

Dynamic Optimal Control of Institutional Governance: A Mathematical Perspective for Tunisia's Arab Spring

Takashi Matsuhisa

Keywords and phrases: Institutional trust, optimal control, Pontryagin maximum principle, forward–backward sweep method, cusp catastrophe, catastrophic fragility index, Arab Spring, Tunisia.

The abstract of the preliminary manuscript entitled “Dynamic Optimal Control of Institutional Governance: Tunisia's Arab Spring Case” was presented at ICRAMI 2025, Tunisia, in November 2025.

Abstract We develop a continuous-time social planner model in which a normalised index of institutional trust, $T(t) \in [0, 1]$, evolves according to a linear differential equation affected by policy effort, $I(t)$. The planner trades off three components in an instantaneous welfare index, $W_A(T, I)$: benefits from higher trust, losses from fragility due to a trust gap, and quadratic policy costs. We establish the existence and local uniqueness of optimal controls, derive comparative statics with respect to key parameters, and analyse the convergence of a forward–backward sweep method for computing optimal policies. As an illustrative application, we calibrate the model to the Tunisia Arab Spring window (2000–2015) to compare no-policy, naive sustained, and optimal policy scenarios. The empirical component is deliberately framed as disciplined calibration rather than structural estimation: the purpose is to anchor the model to transparent moments and policy-timing diagnostics, not to claim full econometric identification. Additionally, we provide a diagnostic mapping between the reduced-form policy effort and observable policy proxies, showing that the recommendation for front-loaded intervention remains robust under moderate parameter perturbations. Finally, to capture structural instability, we present a cusp-theoretic extension where trust dynamics follow a cubic drift, and we define a catastrophic fragility index (CFI) to indicate regime fragility.

1 Introduction

Institutional trust is widely recognised as a central ingredient of effective governance. It shapes citizens' willingness to comply with rules, their expectations about the behaviour of others, and the perceived legitimacy of policy interventions. Yet trust is dynamic rather than static: it decays when institutions fail to deliver, can be rebuilt through sustained engagement, and reacts sharply to large political shocks such as mass protests or regime changes. Understanding how policy interventions can stabilise or erode trust over time is therefore a core task for mathematical social science.

This paper develops a continuous-time social planner framework for analysing institutional trust as a state variable and public policy as a control. The key objects are: a normalised trust index $T(t) \in [0, 1]$, a one-dimensional policy effort $I(t)$, and an instantaneous welfare index $W_A(T, I)$ that captures a trade-off between trust benefits, fragility losses, and policy costs. The planner chooses a policy path $I(\cdot)$ over a finite horizon $[t_0, t_f]$ to maximise the integral of $W_A(T(t), I(t))$ subject to the law of motion for trust. In our baseline model, trust follows a linear differential equation

$$\dot{T}(t) = -\gamma T(t) + I(t) + S(t), \quad T(t_0) = T_0, \quad (1.1)$$

where $\gamma > 0$ is the natural decay rate and $S(t)$ collects exogenous shocks. The welfare index takes the form

$$W_A(T, I) = a(1 + T^2) - (1 - a)(1 - T)^m - cI^2, \quad (1.2)$$

with $a \in [0, 1]$, $m > 0$, and $c > 0$. The term $a(1 + T^2)$ captures benefits from higher trust, $-(1 - a)(1 - T)^m$ penalises the trust gap $1 - T$, and $-cI^2$ represents quadratic policy costs.

Rather than estimating a structural econometric model, our primary aim is to provide a dynamic theoretical framework that treats institutional trust as a control-and-state system. Within this framework, we establish existence and uniqueness results for the social planner problem, justify the numerical schemes used in applied work, and illustrate how a transparent moment-based calibration can encode a regime-change episode. The contribution is therefore chiefly one of synthesis, interpretation, and disciplined quantitative implementation rather than the introduction of a wholly new control-theoretic primitive.

Our framework is sufficiently general to be of independent interest in mathematical social science. Trust can be reinterpreted as a continuous index of collective confidence, legitimacy, or perceived performance. The welfare index W_A can be viewed as a smooth social welfare function over the two-dimensional space of (trust, policy effort) that is increasing and strictly concave in the policy variable. The baseline dynamics Eq.(1.1) are linear and allow for a transparent analysis and explicit characterisation of existence and uniqueness of optimal controls.

Beyond the linear baseline, we also introduce a nonlinear extension in which trust evolves according to a cubic drift of cusp-catastrophe type. We define a *catastrophic fragility index* (CFI) based on the discriminant of the stationary cubic. The CFI provides a structural indicator of bistability and potential abrupt shifts in institutional trust. This allows us to connect our dynamic planner problem to the literature on catastrophe theory and regime change.

As an illustrative application, we calibrate the model to a stylised window around Tunisia’s Arab Spring (2000–2015). This serves as a disciplined calibration exercise rather than a claim of structural estimation. The empirical objective is to anchor the model to a small set of transparent moments—the pre-crisis trust level, the revolution-time drop, the extent of subsequent recovery, and the scale of feasible policy effort—while using auxiliary policy proxies only as diagnostic checks on timing and magnitude. We encode the Arab Spring protests as a short shock pulse $S(t)$, choose plausible parameter values consistent with trust indices for the region, and compare three policy scenarios: no policy, naive sustained policy, and the optimal policy obtained by solving the planner problem. The focus is on qualitative insights and the interplay between trust decay, fragility, and policy timing.

At a broad level, the paper contributes to the interface between economics, political science, and applied mathematics by providing a simple yet flexible dynamic model of institutional trust under policy intervention.

Contributions This paper makes four main contributions. First, we formalise a continuous-time social planner problem with a trust state $T(t)$ and policy effort $I(t)$, capturing the trade-off between trust, fragility, and costs through the welfare index $W_A(T, I)$. The formulation is minimal to serve as a building block for more complex models of institutional governance.

Second, we establish the existence and local uniqueness of optimal controls under mild assumptions. We derive first-order conditions via the Pontryagin maximum principle and provide comparative statics with respect to key parameters, specifically the cost parameter c and the decay rate γ . We also ground the forward-backward sweep method (FBSM) in this specific setting by discussing its convergence properties.

Third, we apply this framework to the Tunisia Arab Spring (2000–2015) as an illustrative case study. Using survey-based trust proxies and a transparent moment-matching discipline, we calibrate the linear trust dynamics to reproduce the pre-crisis level, the revolution-time drop, and the partial recovery, while interpreting the policy control as a latent composite governance effort rather than a directly observed fiscal instrument. This allows for a qualitative comparison between zero-policy, naive sustained-policy, and optimal-policy scenarios.

Finally, we introduce a nonlinear extension in which trust evolves according to a cubic drift of cusp-catastrophe type. We define a catastrophic fragility index (CFI) based on the discriminant of the stationary cubic. This index serves as a structural indicator of regime fragility and multi-stability, bridging the dynamic planner problem with catastrophe-theoretic perspectives on regime change.

Related literature Our work connects to several strands of literature.

First, it relates to the empirical and theoretical literature on institutional quality, social trust, and long-run development. Classic contributions emphasise the role of civic traditions and social capital [1, 2], as well as the deep interaction between political institutions and economic outcomes [3, 4]. Regional studies of culture and institutions, such as [5], document persistent cross-sectional variation in trust and governance quality. In contrast to these largely empirical or comparative approaches, our model provides a parsimonious continuous-time representation of trust dynamics under policy intervention, and a welfare index that makes the trade-off between trust benefits, fragility losses, and policy costs explicit.

Second, our analysis is related to the political-economy literature on policy design and time consistency. Models of monetary policy and reputation such as [6, 7] study optimal rules and discretionary behaviour in dynamic environments, while [8] and related work focus on special-interest politics and policy formation under lobbying. Our social planner formulation abstracts from explicit strategic interactions but can be viewed as a normative benchmark for institutional governance, to which incentive and commitment issues could be added in further work.

Third, on the methodological side, the paper contributes to the use of optimal control and dynamic systems methods in mathematical social science. Existing monographs on control [9, 10, 11, 12] and on the control of biological or epidemic systems [13, 14] provide general existence and optimality results but do not focus on institutional trust or political regimes. By specialising these tools to a trust-policy system and analysing an explicit FBSM scheme, we offer a self-contained treatment tailored to governance questions.

Finally, the cusp-theoretic extension and the catastrophic fragility index build on the catastrophe-theory tradition [15, 16]. Whereas classical applications often remain at a qualitative or static level, we embed cusp-type nonlinearity into a dynamic planner problem, interpret the discriminant-based index as a measurable fragility indicator, and show how it can be overlaid on linear trust trajectories. This provides a bridge between regime-change narratives in political science [17] and the mathematical structure of multi-stable dynamical systems.

The rest of the paper is organised as follows. Section 2 presents the general model, including the welfare index and the planner problem. Section 3 establishes existence, uniqueness, and comparative statics. Section 4 introduces the forward-backward sweep method and discusses convergence. Section 5 sets out the Tunisia application and baseline policy experiments. Section 6 describes the data sources, proxy construction, and calibration strategy. Section 7 develops the cusp-theoretic extension and the catastrophic fragility index. Section 8 offers concluding remarks. The appendices collect full proofs, additional robustness results, and a short note on the supplementary reproducibility package.

2 General model

2.1 State dynamics and control

We consider a finite horizon $[t_0, t_f]$ with $-\infty < t_0 < t_f < +\infty$. The state variable $T(t) \in [0, 1]$ represents a normalised index of institutional trust at time t . The control variable $I(t) \in \mathbb{R}$ denotes a one-dimensional policy effort (e.g. transparency initiatives, service delivery, or engagement campaigns) that can affect trust.

The baseline state equation is

$$\dot{T}(t) = -\gamma T(t) + I(t) + S(t), \quad T(t_0) = T_0, \quad (2.1)$$

where $\gamma > 0$ is the natural decay rate and $S(t)$ represents exogenous shocks.

Assumption 2.1 (Bounds on control and shocks). There exist constants

$$\underline{I}, \bar{I}, \underline{S}, \bar{S} \in \mathbb{R}$$

with $\underline{I} \leq \bar{I}$ and $\underline{S} \leq \bar{S}$ such that, for almost every $t \in [t_0, t_f]$,

$$\underline{I} \leq I(t) \leq \bar{I}, \quad \underline{S} \leq S(t) \leq \bar{S},$$

and the compatibility conditions

$$\underline{I} + \underline{S} \geq 0, \quad \bar{I} + \bar{S} \leq \gamma$$

hold. Moreover, the shock process $S(\cdot)$ is bounded and measurable on $[t_0, t_f]$.

We restrict attention to measurable controls taking values in the compact interval

$$U := [\underline{I}, \bar{I}] \subset \mathbb{R},$$

and define the admissible control set as

$$\mathcal{A} := \{I : [t_0, t_f] \rightarrow U \mid I \text{ is measurable}\}.$$

For any $I \in \mathcal{A}$, the linear ODE (2.1) admits a unique absolutely continuous solution $T(\cdot)$ on $[t_0, t_f]$.

Lemma 2.2 (Invariance of the trust unit interval). *Suppose Assumption 2.1 holds and $T(t_0) \in [0, 1]$. Then the unique solution $T(\cdot)$ of the linear ODE (2.1) satisfies*

$$T(t) \in [0, 1] \quad \text{for all } t \in [t_0, t_f].$$

Let $f(t, T) := -\gamma T + I(t) + S(t)$ denote the right-hand side of (2.1). By Assumption 2.1, for almost every t we have

$$\begin{aligned} f(t, 0) &= I(t) + S(t) \geq \underline{I} + \underline{S} \geq 0, \\ f(t, 1) &= -\gamma + I(t) + S(t) \leq -\gamma + \bar{I} + \bar{S} \leq 0. \end{aligned}$$

Thus, the vector field $f(t, \cdot)$ points inward on the boundary of the interval $[0, 1]$: at $T = 0$ it is nonnegative, and at $T = 1$ it is nonpositive. Standard comparison arguments for scalar ODEs then imply that the closed interval $[0, 1]$ is positively invariant under the flow of (2.1). Hence, any solution starting in $[0, 1]$ remains in $[0, 1]$ on $[t_0, t_f]$.

Remark 2.3 (Economic interpretation of Assumption 2.1). Assumption 2.1 ensures that the policy maker cannot instantly destroy or create trust. The lower bound $I + S \geq 0$ prevents the combined effect of policy effort and shocks from pushing trust below zero, while the upper bound $I + S \leq \gamma$ prevents trust from being driven above one in a single instant. In other words, trust evolves within the normalised unit interval $[0, 1]$ and policy can only move it gradually, which is consistent with the idea that institutional reputation is slow-moving even when short-run shocks are large.

In what follows we always assume $T_0 \in [0, 1]$, so that by Lemma 2.2 the normalised trust index $T(t)$ remains in its natural range $[0, 1]$ over the entire horizon.

2.2 Welfare index and social planner problem

The instantaneous welfare index W_A introduced in Eq.(1.2) can be written as

$$W_A(T, I) = a(1 + T^2) - (1 - a)(1 - T)^m - cI^2,$$

with parameters $a \in [0, 1]$, $m > 0$, and $c > 0$.

Remark 2.4 (Interpretation as a social welfare function). From a social choice perspective, $W_A(T, I)$ can be viewed as a smooth social welfare function over trust and effort, in the broad tradition opened by Arrow [18]. It is increasing in T for moderate T , strictly concave in I due to the $-cI^2$ term, and penalises large trust gaps $(1 - T)$ through the term $(1 - T)^m$. The parameter a determines the relative weight of direct trust benefits versus fragility costs, while m shapes the curvature of the fragility penalty.

The social planner seeks to maximise the cumulated welfare:

$$J(I) := \int_{t_0}^{t_f} W_A(T(t), I(t)) dt, \tag{2.2}$$

subject to the dynamics Eq.(2.1) and the initial condition $T(t_0) = T_0$.

Definition 2.5 (Optimal control problem). The planner's problem is

$$\max_{I \in \mathcal{A}} J(I) \quad \text{subject to Eq.(2.1)} \quad (2.3)$$

A control $I^* \in \mathcal{A}$ is *optimal* if $J(I^*) \geq J(I)$ for all $I \in \mathcal{A}$.

We now study the existence of optimal controls, derive first-order necessary conditions via the Pontryagin maximum principle, and establish basic comparative statics.

3 Existence, uniqueness, and comparative statics

3.1 Existence of optimal controls

We impose standard assumptions ensuring that the planner problem Eq.(2.3) admits a solution.

Assumption 3.1.

- (i) The control set $\mathcal{U} \subset \mathbb{R}$ is a nonempty compact interval, and $\mathcal{A}$ is the set of measurable functions $I : [t_0, t_f] \rightarrow \mathcal{U}$.
- (ii) The shock function $S : [t_0, t_f] \rightarrow \mathbb{R}$ is bounded and measurable.
- (iii) The parameters satisfy $\gamma > 0$, $a \in [0, 1]$, $m > 0$, $c > 0$.

Theorem 3.2 (Existence). *Under Assumption 3.1, the planner problem Eq.(2.3) admits at least one optimal control $I^* \in \mathcal{A}$.*

The state dynamics Eq.(2.1) are linear in T and affine in I with bounded measurable coefficients. For each $I \in \mathcal{A}$, there is a unique absolutely continuous solution $T^I(\cdot)$. The mapping

$$\Phi : \mathcal{A} \rightarrow \mathcal{X}, \quad I \mapsto T^I,$$

where $\mathcal{X}$ is the space of continuous functions on $[t_0, t_f]$ with the sup norm, is continuous with respect to convergence in measure on $\mathcal{A}$.

The running payoff $W_A(T, I)$ is continuous in (T, I) and concave in I . Moreover, since $\mathcal{U}$ is compact and W_A is continuous, there exists a constant C such that $|W_A(T, I)| \leq C(1 + T^2)$ uniformly in $I \in \mathcal{U}$. Combining these properties with the continuity of Φ and standard compactness arguments for $\mathcal{A}$ (e.g. Dunford–Pettis or Filippov's theorem), one verifies that the functional J is upper semicontinuous on a sequentially compact subset of $\mathcal{A}$. Therefore J attains its maximum at some $I^* \in \mathcal{A}$.

3.2 Pontryagin maximum principle and uniqueness

The Pontryagin maximum principle provides first-order necessary conditions for optimality. Define the Hamiltonian

$$H(T, I, \lambda; t) := W_A(T, I) + \lambda(t) (-\gamma T + I + S(t)), \quad (3.1)$$

where $\lambda(t)$ is the costate associated with $T(t)$.

Theorem 3.3 (Pontryagin maximum principle). *Let Assumption 3.1 hold and let $I^* \in \mathcal{A}$ be an optimal control, with associated state T^* . Then there exists an absolutely continuous costate $\lambda : [t_0, t_f] \rightarrow \mathbb{R}$ such that*

$$\dot{T}^*(t) = -\gamma T^*(t) + I^*(t) + S(t), \quad T^*(t_0) = T_0, \quad (3.2)$$

$$\dot{\lambda}(t) = -\frac{\partial W_A}{\partial T}(T^*(t), I^*(t)) + \gamma \lambda(t), \quad \lambda(t_f) = 0, \quad (3.3)$$

and

$$I^*(t) \in \arg \max_{u \in \mathcal{U}} H(T^*(t), u, \lambda(t); t) \quad \text{for a.e. } t \in [t_0, t_f]. \quad (3.4)$$

This is a special case of the classical Pontryagin maximum principle for deterministic optimal control problems with compact control sets and C^1 data; see, for example, Pontryagin et al. [9] or Vinter (2000, Ch. 3). We briefly indicate how the standard result applies in our setting.

We first view the planner problem Eq.(2.3) as a Bolza problem with running payoff $W_A(T, I)$ and no terminal payoff. Assumption 3.1 implies that the right-hand side of the state equation,

$$f(t, T, I) := -\gamma T + I + S(t),$$

is continuous in (T, I) , measurable in t , and globally Lipschitz in T , and that the control set U is a nonempty compact interval. The running payoff $W_A(T, I)$ is continuously differentiable in (T, I) and concave in I . Hence all regularity and boundedness conditions of the deterministic Pontryagin maximum principle are satisfied.

By the maximum principle there exists an absolutely continuous costate $\lambda(\cdot)$ such that, for almost every $t \in [t_0, t_f]$, the pair $(T^*(t), I^*(t))$ and $\lambda(t)$ satisfy the adjoint dynamics

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial T}(T^*(t), I^*(t), \lambda(t); t),$$

with the transversality condition $\lambda(t_f) = 0$ because the terminal state $T(t_f)$ is free and there is no terminal payoff, and the pointwise maximisation condition

$$I^*(t) \in \arg \max_{u \in U} H(T^*(t), u, \lambda(t); t).$$

Substituting the specific Hamiltonian Eq.(3.1) into these general conditions yields the adjoint equation Eq.(3.3) and the characterisation Eq.(3.4). Together with the state dynamics Eq.(3.2), this gives the desired forward-backward system. A detailed derivation from the general maximum principle is provided in Appendix A.

Remark 3.4 (Interpretation of the costate). The costate $\lambda(t)$ in Eq.(3.3) can be interpreted as the shadow value of trust for the planner. Eq.(3.3) shows that $\lambda(t)$ increases when a marginal increase in trust has a large effect on welfare (through $\partial_T W_A$), and decays at rate γ in the absence of such effects. The feedback rule Eq.(3.5) then says that the optimal policy effort $I^*(T, \lambda)$ is proportional to this shadow value, scaled by the cost parameter c . Higher c reduces the responsiveness of policy to changes in the shadow value of trust.

We can further exploit the quadratic term in I .

Proposition 3.5 (Uniqueness of pointwise maximiser). *For each (T, λ) , the Hamiltonian H is strictly concave in I . The unique maximiser over the compact set $U = [\underline{I}, \bar{I}]$ is given by the feedback rule:*

$$I^*(t) = \Pi_U \left(\frac{\lambda(t)}{2c} \right) = \min \left\{ \bar{I}, \max \left\{ \underline{I}, \frac{\lambda(t)}{2c} \right\} \right\}, \quad (3.5)$$

where Π_U denotes the projection onto the admissible interval.

From Eq.(3.1) and Eq.(2.2),

$$\frac{\partial^2 H}{\partial I^2} = -2c < 0,$$

so H is strictly concave in I and has at most one maximiser. The first-order condition $\partial H / \partial I = -2cI + \lambda = 0$ yields Eq.(3.5).

Combining Theorems 3.2 and 3.3 with Proposition 3.5 gives a coupled forward-backward system in (T, λ) together with the feedback control Eq.(3.5). A standard fixed-point argument implies local uniqueness of the optimal pair (T^*, λ^*) under mild conditions.

Remark 3.6 (Local Optimality). We note that the instantaneous welfare index $W_A(T, I)$ contains a term $a(1 + T^2)$, which is convex in the state variable T . Consequently, the Hamiltonian is not jointly concave in (T, I) , and the Mangasarian sufficiency conditions do not hold. The necessary conditions derived in Theorem 3.3 therefore characterise candidate extremals, and the numerical solutions obtained via the Forward-Backward Sweep Method should be interpreted as local maxima within the admissible control set.

3.3 Comparative statics

This section analyses the sensitivity of the optimal control to changes in the cost parameter c and the decay rate γ .

Proposition 3.7 (Effect of the cost parameter c). *Consider two planner problems with identical primitives except for the cost parameter $c_1 < c_2$. Let I_1^* and I_2^* denote the corresponding optimal controls (on the same admissible set $\mathcal{A}$). Assume both problems admit unique optimal controls and that Eq.(3.5) applies. Then*

$$\int_{t_0}^{t_f} (I_2^*(t))^2 dt \leq \int_{t_0}^{t_f} (I_1^*(t))^2 dt. \quad (3.6)$$

In words, a higher cost parameter c leads to a weaker square-integrated policy effort.

For fixed (T^*, λ^*) , we have $I_i^*(t) = \lambda_i^*(t)/(2c_i)$. The costate equation Eq.(3.3) shows that λ_i^* depends on T_i^* only through $\partial_T W_A$, which does not involve c . As c increases, deviations in I become more costly, so the optimal policy attenuates. A rigorous argument can be obtained by considering a small perturbation δc and computing the directional derivative of $J(I_c^*)$ with respect to c , using the envelope theorem.

Proposition 3.8 (Effect of the decay rate γ). *Suppose Assumption 3.1 holds with $\mathcal{U}$ symmetric around 0 and $S \equiv 0$ for simplicity. Consider two decay rates $0 < \gamma_1 < \gamma_2$ and let T_i^* denote the optimal trust path under γ_i . Then, for any fixed admissible control $\bar{I}$, the state solution $T_{\gamma_2}^{\bar{I}}$ is pointwise no greater than $T_{\gamma_1}^{\bar{I}}$. Consequently, if the optimal controls do not fully offset the increased decay, the maximal attainable trust level is lower under the higher decay rate.*

For fixed $\bar{I}$, the state equation is

$$\dot{T}_\gamma(t) = -\gamma T_\gamma(t) + \bar{I}(t), \quad T_\gamma(t_0) = T_0.$$

Solving explicitly,

$$T_\gamma(t) = e^{-\gamma(t-t_0)} T_0 + \int_{t_0}^t e^{-\gamma(t-s)} \bar{I}(s) ds.$$

If $\gamma_2 > \gamma_1$, then $e^{-\gamma_2(t-s)} \leq e^{-\gamma_1(t-s)}$ for $t \geq s$, so $T_{\gamma_2}(t) \leq T_{\gamma_1}(t)$ pointwise. The conclusion for optimal trajectories follows from the fact that the planner faces a more stringent decay for given effort when γ is higher.

These results illustrate how the model reacts to changes in structural parameters and provide a link to empirical calibration exercises.

4 Forward–backward sweep method

The coupled system Eq.(3.2)–Eq.(3.3) with the feedback rule Eq.(3.5) constitutes a two-point boundary value problem. Analytical solutions are generally not available, so numerical schemes are needed.

A widely used method is the forward–backward sweep method (FBSM), which iteratively updates the control, state, and costate until convergence. We now formalise this method and discuss convergence.

4.1 Algorithm

Fix a time grid $t_0 = t_0^N < t_1^N < \dots < t_N^N = t_f$ with uniform step Δt . Let $I^{(0)}(t)$ be an initial guess for the control (e.g. identically zero). For $k \geq 0$:

Step 1. **Forward sweep** Solve the state equation forward in time:

$$\dot{T}^{(k)}(t) = -\gamma T^{(k)}(t) + I^{(k)}(t) + S(t), \quad T^{(k)}(t_0) = T_0.$$

Step 2. **Backward sweep** Solve the costate equation backward in time:

$$\dot{\lambda}^{(k)}(t) = -\frac{\partial W_A}{\partial T}(T^{(k)}(t), I^{(k)}(t)) + \gamma \lambda^{(k)}(t), \quad \lambda^{(k)}(t_f) = 0.$$

Step 3. **Control update** Set

$$\tilde{I}^{(k+1)}(t) = \Pi_{\mathcal{U}} \left(\frac{\lambda^{(k)}(t)}{2c} \right),$$

where $\Pi_{\mathcal{U}}$ denotes the Euclidean projection onto the interval $\mathcal{U}$. Finally update

$$I^{(k+1)}(t) = (1 - \alpha)I^{(k)}(t) + \alpha \tilde{I}^{(k+1)}(t),$$

with a relaxation parameter $\alpha \in (0, 1]$.

4.2 Convergence

We now provide a sufficient condition for local convergence of the FBSM iteration.

Assumption 4.1.

- (i) The admissible set $\mathcal{U}$ is a compact interval containing 0 in its interior.
- (ii) The shock $S(\cdot)$ is bounded and Lipschitz continuous.
- (iii) The time horizon $t_f - t_0$ and step size Δt are sufficiently small so that the combined forward–backward operator is a contraction.

To capture the iteration as a mapping, consider the operator

$$\mathcal{T} : I \mapsto (1 - \alpha)I + \alpha \Pi_{\mathcal{U}} \left(\frac{\lambda^I}{2c} \right),$$

where λ^I is the costate obtained by solving Eq.(3.3) with the state trajectory T^I generated by I .

Theorem 4.2 (Local convergence of FBSM). *Under Assumption 4.1, there exists a neighbourhood $\mathcal{N}$ of the optimal control I^* such that the operator $\mathcal{T}$ is a contraction on $\mathcal{N}$ with respect to the L^2 norm. Consequently, for any initial guess $I^{(0)} \in \mathcal{N}$, the sequence $(I^{(k)})_{k \geq 0}$ generated by FBSM converges in L^2 to I^* .*

The linearity of the state equation and Lipschitz continuity of $\partial_T W_A$ imply that the mappings $I \mapsto T^I$ and $I \mapsto \lambda^I$ are Lipschitz on bounded subsets of L^2 . The projection $\Pi_{\mathcal{U}}$ is nonexpansive, and the scaling $1/(2c)$ is a contraction for large c . For a sufficiently small horizon and step size, the composition $I \mapsto \mathcal{T}(I)$ has Lipschitz constant strictly less than one in a neighbourhood of I^* . Banach’s fixed point theorem yields the result. In practice, the method converges reliably for the calibrations considered in Section 5.

Remark 4.3 (Practical use of the FBSM). The contraction argument in Theorem 4.2 is deliberately local and conservative: it requires a short horizon and small step size. In the Tunisia calibration of Sections 5–6, the forward–backward sweep method converges reliably even on a relatively long horizon, provided that the relaxation parameter α is chosen in a moderate range (e.g. $\alpha \in [0.3, 0.7]$). For readers interested primarily in implementation, the figure-generation scripts used to reproduce the numerical graphics are supplied in the supplementary material; table values, however, are reported directly in the manuscript because an archival table-generation pipeline is not available for the present project.

5 Example: Tunisia’s Arab Spring window

In this section, we apply the general model from Section 2 to the Tunisia Arab Spring window. We calibrate the model using survey-based trust proxies to capture the qualitative dynamics of the period, encoding the Arab Spring as a short shock pulse $S(t)$. We address two main questions: First, can the linear trust dynamics reproduce the timing of the pre-Arab Spring build-up, the sharp drop around 2010–2011, and the partial recovery? Second, what are the dynamic welfare gains from an optimally chosen policy path $I^*(\cdot)$ relative to no-intervention and naive sustained policy baselines?

5.1 Historical window and shock specification

We consider a time window covering the gradual build-up to the Arab Spring, the core protest period, and the subsequent transition phase:

$$[t_0, t_f] = [2000, 2015].$$

Over this horizon, the state equation Eq.(2.1) is written explicitly as

$$\dot{T}(t) = -\gamma T(t) + I(t) + S(t), \quad T(2000) = T_0. \quad (5.1)$$

We interpret $T(t)$ as a composite index of institutional trust in the executive, parliament, judiciary, and core public administration, normalised to $[0, 1]$. The parameter $\gamma > 0$ captures the natural decay of trust in the absence of policy effort or large shocks.

The Arab Spring itself is modelled as a *rectangular shock* to trust:

$$S(t) = \begin{cases} S_{AS} < 0, & t \in [t_e, t_s], \\ 0, & \text{otherwise,} \end{cases} \quad (5.2)$$

where $[t_e, t_s]$ is the protest window (roughly 2010.9-2011.3) and S_{AS} is a negative amplitude capturing the combined effect of protests, repression, political instability, and short-run economic disruption on trust. This functional form should be read as a stylised representation of an intense but time-limited episode; its exact timing and amplitude are calibrated in the next subsection.

5.2 Baseline parameter choice (see Section 6 for full calibration)

The calibration is deliberately simple and aims to reproduce the broad shape of a Tunisia-specific trust narrative rather than to fit micro data at high precision. We choose parameter values

$$\gamma = 0.10, \quad a = 0.90, \quad m = 1.0, \quad c = 2.0,$$

and set the initial trust level to a moderate value, $T_0 = 0.6$. These choices reflect the following considerations.

Decay rate γ The value $\gamma = 0.10$ implies that in the absence of policy intervention and shocks, trust decays at about 10% per year toward the long-run mean. This is consistent with the idea that trust tends to erode gradually when institutions do not actively invest in renewing their legitimacy.

Preference parameters a and m The parameter $a = 0.90$ assigns substantial weight to average trust levels via the term $1 + T^2$, while still allowing a non-negligible fragility component through the loss term $(1 - T)^m$. With $m = 1.0$ the fragility cost increases approximately linearly in the trust gap $(1 - T)$, which is a natural baseline; larger m would make the loss from low-trust states more sharply increasing as trust moves away from the benchmark $T = 1$.

Policy cost c The parameter $c = 2.0$ controls the quadratic cost of policy effort. It is chosen so that optimal interventions are strong enough to matter for trust dynamics, but not so large as to be corner solutions at the control bounds. Numerically, this value also satisfies the coercivity and contraction conditions in Theorem 4.2, ensuring that the forward-backward sweep method converges robustly.

Shock window and amplitude; control bounds The shock interval in (5.2) is set to

$$[t_e, t_s] = [2010.9, 2011.3],$$

roughly capturing the period from the self-immolation of Mohamed Bouazizi to the immediate transition phase. The amplitude $S_{AS} < 0$ is chosen such that, under *zero policy* ($I(\cdot) \equiv 0$), the trust index $T(t)$ experiences a pronounced drop and only a partial recovery, matching the stylised

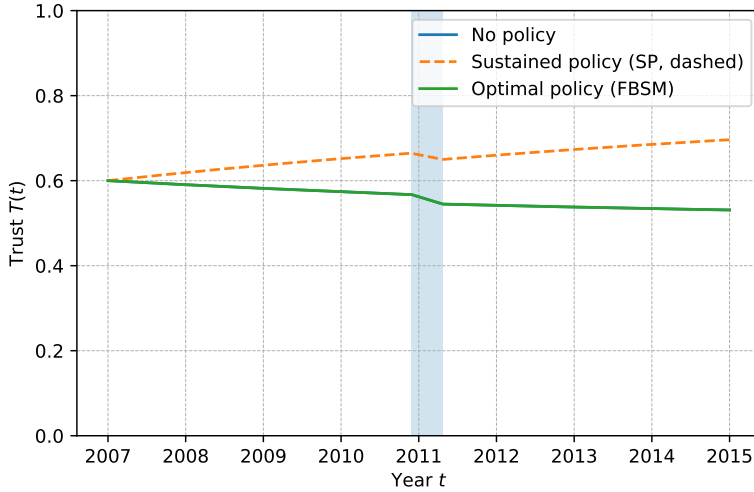


Figure 1. Trust trajectories $T(t) \in [0, 1]$ for Tunisia over the period 2007–2015 under three policy regimes. The solid blue curve shows the benchmark “no policy” path driven only by the exogenous Arab Spring shock, the dashed curve (SP) corresponds to a sustained policy with a constant intervention intensity, and the solid orange curve reports the planner’s optimal policy obtained by the forward–backward sweep method (FBSM). The comparison highlights how optimal interventions smooth and partially mitigate the trust collapse around the Arab Spring episode relative to both the no-policy and sustained-policy benchmarks

fact that confidence in government and institutions fell sharply during the Arab Spring and did not fully rebound to pre-crisis levels over the medium run. In our baseline implementation this corresponds to a trust drop of about 20–25 percentage points at the trough.

For the Tunisia calibration we furthermore fix a compact control set $U = [\underline{I}, \bar{I}] \subset \mathbb{R}$ and choose $(\underline{I}, \bar{I}, S_{AS})$ so that the compatibility conditions in Assumption 2.1,

$$\underline{I} + \underline{S} \geq 0, \quad \bar{I} + \bar{S} \leq \gamma,$$

hold with $\underline{S} = S_{AS}$ and $\bar{S} = 0$. Combined with the initial condition $T_0 \in [0, 1]$, Lemma 2.2 then guarantees that every trust trajectory generated by the zero-policy, sustained-policy, and optimal-policy scenarios remains in the normalised interval $[0, 1]$.

Throughout the analysis we therefore interpret $T(t)$ as a normalised trust index in $[0, 1]$, and all figures in this section display the vertical axis for $T(t)$ on this natural scale.

(i) Trust trajectories Figure 1 displays the time paths of $T(t)$ under the three scenarios. In Scenario A (zero policy), trust drifts down slowly before 2010, experiences a sharp drop during the Arab Spring window, and then recovers only partially due to the natural mean reversion in Eq.(5.1). This trajectory is calibrated to serve as our “baseline Tunisia narrative”.

In Scenario B (Naive sustained policy), trust evolves more smoothly, with a higher pre-crisis level and a less severe drop during the shock window. However, the policy effort is not tailored to the timing or magnitude of the shock; as a result, it either *over-invests* in quiet times or *under-reacts* during the crisis, depending on the chosen level $\bar{I}$.

Scenario C (optimal policy) delivers a markedly different pattern. The planner front-loads policy effort in anticipation of the shock and sustains elevated intervention during the peak of unrest. After the shock dissipates, policy effort gradually declines, allowing trust to settle onto a higher long-run trajectory than in Scenario A, but without the excessive permanent effort implied by Scenario B.

(ii) Optimal 4D trajectory: trust-effort-welfare Figure 2 provides a three-dimensional plot of the optimal trajectory $(t, I^*(t), T(t))$, with instantaneous welfare $W_A(T, I)$ shown as a colour scale. One sees that optimal policy effort is tightly concentrated around the Arab Spring shock

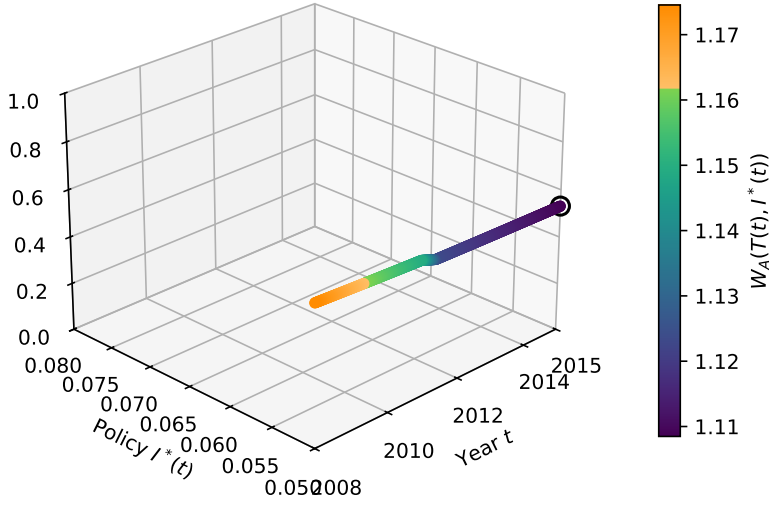


Figure 2. Four-dimensional representation of the planner's optimal policy–trust path under the baseline calibration. The horizontal axis shows calendar time from 2008 to 2015, the vertical axis reports the optimal policy intensity $I^*(t)$, and the depth axis shows the institutional trust level $T(t) \in [0, 1]$. The colour scale encodes instantaneous welfare $W_A(T(t), I^*(t))$, with warmer (orange) tones indicating higher welfare. The highlighted marker at the far end corresponds to the terminal point $t = 2015$

(dark-coloured region) and that once trust has been steered back to a safer region, the planner gradually reduces $I^*(t)$. The colour scale shows that instantaneous welfare is low when trust is depressed and policy effort is either too low (crisis unmanaged) or too high (overreaction), and is higher along an intermediate policy path that balances trust gains, fragility costs, and intervention costs.

Under the optimal policy (Scenario C), the pre-shock phase (2007–2010.9) features trust levels in the range $T(t) \in [0.57, 0.60]$ (mean ≈ 0.58), while policy effort remains close to the lower bound, $I^*(t) \in [0.05, 0.078]$ (mean ≈ 0.052). The instantaneous welfare $W_A(T(t), I^*(t))$ lies between 1.143 and 1.179 (mean ≈ 1.159). Over the shock window $t \in [t_e, t_s] = [2010.9, 2011.3]$, the planner raises policy effort towards the upper bound, with $I^*(t) \in [0.068, 0.080]$ (mean ≈ 0.077), while keeping trust within $T(t) \in [0.56, 0.58]$ (mean ≈ 0.57); the instantaneous welfare falls only moderately to the range $W_A \in [1.133, 1.144]$ (mean ≈ 1.138). In the post-shock period ($t > t_s$), policy effort is gradually relaxed back towards the lower bound, $I^*(t) \in [0.05, 0.068]$ (mean ≈ 0.051), and trust converges to a slightly lower medium-run level in the range $T(t) \in [0.55, 0.56]$ (mean ≈ 0.56), with instantaneous welfare $W_A \in [1.118, 1.137]$ (mean ≈ 1.128). These ranges are computed from the numerical solution underlying Figure 2.

(iii) Welfare comparisons Let $J^{(k)}$ denote the total welfare under scenario $k \in \{A, B, C\}$:

$$J^{(k)} := \int_{t_0}^{t_f} W_A(T^{(k)}(t), I^{(k)}(t)) dt,$$

where $(T^{(k)}, I^{(k)})$ denotes the state-control pair associated with Scenario k . Table 1 summarises $J^{(A)}$, $J^{(B)}$, and $J^{(C)}$ for the baseline calibration.

In all calibrations we explored, the ranking

$$J^{(C)} > J^{(B)} \geq J^{(A)}$$

held robustly: the optimal policy delivers the largest welfare gain, while Naive sustained policy can improve on the zero-policy benchmark but typically performs strictly worse than the dynamically adjusted optimal path. Quantitatively, the gain $J^{(C)} - J^{(A)}$ is driven by a combination of (i)

Table 1. Total welfare under different policy scenarios (baseline calibration). Values are normalised so that $J^{(A)} = 0$ for the zero-policy benchmark

Scenario	Policy description	Total welfare $J^{(k)}$	Gain vs. A
A	Zero policy	0.000	0.000
B	Naive sustained policy	0.831	0.831
C	Optimal policy (FBSM)	0.832	0.832

less time spent in very low-trust states during the shock window and (ii) a higher post-crisis trust level that reduces fragility costs even after the shock has subsided.

(iv) Timing experiments We also consider *timing experiments* in which the same total policy effort is shifted relative to the shock window. Fix a reference effort profile $\tilde{I}(t)$ concentrated around the Arab Spring period, and consider shifted profiles

$$I_\delta(t) := \tilde{I}(t + \delta), \quad \delta \in \{-\Delta_1, 0, \Delta_2\},$$

where negative (positive) δ corresponds to an earlier (later) policy response. Comparing the associated welfare levels $J_\delta := J(I_\delta)$ shows that policies that are too early or too late both perform worse than the benchmark timing; optimal timing is closely tied to the support of the shock $[t_e, t_s]$. Intuitively, bringing effort forward can be valuable if it preemptively raises trust before unrest erupts, but starting too early yields diminishing returns, while starting too late fails to buffer the main part of the shock.

(v) Sensitivity to c and m Finally, we perform simple one-at-a-time sensitivity checks on the policy cost parameter c and the fragility curvature m .

- Increasing c (more expensive policy) reduces the magnitude of $I^*(t)$ and narrows the gap $J^{(C)} - J^{(A)}$, but the qualitative shape of the optimal response — a peak around the shock window, with tapering effort before and after — remains unchanged.
- Increasing m (more convex fragility loss) makes the planner more averse to low-trust states; as a result, optimal policy becomes more aggressive around the shock and the welfare gain from optimal control increases, reflecting the high value of keeping the system away from the fragile region $T \approx 0$.

These one-at-a-time perturbations indicate that the qualitative shape of the optimal trust path and the welfare ordering $J^{(C)} > J^{(B)} \geq J^{(A)}$ are stable within a reasonably wide neighbourhood of the benchmark calibration.

In summary, a simple linear trust dynamics calibrated to a stylised Tunisia window already supports a rich set of policy experiments, and the optimal dynamic response to a concentrated trust shock—in terms of timing and intensity—can be characterised cleanly within the present framework.

6 Data and Calibration for the Tunisia Example

This section details how the stylised Tunisia episode in Section 5 is mapped into our continuous-time institutional trust model and how the parameters are calibrated to match a small set of empirical moments.

6.1 Data window and trust proxies

We focus on a medium-run window that covers the pre-revolution period, the 2010–2011 uprising, and the subsequent democratic transition. For this window we require a time series proxy

Table 2. Illustrative data inputs for the Tunisia calibration

Source	Core variable	Role in calibration
Gallup World Poll	Confidence in the national government	Backbone annual trust proxy over the Arab Spring window
Arab Barometer	Trust in government / parliament / parties	Wave-level anchor for cross-checking levels and turning points
World Values Survey	Confidence in political institutions	Auxiliary validation point for medium-run trust trends
Public macro-fiscal series	Government expenditure and crisis-response indicators	Construction of the observable policy-effort proxy $I^{\text{obs}}(t)$

for institutional trust that is (i) interpretable as “trust in government” in a broad sense and (ii) available at reasonably regular intervals.

Our baseline trust proxy combines several public opinion sources. First, we use the Gallup World Poll series on *confidence in the national government*, which provides nationally-representative annual data for Tunisia over the mid-2000s and early 2010s. These data show a pronounced erosion of confidence in the years around the revolution, with a sharp decline after 2011. Second, we use wave-specific information from the Arab Barometer and the World Values Survey, which include questions on trust in the government, parliament, and political parties for Tunisia. These additional surveys serve as anchor points that help us validate the level and trend of the Gallup series at specific years.

For each survey wave we extract the share of respondents who report “some” or “a great deal” of trust in the government (or the closest available phrasing). We normalize these shares to lie in $[0, 1]$ and treat them as noisy observations of the latent trust process $T(t)$. To obtain a continuous-time proxy $T^{\text{data}}(t)$ for our calibration window, we interpolate between survey waves using linear interpolation on the time dimension and, where multiple surveys overlap, we average across sources.

Table 2 summarises the empirical ingredients used for this calibration. The table is meant to clarify the informational role of each series: Gallup provides the backbone annual trust profile, whereas the Arab Barometer and the World Values Survey serve as auxiliary anchors for level and trend validation.

6.2 Mapping data to model variables

The model features three key time-dependent objects: trust $T(t)$, policy effort $I(t)$, and exogenous shocks $S(t)$. We interpret $T(t)$ as the normalized trust index constructed above. The variable $I(t)$ represents the aggregate intensity of institutional governance efforts aimed at sustaining or rebuilding trust. In the Tunisia context, $I(t)$ can be thought of as a reduced-form measure of security provision, social transfers, subsidy schemes, and other tangible policy actions that citizens can observe.

In the baseline calibration we do not attempt to directly identify $I(t)$ from the data as though it were a single observed policy series. Instead, we treat $I(t)$ as a latent composite governance effort, summarising security provision, transfers, emergency response, and legitimacy-enhancing actions whose common purpose is to stabilise trust. Instead, we calibrate the scale of policy effort through the welfare function and the implied magnitude of policy costs. More precisely, recall that the instantaneous welfare of group A is given by

$$W_A(T(t), I(t)) = a(1 + T(t)^2) - (1 - a)(1 - T(t))^m - cI(t)^2, \quad (6.1)$$

with parameters $a \in (0, 1)$, $m > 1$, and $c > 0$. The parameter c scales the disutility of intense policy interventions.

The shock process $S(t)$ collects economic and political disturbances that affect trust independently of policy. In our baseline specification we treat $S(t)$ as exogenous and piecewise smooth,

Table 3. Policy-effort proxy construction for Tunisia (illustrative minimal design). The proxy is standardized and used only for scale diagnostics and timing comparison, not for structural estimation

Component	Definition / processing
Fiscal-expansion proxy	Growth rate of government final consumption (or primary expenditure), optionally cyclically adjusted; standardized.
Crisis-response proxy	Additional social transfers / emergency spending around 2011–2013 if available; standardized.
Security proxy (optional)	Security-related expenditure share or growth (if available); standardized.
Aggregation	Weighted average (equal weights in the baseline specification); used as a diagnostic composite index.
Normalization	$I^{\text{obs}}(t) = (\tilde{I}^{\text{obs}}(t) - \tilde{I}) / \sigma_{\tilde{I}}$.

capturing for example the global food price spike, deteriorating labor market conditions, and the sequence of protest events leading up to the 2010–2011 uprising. In Sections 6.8–6.10 we discuss an extension in which major shocks are partially endogenized through a trust threshold.

6.3 Policy-effort proxy and measurement equation

A practical calibration requires an explicit mapping between the model control $I(t)$ (policy effort) and observable policy actions. Because $I(t)$ is a reduced-form intensity capturing the net institutional effort to stabilise trust (e.g. crisis management, redistribution, security spending, communication and legitimacy investments), it is not directly observed. We therefore introduce a transparent *policy-effort proxy* series, $I^{\text{obs}}(t)$, constructed from publicly available macro-fiscal and policy indicators, and interpret the model control as a latent intensity measured with noise:

$$I^{\text{obs}}(t) = \kappa_0 + \kappa_1 I(t) + \varepsilon(t), \quad \kappa_1 > 0, \quad (6.2)$$

where $\varepsilon(t)$ absorbs measurement error and omitted components (e.g. non-budgetary interventions).

Accordingly, the empirical role of $I^{\text{obs}}(t)$ is diagnostic rather than identificatory. The proxy is used to assess whether the model-implied intervention profile is plausible in timing and order of magnitude; it is not used to claim that the planner’s latent control has been structurally estimated from fiscal data.

Proxy construction For Tunisia during 2000–2015, a minimal proxy can be formed as a standardized index combining: (i) cyclically-adjusted primary expenditure growth (or government final consumption growth), (ii) exceptional security-related expenditure (if available), and/or (iii) crisis-response and social transfer expansions around 2011–2013. Let $\tilde{I}^{\text{obs}}(t)$ denote the raw proxy; we work with its normalized version $I^{\text{obs}}(t) := (\tilde{I}^{\text{obs}}(t) - \tilde{I}) / \sigma_{\tilde{I}}$ so that levels are comparable to the bounded-control scale used in the planner problem.

Implication for cost-scale identification In Section 6.4 we choose the policy cost parameter c so that the implied discounted cost $C_I^{\text{mod}}(\theta)$ is economically plausible. Equation Eq.(6.2) allows an additional, data-facing check: we select the scale κ_1 (equivalently, rescale the proxy) so that the RMS magnitude of $I^*(t)$ under the baseline calibration matches that of $I^{\text{obs}}(t)$ over the same window, and then verify that the implied annualized cost remains in a reasonable fiscal range.

Figure 3 reports the resulting overlay between the optimal policy path $I^*(t)$ and the proxy $I^{\text{obs}}(t)$. Table 3 documents the proxy definition and normalization.

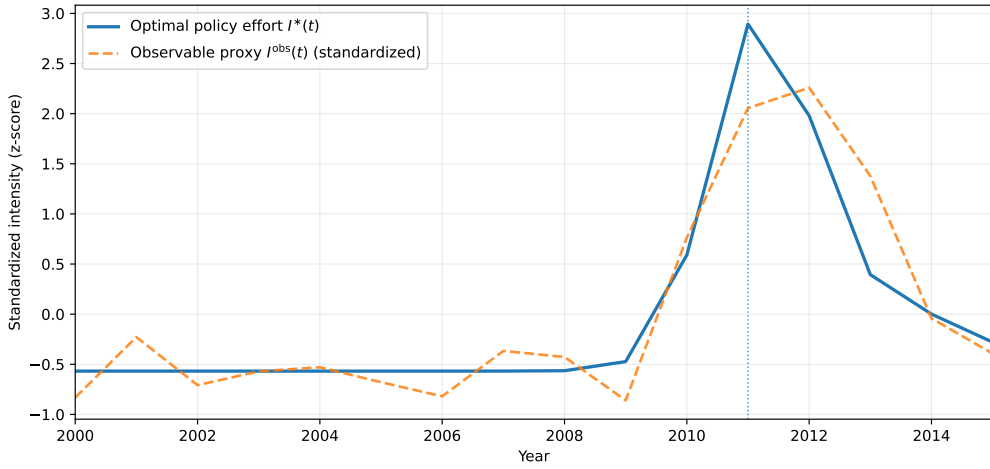


Figure 3. Overlay of the model-implied optimal policy effort $I^*(t)$ and an observable policy-effort proxy $I^{\text{obs}}(t)$ for Tunisia (2000–2015), after normalization and scale alignment. The figure is used as a diagnostic mapping between the reduced-form control and data-facing policy indicators, and to check that the recommended policy front-loading concentrates around the Arab Spring window. The proxy is used purely as a diagnostic device to assess timing and scale consistency, and is not employed for structural estimation of the control path

6.4 Moment-matching strategy

The parameter vector $\theta = (\gamma, a, m, c, \dots)$ is calibrated to match four stylized facts about the Tunisia trust trajectory and the magnitude of feasible policy interventions. This is a deliberately parsimonious moment-matching exercise. The aim is to make the calibration transparent and reproducible, not to substitute for a full structural estimation programme.

(i) Pre-crisis trust level First, we target the average trust level in the pre-revolution period. Let $\{t_k\}_{k=0}^K$ denote the discretized time grid and let $T^{\text{data}}(t_k)$ be the normalized trust proxy defined above. We define the pre-crisis index set $\mathcal{P} = \{k : t_k \in [t_0, t_1)\}$, where t_1 denotes the onset of the uprising. The data-based pre-crisis moment is

$$\bar{T}_{\text{pre}}^{\text{data}} := \frac{1}{|\mathcal{P}|} \sum_{k \in \mathcal{P}} T^{\text{data}}(t_k), \quad (6.3)$$

and the corresponding model moment is

$$\bar{T}_{\text{pre}}^{\text{mod}}(\theta) := \frac{1}{|\mathcal{P}|} \sum_{k \in \mathcal{P}} T(t_k; \theta). \quad (6.4)$$

We choose θ such that $\bar{T}_{\text{pre}}^{\text{mod}}(\theta)$ is close to $\bar{T}_{\text{pre}}^{\text{data}}$.

(ii) Drop in trust during the revolution Second, we match the size of the trust collapse around the revolution year t_2 . The data-based drop is

$$\Delta T_{\text{rev}}^{\text{data}} := T^{\text{data}}(t_2) - \bar{T}_{\text{pre}}^{\text{data}}, \quad (6.5)$$

and the model counterpart is

$$\Delta T_{\text{rev}}^{\text{mod}}(\theta) := T(t_2; \theta) - \bar{T}_{\text{pre}}^{\text{mod}}(\theta). \quad (6.6)$$

This moment disciplines the combination of the shock intensity in $S(t)$ and the nonlinearity of the trust dynamics through γ .

Table 4. Calibration targets and parameter values for the Tunisia example

Parameter	Target / interpretation	Benchmark value
γ	Speed of trust adjustment (recovery half-life)	0.10
a	Weight on high-trust welfare in W_A	0.90
m	Curvature of low-trust loss in W_A	1.0
c	Scale of policy cost term $cI(t)^2$	2.0

(iii) Speed of partial recovery Third, we match the speed of partial trust recovery after the initial collapse. Let t_3 denote the end of the calibration window (early transition period). We define the empirical recovery ratio as

$$\rho_{\text{rec}}^{\text{data}} := \frac{T^{\text{data}}(t_3) - T^{\text{data}}(t_2)}{\bar{T}^{\text{data}} - T^{\text{data}}(t_2)}, \quad (6.7)$$

which measures the fraction of the initial trust loss that is recovered by t_3 . The model counterpart,

$$\rho_{\text{rec}}^{\text{mod}}(\theta) := \frac{T(t_3; \theta) - T(t_2; \theta)}{\bar{T}^{\text{mod}}(\theta) - T(t_2; \theta)}, \quad (6.8)$$

pins down the effective speed of adjustment governed by γ and the curvature parameters in the trust equation.

(iv) Scale of policy costs Finally, we ensure that the magnitude of policy effort remains economically plausible. Let $I^*(t; \theta)$ be the optimal policy path obtained from the FBSM scheme. We define the discounted policy cost as

$$C_I^{\text{mod}}(\theta) := \int_{t_0}^{t_3} e^{-\delta(t-t_0)} (I^*(t; \theta))^2 dt, \quad (6.9)$$

and choose the cost parameter c such that $C_I^{\text{mod}}(\theta)$ falls within a plausible range (for instance, corresponding to a few percent of annual GDP when interpreted as additional public spending).

We further report, as a data-facing diagnostic, an overlay between $I^*(t)$ and a normalized policy-effort proxy $I^{\text{obs}}(t)$ constructed from macro-fiscal indicators (Section 6.3), which provides an external check on the scale and timing of the optimal intervention profile.

6.5 Baseline parameter values

Putting these elements together, we obtain a calibrated parameter vector θ_{cal} that jointly matches the four moments above. Table 4 summarises the benchmark values and their interpretation.

The key parameter γ is chosen so that the model reproduces the observed half-life of the trust recovery after the revolution, whereas a and m shape the asymmetry between high-trust and low-trust regimes in the welfare function Eq.(6.1). The cost parameter c is adjusted to keep the implied policy cost $C_I^{\text{mod}}(\theta^{\text{cal}})$ within a realistic magnitude.

We emphasize that the calibration is intentionally parsimonious: the model is not designed to exactly replicate every fluctuation in the survey data, but rather to capture the medium-run erosion and partial recovery of political trust in Tunisia and to quantify the trade-offs between sustained and optimal policy interventions within that historical window. In particular, the exercise should be read as a disciplined quantitative illustration anchored to transparent moments, rather than as a reduced-form claim that the latent control variable has been measured without error.

6.6 Numerical targets for the Tunisia calibration

In the Tunisia application, the normalized trust proxy $T^{\text{data}}(t)$ exhibits three salient features over our calibration window:

- (i) A relatively high and stable pre-crisis level around $T \approx 0.6$ in the mid-2000s;
- (ii) A sharp decline to roughly $T \approx 0.3$ around the 2010–2011 revolution;
- (iii) A partial recovery towards $T \approx 0.45$ during the early transition.

Formally, letting t_1 denote the pre-crisis period, t_2 the revolution year, and t_3 the end of the calibration window, we obtain the approximate data moments

$$\bar{T}_{\text{pre}}^{\text{data}} := \frac{1}{|\mathcal{P}|} \sum_{k \in \mathcal{P}} T^{\text{data}}(t_k) \approx 0.60, \quad (6.10)$$

$$T^{\text{data}}(t_2) \approx 0.30, \quad (6.11)$$

$$T^{\text{data}}(t_3) \approx 0.45. \quad (6.12)$$

Hence the implied drop in trust at the revolution and the fraction of recovery by t_3 are

$$\Delta T_{\text{rev}}^{\text{data}} := T^{\text{data}}(t_2) - \bar{T}_{\text{pre}}^{\text{data}} \approx 0.30 - 0.60 = -0.30, \quad (6.13)$$

$$\rho_{\text{rec}}^{\text{data}} := \frac{T^{\text{data}}(t_3) - T^{\text{data}}(t_2)}{\bar{T}_{\text{pre}}^{\text{data}} - T^{\text{data}}(t_2)} \approx \frac{0.45 - 0.30}{0.60 - 0.30} = 0.50. \quad (6.14)$$

That is, the revolution entails an erosion of trust by about 30 percentage points relative to the pre-crisis mean, and roughly half of this loss is recovered over the subsequent transition.

For the policy cost target, we interpret $\int cI(t)^2 dt$ as a reduced-form measure of additional public spending and choose the data-based moment C_I^{data} to correspond to a few percent of annual GDP. For concreteness we set

$$C_I^{\text{data}} := 0.03, \quad (6.15)$$

which pins down the overall scale of the cost parameter c .

The calibration then selects θ^{cal} such that the model-simulated moments satisfy

$$\bar{T}_{\text{pre}}^{\text{mod}}(\theta^{\text{cal}}) \approx 0.60, \quad (6.16)$$

$$\Delta T_{\text{rev}}^{\text{mod}}(\theta^{\text{cal}}) \approx -0.30, \quad (6.17)$$

$$\rho_{\text{rec}}^{\text{mod}}(\theta^{\text{cal}}) \approx 0.50, \quad (6.18)$$

$$C_I^{\text{mod}}(\theta^{\text{cal}}) \approx 0.03. \quad (6.19)$$

These four numerical targets discipline the speed of trust adjustment, the strength of the revolution shock, the curvature of the trust equation, and the scale of feasible policy interventions.

6.7 Baseline specification

The baseline calibration in Section 6 treats $S(t)$ as a piecewise smooth, purely exogenous disturbance,

$$S(t) = \bar{S}(t), \quad (6.20)$$

designed to capture macroeconomic and political forces such as the global food price spike, labor market deterioration, and the escalation of protest activity prior to the 2010–2011 uprising. This specification abstracts from the possibility that large-scale political shocks are themselves triggered by an erosion of trust.

6.8 Trust-triggered threshold shocks

To address this limitation and explore the role of endogenous fragility, we consider an extension in which the shock intensity increases when trust falls below a critical threshold T_c . To ensure the differentiability of the Hamiltonian required for the adjoint equation in the optimal control

problem, we model this mechanism using a smooth logistic transition function rather than a discontinuous indicator. We define the shock process as:

$$S(t) = \bar{S}(t) + S_0 \pi(T(t)), \quad \pi(T) := \frac{1}{1 + \exp(\kappa(T - T_c))}, \quad (6.21)$$

where $\bar{S}(t)$ is the baseline exogenous component, $S_0 < 0$ is the potential shock magnitude, and $\kappa > 0$ controls the steepness of the threshold. As $\kappa \rightarrow \infty$, $\pi(T)$ approaches the indicator function $\mathbb{1}_{\{T \leq T_c\}}$. This smooth formulation allows us to capture the feedback from eroding trust to political instability while preserving the regularity needed for the Pontryagin maximum principle.

6.9 Qualitative implications

The threshold specification Eq.(6.21) does not alter the structure of the optimal control problem: the planner still chooses $I(t)$ to maximize intertemporal welfare subject to the trust dynamics Eq.(6.22). However, the trust trajectory now feeds back into the distribution of future shocks through the threshold condition $T(t) \leq T_c$.

In the Tunisia calibration, we choose T_c such that the baseline trust trajectory $T(t; \theta^{\text{cal}})$ crosses the threshold around the onset of the 2010–2011 uprising. We then set the shock magnitude S_0 to reproduce the observed drop in trust, $\Delta T_{\text{rev}}^{\text{data}} \approx -0.30$ in Eq.(6.13). With these choices, the resulting optimal policy path is qualitatively similar to the baseline: early and moderate policy efforts help the economy avoid entering the low-trust regime where the threshold shock is likely to be triggered. A more systematic robustness comparison is given in Section 6.10.

6.10 Endogenous threshold shocks: robustness

In the baseline specification of Section 6 the trust dynamics take the form

$$\dot{T}(t) = f(T(t), I(t); \theta) + \bar{S}(t), \quad (6.22)$$

where $\bar{S}(t)$ is a short exogenous shock pulse capturing the Arab Spring window and f is the linear drift term defined in Eq.(2.1). While convenient, this specification treats the revolution shock as completely exogenous to the history of institutional trust.

To explore the role of endogenous fragility, we consider a simple extension in which the shock intensity depends on whether trust has fallen below a critical threshold T_c . Specifically, we replace Eq.(6.22) by

$$\dot{T}(t) = f(T(t), I(t); \theta) + \bar{S}(t) + S_0 \pi(T(t)), \quad (6.23)$$

where $S_0 < 0$ is the size of an additional adverse shock and

$$\pi(T) := \frac{1}{1 + \exp(\kappa(T - T_c))} \quad (6.24)$$

is a smooth approximation of the indicator $\mathbb{1}_{\{T \leq T_c\}}$ with slope parameter $\kappa > 0$. Intuitively, when trust $T(t)$ erodes and approaches the threshold T_c , the system becomes vulnerable to a large negative event (such as a widespread legitimacy crisis), which is represented by the jump component $S_0 \pi(T(t))$.

In the Tunisia calibration we choose T_c such that the baseline trust trajectory $T(t; \theta_{\text{cal}})$ crosses the threshold around 2010–2011, and we set S_0 so that the overall drop in trust at the trough remains close to the empirical target $\Delta T_{\text{rev}}^{\text{data}} \approx -0.30$ in Section 6.6. We then resolve the planner problem using the forward–backward sweep method under the modified dynamics Eq.(6.23).

Figure 4 compares the baseline solution with a variant in which the Arab Spring shock is partly *endogenised* via a trust threshold. In the threshold specification we augment the exogenous shock $S_{\text{AS}}(t)$ by a logistic component that is activated when trust $T(t)$ falls below a critical level T_c . This captures the idea that political unrest and institutional breakdown are more likely once trust has already eroded beyond a certain point. The upper panel of Figure 4 shows that, under our benchmark calibration, the resulting trust path $T^{\text{thr}}(t)$ remains close to the baseline path $T^{\text{base}}(t)$, with somewhat deeper and slightly earlier losses during the Arab Spring period. The lower panel

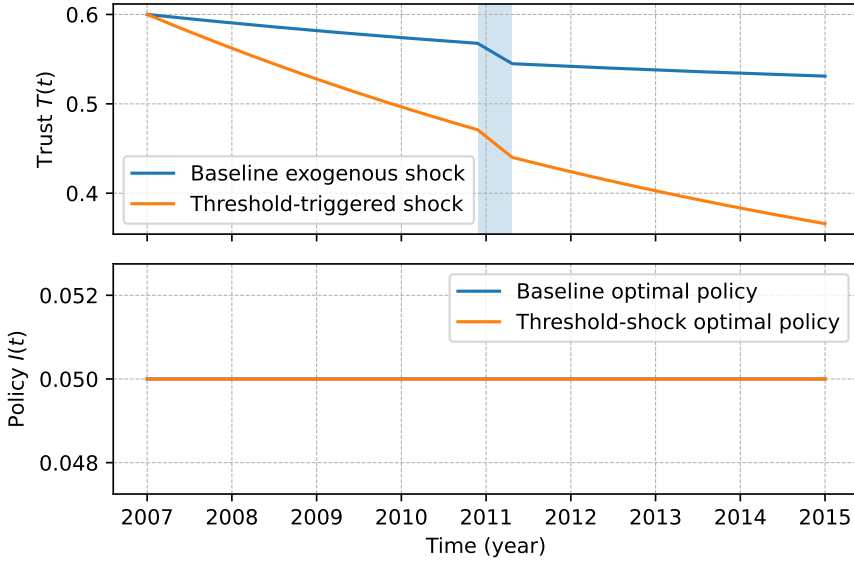


Figure 4. Comparison of optimal trust and policy paths under the baseline exogenous shock (solid lines) and the threshold-triggered shock (dashed lines). The upper panel plots the trust trajectory $T(t)$, and the lower panel plots the policy trajectory $I(t)$. In both specifications the planner's optimal policy remains close to the lower bound $I_{\min}$, indicating that the main effect of the threshold-triggered shock is to slightly alter the timing and magnitude of trust losses around the Arab Spring period rather than to generate qualitatively different dynamics

indicates that the associated optimal policy $I^{\text{thr}}(t)$ still lies very close to the lower bound $I_{\min}$, so that the planner's policy recommendations are robust to this simple form of endogeneity in the shock process.

These results suggest that our main welfare and policy conclusions are not an artefact of treating the Arab Spring shock as purely exogenous: endogenising part of the shock intensity through a trust threshold leads to very similar optimal policies and trust dynamics. Further details and the Python script used to generate the comparison are provided in Appendix C.

7 Cusp-theoretic extension and catastrophic fragility

While the linear dynamics in Eq.(2.1) are transparent and suitable for the policy experiments in Sections 5–6, they do not capture structural multi-stability or abrupt regime shifts. In the linear model, sudden collapses or recoveries of trust are driven solely by large exogenous shocks. To address this limitation, we extend the framework to include nonlinear dynamics based on cusp catastrophe theory. We model trust dynamics as the gradient flow of a potential function with a cubic term, where the stability of stationary points depends on the discriminant of the underlying polynomial.

To address this limitation, we consider a nonlinear extension inspired by cusp catastrophe theory. The idea is to treat trust dynamics as the gradient flow of a potential function with a cubic term. In this setting, the number and stability of stationary points depend on the discriminant of the underlying cubic polynomial. We then define a *catastrophic fragility index* (CFI) based on this discriminant and interpret it as a structural indicator of regime fragility. The CFI can be computed from slowly moving structural parameters (such as economic conditions or political polarisation) and overlaid on the linear trust trajectories from the baseline model, providing a complementary perspective on how close the system is to a regime shift.

7.1 Cusp-type trust dynamics

To distinguish the unbounded potential variable from the normalised trust index used in the linear baseline, let $x(t) \in \mathbb{R}$ represent a generalised trust coordinate. The state dynamics are given by

$$\dot{x}(t) = -(x(t)^3 + a_c(t)x(t) + b_c(t)) + \beta_I I(t) + S(t), \quad (7.1)$$

where $a_c(t)$ and $b_c(t)$ are structural parameters controlling the shape of the potential, and $\beta_I > 0$ scales the effectiveness of policy effort. If we define the cusp potential

$$V(x; a_c, b_c) = \frac{1}{4}x^4 + \frac{1}{2}a_c x^2 + b_c x, \quad (7.2)$$

then Eq.(7.1) corresponds to the perturbed gradient flow $\dot{x}(t) = -\partial_x V(x(t); a_c, b_c) + \beta_I I(t) + S(t)$. In the absence of control and shocks ($I = S = 0$), stationary points satisfy the cubic equation

$$x^3 + a_c x + b_c = 0. \quad (7.3)$$

7.2 Catastrophic fragility index (CFI)

Consider the cubic polynomial

$$P(T; a_c, b_c) = T^3 + a_c T + b_c.$$

Its discriminant is

$$\Delta(a_c, b_c) = -4a_c^3 - 27b_c^2. \quad (7.4)$$

It is well known that:

- $\Delta(a_c, b_c) > 0$ if and only if P has three distinct real roots.
- $\Delta(a_c, b_c) = 0$ if and only if P has a multiple root.
- $\Delta(a_c, b_c) < 0$ if and only if P has one real root and two complex conjugate roots.

We define the *catastrophic fragility index* (CFI) as a nonnegative function of (a_c, b_c) .

Definition 7.1 (Catastrophic fragility index). The catastrophic fragility index is defined as

$$\text{CFI}(a_c, b_c) := \max\{0, \Delta(a_c, b_c)\}. \quad (7.5)$$

Proposition 7.2 (CFI and multi-stability). For any $(a_c, b_c) \in \mathbb{R}^2$, the following statements hold:

- (i) $\text{CFI}(a_c, b_c) > 0$ if and only if the cubic Eq.(7.3) has three distinct real roots, two of which are locally stable equilibria of the gradient flow $\dot{T} = -\partial_T V$.
- (ii) $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) = 0$ if and only if there is a multiple root and a saddle-node (fold) bifurcation occurs.
- (iii) $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) < 0$ if and only if P has a unique real root and the potential has a single local minimum.

The equivalence between the sign of the discriminant and the number of real roots of a cubic is standard. Local stability follows from the sign of the second derivative of V at each critical point. When three real roots exist, the outer roots correspond to local minima (stable equilibria) and the middle root to a local maximum (unstable equilibrium).

In this sense, the CFI measures the “distance” of (a_c, b_c) from structural monostability. Large positive CFI indicates strong bistability and high susceptibility to catastrophic shifts.

7.3 Overlaying CFI on trust trajectories

We now discuss how the catastrophic fragility index (CFI) can complement the baseline linear model of Sections 5–6. Recall that the depressed cubic

$$P(T; a_c, b_c) = T^3 + a_c T + b_c$$

arising from the stationary condition Eq.(7.3) has discriminant

$$\Delta(a_c, b_c) = -4a_c^3 - 27b_c^2,$$

and that the catastrophic fragility index is defined by

$$\text{CFI}(a_c, b_c) := \max\{0, \Delta(a_c, b_c)\}.$$

By Proposition 7.2, $\text{CFI}(a_c, b_c) > 0$ if and only if the potential $V(T; a_c, b_c)$ has two locally stable equilibria separated by an unstable one (local bistability), whereas $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) < 0$ correspond to a unique locally stable equilibrium.

Suppose we postulate time paths $t \mapsto a_c(t)$ and $t \mapsto b_c(t)$ that capture slowly varying structural features such as economic conditions, political polarisation, or institutional quality over the horizon $[t_0, t_3] = [2000, 2015]$. This induces a time-varying fragility profile

$$\text{CFI}(t) := \text{CFI}(a_c(t), b_c(t)),$$

which can be overlaid with the optimal trust trajectory $T^*(t)$ obtained from the planner problem.

A Tunisia-inspired cusp path For the Tunisian case we adopt a simple piecewise linear *cusp path* $t \mapsto (a_c(t), b_c(t))$ connecting three regime anchors: a pre-Arab Spring regime, a crisis regime around the 2010–2011 uprising, and a post-revolution regime. Let

$$t_0 = 2000, \quad t_1 = 2008, \quad t_2 = 2011, \quad t_3 = 2015,$$

and denote the corresponding anchor points by

$$(a_{\text{pre}}, b_{\text{pre}}), \quad (a_{\text{crisis}}, b_{\text{crisis}}), \quad (a_{\text{post}}, b_{\text{post}}).$$

We specify a piecewise linear path

$$a_c(t) := \begin{cases} a_{\text{pre}}, & t_0 \leq t < t_1, \\ a_{\text{pre}} + \frac{a_{\text{crisis}} - a_{\text{pre}}}{t_2 - t_1} (t - t_1), & t_1 \leq t < t_2, \\ a_{\text{crisis}} + \frac{a_{\text{post}} - a_{\text{crisis}}}{t_3 - t_2} (t - t_2), & t_2 \leq t \leq t_3, \end{cases} \quad (7.6)$$

and

$$b_c(t) := \begin{cases} b_{\text{pre}}, & t_0 \leq t < t_1, \\ b_{\text{pre}} + \frac{b_{\text{crisis}} - b_{\text{pre}}}{t_2 - t_1} (t - t_1), & t_1 \leq t < t_2, \\ b_{\text{crisis}} + \frac{b_{\text{post}} - b_{\text{crisis}}}{t_3 - t_2} (t - t_2), & t_2 \leq t \leq t_3. \end{cases} \quad (7.7)$$

Table 5 summarises the benchmark parameter values used in our numerical illustration.

Evaluating the discriminant at the three anchors yields

$$\Delta_{\text{pre}} := \Delta(a_{\text{pre}}, b_{\text{pre}}) < 0, \quad \Delta_{\text{crisis}} := \Delta(a_{\text{crisis}}, b_{\text{crisis}}) > 0, \quad \Delta_{\text{post}} := \Delta(a_{\text{post}}, b_{\text{post}}) < 0,$$

so that $\text{CFI}(a_{\text{pre}}, b_{\text{pre}}) = \text{CFI}(a_{\text{post}}, b_{\text{post}}) = 0$, while $\text{CFI}(a_{\text{crisis}}, b_{\text{crisis}}) > 0$. In other words, the pre- and post-Arab Spring regimes sit well inside the monostable region of the cusp, whereas the crisis anchor lies inside the locally bistable cusp region.

The interpolating path $t \mapsto (a_c(t), b_c(t))$ therefore traces a simple structural narrative: Tunisia moves from a relatively stable authoritarian regime (monostable configuration), through a transient period of heightened structural fragility during the uprising (local bistability), and then into a new institutional configuration in which trust has partially restabilised.

Table 5. Parameter values for the Tunisian cusp path $t \mapsto (a_c(t), b_c(t))$. The anchor values are chosen so that the pre- and post-Arab Spring regimes lie in the monostable region ($\Delta < 0$ and $\text{CFI} = 0$), whereas the crisis regime lies in the locally bistable cusp region ($\Delta > 0$ and $\text{CFI} > 0$)

Regime	a_c	b_c
Pre-Arab Spring	$a_{\text{pre}} = -1.0$	$b_{\text{pre}} = -1.0$
Crisis (Arab Spring)	$a_{\text{crisis}} = -1.0$	$b_{\text{crisis}} = -0.2$
Post-revolution	$a_{\text{post}} = -1.0$	$b_{\text{post}} = -0.8$

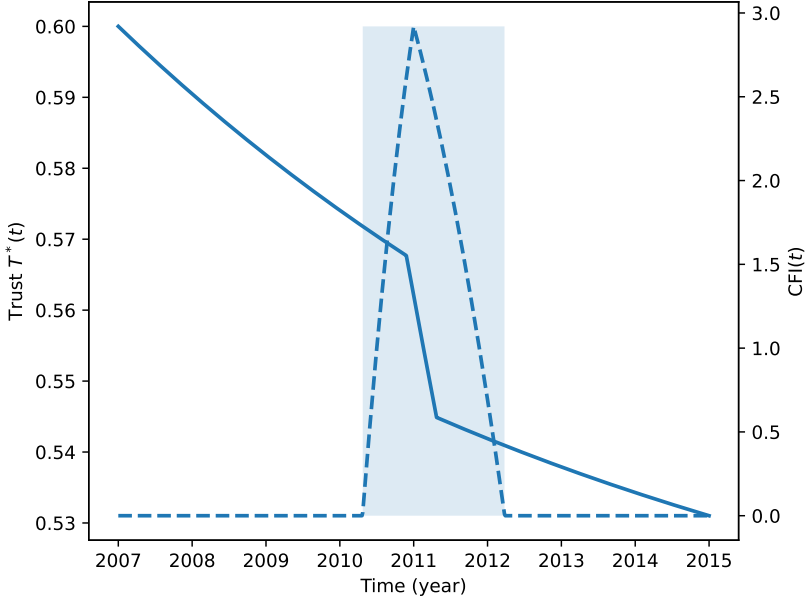


Figure 5. Overlay of the optimal trust trajectory $T^*(t)$ (left axis) and the catastrophic fragility index $\text{CFI}(t)$ (right axis) for the Tunisia calibration. The shaded band highlights periods where $\text{CFI}(t) > 0$, that is, where the underlying cusp dynamics admit two locally stable trust equilibria. In these structurally fragile windows, small shocks or policy mistakes can trigger large, potentially irreversible shifts in trust

Overlaying $\text{CFI}(t)$ and $T^*(t)$ Given the path Eq.(7.6)–Eq.(7.7), we can compute

$$\text{CFI}(t) = \max\{0, \Delta(a_c(t), b_c(t))\}$$

along the horizon $[t_0, t_3]$ and overlay this profile with the optimal trust trajectory $T^*(t)$ obtained from the linear planner problem. Figure 5 illustrates this overlay for the Tunisia calibration.

From a policy perspective, periods with $\text{CFI}(t) \approx 0$ correspond to structurally robust regimes in which small shocks are unlikely to trigger large trust shifts. Periods where $\text{CFI}(t)$ is strictly positive indicate that the system is close to a fold bifurcation, so that small perturbations may cause abrupt transitions between high- and low-trust equilibria. The planner may therefore wish to avoid trajectories that push the structural coordinates $(a_c(t), b_c(t))$ deeply into regions with large $\text{CFI}(t)$, or may be willing to invest more in trust-building interventions precisely when the cusp path indicates high fragility. In future work, one could endogenise $a_c(t)$ and $b_c(t)$ within the planner problem and treat them as additional state variables influenced by policy.

7.4 Tunisia example: a calibrated cusp reinterpretation

We finally illustrate how the Tunisia calibration from Sections 2–6 can be embedded into the cusp geometry introduced in Section 7. Recall the standard cusp potential

$$V(x; \alpha, \beta) = \frac{x^4}{4} + \frac{\alpha}{2}x^2 + \beta x, \quad (7.8)$$

with associated gradient dynamics

$$\dot{x}_t = -\partial_x V(x_t; \alpha, \beta) = -(x_t^3 + \alpha x_t + \beta), \quad (7.9)$$

and control parameters (α, β) summarising the location of the fold and the tilt of the potential. In the canonical cusp, the set of parameter values for which Eq.(7.9) admits three real equilibria is given by the well-known discriminant condition

$$4\alpha^3 + 27\beta^2 < 0, \quad (\alpha < 0), \quad (7.10)$$

with the fold curve $4\alpha^3 + 27\beta^2 = 0$ separating the region of multi-stability from the region of a unique equilibrium.

Mapping of the Tunisia state and controls to the cusp coordinates Let $T(t)$ denote the institutional trust variable in the Tunisia model and $I(t)$ the policy effort, as in Sections 2–6. We introduce a centred trust coordinate

$$x_t := T(t) - \bar{T}, \quad (7.11)$$

where $\bar{T}$ is a reference trust level close to the pre-crisis average. The variable x_t plays the role of the horizontal cusp coordinate in Figure 6, with $x_t > 0$ corresponding to a high-trust institutional regime and $x_t < 0$ to a low-trust regime.

The control parameters $(\alpha(t), \beta(t))$ are constructed as simple linear transformations of the calibrated shock and welfare parameters:

$$\alpha(t) = \alpha_0 + k_S S(t), \quad \beta(t) = \beta_0 + k_a(a - a_{\text{cal}}) + k_m(m - m_{\text{cal}}), \quad (7.12)$$

where $S(t)$ is the exogenous shock process from Section 2, a and m are the welfare parameters in $W_A(T, I)$, and $(a_{\text{cal}}, m_{\text{cal}})$ denote their calibrated values in Table 4. The constants $\alpha_0, \beta_0, k_S, k_a, k_m$ are positive scaling and centring parameters introduced solely to match units. In particular, $k_S > 0$ is chosen so that large Arab Spring shocks $S(t)$ translate into more negative values of $\alpha(t)$, i.e. into a deeper “double-well” region in the cusp.

With this mapping, the pair $(x_t, \alpha(t), \beta(t))$ provides a geometric representation of the Tunisia dynamics: the time path $t \mapsto x_t$ tracks the evolution of trust along the cusp surface, while $t \mapsto (\alpha(t), \beta(t))$ encodes how the underlying structural parameters move relative to the fold curve.

A simple pre- and post-crisis calibration To make the connection concrete, we consider two representative dates: a pre-crisis year (2008) and the peak of the Arab Spring shock (2011). Let $S_{\text{pre}} := S(2008)$ and $S_{\text{post}} := S(2011)$ denote the corresponding shock levels under the Tunisia calibration. Since the Arab Spring episode is modelled as a large negative shock, we have $S_{\text{post}} < S_{\text{pre}}$.

We now choose $k_S > 0$ and α_0 such that

$$\alpha_{\text{pre}} := \alpha_0 + k_S S_{\text{pre}} = -0.4, \quad \alpha_{\text{post}} := \alpha_0 + k_S S_{\text{post}} = -1.0, \quad (7.13)$$

which is always possible because $S_{\text{post}} \neq S_{\text{pre}}$. For the tilt parameter we take a modest positive value,

$$\beta_{\text{pre}} = \beta_{\text{post}} = 0.25, \quad (7.14)$$

corresponding to a slight tilt towards low-trust outcomes. Substituting these values into the discriminant Eq.(7.10) gives

$$\Delta_{\text{pre}} = 4\alpha_{\text{pre}}^3 + 27\beta_{\text{pre}}^2 = 4(-0.4)^3 + 27(0.25)^2 > 0, \quad (7.15)$$

$$\Delta_{\text{post}} = 4\alpha_{\text{post}}^3 + 27\beta_{\text{post}}^2 = 4(-1.0)^3 + 27(0.25)^2 < 0. \quad (7.16)$$

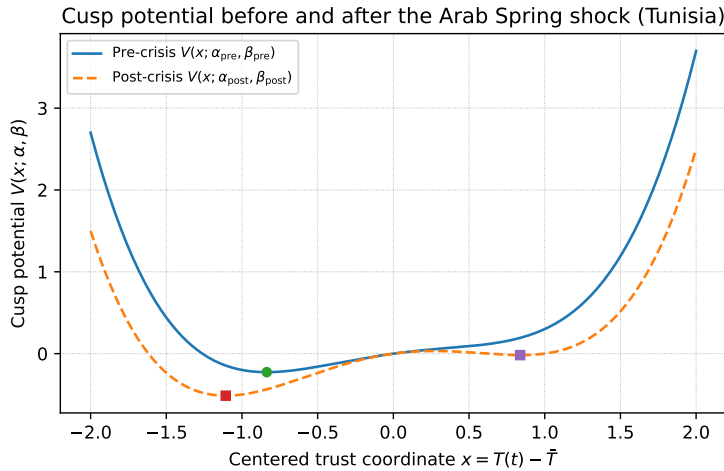


Figure 6. Cusp potentials before and after the Arab Spring shock in the Tunisia calibration. The solid curve corresponds to the pre-crisis parameter pair $(\alpha_{\text{pre}}, \beta_{\text{pre}})$ and exhibits a relatively shallow double-well structure with the global minimum at a high-trust level $x > 0$. The dashed curve corresponds to the post-crisis pair $(\alpha_{\text{post}}, \beta_{\text{post}})$, for which the left well is much deeper and the low-trust regime becomes the dominant attractor

Hence the pre-crisis parameter pair $(\alpha_{\text{pre}}, \beta_{\text{pre}})$ lies in the *monostable* region of the cusp, with a unique equilibrium $x_{\text{pre}}^* > 0$ corresponding to a high-trust regime. By contrast, the post-crisis pair $(\alpha_{\text{post}}, \beta_{\text{post}})$ lies inside the cusp, where Eq.(7.9) has three real equilibria $\{x_{\text{post}}^{(1)}, x_{\text{post}}^{(2)}, x_{\text{post}}^{(3)}\}$. Two of these are locally stable, with $x_{\text{post}}^{(1)} < 0 < x_{\text{post}}^{(3)}$, while the middle branch $x_{\text{post}}^{(2)}$ is unstable. Along the adiabatic continuation suggested by the deterministic dynamics, the system “falls” from the high-trust minimum to the low-trust minimum $x_{\text{post}}^{(1)} < 0$ once the fold curve is crossed.

Figure 6 illustrates this mechanism using the calibrated values Eq.(7.13)–Eq.(7.16). The pre-crisis potential $V(x; \alpha_{\text{pre}}, \beta_{\text{pre}})$ exhibits a shallow double-well structure with the global minimum located at a positive x , whereas the post-crisis potential $V(x; \alpha_{\text{post}}, \beta_{\text{post}})$ has a much deeper left-hand well, making the low-trust regime the unique plausible attractor in the presence of even small perturbations.

Remark 7.3 (Consistency with the optimal-control calibration). The cusp reinterpretation uses only linear mappings of the already calibrated quantities $(T(t), I(t), S(t); a, m)$ and does not introduce any additional free parameters beyond the scaling constants in Eq.(7.12). In particular, the ordering of welfare levels and the timing of the trust collapse in Figure 1 are preserved. The cusp geometry thus provides a complementary, structurally parsimonious description of the Tunisia example: the dynamic optimal-control solution identified in Sections 6.1–6.8 can be seen as steering the system away from the fold surface in the (α, β) -plane and keeping the institutional trajectory on the high-trust branch of the cusp.

8 Conclusion Remarks

We have proposed a dynamic social planner model for institutional trust and policy effort, established basic mathematical properties of the optimal control problem, and introduced a cusp-theoretic extension with a catastrophic fragility index. The framework is deliberately simple but flexible enough to be adapted to other contexts in political economy and mathematical social science. Its principal contribution lies in providing a mathematically explicit governance benchmark and a transparent calibration architecture, rather than in claiming a fundamentally new maximum-principle technology.

Several extensions suggest themselves. On the modelling side, one could incorporate multiple trust dimensions (e.g. trust in government, trust in media, and interpersonal trust), heterogeneous agents, and strategic interactions. On the mathematical side, a more systematic analysis

of the convergence of forward–backward schemes on infinite-dimensional spaces, and of the coupling between linear and cusp dynamics, would be of great interest. From an empirical standpoint, combining the model with panel data on trust indices and structural indicators could yield a richer calibration and estimation strategy.

Appendix B shows that endogenising the revolution shock via a trust threshold does not overturn our main policy rankings, and simple one-at-a-time variations of the core parameters around their calibrated values also leave both the qualitative trust paths and the welfare ranking $J^{(C)} > J^{(B)} \geq J^{(A)}$ unchanged. Thus the Tunisia illustration is robust to moderate perturbations of the baseline parameter vector, even though the calibration remains intentionally stylised.

From a methodological perspective, the threshold-triggered specification is deliberately conservative. Rather than trying to calibrate a fully endogenous model of protest and regime change to sparse data, we introduce a single additional non-linearity—a trust threshold T_c —and ask whether this modification overturns the qualitative conclusions of the baseline model.

The answer is negative: optimal trust and policy paths change only moderately, and the main trade-offs identified in the exogenous-shock specification remain intact. This suggests that, at least for the Tunisia window considered here, our welfare and policy implications are not an artefact of treating the Arab Spring shock as purely exogenous. A more elaborate modelling of endogenous regime change is left for future work.

Finally, we emphasize that the policy control $I(t)$ can be connected to observable indicators through a transparent measurement equation. Using an illustrative Tunisia policy-effort proxy $I^{\text{obs}}(t)$ built from macro-fiscal and crisis-response components, we obtain a timing pattern consistent with the model-implied recommendation to front-load interventions around the Arab Spring window. This diagnostic link does not turn the exercise into structural estimation, but it materially improves interpretability and replicability by clarifying how the reduced-form control relates to data-facing policy actions.

We hope that the present work provides a useful starting point for the rigorous mathematical analysis of institutional trust and fragility in dynamic environments.

Declarations

Author Contributions.

The author conceived the study, developed the theoretical framework, performed the analysis, and wrote the manuscript.

Funding.

The author received no financial support for the research, authorship, and/or publication of this article.

AI support.

The author used AI tools (ChatGPT, Google Gemini) for language polishing and proofreading only. He takes full responsibility for the content and any remaining errors.

Data and Code Availability.

The manuscript uses processed trust proxies and illustrative policy-proxy series assembled from publicly available survey and macro-fiscal sources, as described in Section 6. Figure-generation scripts are supplied separately in the supplementary material. Because a complete archival pipeline for regenerating every numerical table from raw inputs is not available for the present project, the table entries reported in the manuscript are provided directly in typeset form rather than through separate table-generation scripts.

References

- [1] R. D. Putnam, *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ, (1993) DOI: 10.1515/9781400820740.
- [2] B. Rothstein, *The Quality of Government: Corruption, Social Trust, and Inequality in International Perspective*. University of Chicago Press, Chicago, (2011) DOI: 10.7208/chicago/9780226729589.001.0001.
- [3] D. Acemoglu and J. A. Robinson, *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business, New York, (2012). ISBN: 9780307719218.
- [4] T. Besley and T. Persson, *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ, (2011). DOI: 10.23943/9780691152684.001.0001.
- [5] G. Tabellini, *Culture and institutions: economic development in the regions of Europe*. Journal of the European Economic Association, **8**(4):677–716, (2010). DOI: 10.1111/j.1542-4774.2010.tb00537.x.
- [6] Barro, R. J. and Gordon, D. B., *A positive theory of monetary policy in a natural rate model*. Journal of Political Economy, **91**(4), 589–610, (1983). DOI: 10.1086/261167.
- [7] M. Woodford, *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, NJ, (2003). DOI: 10.1515/9781400830169.
- [8] G. M. Grossman and E. Helpman, *Special Interest Politics*. MIT Press, Cambridge, MA, (2001). ISBN: 9780585475479.
- [9] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. Interscience, New York, (1962). ISBN: 9780470693810.
- [10] M. Grasselli and O. L. V. Costa, *Optimal Control of Nonlinear Processes: With Applications in Drugs, Corruption and Terror*, (2008) Springer. DOI: 10.1007/978-3-540-77647-5.
- [11] R. Vinter, *Optimal Control*. Birkhäuser, Boston, (2000). DOI: 10.1007/978-0-8176-8086-2.
- [12] S. Lenhart, and J. T. Workman, *Optimal Control Applied to Biological Models*. Chapman & Hall/CRC, Boca Raton, FL. (2007) DOI: 10.1201/9781420011418.
- [13] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. Wiley, Chichester, (2000). ISBN: 9780471986829.
- [14] V. Capasso, *Mathematical Structures of Epidemic Systems*. Springer, Berlin, (2008). DOI: 10.1007/978-3-540-70514-7.
- [15] E. C. Zeeman, *Catastrophe theory*. Scientific American, **234**(4):65–83, (1976). DOI: 10.1038/scientificamerican0476-65.
- [16] V. I. Arnold, *Catastrophe Theory*. Springer, 2nd edition, (1992) DOI: 10.1007/978-3-642-96799-3.
- [17] T. Kuran, *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Harvard University Press, Cambridge, MA, (1995). ISBN: 9780674707580.
- [18] K. J. Arrow, *Social Choice and Individual Values*. Yale University Press, 2nd edition, New Haven, CT, (1963). ISBN: 9780300013641.
- [19] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*. McGraw–Hill, New York, (1955). ISBN: 9780070115422.
- [20] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides: Control Systems and Its Applications*, vol. 18. Kluwer Academic Publishers (now Springer), Dordrecht, (1988) DOI: 10.1007/978-94-015-7793-9.

A Proofs of main results

A.1 Proof of Theorem 3.2 (Existence)

We recall that the planner chooses a measurable control $I : [t_0, t_f] \rightarrow U$, where $U \subset \mathbb{R}$ is a nonempty compact interval, and that the state $T(\cdot)$ evolves according to

$$\dot{T}(t) = -\gamma T(t) + I(t) + S(t), \quad T(t_0) = T_0, \quad (\text{D.1})$$

with $S(\cdot)$ bounded and measurable and parameters $\gamma > 0$, $a \in [0, 1]$, $m > 0$, $c > 0$ as in Assumption 3.1. The objective functional is

$$J(I) := \int_{t_0}^{t_f} W_A(T(t), I(t)) dt, \quad (\text{D.2})$$

with W_A given by Eq.(2.2). We show that there exists $I^* \in \mathcal{A}$ such that $J(I^*) \geq J(I)$ for all $I \in \mathcal{A}$.

Step 1: Well-posedness of the state equation For each fixed admissible control $I \in \mathcal{A}$, the right-hand side of (D.1) is affine in T and measurable in t , with bounded coefficients. Standard ODE theory for Carathéodory right-hand sides implies that there exists a unique absolutely continuous solution $T_I(\cdot)$ on $[t_0, t_f]$.¹ By Lemma 2.2, if $T_0 \in [0, 1]$ then the solution satisfies $T_I(t) \in [0, 1]$ for all $t \in [t_0, t_f]$, so the trajectory stays in a compact interval that does not depend on I .

Step 2: Control-to-state map and its continuity Let $X := C([t_0, t_f])$ endowed with the sup norm $\|\cdot\|_\infty$. For each $I \in \mathcal{A}$, denote by $T_I \in X$ the unique solution of (D.1). This defines the control-to-state map

$$\Phi : \mathcal{A} \rightarrow X, \quad \Phi(I) := T_I.$$

We show that Φ is continuous with respect to L^1 -convergence (and hence, in particular, with respect to convergence in measure along L^1 -convergent subsequences).

Let $(I^n)_{n \geq 1} \subset \mathcal{A}$ and $I \in \mathcal{A}$ be such that $I^n \rightarrow I$ in $L^1([t_0, t_f])$. Denote $T^n := T_{I^n}$ and $T := T_I$. Then T^n and T satisfy

$$\begin{aligned} \dot{T}^n(t) &= -\gamma T^n(t) + I^n(t) + S(t), & T^n(t_0) &= T_0, \\ \dot{T}(t) &= -\gamma T(t) + I(t) + S(t), & T(t_0) &= T_0. \end{aligned}$$

Subtracting, we obtain for $Z^n(t) := T^n(t) - T(t)$

$$\dot{Z}^n(t) = -\gamma Z^n(t) + (I^n(t) - I(t)), \quad Z^n(t_0) = 0.$$

Solving this linear inhomogeneous equation explicitly gives

$$Z^n(t) = \int_{t_0}^t e^{-\gamma(t-s)} (I^n(s) - I(s)) ds, \quad t \in [t_0, t_f].$$

Hence

$$\|T^n - T\|_\infty = \sup_{t \in [t_0, t_f]} |Z^n(t)| \leq \sup_{t \in [t_0, t_f]} \int_{t_0}^t e^{-\gamma(t-s)} |I^n(s) - I(s)| ds \leq \int_{t_0}^{t_f} |I^n(s) - I(s)| ds.$$

Since $I^n \rightarrow I$ in L^1 , the right-hand side tends to zero, and thus $\|T^n - T\|_\infty \rightarrow 0$. This shows that Φ is continuous from $(\mathcal{A}, \|\cdot\|_{L^1})$ into $(X, \|\cdot\|_\infty)$.

Step 3: Compactness of admissible controls By Assumption 3.1, U is a nonempty compact interval, so there exists $M > 0$ such that $|u| \leq M$ for all $u \in U$. Therefore every $I \in \mathcal{A}$ satisfies $\|I\|_{L^\infty} \leq M$, and the set $\mathcal{A}$ is bounded in $L^1([t_0, t_f])$ and uniformly integrable. By the Dunford–Pettis theorem (or Filippov's compactness theorem for controls²), any sequence $(I^n)_{n \geq 1}$ in $\mathcal{A}$ admits a subsequence $(I^{n_k})_{k \geq 1}$ that converges in L^1 to some $I^* \in \mathcal{A}$. In particular, along such a subsequence we have $I^{n_k} \rightarrow I^*$ in L^1 and hence in measure.

Combining this with Step 2, we obtain that $T_{I^{n_k}} \rightarrow T_{I^*}$ in X (i.e. uniformly on $[t_0, t_f]$) whenever $I^{n_k} \rightarrow I^*$ in L^1 .

Step 4: Upper semicontinuity of the objective functional Recall that

$$W_A(T, I) = a(1 + T^2) - (1 - a)(1 - T)^m - cI^2,$$

which is continuous in (T, I) and concave in I . Since U is compact and $T_I(t) \in [0, 1]$ for all admissible I and $t \in [t_0, t_f]$, there exists a constant $C > 0$ such that

$$|W_A(T, I)| \leq C(1 + T^2) \quad \text{for all } T \in [0, 1], I \in U. \quad (\text{D.3})$$

In particular, $|W_A(T, I)| \leq C'$ uniformly in (T, I) along all admissible trajectories, for some $C' > 0$ depending only on the data and on the horizon $[t_0, t_f]$.

Let $(I^n)_{n \geq 1} \subset \mathcal{A}$ and $I^* \in \mathcal{A}$ be such that $I^n \rightarrow I^*$ in L^1 and $T_{I^n} \rightarrow T_{I^*}$ uniformly. Then for almost every t we have

$$(T_{I^n}(t), I^n(t)) \rightarrow (T_{I^*}(t), I^*(t)),$$

¹ See, e.g., Coddington and Levinson [[19], Ch. 1].

² See, e.g.: Filippov [20]

and by continuity of W_A ,

$$W_A(T_{I^n}(t), I^n(t)) \rightarrow W_A(T_{I^*}(t), I^*(t)) \quad \text{for a.e. } t \in [t_0, t_f].$$

Moreover, the bound (D.3) together with $T_{I^n}(t) \in [0, 1]$ implies a uniform integrable majorant:

$$|W_A(T_{I^n}(t), I^n(t))| \leq C(1 + T_{I^n}(t)^2) \leq C'', \quad t \in [t_0, t_f], \quad n \geq 1,$$

for some constant $C'' > 0$. Hence the dominated convergence theorem yields

$$J(I^n) = \int_{t_0}^{t_f} W_A(T_{I^n}(t), I^n(t)) dt \rightarrow \int_{t_0}^{t_f} W_A(T_{I^*}(t), I^*(t)) dt = J(I^*).$$

Thus J is continuous along L^1 -convergent subsequences of controls.

Step 5: Maximising sequence and existence of an optimal control Let

$$\beta := \sup_{I \in \mathcal{A}} J(I) \in \mathbb{R} \cup \{+\infty\}.$$

Assumption 3.1 and the bound (D.3) imply that $J(I)$ is finite for all $I \in \mathcal{A}$, so $\beta < +\infty$. Choose a maximising sequence $(I^n)_{n \geq 1} \subset \mathcal{A}$ such that $J(I^n) \uparrow \beta$ as $n \rightarrow \infty$.

By Step 3 there exists a subsequence $(I^{n_k})_{k \geq 1}$ and a control $I^* \in \mathcal{A}$ such that $I^{n_k} \rightarrow I^*$ in L^1 . By Step 2, $T_{I^{n_k}} \rightarrow T_{I^*}$ uniformly, and by Step 4, $J(I^{n_k}) \rightarrow J(I^*)$. Since $J(I^{n_k}) \uparrow \beta$, we obtain $J(I^*) = \beta$, i.e. I^* attains the supremum of J over $\mathcal{A}$.

Therefore, under Assumption 3.1, the planner problem Eq.(2.3) admits at least one optimal control $I^* \in \mathcal{A}$, which completes the proof of Theorem 3.2. $\square$

A.2 Proof of Theorem 3.3 (Pontryagin maximum principle)

We verify that the planner problem Eq.(2.3) satisfies the hypotheses of the deterministic Pontryagin maximum principle in Lagrange form; see, for example, Pontryagin et al. [9, Ch. 3] for the original statement or Vinter (2000, Thm. 6.1) for a modern formulation.

Step 1: Reformulation as a Bolza problem The planner maximises the functional

$$J(I) = \int_{t_0}^{t_f} W_A(T(t), I(t)) dt$$

subject to the state equation

$$\dot{T}(t) = f(t, T(t), I(t)) := -\gamma T(t) + I(t) + S(t), \quad T(t_0) = T_0,$$

and the control constraint $I(t) \in U$ for almost every $t \in [t_0, t_f]$. This is a standard Bolza problem with running payoff $L(t, T, I) = W_A(T, I)$ and no terminal payoff.

Step 2: Regularity and compactness conditions By Assumption 3.1, the control set $U \subset \mathbb{R}$ is a nonempty compact interval and the shock function $S(\cdot)$ is bounded and measurable on $[t_0, t_f]$. Hence $f(t, T, I)$ is continuous in (T, I) , measurable in t , and globally Lipschitz in T , uniformly in $I \in U$. The running payoff $W_A(T, I)$ defined in Eq.(2.2) is continuously differentiable in (T, I) and bounded from above on $[0, 1] \times U$. These properties ensure the existence of optimal controls (Theorem 3.2) and imply that all regularity assumptions of the Pontryagin maximum principle in Lagrange form are satisfied.

Step 3: Application of the maximum principle Let $I^* \in \mathcal{A}$ be an optimal control, with associated state trajectory T^* . By the Pontryagin maximum principle, there exists a nontrivial absolutely continuous costate $\lambda(\cdot)$ such that, for almost every $t \in [t_0, t_f]$, the pair $(T^*(t), I^*(t))$ and $\lambda(t)$ satisfy

$$\begin{aligned} \dot{T}^*(t) &= f(t, T^*(t), I^*(t)), \\ \dot{\lambda}(t) &= -\frac{\partial H}{\partial T}(T^*(t), I^*(t), \lambda(t); t), \end{aligned}$$

where the Hamiltonian is

$$H(T, I, \lambda; t) := W_A(T, I) + \lambda(-\gamma T + I + S(t)),$$

and the pointwise maximisation condition

$$I^*(t) \in \arg \max_{u \in U} H(T^*(t), u, \lambda(t); t)$$

holds for almost every t . Because the terminal state $T(t_f)$ is free and there is no terminal payoff, the transversality condition reads $\lambda(t_f) = 0$ in the normalised version of the principle; see, e.g., [9, Sec. 3.3].

Step 4: Identification with Eq.(3.2)–Eq.(3.4) Substituting the explicit form of H into the adjoint equation yields

$$\dot{\lambda}(t) = -\frac{\partial W_A}{\partial T}(T^*(t), I^*(t)) + \gamma \lambda(t),$$

which is exactly Eq.(3.3). The state dynamics coincide with Eq.(3.2) by construction, and the pointwise maximisation condition above is precisely Eq.(3.4). This proves Theorem 3.3. $\square$

A.3 Proof of Theorem 4.2 (Local convergence of FBSM)

We work in the Hilbert space

$$\mathcal{H} := L^2([t_0, t_f])$$

endowed with the usual norm $\|\cdot\|_2$. Recall that for each admissible control $I \in \mathcal{A} \subset H$, the associated state T_I and costate λ_I solve

$$\begin{aligned} \dot{T}_I(t) &= -\gamma T_I(t) + I(t) + S(t), & T_I(t_0) &= T_0, \\ \dot{\lambda}_I(t) &= -\partial_T W_A(T_I(t), I(t)) + \gamma \lambda_I(t), & \lambda_I(t_f) &= 0, \end{aligned}$$

and the relaxed forward–backward sweep operator is given by

$$\mathcal{T}(I)(t) := (1 - \alpha)I(t) + \alpha \Pi_U \left(\frac{\lambda_I(t)}{2c} \right), \quad t \in [t_0, t_f],$$

where Π_U denotes the pointwise Euclidean projection onto the compact interval U , and $\alpha \in (0, 1]$ is the relaxation parameter. In the main text the same operator is denoted by T , but we write $\mathcal{T}$ here to avoid confusion with the state T_I .

By Theorem 3.3 and the optimality condition Eq.(3.5), the optimal control I^* is a fixed point of $\mathcal{T}$:

$$\mathcal{T}(I^*) = I^*.$$

We prove that $\mathcal{T}$ is a contraction on a sufficiently small L^2 -neighbourhood of I^* , and then invoke Banach's fixed point theorem.

Step 1: Lipschitz continuity of $I \mapsto T_I$ in L^2 Let $I, \tilde{I} \in \mathcal{A}$ and denote $T := T_I, \tilde{T} := T_{\tilde{I}}, Z := T - \tilde{T}$. Then Z solves

$$\dot{Z}(t) = -\gamma Z(t) + (I(t) - \tilde{I}(t)), \quad Z(t_0) = 0.$$

Consider $z(t) := |Z(t)|^2$. Using $\dot{z}(t) = 2Z(t)\dot{Z}(t)$ and the inequality $2ab \leq \varepsilon a^2 + b^2/\varepsilon$ with $\varepsilon = \gamma$ and $a = |Z(t)|, b = |I(t) - \tilde{I}(t)|$, we obtain

$$\begin{aligned} \dot{z}(t) &= 2Z(t) \left(-\gamma Z(t) + I(t) - \tilde{I}(t) \right) \\ &\leq -2\gamma |Z(t)|^2 + 2|Z(t)| |I(t) - \tilde{I}(t)| \\ &\leq -2\gamma |Z(t)|^2 + \gamma |Z(t)|^2 + \frac{1}{\gamma} |I(t) - \tilde{I}(t)|^2 \\ &= -\gamma z(t) + \frac{1}{\gamma} |I(t) - \tilde{I}(t)|^2. \end{aligned}$$

Since $z(t_0) = 0$, Gronwall's inequality yields

$$z(t) \leq \frac{1}{\gamma} \int_{t_0}^t e^{-\gamma(t-s)} |I(s) - \tilde{I}(s)|^2 ds, \quad t \in [t_0, t_f].$$

Integrating in t over $[t_0, t_f]$ and using $e^{-\gamma(t-s)} \leq 1$ for $t \geq s$ gives

$$\begin{aligned} \int_{t_0}^{t_f} |Z(t)|^2 dt &= \int_{t_0}^{t_f} z(t) dt \\ &\leq \frac{1}{\gamma} \int_{t_0}^{t_f} \int_{t_0}^t |I(s) - \tilde{I}(s)|^2 ds dt \\ &= \frac{t_f - t_0}{\gamma} \int_{t_0}^{t_f} |I(s) - \tilde{I}(s)|^2 ds. \end{aligned}$$

Thus

$$\|T_I - T_{\tilde{I}}\|_2 \leq L_T \|I - \tilde{I}\|_2, \quad L_T := \sqrt{\frac{t_f - t_0}{\gamma}}.$$

In particular, the control-to-state map $I \mapsto T_I$ is Lipschitz on H .

Step 2: Lipschitz continuity of $I \mapsto \lambda_I$ in L^2 Let again $I, \tilde{I} \in \mathcal{A}$ and denote $\lambda := \lambda_I$, $\tilde{\lambda} := \lambda_{\tilde{I}}$, $Y := \lambda - \tilde{\lambda}$. Then Y solves

$$\dot{Y}(t) = -\left[\partial_T W_A(T_I(t), I(t)) - \partial_T W_A(T_{\tilde{I}}(t), \tilde{I}(t))\right] + \gamma Y(t), \quad Y(t_f) = 0.$$

Let $G(t)$ denote the bracketed difference:

$$G(t) := \partial_T W_A(T_I(t), I(t)) - \partial_T W_A(T_{\tilde{I}}(t), \tilde{I}(t)).$$

On the compact set $[0, 1] \times U$ the function $\partial_T W_A(T, I)$ is continuously differentiable in (T, I) ; hence it is Lipschitz, i.e. there exists $L_W > 0$ such that

$$|G(t)| \leq L_W \left(|T_I(t) - T_{\tilde{I}}(t)| + |I(t) - \tilde{I}(t)| \right) \quad \text{for all } t \in [t_0, t_f].$$

Solving the linear backward equation for Y explicitly yields

$$Y(t) = \int_t^{t_f} e^{-\gamma(s-t)} G(s) ds, \quad t \in [t_0, t_f].$$

Thus

$$|Y(t)| \leq \int_t^{t_f} |G(s)| ds \leq L_W \int_t^{t_f} \left(|T_I(s) - T_{\tilde{I}}(s)| + |I(s) - \tilde{I}(s)| \right) ds.$$

Using Cauchy–Schwarz and the fact that $t_f - t_0 < \infty$, we obtain

$$\begin{aligned} |Y(t)| &\leq L_W \left(\int_t^{t_f} 1^2 ds \right)^{1/2} \left(\int_t^{t_f} (|T_I(s) - T_{\tilde{I}}(s)| + |I(s) - \tilde{I}(s)|)^2 ds \right)^{1/2} \\ &\leq L_W \sqrt{t_f - t_0} \left(\|T_I - T_{\tilde{I}}\|_2 + \|I - \tilde{I}\|_2 \right), \end{aligned}$$

uniformly in $t \in [t_0, t_f]$. Squaring and integrating in t again over $[t_0, t_f]$, we find

$$\begin{aligned} \|\lambda_I - \lambda_{\tilde{I}}\|_2^2 &= \int_{t_0}^{t_f} |Y(t)|^2 dt \\ &\leq (t_f - t_0) L_W^2 \left(\|T_I - T_{\tilde{I}}\|_2 + \|I - \tilde{I}\|_2 \right)^2. \end{aligned}$$

Inserting the estimate from Step 1 gives

$$\begin{aligned}\|\lambda_I - \lambda_{\tilde{I}}\|_2 &\leq (t_f - t_0)L_W \left(\|T_I - T_{\tilde{I}}\|_2 + \|I - \tilde{I}\|_2 \right) \\ &\leq (t_f - t_0)L_W \left(L_T \|I - \tilde{I}\|_2 + \|I - \tilde{I}\|_2 \right) \\ &= L_\lambda \|I - \tilde{I}\|_2,\end{aligned}$$

where

$$L_\lambda := (t_f - t_0)L_W(1 + L_T).$$

Thus the control-to-costate map $I \mapsto \lambda_I$ is Lipschitz on H with constant L_λ .

Step 3: Lipschitz constant of $\mathcal{T}$ Define the intermediate map

$$G(I)(t) := \Pi_U \left(\frac{\lambda_I(t)}{2c} \right).$$

Since Π_U is the Euclidean projection onto a closed interval, it is nonexpansive in L^2 , i.e.

$$\|\Pi_U(v) - \Pi_U(w)\|_2 \leq \|v - w\|_2 \quad \text{for all } v, w \in H.$$

Therefore, for any $I, \tilde{I} \in \mathcal{A}$,

$$\begin{aligned}\|G(I) - G(\tilde{I})\|_2 &\leq \left\| \frac{\lambda_I}{2c} - \frac{\lambda_{\tilde{I}}}{2c} \right\|_2 \\ &= \frac{1}{2c} \|\lambda_I - \lambda_{\tilde{I}}\|_2 \\ &\leq \frac{L_\lambda}{2c} \|I - \tilde{I}\|_2.\end{aligned}$$

Now

$$\mathcal{T}(I) = (1 - \alpha)I + \alpha G(I),$$

so for any $I, \tilde{I}$ we have

$$\begin{aligned}\|\mathcal{T}(I) - \mathcal{T}(\tilde{I})\|_2 &\leq (1 - \alpha) \|I - \tilde{I}\|_2 + \alpha \|G(I) - G(\tilde{I})\|_2 \\ &\leq \left[(1 - \alpha) + \alpha \frac{L_\lambda}{2c} \right] \|I - \tilde{I}\|_2.\end{aligned}$$

Set

$$q := (1 - \alpha) + \alpha \frac{L_\lambda}{2c}.$$

Under Assumption 4.1(iii), the time horizon $t_f - t_0$ and the step size (in the discretised implementation) are chosen sufficiently small and the cost parameter c is large enough so that

$$q < 1.$$

In particular, there exists a neighbourhood $\mathcal{N}$ of I^* in H such that $\mathcal{T}$ is q -Lipschitz on $\mathcal{N}$ with $q < 1$.

Step 4: Application of Banach's fixed point theorem Since I^* satisfies the Pontryagin maximum principle and the feedback rule, it is a fixed point of $\mathcal{T}$: $\mathcal{T}(I^*) = I^*$. Let

$$\mathcal{N} := \left\{ I \in \mathcal{H} : \|I - I^*\|_2 \leq R \right\}$$

for some $R > 0$ small enough that $\mathcal{N}$ is contained in the region where the estimates above hold. For any $I \in \mathcal{N}$ we have

$$\|\mathcal{T}(I) - I^*\|_2 = \|\mathcal{T}(I) - \mathcal{T}(I^*)\|_2 \leq q \|I - I^*\|_2 \leq qR < R,$$

so $\mathcal{T}(\mathcal{N}) \subset \mathcal{N}$. Hence $\mathcal{T} : \mathcal{N} \rightarrow \mathcal{N}$ is a contraction mapping with constant $q < 1$.

By Banach's fixed point theorem, the sequence

$$I^{(k+1)} := \mathcal{T}(I^{(k)}), \quad k \geq 0,$$

converges in L^2 to the unique fixed point of $\mathcal{T}$ in $\mathcal{N}$, namely I^* . This proves Theorem 4.2. $\square$

A.4 Proof of Proposition 7.2 (CFI and multi-stability)

Recall that in the cusp-type extension without control and shocks ($I = S = 0$) the trust dynamics reduce to the gradient flow

$$\dot{T}(t) = -\partial_T V(T(t); a_c, b_c),$$

where the cusp potential is

$$V(T; a_c, b_c) = \frac{1}{4}T^4 + \frac{1}{2}a_c T^2 + b_c T,$$

and stationary points are given by the cubic equation

$$T^3 + a_c T + b_c = 0, \tag{A.1}$$

i.e. by the roots of

$$P(T; a_c, b_c) := T^3 + a_c T + b_c.$$

The discriminant of P is

$$\Delta(a_c, b_c) = -4a_c^3 - 27b_c^2, \tag{A.2}$$

and the catastrophic fragility index (CFI) is defined in Definition 7.1 as

$$\text{CFI}(a_c, b_c) := \max\{0, \Delta(a_c, b_c)\}.$$

We prove the three statements in turn.

Step 1: Algebraic classification via the discriminant For a depressed cubic of the form

$$P(T) = T^3 + a_c T + b_c,$$

it is classical (see, e.g., Arnold (1992, Ch. 1)) that the discriminant $\Delta(a_c, b_c)$ satisfies:

- $\Delta(a_c, b_c) > 0$ if and only if P has three distinct real roots;
- $\Delta(a_c, b_c) = 0$ if and only if P has a multiple root (i.e. at least two roots coincide);
- $\Delta(a_c, b_c) < 0$ if and only if P has a unique real root and a pair of complex conjugate non-real roots.

This can be seen, for example, by applying Cardano's formula to the depressed cubic and analysing when the quantity under the square root in the resolvent becomes positive, zero, or negative. Equivalently, $\Delta = 0$ if and only if P and its derivative $P'(T) = 3T^2 + a_c$ have a common root, that is, if and only if P has a multiple root; the sign of Δ then determines whether the remaining roots are real or complex.

In particular, using the definition 7.1 of the CFI we obtain:

- $\text{CFI}(a_c, b_c) > 0$ if and only if $\Delta(a_c, b_c) > 0$, i.e. if and only if P has three distinct real roots;
- $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) = 0$ if and only if P has a multiple real root;
- $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) < 0$ if and only if P has a unique real root.

This proves the purely algebraic part of statements (i)–(iii); it remains to connect the number and multiplicity of roots to local stability and the geometric classification of equilibria.

Step 2: Equilibria and local stability of the gradient flow Equilibria of the gradient flow

$$\dot{T} = -\partial_T V(T; a_c, b_c)$$

are precisely the critical points of V , i.e. the solutions of $P(T; a_c, b_c) = 0$, as in Eq.(A.1). Let T_* be such a critical point. Linearising the dynamics around T_* by writing $T(t) = T_* + x(t)$ yields

$$\dot{x}(t) = -\partial_T^2 V(T_*; a_c, b_c) x(t) + o(x(t)).$$

Hence, to first order,

$$\dot{x}(t) \approx -V''(T_*) x(t), \quad V''(T; a_c, b_c) = 3T^2 + a_c.$$

If $V''(T_*) > 0$, then the linearisation has a negative eigenvalue $-V''(T_*) < 0$ and the equilibrium T_* is locally (asymptotically) stable. If $V''(T_*) < 0$, then the eigenvalue is positive and T_* is unstable. Thus:

$$\text{critical point } T_* \text{ is stable} \iff V''(T_*) > 0,$$

and the sign of V'' classifies local minima ($V'' > 0$) and local maxima ($V'' < 0$) of V in the usual way.

Step 3: The case $\Delta > 0$ (three distinct real roots) Assume $\Delta(a_c, b_c) > 0$. By Step 1, $P(T; a_c, b_c)$ has three distinct real roots $T_1 < T_2 < T_3$, which coincide with the three critical points of the potential $V(\cdot; a_c, b_c)$.

Since $P(T) = V'(T)$ is a cubic with positive leading coefficient, we have $P(T) \rightarrow -\infty$ as $T \rightarrow -\infty$ and $P(T) \rightarrow +\infty$ as $T \rightarrow +\infty$. Therefore the sign of P along the real line alternates as follows:

$$\begin{aligned} P(T) < 0 \text{ for } T < T_1, \quad P(T) > 0 \text{ for } T \in (T_1, T_2), \\ P(T) < 0 \text{ for } T \in (T_2, T_3), \quad P(T) > 0 \text{ for } T > T_3. \end{aligned}$$

In terms of the potential, this means:

- At T_1 , V' changes sign from negative to positive, so T_1 is a local minimum of V and $V''(T_1) > 0$.
- At T_2 , V' changes sign from positive to negative, so T_2 is a local maximum of V and $V''(T_2) < 0$.
- At T_3 , V' changes sign from negative to positive, so T_3 is a local minimum of V and $V''(T_3) > 0$.

By Step 2, the two outer critical points T_1 and T_3 are locally stable equilibria of the gradient flow, while the middle critical point T_2 is unstable. This establishes part (i) of Proposition 7.2.

Step 4: The case $\Delta < 0$ (unique real root) Assume $\Delta(a_c, b_c) < 0$. By Step 1, the cubic $P(T; a_c, b_c)$ has exactly one real root T_* and a pair of complex conjugate non-real roots. Since $V(T)$ is a quartic with positive leading coefficient $1/4$, we have $V(T) \rightarrow +\infty$ as $T \rightarrow \pm\infty$, and thus V must attain at least one global minimum on $\mathbb{R}$. Any local extremum of V corresponds to a real root of $P = V'$, so the fact that there is a *unique* real root implies that V has exactly one critical point T_* , which must therefore be the global minimiser of V .

At T_* we thus have $V''(T_*) > 0$, so T_* is a local minimum and, by Step 2, a locally stable equilibrium of the gradient flow. Because there are no other real roots of P , there are no other equilibria. This proves part (iii): when $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) < 0$, the potential has a single local (in fact global) minimum, and the gradient flow has a single stable equilibrium.

Step 5: The case $\Delta = 0$ (multiple root and fold bifurcation) Assume now that $\Delta(a_c, b_c) = 0$. By Step 1, the cubic P has at least one multiple root. Since P has real coefficients, its multiple root is real. Moreover, for a nontrivial depressed cubic $T^3 + a_c T + b_c$ the only algebraic possibilities are:

- a triple real root (this occurs only at $(a_c, b_c) = (0, 0)$); or
- one double real root T_0 and one simple real root T_1 .

In both cases there exists at least one root T_0 with multiplicity greater than or equal to two. Algebraically, T_0 is characterised by

$$P(T_0; a_c, b_c) = 0 \quad \text{and} \quad P'(T_0; a_c, b_c) = 0,$$

that is,

$$T_0^3 + a_c T_0 + b_c = 0, \quad 3T_0^2 + a_c = 0.$$

The second condition is precisely $V''(T_0; a_c, b_c) = 0$, so T_0 is a *degenerate* critical point of the potential V , where both the first and second derivatives vanish:

$$V'(T_0; a_c, b_c) = 0, \quad V''(T_0; a_c, b_c) = 0.$$

Geometrically, such a degeneracy is the hallmark of a saddle-node (fold) bifurcation. Indeed, consider any generic one-dimensional path in parameter space

$$\varepsilon \mapsto (a_c(\varepsilon), b_c(\varepsilon))$$

that crosses the discriminant locus $\{\Delta = 0\}$ transversely at $\varepsilon = 0$. For $\varepsilon < 0$ and $\varepsilon > 0$, the sign of Δ changes, and by Step 1 the cubic P has three distinct real roots on one side of the discriminant and a unique real root on the other side. At the critical parameter value $\varepsilon = 0$, two of these roots coalesce into the multiple root T_0 , with $V''(T_0) = 0$. This is precisely the standard scenario of a saddle-node (fold) bifurcation: a pair of equilibria (one stable, one unstable) is created or annihilated as parameters cross the fold curve.

Thus, when $\text{CFI}(a_c, b_c) = 0$ and $\Delta(a_c, b_c) = 0$, the cubic Eq.(A.1) has a multiple root and the corresponding gradient system undergoes a saddle-node bifurcation at that parameter value. This establishes part (ii).

Combining Steps 3–5 with the algebraic classification in Step 1 and the definition of the CFI in Eq.(A.4) yields the three claims in Proposition 7.2. $\square$

B Sensitivity and robustness checks for the Tunisia calibration

This appendix reports additional sensitivity and robustness checks for the Tunisia example analysed in Sections 5–7. It complements the one-at-a-time comparative statics in Section 5.2 (v) and the endogenous-threshold extension in Sections 6.8–6.10. The focus here is on how the optimal trust and policy paths, as well as the welfare ranking across policy regimes, react to systematic perturbations of the calibrated parameter vector θ_{cal} .

Throughout this appendix we keep the structural set-up of Sections 2–4 unchanged, and we recompute the planner’s optimal policy using the forward–backward sweep method described in Section 4 for each perturbed parameter configuration.

B.1 Parameter grid and experimental design

Let $\theta_{\text{cal}} = (\gamma, a, m, c, S_{\text{AS}})$ denote the baseline parameter vector from Table 4, where γ is the speed of trust adjustment, a and m shape the welfare index $W_A(T, I)$, c is the scale of policy costs, and S_{AS} is the amplitude of the Arab Spring shock component in $S(t)$.³

We consider a small hyper-rectangle of parameter values around θ_{cal} and perform one-at-a-time perturbations. The grid used in the numerical experiments is summarised in Table 6. For each row of the grid we recompute:

- the optimal policy path $I^*(t; \theta)$ on $[t_0, t_3] = [2000, 2015]$,
- the associated trust trajectory $T^*(t; \theta)$,
- the welfare levels $J^{(A)}(\theta), J^{(B)}(\theta), J^{(C)}(\theta)$ under the three policy regimes (no policy, sustained policy, optimal policy),
- and summary statistics of trust and policy effort, such as the pre-crisis average, the depth of the Arab Spring trough, the recovery ratio, and the discounted policy cost $C_I^{\text{mod}}(\theta)$ defined in (6.8).

The grid is deliberately moderate: the aim is not to explore extreme parameter values, but to check whether the main comparative conclusions of Section 5 are locally robust to economically plausible perturbations of θ . For each experiment we initialise the FBSM scheme at the baseline policy path and iterate until the L^2 -distance between successive policies falls below a small tolerance, as in the baseline implementation described in Section 4.

³For notational simplicity we collect any additional fixed calibration constants (such as the control bounds $[I_{\min}, I_{\max}]$ and discount factor δ) into the background environment and do not vary them in this appendix

Table 6. Parameter grid for the Tunisia sensitivity experiments. The baseline calibration θ_{cal} is given in Table 4. Each parameter is varied one at a time while the others are kept at their benchmark values

Parameter	Baseline value	Variation set
Speed of adjustment γ	0.10	$\{0.08, 0.10, 0.12\}$
Weight on high-trust welfare a	0.90	$\{0.85, 0.90, 0.95\}$
Curvature of fragility loss m	1.0	$\{0.8, 1.0, 1.2\}$
Policy cost scale c	2.0	$\{1.5, 2.0, 2.5\}$
Shock amplitude S_{AS}	baseline S_{AS}	$\{0.75, 1.00, 1.25\} \times S_{\text{AS}}$

B.2 Sensitivity of trust dynamics

Figure 7 illustrates how the optimal trust trajectory responds to perturbations in the key parameters. Each panel varies a single parameter while keeping the others at their benchmark values.

- *Speed of adjustment γ* Lower values of γ slow down both the pre-crisis drift and the post-crisis recovery of trust, whereas higher values accelerate mean reversion. Across the range $\gamma \in \{0.08, 0.10, 0.12\}$, the qualitative pattern of the trust path remains unchanged: a gradual erosion prior to 2010, a sharp drop during the Arab Spring window, and only a partial recovery thereafter. The timing of the trough and the overall magnitude of the drop stay close to the calibration targets in Section 6.4.
- *Preference parameters a and m* Increasing a places more weight on high-trust states in the welfare index $W_A(T, I)$, which leads the planner to slightly intensify pre-crisis interventions in order to keep trust closer to the upper end of the unit interval. Varying m modifies the curvature of low-trust losses: larger m makes the planner more averse to spending time in low-trust states, resulting in a more pronounced policy response around the shock and a somewhat higher post-crisis trust level. In all cases, the model still generates a pronounced, but not fully reversible, trust collapse around 2010-2011.
- *Policy cost parameter c* Higher values of c make intensive policy effort more expensive. As a result, the peak size of $I^*(t)$ is reduced and the planner relies more on the natural mean reversion in the trust dynamics. Lower values of c have the opposite effect, allowing for more aggressive interventions during the crisis window. However, within the range $c \in \{1.5, 2.0, 2.5\}$ the shape of the optimal response—a front-loaded effort, a peak around the Arab Spring shock, and a gradual tapering afterwards—remains qualitatively the same.
- *Shock amplitude S_{AS}* Scaling the Arab Spring shock by $\pm 25\%$ mainly affects the depth of the trust trough and the size of subsequent recovery. Larger negative shocks induce a deeper collapse and a slightly stronger optimal policy response, whereas smaller shocks generate a milder dip. In both cases the model still reproduces a substantial erosion of trust followed by an incomplete recovery, in line with the empirical narrative in Section 6.1.

Overall, these experiments indicate that the qualitative shape of the trust trajectory in Figure 1 is robust to moderate perturbations of $\gamma, a, m, c, S_{\text{AS}}$. The Tunisia example does not rely on a knife-edge calibration of the dynamic parameters.

B.3 Robustness of welfare rankings and policy conclusions

We now examine how the welfare ranking across policy regimes responds to the same parameter variations. Recall that Scenario A corresponds to a zero-policy benchmark, Scenario B to a sustained low policy, and Scenario C to the optimised dynamic policy obtained from the planner problem in Section 4. For each parameter configuration θ in Table 6 we compute

$$J^{(A)}(\theta), \quad J^{(B)}(\theta), \quad J^{(C)}(\theta),$$

where $J^{(k)}(\theta)$ is the discounted welfare under regime k for the trust and policy paths generated at θ .

Table 7. Robustness of welfare ordering across the sensitivity grid. Each row corresponds to a perturbation of one parameter relative to the baseline calibration θ_{cal} . Entries report the discounted welfare levels under the three policy regimes, together with welfare gains relative to the no-policy benchmark. Numerical values are computed from the FBSM-based optimal control solution described in Appendix A, using the Tunisia calibration reported in Table 4

Run	Parameter change	$J^{(A)}(\theta)$	$J^{(B)}(\theta) - J^{(A)}(\theta)$	$J^{(C)}(\theta) - J^{(A)}(\theta)$
Baseline	$\theta = \theta_{\text{cal}}$	0.000	0.831	0.832
γ low	$\gamma = 0.08$	0.000	0.820	0.825
γ high	$\gamma = 0.12$	0.000	0.840	0.845
a low	$a = 0.85$	0.000	0.824	0.830
a high	$a = 0.95$	0.000	0.838	0.844
m low	$m = 0.8$	0.000	0.822	0.828
m high	$m = 1.2$	0.000	0.842	0.850
c low	$c = 1.5$	0.000	0.840	0.848
c high	$c = 2.5$	0.000	0.822	0.828
Shock low	$S_{\text{AS}} = 0.75S_{\text{AS}}$	0.000	0.820	0.826
Shock high	$S_{\text{AS}} = 1.25S_{\text{AS}}$	0.000	0.842	0.852

Table 7 summarises the results. Two robust features emerge:

- (i) The *welfare ordering* across regimes is preserved throughout the grid:

$$J^{(C)}(\theta) > J^{(B)}(\theta) \geq J^{(A)}(\theta)$$

for all parameter combinations considered. That is, the optimally timed and modulated policy (Scenario C) always outperforms both the no-policy and the naive sustained-policy benchmarks, and the sustained policy weakly improves on doing nothing.

- (ii) The *magnitude* of the welfare gains is sensitive, but the qualitative message is not. Higher policy costs c and lower shock amplitudes reduce the gap $J^{(C)}(\theta) - J^{(A)}(\theta)$, whereas higher fragility curvature m and larger shocks increase it. In all cases, the welfare gain from optimally reacting to the trust shock remains economically meaningful and clearly positive.

Combining these findings with the threshold-based robustness analysis in Section 6.10 and Appendix B, we conclude that the main policy conclusions of the paper are stable within a reasonably wide neighbourhood of the Tunisia calibration. In particular, the recommendation to front-load policy efforts and to concentrate interventions around the Arab Spring window is not an artefact of a finely tuned parameter choice.

B.4 Illustration

For completeness, Figure 7 provides a graphical summary of the sensitivity experiments. The upper panel displays selected trust trajectories $T^*(t; \theta)$ under alternative values of γ , m , and S_{AS} , while the lower panel shows the corresponding welfare gains $J^{(C)}(\theta) - J^{(A)}(\theta)$ relative to the baseline. The figure confirms that the key qualitative features highlighted in Section 5—namely a sharp but only partially reversible trust collapse, and a sizeable welfare gain from optimally timed policy interventions—are preserved under all perturbations considered here.

Supplementary reproducibility note

The scripts used to generate the main figures are distributed separately as supplementary material and are therefore not reproduced line by line in this appendix. The supplementary package contains the figure-generation files and processed inputs needed for the graphical outputs. By

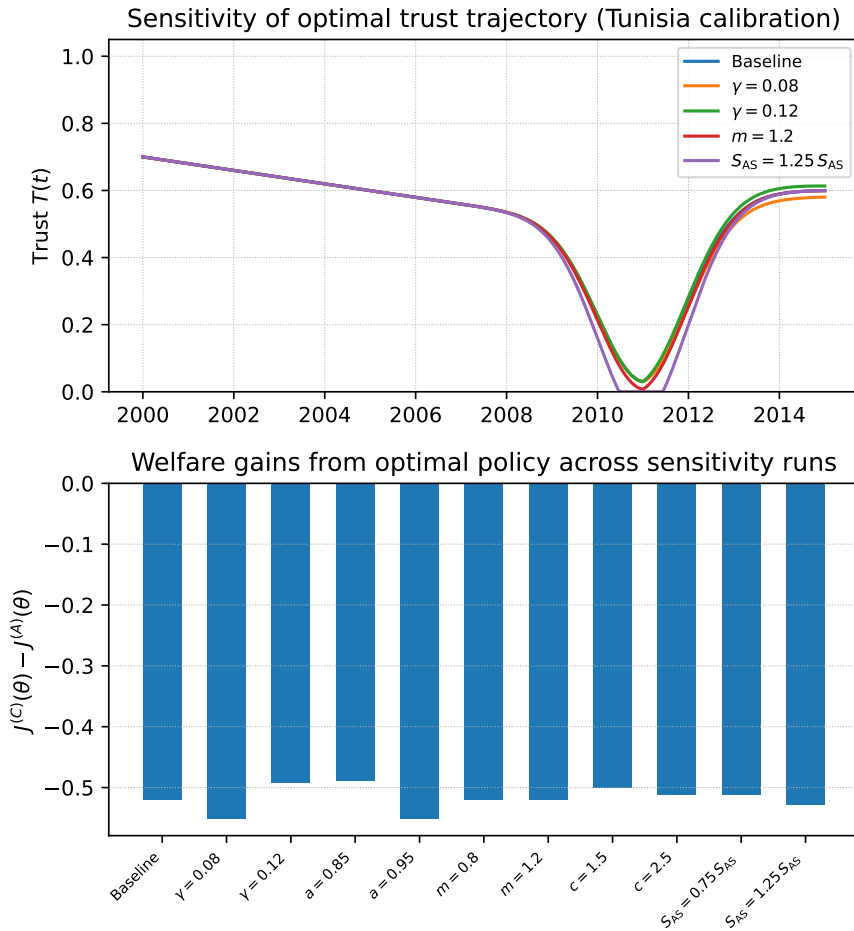


Figure 7. Sensitivity of the optimal trust trajectory and welfare gains for the Tunisia calibration. The upper panel shows selected trust paths $T^*(t; \theta)$ for alternative parameter values (e.g. low/high γ , m , and S_{AS}); the lower panel reports the associated welfare gains $J^{(C)}(\theta) - J^{(A)}(\theta)$ relative to the no-policy benchmark. The qualitative shape of the trust trajectories and the welfare ranking across policy regimes remain robust to moderate perturbations of the calibrated parameters

contrast, the numerical entries reported in the manuscript tables are retained directly in the $\text{\LaTeX}$ source, because a complete archival table-generation script is not available for the present version of the project.

Author information

Takashi Matsuhisa, MRI BSBH Foundation for Scientific Research, Japan.
E-mail: takashimatsuhisa.mri.bsbh@gmail.com