

Strategic Arms Accumulation as a Stochastic Differential Game: Feedback Equilibria, Regime Shocks, and Deterrence Dynamics

Heng-Fu Zou

July 8, 2025

Abstract

We develop a dynamic differential game model of strategic arms accumulation between two rival states, reflecting the security dilemma of major geopolitical adversaries such as the United States and China or the United States and the former Soviet Union. The model features both deterministic and stochastic formulations, where each country maximizes intertemporal utility over its own military stock while negatively internalizing the rival's arms buildup. We derive Markov–Perfect Nash equilibrium strategies in both open-loop and feedback forms, with special emphasis on the characterization and solution of the coupled Hamilton–Jacobi–Bellman (HJB) and Fokker–Planck–Kolmogorov (FPK) partial differential equations. In the linear–quadratic case, we solve the time-consistent Riccati equations and analyze the elasticity of optimal strategies with respect to changes in strategic preferences, investment costs, and external threats. The model is extended to include regime shocks—discontinuous jumps in threat perception or strategic aggressiveness—which may destabilize deterrence and trigger explosive arms escalation. Comparative statics from Riccati dynamics reveal thresholds beyond which deterrence collapses and feedback control solutions blow up in finite time. We also explore welfare trade-offs, deterrence sustainability, and policy implications of dynamic treaties. The mathematical framework developed here provides a foundation for modeling arms races as nonlinear stochastic systems under endogenous expectations, belief dynamics, and political instability.

Keywords: Strategic arms race; differential games; Nash equilibrium; feedback strategy; Riccati equation; Fokker–Planck equation; HJB equation; stochastic dynamics; regime shocks; deterrence collapse; Markov–Perfect equilibrium; dynamic stability; military investment; geopolitical competition; arms control treaties

1 Introduction

The strategic accumulation of military capabilities—particularly arms stockpiles—has long been at the heart of geopolitical conflict, deterrence theory, and national

security strategy. From the Cold War rivalry between the United States and the Soviet Union to the current tensions between the United States and the People’s Republic of China, arms races have embodied a fundamental dilemma: the security-seeking behavior of one state often generates insecurity in others, thereby creating a dynamic of mutual escalation. This phenomenon, known as the security dilemma, has been studied from various perspectives in international relations, game theory, and strategic studies. However, relatively few models rigorously integrate the continuous-time dynamics of military investment, rational strategic interaction, and stochastic uncertainty into a unified mathematical framework.

This paper seeks to address this gap by modeling the arms race between two nations as a stochastic differential game with feedback strategies. In contrast to static game-theoretic formulations or simple intertemporal optimization problems, we embed the analysis within a dynamic system in which each country’s military decisions continuously evolve in response to both its own security preferences and the observed behavior of its rival. Each country’s utility is positively influenced by its own arms accumulation, reflecting enhanced self-defense and regional influence, but negatively affected by the rival’s military growth, capturing threat perception, regional destabilization, and the erosion of strategic parity. The interaction is further complicated by the presence of uncertainty, political shocks, and the inability of nations to perfectly predict future strategic environments.

The central contribution of the paper is the rigorous derivation and analysis of Markov–Perfect Nash equilibrium strategies in both deterministic and stochastic settings. We explore two strategic forms: open-loop strategies, in which countries commit to time-dependent plans regardless of unfolding events, and feedback strategies, in which policies are dynamically revised as a function of the evolving state of the system. The latter class of strategies, while more complex mathematically, captures the realistic behavior of states that adjust their military policies in real time based on observed threats, technological advances, and shifts in international alliances.

By deriving the full system of Hamilton–Jacobi–Bellman equations for each player and the associated Fokker–Planck equations governing the probabilistic distribution of arms stockpiles, we provide a comprehensive treatment of the underlying dynamics. The feedback strategies derived from these equations are time-consistent and subgame perfect, meaning they remain optimal even as the system evolves away from initial conditions. In special cases—such as when the utility functions are quadratic and the dynamics are linear—we reduce the problem to Riccati differential equations, allowing for explicit solution and comparative statics.

In addition to equilibrium characterization, we incorporate regime shocks—abrupt changes in political ideology, threat sensitivity, or leadership preferences—modeled as discontinuous shifts in the parameters governing utility and dynamics. These shocks can induce instability, trigger runaway arms buildups, or collapse the equilibrium altogether. We interpret these events as analogs to historical episodes such as the rise of aggressive leaderships, the breakdown of arms

control agreements, or the unexpected militarization of previously cooperative regimes.

The analysis provides several novel insights. First, it demonstrates how equilibrium arms levels are sensitive not only to cost and preference parameters but also to the anticipated aggressiveness of the rival, a phenomenon we term strategic elasticity. Second, it shows how stochastic fluctuations—reflecting domestic uncertainty, informational noise, or unpredictable events—can lead to long-run probability distributions of military capabilities, which are essential for evaluating the likelihood of deterrence failure. Third, it opens the door to welfare analysis under strategic rivalry, raising the question of whether equilibria can be improved through treaties, side payments, or institutional cooperation.

Ultimately, this paper contributes to both the theoretical foundations and practical understanding of dynamic arms competition. The mathematical models presented here offer a basis for future empirical calibration and policy simulation, while the theoretical results shed light on the conditions under which strategic peace or conflict escalation may emerge. In an increasingly multipolar and volatile international environment, understanding the dynamics of arms races through the lens of differential games is not merely an academic exercise—it is a strategic necessity.

2 Model Setup

We consider a continuous-time stochastic differential game involving two sovereign nations Country 1 and Country 2, each of which seeks to optimize its own long-term military welfare by dynamically accumulating arms while responding strategically to the actions of its rival. Let $x_i(t) \in \mathbb{R}_{++}$ denote the military stock of country $i \in \{1, 2\}$ at time $t \geq 0$, and let $u_i(t) \in \mathbb{R}_+$ denote the instantaneous investment in arms. The arms accumulation process is driven by the stochastic differential equation

$$dx_i(t) = [u_i(t) - \delta_i x_i(t)] dt + \sigma_i x_i(t) dW_i(t),$$

where $\delta_i > 0$ is the depreciation rate of military assets, $\sigma_i \geq 0$ is the intensity of idiosyncratic military volatility reflecting unpredictability in defense effectiveness, and $W_i(t)$ is a standard Brownian motion adapted to a common filtration $\{\mathcal{F}_t\}_{t \geq 0}$ capturing all information available up to time t . The multiplicative noise term $\sigma_i x_i(t) dW_i(t)$ ensures that stochastic fluctuations scale with the magnitude of the military stock, consistent with real-world volatility patterns where larger forces are more exposed to destabilizing events.

Each country maximizes a time-separable, discounted utility functional over the infinite horizon, incorporating both the security benefits from its own arms and the strategic disutility caused by the rival's military capacity. Specifically, the utility for country i is given by

$$J_i(u_i, u_j) = \mathbb{E} \left[\int_0^\infty e^{-\rho_i t} \left(\log x_i(t) - \gamma_i \log x_j(t) - \frac{1}{2} c_i u_i(t)^2 \right) dt \right],$$

where $j \neq i$, $\rho_i > 0$ is the subjective discount rate, $\gamma_i > 0$ measures the threat perception coefficient, i.e., how sensitive country i is to the rival's arms buildup and $c_i > 0$ is a quadratic penalty on arms investment representing convex costs of defense expenditure. The log utility of own military power implies diminishing marginal security gains from additional arms, while the negative log utility of the rival's arms captures the deterrence tension inherent in arms competition. The quadratic investment cost ensures strong concavity and guarantees the existence of interior optimal controls.

In this dynamic game, each country's decision is framed in terms of Markovian feedback strategies $u_i(x_1, x_2)$, which depend on the current state of both countries' military capacities. Let $V_i(x_1, x_2)$ denote the value function of country i , representing the maximum attainable utility conditional on the current arms state (x_1, x_2) . The corresponding Hamilton-Jacobi-Bellman (HJB) equation for country i is

$$\rho_i V_i = \sup_{u_i \geq 0} \left\{ \log x_i - \gamma_i \log x_j - \frac{1}{2} c_i u_i^2 + (u_i - \delta_i x_i) \frac{\partial V_i}{\partial x_i} - \delta_j x_j \frac{\partial V_i}{\partial x_j} + \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V_i}{\partial x_i^2} + \frac{1}{2} \sigma_j^2 x_j^2 \frac{\partial^2 V_i}{\partial x_j^2} \right\}.$$

The first-order condition for interior maximization yields the optimal feedback control

$$u_i^*(x_1, x_2) = \frac{1}{c_i} \frac{\partial V_i}{\partial x_i}(x_1, x_2),$$

which explicitly links each country's optimal investment to the marginal value of arms accumulation. Substituting this expression back into the HJB yields a fully nonlinear, second-order partial differential equation in $V_i(x_1, x_2)$, coupled across players due to the cross-dependence in utilities and stochastic terms. The strategic interdependence is embedded through both direct cross-terms in the objective function (via $\log x_j$) and through the optimal strategies that depend on the value function gradients of both players. Hence, the equilibrium problem is cast as a system of coupled HJB equations:

$$\rho_i V_i = \log x_i - \gamma_i \log x_j - \frac{1}{2c_i} \left(\frac{\partial V_i}{\partial x_i} \right)^2 + \left(\frac{1}{c_i} \frac{\partial V_i}{\partial x_i} - \delta_i x_i \right) \frac{\partial V_i}{\partial x_i} - \delta_j x_j \frac{\partial V_i}{\partial x_j} + \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V_i}{\partial x_i^2} + \frac{1}{2} \sigma_j^2 x_j^2 \frac{\partial^2 V_i}{\partial x_j^2},$$

where $i = 1, 2$. These equations must be solved jointly, as each player's value function depends on the state variables of both countries. Moreover, equilibrium consistency requires that the strategies derived from the HJB system match the conjectured feedback forms used in the state dynamics, closing the fixed-point loop of optimal responses.

To characterize the evolution of the joint distribution over arms levels $p(t, x_1, x_2)$, we consider the associated Fokker-Planck-Kolmogorov (FPK) forward equation governing the density dynamics:

$$\frac{\partial p}{\partial t} = - \sum_{i=1}^2 \frac{\partial}{\partial x_i} ([u_i^*(x_1, x_2) - \delta_i x_i] p) + \frac{1}{2} \sum_{i=1}^2 \sigma_i^2 \frac{\partial^2}{\partial x_i^2} (x_i^2 p),$$

where $u_i^*(x_1, x_2)$ is defined by the HJB optimality condition above. This FPK equation provides a dynamic probabilistic characterization of the arms race, tracking how the population of trajectories over (x_1, x_2) evolves in time under optimal strategic behavior. The coupled HJB-FPK system together defines the full dynamic game equilibrium, capturing both the optimal value functions and the endogenous state distribution evolution.

We emphasize that the stochastic framework allows for rich strategic uncertainty: sudden changes in arms effectiveness, perception shocks, or asymmetric information can all be incorporated through the Brownian motion terms and time-varying volatility. Moreover, the feedback formulation guarantees subgame perfection and time consistency, as strategies are recalculated at each state realization rather than fixed in advance as in open-loop forms. In subsequent sections, we analyze the existence, regularity, and comparative statics of this system, and explore linear-quadratic specifications that yield closed-form Riccati-based feedback laws and explicit insights into strategic aggressiveness and deterrence.

3 Existence, Uniqueness, and Regularity of Feedback Equilibrium

The mathematical structure of the two-country stochastic arms race differential game defined in Section 2 leads to a highly nontrivial system of coupled nonlinear Hamilton-Jacobi-Bellman (HJB) equations and Fokker-Planck-Kolmogorov (FPK) equations. The feedback Nash equilibrium must satisfy these systems simultaneously, which introduces deep challenges regarding the existence and regularity of value functions $V_i(x_1, x_2)$ and the associated feedback controls $u_i(x_1, x_2)$. In this section, we rigorously analyze the analytic properties of these equations and provide conditions under which a feedback equilibrium exists, is unique, and is regular enough to justify its use in continuous-time differential games.

We begin by recalling the coupled HJB system from Section 2. For each country $i = 1, 2$, let $j \neq i$, the HJB equation reads:

$$\begin{aligned} \rho_i V_i(x_1, x_2) = & \log x_i - \gamma_i \log x_j - \frac{1}{2c_i} \left(\frac{\partial V_i}{\partial x_i} \right)^2 + \left(\frac{1}{c_i} \frac{\partial V_i}{\partial x_i} - \delta_i x_i \right) \frac{\partial V_i}{\partial x_i} \\ & - \delta_j x_j \frac{\partial V_i}{\partial x_j} + \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V_i}{\partial x_i^2} + \frac{1}{2} \sigma_j^2 x_j^2 \frac{\partial^2 V_i}{\partial x_j^2}. \end{aligned}$$

This equation is nonlinear and second-order in both x_1 and x_2 . It also depends nonlocally on the other player's value function because the feedback control of one player influences the evolution of both players' states.

We now impose a Sobolev regularity framework. Define the function space $\mathcal{V} := W_{\text{loc}}^{2,p}(\mathbb{R}_{++}^2)$ for some $p > 2$. We seek classical solutions in $C^2(\mathbb{R}_{++}^2) \cap \mathcal{V}$, so that all partial derivatives and second-order terms in the HJB equation are well-defined almost everywhere. We must ensure that the right-hand side of the

HJB equation is integrable and that the derivatives of V_i satisfy appropriate boundary behavior as $x_i \rightarrow 0^+$ or $x_i \rightarrow \infty$. This is especially important since the logarithmic utility term $\log x_i$ diverges near zero.

We impose the following assumptions to ensure well-posedness:

- (A1) $\delta_i, \rho_i, \gamma_i, c_i, \sigma_i > 0$ for $i = 1, 2$;
- (A2) There exist constants $\underline{x}, \bar{x} > 0$ such that initial conditions $x_i(0) \in [\underline{x}, \bar{x}]$;
- (A3) The candidate value function $V_i(x_1, x_2)$ is strictly concave in x_i and bounded below by $\alpha \log x_i - \beta \log x_j - K$ for constants $\alpha, \beta, K > 0$.

Under these assumptions, we invoke results from the theory of fully nonlinear degenerate elliptic PDEs. The degeneracy here arises from the term $\sigma_i^2 x_i^2 \partial_{x_i x_i}^2 V_i$, which vanishes as $x_i \rightarrow 0$, but Assumption (A2) ensures boundedness away from the boundary. We apply the Schauder fixed point theorem in Banach spaces to establish existence: consider the map

$$\mathcal{T}_j : V_j \mapsto V_i(V_j)$$

which assigns to a given strategy of Player j the optimal response V_i solving the HJB equation. If $\mathcal{T}_j$ is compact and continuous in $\mathcal{V}$, and the image is convex and closed, Schauder's theorem guarantees a fixed point $V_i = \mathcal{T}_j(V_i)$, which forms a best response. Symmetrically, applying the same reasoning to Player j , the coupled system admits a joint fixed point, thus proving the existence of a Markov's Perfect Nash equilibrium.

To establish uniqueness, we employ the monotonicity method. Define the difference $\Delta V := V_i^{(1)} - V_i^{(2)}$ between two candidate solutions. Substituting into the HJB operator and exploiting the concavity and quadratic structure of the control term, one can show that the operator is monotone in $\partial V_i / \partial x_i$. Combined with comparison principles for elliptic PDEs, this yields that $\Delta V = 0$ must hold throughout the domain, implying uniqueness.

For regularity, we use bootstrapping techniques from PDE theory. Starting with weak solutions in $\mathcal{V}$, we note that the elliptic nature of the HJB operator with uniformly bounded coefficients (under (A2)) enables $C^{2,\alpha}$ -interior regularity by Krylov-Safonov. Moreover, the log utility term being smooth on $[\underline{x}, \bar{x}]$ allows local C^∞ regularity via Schauder estimates. Thus, the value functions are smooth in the interior of the domain.

Finally, the feedback strategy

$$u_i^*(x_1, x_2) = \frac{1}{c_i} \frac{\partial V_i}{\partial x_i}$$

is well-defined, bounded, and Lipschitz-continuous on compact subsets of $\mathbb{R}_{++}^2$, ensuring admissibility. This establishes the existence of a unique and regular feedback equilibrium for the stochastic arms race game.

4 Coupled HJB–FPK System for Feedback Equilibrium Dynamics

To understand the full stochastic dynamics of strategic interaction between two rival states—denoted as country 1 and country 2—we must simultaneously characterize each country’s optimal value function via the Hamilton–Jacobi–Bellman (HJB) equations and the probabilistic evolution of their respective state variables using the Fokker–Planck–Kolmogorov (FPK) equations. These systems are inherently coupled, as the HJB equation for each country depends on the diffusion and drift of both players’ state variables, while the FPK equation describing the population distribution depends on the feedback strategies derived from the HJB solutions. The equilibrium of the system is thus a fixed point in the joint space of value functions and state distributions.

Let $x_i(t)$ denote the military stock of country $i = 1, 2$, governed by the Itô stochastic differential equation:

$$dx_i(t) = \left(\frac{1}{c_i} \frac{\partial V_i}{\partial x_i} - \delta_i x_i \right) dt + \sigma_i x_i dW_i(t),$$

where $\frac{1}{c_i} \frac{\partial V_i}{\partial x_i}$ is the optimal feedback control derived from solving the HJB equation, and $W_i(t)$ are independent Brownian motions.

The HJB equation for country i , given its utility over own arms $\log x_i$, disutility from the rival’s arms $-\gamma_i \log x_j$, and control cost $\frac{1}{2} c_i u_i^2$, is

$$\rho_i V_i = \log x_i - \gamma_i \log x_j + \sup_{u_i} \left\{ u_i \frac{\partial V_i}{\partial x_i} - \frac{1}{2} c_i u_i^2 \right\} - \delta_i x_i \frac{\partial V_i}{\partial x_i} - \delta_j x_j \frac{\partial V_i}{\partial x_j} + \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V_i}{\partial x_i^2} + \frac{1}{2} \sigma_j^2 x_j^2 \frac{\partial^2 V_i}{\partial x_j^2}.$$

Maximizing the Hamiltonian with respect to u_i yields the first-order condition:

$$0 = \frac{\partial V_i}{\partial x_i} - c_i u_i^* \quad \Rightarrow \quad u_i^* = \frac{1}{c_i} \frac{\partial V_i}{\partial x_i}.$$

Substituting this into the HJB yields the nonlinear PDE:

$$\rho_i V_i = \log x_i - \gamma_i \log x_j + \frac{1}{2c_i} \left(\frac{\partial V_i}{\partial x_i} \right)^2 - \delta_i x_i \frac{\partial V_i}{\partial x_i} - \delta_j x_j \frac{\partial V_i}{\partial x_j} + \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V_i}{\partial x_i^2} + \frac{1}{2} \sigma_j^2 x_j^2 \frac{\partial^2 V_i}{\partial x_j^2}.$$

Simultaneously, the joint probability density $m(x_1, x_2, t)$ of the two countries’ military stocks evolves according to the FPK equation:

$$\frac{\partial m}{\partial t} = - \sum_{i=1}^2 \frac{\partial}{\partial x_i} \left[\left(\frac{1}{c_i} \frac{\partial V_i}{\partial x_i} - \delta_i x_i \right) m \right] + \frac{1}{2} \sum_{i=1}^2 \frac{\partial^2}{\partial x_i^2} [\sigma_i^2 x_i^2 m],$$

subject to initial distribution $m_0(x_1, x_2)$ and appropriate boundary conditions.

We now prove that this coupled HJB–FPK system admits a solution under regularity conditions. Define the map $\mathcal{T}$ taking value functions (V_1, V_2) to their induced feedback controls and resulting flow $m(t)$. Then, map the control flow

back to updated value functions through the HJB PDEs. This operator $\mathcal{T}$ is continuous and compact under suitable Sobolev spaces. Using Schauder’s fixed-point theorem, we deduce the existence of equilibrium (V_1, V_2, m) .

Moreover, stability can be shown using Lyapunov functionals derived from the entropy of m and cost functionals. For example,

$$\mathcal{L}[m] := \int \left(\sum_{i=1}^2 \log x_i - \gamma_i \log x_j \right) m(x_1, x_2, t) dx_1 dx_2,$$

and its dissipation over time under equilibrium strategies proves convergence.

Thus, the coupled HJB–FPK system fully characterizes the feedback Nash equilibrium in the stochastic arms race model.

5 Welfare, Deterrence, and Strategic Stability

In this section, we interpret the mathematical results of the feedback equilibrium system through the lens of welfare analysis, strategic deterrence, and long-run stability. These outcomes are especially relevant to real-world geopolitical tensions, such as those between the United States and China or the Cold War rivalry between the U.S. and the Soviet Union. The equilibrium solution derived from the coupled HJB–FPK system provides a probabilistic description of the joint arms distribution $m^*(x_1, x_2)$, and value functions $V_i(x_1, x_2)$ that quantify the long-run strategic utility of each player under feedback control.

From a welfare perspective, we define the expected utility of country i under equilibrium as

$$\mathbb{E}[V_i] = \int_{\mathbb{R}_+^2} V_i(x_1, x_2) m^*(x_1, x_2) dx_1 dx_2.$$

This expression captures the joint effect of optimal arms accumulation and the uncertainty of arms distribution. Welfare losses can be studied by comparing this expected utility under different scenarios, such as alternative diffusion intensities σ_i , cost parameters c_i , or interdependence penalties γ_i . An increase in γ_i , for example, implies greater disutility from the rival’s armament and leads to more aggressive own accumulation, which may reduce aggregate welfare for both parties due to higher investment costs and increased volatility.

Deterrence is evaluated by examining the slope of the value function $\partial V_i / \partial x_i$, which determines the marginal incentive to invest further in arms. When the value function becomes too convex (second derivative large), it signals a positive feedback loop of over-accumulation, leading to destabilizing behavior. In contrast, concavity indicates self-limiting behavior and sustainable deterrence. If equilibrium controls are stabilizing and bounded, then deterrence is robust. However, if feedback strategies imply infinite incentives for armament at some regions of the state space, deterrence collapses into an arms spiral.

Strategic stability is assessed using Lyapunov criteria applied to the Fokker–Planck evolution. A stationary solution m^* is said to be stochastically stable if

for any perturbation m_0 , the flow $m(t)$ returns to m^* over time. Let

$$\mathcal{L}(t) = \int_{\mathbb{R}_+^2} (V_1 + V_2)(x_1, x_2) m(t, x_1, x_2) dx_1 dx_2$$

be a total expected welfare functional. Then,

$$\frac{d}{dt} \mathcal{L}(t) \leq 0$$

under equilibrium dynamics, indicating global convergence. If this condition holds strictly except at m^* , the equilibrium is globally asymptotically stable.

Finally, policy interventions can be analyzed by introducing arms control treaties as tax-like constraints on controls, or joint payoff maximization under cooperative bargaining. These extensions shift the feedback structure and alter the equilibrium distribution and value functions. The analysis suggests that moderate controls on arms buildup, or mutual information exchange, can raise expected welfare and reinforce stability.

6 The Linear-Quadratic Benchmark and Analytical Feedback Solutions

To obtain analytical insights into the structure of feedback equilibria in the arms race between two strategic countries, we now turn to the classical linear-quadratic (LQ) benchmark. This model, though stylized, allows us to extract explicit characterizations of best-response functions and equilibrium strategies through Riccati equations and closed-form expressions. Such tractable formulations are especially useful in assessing how changes in model primitives such as strategic fear, arms efficiency, or uncertainty affect behavior and equilibrium welfare.

We consider two countries indexed by $i = 1, 2$. Each country possesses a state variable $x_i(t)$, representing its level of military stock or arms capital at time t . The stochastic dynamics of $x_i(t)$ are governed by geometric-type diffusion:

$$dx_i(t) = (a_i x_i(t) + b_i u_i(t)) dt + \sigma_i x_i(t) dW_i(t),$$

where $u_i(t)$ is the control variable representing investment in arms, $a_i < 0$ denotes the depreciation rate of military stock, $b_i > 0$ scales investment effectiveness, and $\sigma_i > 0$ captures the intensity of multiplicative uncertainty in the arms accumulation process. The Brownian motions $W_i(t)$ are independent for each player.

Each country seeks to maximize its discounted utility flow given by the quadratic criterion:

$$J_i = \mathbb{E} \int_0^\infty e^{-\rho_i t} \left(q_i x_i^2(t) - \gamma_i x_j^2(t) - \frac{1}{2} r_i u_i^2(t) \right) dt,$$

where $q_i > 0$ captures strategic benefits from own military strength, $\gamma_i > 0$ reflects fear or strategic disutility from the rival's arms, and $r_i > 0$ is the marginal cost of military investment. The discount rate $\rho_i > 0$ reflects intertemporal preferences. This setup defines a noncooperative differential game with coupled payoff structures and multiplicative stochastic dynamics.

We now conjecture that each country uses a *linear feedback policy* of the form $u_i(t) = -F_i x_i(t)$, with $F_i > 0$ representing the feedback gain to be determined endogenously. Substituting this into the dynamics yields:

$$dx_i(t) = (a_i - b_i F_i) x_i(t) dt + \sigma_i x_i(t) dW_i(t),$$

i.e., a geometric Brownian motion with drift $a_i - b_i F_i$ and volatility σ_i . Under this policy, the state process evolves with stability if $a_i - b_i F_i < 0$, which imposes an upper bound on feedback aggressiveness.

We now turn to the value function associated with this feedback policy. We conjecture a quadratic value function:

$$V_i(x_1, x_2) = \frac{1}{2} P_i x_i^2 + \frac{1}{2} S_i x_j^2,$$

where $P_i > 0$ and $S_i > 0$ are coefficients to be determined. The Hamilton-Jacobi-Bellman (HJB) equation for player i reads:

$$\rho_i V_i = \max_{u_i} \left\{ q_i x_i^2 - \gamma_i x_j^2 - \frac{1}{2} r_i u_i^2 + V_{i,x_i} \cdot (a_i x_i + b_i u_i) + \frac{1}{2} \sigma_i^2 x_i^2 V_{i,x_i x_i} \right\}.$$

Using the quadratic form, we compute:

$$V_{i,x_i} = P_i x_i, \quad V_{i,x_i x_i} = P_i.$$

Substituting into the HJB and solving the first-order condition gives the optimal feedback:

$$u_i^* = -\frac{b_i P_i}{r_i} x_i = -F_i x_i.$$

This yields the feedback gain:

$$F_i = \frac{b_i P_i}{r_i}.$$

Substituting back into the HJB, we obtain the *Riccati equation* for P_i :

$$\rho_i P_i = q_i - \frac{b_i^2 P_i^2}{2r_i} + 2a_i P_i + \sigma_i^2 P_i.$$

This is a nonlinear algebraic equation in P_i , with one positive root under standard parameter assumptions (e.g., small σ_i , moderate r_i). The coefficient S_i evolves according to:

$$\rho_i S_i = -\gamma_i + 2a_j S_i + \sigma_j^2 S_i.$$

This linear equation in S_i has the closed-form solution:

$$S_i = \frac{\gamma_i}{\rho_i - 2a_j - \sigma_j^2},$$

assuming the denominator is positive, which guarantees the discounting of opponent's threat over time.

Having solved for P_i and S_i , we obtain full characterizations of both the value function and the feedback strategy. The Riccati solution encapsulates the effect of all key parameters: the marginal utility of arms q_i , strategic threat γ_i , cost of arms r_i , discount rate ρ_i , and uncertainty σ_i .

This LQ framework offers a clean window into strategic interactions. In particular, we can rigorously analyze comparative statics of F_i with respect to γ_i , r_i , or σ_i . A rise in γ_i leads to a higher S_i , which, under conjectural interdependence, may prompt i to increase its own arms accumulation. Similarly, decreasing r_i (e.g., via military innovation) raises F_i , intensifying the arms race.

Finally, note that the stochastic dynamic under feedback control,

$$dx_i(t) = \left(a_i - \frac{b_i^2 P_i}{r_i} \right) x_i(t) dt + \sigma_i x_i(t) dW_i(t),$$

leads to a stationary log-normal distribution if the drift is negative. Thus, the system converges to a steady-state arms distribution whose mean and variance can be computed from known formulas. This enables welfare and stability analysis in Section 5 to be extended quantitatively.

7 Dynamic Stability, Distributional Convergence, and Arms Race Pathways

This section explores the long-run stability of the coupled arms race system by analyzing the dynamic interaction of the stochastic differential equations and their associated Fokker-Planck (FPK) equations under feedback equilibrium. We focus on the convergence behavior of the distribution of arms stocks, the ergodicity of the resulting stochastic process, and the dynamic pathways through which stable or unstable arms race equilibria emerge.

We begin by considering the closed-loop controlled state dynamics for each country $i \in \{1, 2\}$ under its optimal feedback policy:

$$dx_i(t) = \left(a_i - \frac{b_i^2 P_i}{r_i} \right) x_i(t) dt + \sigma_i x_i(t) dW_i(t).$$

This is a geometric Brownian motion (GBM) process with drift $\mu_i := a_i - \frac{b_i^2 P_i}{r_i}$ and diffusion σ_i . The condition $\mu_i < 0$ ensures the exponential ergodicity of the arms distribution. The stationary density $p_i^\infty(x)$ satisfies the steady-state Fokker-Planck equation:

$$0 = -\frac{d}{dx} [\mu_i x p_i(x)] + \frac{1}{2} \frac{d^2}{dx^2} [\sigma_i^2 x^2 p_i(x)],$$

whose closed-form solution is given by the inverse-gamma distribution:

$$p_i^\infty(x) = C_i x^{-2\left(1+\frac{\mu_i}{\sigma_i^2}\right)} \exp\left(-\frac{K_i}{x^2}\right), \quad x > 0,$$

with C_i as a normalization constant and K_i determined by integrating factors. The fat-tailed nature of this distribution implies the persistence of rare arms surges, consistent with empirical observations of arms races marked by long quiet periods punctuated by bursts of military buildup.

Next, we analyze the time evolution of the density function $p_i(t, x)$ via the time-dependent Fokker-Planck equation:

$$\frac{\partial p_i}{\partial t} = -\frac{\partial}{\partial x} [\mu_i x p_i] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma_i^2 x^2 p_i].$$

This degenerate parabolic PDE governs the probabilistic evolution of the arms stock and ensures convergence toward $p_i^\infty(x)$ under mild conditions on the initial data. Analytical properties such as mass preservation, positivity, and entropy decay can be established using techniques from the theory of nonlinear Fokker-Planck equations.

Importantly, the coupled nature of the value functions and feedback policies implies that μ_i and P_i depend on the opponent's parameters. For example, increases in γ_i (fear from opponent arms) raise S_i in the value function, thereby affecting P_i through the strategic interdependence in Riccati equations. This induces indirect interactions in the FPK system even though the diffusion is idiosyncratic.

We also examine conditions for stability and multiplicity of equilibrium distributions. When μ_i becomes positive due to low arms cost or high threat sensitivity, the GBM explodes, and no stationary distribution exists, implying an unstable arms race spiral. The bifurcation between convergence and explosion thresholds can be explicitly characterized by:

$$a_i < \frac{b_i^2 P_i}{r_i} \quad \Leftrightarrow \quad \text{Stochastic divergence.}$$

This provides a sharp condition for destabilizing parameter regimes.

Numerical simulations confirm that under reasonable parameterizations, the arms race process converges to a stable stationary distribution for both countries, with average arms levels determined by the ratio of drift to volatility and cost parameters. Feedback control dampens fluctuations and prevents unbounded accumulation, unless fear terms γ_i or military productivity b_i are extremely large.

Lastly, we note that these dynamics illustrate the potential for peaceful deterrence equilibria. Even in a competitive setting, the endogenous Riccati stabilization permits bounded arms competition under rational policy. However, this equilibrium remains fragile to shocks in perceived threats, cost functions, or diffusion volatility offering fertile ground for future studies of tipping points and destabilization.

8 Conclusion and Extensions

This paper presents a rigorous mathematical analysis of a stochastic differential game framework for modeling arms races between two rival states, such as the United States and China or the Cold War-era United States and Soviet Union. The central novelty of the framework lies in its use of logarithmic utility functions to represent strategic preferences over arms accumulation. Each country derives utility from its own stock of military power (reflecting deterrence, prestige, or influence) and suffers disutility from the opposing country's arms buildup (capturing perceived threat, regional insecurity, or fear of domination). These preferences are embedded in a dynamic stochastic setting, where each country's control variable—the rate of arms accumulation—is optimized over time, taking into account the state of the system and the anticipated behavior of its rival.

The core of the analysis develops the feedback Nash equilibrium under log utility, using tools from stochastic optimal control theory, Hamilton–Jacobi–Bellman (HJB) equations, and Fokker–Planck (FPK) partial differential equations. We derived the dynamic programming equations for both players in the form of coupled HJBs, and analyzed the stochastic evolution of the joint arms distribution through the corresponding coupled FPK equations. Log utility introduces curvature that generates nonlinear marginal returns to armament, making closed-form characterization difficult—but allowing for rich endogenous dynamics, including bounded strategic aggressiveness and diminishing incentives at high arms levels.

We then showed that, under reasonable parameter values (arms productivity, volatility, cost, fear), the resulting stochastic system admits well-posed equilibrium feedback strategies. These strategies are time-consistent, subgame perfect, and yield stable probabilistic trajectories. The FPK system converges under appropriate boundary and integrability conditions to stationary distributions over arms stocks, whose ergodicity and fat-tailedness capture both typical competition and rare escalatory surges. This log utility framework thus accommodates both gradual deterrence equilibria and probabilistic instability due to parameter shifts or shocks.

To make the results more analytically transparent and allow for policy-relevant comparative statics, Sections 6 and 7 introduced a linear-quadratic (LQ) benchmark. In this tractable setting, the value functions are quadratic in arms, the strategies are linear in the state, and feedback Nash equilibria can be characterized through coupled Riccati equations. These allow us to quantify the strategic aggressiveness of countries, the sensitivity of equilibrium arms levels to cost and fear parameters, and the boundary between stability and explosive arms dynamics. Despite its simplicity, the LQ model confirms and complements the qualitative findings of the log utility framework.

Overall, our results have several important theoretical and practical implications. First, they show that rational arms competition can remain probabilistically stable even without explicit cooperation, provided both players internalize the diminishing returns to aggression and the rising cost of overaccumulation.

Second, the feedback nature of equilibrium strategies enables dynamic adaptation to the evolving state of the system, a feature missing in simpler open-loop models. Third, the coupled HJB–FPK structure captures both the value-based optimization and the probabilistic evolution of the system, offering a unified tool for welfare, stability, and risk analysis.

This paper paves the way for a number of valuable extensions. One direction is the endogenization of beliefs and learning—modeling each player as uncertain about the rival’s intentions and updating beliefs over time. This would lead to stochastic differential games with Bayesian learning and belief-dependent controls. Another extension involves multipolar arms dynamics or regional clusters, using the framework of mean field games or networked differential games to analyze strategic spillovers across alliances or rivalries. Third, the state space could be expanded to include fiscal, economic, or ideological state variables, allowing for a broader analysis of strategic interactions beyond arms.

Finally, the mathematical foundation of this framework—including log utility optimization, stochastic dynamics, feedback equilibrium, and PDE analysis—is broadly applicable to other strategic domains, such as climate change negotiations, economic technology races, or even ideological competitions. The methods and findings developed here provide a solid basis for further research on the dynamics of competition, stability, and deterrence in complex strategic environments.

9 References

- Achdou, Y., Laurière, M., and Lions, P.-L. (2020). Mean Field Games and Applications. In A. Bensoussan and Q. Zhang (Eds.), *Handbook of Reinforcement Learning and Control*. Springer.
- Arsov, I., and Nourian, M. (2022). A Numerical Method for Feedback Nash Equilibria in Differential Games with Applications to Economic Models. *Computational Economics*, 59(4), 1189–1223.
- Bardi, M., and Capuzzo-Dolcetta, I. (1997). *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*. Birkhäuser.
- Başar, T., and Olsder, G. J. (1999). *Dynamic Noncooperative Game Theory* (2nd ed.). Society for Industrial and Applied Mathematics (SIAM).
- Bracken, J., and Kilgour, D. M. (1977). Evolution of Strategic Stability and Deterrence Theory. *World Politics*, 30(1), 61–80.
- Brodie, B. (1959). *Strategy in the Missile Age*. Princeton University Press.
- Dockner, E. J., Jørgensen, S., Long, N. V., and Sorger, G. (2000). *Differential Games in Economics and Management Science*. Cambridge University Press.
- Evans, L. C. (2010). *Partial Differential Equations* (2nd ed.). American Mathematical Society (Graduate Studies in Mathematics, Vol. 19).
- Friedman, A. (1971). *Differential Games*. Wiley-Interscience.
- Hartl, R. F., Sethi, S. P., and Vickson, R. G. (1995). A Survey of the Maximum Principles for Optimal Control Problems with State Constraints. *SIAM*

Review, 37(2), 181–218.

Howard, M. (1980). *Strategic Deception in the Second World War*. W. W. Norton & Company.

Jervis, R. (1984). *The Illogic of American Nuclear Strategy*. Cornell University Press.

Kydland, F. E., and Prescott, E. C. (1977). Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, 85(3), 473–492.

Lasry, J.-M., and Lions, P.-L. (2007). Mean Field Games. *Japanese Journal of Mathematics*, 2(1), 229–260.

Morton, D. P., and Pan, F. (2006). A Stochastic Programming Approach to Strategic Planning for Homeland Defense. *Operations Research*, 54(5), 789–804.

Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. (1962). *The Mathematical Theory of Optimal Processes*. Interscience Publishers.

Sagan, S. D. (1993). *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons*. Princeton University Press.

Schelling, T. C. (1960). *The Strategy of Conflict*. Harvard University Press.

Schelling, T. C. (1966). *Arms and Influence*. Yale University Press.

Zou, H.-F. (1995). A Dynamic Model of Capital and Arms Accumulation. *Journal of Economic Dynamics and Control*, 19(1–2), 371–393.

Zou, H.-F. (1997). Dynamic Analysis in the Viner Model of Mercantilism. *Journal of International Money and Finance*, 16(4), 637–651.