Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Trust-Building Game: Dynamic Signaling, Markov Equilibrium, and Threshold Cooperation

Madjid Eshaghi Gordji
Semnan University
Mohamadali Berahman
mohamadali_berahman@semnan.ac.ir
Semnan University
Zeinab Shahini
Semnan University

Research Article

Keywords: trust-building game, dynamic game, repeated interaction, Markov strategies, dynamic programming, optimal signaling, threshold cooperation, state persistence

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11101808/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

The Trust-Building Game: Dynamic Signaling, Markov Equilibrium, and Threshold Cooperation

Madjid Eshaghi Gordji[1] Mohamadali Berahman[1,*] Zeinab Shahini[1]

[1]Faculty of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran
[*]Corresponding author: mohamadali_berahman@semnan.ac.ir

Abstract

Trust in repeated interaction depends on how observed behavior changes future incentives as well as current responses. This paper develops a discounted dynamic Trust-Building Game in which a trust builder chooses observable binary signals, a public state aggregates past signals through a persistence-weighted recursion, and a trustee observes the updated state before deciding whether to cooperate. The trustee's cooperation cutoff is derived as a dynamic best response under stationary Markov strategies, while the builder's problem is characterized by a Bellman equation. We establish existence of a pure stationary Markov equilibrium, boundedness and monotonicity of the builder's value function, and exact bounds on the gain from positive signaling. These results identify a parameter regime in which the builder's unique optimal stationary policy has an activation-band structure: signaling is optimal only over an intermediate interval of public states, demonstrating that a monotone one-sided threshold policy cannot be imposed in general. We further derive the weighted-history representation of the public state, exact first-passage and cooperation times, finite-horizon cooperation fractions, geometric decay after signaling stops, and the minimum current signaling cost required to preserve cooperation for one period. Deterministic Bellman computations recover the analytically derived activation band without policy mismatches on the numerical grid and examine optimal signaling outside the analytically characterized regime. Additional computations verify the distinction between the threshold-state index and the first cooperative period and preserve cooperation at threshold equality. The framework is deliberately reduced form: the public state is neither a Bayesian posterior nor an empirically calibrated measure of psychological trust, and the numerical results are model evaluations rather than empirical evidence.

Keywords: trust-building game; dynamic game; repeated interaction; Markov strategies; dynamic programming; optimal signaling; threshold cooperation; state persistence.

1 Introduction

The investment game of Berg, Dickhaut, and McCabe established a widely used experimental framework for studying decisions involving interpersonal vulnerability and subsequent reciprocity.[1] Meta-analytic evidence shows that behavior in this paradigm is sensitive to experimental design and context, cautioning against interpreting observed transfers as a context-free measure of a single latent characteristic.[2] More recent meta-analytic evidence from repeated

trust games indicates that trial-by-trial learning is strongly associated with a partner's observed reciprocation behavior.[3] These findings motivate models in which current behavior depends not only on the present interaction but also on a state summarizing previously observed conduct.

Repeated versions of the investment game provide further motivation for an explicit dynamic formulation. Cochard, Nguyen Van, and Willinger document temporal variation and an end-game decline in a finitely repeated investment game,[4] while Engle-Warnick and Slonim show that strategies evolve differently when the duration of a relationship is definite rather than indefinite.[5] King-Casas and coauthors provide complementary evidence that repeated exchange generates adaptation to a partner's previous behavior.[6] Such evidence supports the use of history-dependent state variables, but it does not establish any particular deterministic updating equation as an empirical law. The state recursion studied in this paper is therefore treated as a reduced-form mathematical mechanism whose implications can be analyzed exactly.

A separate literature shows why observed cooperative behavior should not be identified mechanically with trust. Promises and higher-order beliefs can affect cooperation,[7] and expressed trust may itself influence subsequent trustworthy behavior.[8] Cox emphasizes the need to distinguish trust and reciprocity from unconditional other-regarding preferences,[9] while theories of sequential reciprocity, intention-based reciprocity, and guilt aversion allow history, intentions, and beliefs to enter strategic evaluation in ways that are absent from purely material stage-payoff models.[10–12]

The same caution is reinforced by the broader social-preference literature. Inequality aversion and relative-payoff concerns can affect observed behavior,[13, 14] while betrayal aversion distinguishes interpersonal vulnerability from formally comparable nonsocial risk.[15, 16] Experimental evidence also documents heterogeneity associated with expectations, reciprocity, and context.[17, 18] Biological interventions have likewise been studied as possible influences on trusting behavior, but such results concern empirical mechanisms rather than the deterministic public state considered here.[19] Fehr emphasizes that trust is related to, but not reducible to, ordinary risk taking,[20] and Fehr and Charness place reciprocity, guilt aversion, fairness, and related social preferences within a broader account of social preferences in economics.[21]

Trust and reputation also arise in theoretical settings beyond the bilateral laboratory investment game. Evolutionary models identify direct and indirect reciprocity as mechanisms capable of supporting cooperation,[22, 23] and Manapat, Nowak, and Rand analyze how information and partner choice can sustain trust and trustworthiness.[24] The corporate-reputation literature interprets reputation as an inference from past actions,[25] whereas classical game-theoretic reputation models frequently rely on incomplete information about persistent types.[26, 27] Computational trust systems provide another point of comparison because they often summarize historical information in an evolving score.[28] The public state studied below is closest in spirit to such a reduced-form summary, but it is neither a Bayesian posterior nor an empirically calibrated reputation measure.

Applications of trust games and repeated trust interactions extend to settings involving institutions and artificial agents. Evolutionary analyses have considered repeated trust involving intelligent machines,[29] immersive experiments have adapted the investment game to human-robot interaction,[30] and repeated group models have examined conditional investment based on

perceived trustworthiness.[31] Framing and order effects can alter observed trust and reciprocity, [32] while institution-specific experiments show that generalized trust and trust in a particular institution need not coincide.[33] Trust has also been associated with financial participation, illustrating the economic relevance of repeated trust-related decisions beyond laboratory exchange. [34]

The present paper does not claim novelty for exponentially weighted state updating itself, nor does it attempt to replace Bayesian reputation theory or empirical models of psychological trust. Its contribution is instead to embed a simple public-state recursion in a fully specified discounted sequential game and to characterize the strategic consequences that follow from that structure. Player 1, the trust builder, chooses a costly observable binary signal that changes the public state. Player 2, the trustee, observes the updated state before choosing whether to cooperate. Because the trustee's current action does not affect the state transition, the continuation term cancels from the trustee's action comparison, yielding an exact cooperation threshold as a stationary Markov best response.

The builder's problem is then reduced to a discounted Bellman equation rather than being assigned an assumed behavioral rule. We establish existence of a pure stationary Markov equilibrium and prove that the builder's optimal value function is bounded and nondecreasing. By comparing the Bellman values of the two signaling actions, we derive exact bounds on the signaling gain and identify a nonempty parameter regime in which the unique optimal stationary policy has an activation-band form. In that regime, signaling is optimal only over an intermediate interval of public states. Under additional admissible parameter restrictions, the optimal policy therefore has a no-signal–signal–no-signal structure, showing that a monotone one-sided threshold policy cannot be imposed in general.

A second set of results concerns state dynamics independently of optimality. We derive the weighted-history representation of the public state, order preservation with respect to signal histories, the earliest attainable threshold-crossing time, the corresponding first cooperative period, and the exact finite-horizon cooperation fraction under persistent positive signaling. We also characterize geometric state decay after positive signaling stops and the minimum current signaling cost required to preserve cooperation for one period. A distinction is maintained throughout between the first state index τ_T satisfying the cooperation threshold and the first cooperative period τ_C, with

$$\tau_C = \tau_T - 1,$$

because the trustee acts after observing the updated state.

The analytical results are complemented by deterministic numerical dynamic programming rather than by empirical calibration or stochastic simulation. First, value iteration independently recovers the analytically derived activation-band policy for a parameter configuration satisfying the theorem's conditions, with no policy mismatch on the evaluated state grid. Second, a broader parameter scan examines how the Bellman-optimal signaling region changes with the builder's discount factor and signaling cost outside the analytically characterized regime. The numerical scan is treated as computational evidence, not as a substitute for a global theorem. Finally, exact state calculations verify the correct cooperation timing, finite-horizon

cooperation fractions, and treatment of threshold equality.

Figure 1 summarizes the architecture of the model and the logical connection between its main results. Panel (A) shows the within-period sequence from the builder's signal to the public-state update and the trustee's threshold response. Panel (B) separates the builder's Bellman optimization from the trustee's current cooperation rule. Panel (C) collects the state-dynamic implications for weighted history, threshold-crossing timing, and post-signal decay. Panel (D) contrasts the trustee's one-sided cooperation threshold with the builder's activation-band policy, which is the central structural distinction established by the analysis.

2 Dynamic Game Formulation

The purpose of this section is to formulate the Trust-Building Game as a well-defined discounted dynamic game rather than as a state recursion supplemented by an exogenous response rule. The model has one-sided state control: Player 1 chooses the action that changes the public trust score, whereas Player 2 observes the updated score and decides whether to cooperate. Both players evaluate an infinite stream of payoffs. Because Player 2's current action does not affect the law of motion of the state, the continuation component of Player 2's optimization does not alter the current cooperation decision. This feature will allow the cooperation threshold to be derived as a genuine dynamic best response.

2.1 Timing, information, and public state

Time is discrete, $t = 0, 1, 2, \ldots$. At the beginning of period t, both players observe the public state

$$T_t \in [0, 1], \tag{1}$$

which summarizes the accumulated publicly observable record relevant to the interaction.

The variable T_t is a reduced-form public trust score. It is not interpreted as a Bayesian posterior probability, nor is it claimed to be an empirically identified measure of psychological trust. Its role is to provide a scalar state variable through which past observable behavior affects future cooperation.

Player 1, called the *trust builder*, moves first and chooses

$$s_t \in \{0, 1\}. \tag{2}$$

The action $s_t = 1$ denotes a publicly observable action classified by the model as favorable evidence, whereas $s_t = 0$ denotes the absence of such favorable input. The model does not derive the credibility of this classification endogenously. In particular, $s_t = 0$ should not be interpreted as an active betrayal; under the exponential updating rule below, it simply produces passive decay of the public score.

After Player 1 acts, the within-period post-signal state is

$$\widehat{T}_t = F(T_t, s_t) := \lambda T_t + (1 - \lambda) s_t, \qquad 0 \leq \lambda < 1. \tag{3}$$

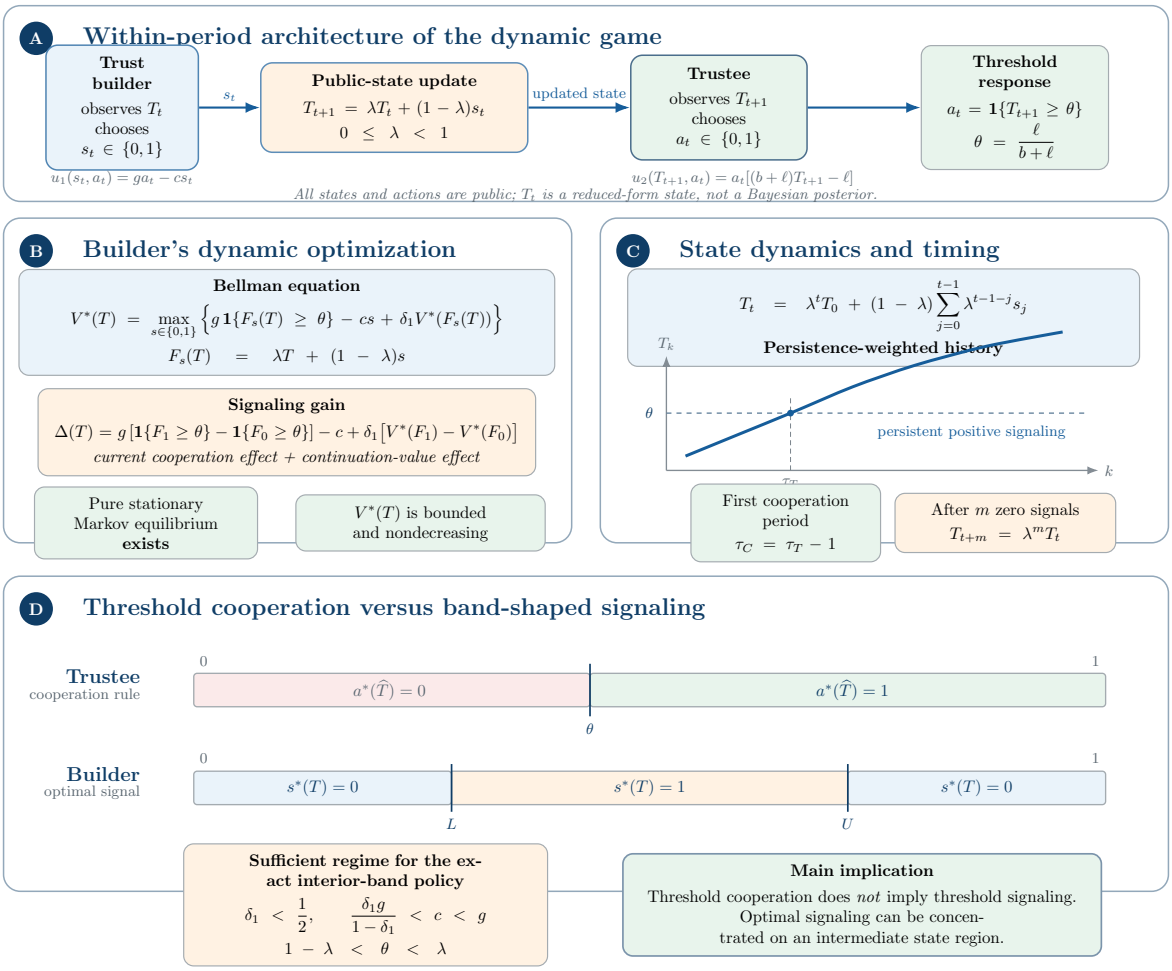


Figure 1: Conceptual architecture and principal results of the Trust-Building Game. Panel (A) summarizes the within-period sequence: the trust builder chooses an observable binary signal s_t, the public state evolves according to $T_{t+1} = \lambda T_t + (1-\lambda)s_t$, and the trustee subsequently cooperates when the updated state satisfies $T_{t+1} \ge \theta$, where $\theta = \ell/(b+\ell)$. Panel (B) presents the builder's Bellman problem and decomposes the gain from positive signaling into its current cooperation effect, signaling cost, and continuation-value effect. Panel (C) summarizes the state dynamics, including the persistence-weighted history representation, the timing relation $\tau_C = \tau_T - 1$, and geometric state decay after signaling stops. Panel (D) highlights the main structural result: although the trustee's cooperation decision follows a one-sided threshold rule, the builder's optimal stationary signaling policy can have an interior activation-band structure, with no signaling at low and high states and positive signaling only at intermediate states under the stated sufficient conditions.

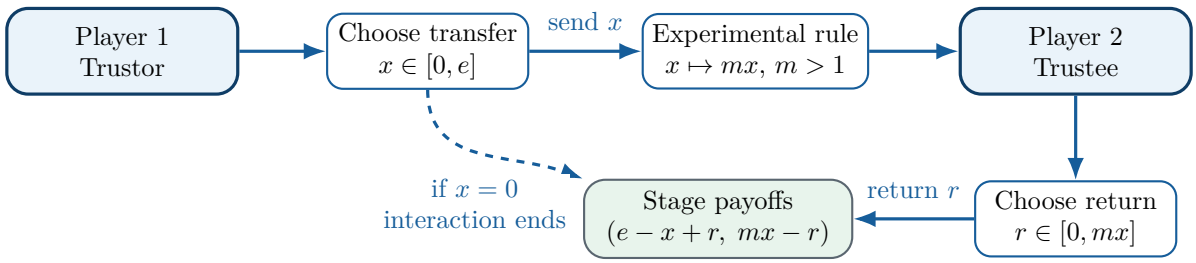


Figure 2: Conceptual structure of the classical investment game used as a point of comparison. Player 1 transfers x from an initial endowment e, the experimental rule multiplies the transfer by $m > 1$, and Player 2 subsequently chooses a return r. The dynamic model developed in this paper replaces this within-stage transfer-return mechanism with an observable signaling action, a public-state transition, and a subsequent cooperation decision.

The parameter λ is therefore a persistence parameter. A larger value of λ assigns more weight to the previous public state and less weight to the current action. No psychological interpretation of λ is imposed.

Player 2, called the *trustee*, observes $\widehat{T}_t$ and subsequently chooses

$$a_t \in \{0, 1\}, \tag{4}$$

where $a_t = 1$ denotes cooperation and $a_t = 0$ denotes noncooperation.

After the two actions and the period payoffs have been realized, the next-period state is

$$T_{t+1} = \widehat{T}_t. \tag{5}$$

Hence the timing within every period is

$$T_t \longrightarrow s_t \longrightarrow \widehat{T}_t \longrightarrow a_t \longrightarrow \text{period payoffs} \longrightarrow T_{t+1} = \widehat{T}_t.$$

All actions and states are publicly observed. Thus there are no hidden types, private signals, or Bayesian beliefs in the baseline model.

2.2 Payoffs and discounted objectives

The trustee's payoff must specify why cooperation becomes desirable only after the public score is sufficiently high. Let $b > 0$ and $\ell > 0$. At a post-signal state $\widehat{T}$, Player 2 receives the stage payoff

$$u_2(\widehat{T}, a) = a\big[(b + \ell)\widehat{T} - \ell\big]. \tag{6}$$

Noncooperation is normalized to zero. Equation (6) is an affine reduced-form utility index. Although it can be written as $b\widehat{T} - \ell(1 - \widehat{T})$ when $a = 1$, the variable $\widehat{T}$ is not assumed to be a probability.

Player 1 obtains a benefit $g > 0$ whenever Player 2 cooperates and incurs a cost $c > 0$ whenever a positive signal is supplied:

$$u_1(s, a) = ga - cs. \tag{7}$$

Let $\delta_1, \delta_2 \in (0, 1)$ denote the discount factors of Players 1 and 2, respectively. Starting from T_0, their discounted objectives are

$$J_1 = \sum_{t=0}^{\infty} \delta_1^t (ga_t - cs_t), \tag{8}$$

$$J_2 = \sum_{t=0}^{\infty} \delta_2^t a_t \big[(b + \ell)\widehat{T}_t - \ell\big]. \tag{9}$$

Because g, c, b, and ℓ are finite and both discount factors lie in $(0, 1)$, the discounted payoffs are bounded along every feasible path.

2.3 Stationary Markov strategies

The state variable T_t is intended to summarize the payoff-relevant public history. We therefore study stationary Markov strategies.

A stationary Markov strategy for Player 1 is a mapping

$$\sigma_1 : [0,1] \to \{0,1\}, \tag{10}$$

where $\sigma_1(T)$ specifies the signal chosen when the current state is T.

A stationary Markov strategy for Player 2 is a mapping

$$\sigma_2 : [0,1] \to \{0,1\}, \tag{11}$$

where $\sigma_2(\widehat{T})$ specifies the cooperation decision after observing the post-signal state $\widehat{T}$.

The equilibrium concept used below is stationary Markov perfect equilibrium: each player's strategy must be optimal at every payoff-relevant state against the other player's stationary Markov strategy.

This restriction is substantive. The paper does not claim a subgame-perfect equilibrium with unrestricted history-dependent strategies. If Player 1 were allowed to condition future signaling on Player 2's past actions, then Player 2's current action could influence future behavior indirectly, and the best-response argument below would require a different analysis.

2.4 The trustee's dynamic threshold response

The first equilibrium component can be derived exactly. Although Player 2 has an infinite-horizon objective, the current cooperation decision does not affect the next state. Hence the continuation term is identical under $a = 0$ and $a = 1$, and the dynamic comparison reduces to the current payoff difference.

Proposition 2.1 (Dynamic threshold best response)**.** *Suppose ties are resolved in favor of cooperation. For every stationary Markov strategy of Player 1 and every post-signal state $\widehat{T} \in [0,1]$, Player 2's stationary Markov best response is*

$$\sigma_2^*(\widehat{T}) = \mathbf{1}\{\widehat{T} \geq \theta\}, \qquad \theta := \frac{\ell}{b+\ell} \in (0,1). \tag{12}$$

The result is independent of δ_2.

Proof. Fix a post-signal state $\widehat{T}$. Since Player 2's current action does not enter the transition equation (5), choosing $a = 0$ or $a = 1$ leads to the same next-period state and therefore to the same continuation value under stationary Markov strategies. The difference between the total discounted values of the two current actions is consequently equal to the difference in their current payoffs.

For $a = 0$, the current payoff is zero. For $a = 1$, it is

$$(b+\ell)\widehat{T} - \ell.$$

Thus cooperation is strictly optimal when

$$(b+\ell)\widehat{T}-\ell>0,$$

noncooperation is strictly optimal when the expression is negative, and the player is indifferent at equality. Solving the equality yields

$$\widehat{T}=\frac{\ell}{b+\ell}.$$

The stated tie-breaking rule therefore gives (12). Since the continuation terms cancel, the argument does not depend on δ_2. □

Proposition 2.1 is stronger than imposing a threshold as an exogenous behavioral rule. Within the stationary Markov formulation, the threshold is the trustee's dynamically optimal response to the stated payoff function.

2.5 Reduction of the builder's problem

Once Player 2 uses Proposition 2.1, Player 1 faces a well-defined discounted control problem. For $s \in \{0,1\}$, define

$$F_s(T) := \lambda T + (1-\lambda)s. \tag{13}$$

The current payoff to Player 1 from choosing s at state T is therefore

$$r(T,s) := g\,\mathbf{1}\{F_s(T) \geq \theta\} - cs. \tag{14}$$

The associated Bellman operator acting on bounded functions $V : [0,1] \to \mathbb{R}$ is

$$(\mathcal{B}V)(T) = \max_{s\in\{0,1\}} \left\{ g\,\mathbf{1}\{F_s(T) \geq \theta\} - cs + \delta_1 V\big(F_s(T)\big) \right\}. \tag{15}$$

Equation (15) is the central optimization equation of the model. Unlike the elementary state recursion alone, it incorporates the trade-off between the immediate cost of trust-building activity, the current gain from cooperation, and the effect of the action on future states.

2.6 Existence of a stationary Markov equilibrium

Before attempting to characterize the shape of the builder's optimal policy, we first establish that the optimization problem is well posed and that an optimal stationary strategy exists.

Theorem 2.2 (Existence of a pure stationary Markov equilibrium)**.** *Assume*

$$g>0, \qquad c>0, \qquad b>0, \qquad \ell>0,$$

$$0\leq\lambda<1, \qquad \delta_1,\delta_2\in(0,1).$$

Then the Trust-Building Game admits at least one pure stationary Markov perfect equilibrium.

More precisely, Player 2 uses the threshold strategy

$$\sigma_2^*(\widehat{T}) = \mathbf{1}\{\widehat{T} \geq \theta\}, \qquad \theta = \frac{\ell}{b+\ell},$$

and Player 1 has at least one pure stationary optimal strategy $\sigma_1^* : [0,1] \to \{0,1\}$ *satisfying the Bellman equation* (15).

Proof. Let $\mathcal{B}_b([0,1])$ denote the space of bounded Borel-measurable real-valued functions on $[0,1]$, equipped with the supremum norm $\|\cdot\|_\infty$. This is a complete normed space.

For $V, W \in \mathcal{B}_b([0,1])$, the Bellman operator (15) satisfies

$$\begin{aligned} |(\mathcal{B}V)(T) - (\mathcal{B}W)(T)| &\leq \max_{s\in\{0,1\}} \delta_1 \, |V(F_s(T)) - W(F_s(T))| \\ &\leq \delta_1 \|V - W\|_\infty. \end{aligned}$$

Taking the supremum over $T \in [0,1]$ gives

$$\|\mathcal{B}V - \mathcal{B}W\|_\infty \leq \delta_1 \|V - W\|_\infty. \tag{16}$$

Since $\delta_1 \in (0,1)$, $\mathcal{B}$ is a contraction. Therefore it has a unique fixed point $V^* \in \mathcal{B}_b([0,1])$.

Because the action set $\{0,1\}$ is finite, the maximum in (15) is attained at every state. Define $\sigma_1^*(T)$ to be a maximizing action, using a fixed tie-breaking rule whenever both actions have equal value. The two action-value functions are Borel measurable, so this selector can be chosen Borel measurable.

The Bellman fixed-point equation then gives

$$V^*(T) = r(T, \sigma_1^*(T)) + \delta_1 V^*\Big(F_{\sigma_1^*(T)}(T)\Big)$$

for every $T \in [0,1]$. Standard iteration of this equality, together with boundedness of V^* and $\delta_1 < 1$, shows that σ_1^* attains the discounted value V^* and is therefore optimal within the stationary Markov control problem.

By Proposition 2.1, Player 2's threshold strategy is a best response at every post-signal state. Player 1's strategy σ_1^* is a best response to that threshold strategy at every current state. Hence (σ_1^*, σ_2^*) is a pure stationary Markov perfect equilibrium. □

Theorem 2.2 establishes equilibrium existence without claiming a particular closed-form shape for Player 1's optimal strategy. The next analytical task is therefore not to prove existence again, but to determine the geometry of the maximizing action in (15): whether the optimal policy is characterized by one threshold, multiple regions, or a more complex state-dependent pattern.

3 Dynamic Optimization of the Trust Builder

Section 2 establishes that, after the trustee's dynamic best response is substituted into the game, Player 1 faces the discounted Bellman problem in (15). The purpose of the present section is to

study the structure of that optimization problem without imposing a threshold policy a priori. In particular, we first establish basic properties of the value function, then derive exact bounds on the gain from sending a positive signal, and finally identify a parameter regime in which the optimal stationary policy can be characterized completely.

Throughout this section, let V^* denote the unique fixed point of the Bellman operator in (15), and write δ_1 for Player 1's discount factor.

3.1 Bounds and monotonicity of the value function

Before comparing the two signaling actions, we need two structural properties of the optimal value function. The first gives uniform bounds on attainable discounted payoffs. The second shows that a higher public state cannot reduce Player 1's optimal continuation value.

Lemma 3.1 (Bounds on the optimal value)**.** *For every $T \in [0,1]$,*

$$0 \le V^*(T) \le \frac{g}{1-\delta_1}. \tag{17}$$

Proof. Player 1 can always choose $s_t = 0$ in every period. Under this policy the one-period payoff is

$$g\,\mathbf{1}\{F_0(T_t) \ge \theta\} \ge 0,$$

so the value generated by this feasible policy is nonnegative. Since V^* is the optimal value, $V^*(T) \ge 0$.

For the upper bound, observe from (14) that for every state and every action,

$$r(T,s) = g\,\mathbf{1}\{F_s(T) \ge \theta\} - cs \le g.$$

Therefore the discounted payoff of every feasible policy is bounded above by

$$\sum_{t=0}^{\infty} \delta_1^t g = \frac{g}{1-\delta_1}.$$

Taking the supremum over feasible policies proves (17). □

The next result formalizes the intuitive but nontrivial fact that a larger public state cannot reduce the builder's optimal value.

Proposition 3.2 (Monotonicity of the optimal value)**.** *The optimal value function V^* is nondecreasing on $[0,1]$. Thus, for $T, T' \in [0,1]$,*

$$T \le T' \quad \Longrightarrow \quad V^*(T) \le V^*(T').$$

Proof. Let V be any bounded nondecreasing function on $[0,1]$. For each fixed $s \in \{0,1\}$, the transition

$$F_s(T) = \lambda T + (1-\lambda)s$$

is nondecreasing in T. Hence $V(F_s(T))$ is nondecreasing in T. Moreover,

$$\mathbf{1}\{F_s(T) \geq \theta\}$$

is also nondecreasing in T. Consequently, for each fixed action s, the function

$$T \longmapsto g\,\mathbf{1}\{F_s(T) \geq \theta\} - cs + \delta_1 V(F_s(T))$$

is nondecreasing.

The maximum of two nondecreasing functions is nondecreasing. Therefore the Bellman operator $\mathcal{B}$ maps bounded nondecreasing functions into bounded nondecreasing functions.

Now initialize value iteration with

$$V_0(T) \equiv 0$$

and define

$$V_{n+1} = \mathcal{B}V_n.$$

By induction, every V_n is nondecreasing. By the contraction property proved in Theorem 2.2,

$$V_n \longrightarrow V^*$$

uniformly on $[0, 1]$. The pointwise limit of nondecreasing functions is nondecreasing, so V^* has the stated property. □

3.2 The signaling gain

The optimal action at a state T is determined by the difference between the Bellman values of signaling and not signaling. Define the optimal action-value functions

$$Q_0(T) := g\,\mathbf{1}\{F_0(T) \geq \theta\} + \delta_1 V^*(F_0(T)), \tag{18}$$

$$Q_1(T) := g\,\mathbf{1}\{F_1(T) \geq \theta\} - c + \delta_1 V^*(F_1(T)). \tag{19}$$

The net gain from sending a positive signal is

$$\Delta(T) := Q_1(T) - Q_0(T). \tag{20}$$

Thus,

$$\Delta(T) = g\left[\mathbf{1}\{F_1(T) \geq \theta\} - \mathbf{1}\{F_0(T) \geq \theta\}\right] - c + \delta_1\left[V^*(F_1(T)) - V^*(F_0(T))\right]. \tag{21}$$

A positive signal is strictly optimal when $\Delta(T) > 0$, no signal is strictly optimal when $\Delta(T) < 0$, and both actions are optimal when $\Delta(T) = 0$.

Because

$$F_1(T) - F_0(T) = 1 - \lambda \geq 0,$$

Proposition 3.2 implies

$$V^*(F_1(T)) - V^*(F_0(T)) \geq 0. \tag{22}$$

Combining this observation with Lemma 3.1 gives the uniform bound

$$0 \leq \delta_1 \left[V^*(F_1(T)) - V^*(F_0(T))\right] \leq \frac{\delta_1 g}{1-\delta_1}. \tag{23}$$

Equation (23) is useful because it allows several optimality statements to be obtained without knowing V^* in closed form.

3.3 The cooperation-activation region

A positive signal has an immediate cooperation effect only when it moves the post-signal state from below the trustee's cutoff to or above that cutoff. Accordingly, define the *cooperation-activation region*

$$\mathcal{A} := \{T \in [0,1] : F_0(T) < \theta \leq F_1(T)\}. \tag{24}$$

For $0 < \lambda < 1$, this set can be written explicitly as

$$\mathcal{A} = [0,1] \cap \left[\frac{\theta + \lambda - 1}{\lambda}, \frac{\theta}{\lambda}\right). \tag{25}$$

For $\lambda = 0$, one has $F_0(T) = 0$ and $F_1(T) = 1$ for every state, and hence

$$\mathcal{A} = [0,1], \qquad \lambda = 0. \tag{26}$$

The set $\mathcal{A}$ has a precise economic interpretation. At states inside $\mathcal{A}$, signaling changes Player 2's current action from noncooperation to cooperation. Outside $\mathcal{A}$, signaling does not change the current cooperation status: either both actions fail to induce cooperation or both actions induce it.

The following proposition separates these two cases.

Proposition 3.3 (Bounds on the signaling gain)**.** *For every $T \in [0,1]$:*

(i) If $T \notin \mathcal{A}$, then

$$-c \leq \Delta(T) \leq -c + \frac{\delta_1 g}{1-\delta_1}. \tag{27}$$

(ii) If $T \in \mathcal{A}$, then

$$g - c \leq \Delta(T) \leq g - c + \frac{\delta_1 g}{1-\delta_1}. \tag{28}$$

Proof. Since $F_1(T) \geq F_0(T)$, the difference

$$\mathbf{1}\{F_1(T) \geq \theta\} - \mathbf{1}\{F_0(T) \geq \theta\}$$

can take only the values 0 and 1.

If $T \notin \mathcal{A}$, this difference is zero. Substituting into (21) and applying (23) gives (27).

If $T \in \mathcal{A}$, the indicator difference equals one. Substituting into (21) and again applying (23) gives (28). □

3.4 Parameter regimes with exact optimal actions

The preceding bounds yield optimal actions for several parameter regimes without requiring an explicit solution of the Bellman equation.

Corollary 3.4 (A high-cost no-signaling regime)**.** *If*

$$c > \frac{g}{1-\delta_1}, \tag{29}$$

then

$$\sigma_1^*(T) = 0$$

is the unique optimal action at every $T \in [0,1]$*.*

If equality holds in (29)*, no signaling is optimal at every state, although ties cannot be excluded from the present bounds.*

Proof. For $T \notin \mathcal{A}$, Proposition 3.3 gives

$$\Delta(T) \leq -c + \frac{\delta_1 g}{1-\delta_1} < 0.$$

For $T \in \mathcal{A}$,

$$\Delta(T) \leq g - c + \frac{\delta_1 g}{1-\delta_1} = \frac{g}{1-\delta_1} - c < 0.$$

Thus $Q_1(T) < Q_0(T)$ at every state.

When $c = g/(1-\delta_1)$, the same calculation yields $\Delta(T) \leq 0$, which establishes weak optimality of $s = 0$ but does not rule out equality at some states. □

The next two consequences distinguish states where signaling changes current cooperation from states where it does not.

Corollary 3.5 (No signaling outside the activation region)**.** *If*

$$c > \frac{\delta_1 g}{1-\delta_1}, \tag{30}$$

then $s = 0$ *is uniquely optimal at every* $T \notin \mathcal{A}$*.*

Proof. For $T \notin \mathcal{A}$, (27) gives

$$\Delta(T) \leq -c + \frac{\delta_1 g}{1-\delta_1} < 0.$$

□

Corollary 3.6 (Signaling inside the activation region)**.** *If*

$$c < g, \tag{31}$$

then $s = 1$ is uniquely optimal at every $T \in \mathcal{A}$.

Proof. For $T \in \mathcal{A}$, (28) gives

$$\Delta(T) \geq g - c > 0.$$

□

The two previous corollaries can be combined to obtain an exact policy characterization over a nonempty parameter regime.

Theorem 3.7 (Exact activation-band policy)**.** *Suppose*

$$\delta_1 < \frac{1}{2} \tag{32}$$

and

$$\frac{\delta_1 g}{1 - \delta_1} < c < g. \tag{33}$$

Then Player 1's unique optimal stationary action at every state is

$$\sigma_1^*(T) = \begin{cases} 1, & T \in \mathcal{A}, \\ 0, & T \notin \mathcal{A}. \end{cases} \tag{34}$$

Proof. Condition (33) implies

$$c > \frac{\delta_1 g}{1 - \delta_1},$$

so Corollary 3.5 establishes that $s = 0$ is uniquely optimal outside $\mathcal{A}$.

The same condition also gives $c < g$, so Corollary 3.6 establishes that $s = 1$ is uniquely optimal inside $\mathcal{A}$.

These two regions partition $[0, 1]$, yielding (34). The condition $\delta_1 < 1/2$ guarantees that the open interval in (33) is nonempty, since

$$\frac{\delta_1}{1 - \delta_1} < 1 \quad \Longleftrightarrow \quad \delta_1 < \frac{1}{2}.$$

□

Theorem 3.7 is useful for two reasons. First, it provides an exact stationary optimal policy without requiring an explicit closed-form solution for V^*. Second, it shows that a one-sided threshold policy cannot be assumed in general.

This point becomes explicit when the activation region lies strictly inside the state space.

Corollary 3.8 (An interior nonmonotone policy regime)**.** *Assume the conditions of Theorem 3.7, and suppose*

$$0 < \lambda < 1, \qquad 1 - \lambda < \theta < \lambda. \tag{35}$$

Then necessarily $\lambda > 1/2$, and

$$\mathcal{A} = \left[\frac{\theta + \lambda - 1}{\lambda}, \frac{\theta}{\lambda} \right) \subset (0, 1). \tag{36}$$

Consequently,

$$\sigma_1^*(T) = \begin{cases} 0, & 0 \le T < \dfrac{\theta + \lambda - 1}{\lambda}, \\ 1, & \dfrac{\theta + \lambda - 1}{\lambda} \le T < \dfrac{\theta}{\lambda}, \\ 0, & \dfrac{\theta}{\lambda} \le T \le 1. \end{cases} \tag{37}$$

Thus the optimal stationary policy in this regime is not a monotone single-threshold rule in the state variable.

Proof. The inequalities

$$1 - \lambda < \theta < \lambda$$

imply $1 - \lambda < \lambda$, hence $\lambda > 1/2$.

Moreover,

$$\theta > 1 - \lambda \quad \Longrightarrow \quad \frac{\theta + \lambda - 1}{\lambda} > 0,$$

while

$$\theta < \lambda \quad \Longrightarrow \quad \frac{\theta}{\lambda} < 1.$$

Therefore the truncation by $[0, 1]$ in (25) is inactive, and (36) follows.

Applying Theorem 3.7 to this explicit activation interval gives (37). Since the optimal action changes from 0 to 1 and then back to 0 as T increases, it cannot be represented by a monotone one-sided threshold rule. □

3.5 What is and is not established

The preceding analysis establishes several properties of the trust builder's optimization problem without imposing an unproved policy structure. The value function is bounded and nondecreasing; the exact gain from signaling is given by (21); and the activation region (24) identifies precisely where signaling changes the trustee's current cooperation decision.

Most importantly, Theorem 3.7 provides a parameter regime in which the optimal stationary policy is characterized completely. Corollary 3.8 further demonstrates that, for admissible parameters, the optimal action can have a genuine no-signal–signal–no-signal structure.

Outside the parameter restrictions used in Theorem 3.7, the present arguments do not establish a universal threshold, band, or cycling policy. In particular, when

$$c \le \frac{\delta_1 g}{1 - \delta_1},$$

the future value generated by raising the public state may be large enough to justify signaling even when it does not change current cooperation. Likewise, when $c \ge g$, the immediate gain from activating current cooperation is not, by itself, sufficient to guarantee signaling.

Accordingly, no stronger global characterization is asserted here without an additional proof. The remaining parameter regimes require either a sharper analytical characterization of the continuation-value difference

$$V^*(F_1(T)) - V^*(F_0(T))$$

or a separately verified numerical solution of the Bellman equation.

4 State Dynamics and Cooperation Timing

The equilibrium analysis in Section 3 concerns Player 1's discounted optimization problem. The present section addresses a different question. Conditional on a given admissible signal sequence, we characterize the evolution of the public state, determine the earliest possible time at which cooperation can occur, quantify finite-horizon cooperation under persistent positive signaling, and derive the rate at which the state decays when positive signaling stops.

This distinction is important. A signal sequence that minimizes the time to reach the cooperation threshold need not maximize Player 1's discounted payoff. Accordingly, the results below concern the state dynamics implied by the model and do not replace the optimal-policy analysis of Section 3.

Combining (3) and (5), the public state satisfies

$$T_{t+1} = \lambda T_t + (1-\lambda)s_t, \qquad 0 \le \lambda < 1, \qquad s_t \in \{0,1\}. \tag{38}$$

Since the trustee acts after observing the post-signal state,

$$\widehat{T}_t = T_{t+1},$$

cooperation in period t occurs if and only if

$$T_{t+1} \ge \theta. \tag{39}$$

4.1 State Invariance and History Representation

We first verify that the state remains in its intended domain and derive its exact dependence on the preceding signal history.

Lemma 4.1 (State invariance)**.** *If $T_0 \in [0,1]$, then*

$$T_t \in [0,1] \qquad \text{for every } t \ge 0 \tag{40}$$

under every admissible binary signal sequence.

Proof. Suppose $T_t \in [0,1]$. Since $s_t \in \{0,1\} \subset [0,1]$ and $0 \le \lambda < 1$, equation (38) expresses T_{t+1} as a convex combination of T_t and s_t. Hence $T_{t+1} \in [0,1]$. Induction from $T_0 \in [0,1]$ proves the result. □

The recursion can be unfolded explicitly, allowing every previous signal to be assigned its exact weight in the current state.

Proposition 4.2 (Weighted-history representation)**.** *For every integer $t \ge 1$,*

$$T_t = \lambda^t T_0 + (1-\lambda)\sum_{j=0}^{t-1} \lambda^{t-1-j} s_j. \tag{41}$$

Consequently, a signal that occurred k periods before state index t enters T_t with coefficient

$$(1-\lambda)\lambda^k. \tag{42}$$

Proof. For $t = 1$, equation (41) coincides with (38). Suppose that (41) holds for some $t \geq 1$. Then

$$\begin{aligned} T_{t+1} &= \lambda T_t + (1-\lambda)s_t \\ &= \lambda \left[\lambda^t T_0 + (1-\lambda) \sum_{j=0}^{t-1} \lambda^{t-1-j} s_j \right] + (1-\lambda)s_t \\ &= \lambda^{t+1} T_0 + (1-\lambda) \sum_{j=0}^{t} \lambda^{t-j} s_j, \end{aligned}$$

which is (41) with t replaced by $t+1$. □

The weighted-history representation also gives an immediate order comparison between any two admissible signal histories.

Corollary 4.3 (Order preservation with respect to signals)**.** *Consider two admissible signal sequences (s_j) and (s'_j) starting from the same initial state T_0. If*

$$s_j \leq s'_j \qquad \text{for every } j < t, \tag{43}$$

then their associated states satisfy

$$T_t \leq T'_t. \tag{44}$$

Proof. Subtracting the two versions of (41) yields

$$T'_t - T_t = (1-\lambda) \sum_{j=0}^{t-1} \lambda^{t-1-j} (s'_j - s_j).$$

Every coefficient and every difference on the right-hand side is nonnegative. Hence $T'_t - T_t \geq 0$. □

4.2 Threshold Crossing and the First Cooperative Period

The within-period timing established in Section 2 requires a distinction between the first state index at which the threshold is reached and the first period in which the trustee cooperates. These two quantities differ by one index because Player 2 acts after the state update.

Suppose first that

$$T_0 < \theta.$$

Define the first threshold-crossing state index by

$$\tau_T := \inf \{k \geq 1 : T_k \geq \theta\}, \tag{45}$$

with the convention $\tau_T = \infty$ when the set is empty.

Define the first cooperative period by

$$\tau_C := \inf\{t \geq 0 : a_t = 1\}. \tag{46}$$

Because cooperation in period t is determined by T_{t+1}, whenever threshold crossing occurs,

$$\tau_C = \tau_T - 1. \tag{47}$$

This distinction removes the indexing ambiguity that would arise from using a single symbol for both the state-crossing index and the cooperation period.

Theorem 4.4 (Earliest attainable threshold crossing)**.** *Let*

$$T_0 < \theta \qquad \text{and} \qquad 0 < \lambda < 1.$$

Among all admissible binary signal sequences, the persistent positive sequence

$$s_t = 1, \qquad t = 0, 1, 2, \ldots, \tag{48}$$

attains the minimum possible value of τ_T. Other signal sequences may attain the same minimum, but no admissible sequence can reach the threshold earlier.

Under (48),

$$T_t = 1 - \lambda^t(1 - T_0), \qquad t \geq 0, \tag{49}$$

and the first threshold-crossing state index is

$$\tau_T = \left\lceil \frac{\log\left(\dfrac{1-\theta}{1-T_0}\right)}{\log \lambda} \right\rceil. \tag{50}$$

Consequently, the first cooperative period is

$$\tau_C = \left\lceil \frac{\log\left(\dfrac{1-\theta}{1-T_0}\right)}{\log \lambda} \right\rceil - 1. \tag{51}$$

For $\lambda = 0$, one positive signal gives $T_1 = 1$. Hence

$$\tau_T = 1, \qquad \tau_C = 0. \tag{52}$$

Proof. By Corollary 4.3, the persistent positive signal sequence weakly dominates every other admissible binary sequence at every state index. Therefore no alternative sequence can reach $[\theta, 1]$ earlier.

Substituting $s_j = 1$ into (41) gives

$$\begin{aligned} T_t &= \lambda^t T_0 + (1-\lambda)\sum_{j=0}^{t-1} \lambda^{t-1-j} \\ &= \lambda^t T_0 + 1 - \lambda^t \\ &= 1 - \lambda^t(1 - T_0), \end{aligned}$$

which establishes (49).

The threshold condition $T_t \geq \theta$ is equivalent to

$$\lambda^t(1 - T_0) \leq 1 - \theta,$$

and therefore to

$$\lambda^t \leq \frac{1-\theta}{1-T_0}.$$

Since $0 < \lambda < 1$, one has $\log \lambda < 0$. Taking logarithms and dividing by $\log \lambda$ therefore gives

$$t \geq \frac{\log((1-\theta)/(1-T_0))}{\log \lambda}.$$

The smallest admissible integer is (50). Equation (47) then gives (51). □

The formula also yields comparative statics without requiring differentiation of the ceiling function.

Corollary 4.5 (Comparative statics of threshold acquisition)**.** *For $T_0 < \theta$ and $0 \leq \lambda < 1$, the minimum attainable τ_T is nondecreasing in λ, nondecreasing in θ, and nonincreasing in T_0. Whenever crossing occurs, the same monotonicity statements hold for τ_C.*

Proof. For every fixed integer $t \geq 1$, the persistent-positive trajectory

$$T_t = 1 - \lambda^t(1 - T_0)$$

is nonincreasing in λ and nondecreasing in T_0. Thus increasing λ cannot make a fixed state index reach the threshold when it did not already do so, whereas increasing T_0 cannot destroy an existing crossing. Increasing θ weakly strengthens the threshold requirement. Therefore the first crossing index is nondecreasing in λ and θ and nonincreasing in T_0. Equation (47) transfers the same ordering to τ_C. □

The comparative static with respect to λ has a strictly mathematical interpretation. A larger λ makes the previous state more persistent and therefore slows adjustment toward a sequence of new positive signals. No psychological interpretation of λ is required for this result.

4.3 Finite-Horizon Cooperation Under Persistent Positive Signaling

The distinction between τ_T and τ_C also determines the correct finite-horizon cooperation count.

Let $H \geq 1$ and consider the first H periods,

$$t = 0, 1, \ldots, H-1.$$

Suppose that Player 1 supplies a positive signal in each of these periods.

Corollary 4.6 (Exact finite-horizon cooperation count)**.** *Suppose* $T_0 < \theta$ *and*

$$s_t = 1, \qquad t = 0, \ldots, H-1. \tag{53}$$

Then the number of cooperative periods among the first H *periods is*

$$N_H = \max\{0, H - \tau_C\} = \max\{0, H - \tau_T + 1\}. \tag{54}$$

The corresponding cooperation fraction is

$$q_H := \frac{N_H}{H}. \tag{55}$$

Proof. If $\tau_C \geq H$, the first cooperative period lies outside the stated horizon and hence $N_H = 0$.

If $\tau_C < H$, cooperation occurs in the periods

$$\tau_C, \tau_C + 1, \ldots, H-1.$$

The number of integers in this set is

$$H - \tau_C.$$

Using $\tau_C = \tau_T - 1$ yields

$$H - \tau_C = H - \tau_T + 1,$$

which proves (54). □

This result is particularly important for the computational analysis. Numerical cooperation indicators must be evaluated from the updated state T_{t+1}, not from the pre-signal state T_t.

4.4 Decay After Positive Signaling Stops

Threshold acquisition does not imply permanent cooperation. When positive signaling stops, the public state follows a geometric decay path.

Proposition 4.7 (State decay under consecutive zero signals)**.** *Let* $0 < \lambda < 1$ *and suppose that*

$$T_t \geq \theta.$$

If Player 1 sends zero signals for the next m *periods,*

$$s_t = s_{t+1} = \cdots = s_{t+m-1} = 0, \tag{56}$$

then

$$T_{t+m} = \lambda^m T_t. \tag{57}$$

The largest number of consecutive zero-signal periods beginning from state T_t during which cooperation occurs in every period is

$$m_0(T_t) = \left\lfloor \frac{\log(\theta/T_t)}{\log \lambda} \right\rfloor. \tag{58}$$

For $\lambda = 0$ and $\theta > 0$, the first zero signal sends the updated state to zero, so no zero-signal period can sustain cooperation.

Proof. Repeated application of (38) under (56) gives

$$T_{t+j} = \lambda^j T_t, \qquad j = 1, \ldots, m.$$

Since $0 < \lambda < 1$, this sequence is nonincreasing in j. Thus cooperation occurs throughout all m zero-signal periods if and only if the last relevant updated state remains above the threshold:

$$\lambda^m T_t \geq \theta.$$

Equivalently,

$$\lambda^m \geq \frac{\theta}{T_t}.$$

Taking logarithms and using $\log \lambda < 0$ gives

$$m \leq \frac{\log(\theta/T_t)}{\log \lambda}.$$

The largest admissible nonnegative integer is therefore (58). □

The quantity $m_0(T_t)$ is purely a state-dynamic persistence measure. It is not an optimal stopping time and does not incorporate the future signaling incentives contained in the Bellman problem.

4.5 Minimum Current Cost of Preserving Cooperation

The final result concerns a deliberately narrower problem than the infinite-horizon optimization analyzed in Section 3. Suppose that the current state has already reached the threshold and that Player 1's only requirement is to preserve cooperation in the current period. We ask which of the two available signals satisfies that constraint at the smallest immediate signaling cost.

Formally, consider

$$s_t^{\text{pres}} \in \arg\min_{s \in \{0,1\}} cs \quad \text{subject to} \quad F_s(T_t) \geq \theta. \tag{59}$$

Proposition 4.8 (Minimum-cost action for one-period preservation)**.** *Suppose*

$$T_t \geq \theta \qquad \textit{and} \qquad c > 0.$$

Then the unique solution of (59) *is*

$$s_t^{\text{pres}} = \begin{cases} 0, & \lambda T_t \geq \theta, \\ 1, & \lambda T_t < \theta. \end{cases} \tag{60}$$

Proof. If $s_t = 0$, then

$$F_0(T_t) = \lambda T_t.$$

Therefore zero signaling satisfies the cooperation constraint if and only if

$$\lambda T_t \geq \theta.$$

If $s_t = 1$, then

$$F_1(T_t) = \lambda T_t + (1 - \lambda).$$

Since $T_t \geq \theta$,

$$\begin{aligned} F_1(T_t) - \theta &= \lambda(T_t - \theta) + (1 - \lambda)(1 - \theta) \\ &\geq 0. \end{aligned}$$

Thus $s_t = 1$ always satisfies the one-period cooperation constraint whenever the current state is at or above θ.

The immediate signaling cost equals zero under $s_t = 0$ and equals $c > 0$ under $s_t = 1$. Hence $s_t = 0$ is uniquely cost minimizing whenever it is feasible. When it is not feasible, $s_t = 1$ is the only feasible binary action. □

Remark 4.9. Proposition 4.8 is a constrained one-period cost-minimization result. It does not characterize Player 1's optimal stationary strategy for the discounted game. The dynamically optimal action is determined by the Bellman comparison in Section 3, which incorporates both the current payoff and the continuation value.

The results of this section provide the analytical quantities required for the subsequent computational analysis. In particular, (49), the distinction between τ_T and τ_C, the finite-horizon cooperation fraction (55), and the decay relation (57) can all be evaluated directly from the model without introducing additional behavioral assumptions.

5 Computational Analysis of Optimal Policies and State Dynamics

The analytical results of Sections 3 and 4 provide several exact characterizations, but they also define a Bellman problem whose optimal action can be evaluated numerically. The purpose of this section is not to replace the analytical results by simulation. Instead, we use deterministic numerical dynamic programming to verify the theoretically characterized policy and, in subsequent analyses, to examine parameter regimes for which the preceding sufficient conditions do not provide a complete closed-form policy characterization.

All computations below are direct evaluations of the model equations. No external data are used.

5.1 Numerical Validation of the Activation-Band Policy

We first use the Bellman equation to validate numerically the exact policy derived in Theorem 3.7 and Corollary 3.8. This provides a consistency check between the analytical characterization and an independently implemented value-iteration solution of the dynamic optimization problem.

We consider the parameter configuration

$$g = 1, \qquad c = 0.6, \qquad \delta_1 = 0.3, \qquad \lambda = 0.75, \qquad \theta = 0.5. \tag{61}$$

These values are chosen because they satisfy the exact parameter restrictions of Theorem 3.7. In particular,

$$\frac{\delta_1 g}{1-\delta_1} = \frac{0.3}{0.7} \approx 0.428571 < 0.6 < 1 = g, \tag{62}$$

and

$$1 - \lambda = 0.25 < \theta = 0.5 < 0.75 = \lambda. \tag{63}$$

Hence Corollary 3.8 predicts an interior activation region of the form

$$\mathcal{A} = \left[\frac{\theta + \lambda - 1}{\lambda}, \frac{\theta}{\lambda}\right) = \left[\frac{1}{3}, \frac{2}{3}\right), \tag{64}$$

and therefore the analytical optimal policy is

$$\sigma_1^*(T) = \begin{cases} 0, & 0 \le T < 1/3, \\ 1, & 1/3 \le T < 2/3, \\ 0, & 2/3 \le T \le 1. \end{cases} \tag{65}$$

5.1.1 Numerical procedure

The Bellman equation

$$V(T) = \max_{s \in \{0,1\}} \{g\,\mathbf{1}\{F_s(T) \ge \theta\} - cs + \delta_1 V(F_s(T))\} \tag{66}$$

was solved by value iteration on a uniform grid of 8001 state points over $[0, 1]$. Continuation values at off-grid transition states were evaluated by linear interpolation.

Iteration was terminated when the maximum absolute difference between successive value iterates was below 10^{-12}. The computation converged in 24 iterations, with final iteration error

$$9.415 \times 10^{-13}. \tag{67}$$

The resulting maximum Bellman residual was

$$2.824 \times 10^{-13}. \tag{68}$$

Because the trustee cooperates at equality,

$$F_s(T) = \theta,$$

the numerical implementation preserves the tie-breaking convention of Proposition 2.1. A numerical tolerance was used only to prevent floating-point representation error from incorrectly classifying states lying at the cooperation boundary.

For the builder, the computed action-value difference is

$$\Delta(T) = Q_1(T) - Q_0(T), \tag{69}$$

where

$$Q_0(T) = g\,\mathbf{1}\{F_0(T) \geq \theta\} + \delta_1 V^*(F_0(T)), \tag{70}$$

$$Q_1(T) = g\,\mathbf{1}\{F_1(T) \geq \theta\} - c + \delta_1 V^*(F_1(T)). \tag{71}$$

Thus the numerical policy selects $s = 1$ when $\Delta(T) > 0$ and $s = 0$ when $\Delta(T) < 0$.

5.1.2 Agreement between the analytical and numerical policies

Figure 3 provides several internal consistency checks. Panel (A) shows that the computed value function is nondecreasing in the state, in agreement with Proposition 3.2. The value function need not be smooth, because the current payoff contains the discontinuous cooperation indicator $\mathbf{1}\{F_s(T) \geq \theta\}$.

Panel (B) gives a more direct characterization of the signaling decision. The computed signaling advantage is positive precisely within the analytical activation interval

$$\left[\frac{1}{3}, \frac{2}{3}\right)$$

and negative outside it. Thus the numerical Bellman comparison reproduces the sign pattern required by the analytical policy.

The strongest validation is shown in Panel (C). Across all 8001 evaluated states, the numerical policy and the exact policy in (65) produce

$$\text{number of policy mismatches} = 0. \tag{72}$$

No numerical action ties were detected on the grid.

The result is therefore not merely a qualitative resemblance between the analytical and computational solutions. For the parameter configuration (61), the independently evaluated Bellman policy coincides with the analytically characterized activation-band policy at every

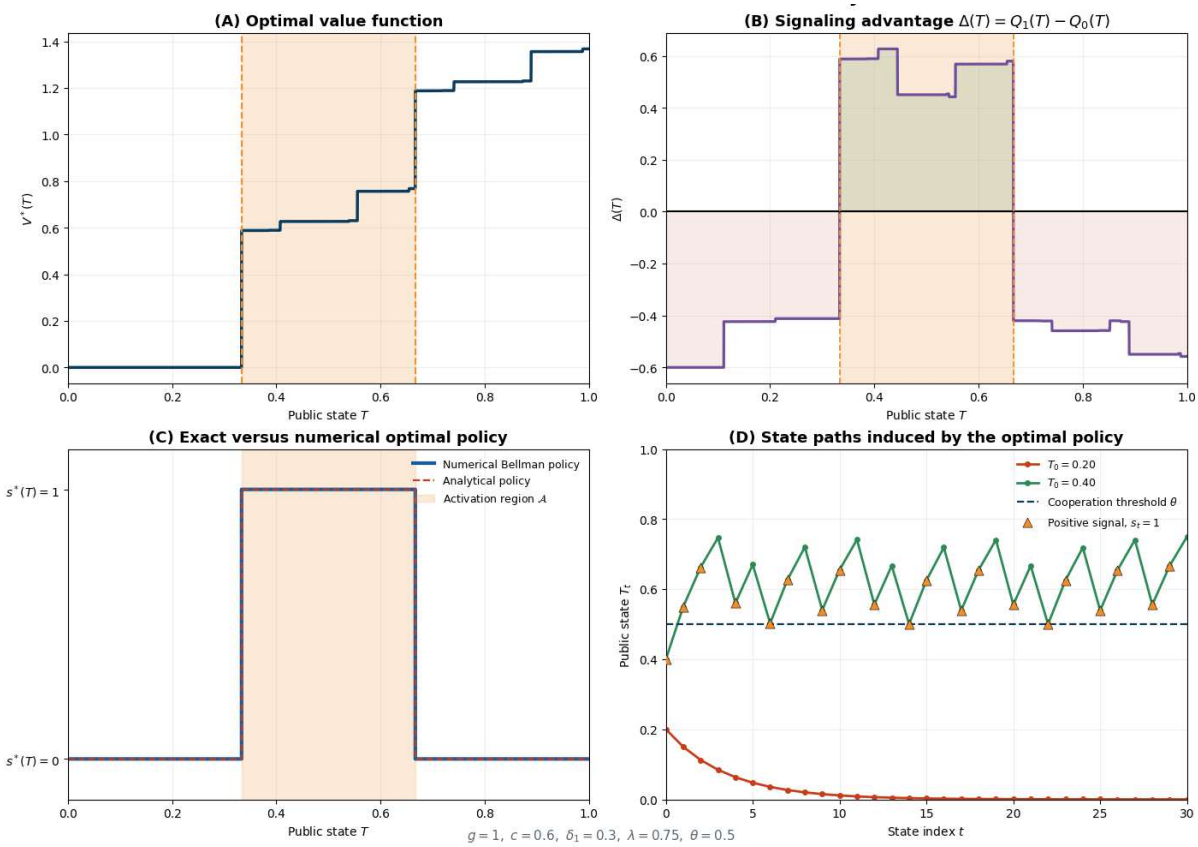


Figure 3: Numerical validation of the activation-band policy for $g = 1$, $c = 0.6$, $\delta_1 = 0.3$, $\lambda = 0.75$, and $\theta = 0.5$. Panel (A) shows the converged optimal value function $V^*(T)$. Panel (B) reports the signaling advantage $\Delta(T) = Q_1(T) - Q_0(T)$; signaling is optimal precisely where this quantity is positive. Panel (C) compares the numerical Bellman policy with the analytical policy derived in Theorem 3.7 and Corollary 3.8. The two policies coincide at all 8001 evaluated state points. Panel (D) shows state paths generated by the computed optimal policy from two initial states. Starting from $T_0 = 0.20$, no positive signal is selected and no cooperation occurs over the displayed 30-period horizon. Starting from $T_0 = 0.40$, the optimal policy selects 20 positive signals while cooperation occurs in all 30 periods. The dashed horizontal line denotes the cooperation threshold $\theta = 0.5$, and triangular markers indicate state indices at which $s_t = 1$.

state represented on the numerical grid.

5.1.3 State paths under the optimal policy

Panel (D) illustrates why the activation-band policy should not be interpreted as either "always signal" or "never signal." When the initial state is

$$T_0 = 0.20 < \frac{1}{3},$$

the computed optimal action remains $s_t = 0$. Over the displayed 30 periods, the calculation therefore produces

$$\sum_{t=0}^{29} s_t = 0, \qquad \sum_{t=0}^{29} a_t = 0. \tag{73}$$

By contrast, when

$$T_0 = 0.40 \in \mathcal{A},$$

the optimal policy generates repeated movement between signaling and non-signaling states. Over the same 30-period horizon, the numerical path contains

$$\sum_{t=0}^{29} s_t = 20, \qquad \sum_{t=0}^{29} a_t = 30. \tag{74}$$

This trajectory demonstrates an important distinction between the signaling decision and the cooperation decision. A positive signal need not be supplied in every period in which cooperation occurs. Once the state is sufficiently high, the persistence term can permit temporary signal omission while the post-signal state remains at or above the trustee's cooperation threshold. This behavior is generated by the Bellman-optimal policy for the stated parameters and is distinct from the persistent-positive signaling path studied in Section 4.

The present computation serves as a validation exercise for a parameter regime whose policy has already been characterized analytically. It does not establish that the same activation-band structure holds outside the conditions of Theorem 3.7. Parameter regimes not covered by that theorem must therefore be examined separately rather than inferred from Figure 3.

5.2 Optimal Signaling Beyond the Analytically Characterized Regime

The preceding computation validates the analytical activation-band result in a parameter regime for which the optimal policy is known exactly. We now examine the Bellman problem outside that restricted regime. The purpose of this calculation is exploratory but fully model-based: no threshold, band, or monotonicity structure is imposed on the builder's optimal policy before the Bellman equation is solved.

We normalize

$$g = 1 \tag{75}$$

and hold the state-dynamic parameters fixed at

$$\lambda = 0.75, \qquad \theta = 0.5. \tag{76}$$

For these values, the immediate cooperation-activation region defined in (24) has boundaries

$$L = \frac{\theta + \lambda - 1}{\lambda} = \frac{1}{3}, \qquad U = \frac{\theta}{\lambda} = \frac{2}{3}. \tag{77}$$

The Bellman equation was solved independently over a two-dimensional parameter grid consisting of 37 discount factors,

$$\delta_1 \in [0.05, 0.95],$$

and 61 cost-benefit ratios,

$$c/g \in [0.05, 1.40].$$

The state space was discretized using 1201 equally spaced points on $[0, 1]$. In total, the computation therefore solved

$$37 \times 61 = 2257 \tag{78}$$

separate Bellman problems.

For each parameter pair, a state was classified as a strict signaling state when

$$\Delta(T) = Q_1(T) - Q_0(T) > 0,$$

as a strict no-signaling state when $\Delta(T) < 0$, and as a numerical tie when $|\Delta(T)|$ was below the numerical policy tolerance. The maximum Bellman residual over the entire parameter scan was

$$9.492 \times 10^{-11}. \tag{79}$$

Across all state-parameter combinations, 266 state-grid points were classified as numerical ties.

Figure 4 shows that the region in which signaling is strictly optimal changes substantially with both the discount factor and the signal cost. In particular, increasing the discount factor can make positive signaling optimal at lower public states because the builder places greater weight on the continuation value generated by a higher future state.

This effect is visible directly in the selected numerical boundaries. For a cost-benefit ratio close to

$$c/g = 0.59,$$

the lowest state with strictly optimal signaling is approximately

$$0.3333$$

when $\delta_1 = 0.30$,

$$0.1117$$

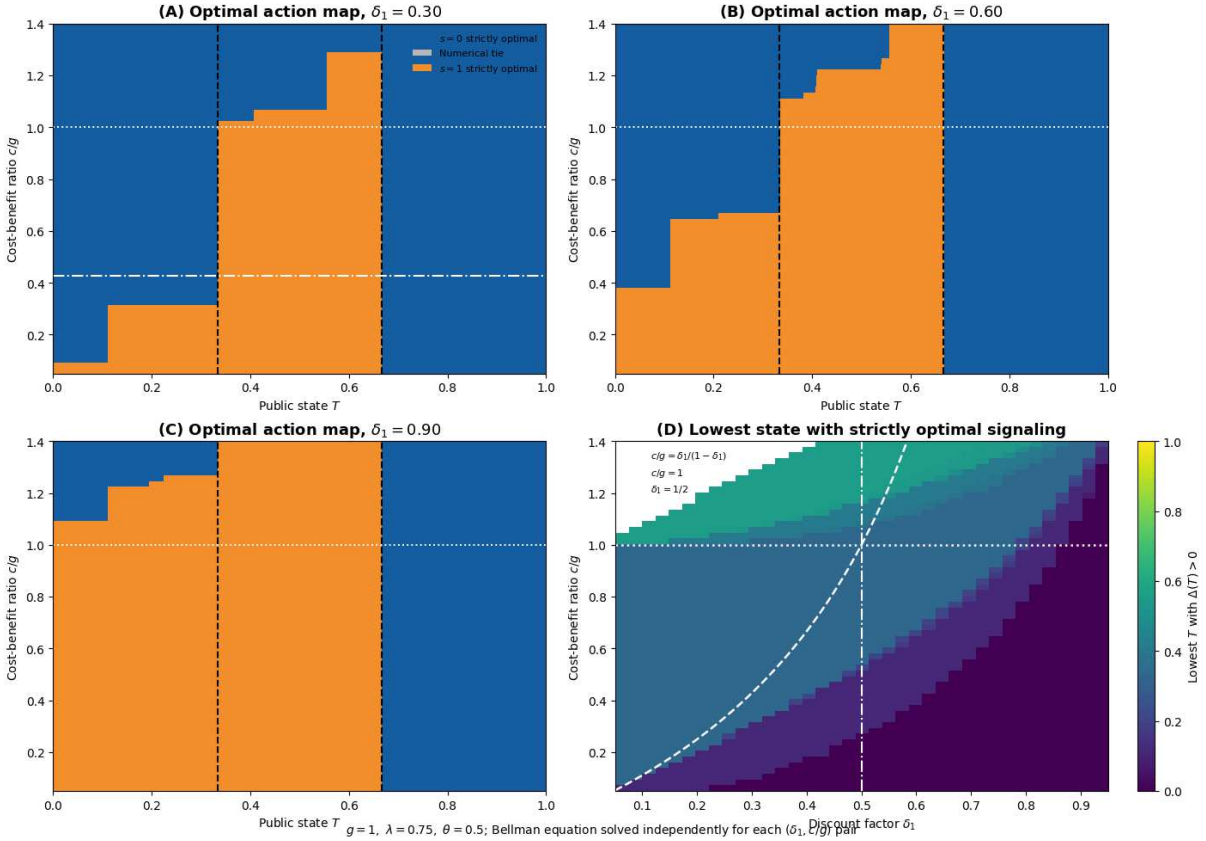


Figure 4: Numerical structure of the optimal signaling policy outside the analytically characterized activation-band regime, with $g = 1$, $\lambda = 0.75$, and $\theta = 0.5$. Panels (A)–(C) show the optimal Bellman action over the public state T and the cost-benefit ratio c/g for $\delta_1 = 0.30$, 0.60, and 0.90, respectively. Blue denotes states at which $s = 0$ is strictly optimal, orange denotes states at which $s = 1$ is strictly optimal, and gray denotes numerical ties. The vertical dashed lines mark the immediate activation-region boundaries $L = 1/3$ and $U = 2/3$. The horizontal dotted line marks $c/g = 1$, while the dash-dotted boundary, where visible, corresponds to $c/g = \delta_1/(1 - \delta_1)$. Panel (D) reports, for every scanned pair $(\delta_1, c/g)$ for which strict signaling occurs, the lowest state at which $\Delta(T) > 0$. The numerical results are descriptive of the scanned parameter grid and are not asserted as a global analytical characterization.

when $\delta_1 = 0.60$, and

$$0$$

when $\delta_1 = 0.90$. Thus, for this cost level, increasing patience enlarges the signaling region toward lower trust states.

The same computation also shows the opposite effect of higher signaling cost. For example, at $\delta_1 = 0.30$, increasing the sampled cost-benefit ratio from approximately 0.297 to 1.198 moves the lowest state with strictly optimal signaling from approximately

$$0.1117$$

to

$$0.5558.$$

At $\delta_1 = 0.60$, the corresponding lower boundary moves from

$$0$$

at $c/g \approx 0.297$ to approximately

$$0.4075$$

at $c/g \approx 1.198$.

A further numerical regularity appears in the topology of the strict-signaling set. Among the 2257 parameter combinations examined, 126 contain no state at which signaling is strictly optimal. For the remaining 2131 parameter combinations, the strict-signaling states form exactly one connected component on the 1201-point state grid. No scanned parameter pair produced two or more disconnected strict-signaling components.

This finding is potentially informative, but it must be interpreted with care. The computation does not prove that the optimal signaling set is always an interval in the continuous-state model. It establishes only that no disconnected signaling set was detected over the stated parameter and state grids. A global interval or threshold theorem would require an additional analytical argument about the shape of

$$\Delta(T) = Q_1(T) - Q_0(T),$$

which is not supplied by the numerical scan alone.

The numerical results therefore extend the analytical discussion in a limited but useful way. Theorem 3.7 identifies a parameter regime in which the signaling region is known exactly. Figure 4 shows that outside that regime the lower boundary of the strict-signaling region can move substantially with both c/g and δ_1. In particular, future-state incentives can make signaling optimal even at states below the one-period activation boundary $L = 1/3$. Conversely, sufficiently high signaling costs can shrink the signaling region toward states closer to the cooperation threshold.

These observations concern the Bellman-optimal action itself rather than the persistent-positive signal path considered in Section 4. They therefore provide a numerical complement to the analytical distinction between immediate cooperation activation and intertemporal trust-

building incentives.

5.3 State Persistence, Cooperation Timing, and Finite-Horizon Cooperation

We next evaluate the exact state-dynamic results derived in Section 4. Unlike the Bellman computations in the preceding subsections, the calculations below do not solve an optimization problem. They evaluate the closed-form state trajectory and cooperation timing under persistent positive signaling,

$$s_t = 1 \qquad \text{for all } t. \tag{80}$$

The purpose of this analysis is threefold. First, it illustrates the role of the state-persistence parameter λ in the speed of state accumulation. Second, it implements the trustee's decision using the correct within-period timing,

$$a_t = \mathbf{1}\{T_{t+1} \geq \theta\}, \tag{81}$$

rather than evaluating cooperation from the pre-signal state T_t. Third, it evaluates finite-horizon cooperation using the exact distinction between the first threshold-state index τ_T and the first cooperative period $\tau_C = \tau_T - 1$.

Throughout Panels (A), (B), and (D) of Figure 5, we use

$$T_0 = 0, \qquad \theta = 0.5, \tag{82}$$

and compare the persistence levels

$$\lambda \in \{0.3, 0.6, 0.9\}. \tag{83}$$

Under persistent positive signaling, Theorem 4.4 gives

$$T_k = 1 - \lambda^k (1 - T_0). \tag{84}$$

For $T_0 = 0$, this simplifies to

$$T_k = 1 - \lambda^k. \tag{85}$$

The exact timing calculations for the three displayed persistence levels are

λ	τ_T	τ_C	q_{30}
0.3	1	0	1.000000
0.6	2	1	0.966667
0.9	7	6	0.800000

(86)

Thus, increasing λ slows adjustment toward the persistent sequence of positive signals. This is a consequence of greater state persistence, rather than an empirical claim about psychological memory.

Panel (A) of Figure 5 visualizes the state trajectories in (85). For $\lambda = 0.3$, the first positive

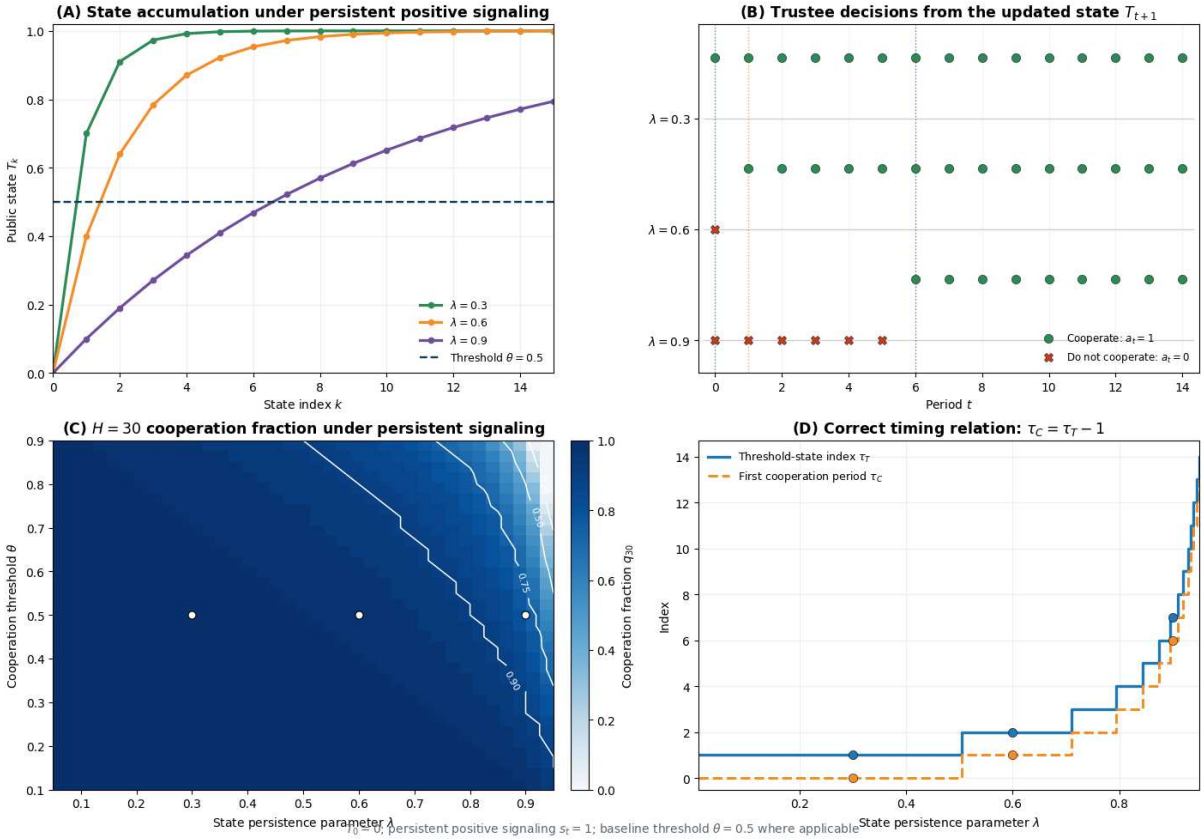


Figure 5: State persistence, cooperation timing, and finite-horizon cooperation under persistent positive signaling. Panel (A) shows the public-state trajectories generated by $T_k = 1-\lambda^k$ for $\lambda = 0.3$, 0.6, and 0.9, with $T_0 = 0$ and cooperation threshold $\theta = 0.5$. Panel (B) displays the corresponding trustee decisions using the updated state T_{t+1}, as required by the within-period timing of the model. Cooperation therefore begins in periods 0, 1, and 6 for $\lambda = 0.3$, 0.6, and 0.9, respectively. Panel (C) reports the exact 30-period cooperation fraction q_{30} over combinations of the state-persistence parameter λ and the cooperation threshold θ. Panel (D) distinguishes the first threshold-state index τ_T from the first cooperative period τ_C and displays their exact relation $\tau_C = \tau_T - 1$. The calculations use the analytical first-passage formulas derived in Section 4; they are not stochastic simulations or empirical estimates.

signal moves the state directly to

$$T_1 = 1 - 0.3 = 0.7,$$

which already exceeds the threshold $\theta = 0.5$. For $\lambda = 0.6$, the first two updated states are

$$T_1 = 0.4, \qquad T_2 = 0.64,$$

so threshold crossing occurs at state index $\tau_T = 2$. For $\lambda = 0.9$, the state evolves more slowly and first reaches the threshold at

$$\tau_T = 7.$$

These values agree exactly with the analytical expression

$$\tau_T = \left\lceil \frac{\log((1-\theta)/(1-T_0))}{\log \lambda} \right\rceil, \qquad 0 < \lambda < 1, \tag{87}$$

derived in Theorem 4.4.

5.3.1 Correct timing of the trustee's decision

Panel (B) implements the trustee's action according to (81). Because Player 2 acts after observing the updated state, the first cooperative period is one index earlier than the first threshold-state index:

$$\tau_C = \tau_T - 1. \tag{88}$$

Thus, for the three persistence levels considered above,

$$(\tau_T, \tau_C) = (1, 0),\ (2, 1),\ (7, 6) \tag{89}$$

for $\lambda = 0.3$, 0.6, and 0.9, respectively.

In particular, when $\lambda = 0.3$, cooperation occurs already in period $t = 0$ because Player 2 observes

$$T_1 = 0.7 > \theta.$$

When $\lambda = 0.6$, the updated state in period 0 remains below the threshold, while the updated state in period 1 exceeds it. Hence the first cooperative period is $t = 1$. For $\lambda = 0.9$, the first cooperative period is $t = 6$ because Player 2 then observes $T_7 \geq \theta$.

Panel (D) shows this indexing relation directly across a range of persistence values. The two step functions remain separated by exactly one index wherever threshold crossing occurs, as required by (88).

5.3.2 Finite-horizon cooperation fraction

For a horizon of H periods under persistent positive signaling, Corollary 4.6 gives

$$N_H = \max\{0, H - \tau_T + 1\}, \tag{90}$$

and therefore

$$q_H = \frac{\max\{0, H - \tau_T + 1\}}{H}. \tag{91}$$

For the 30-period horizon used in Panel (C),

$$q_{30} = \frac{\max\{0, 31 - \tau_T\}}{30}. \tag{92}$$

Substituting the exact crossing indices gives

$$q_{30} = 1, \qquad \lambda = 0.3, \tag{93}$$

$$q_{30} = \frac{29}{30} \approx 0.966667, \qquad \lambda = 0.6, \tag{94}$$

$$q_{30} = \frac{24}{30} = 0.8, \qquad \lambda = 0.9. \tag{95}$$

Panel (C) extends this calculation over the displayed (λ, θ) grid. The resulting pattern follows directly from the first-passage formula. Increasing λ slows adjustment toward the persistent positive-signal state, while increasing θ raises the state required for cooperation. Either change can therefore delay the first cooperative period and reduce the cooperation fraction observed over a fixed finite horizon.

The heatmap should consequently be interpreted as a deterministic evaluation of the exact model equations, not as a Monte Carlo simulation or an empirical estimate.

5.3.3 Boundary equality and numerical implementation

The computation also preserves the tie-breaking convention specified in Proposition 2.1. In particular, cooperation occurs whenever

$$T_{t+1} \geq \theta, \tag{96}$$

including the equality case.

A useful numerical boundary check is obtained by setting

$$\lambda = 0.9, \qquad \theta = 0.1, \qquad T_0 = 0.$$

The exact model gives

$$T_1 = 1 - \lambda = 0.1 = \theta. \tag{97}$$

Therefore

$$\tau_T = 1, \qquad \tau_C = 0, \qquad q_{30} = 1. \tag{98}$$

In floating-point arithmetic, the same theoretical quantity may be represented numerically as a value such as 0.09999999999999998. The implementation therefore uses a numerical tolerance only when evaluating threshold equality. This prevents machine representation error from changing the mathematically specified tie-breaking rule; it does not modify the underlying model.

Taken together, Figure 5 and equations (87)–(98) resolve the computational timing issues associated with the state recursion. The trustee responds to the updated state T_{t+1}, the threshold-state index τ_T is kept distinct from the cooperative-period index τ_C, and λ is interpreted consistently as a state-persistence parameter. The finite-horizon cooperation fractions are consequently aligned with the analytical results of Section 4.

6 Discussion

The analysis developed in this paper separates three objects that are often conflated in discussions of repeated trust: the evolution of a public state, the trustee's cooperation decision, and the trust builder's intertemporal signaling decision. In the present model, these objects are related but distinct. The public state evolves according to

$$T_{t+1} = \lambda T_t + (1 - \lambda)s_t,$$

the trustee responds to the updated state through

$$a_t = \mathbf{1}\{T_{t+1} \geq \theta\},$$

and the builder chooses the signal by comparing current signaling cost with both the current cooperation benefit and the continuation value generated by the resulting future state. This separation is important because a state can increase without making signaling dynamically optimal, and cooperation can persist even when no positive signal is supplied in the current period.

The trustee's decision has a particularly simple dynamic structure. Because Player 2's action does not enter the transition equation, cooperation and noncooperation lead to the same continuation state. Consequently, the continuation component cancels from the current action comparison and Proposition 2.1 yields the cutoff

$$\theta = \frac{\ell}{b + \ell}.$$

Thus, within the stationary Markov formulation, the trustee's threshold rule is derived from the stated payoff specification rather than imposed as an independent behavioral assumption. The trustee's discount factor does not affect this cutoff because the current action has no effect on future states.

The builder's problem is fundamentally different. Player 1 controls the state transition and therefore faces an intertemporal trade-off. A positive signal has an immediate cost c, may change the trustee's current cooperation decision, and also raises the state from $F_0(T)$ to $F_1(T)$. The relevant comparison is therefore

$$\Delta(T) = Q_1(T) - Q_0(T),$$

where

$$\Delta(T) = g\left[\mathbf{1}\{F_1(T) \geq \theta\} - \mathbf{1}\{F_0(T) \geq \theta\}\right] - c + \delta_1\left[V^*(F_1(T)) - V^*(F_0(T))\right].$$

The first part of this expression captures the immediate effect of signaling on cooperation, whereas the final term captures the value of improving the future public state. This decomposition explains why current cooperation alone is insufficient to determine the optimal signal.

Theorem 3.7 identifies a parameter regime in which this trade-off can be characterized exactly. When

$$\delta_1 < \frac{1}{2}$$

and

$$\frac{\delta_1 g}{1-\delta_1} < c < g,$$

the builder signals precisely on the cooperation-activation region $\mathcal{A}$. Under the additional conditions of Corollary 3.8, this region lies strictly inside the state space, so the unique optimal stationary policy has the structure

$$0 \longrightarrow 1 \longrightarrow 0.$$

The policy is therefore not a monotone one-sided threshold rule.

The economic logic of this activation band follows directly from the interaction between cost, threshold activation, and continuation value. At sufficiently low states, a positive signal can be too costly relative to the current and future benefit it generates. At intermediate states, signaling is valuable because it moves the updated state from below the cooperation cutoff to or above it. At sufficiently high states, cooperation can already be obtained without paying the current signaling cost. Signaling is therefore concentrated in an intermediate region rather than necessarily increasing monotonically with the public state.

This result also shows why a high public state does not necessarily imply a stronger incentive to signal. A higher state increases the builder's value function, as established by Proposition 3.2, but it can simultaneously reduce the marginal value of an additional positive signal. In particular, once $F_0(T)$ already exceeds the trustee's threshold, signaling no longer creates an immediate cooperation gain. Its value then depends entirely on the continuation effect relative to the signaling cost.

The computational results provide an independent check of this analytical structure. For

$$g = 1, \qquad c = 0.6, \qquad \delta_1 = 0.3, \qquad \lambda = 0.75, \qquad \theta = 0.5,$$

the analytical activation region is

$$\mathcal{A} = \left[\frac{1}{3}, \frac{2}{3}\right).$$

As shown in Figure 3, the numerically computed Bellman policy coincides with this analytical policy at all 8001 evaluated state points. The zero-mismatch result is therefore a consistency check between the theoretical characterization and the numerical implementation rather than

evidence for an additional theorem.

The optimal trajectory in Figure 3 also illustrates the distinction between signaling and cooperation. Starting from $T_0 = 0.40$, the computed policy produces cooperation in all 30 displayed periods while using only 20 positive signals. Thus a positive signal is not required in every period in which the trustee cooperates. State persistence can allow cooperation to survive temporary signal omission, while the builder avoids unnecessary current signaling costs.

The broader parameter scan in Figure 4 provides additional information about regimes not covered by Theorem 3.7. Holding

$$g = 1, \qquad \lambda = 0.75, \qquad \theta = 0.5$$

fixed, the numerical solutions show that greater patience can make signaling optimal at lower public states. For example, at a cost-benefit ratio near $c/g = 0.59$, the lowest state with strictly optimal signaling moves from approximately 0.3333 at $\delta_1 = 0.30$ to 0.1117 at $\delta_1 = 0.60$, and to 0 at $\delta_1 = 0.90$. This pattern is consistent with the Bellman equation because increasing δ_1 raises the relative importance of future state improvements.

The same scan shows the opposite effect of higher signaling cost. At $\delta_1 = 0.30$, increasing the sampled cost-benefit ratio from approximately 0.297 to 1.198 moves the lower boundary of strict signaling from approximately 0.1117 to 0.5558. Thus the numerical results support the basic economic interpretation of the Bellman problem: patience strengthens the importance of continuation value, whereas signaling cost works against positive state investment.

A further numerical regularity is that no disconnected strict-signaling region was detected over the scanned parameter grid. Among the 2257 parameter combinations considered, 126 produced no strict-signaling state, while the remaining 2131 produced one connected strict-signaling component on the 1201-point state grid. This observation is potentially informative, but it should not be elevated to an analytical claim. The computation does not prove that the strict-signaling set is always an interval in the continuous-state model. A global connectedness theorem would require additional control over the shape of

$$V^*(F_1(T)) - V^*(F_0(T))$$

and therefore over $\Delta(T)$ itself.

The state-dynamic results answer a different question from the Bellman optimization. Under persistent positive signaling, Theorem 4.4 shows that

$$T_t = 1 - \lambda^t(1 - T_0)$$

and that this path reaches the cooperation threshold as early as any admissible binary signal sequence. This result concerns the fastest attainable movement of the public state. It does not imply that signaling in every period is optimal for Player 1, because the builder's objective includes signaling costs.

The distinction becomes especially important when interpreting λ. A larger λ assigns more weight to the previous state and less weight to the current signal. Starting below the cooper-

ation threshold, this slows adjustment toward a persistent sequence of positive signals. At the same time, once the state is high, greater persistence can delay decay when positive signaling stops. The parameter therefore governs persistence of the state in both directions. It should not be interpreted as a directly observed psychological memory parameter without empirical calibration.

The timing analysis further shows that the state-crossing index and the first cooperative period must be kept distinct. Because Player 2 acts after observing the updated state,

$$a_t = \mathbf{1}\{T_{t+1} \geq \theta\},$$

and hence

$$\tau_C = \tau_T - 1.$$

For the parameter values displayed in Figure 5, the threshold-state indices

$$\tau_T = 1,\ 2,\ 7$$

for $\lambda = 0.3$, 0.6, and 0.9 correspond to first cooperative periods

$$\tau_C = 0,\ 1,\ 6.$$

This indexing distinction is not cosmetic: it directly affects the finite-horizon cooperation count in Corollary 4.6.

The one-period preservation result should also be distinguished from the infinite-horizon signaling problem. Proposition 4.8 asks for the least costly current action subject to maintaining cooperation for one period. If

$$\lambda T_t \geq \theta,$$

zero signaling is sufficient for that immediate objective. The Bellman problem, however, may still assign positive value to signaling because a larger future state can improve subsequent payoffs. Conversely, Player 1 may choose not to preserve cooperation if its discounted value is lower than the required signaling cost. The preservation rule is therefore a constrained static comparison embedded within, but not equivalent to, the dynamic optimization problem.

The public state itself should be interpreted narrowly. Although the trustee's payoff is affine in T, the model does not treat T as a Bayesian probability of a hidden type. Classical reputation models based on persistent private information therefore provide a conceptual comparison rather than the probabilistic foundation of the present framework.[26, 27] The public state is instead a reduced-form score summarizing observable past signals.

Similarly, the model does not claim that observed cooperation is a pure measure of psychological trust. Experimental behavior can reflect reciprocity, intentions, fairness, guilt, betrayal aversion, expectations, and contextual features in addition to trust itself.[9, 10, 12, 15] The binary signaling mechanism and affine trustee payoff intentionally abstract from these channels. Their purpose is to isolate the mathematical interaction between observable signaling, public-state persistence, threshold cooperation, and intertemporal incentives.

Overall, the results indicate that even a deliberately reduced-form dynamic game can generate a nontrivial separation between trust-state accumulation, cooperation activation, state persistence, and optimal signaling. The main analytical contribution is therefore not a universal claim that trust building follows a single threshold rule. Rather, the model identifies explicit conditions under which an activation-band policy is optimal, demonstrates that a monotone one-sided threshold need not describe optimal signaling, and shows through deterministic Bellman calculations how the signaling region can change outside the analytically characterized regime.

7 Limitations and Extensions

The results of this paper are obtained within a deliberately reduced-form dynamic game. This structure makes several analytical statements possible in closed form, but it also imposes restrictions on the interpretation and scope of the conclusions. The limitations described below therefore identify both the boundaries of the present results and natural directions for extending the model.

First, the public state

$$T_t \in [0, 1]$$

is a deterministic summary of past observable signals. Its evolution,

$$T_{t+1} = \lambda T_t + (1 - \lambda) s_t,$$

is specified directly rather than derived from Bayesian learning, statistical filtering, or an empirically estimated behavioral process. Accordingly, T_t should not be interpreted as a posterior probability that Player 1 has a particular hidden type, nor as a directly measured psychological level of trust. The analysis establishes the implications of this state equation conditional on the model, not the empirical validity of the updating rule itself.

Second, the signal space is binary. The action $s_t = 1$ represents observable behavior classified by the model as favorable evidence, whereas $s_t = 0$ represents the absence of such favorable input. This specification intentionally does not distinguish between several conceptually different events, such as neutral inaction, failed positive behavior, adverse evidence, and deliberate betrayal. In particular, a zero signal produces passive geometric decay rather than an explicitly modeled negative shock. A richer formulation could introduce a three-state or continuous signal space, for example

$$s_t \in \{-1, 0, 1\},$$

allowing favorable, neutral, and adverse observations to have different effects on the public state.

Third, the favorable interpretation of $s_t = 1$ is exogenous. Once a positive signal is observed, it mechanically increases the public state according to the transition equation. The trustee does not question whether the signal is truthful, manipulated, strategically misleading, or costly to verify. The model therefore abstracts from endogenous credibility. Introducing private types, hidden actions, noisy observation, verification costs, or signal precision would turn credibility itself into a strategic object and could substantially alter both the trustee's response and the

builder's optimal policy.

Fourth, the trustee's payoff is intentionally affine in the public state:

$$u_2(\widehat{T}, a) = a\big[(b+\ell)\widehat{T} - \ell\big].$$

This specification makes the cooperation cutoff transparent and yields

$$\theta = \frac{\ell}{b+\ell}.$$

The resulting threshold is therefore a structural consequence of the assumed reduced-form payoff. The paper does not claim that all trust-related decisions must have an affine payoff representation or a sharp deterministic cutoff. Nonlinear payoff functions, heterogeneous thresholds, probabilistic response rules, or smooth choice models could produce different cooperation behavior.

Fifth, the equilibrium analysis is restricted to stationary Markov strategies. The public state is treated as the payoff-relevant state variable, and future strategies are not allowed to depend separately on the full history once the current state is given. This restriction is important for the trustee's dynamic best-response argument. Under the stationary Markov formulation, Player 2's current action does not affect the physical state transition, so the continuation term cancels from the comparison between cooperation and noncooperation.

With unrestricted history-dependent strategies, however, Player 1 could condition future signaling on Player 2's past actions. In such a formulation, the trustee's present action could influence future behavior indirectly even though it does not enter the transition equation mechanically. The present analysis therefore does not characterize the full set of history-dependent subgame-perfect equilibria. Extending the game to unrestricted public histories would require a different equilibrium analysis and could permit reward, punishment, or reputation mechanisms that are absent from the current model.

Sixth, the exact activation-band result is established only under explicit sufficient parameter conditions. Theorem 3.7 proves the policy

$$\sigma_1^*(T) = \begin{cases} 1, & T \in \mathcal{A}, \\ 0, & T \notin \mathcal{A}, \end{cases}$$

when

$$\delta_1 < \frac{1}{2}$$

and

$$\frac{\delta_1 g}{1-\delta_1} < c < g.$$

These conditions are sufficient for the exact result; they are not claimed to be necessary. Outside this regime, the continuation-value difference

$$V^*(F_1(T)) - V^*(F_0(T))$$

can become decisive, and the analytical bounds used in the theorem are not sufficient to determine the optimal action globally.

The numerical parameter scan partially explores these remaining cases, but it does not replace an analytical characterization. In particular, the observation that all nonempty strict-signaling sets detected in the scan were connected does not prove that the signaling set must always be an interval in the continuous-state model. Establishing connectedness, monotonicity of the signaling boundaries, or necessary and sufficient conditions for alternative policy geometries would require additional theoretical analysis of $\Delta(T)$.

Seventh, the computational results are model evaluations rather than empirical evidence. The value-iteration calculations verify the analytical activation-band policy for one parameter configuration and explore a finite grid of additional values. The state-dynamic calculations similarly evaluate the exact deterministic formulas derived in Section 4. No human-subject, institutional, market, or experimental dataset is used to calibrate or validate the parameters.

The parameter scan itself is finite. It uses

$$37$$

values of δ_1,

$$61$$

values of c/g, and

$$1201$$

state-grid points for fixed

$$g = 1, \qquad \lambda = 0.75, \qquad \theta = 0.5.$$

Consequently, numerical regularities observed in Figure 4 should be interpreted only over the stated grid. They do not establish global comparative statics over the entire admissible parameter space.

Eighth, empirical implementation of the trustee's payoff parameters presents an identification problem. The trustee's binary decision depends on b and ℓ only through

$$\theta = \frac{\ell}{b + \ell}.$$

Thus, observations of binary cooperation decisions generated by the baseline model cannot separately identify the absolute magnitudes of b and ℓ. Indeed, multiplying both parameters by the same positive constant leaves θ, and therefore the trustee's action rule, unchanged.

Equivalently,

$$\frac{\ell}{b} = \frac{\theta}{1 - \theta}.$$

Hence sufficiently informative cooperation data could in principle identify the threshold θ, or the relative payoff ratio ℓ/b, but not b and ℓ separately without additional cardinal payoff information, experimental variation, or further structural restrictions. Any empirical estimation strategy should therefore parameterize the observable cooperation rule in terms of θ unless richer

payoff data are available.

Identification of the persistence parameter λ also depends on what is observed. If both signals and the public state are directly measured, the transition equation provides a natural basis for estimating λ. If T_t is latent and only the binary signals and cooperation decisions are observed, identification becomes substantially more demanding because the latent state path and the cooperation threshold must be inferred jointly. The present paper does not attempt such an estimation exercise.

Ninth, the trustee's current action does not affect the future public state. This one-sided state control is central to the tractability of the model and is the reason the trustee's continuation value cancels from the current cooperation comparison. A natural extension would allow

$$T_{t+1} = F(T_t, s_t, a_t),$$

so that cooperation or noncooperation itself changes the subsequent state. Such a formulation would create genuinely two-sided dynamic incentives and would generally make Player 2's discount factor relevant to the current cooperation decision.

Several further extensions follow from these limitations. The binary signaling action could be replaced by a continuous costly investment, allowing the builder to choose the intensity of trust-building behavior. The public state could be stochastic, signals could be observed with noise, or trustees could have heterogeneous payoff parameters and therefore heterogeneous thresholds. The model could also incorporate asymmetric state adjustment, so that adverse evidence destroys accumulated trust faster than favorable evidence builds it.

A more strategic informational extension would introduce privately known builder types or endogenous signal reliability. In such a game, the public state could become a genuine belief or reputation process, and signaling would serve both to alter current incentives and to convey information. This would connect the reduced-form mechanism studied here more directly to dynamic signaling and reputation models while preserving the distinction between observable behavior and inferred type.

Finally, an empirical extension could test whether a persistence-weighted public state provides a useful reduced-form description of repeated interaction data. Such an analysis would require independently specifying observable counterparts for the signal, the public state, and the cooperation decision, together with an identification strategy appropriate to the available data. This empirical question is intentionally left separate from the theoretical and computational results established in the present paper.

8 Conclusion

This paper develops a discounted dynamic Trust-Building Game in which a trust builder chooses costly observable binary signals, a public state summarizes their accumulated effect, and a trustee observes the updated state before deciding whether to cooperate. The formulation separates three elements that must be distinguished in a repeated trust environment: the dynamics of the public state, the trustee's cooperation response, and the builder's intertemporal signaling

decision.

Within the stationary Markov framework, Proposition 2.1 shows that the trustee's optimal response has the threshold form

$$\sigma_2^*(\widehat{T}) = \mathbf{1}\{\widehat{T} \geq \theta\}, \qquad \theta = \frac{\ell}{b+\ell}.$$

The builder's problem is then characterized by a discounted Bellman equation. Theorem 2.2 establishes existence of a pure stationary Markov equilibrium, while Lemma 3.1 and Proposition 3.2 show that the builder's value function is bounded and nondecreasing. Proposition 3.3 further isolates the immediate and continuation-value components of the gain from positive signaling.

The central strategic result is Theorem 3.7. Under explicit restrictions on the builder's discount factor and signaling cost, the unique optimal stationary action is to signal exactly on the cooperation-activation region. Under the additional conditions of Corollary 3.8, this region lies strictly inside the state space and the optimal policy takes the form

$$0 \longrightarrow 1 \longrightarrow 0.$$

Thus, even though the trustee follows a threshold cooperation rule, the builder's optimal signaling strategy need not be a monotone one-sided threshold policy.

The state-dynamic analysis complements the equilibrium results. Proposition 4.2 gives the exact weighted-history representation of the public state, and Theorem 4.4 characterizes the earliest attainable threshold-crossing state index under persistent positive signaling. Because the trustee acts after observing the updated state, the first cooperative period satisfies

$$\tau_C = \tau_T - 1,$$

which leads directly to the exact finite-horizon cooperation count in Corollary 4.6. The analysis also establishes geometric decay after positive signaling stops through Proposition 4.7 and identifies the minimum current signaling cost required to preserve cooperation for one period in Proposition 4.8. These state-dynamic results are distinct from the builder's infinite-horizon optimal policy and should not be interpreted as alternative optimality conditions.

The computational analysis provides two forms of model-based verification. First, value iteration reproduces the analytically derived activation-band policy for a parameter configuration satisfying the theorem's assumptions, with no policy mismatch on the evaluated state grid. Second, a broader Bellman scan shows how the strict-signaling region changes with patience and signaling cost outside the analytically characterized regime. The numerical calculations also verify the distinction between the threshold-state index and the first cooperative period and preserve the specified cooperation rule at threshold equality. These computations are deterministic evaluations of the model rather than empirical evidence or substitutes for global analytical results.

The framework remains deliberately reduced form. The public state is neither a Bayesian posterior nor an empirically calibrated measure of psychological trust; signal credibility is ex-

ogenous; the trustee's payoff has an affine threshold structure; and the equilibrium analysis is restricted to stationary Markov strategies. Moreover, binary cooperation behavior identifies the threshold

$$\theta = \frac{\ell}{b + \ell},$$

rather than the absolute payoff parameters b and ℓ separately. These limitations define natural extensions involving richer signal spaces, noisy or strategic information, endogenous credibility, two-sided state dynamics, history-dependent strategies, and empirical identification.

Within these stated boundaries, the model provides a tractable mathematical link between observable signaling, state persistence, threshold cooperation, and dynamic incentives. Its principal implication is not that trust building is governed by a universal threshold rule, but that the interaction between current cooperation gains, signaling costs, and continuation values can produce an optimal signaling region with a fundamentally different geometry from the trustee's own cooperation threshold.

Declarations

Funding. The authors received no specific funding for this work.

Competing interests. The authors declare that they have no competing interests.

Data availability. No external dataset was used. All numerical values shown in the figures and table are deterministic evaluations of the equations stated in the manuscript.

Ethical approval. This article does not contain any studies with human participants or animals performed by any of the authors.

AI-assisted language editing. Generative or AI-assisted tools were used only for language editing, proofreading, formatting assistance, and editorial refinement. The research idea, mathematical formulation, scientific reasoning, theoretical development, interpretation, and intellectual contribution remain the responsibility of the authors.

References

[1] Berg, Joyce, Dickhaut, John and McCabe, Kevin. Trust, Reciprocity, and Social History. *Games and Economic Behavior*, **10**(1): 122–142, 1995. DOI: 10.1006/game.1995.1027.

[2] Johnson, Noel D. and Mislin, Alexandra A. Trust Games: A Meta-Analysis. *Journal of Economic Psychology*, **32**(5): 865–889, 2011. DOI: 10.1016/j.joep.2011.05.007.

[3] Duncan, Caitlin, Sganzerla, Lorena, Kaltwasser, Laura and Dziobek, Isabel. Trust Learning in the Repeated Trust Game: A Meta-Analytic Study. *British Journal of Psychology*, **117**(3): 1026–1055, 2026. DOI: 10.1111/bjop.70045.

[4] Cochard, François, Nguyen Van, Phu and Willinger, Marc. Trusting Behavior in a Repeated Investment Game. *Journal of Economic Behavior & Organization*, **55**(1): 31–44, 2004. DOI: 10.1016/j.jebo.2003.07.004.

[5] Engle-Warnick, Jim and Slonim, Robert L. The Evolution of Strategies in a Repeated Trust Game. *Journal of Economic Behavior & Organization*, **55**(4): 553–573, 2004. DOI: 10.1016/j.jebo.2003.11.008.

[6] King-Casas, Brooks, Tomlin, Damon, Anen, Cedric, Camerer, Colin F., Quartz, Steven R. and Montague, P. Read. Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange. *Science*, **308**(5718): 78–83, 2005. DOI: 10.1126/science.1108062.

[7] Charness, Gary and Dufwenberg, Martin. Promises and Partnership. *Econometrica*, **74**(6): 1579–1601, 2006. DOI: 10.1111/j.1468-0262.2006.00719.x.

[8] Bacharach, Michael, Guerra, Gerardo and Zizzo, Daniel J. The Self-Fulfilling Property of Trust: An Experimental Study. *Theory and Decision*, **63**(4): 349–388, 2007. DOI: 10.1007/s11238-007-9043-5.

[9] Cox, James C. How to Identify Trust and Reciprocity. *Games and Economic Behavior*, **46**(2): 260–281, 2004. DOI: 10.1016/S0899-8256(03)00119-2.

[10] Dufwenberg, Martin and Kirchsteiger, Georg. A Theory of Sequential Reciprocity. *Games and Economic Behavior*, **47**(2): 268–298, 2004. DOI: 10.1016/j.geb.2003.06.003.

[11] Falk, Armin and Fischbacher, Urs. A Theory of Reciprocity. *Games and Economic Behavior*, **54**(2): 293–315, 2006. DOI: 10.1016/j.geb.2005.03.001.

[12] Battigalli, Pierpaolo and Dufwenberg, Martin. Guilt in Games. *American Economic Review*, **97**(2): 170–176, 2007. DOI: 10.1257/aer.97.2.170.

[13] Fehr, Ernst and Schmidt, Klaus M. A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, **114**(3): 817–868, 1999. DOI: 10.1162/003355399556151.

[14] Bolton, Gary E. and Ockenfels, Axel. ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, **90**(1): 166–193, 2000. DOI: 10.1257/aer.90.1.166.

[15] Bohnet, Iris and Zeckhauser, Richard. Trust, Risk and Betrayal. *Journal of Economic Behavior & Organization*, **55**(4): 467–484, 2004. DOI: 10.1016/j.jebo.2003.11.004.

[16] Bohnet, Iris, Greig, Fiona, Herrmann, Benedikt and Zeckhauser, Richard. Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review*, **98**(1): 294–310, 2008. DOI: 10.1257/aer.98.1.294.

[17] Ashraf, Nava, Bohnet, Iris and Piankov, Nikita. Decomposing Trust and Trustworthiness. *Experimental Economics*, **9**: 193–208, 2006. DOI: 10.1007/s10683-006-9122-4.

[18] Croson, Rachel and Buchan, Nancy. Gender and Culture: International Experimental Evidence from Trust Games. *American Economic Review*, **89**(2): 386–391, 1999. DOI: 10.1257/aer.89.2.386.

[19] Kosfeld, Michael, Heinrichs, Markus, Zak, Paul J., Fischbacher, Urs and Fehr, Ernst. Oxytocin Increases Trust in Humans. *Nature*, **435**(7042): 673–676, 2005. DOI: 10.1038/nature03701.

[20] Fehr, Ernst. On the Economics and Biology of Trust. *Journal of the European Economic Association*, **7**(2-3): 235–266, 2009. DOI: 10.1162/JEEA.2009.7.2-3.235.

[21] Fehr, Ernst and Charness, Gary. Social Preferences: Fundamental Characteristics and Economic Consequences. *Journal of Economic Literature*, **63**(2): 440–514, 2025. DOI: 10.1257/jel.20241391.

[22] Nowak, Martin A. Five Rules for the Evolution of Cooperation. *Science*, **314**(5805): 1560–1563, 2006. DOI: 10.1126/science.1133755.

[23] Nowak, Martin A. and Sigmund, Karl. Evolution of Indirect Reciprocity by Image Scoring. *Nature*, **393**(6685): 573–577, 1998. DOI: 10.1038/31225.

[24] Manapat, Michael L., Nowak, Martin A. and Rand, David G. Information, Irrationality, and the Evolution of Trust. *Journal of Economic Behavior & Organization*, **90**: S57–S75, 2013. DOI: 10.1016/j.jebo.2012.10.018.

[25] Weigelt, Keith and Camerer, Colin. Reputation and Corporate Strategy: A Review of Recent Theory and Applications. *Strategic Management Journal*, **9**(5): 443–454, 1988. DOI:

10.1002/smj.4250090505.

[26] Kreps, David M. and Wilson, Robert. Reputation and Imperfect Information. *Journal of Economic Theory*, **27**(2): 253–279, 1982. DOI: 10.1016/0022-0531(82)90030-8.

[27] Mailath, George J. and Samuelson, Larry. *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, 2006. DOI: 10.1093/acprof:oso/9780195300796.001.0001.

[28] Sabater, Jordi and Sierra, Carles. Review on Computational Trust and Reputation Models. *Artificial Intelligence Review*, **24**(1): 33–60, 2005. DOI: 10.1007/s10462-004-0041-5.

[29] Han, The Anh, Perret, Cedric and Powers, Simon T. When to (or Not to) Trust Intelligent Machines: Insights from an Evolutionary Game Theory Analysis of Trust in Repeated Games. *Cognitive Systems Research*, **68**: 111–124, 2021. DOI: 10.1016/j.cogsys.2021.02.003.

[30] Zörner, Sebastian, Arts, Emy, Vasiljevic, Brenda, Srivastava, Ankit, Schmalzl, Florian, Mir, Glareh, Bhatia, Kavish, Strahl, Erik, Peters, Annika, Alpay, Tayfun and Wermter, Stefan. An Immersive Investment Game to Study Human-Robot Trust. *Frontiers in Robotics and AI*, **8**: 644529, 2021. DOI: 10.3389/frobt.2021.644529.

[31] Liu, Linjie and Chen, Xiaojie. Conditional Investment Strategy in Evolutionary Trust Games with Repeated Group Interactions. *Information Sciences*, **609**: 1694–1705, 2022. DOI: 10.1016/j.ins.2022.07.073.

[32] Bayat, Davood, Mohamadpour, Hadi, Fang, Huihua, Xu, Pengfei and Krueger, Frank. The Impact of Order Effects on the Framing of Trust and Reciprocity Behaviors. *Games*, **14**(2): 21, 2023. DOI: 10.3390/g14020021.

[33] Carlsson, Fredrik, Demeke, Eyoual, Martinsson, Peter and Tesemma, Tewodros. Measuring Trust in Institutions. *Oxford Economic Papers*, **76**(1): 22–40, 2024. DOI: 10.1093/oep/gpac047.

[34] Guiso, Luigi, Sapienza, Paola and Zingales, Luigi. Trusting the Stock Market. *The Journal of Finance*, **63**(6): 2557–2600, 2008. DOI: 10.1111/j.1540-6261.2008.01408.x.