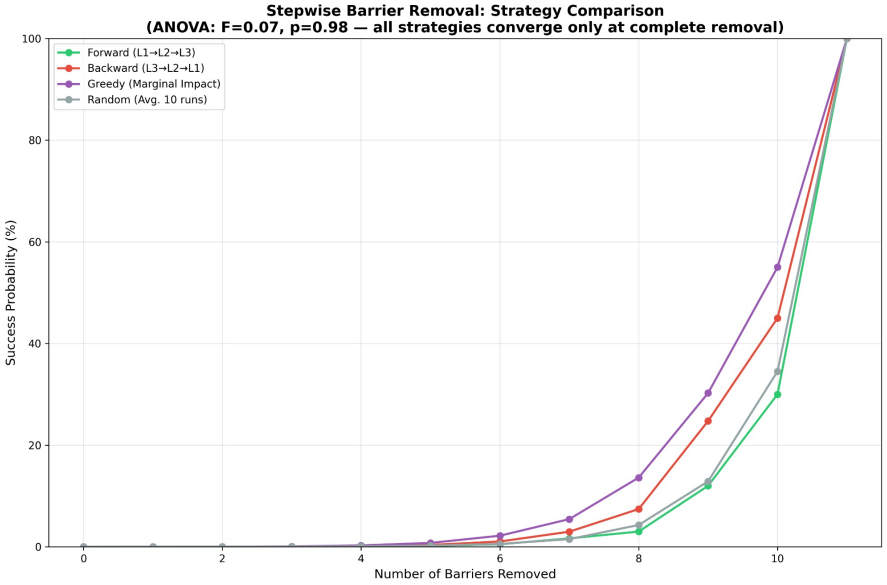


Fig. 2 Strategy comparison for barrier removal. Barrier removal strategies (forward, backward, greedy, random) showing threshold trajectories. Strategy equivalence (ANOVA: $F = 0.07$, $P = 0.98$) confirms removal ordering is irrelevant; only completeness determines outcome.

2.4 Cross-layer interaction dominance

Factorial analysis revealed interaction-dominant dynamics as predicted by Proposition 2 (Fig. 3, Table 1). The three-way interaction accounted for 87.6% of total effect variance. Main effects contributed minimal variance: Layer 1 (0.0%), Layer 2 (0.0%), Layer 3 (0.3%).

Table 1 ANOVA decomposition of layer removal effects.

Effect	Δ Success (%)	% of Total
Layer 1 (Data Integration)	0.0	0.0
Layer 2 (Data Accuracy)	0.0	0.0
Layer 3 (Institutional)	0.3	0.3
L1 × L2	3.2	3.4
L1 × L3	7.6	8.0
L2 × L3	0.5	0.5
L1 × L2 × L3 (Three-way)	83.4	87.6
Total	95.0	100.0

Percentages normalized to total explainable variance (95.0% of total modeled effect).

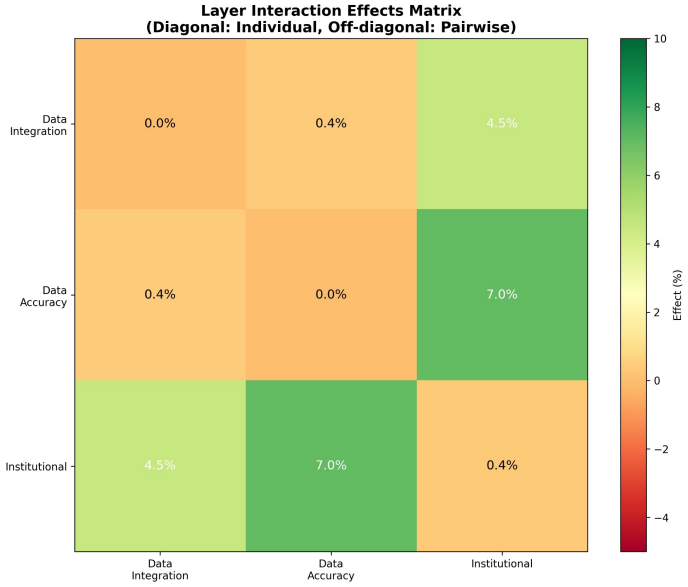


Fig. 3 Layer interaction effects. Heatmap of main effects (diagonal) and pairwise interactions (off-diagonal). The three-way interaction (87.6%, not shown) dominates, consistent with Proposition 2.

2.5 Sensitivity analysis and robustness

Sobol indices confirmed interaction dominance, with total-order indices substantially exceeding first-order indices for all barriers (Fig. 4). Morris screening identified Legal Knowledge Gap and Systemic Bias in Algorithms as having highest nonlinear/interaction involvement. Signal-to-noise ratio remained positive (>0 dB) up to 25% parameter noise (Fig. 5). Bootstrap validation ($n = 1{,}000$) demonstrated 100% robustness of all principal findings across plausible parameterizations.

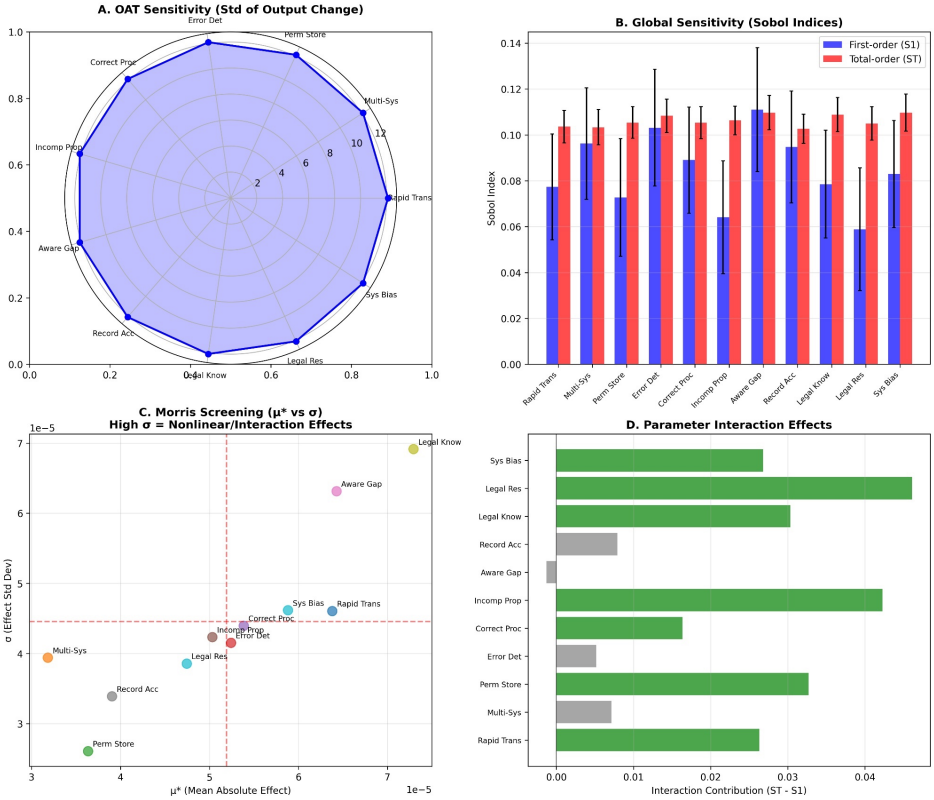


Fig. 4 Global sensitivity analysis. a) OAT sensitivity radar. **b)** Sobol S1 and ST indices with 95% CI. **c)** Morris elementary effects. **d)** Interaction contribution (ST − S1) per barrier.

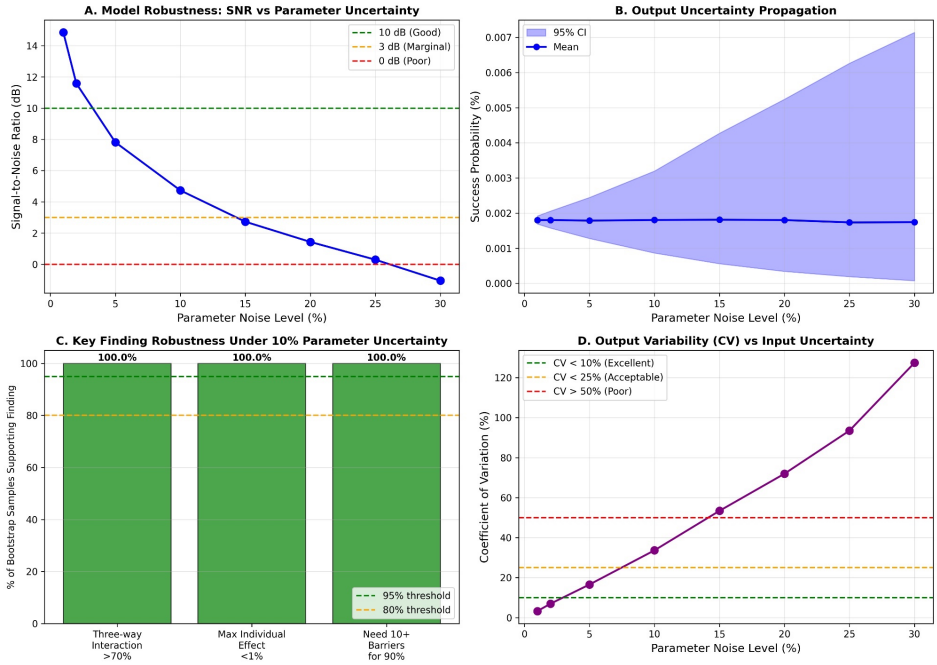


Fig. 5 Robustness of model findings under parameter uncertainty. a) Signal-to-noise ratio (SNR) as a function of parameter noise level; SNR remains positive (>0 dB) up to approximately 25% noise. **b)** Output uncertainty propagation showing mean success probability and 95% confidence interval across noise levels. **c)** Bootstrap robustness showing percentage of 1,000 resamples supporting each key finding at ±10% perturbation; all findings achieve 100% robustness. **d)** Coefficient of variation (CV) of output as a function of input uncertainty, with quality thresholds indicated.

2.6 Shapley value attribution

Shapley decomposition assigned fair responsibility accounting for all possible removal orderings. Top contributors: Legal Knowledge Gap (11.3%), Rapid Data Transmission (11.1%), Correction Process Barriers (10.7%), Systemic Bias in Algorithms (10.5%) and Awareness Gap (10.0%). Unlike marginal effects—which approach zero under multiplicative blocking (Proposition 1)—Shapley values reveal causal contribution under interaction-dominant dynamics.

3 Discussion

Under a series-structured barrier model with empirically derived parameters, algorithmic discrimination operates as a synergistic system where barriers reinforce each other, rendering incremental reform structurally insufficient. The dominant three-way interaction (87.6%) provides a quantitative explanation for the observed pattern whereby incremental policy reforms—including targeted improvements to data accuracy, legal access and algorithmic auditing—have had limited impact: under synergistic barrier conditions, targeting any subset of barriers leaves the multiplicative blocking structure substantially intact.

The closed-form bound established in Proposition 1 makes precise what has been intuitive in advocacy contexts: when system-level success probability is extremely low, no single-barrier intervention can produce meaningful absolute improvement. This is an algebraic property of series systems, not contingent on specific parameter values. When health systems propose that improving algorithmic audit quality or adding patient notification requirements will substantially reduce algorithmic harm, Proposition 1 demonstrates that these claims are mathematically bounded by baseline success probability regardless of the magnitude of the barrier addressed.

The distinction between probability-scale and log-scale analysis (Proposition 2) is substantively important for clinical AI governance. On the log scale, barriers contribute additively—a framing suggesting incremental reform could accumulate benefits. On the probability scale, where patient outcomes and equity metrics are evaluated, multiplicative stacking produces threshold behavior: improvement remains negligible until most barriers are removed, then increases rapidly. This threshold pattern is a generic property of series systems[8] and predicts that single-lever fairness interventions will yield small absolute effects whenever baseline recourse probability is low.

The synergistic structure parallels combination therapy in infectious disease, where incomplete antiretroviral therapy is insufficient because the virus exploits any untargeted mechanism[11]. Similarly, incomplete policy intervention addresses visible symptoms while leaving the underlying multiplicative structure intact. For healthcare AI specifically, this suggests that fairness interventions limited to the algorithmic layer (bias mitigation, audit requirements) will have negligible population-level impact without simultaneous attention to data integration practices and institutional access to recourse mechanisms. The regulatory fragmentation governing algorithmic systems—CFPB for credit reporting, EEOC for employment, HHS for healthcare

algorithms—mirrors the barrier fragmentation in our model and may structurally impede the coordinated approach our analysis indicates is necessary.

Interaction-dominant barrier structures likely extend to other clinical AI domains where administrative, technical and institutional processes create series-structured requirements for patient recourse. Sociotechnical systems in which fairness depends on alignment of multiple organizational and technical components[7] may exhibit analogous dynamics. Recent work on clinical AI fairness frameworks[12] and bias recognition strategies[13] has emphasized multi-level approaches; our quantitative framework provides formal support for this intuition and a measurement methodology for evaluating whether proposed interventions address sufficient layers of the barrier system.

Several limitations should be noted. This model is a formal abstraction and does not directly observe individual-level resolution trajectories. Barrier probabilities are derived or mapped (when direct measurement is unavailable) from heterogeneous empirical sources rather than a single integrated dataset; however, each parameter is traceable to publicly available federal reports or peer-reviewed studies (Tables 2–3), with explicit derivation logic documented in the repository. The principal conclusions are analytical properties of the series structure (Propositions 1–2) that hold across wide parameter perturbations, as confirmed by bootstrap resampling and global sensitivity analysis. The model assumes multiplicative independence between barriers; correlation between barriers could produce either stronger or weaker synergistic effects depending on correlation structure. The model does not capture feedback loops or heterogeneity across populations. Future work should integrate longitudinal administrative data—including healthcare claims, appeal outcomes and algorithmic audit records—to empirically estimate joint barrier transition probabilities and validate the series-system structure against observed recourse outcomes.

In summary, under a series-structured barrier model, algorithmic discrimination involves synergistic dynamics that render incremental reform insufficient in isolation. The dominant three-way interaction provides a mathematical basis for coordinated, multi-layer intervention. Effective mitigation of algorithmic harm in healthcare and beyond likely requires simultaneous intervention across data integration, data accuracy and institutional access barriers rather than isolated single-layer reforms.

4 Methods

4.1 Barrier model specification

We modeled algorithmic discrimination as requiring successful passage through 11 sequential barriers organized into three layers (Table 2). Layer 1 (Data Integration) comprises barriers related to speed and breadth of adverse data propagation across interconnected systems. Layer 2 (Data Accuracy) comprises barriers to detecting, correcting and propagating corrections of erroneous data. Layer 3 (Institutional) comprises barriers to awareness, record access, legal knowledge, legal resources and algorithmic bias in decision systems.

Table 2 Barrier parameters, layer assignments and primary sources.

Layer	Barrier	Pass prob.	Primary source(s)
Data Integration	Rapid Data Transmission	0.30	CFPB 2022[3, 14]
	Multi-System Integration	0.55	FTC 2013[15]
	Permanent Storage	0.45	CFPB 2022; FCRA §605[3, 14]
Data Accuracy	Error Detection Difficulty	0.35	FTC 2013[15]
	Correction Process Barriers	0.35	FTC 2013; CFPB 2022[15, 16]
	Incomplete Correction Propagation	0.40	CFPB 2022; FTC 2015[16]
Institutional	Awareness Gap	0.30	LSC 2022[17]
	Record Access Barriers	0.55	CFPB 2022; FCRA §612[15]
	Legal Knowledge Gap	0.25	LSC 2022[17]
	Legal Resource Barriers	0.40	LSC 2022[17]
	Systemic Bias in Algorithms	0.30	Obermeyer *et al.* 2019[1]

Each row cites its primary empirical source. Full derivation logic, key statistics, mapping rationale and plausible uncertainty ranges are provided in Table 3.

Table 3 Barrier definitions, empirical sources and derivation logic.

Barrier	p	Range	Primary source	Key statistic	Derivation logic	Layer
Rapid Data Trans.	0.30	0.20–0.40	CFPB 2022	1–2 day transmission	∼30% intervene	L1
Multi-Sys. Integ.	0.55	0.40–0.65	FTC 2013	20% error on $\geq$1 report	Composite non-propagation	L1
Permanent Storage	0.45	0.35–0.55	CFPB; FCRA	7-yr retention	∼45% unexpired	L1
Error Detection	0.35	0.25–0.45	FTC 2013	26% material errors	∼35% identify	L2
Correction Process	0.35	0.25–0.45	FTC; CFPB	37% fully resolved	∼35% corrected	L2
Incompl. Corr. Prop.	0.40	0.30–0.50	CFPB; FTC	No auto-propagation	∼40% propagate	L2
Awareness Gap	0.30	0.20–0.40	LSC 2022	92% justice gap	∼30% aware	L3
Record Access	0.55	0.45–0.65	CFPB; FCRA	Free annual reports	∼55% navigate	L3
Legal Knowledge	0.25	0.15–0.35	LSC 2022	39% trust legal system	∼25% know rights	L3
Legal Resources	0.40	0.30–0.50	LSC 2022	46% cite cost	∼40% access	L3
Systemic Bias	0.30	0.20–0.40	Obermeyer 2019	62% undertriage	∼30% unbiased	L3

Mapping decisions were conservative, favouring higher pass probabilities to avoid overstating interaction effects. L1 = Data Integration; L2 = Data Accuracy; L3 = Institutional.

4.2 Empirical parameter derivation

Pass probabilities were derived or mapped (when direct measurement is unavailable) from publicly available federal datasets and peer-reviewed studies, with explicit uncertainty ranges (Table 3). Mapping decisions were conservative, favoring higher pass probabilities where ambiguity existed to avoid overstating interaction effects. Layer 1 probabilities were derived from CFPB and FTC data on credit reporting system dynamics[3, 14]. Layer 2 probabilities were derived from FTC and CFPB dispute outcome data[15, 16]. Layer 3 probabilities were derived from Legal Services Corporation data on civil legal needs and the Obermeyer *et al.* algorithmic audit[1, 17]. Full derivation logic for all 11 parameters is documented in the public repository.

4.3 Mathematical framework

Under the multiplicative barrier model, the probability of successful cascade completion is:

$$P(\text{success}) = \prod_{i=1}^{11} p_i \tag{1}$$

The effect of removing barrier j is $\Delta_j = P(\text{success} \mid p_j = 1) - P(\text{success})$.

Proposition 1 (Single-barrier improvement bound) *Removing barrier j yields $\Delta_j = P \cdot (1/p_j - 1)$. Thus absolute improvement is linear in P, the baseline success. When $P \approx 1.8 \times 10^{-5}$, every Δ_j is necessarily small regardless of which barrier is targeted. This is an algebraic property of series systems, not a simulation finding.*

Proposition 2 (Interaction dominance on the probability scale) *Define layer aggregate probabilities $L_k = \prod_{i \in k} p_i$ for layers $k = 1, 2, 3$, so that $P = L_1 \cdot L_2 \cdot L_3$. On the probability scale, factorial contrasts generate higher-order interactions because the response function is multiplicative. On the log scale, $\log P = \sum \log p_i$ is additive and interactions vanish; however, policy outcomes and equity metrics are evaluated on the probability scale, where multiplicative stacking produces threshold behavior and interaction dominance[8].*

4.4 Shapley value attribution

To fairly attribute system effect to individual barriers, we employed Shapley value decomposition from cooperative game theory[18], assigning each barrier its average marginal contribution across all possible removal orderings.

4.5 Sensitivity analysis

We conducted comprehensive sensitivity analysis including one-at-a-time sensitivity ($\pm 10\%$), Sobol global sensitivity indices[19] with Saltelli sampling (n = 1,024) and Morris screening[20]. All barrier probabilities were perturbed across $\pm 33\%$ multiplicative ranges and across empirically bounded intervals (Table 3). Bootstrap resampling (n = 1,000) confirmed robustness of all principal findings.

4.6 Software and reproducibility

All analyses were conducted in Python 3.10+ using NumPy, SciPy, Matplotlib and SALib[21]. Code is available at https://github.com/Nyx-Dynamics/algorithmic-bias-epidemiology-academic (random seed: 42).

4.7 Use of artificial intelligence

Large language models (Anthropic Claude and OpenAI ChatGPT) were used for literature search support and language refinement. No generative-AI–generated images or figures were used; all figures were produced programmatically using Matplotlib. Jetbrains Junie was utilized for code correction, Zotero AI for reference management

and the manuscript was prepared using Overleaf AI for LaTeX code correction. All analyses, interpretation and conclusions were conducted independently by the author.

4.8 Ethics statement

This study used publicly available, aggregate federal datasets and published peer-reviewed statistics. No individual-level data were collected or analyzed and no human subjects were enrolled. Institutional review board approval was not required.

Data availability

All data files are available in CSV format at https://github.com/Nyx-Dynamics/algorithmic-bias-epidemiology-academic/tree/main/data/processed.

Code availability

All analysis code is available at https://github.com/Nyx-Dynamics/algorithmic-bias-epidemiology-academic. Key scripts: `sensitivity_analysis.py` (Sobol, Morris, bootstrap); `barrier_visualization.py` (figure generation and counterfactual analysis). Environment: Python 3.10+, NumPy, SciPy, Matplotlib, SALib. Random seed: 42.

Acknowledgements

The author thanks the open-source software communities whose tools made this analysis possible. This work received no external funding.

Author contributions

A.C.D. conceived the study, developed the mathematical framework, conducted all analyses, interpreted results and wrote the manuscript.

Competing interests

A.C.D. reports prior employment with Gilead Sciences from January 2020 through November 2024, and prior ownership of company stock, which was fully divested in December 2024. Gilead Sciences had no role in this study. A.C.D. is the owner of Nyx Dynamics LLC, a consulting company. This research was conducted independently and was not produced as part of any paid consulting engagement. No other competing interests are declared.

References

[1] Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).

[2] Fuller, J. B., Raman, M., Sage-Gavin, E. *et al.* Hidden Workers: Untapped Talent (Harvard Business School Project on Managing the Future of Work, 2021).

[3] Consumer Financial Protection Bureau. Consumer Credit Reports: A Study of Medical and Non-Medical Collections (CFPB, 2022).

[4] Wachter, S., Mittelstadt, B. & Russell, C. Counterfactual explanations without opening the black box: automated decisions and the GDPR. *Harv. J. Law Technol.* **31**, 841–887 (2018).

[5] Ustun, B., Spangher, A. & Liu, Y. Actionable recourse in linear classification. In *Proc. Conference on Fairness, Accountability, and Transparency* 10–19 (ACM, 2019).

[6] Equal Employment Opportunity Commission. Enforcement Guidance on the Consideration of Arrest and Conviction Records in Employment Decisions (EEOC, 2012).

[7] Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S. & Vertesi, J. Fairness and abstraction in sociotechnical systems. In *Proc. Conference on Fairness, Accountability, and Transparency* 59–68 (ACM, 2019).

[8] Rausand, M. & Høyland, A. *System Reliability Theory: Models, Statistical Methods, and Applications* 2nd edn (Wiley, 2004).

[9] Billinton, R. & Allan, R. N. *Reliability Evaluation of Engineering Systems: Concepts and Techniques* 2nd edn (Springer, 1992).

[10] Reason, J. Human error: models and management. *BMJ* **320**, 768–770 (2000).

[11] Volberding, P. A. & Deeks, S. G. Antiretroviral therapy and management of HIV infection. *Lancet* **376**, 49–62 (2010).

[12] Hoche, M., Mineeva, O., Rätsch, G., Vayena, E. & Blasimme, A. What makes clinical machine learning fair? A practical ethics framework. *PLOS Digit. Health* **4**, e0000728 (2025).

[13] Alderman, J. E. *et al.* Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together Consensus Recommendations. *NEJM AI* **2**, 10.1056/AIp2401088 (2025).

[14] Consumer Financial Protection Bureau. Disputes on Consumer Credit Reports (CFPB, 2022).

[15] Federal Trade Commission. Report to Congress Under Section 319 of the Fair and Accurate Credit Transactions Act of 2003: Fifth Interim Report (FTC, 2013).

[16] Consumer Financial Protection Bureau. CFPB Report on Consumer Complaint Response Deficiencies of the Big Three Credit Bureaus (CFPB, 2022).

[17] Legal Services Corporation. The Justice Gap: The Unmet Civil Legal Needs of Low-Income Americans (LSC, 2022).

[18] Shapley, L. S. A value for n-person games. In *Contributions to the Theory of Games* Vol. 2 (eds Kuhn, H. W. & Tucker, A. W.) 307–317 (Princeton Univ. Press, 1953).

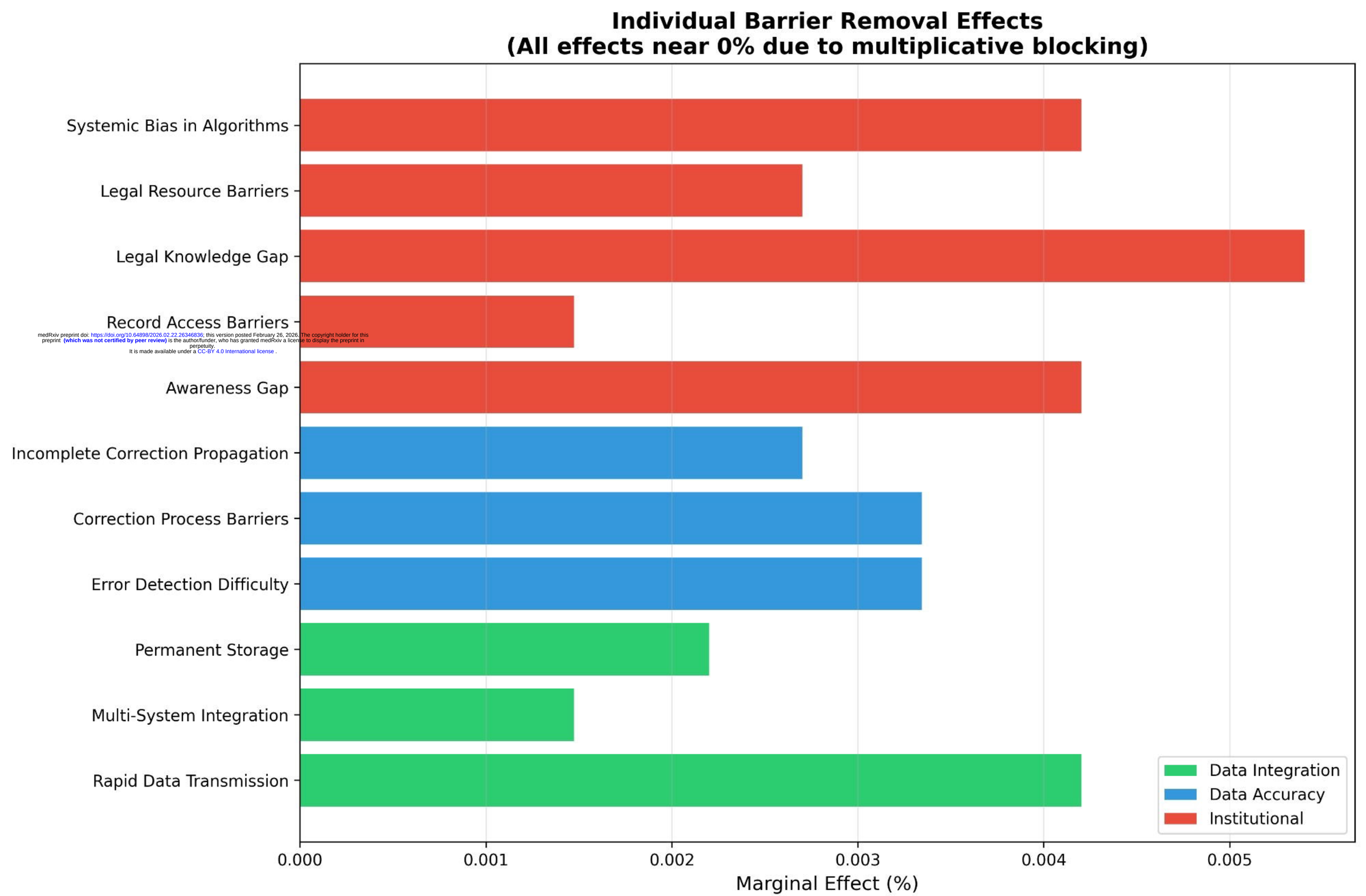
[19] Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **55**, 271–280 (2001).

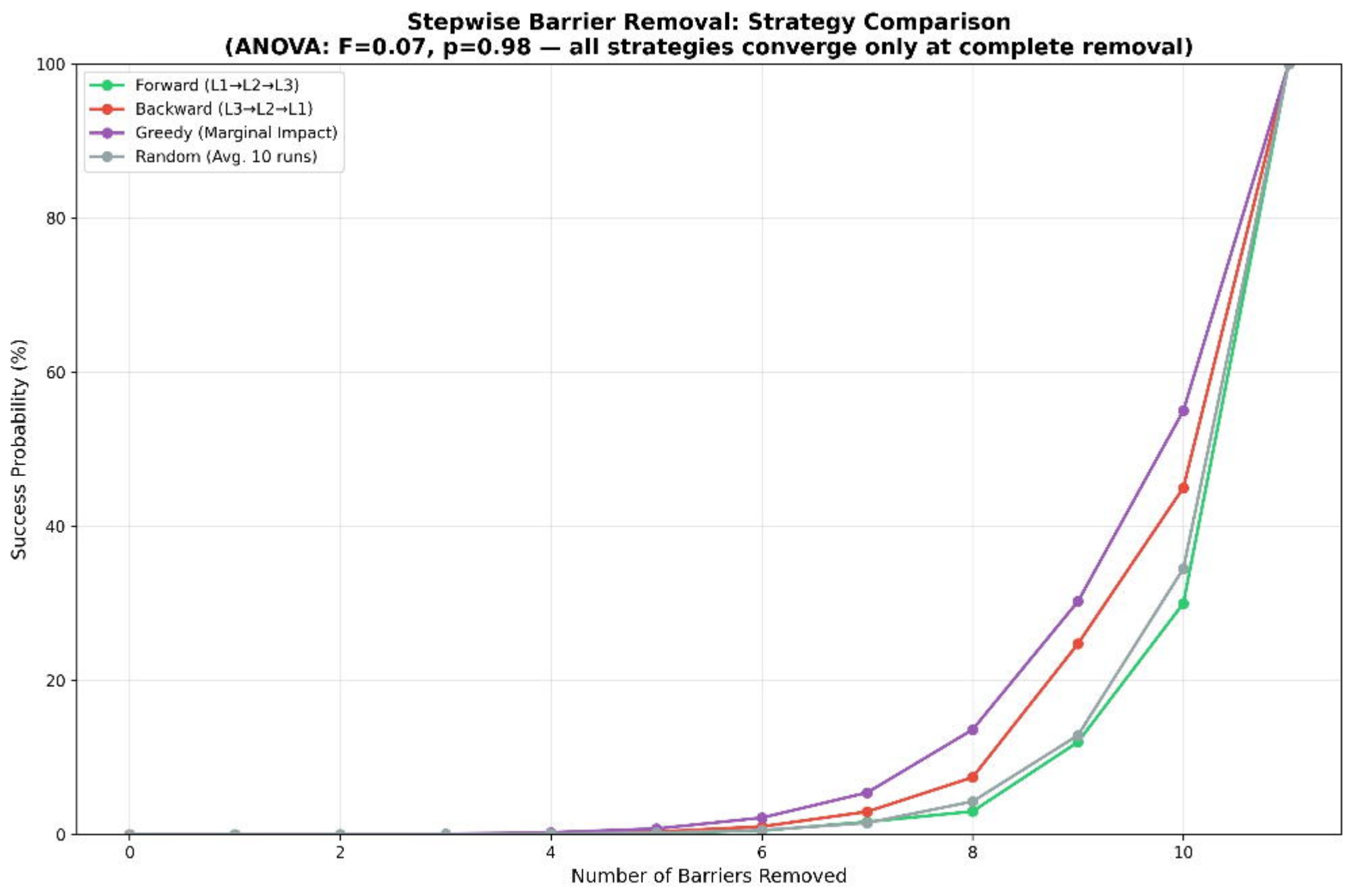
[20] Morris, M. D. Factorial sampling plans for preliminary computational experiments. *Technometrics* **33**, 161–174 (1991).

[21] Herman, J. & Usher, W. SALib: an open-source Python library for sensitivity analysis. *J. Open Source Softw.* **2**, 97 (2017).

[22] Ajunwa, I. The paradox of automation as anti-bias intervention. *Cardozo Law Rev.* **41**, 1671–1742 (2020).

[23] Raghavan, M., Barocas, S., Kleinberg, J. & Levy, K. Mitigating bias in algorithmic hiring: evaluating claims and practices. In *Proc. Conference on Fairness, Accountability, and Transparency* 469–481 (ACM, 2020).




Stepwise Barrier Removal: Strategy Comparison
(ANOVA: F=0.07, p=0.98 — all strategies converge only at complete removal)
Forward (L1→L2→L3)
Backward (L3→L2→L1)
Greedy (Marginal Impact)
Random (Avg. 10 runs)
Success Probability (%)
Number of Barriers Removed
0
20
40
60
80
100
0
2
4
6
8
10

Layer Interaction Effects Matrix
(Diagonal: Individual, Off-diagonal: Pairwise)