Thresholds and traps in AI reliance: an ODE and network model with relational capital

Weinan Wang
Department of Mathematics, University of Oklahoma, Norman, OK, USA
ww@ou.edu

Abstract

In this paper, we propose and analyze a deterministic ODE and network model for population-level AI reliance, treated as a behavioral contagion coupled to human relational capital. The three state variables are casual AI use, AI dependence, and relational capital; adoption is driven by social normalization, imitation, and the substitution of AI support for human support, and a single relational variable governs both the adoption pressure and the rate of exit from dependence. We prove positive invariance of the natural state polytope, derive an explicit invasion threshold for the AI-free human-centered equilibrium, and obtain an explicit backward-bifurcation criterion at that threshold through a center-manifold reduction. Reducing the positive equilibria to a scalar equation, we exhibit an explicit parameter regime containing a saddle and a linearly stable high-reliance trap, so that a locally stable human-centered state coexists with a stable high-reliance attractor. Placing the same local kinetics on a social network, we obtain a graph-dependent invasion threshold governed by the spectral radius of the exposure matrix, together with a degraded-state condition for global recovery. These thresholds separate four distinct effects: local invasion, backward bifurcation, graph-amplified invasion, and degraded-state recovery.

1 Introduction

Conversational AI systems increasingly provide advice, reassurance, emotional support, and social feedback. Recent evidence on sycophantic AI reports that such systems make users feel understood while making later human interaction feel more effortful and less satisfying [16]. This dynamic suggests a feedback mechanism: AI support becomes more attractive when human support is weak, and repeated reliance on AI in turn erodes the availability and perceived value of that human support.

A growing body of empirical work on AI companions and social chatbots documents supportive use alongside maladaptive patterns. A survey of active companion-chatbot users found that usage enhanced social confidence for some users while raising isolation risk for others, depending on user characteristics and problematic-use patterns [19]. A four-week randomized controlled study of extended chatbot use found that heavier voluntary use was associated with higher loneliness, emotional dependence, problematic use, and lower real-world socialization [8]. Editorial and policy discussions similarly emphasize that companion systems offer perceived support while creating risks of maladaptive attachment [20]. Taken together, these studies indicate that supportive use and self-reinforcing substitutive reliance can arise from a single feedback structure.

This feedback invites a dynamical treatment, for which behavioral epidemic models supply the template: social exposure increases entry into AI use, while the behavioral state itself reshapes the

effective reproduction landscape. A recent systematic review of behavior-change models in infectious disease identifies prevalence feedback, awareness and media dynamics, policy triggers, and social learning as the dominant post-2020 mechanisms [17]. Multilayer network models couple disease spread to competing opinions, with opinion states altering disease-transition rates [26], and adaptive disinformation-diffusion models show how history-dependent link influence produces delayed activation, nonlinear reinforcement, and tipping [21]. Closely related noncompliance models treat public-health behavior as a coupled social contagion across ODE, stochastic, reaction-diffusion, and optimal-control formulations [23, 24, 3, 22]. We follow this template: an ODE governs the local population dynamics, and a weighted graph carries social coupling between groups.

The mechanism specific to AI is an endogenous substitution. AI systems offer an always-available, personalized, and socially responsive channel for advice, reassurance, and feedback; as users substitute that channel for human contact, relational capital weakens. Weaker relational capital then raises adoption pressure and slows recovery from dependence, so AI reliance reshapes the very environment that governs its own spread. Several modeling choices encode this structure: a distinction between casual use and dependence, a weaker relational erosion from casual use than from dependence (the parameter $\theta < 1$), and a single relational variable that, when low, simultaneously amplifies adoption and suppresses exit from dependence. Threshold and bifurcation tools from behavioral epidemic modeling then make precise how this relational feedback decides whether reliance fades, persists, or becomes self-sustaining.

Three thresholds organize the analysis, and keeping them separate is the main conceptual point. The first governs local invasion: whether a small amount of reliance grows near a population with intact relational capital. The second is a backward bifurcation, under which a stable high-reliance trap appears before that local threshold is crossed, so that a locally stable human-centered state coexists with a stable trap; such a population recovers from small perturbations but tips to high reliance after a large relational shock. The third is a network effect: social coupling can destabilize the all-human state even when every isolated group is subcritical, after which capture by the trap is a nonlinear consequence of the local dynamics. In both the well-mixed and network settings, a degraded-state condition then guarantees that the human-centered state remains globally attracting once relational capital has eroded across the population.

2 Related work

AI reliance with relational feedback fits into a broad class of epidemic-behavior systems in which behavioral states alter transmission, recovery, or contact structure. Reviews and perspectives on behavior-dependent epidemic modeling emphasize that awareness, risk perception, vaccination behavior, and social response alter the effective reproduction landscape [10, 1]. The same mathematical structure applies to behavioral reliance. A closely related example is the epidemic-behavior model of Parkinson and Wang [23], where noncompliance evolves as a behavioral process coupled to disease transmission. The relational-capital variable is also grounded in the literature on social support, social integration, and health. Social relationships are associated with morbidity and mortality risk across epidemiological studies [15, 14]. The variable R abstracts the availability, trust, and perceived satisfaction of human support across possible survey instruments. Relationship-science accounts of AI chatbots emphasize a related asymmetry: chatbots are perceived as responsive and supportive, but they differ from human relationships in reciprocal influence, mutuality, and the broader social functions of close relationships [27].

The network dynamics are related to threshold and cascade models of collective behavior. Granovetter's threshold model [11] and Watts' random-network cascade model [30] show how small

shocks trigger large aggregate changes when social influence and network structure align. Disease-opinion multilayer models provide a closer epidemic-behavior analogue, with competing opinions on one layer changing disease spread on another layer [26]. Adaptive disinformation diffusion provides an information-contagion analogue in which directed link influence evolves through reinforcement and decay, so cumulative exposure produces delayed activation and tipping behavior [21]. AI-reliance adoption pressure is continuous, and the threshold emerges from the interaction between AI reliance and relational capital.

The threshold calculations also parallel standard reproduction-number theory in epidemic models. The next-generation approach of van den Driessche and Watmough [28] provides the classical local-invasion interpretation of $\mathcal{R}_0$. On networks, spectral quantities play an analogous role in epidemic and metapopulation thresholds [25, 7, 29]. In the metapopulation AI-reliance system, the graph threshold $\mathcal{R}_G$ is a behavioral analogue: the spectral radius of the group-level exposure matrix amplifies local adoption pressure.

The local bifurcation calculation uses standard center-manifold and normal-form arguments for finite-dimensional dynamical systems [5, 18]. The specific left-null-vector reduction is the same one used in backward-bifurcation analyses of epidemic models [6]. Backward bifurcation has a long history in epidemic control and vaccination models [13, 4, 12]. Behavior-disease coupling, awareness, and imitation dynamics provide further modeling precedents [2, 9]. The same tools isolate the mechanism produced by coupling AI reliance to relational capital and network exposure.

3 ODE model and analysis

Let

$$e(t) \geq 0, \qquad a(t) \geq 0, \qquad h(t) = 1 - e(t) - a(t) \geq 0$$

denote the fractions of a population in casual AI use, AI dependence, and primarily human-centered support, respectively. Let

$$R(t) \in [0, 1]$$

denote relational capital: an index of available, trusted, and satisfying human support. The ODE model is

$$\dot{e} = \Lambda(R)(e + a)h - (\kappa + \delta)e, \tag{1}$$

$$\dot{a} = \kappa e - \Gamma(R)a, \tag{2}$$

$$\dot{R} = \rho h(1 - R) - \sigma(a + \theta e)R, \tag{3}$$

where

$$\Lambda(R) = \beta_0 + \beta_1(1 - R), \qquad \Gamma(R) = \gamma_0 + \gamma_1 R. \tag{4}$$

All parameters are nonnegative, with

$$\kappa, \delta, \rho, \sigma > 0, \qquad \gamma_0 > 0, \qquad 0 \leq \theta \leq 1.$$

Remark 3.1. *The term $\Lambda(R)(e + a)h$ is a mass-action behavioral incidence term. Existing AI users increase exposure and normalization. A human-centered individual enters casual AI use at a rate that increases when relational capital is weaker.*

Remark 3.2. *The relational capital equation tracks rebuilding through human-centered interaction, $\rho h(1 - R)$, and erosion through AI reliance, $\sigma(a + \theta e)R$. The parameter θ allows casual use to be less relationally erosive than dependence.*

Parameter	Interpretation
β_0	baseline social adoption pressure from existing AI users
β_1	extra adoption pressure when relational capital is low
κ	transition rate from casual AI use to dependence
δ	exit rate from casual AI use back to human-centered support
γ_0	spontaneous exit rate from AI dependence
γ_1	relational-capital-assisted exit rate from AI dependence
ρ	rebuilding rate of relational capital through human-centered interaction
σ	erosion rate of relational capital through AI reliance
θ	relative erosion caused by casual AI use compared with dependence

Table 1: Parameter interpretation.

Remark 3.3. *The same variable R controls both sides of the feedback. When R is low, the adoption rate $\Lambda(R)$ is higher and the dependence-exit rate $\Gamma(R)$ is lower. This couples AI's availability as a substitute support channel to the erosion of the human support network that helps users leave dependence.*

Quantity	Meaning
$\mathcal{R}_H = \mathcal{R}_{\mathrm{AI}}(1)$	local invasion number at the human-centered equilibrium, where relational capital is maximal
$\mathcal{R}_{\mathrm{AI}}(R)$	frozen-R invasion number obtained by holding relational capital fixed at R in the (e, a) subsystem
$\mathcal{R}_{\mathrm{AI}}(0)$	low-relational-capital invasion number for trap removal and global recovery
$\mathcal{R}_G$	graph invasion number at the all-human network equilibrium, with spectral amplification through $\varrho(A)$
$\mathcal{R}_G^0$	low-relational-capital graph invasion number for global network recovery

Table 2: Threshold quantities used in the ODE and network analysis.

3.1 Comparison principles

Two elementary comparison facts are used repeatedly. Vector inequalities are understood componentwise.

Lemma 3.4 (Comparison with a Metzler linear system)**.** *Let $M \in \mathbb{R}^{m\times m}$ be Metzler, meaning that all off-diagonal entries of M are nonnegative. Suppose $x : [0, T] \to \mathbb{R}^m$ is continuously differentiable and satisfies*

$$\dot{x}(t) \le Mx(t) \qquad \text{for } 0 \le t \le T.$$

If y solves $\dot{y} = My$ and $x(0) \le y(0)$, then

$$x(t) \le y(t) \qquad \text{for } 0 \le t \le T.$$

In particular, if M is Hurwitz and $x(0) \le y(0)$, then $x(t)$ is bounded above by an exponentially decaying function whenever $y(0) \ge 0$.

Proof. For a Metzler matrix M, the semigroup e^{Mt} is nonnegative for every $t \geq 0$. Indeed, choose $\alpha > 0$ so that $M + \alpha I$ is nonnegative. Then

$$e^{Mt} = e^{-\alpha t} e^{(M+\alpha I)t},$$

and the exponential of a nonnegative matrix is nonnegative term by term.

Let $r(t) = Mx(t) - \dot{x}(t) \geq 0$. Variation of constants gives

$$x(t) = e^{Mt}x(0) - \int_0^t e^{M(t-s)} r(s)\, ds.$$

Since e^{Mt} is nonnegative and $r(s) \geq 0$, it follows that

$$x(t) \leq e^{Mt}x(0) \leq e^{Mt}y(0) = y(t).$$

If M is Hurwitz, there exist constants $C, \alpha > 0$ such that $\|e^{Mt}\| \leq Ce^{-\alpha t}$. When $y(0) \geq 0$, the comparison solution $y(t) = e^{Mt}y(0)$ is componentwise nonnegative and satisfies

$$0 \leq y_j(t) \leq Ce^{-\alpha t}\|y(0)\|, \qquad j = 1, \ldots, m,$$

after increasing C if necessary. This gives the final claim. □

Lemma 3.5 (A Perron-Frobenius block threshold)**.** *Let $C \in \mathbb{R}^{n\times n}$ be nonnegative, let $K > 0$, $\Gamma > 0$, and $\kappa > 0$, and define*

$$M(C, \Gamma) = \begin{pmatrix} C - KI & C \\ \kappa I & -\Gamma I \end{pmatrix}.$$

Let $\mu_* = \varrho(C)$ be the spectral radius of C. Then $M(C, \Gamma)$ is Hurwitz if

$$\frac{\mu_*(\Gamma + \kappa)}{K\Gamma} < 1,$$

and has a positive real eigenvalue if

$$\frac{\mu_*(\Gamma + \kappa)}{K\Gamma} > 1.$$

At equality, zero is an eigenvalue of $M(C, \Gamma)$.

Proof. The matrix $M(C, \Gamma)$ is Metzler. Choose $q > 0$ large enough that $N := M(C, \Gamma) + qI$ is nonnegative. By the Perron-Frobenius theorem, $\varrho(N)$ is an eigenvalue of N with a nonnegative eigenvector. If ζ is any other eigenvalue of N, then $\operatorname{Re}\zeta \leq |\zeta| \leq \varrho(N)$. Hence the spectral bound of $M(C, \Gamma)$ is

$$s(M(C, \Gamma)) = \varrho(N) - q,$$

and is a real eigenvalue with a nonnegative eigenvector. Assume first that

$$\mu_* < \frac{K\Gamma}{\Gamma + \kappa}.$$

Suppose, toward a contradiction, that the spectral bound s of $M(C, \Gamma)$ is nonnegative. Let $(e, a) \geq 0$ be a corresponding nonzero eigenvector. From the second block equation,

$$sa = \kappa e - \Gamma a,$$

so

$$a = \frac{\kappa}{s+\Gamma} e.$$

The vector e is nonzero; otherwise $a = 0$, contradicting nontriviality of the eigenvector. Substituting this relation into the first block equation gives

$$Ce = \frac{(s+K)(s+\Gamma)}{s+\Gamma+\kappa}\, e.$$

Thus

$$\mu(s) := \frac{(s+K)(s+\Gamma)}{s+\Gamma+\kappa}$$

is an eigenvalue of C associated with a nonnegative nonzero eigenvector. Hence $\mu(s) \le \varrho(C) = \mu_*$. On the other hand, for $s \ge 0$,

$$\mu(s) \ge \mu(0) = \frac{K\Gamma}{\Gamma+\kappa}.$$

To see this directly, differentiate:

$$\mu'(s) = \frac{(s+\Gamma)^2 + \kappa(2s+K+\Gamma)}{(s+\Gamma+\kappa)^2} > 0.$$

Therefore $\mu_* \ge \mu(s) \ge K\Gamma/(\Gamma+\kappa)$, contradicting the assumed strict inequality. Hence the spectral bound is negative, and $M(C,\Gamma)$ is Hurwitz. Now assume

$$\mu_* > \frac{K\Gamma}{\Gamma+\kappa}.$$

Let $v \ge 0$, $v \ne 0$, be a Perron-Frobenius eigenvector of C, so that $Cv = \mu_* v$. Consider

$$p(\lambda) = (\lambda+K)(\lambda+\Gamma) - \mu_*(\lambda+\Gamma+\kappa).$$

Then

$$p(0) = K\Gamma - \mu_*(\Gamma+\kappa) < 0,$$

while $p(\lambda) \to +\infty$ as $\lambda \to +\infty$. Thus p has a positive root $\lambda > 0$. Setting

$$e = v, \qquad a = \frac{\kappa}{\lambda+\Gamma} v$$

gives

$$M(C,\Gamma)\begin{pmatrix} e \\ a \end{pmatrix} = \lambda \begin{pmatrix} e \\ a \end{pmatrix}.$$

Thus $M(C,\Gamma)$ has a positive eigenvalue. At equality, the same construction with $\lambda = 0$ gives a zero eigenvalue. □

3.2 Invariant region and global well-posedness

Define the closed state space

$$\Delta_R = \{(e,a,R) : e \ge 0,\ a \ge 0,\ e+a \le 1,\ 0 \le R \le 1\}.$$

Proposition 3.6 (Positive invariance)**.** *For any initial data $(e(0),a(0),R(0)) \in \Delta_R$, the solution of (1)–(3) remains in Δ_R for all forward time for which it exists.*

Proof. The right-hand side of (1)–(3) is polynomial in (e, a, R) after substituting $h = 1 - e - a$, and is therefore locally Lipschitz on $\mathbb{R}^3$. Hence a unique local solution exists. It remains to prove that no solution can leave the closed polytope Δ_R. Each boundary face has an inward or tangent vector field. On $\{e = 0\}$,

$$\dot{e} = \Lambda(R)ah \geq 0.$$

On $\{a = 0\}$,

$$\dot{a} = \kappa e \geq 0.$$

Let $h = 1 - e - a$. Adding (1) and (2) gives

$$\dot{e} + \dot{a} = \Lambda(R)(e + a)h - \delta e - \Gamma(R)a.$$

Therefore, on the face $\{h = 0\}$,

$$\dot{h} = -(\dot{e} + \dot{a}) = \delta e + \Gamma(R)a \geq 0.$$

Finally, on $\{R = 0\}$,

$$\dot{R} = \rho h \geq 0,$$

and on $\{R = 1\}$,

$$\dot{R} = -\sigma(a + \theta e) \leq 0.$$

Thus the vector field belongs to the tangent cone of Δ_R at every boundary point, which is the subtangentiality condition in Nagumo's invariance theorem for a closed convex set.

Equivalently, suppose for contradiction that the solution first leaves Δ_R at time t_*. Immediately before t_* all four defining inequalities hold. At t_* at least one boundary constraint is active, and for the solution to leave, the derivative of that active constraint must point outside the feasible set. The face computations show the opposite on every possible active face, including intersections of faces. This contradiction proves positive invariance. □

Corollary 3.7 (Global solutions)**.** *Every solution with initial data in Δ_R exists uniquely for all $t \geq 0$.*

Proof. Let $[0, T_{\max})$ be the maximal interval of existence. By positive invariance, the solution remains in the compact set Δ_R for all $t < T_{\max}$. The vector field is smooth on an open neighborhood of Δ_R, hence bounded and locally Lipschitz there. The standard continuation theorem for ODEs then implies that a finite maximal time could occur only if the solution left every compact subset of the domain. This is impossible because the trajectory stays in Δ_R. Therefore $T_{\max} = \infty$. □

3.3 The human-centered equilibrium

The AI-free, human-centered equilibrium is

$$E_H = (e, a, R) = (0, 0, 1).$$

Set

$$K = \kappa + \delta, \qquad \Gamma_1 = \Gamma(1) = \gamma_0 + \gamma_1.$$

Linearizing the (e, a) subsystem at E_H gives

$$\frac{d}{dt}\begin{pmatrix} e \\ a \end{pmatrix} = \begin{pmatrix} \beta_0 - K & \beta_0 \\ \kappa & -\Gamma_1 \end{pmatrix}\begin{pmatrix} e \\ a \end{pmatrix}.$$

This leads to the basic AI-reliance number

$$\mathcal{R}_H = \beta_0 \left(\frac{1}{\kappa + \delta} + \frac{\kappa}{(\kappa + \delta)(\gamma_0 + \gamma_1)} \right) = \frac{\beta_0(\gamma_0 + \gamma_1 + \kappa)}{(\kappa + \delta)(\gamma_0 + \gamma_1)}. \tag{5}$$

Proposition 3.8 (Local threshold)**.** *The human-centered equilibrium E_H is locally asymptotically stable if $\mathcal{R}_H < 1$ and unstable if $\mathcal{R}_H > 1$.*

Proof. The full Jacobian at E_H is

$$J(E_H) = \begin{pmatrix} \beta_0 - K & \beta_0 & 0 \\ \kappa & -\Gamma_1 & 0 \\ -\sigma\theta & -\sigma & -\rho \end{pmatrix}.$$

This matrix is block lower triangular with respect to the splitting (e, a) and R. Hence one eigenvalue is $-\rho < 0$, and the remaining two eigenvalues are those of

$$B = \begin{pmatrix} \beta_0 - K & \beta_0 \\ \kappa & -\Gamma_1 \end{pmatrix}.$$

Its determinant is

$$\det B = \Gamma_1 K - \beta_0(\Gamma_1 + \kappa) = \Gamma_1 K(1 - \mathcal{R}_H).$$

Also

$$\operatorname{tr} B = \beta_0 - K - \Gamma_1.$$

If $\mathcal{R}_H < 1$, then

$$\beta_0(\Gamma_1 + \kappa) < K\Gamma_1,$$

so in particular $\beta_0 < K\Gamma_1/(\Gamma_1 + \kappa) < K$. Therefore $\operatorname{tr} B < -\Gamma_1 < 0$, while $\det B > 0$. For a 2×2 real matrix, negative trace and positive determinant are exactly the Routh-Hurwitz conditions for both eigenvalues to have negative real part. Together with the eigenvalue $-\rho$, this proves local asymptotic stability of E_H.

If $\mathcal{R}_H > 1$, then $\det B < 0$. The two eigenvalues of B have product $\det B < 0$, so one is positive and one is negative. Hence the full Jacobian has a positive eigenvalue and E_H is unstable. At the threshold $\mathcal{R}_H = 1$, $\det B = 0$, so the equilibrium is nonhyperbolic. □

Remark 3.9. *Formula (5) has the usual next-generation interpretation. A casual user produces new casual users for an average time $1/(\kappa + \delta)$. With probability $\kappa/(\kappa + \delta)$ the casual user becomes dependent, and then produces new casual users for an average time $1/(\gamma_0 + \gamma_1)$. Multiplying the total expected adoption-producing time by the baseline adoption rate β_0 gives $\mathcal{R}_H$.*

3.4 Backward bifurcation at the human-centered threshold

At $\mathcal{R}_H = 1$, equivalently $\beta_0 = \beta_*$, the human-centered equilibrium has a simple zero eigenvalue. When the quadratic coefficient in the center-manifold reduction is nonzero, the local branch has the standard transcritical form. The sign of this coefficient determines whether the positive branch lies on the locally stable or unstable side of E_H.

Theorem 3.10 (Backward-bifurcation criterion)**.** *Fix all parameters except β_0, and set*

$$K = \kappa + \delta, \qquad \Gamma_1 = \gamma_0 + \gamma_1, \qquad n = \Gamma_1 + \kappa, \qquad \beta_* = \frac{K\Gamma_1}{n}.$$

Thus $\mathcal{R}_H = 1$ exactly when $\beta_0 = \beta_*$. Define

$$\mathcal{B} = \frac{\sigma}{\rho}(\theta\Gamma_1 + \kappa)\left[\beta_1 n^2 + \gamma_1 \kappa K\right] - K\Gamma_1 n^2. \tag{6}$$

The case $\mathcal{B} \neq 0$ is the nondegenerate quadratic case. If $\mathcal{B} > 0$, then the bifurcation at $\beta_0 = \beta_*$ is backward: for all sufficiently small $\beta_* - \beta_0 > 0$, the ODE has a positive equilibrium near E_H even though E_H is locally asymptotically stable. If $\mathcal{B} < 0$, the local positive branch is forward and lies on the $\beta_0 > \beta_*$ side. At $\mathcal{B} = 0$, the quadratic center-manifold coefficient vanishes.

Proof. Use the variables

$$s = 1 - R, \qquad \mu = \beta_0 - \beta_*,$$

so that E_H becomes $(e, a, s) = (0, 0, 0)$. At $\mu = 0$,

$$\Lambda = \beta_* + \beta_1 s, \qquad \Gamma = \Gamma_1 - \gamma_1 s,$$

and the system is

$$\begin{aligned}
\dot{e} &= (\beta_* + \mu + \beta_1 s)(e + a)(1 - e - a) - Ke, \\
\dot{a} &= \kappa e - (\Gamma_1 - \gamma_1 s)a, \\
\dot{s} &= -\rho(1 - e - a)s + \sigma(a + \theta e)(1 - s).
\end{aligned}$$

The Jacobian at the origin and at $\mu = 0$ is

$$J_* = \begin{pmatrix} \beta_* - K & \beta_* & 0 \\ \kappa & -\Gamma_1 & 0 \\ \sigma\theta & \sigma & -\rho \end{pmatrix} = \begin{pmatrix} -K\kappa/n & K\Gamma_1/n & 0 \\ \kappa & -\Gamma_1 & 0 \\ \sigma\theta & \sigma & -\rho \end{pmatrix}.$$

The zero eigenvalue has right eigenvector

$$w = \begin{pmatrix} \Gamma_1 \\ \kappa \\ M \end{pmatrix}, \qquad M = \frac{\sigma}{\rho}(\theta\Gamma_1 + \kappa),$$

because the third row gives

$$\sigma\theta\Gamma_1 + \sigma\kappa - \rho M = 0.$$

A left eigenvector for the zero eigenvalue is

$$v = \begin{pmatrix} 1 \\ K/n \\ 0 \end{pmatrix},$$

and

$$D := v \cdot w = \Gamma_1 + \frac{K\kappa}{n} > 0.$$

The remaining eigenvalues have negative real part: the (e, a) block has one zero eigenvalue and trace $\beta_* - K - \Gamma_1 < 0$, while the relational-capital direction contributes the eigenvalue $-\rho$ after triangularization. Thus the center manifold at the threshold is one-dimensional.

The standard center-manifold projection formula [6, 18], with the left eigenvector normalized by D, has the form

$$\dot{z} = Az^2 + c\mu z + O(|z|^3 + |\mu|z^2 + \mu^2|z|), \tag{7}$$

where

$$c = \frac{1}{D}\, v \cdot \frac{\partial^2 f}{\partial x\, \partial \mu}(0, 0)w = \frac{\Gamma_1 + \kappa}{D} = \frac{n}{D} > 0.$$

It remains to compute the quadratic coefficient A. The quadratic terms in the first two equations at $\mu = 0$ are

$$-\beta_*(e+a)^2 + \beta_1 s(e+a), \qquad \gamma_1 sa.$$

Although the left null vector has third component zero, the relational component enters the coefficient through $w_3 = M$ in the products $s(e+a)$ and sa. The quadratic correction to the center manifold does is annihilated in the projection for A, because it is multiplied by J_* and $v^\top J_* = 0$. The Taylor expansion contains the factor $\frac{1}{2}D_x^2 f(0,0)[x,x]$, so

$$\begin{aligned} DA &= \frac{1}{2}\, v \cdot D_x^2 f(0,0)[w,w] \\ &= \left[-\beta_* n^2 + \beta_1 Mn\right] + \frac{K}{n}\gamma_1 M\kappa \\ &= \frac{1}{n}\left\{M\left[\beta_1 n^2 + \gamma_1 \kappa K\right] - K\Gamma_1 n^2\right\}. \end{aligned}$$

Since $D > 0$ and $n > 0$, the sign of A is the sign of $\mathcal{B}$. When $\mathcal{B} > 0$, one has $A > 0$. The nonzero equilibrium of (7) satisfies

$$z(\mu) = -\frac{c}{A}\mu + O(\mu^2).$$

For $\mu < 0$ sufficiently small, $z(\mu) > 0$. In the original variables,

$$e = \Gamma_1 z + O(z^2), \qquad a = \kappa z + O(z^2), \qquad s = Mz + O(z^2),$$

so this branch corresponds to a positive equilibrium with $0 < R = 1 - s < 1$. For these same $\mu < 0$, the local-threshold proposition gives local asymptotic stability of E_H. This proves the backward-bifurcation claim. If $\mathcal{B} < 0$, then $A < 0$, and the same expansion gives the positive branch for $\mu > 0$, which is the forward case. □

Remark 3.11 (Stability of the local positive branch)**.** *The backward branch in Theorem 3.10 is the local unstable threshold branch near the human-centered equilibrium. In the center-manifold normal form*

$$\dot{z} = Az^2 + c\mu z + \textit{higher-order terms}, \qquad c > 0,$$

the nonzero branch satisfies

$$z(\mu) = -\frac{c}{A}\mu + O(\mu^2).$$

Along this branch,

$$\frac{\partial}{\partial z}\left(Az^2 + c\mu z\right) = -c\mu + O(\mu^2).$$

Hence, when $\mathcal{B} > 0$, the positive branch on the $\mu < 0$ side is unstable in the center direction. In the bistable regime of Figure 2, this branch is the local basin-threshold branch; the stable trap branch appears after the saddle-node fold.

Remark 3.12. *In the symmetric timescale case $\sigma = \rho$, the large-κ limit of (6) is the simple condition*

$$\beta_1 > \gamma_0.$$

Thus strong low-relational-capital adoption pressure relative to the spontaneous recovery floor drives the branch into the locally stable side of the human-centered threshold.

3.5 Endogenous relational feedback and trap states

The key difference from a one-threshold adoption model is that a frozen-R invasion number depends on relational capital. Holding R fixed in the (e, a) subsystem gives

$$\mathcal{R}_{\mathrm{AI}}(R) = \Lambda(R)\left(\frac{1}{\kappa+\delta} + \frac{\kappa}{(\kappa+\delta)\Gamma(R)}\right) = \frac{\Lambda(R)(\Gamma(R)+\kappa)}{(\kappa+\delta)\Gamma(R)}. \tag{8}$$

Proposition 3.13 (Relational threshold)**.** *For fixed relational capital $\bar{R} \in [0,1]$, freeze $R = \bar{R}$ in the (e, a) equations and linearize at $(e, a) = (0, 0)$. The resulting AI-free state is locally asymptotically stable if $\mathcal{R}_{\mathrm{AI}}(\bar{R}) < 1$ and unstable if $\mathcal{R}_{\mathrm{AI}}(\bar{R}) > 1$. Moreover, $\mathcal{R}_{\mathrm{AI}}(R)$ is nonincreasing in R, with*

$$\frac{d}{dR}\mathcal{R}_{\mathrm{AI}}(R) = -\frac{\beta_1(\Gamma(R)+\kappa)}{(\kappa+\delta)\Gamma(R)} - \frac{\Lambda(R)\kappa\gamma_1}{(\kappa+\delta)\Gamma(R)^2} \le 0. \tag{9}$$

If $\mathcal{R}_{\mathrm{AI}}(1) < 1 < \mathcal{R}_{\mathrm{AI}}(0)$ and at least one of β_1, γ_1 is positive, then there is a unique $R_c \in (0,1)$ such that

$$\mathcal{R}_{\mathrm{AI}}(R_c) = 1.$$

For frozen relational capital, small AI reliance decays when $R > R_c$ and grows when $R < R_c$.

Proof. For fixed $R = \bar{R}$, the linearized (e, a) system at the AI-free state is

$$\frac{d}{dt}\begin{pmatrix} e \\ a \end{pmatrix} = \begin{pmatrix} \Lambda(\bar{R}) - (\kappa+\delta) & \Lambda(\bar{R}) \\ \kappa & -\Gamma(\bar{R}) \end{pmatrix}\begin{pmatrix} e \\ a \end{pmatrix}.$$

The determinant of this matrix is

$$(\kappa+\delta)\Gamma(\bar{R}) - \Lambda(\bar{R})(\Gamma(\bar{R})+\kappa) = (\kappa+\delta)\Gamma(\bar{R})(1 - \mathcal{R}_{\mathrm{AI}}(\bar{R})).$$

The trace is

$$\Lambda(\bar{R}) - (\kappa+\delta) - \Gamma(\bar{R}).$$

If $\mathcal{R}_{\mathrm{AI}}(\bar{R}) < 1$, then

$$\Lambda(\bar{R})(\Gamma(\bar{R})+\kappa) < (\kappa+\delta)\Gamma(\bar{R}),$$

and hence $\Lambda(\bar{R}) < \kappa + \delta$. The trace is therefore negative, and the determinant is positive, so the frozen AI-free state is locally asymptotically stable. If $\mathcal{R}_{\mathrm{AI}}(\bar{R}) > 1$, then the determinant is negative, so the linearized system has one positive eigenvalue and the frozen AI-free state is unstable. It remains to justify the monotonicity statement. Since $\Lambda'(R) = -\beta_1$ and $\Gamma'(R) = \gamma_1$,

$$\frac{d}{dR}\left[\frac{\Lambda(R)(\Gamma(R)+\kappa)}{(\kappa+\delta)\Gamma(R)}\right] = \frac{1}{\kappa+\delta}\left[-\beta_1\frac{\Gamma(R)+\kappa}{\Gamma(R)} - \Lambda(R)\frac{\kappa\gamma_1}{\Gamma(R)^2}\right],$$

which is (9). The derivative is nonpositive on $[0,1]$. Under the crossing hypothesis $\mathcal{R}_{\mathrm{AI}}(1) < 1 < \mathcal{R}_{\mathrm{AI}}(0)$, strict monotonicity follows as soon as at least one of β_1, γ_1 is positive. Indeed, if $\beta_1 > 0$, then the first term in (9) is strictly negative. If $\beta_1 = 0$ and $\gamma_1 > 0$, then $\mathcal{R}_{\mathrm{AI}}(0) > 1$ forces $\Lambda(R) = \beta_0 > 0$, so the second term is strictly negative. Continuity and strict monotonicity then give a unique crossing $R_c \in (0,1)$. □

The crossing condition for trap formation is

$$\mathcal{R}_{\mathrm{AI}}(1) < 1 < \mathcal{R}_{\mathrm{AI}}(0).$$

It makes the human-centered state locally stable while low relational capital makes AI reliance self-amplifying. Stable trap formation is then determined by the positive equilibria of the scalar equation and their Jacobian stability.

3.6 Positive equilibria

Positive equilibria reduce to a scalar equation. On the set where $\Lambda(R) > 0$, define

$$H(R) = \frac{(\kappa+\delta)\Gamma(R)}{\Lambda(R)(\kappa+\Gamma(R))} = \frac{1}{\mathcal{R}_{\mathrm{AI}}(R)}.$$

If $\Lambda(R) = 0$, then $\dot{e} = 0$ implies $(\kappa+\delta)e = 0$, excluding positive equilibria at relational-capital value R.

A *trap equilibrium* is a linearly stable positive equilibrium. The following scalar equation characterizes all positive equilibria, including both stable trap equilibria and unstable threshold equilibria.

Proposition 3.14 (Scalar characterization of positive equilibria)**.** *There is a one-to-one correspondence between positive equilibria (e, a, R) with $e > 0$, $a > 0$, and $h = 1 - e - a > 0$ and roots $R \in (0, 1)$ with $\Lambda(R) > 0$ of*

$$R = \frac{\rho H(R)}{\rho H(R) + \sigma(1 - H(R))\dfrac{\kappa + \theta\Gamma(R)}{\kappa + \Gamma(R)}}, \tag{10}$$

with $0 < H(R) < 1$. Given such a root, the equilibrium is

$$h = H(R), \tag{11}$$

$$e = \frac{\Gamma(R)}{\kappa + \Gamma(R)}(1 - H(R)), \tag{12}$$

$$a = \frac{\kappa}{\kappa + \Gamma(R)}(1 - H(R)). \tag{13}$$

Proof. At a positive equilibrium, equation (2) gives

$$\kappa e = \Gamma(R)a.$$

Thus

$$e = \frac{\Gamma(R)}{\kappa}a, \qquad e + a = \frac{\kappa + \Gamma(R)}{\kappa}a.$$

Using (1) and dividing by $e + a > 0$ gives

$$\Lambda(R)h = \frac{(\kappa+\delta)e}{e+a} = \frac{(\kappa+\delta)\Gamma(R)}{\kappa+\Gamma(R)}.$$

The right-hand side is strictly positive, so $\Lambda(R) > 0$. Therefore $h = H(R)$. Since $e + a = 1 - H(R)$, equations (12)–(13) follow. Finally, imposing $\dot{R} = 0$ yields

$$R = \frac{\rho h}{\rho h + \sigma(a + \theta e)},$$

and substituting (11)–(13) gives (10).

The positivity inequalities also imply $0 < H(R) < 1$. Indeed $h = H(R) > 0$, while $e + a = 1 - h > 0$, so $H(R) < 1$. Moreover, at equilibrium,

$$R = \frac{\rho h}{\rho h + \sigma(a + \theta e)}.$$

Since $\rho h > 0$ and $a + \theta e > 0$ because $a > 0$, the denominator is strictly larger than the numerator. Hence $R \in (0, 1)$.

Conversely, let $R \in (0, 1)$ satisfy (10) with $0 < H(R) < 1$, and define h, e, a by (11)–(13). Since $\Gamma(R) > 0$, $\kappa > 0$, and $0 < H(R) < 1$, the constructed point satisfies $e > 0$, $a > 0$, and $h > 0$. Also,

$$e + a = \frac{\Gamma(R) + \kappa}{\kappa + \Gamma(R)}(1 - H(R)) = 1 - H(R),$$

so $h + e + a = 1$. Next,

$$\kappa e = \frac{\kappa\Gamma(R)}{\kappa + \Gamma(R)}(1 - H(R)) = \Gamma(R)a,$$

and therefore $\dot{a} = 0$. For the e equation, use

$$\frac{e}{e + a} = \frac{\Gamma(R)}{\kappa + \Gamma(R)}$$

and the definition of H:

$$\Lambda(R)h = \Lambda(R)H(R) = \frac{(\kappa + \delta)\Gamma(R)}{\kappa + \Gamma(R)} = (\kappa + \delta)\frac{e}{e + a}.$$

Multiplying by $e + a$ gives $\Lambda(R)(e + a)h = (\kappa + \delta)e$, hence $\dot{e} = 0$. Finally, equation (10) is precisely the condition

$$R = \frac{\rho h}{\rho h + \sigma(a + \theta e)},$$

which is equivalent to $\dot{R} = 0$. Thus the constructed point is a positive equilibrium. The formulas determine (h, e, a) uniquely from R, so the correspondence is one-to-one. □

Remark 3.15. *The condition $0 < H(R) < 1$ is equivalent to $\mathcal{R}_{\mathrm{AI}}(R) > 1$. Thus positive equilibria occur at relational-capital levels where AI reliance is locally self-amplifying.*

Remark 3.16. *Proposition 3.14 gives an exact scalar reduction. Root counts and stability types depend on the parameter regime and are obtained by solving the scalar equation and classifying the corresponding full Jacobian. For the parameter set in Table 3, this procedure gives exactly two positive equilibria, one saddle and one linearly stable trap.*

Proposition 3.17 (Exhaustion of ODE equilibria)**.** *Every equilibrium of (1)–(3) in Δ_R is either the human-centered equilibrium*

$$E_H = (0, 0, 1),$$

or a positive equilibrium characterized by Proposition 3.14. In particular, there are no nontrivial equilibria on the boundary faces $e = 0$, $a = 0$, $h = 0$, $R = 0$, or $R = 1$.

Proof. Let $(e, a, R) \in \Delta_R$ be an equilibrium. If $a = 0$, then $\dot{a} = \kappa e = 0$, so $e = 0$. With $e = a = 0$, the R equation becomes

$$\dot{R} = \rho(1 - R),$$

and hence $R = 1$. Thus $a = 0$ gives only E_H. If $e = 0$, then $\dot{a} = -\Gamma(R)a = 0$. Since $\Gamma(R) > 0$ on $[0, 1]$, this implies $a = 0$, and the previous paragraph again gives E_H. It remains to rule out boundary equilibria with $e > 0$ and $a > 0$. On the face $h = 0$, equivalently $e + a = 1$,

$$\dot{h} = -(\dot{e} + \dot{a}) = \delta e + \Gamma(R)a > 0,$$

so no equilibrium lies on $h = 0$ with $e > 0$ and $a > 0$. If $R = 0$ and $h > 0$, then

$$\dot{R} = \rho h > 0,$$

so no equilibrium lies on $R = 0$ with $h > 0$. If $R = 1$ and $e > 0$ or $a > 0$, then

$$\dot{R} = -\sigma(a + \theta e) < 0$$

provided $a + \theta e > 0$. If $\theta = 0$ and $a = 0$, then the first paragraph applies; otherwise the inequality is strict. Hence no nontrivial boundary equilibrium lies on $R = 1$. Therefore any equilibrium different from E_H must satisfy

$$e > 0, \qquad a > 0, \qquad h > 0, \qquad 0 < R < 1.$$

Such equilibria are precisely the positive equilibria characterized by the scalar equation (10). □

For numerical continuation and stability classification, write the scalar equation as

$$F(R) := R - \Psi(R) = 0, \tag{14}$$

where

$$\Psi(R) = \frac{\rho H(R)}{\rho H(R) + \sigma(1 - H(R))\dfrac{\kappa + \theta\Gamma(R)}{\kappa + \Gamma(R)}}.$$

Simple roots of F correspond to isolated positive equilibria. Along a one-parameter family, a saddle-node fold satisfies $F(R) = F'(R) = 0$, together with the usual nondegeneracy conditions such as a nonzero second derivative in the fold direction. The parameter family in Figure 2 has a fold satisfying these nondegeneracy conditions.

Proposition 3.18 (Jacobian at an equilibrium)**.** *At any point $(e, a, R) \in \Delta_R$, the Jacobian matrix of (1)–(3) is*

$$J(e, a, R) = \begin{pmatrix} \Lambda(R)(h - e - a) - K & \Lambda(R)(h - e - a) & -\beta_1(e + a)h \\ \kappa & -\Gamma(R) & -\gamma_1 a \\ -\rho(1 - R) - \sigma\theta R & -\rho(1 - R) - \sigma R & -\rho h - \sigma(a + \theta e) \end{pmatrix}, \tag{15}$$

where $h = 1 - e - a$ and $K = \kappa + \delta$. In particular, a positive equilibrium obtained from a root of (14) is linearly asymptotically stable if the spectral abscissa of (15) is negative, and unstable if the spectral abscissa is positive.

Proof. The formula follows by differentiating the three components of the vector field. For the first component, write $x = e + a$ and $h = 1 - e - a$. Then

$$\frac{\partial}{\partial e}\big[\Lambda(R)xh - Ke\big] = \Lambda(R)(h - x) - K, \qquad \frac{\partial}{\partial a}\big[\Lambda(R)xh - Ke\big] = \Lambda(R)(h - x),$$

and

$$\frac{\partial}{\partial R}\big[\Lambda(R)xh - Ke\big] = \Lambda'(R)xh = -\beta_1(e + a)h.$$

The remaining entries are obtained directly from $\kappa e - \Gamma(R)a$ and $\rho h(1 - R) - \sigma(a + \theta e)R$. The stability statement is the standard linearization criterion for hyperbolic equilibria. □

Corollary 3.19 (Trap-removal condition)**.** *Assume $\gamma_0 > 0$. If*

$$\mathcal{R}_{\mathrm{AI}}(0) \le 1,$$

then the ODE model has no positive equilibrium with $e > 0$, $a > 0$, and $h > 0$. In terms of the low-relational-capital amplification parameter, this condition is

$$\beta_1 \le \beta_{1,c}^{\mathrm{trap}} := \frac{(\kappa+\delta)\gamma_0}{\gamma_0+\kappa} - \beta_0. \tag{16}$$

Indeed, $\mathcal{R}_{\mathrm{AI}}(0) \le 1$ is exactly the inequality $\beta_1 \le \beta_{1,c}^{\mathrm{trap}}$. If the right-hand side of (16) is negative, then baseline adoption pressure β_0 already places the fully degraded relational state above the trap-removal bound.

Proof. By the scalar characterization, every positive equilibrium has $0 < H(R) < 1$ at its relational-capital value. Since $H(R) = 1/\mathcal{R}_{\mathrm{AI}}(R)$, this is equivalent to $\mathcal{R}_{\mathrm{AI}}(R) > 1$. The function $\mathcal{R}_{\mathrm{AI}}$ is nonincreasing in R, so its maximum on $[0,1]$ is attained at $R = 0$. Thus $\mathcal{R}_{\mathrm{AI}}(0) \le 1$ excludes positive equilibria. Finally,

$$\mathcal{R}_{\mathrm{AI}}(0) = \frac{(\beta_0+\beta_1)(\gamma_0+\kappa)}{(\kappa+\delta)\gamma_0},$$

so $\mathcal{R}_{\mathrm{AI}}(0) \le 1$ is equivalent to (16). □

Theorem 3.20 (Global stability under low-relational-capital subcriticality)**.** *Assume $\gamma_0 > 0$ and*

$$\mathcal{R}_{\mathrm{AI}}(0) < 1.$$

Then the human-centered equilibrium $E_H = (0,0,1)$ is globally asymptotically stable in Δ_R. In particular, for every initial condition in Δ_R,

$$e(t) \to 0, \qquad a(t) \to 0, \qquad R(t) \to 1 \qquad \text{as } t \to \infty.$$

Proof. Set

$$\Lambda_0 = \Lambda(0) = \beta_0 + \beta_1, \qquad \Gamma_0 = \Gamma(0) = \gamma_0, \qquad K = \kappa + \delta.$$

Since $0 \le h \le 1$, $\Lambda(R) \le \Lambda_0$, and $\Gamma(R) \ge \Gamma_0$, the (e,a) variables satisfy the componentwise differential inequality

$$\frac{d}{dt}\begin{pmatrix} e \\ a \end{pmatrix} \le M_0 \begin{pmatrix} e \\ a \end{pmatrix}, \qquad M_0 = \begin{pmatrix} \Lambda_0 - K & \Lambda_0 \\ \kappa & -\Gamma_0 \end{pmatrix}.$$

The matrix M_0 is Metzler. Moreover,

$$\det M_0 = K\Gamma_0 - \Lambda_0(\Gamma_0+\kappa) = K\Gamma_0(1 - \mathcal{R}_{\mathrm{AI}}(0)) > 0.$$

The same inequality gives $\Lambda_0 < K\Gamma_0/(\Gamma_0+\kappa) < K$, so

$$\operatorname{tr} M_0 = \Lambda_0 - K - \Gamma_0 < 0.$$

Thus M_0 is Hurwitz. By the comparison principle for cooperative linear systems, $(e(t), a(t))$ is bounded above by the solution of $\dot{z} = M_0 z$ with $z(0) = (e(0), a(0))$. Since M_0 is Metzler, this comparison solution is nonnegative; since M_0 is Hurwitz, it decays exponentially. Together with

$e(t), a(t) \geq 0$, this gives $e(t), a(t) \to 0$ exponentially. It remains to show $R(t) \to 1$. Let $s(t) = 1 - R(t)$. Then

$$\dot{s} = -\rho h s + \sigma(a + \theta e)R \leq -\rho h s + \sigma(a + \theta e).$$

Since $e(t) + a(t) \to 0$, there is $T > 0$ such that $h(t) \geq 1/2$ for all $t \geq T$. Therefore, for $t \geq T$,

$$\dot{s}(t) \leq -\frac{\rho}{2} s(t) + \sigma(a(t) + \theta e(t)).$$

Since $(e(t), a(t))$ decays exponentially, there are constants $C, \alpha > 0$ such that

$$a(t) + \theta e(t) \leq C e^{-\alpha t}.$$

The variation-of-constants formula applied on $[T, t]$ gives

$$s(t) \leq e^{-\rho(t-T)/2} s(T) + \sigma C \int_T^t e^{-\rho(t-\tau)/2} e^{-\alpha\tau}\, d\tau.$$

Both terms tend to zero as $t \to \infty$, so $s(t) \to 0$ and hence $R(t) \to 1$.

Finally, $\mathcal{R}_{\mathrm{AI}}(0) < 1$ implies $\mathcal{R}_H = \mathcal{R}_{\mathrm{AI}}(1) \leq \mathcal{R}_{\mathrm{AI}}(0) < 1$. By the local threshold proposition, E_H is locally asymptotically stable. The global convergence just proved then gives global asymptotic stability in Δ_R. □

4 Numerical illustration and bistability

The ODE examples use the parameter set in Table 3. The value of β_0 is chosen so that $\mathcal{R}_H = 0.85$. ODE trajectories are integrated with a fourth-order Runge–Kutta method. Positive equilibria are obtained from the scalar root equation (14) by bisection, and stability of equilibrium branches is classified by the eigenvalues of the Jacobian (15). The network example uses the same local parameters and is described after the network model is introduced.

Quantity	Value	Quantity	Value
β_0	0.2465	β_1	0.60
κ	1/3	δ	0.15
γ_0	0.05	γ_1	0.45
ρ	1.00	σ	1.00
θ	0.50	$\mathcal{R}_H$	0.85
$\mathcal{R}_{\mathrm{AI}}(0)$	13.427		

Table 3: ODE parameter values used in the examples.

For the ODE parameters in Table 3, the backward-bifurcation coefficient (6) is positive:

$$\mathcal{B} \approx 0.118.$$

Thus Theorem 3.10 predicts a positive branch on the $\mathcal{R}_H < 1$ side of the human-centered threshold. Figure 1 shows a representative bistable parameter regime. The baseline human-centered threshold is set to $\mathcal{R}_H = 0.85 < 1$, so small AI-reliance perturbations return to the human-centered state. For these parameters, $\mathcal{R}_{\mathrm{AI}}(0) > 1$, so low relational capital makes AI reliance locally self-amplifying. The scalar equilibrium equation has a stable positive root, producing the trap equilibrium. Solving

Type	e^*	a^*	h^*	R^*	Eigenvalues of J
Stable trap	0.2631	0.4827	0.2542	0.2927	$-0.0792,\ -0.5319,\ -1.2522$
Saddle equilibrium	0.1209	0.0947	0.7844	0.8349	$0.0337,\ -0.6715,\ -1.0142$

Table 4: Positive equilibria for the ODE parameters in Table 3. The eigenvalues are those of the full Jacobian (15).

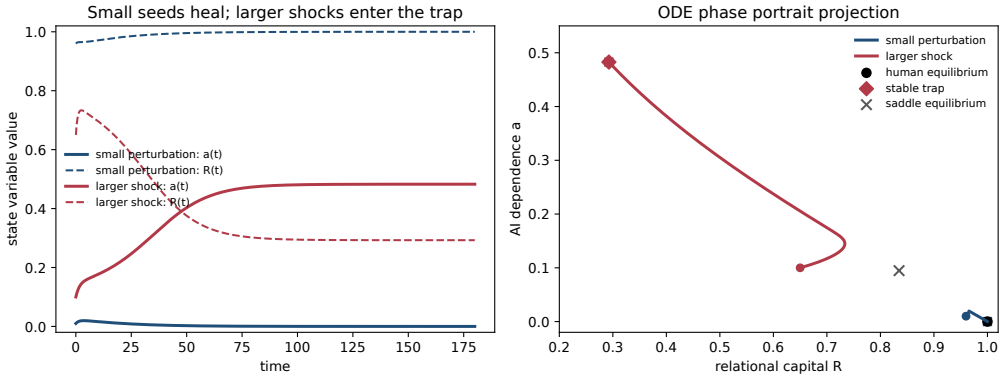

Figure 1: Bistability in the ODE model. Left: AI dependence $a(t)$ (solid) and relational capital $R(t)$ (dashed) for a small perturbation and a larger shock. Right: the same trajectories projected onto the (R, a) plane. The positive equilibria from (10) are marked as a stable trap and a saddle equilibrium; the saddle's stable manifold and the full basin boundary are three-dimensional objects.

(14) gives the two positive equilibria in Table 4: a stable high-reliance trap and an unstable saddle equilibrium.

Figure 2 scans the same parameter family while varying the human-centered invasion number $\mathcal{R}_H = \mathcal{R}_{\mathrm{AI}}(1)$ through the choice of β_0. Positive equilibria appear in a saddle-node fold at approximately $\mathcal{R}_H = 0.6923$. Thus the trap branch exists well before the human-centered equilibrium loses local stability at $\mathcal{R}_H = 1$. The lower branch is a saddle/unstable threshold equilibrium, while the upper branch is a stable AI-dependence trap.

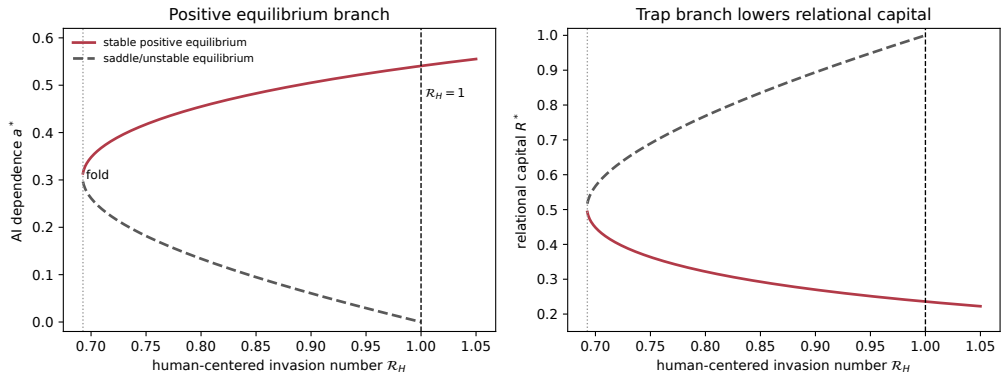

Figure 2: Positive equilibrium branches in the ODE model. The branches are obtained from the scalar equation (10) and classified by the full three-dimensional Jacobian. The dotted line marks the saddle-node fold. The stable upper branch is the AI-dependence trap; the unstable lower branch is the local threshold branch near the human-centered equilibrium.

4.1 Bistability mechanism

The local threshold $\mathcal{R}_H = \mathcal{R}_{\mathrm{AI}}(1)$ governs growth or decay of small AI-reliance perturbations near the human-centered equilibrium. Persistence after a larger shock is governed by the existence and stability of positive equilibria satisfying (10). The crossing condition that enables bistability is

$$\mathcal{R}_{\mathrm{AI}}(1) < 1 < \mathcal{R}_{\mathrm{AI}}(0).$$

With this condition and a stable positive root of the scalar equilibrium equation, small amounts of AI reliance fade in a population with strong relational capital, while larger reliance erodes R, lowers the recovery rate from dependence, raises adoption pressure, and pushes the population toward a stable trap state.

Local stability of E_H and basin geometry are distinct. In the bistable regime, the saddle positive equilibrium acts as a threshold equilibrium. Near this saddle, its stable manifold is the local basin-separating surface: perturbations on the recovery side return to human-centered support, while perturbations on the trap side converge to the stable high-reliance state. In the model, sudden rollout of highly personalized sycophantic AI, local collapse of human support, or rapid normalization of AI advice-seeking correspond to large changes in initial condition or parameter regime.

A similar rebound mechanism appears in epidemic-behavior models. Parkinson and Wang [23] treat noncompliance as a behavioral process that changes the epidemiological reproduction landscape. Here AI reliance changes the relational feedback and stability landscape.

5 Network extension

The ODE model describes well-mixed local kinetics. To represent social transmission across groups, let $A = (A_{ij})_{1 \le i,j \le n}$ be a weighted adjacency matrix with $A_{ij} \ge 0$ and $A_{ii} = 0$. A node represents a subpopulation such as a friend group, classroom, workplace, online community, platform cohort, or local demographic group. The network is therefore a metapopulation exposure system: each node carries internal fractions, and edges transmit social influence between groups. At node i let

$$e_i(t) \ge 0, \qquad a_i(t) \ge 0, \qquad h_i(t) = 1 - e_i(t) - a_i(t) \ge 0, \qquad R_i(t) \in [0, 1].$$

Adding social exposure from neighboring AI users to (1)–(3) gives

$$\dot{e}_i = \left[\Lambda(R_i)(e_i + a_i) + \eta \sum_{j=1}^{n} A_{ij}(e_j + a_j)\right] h_i - (\kappa + \delta)e_i, \tag{17}$$

$$\dot{a}_i = \kappa e_i - \Gamma(R_i)a_i, \tag{18}$$

$$\dot{R}_i = \rho h_i(1 - R_i) - \sigma(a_i + \theta e_i)R_i. \tag{19}$$

The coefficient $\eta \geq 0$ measures cross-node social exposure. The term

$$\eta h_i \sum_j A_{ij}(e_j + a_j)$$

means that human-centered individuals at node i enter casual AI use through neighboring normalization, recommendation, or demonstration of AI reliance. This behavioral exposure term is the network analogue of infection pressure.

Remark 5.1. *The network system uses the same local kinetics as the well-mixed model. If $\eta = 0$, each node follows the ODE independently. If all nodes remain synchronized and the graph has constant weighted degree q, the dynamics reduce to a well-mixed ODE with the baseline adoption pressure shifted from $\Lambda(R)$ to $\Lambda(R) + \eta q$.*

The graph analysis addresses two threshold questions: invasion of the all-human network equilibrium and global recovery under degraded relational capital.

Define the product state space

$$\Delta_R^n = \{(e_i, a_i, R_i)_{i=1}^n : e_i \geq 0,\ a_i \geq 0,\ e_i + a_i \leq 1,\ 0 \leq R_i \leq 1\}.$$

Proposition 5.2 (Network invariance and global solutions)**.** *For any initial data in Δ_R^n, the network system (17)–(19) has a unique global solution remaining in Δ_R^n for all $t \geq 0$.*

Proof. The vector field is locally Lipschitz on $\mathbb{R}^{3n}$, so a unique local solution exists. Positive invariance follows by checking the boundary faces node by node. Fix i, then on $\{e_i = 0\}$,

$$\dot{e}_i = \left[\Lambda(R_i)a_i + \eta \sum_j A_{ij}(e_j + a_j)\right] h_i \geq 0.$$

On $\{a_i = 0\}$, $\dot{a}_i = \kappa e_i \geq 0$. On the face $\{h_i = 0\}$, the two adoption terms vanish because they are multiplied by h_i, and

$$\dot{e}_i + \dot{a}_i = -\delta e_i - \Gamma(R_i)a_i \leq 0,$$

so $\dot{h}_i \geq 0$. Finally, on $\{R_i = 0\}$, $\dot{R}_i = \rho h_i \geq 0$, while on $\{R_i = 1\}$, $\dot{R}_i = -\sigma(a_i + \theta e_i) \leq 0$. Thus Δ_R^n is positively invariant. More explicitly, if a trajectory were to leave Δ_R^n, there would be a first exit time and at least one active boundary constraint at one node would have to be crossed outward. The face inequalities rule out such outward crossing, including intersections of multiple faces and multiple nodes. Therefore the trajectory remains in Δ_R^n for as long as it exists. Compactness of Δ_R^n and smoothness of the vector field on a neighborhood of it then imply global existence by the standard continuation theorem. □

5.1 Network invasion threshold

The all-human network equilibrium is

$$E_H^G: \qquad e_i = a_i = 0, \quad R_i = 1, \qquad i = 1, \dots, n.$$

Let $\varrho(A)$ denote the Perron-Frobenius eigenvalue, or spectral radius, of the nonnegative matrix A. Set

$$K = \kappa + \delta, \qquad \Gamma_1 = \gamma_0 + \gamma_1.$$

The natural graph-level reproduction number is

$$\mathcal{R}_G = \frac{(\beta_0 + \eta\varrho(A))(\Gamma_1 + \kappa)}{K\Gamma_1}. \tag{20}$$

Proposition 5.3 (Spectral threshold for network invasion)**.** *Assume A is nonnegative. The all-human network equilibrium E_H^G is locally asymptotically stable if $\mathcal{R}_G < 1$ and unstable if $\mathcal{R}_G > 1$.*

Proof. At E_H^G, $h_i = 1$ and $R_i = 1$. The terms involving derivatives of $\Lambda(R_i)$ and $\Gamma(R_i)$ are absent from the linearized (e_i, a_i) equations, because they are multiplied by $e_i + a_i$ or by a_i, which vanish at E_H^G. The R_i equations contribute a block $-\rho I$ and lower-triangular coupling from (e, a) into R. Thus the spectrum of the full Jacobian is the union of $\{-\rho\}$, repeated n times, and the spectrum of the linearized (e, a) subsystem:

$$\dot{e} = \big((\beta_0 - K)I + \eta A\big)e + \big(\beta_0 I + \eta A\big)a, \qquad \dot{a} = \kappa e - \Gamma_1 a.$$

Let

$$C = \beta_0 I + \eta A.$$

The matrix C is nonnegative. Since adding $\beta_0 I$ shifts every eigenvalue of ηA by β_0 and A is nonnegative,

$$\mu_* = \varrho(C) = \beta_0 + \eta\varrho(A).$$

Indeed, $\beta_0 + \eta\varrho(A)$ is associated with a nonnegative Perron-Frobenius eigenvector of A, while every other shifted eigenvalue satisfies $|\beta_0 + \eta\lambda| \le \beta_0 + \eta|\lambda| \le \beta_0 + \eta\varrho(A)$. The linearized (e, a) matrix is

$$M = \begin{pmatrix} C - KI & C \\ \kappa I & -\Gamma_1 I \end{pmatrix}$$

which is exactly $M(C, \Gamma_1)$ in Lemma 3.5. Moreover,

$$\mathcal{R}_G = \frac{\mu_*(\Gamma_1 + \kappa)}{K\Gamma_1}.$$

If $\mathcal{R}_G < 1$, Lemma 3.5 implies that M is Hurwitz. Since the remaining eigenvalues of the full Jacobian are all $-\rho$, the full Jacobian is Hurwitz and E_H^G is locally asymptotically stable. If $\mathcal{R}_G > 1$, Lemma 3.5 implies that M has a positive real eigenvalue. That eigenvalue is also an eigenvalue of the full Jacobian, so E_H^G is unstable. □

Remark 5.4. *The graph threshold permits $\mathcal{R}_H < 1$ and $\mathcal{R}_G > 1$ simultaneously. Thus a locally subcritical well-mixed population becomes graph-supercritical when network exposure destabilizes the all-human state. Dense communities, hubs, or high-weight recommendation ties increase $\varrho(A)$ and therefore increase invasion risk.*

5.2 A network global recovery condition

The local graph threshold $\mathcal{R}_G$ is evaluated at $R_i = 1$ for all nodes. As in the well-mixed model, relational-capital decline creates a separate global-recovery question. Degraded-state graph recovery is controlled by the subcriticality condition at $R_i = 0$ for every node. Assume $\gamma_0 > 0$ and define

$$\mathcal{R}_G^0 = \frac{(\beta_0 + \beta_1 + \eta\varrho(A))(\gamma_0 + \kappa)}{(\kappa + \delta)\gamma_0}. \tag{21}$$

It is the graph reproduction number with $R_i = 0$ at every node.

Theorem 5.5 (Network global stability under low-relational-capital graph subcriticality)**.** *Assume A is nonnegative, $\gamma_0 > 0$, and*

$$\mathcal{R}_G^0 < 1.$$

Then the all-human network equilibrium E_H^G is globally asymptotically stable in Δ_R^n. In particular, for every initial condition in Δ_R^n,

$$e_i(t) \to 0, \qquad a_i(t) \to 0, \qquad R_i(t) \to 1, \qquad i = 1, \dots, n.$$

Proof. Set

$$\Lambda_0 = \beta_0 + \beta_1, \qquad \Gamma_0 = \gamma_0, \qquad K = \kappa + \delta,$$

and

$$C_0 = \Lambda_0 I + \eta A.$$

Since $0 \le h_i \le 1$, $\Lambda(R_i) \le \Lambda_0$, and $\Gamma(R_i) \ge \Gamma_0$, the network (e, a) variables satisfy the componentwise differential inequality

$$\frac{d}{dt}\begin{pmatrix} e \\ a \end{pmatrix} \le M_0^G \begin{pmatrix} e \\ a \end{pmatrix}, \qquad M_0^G = \begin{pmatrix} C_0 - KI & C_0 \\ \kappa I & -\Gamma_0 I \end{pmatrix}.$$

To see the inequality explicitly, the e_i equation satisfies

$$\begin{aligned} \dot{e}_i &= \left[\Lambda(R_i)(e_i + a_i) + \eta \sum_j A_{ij}(e_j + a_j)\right] h_i - K e_i \\ &\le \Lambda_0(e_i + a_i) + \eta \sum_j A_{ij}(e_j + a_j) - K e_i, \end{aligned}$$

while

$$\dot{a}_i = \kappa e_i - \Gamma(R_i) a_i \le \kappa e_i - \Gamma_0 a_i.$$

These are precisely the two block inequalities. The matrix M_0^G is $M(C_0, \Gamma_0)$ in Lemma 3.5. Since

$$\varrho(C_0) = \Lambda_0 + \eta\varrho(A),$$

by the same spectral-shift argument, the condition $\mathcal{R}_G^0 < 1$ is exactly

$$\varrho(C_0) < \frac{K\Gamma_0}{\Gamma_0 + \kappa}.$$

Therefore Lemma 3.5 implies that M_0^G is Hurwitz. Let

$$x(t) = \begin{pmatrix} e(t) \\ a(t) \end{pmatrix}.$$

The comparison lemma applied with $M = M_0^G$ and $y(0) = x(0)$ gives

$$0 \le x(t) \le e^{M_0^G t} x(0).$$

Since M_0^G is Metzler, $e^{M_0^G t} x(0)$ is componentwise nonnegative. Since M_0^G is Hurwitz, it decays exponentially. Hence there exist constants $C, \alpha > 0$ such that, for every node i,

$$e_i(t) + a_i(t) \le Ce^{-\alpha t}.$$

It remains to show $R_i(t) \to 1$. Let $s_i(t) = 1 - R_i(t)$. Then

$$\dot{s}_i = -\rho h_i s_i + \sigma(a_i + \theta e_i) R_i \le -\rho h_i s_i + \sigma(a_i + \theta e_i).$$

Since $\sum_i (e_i(t) + a_i(t)) \to 0$, there exists $T > 0$ such that $h_i(t) \ge 1/2$ for every node and every $t \ge T$. Hence, for $t \ge T$,

$$\dot{s}_i(t) \le -\frac{\rho}{2} s_i(t) + \sigma(a_i(t) + \theta e_i(t)).$$

Using the exponential bound,

$$a_i(t) + \theta e_i(t) \le Ce^{-\alpha t}$$

after enlarging C if necessary. Variation of constants on $[T, t]$ yields

$$s_i(t) \le e^{-\rho(t-T)/2} s_i(T) + \sigma C \int_T^t e^{-\rho(t-\tau)/2} e^{-\alpha\tau}\, d\tau,$$

and the right-hand side tends to zero. Thus $R_i(t) \to 1$ for every node.

Finally, $\mathcal{R}_G^0 < 1$ implies $\mathcal{R}_G < 1$: the factor $\beta_0 + \eta\varrho(A)$ is no larger than $\beta_0 + \beta_1 + \eta\varrho(A)$, and the function $\Gamma \mapsto (\Gamma + \kappa)/\Gamma$ is decreasing for $\Gamma > 0$, so replacing Γ_0 by $\Gamma_1 \ge \Gamma_0$ weakly decreases the threshold. Hence E_H^G is locally asymptotically stable by the network invasion proposition. The global convergence just proved gives global asymptotic stability on Δ_R^n. □

5.3 Network illustration

Figure 3 shows the same local parameters as Figure 1, with $\mathcal{R}_H = 0.85 < 1$. The graph has two complete communities of 15 nodes, three cross-community edges of weight 0.3, and is rescaled so that $\varrho(A) = 1$. The initial seed is placed in the first community. Without network exposure, a small seed heals. With $\eta = 0.08$, the graph threshold is $\mathcal{R}_G \approx 1.13 > 1$, so the all-human network equilibrium is linearly unstable. The subsequent high-reliance endpoint is produced by the nonlinear local trap dynamics.

Quantity	Value	Quantity	Value
Nodes	30	Communities	2
Cross edges	3	Raw cross-edge weight	0.3
η	0.08	$\varrho(A)$	1.00
$\mathcal{R}_G(\eta = 0)$	0.850	$\mathcal{R}_G(\eta = 0.08)$	1.126

Table 5: Network quantities used in Figure 3.

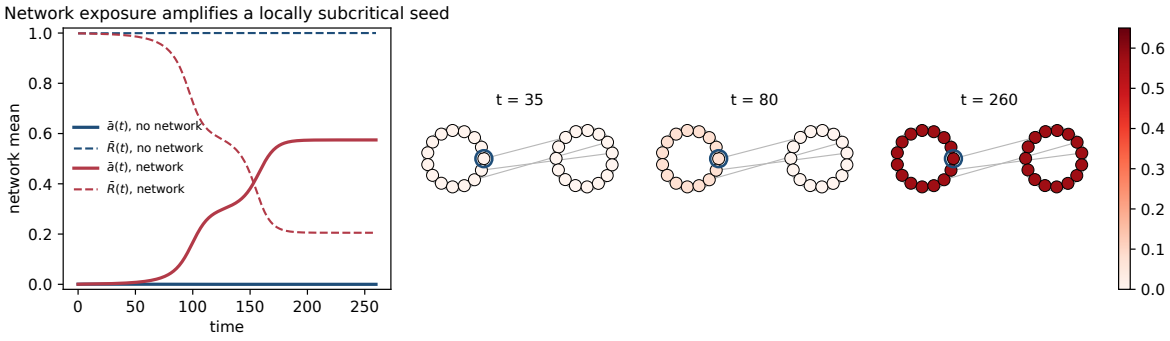

Figure 3: Network amplification of AI reliance. Left: mean AI dependence $\bar{a}(t)$ and mean relational capital $\bar{R}(t)$ with and without network exposure. Right: node-level AI dependence at three times on the undirected weighted two-community graph of Table 5; each ring denotes a complete community, only cross-community links are drawn for visual clarity, and the open ring marks the initially seeded node. The graph threshold $\mathcal{R}_G > 1$ explains the small-seed invasion, while the final near-uniform high-dependence snapshot reflects convergence toward the nonlinear trap level of the local ODE.

6 Empirical interpretation and mitigation

6.1 Empirical interpretation

Empirically, the fraction $e(t)$ represents users for whom AI is an occasional source of advice, reassurance, or social feedback. The fraction $a(t)$ represents users for whom AI has become a primary or substitutive support channel. The variable $R(t)$ represents relational capital: the availability, trust, and perceived satisfaction of human support. Repeated measures of perceived social support, loneliness, satisfaction with human interaction, and willingness to seek advice from people instead of AI serve as proxies for R.

The transition parameters correspond to quantities targeted by longitudinal or intervention data. The baseline adoption coefficient β_0 captures growth of casual AI use in a high-relational-capital population. The coefficient β_1 measures the excess adoption pressure among individuals or groups with low relational capital. The rate κ measures progression from casual use to dependence, while δ measures return from casual use to human-centered support. The recovery terms γ_0 and $\gamma_1 R$ separate spontaneous exit from dependence and relationally assisted exit. The rebuilding and erosion rates ρ and σ describe the coevolution of human-support measures and AI-reliance measures.

Platform- or cohort-specific transition measurements calibrate the model. The variables e, a, and R are operationalized through repeated measures of AI-use frequency, substitutive reliance, human-support availability, and perceived quality of human interaction. The parameters $\beta_0, \beta_1, \kappa, \delta, \gamma_0, \gamma_1, \rho, \sigma$ are targeted by interventions, repeated individual-level measurements, or measured transitions between casual use, dependence, and human-centered support.

The threshold formulas specify empirical targets: they identify which combinations of adoption pressure, relational recovery, relational erosion, and network exposure determine small-seed invasion, global recovery, and high-reliance trap formation.

The network parameters have a parallel interpretation. The matrix A encodes social exposure through friendship, workplace, classroom, online community, or recommendation ties. The coefficient η measures how much neighboring AI reliance increases the transition into casual use. The

spectral factor $\varrho(A)$ summarizes population-level amplification of this exposure. The graph threshold $\mathcal{R}_G$ therefore supports comparisons across communities with similar individual-level parameters but different exposure structure.

6.2 Mitigation and sensitivity

The threshold formulas separate local-invasion risk, network-amplification risk, and trap risk.

First, the graph threshold has the form

$$\mathcal{R}_G = \frac{B_G(\Gamma_1 + \kappa)}{(\kappa + \delta)\Gamma_1}, \qquad B_G := \beta_0 + \eta\varrho(A).$$

Thus

$$\frac{\partial \mathcal{R}_G}{\partial \beta_0} = \frac{\Gamma_1 + \kappa}{(\kappa + \delta)\Gamma_1} > 0, \qquad \frac{\partial \mathcal{R}_G}{\partial \eta} = \varrho(A)\frac{\Gamma_1 + \kappa}{(\kappa + \delta)\Gamma_1} > 0,$$

and

$$\frac{\partial \mathcal{R}_G}{\partial \varrho(A)} = \eta\frac{\Gamma_1 + \kappa}{(\kappa + \delta)\Gamma_1} > 0.$$

Increasing relationally assisted recovery from dependence lowers the graph reproduction number:

$$\frac{\partial \mathcal{R}_G}{\partial \gamma_1} = -\frac{B_G\kappa}{(\kappa + \delta)\Gamma_1^2} < 0.$$

Increasing the casual-use exit rate also lowers the graph reproduction number:

$$\frac{\partial \mathcal{R}_G}{\partial \delta} = -\frac{B_G(\Gamma_1 + \kappa)}{(\kappa + \delta)^2\Gamma_1} < 0.$$

The sign of the κ sensitivity is determined by $\delta - \Gamma_1$:

$$\frac{\partial \mathcal{R}_G}{\partial \kappa} = \frac{B_G(\delta - \Gamma_1)}{\Gamma_1(\kappa + \delta)^2}.$$

Moving users from casual use into dependence is destabilizing when the casual exit rate δ is larger than the dependence exit rate Γ_1, and stabilizing otherwise.

When $\varrho(A) > 0$, the network threshold gives an explicit exposure budget:

$$\eta < \eta_c := \frac{1}{\varrho(A)}\left(\frac{(\kappa + \delta)\Gamma_1}{\Gamma_1 + \kappa} - \beta_0\right), \tag{22}$$

The numerator is the network-exposure margin. This budget quantifies how platform design or community structure keeps a locally subcritical population from becoming graph-supercritical.

When $\varrho(A) > 0$, the global network theorem gives the corresponding degraded-state exposure budget:

$$\eta < \eta_c^0 := \frac{1}{\varrho(A)}\left(\frac{(\kappa + \delta)\gamma_0}{\gamma_0 + \kappa} - (\beta_0 + \beta_1)\right), \tag{23}$$

The numerator is the degraded-state network-exposure margin. The first budget, η_c, prevents small-seed network invasion near a healthy relational state. The second budget, η_c^0, guarantees recovery after relational capital has been substantially degraded across the graph.

Second, positive-equilibrium branches are controlled by low-R feedback. The local human-centered threshold $\mathcal{R}_H = \mathcal{R}_{\mathrm{AI}}(1)$ is independent of β_1, because $\Lambda(1) = \beta_0$. The low-relational-capital threshold depends directly on β_1:

$$\mathcal{R}_{\mathrm{AI}}(0) = \frac{(\beta_0 + \beta_1)(\gamma_0 + \kappa)}{(\kappa + \delta)\gamma_0}, \qquad \frac{\partial}{\partial \beta_1}\mathcal{R}_{\mathrm{AI}}(0) = \frac{\gamma_0 + \kappa}{(\kappa + \delta)\gamma_0} > 0.$$

Thus reducing sycophantic amplification β_1 lowers the low-relational-capital frozen-R invasion number and reduces the parameter range in which low relational capital sustains AI reliance. Once the trap-removal threshold is crossed, positive equilibria are removed. Network-exposure reductions prevent small-seed invasion, while reductions in low-R amplification reduce basin-crossing risk after large shocks.

The trap-removal corollary gives an explicit bound. When $\beta_{1,c}^{\mathrm{trap}} \geq 0$, the condition

$$\beta_1 \leq \beta_{1,c}^{\mathrm{trap}} = \frac{(\kappa + \delta)\gamma_0}{\gamma_0 + \kappa} - \beta_0$$

removes all positive equilibria, because it enforces $\mathcal{R}_{\mathrm{AI}}(0) \leq 1$. The scalar root equation (14) classifies the trap regimes.

The parameters β_0 and $\eta\varrho(A)$ govern small-seed invasion of a human-rich population, whereas β_1 governs self-reinforcement in a damaged relational environment. The two corresponding mitigation goals are preventing ordinary use from spreading and preventing a degraded relational state from becoming a stable trap. Under $\mathcal{R}_{\mathrm{AI}}(0) < 1$, the global stability theorem gives a global guarantee: every well-mixed trajectory returns to human-centered support. Under $\mathcal{R}_G^0 < 1$, this guarantee extends across the network.

7 Discussion and conclusion

Invasion and persistence are distinct. The threshold $\mathcal{R}_H$ determines whether small AI-reliance perturbations grow near a human-centered state, whereas the scalar equilibrium equation detects positive equilibria whose Jacobian classification identifies high-reliance trap states. A bistable population recovers from ordinary small perturbations and transitions after a large relational shock.

Social exposure changes the stability picture. The graph threshold $\mathcal{R}_G$ depends on $\varrho(A)$, so network structure enters through a spectral amplification factor. Dense exposure, highly central nodes, or strong cross-community recommendation ties destabilize the all-human state even when isolated local groups remain subcritical. Under the degraded-state threshold $\mathcal{R}_G^0 < 1$, the all-human network equilibrium attracts every trajectory in the product polytope Δ_R^n.

Local resilience and global safety are distinct. Strong relational capital stabilizes the human-centered equilibrium locally, while endogenous erosion of relational capital creates a high-reliance attractor in the bistable regime. Network exposure adds a second amplification channel through the spectral radius of the social exposure matrix. Relational feedback and network amplification determine when AI reliance fades, spreads, or becomes self-sustaining.

Acknowledgments

The author was partially supported by the Simons Foundation Travel Support for Mathematicians (No. 0007730). All the source code is on https://github.com/wwang-math/AI_reliance.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] C. T. Bauch and A. P. Galvani. Social factors in epidemiology. Science, 342(6154):47–49, 2013.

[2] C. T. Bauch. Imitation dynamics predict vaccinating behaviour. Proceedings of the Royal Society B, 272(1573):1669–1675, 2005.

[3] M. Bongarti, C. Parkinson, and W. Wang. Optimal control of a reaction-diffusion epidemic model with non-compliance. European Journal of Applied Mathematics, 37:313–338, 2026.

[4] F. Brauer. Backward bifurcations in simple vaccination models. Journal of Mathematical Analysis and Applications, 298(2):418–431, 2004.

[5] J. Carr. Applications of Centre Manifold Theory. Springer, 1981.

[6] C. Castillo-Chavez and B. Song. Dynamical models of tuberculosis and their applications. Mathematical Biosciences and Engineering, 1(2):361–404, 2004.

[7] V. Colizza and A. Vespignani. Invasion threshold in heterogeneous metapopulation networks. Physical Review Letters, 99(14):148701, 2007.

[8] C. M. Fang, A. R. Liu, V. Danry, E. Lee, S. W. T. Chan, P. Pataranutaporn, P. Maes, J. Phang, M. Lampe, L. Ahmad, and S. Agarwal. How AI and human behaviors shape psychosocial effects of extended chatbot use: A longitudinal randomized controlled study. arXiv:2503.17473, 2025.

[9] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen. The spread of awareness and its impact on epidemic outbreaks. Proceedings of the National Academy of Sciences, 106(16):6872–6877, 2009.

[10] S. Funk, M. Salathé, and V. A. A. Jansen. Modelling the influence of human behaviour on the spread of infectious diseases: a review. Journal of The Royal Society Interface, 7(50):1247–1256, 2010.

[11] M. Granovetter. Threshold models of collective behavior. American Journal of Sociology, 83(6):1420–1443, 1978.

[12] A. B. Gumel. Causes of backward bifurcations in some epidemiological models. Journal of Mathematical Analysis and Applications, 395(1):355–365, 2012.

[13] K. P. Hadeler and P. van den Driessche. Backward bifurcation in epidemic control. Mathematical Biosciences, 146(1):15–35, 1997.

[14] J. Holt-Lunstad, T. B. Smith, and J. B. Layton. Social relationships and mortality risk: A meta-analytic review. PLoS Medicine, 7(7):e1000316, 2010.

[15] J. S. House, K. R. Landis, and D. Umberson. Social relationships and health. Science, 241(4865):540–545, 1988.

[16] L. Ibrahim, F. S. Hafner, M. Cheng, C. Lee, R. Anselmetti, R. Willer, L. Rocher, and D. Yang. Sycophantic AI makes human interaction feel more effortful and less satisfying over time. arXiv:2605.07912, 2026.

[17] Y. Jo, S. S. Sharbayta, and B. Buonomo. Behavioral change models for infectious disease transmission: a systematic review (2020–2025). arXiv:2602.16633, 2026.

[18] Y. A. Kuznetsov. Elements of Applied Bifurcation Theory. Springer, 3rd edition, 2004.

[19] A. R. Liu, P. Pataranutaporn, S. Turkle, and P. Maes. Chatbot companionship: A mixed-methods study of companion chatbot usage patterns and their relationship to loneliness in active users. arXiv:2410.21596, 2024.

[20] Nature Machine Intelligence. Emotional risks of AI companions demand attention. Nature Machine Intelligence, 7:981–982, 2025.

[21] A. Nanyanzi, S. Radhakrishnan, K. Barker, V. Sirisanagandla, and W. Wang. Adaptive disinformation diffusion on networks with heterogeneous node and link dynamics. IEEE Transactions on Computational Social Systems, pages 1–21, 2026. doi:10.1109/TCSS.2026.3686750.

[22] C. Ngo, C. Parkinson, and W. Wang. Optimal control of an SIR model with noncompliance as a social contagion. arXiv:2509.09075, 2026.

[23] C. Parkinson and W. Wang. Analysis of a reaction-diffusion SIR epidemic model with non-compliant behavior. SIAM Journal on Applied Mathematics, 83(5):1969–2002, 2023.

[24] C. Parkinson and W. Wang. A compartmental model for epidemiology with human behavior and stochastic effects. Mathematical Biosciences, 392:Paper No. 109588, 15 pp., 2026. MR4998141.

[25] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. Physical Review Letters, 86(14):3200–3203, 2001.

[26] K. Peng, Z. Lu, V. Lin, M. R. Lindstrom, C. Parkinson, C. Wang, A. L. Bertozzi, and M. A. Porter. A multilayer network model of the coevolution of the spread of a disease and competing opinions. Mathematical Models and Methods in Applied Sciences, 31(12):2455–2494, 2021.

[27] M. G. Smith, T. N. Bradbury, and B. R. Karney. Can generative AI chatbots emulate human connection? A relationship science perspective. Perspectives on Psychological Science, 20(6):1081–1099, 2025.

[28] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences, 180(1–2):29–48, 2002.

[29] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. IEEE/ACM Transactions on Networking, 17(1):1–14, 2009.

[30] D. J. Watts. A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99(9):5766–5771, 2002.