Bitcoin Exit Dominance in Monetary Coordination Games

Sean Hash
bitcoingametheory.com
sean@bitcoingametheory.com

February 2026

Working Paper

JEL Codes: C72, C73, D83, E42, E51, G11, L14

Keywords: Bitcoin, game theory, Nash equilibrium, monetary coordination, neutral settlement, absorbing state, coordination failure, threshold strategy, digital scarcity, Schelling point, network effects

Abstract

In a multipolar world with no trusted monetary coordinator, how do rational actors settle large-value transactions across trust boundaries? We model this as a non-cooperative game—the "Exit Game"—in which capital allocators choose between capturable settlement systems ("Stay") and neutral settlement ("Exit"). The model rests on four empirical axioms: persistent multipolarity, rational self-interest, computational hardness, and network effect persistence. We prove three results. First, the payoff advantage of Exit over Stay is strictly increasing in adoption: every term in the payoff differential favors Exit under maintained monotonicity conditions, and each actor's adoption threshold approaches zero under structural debasement (Theorem 1). Second, no coalition can sustain coordinated Stay, because permissionless access makes defection costless and the first defector captures fleeing capital (Theorem 2). Third, the resulting equilibrium is absorbing: the monotone adoption process converges to full adoption once a critical mass is reached, because trust conditions required for coordinated return cannot be re-established (Theorem 3). The model is explicitly falsifiable: six conditions are identified under which the central claims would fail. Bitcoin is the unique asset satisfying the necessary properties for neutral settlement—a result proved by systematic elimination across seven asset classes in Hash (2026b).

1. Introduction

The global monetary system has no referee. Multiple sovereign powers compete for economic influence, each controlling monetary instruments that serve domestic policy objectives at the expense of foreign holders. This creates a coordination problem that is easy to state and impossible to solve within existing institutions: in a world where no single entity can credibly commit to monetary neutrality, how do rational actors settle large-value transactions across trust boundaries?

This question is structural, not conjunctural. Sovereign debt levels have reached historical extremes—global debt-to-GDP exceeded 300% in 2024 (IMF, 2024). No nation has repaid debt exceeding 120% of GDP through taxation alone. Debasement is the historically observed path, and debasement is value extraction from holders. The coordination problem is not whether states *will* debase, but whether any mechanism exists through which actors can settle *despite* debasement.

The literature addresses components of this problem—mining incentives (Biais et al., 2019; Eyal and Sirer, 2014), attack economics (Budish, 2018), price-security dynamics (Pagnotta, 2022; Chen, 2025)—but does not integrate these into a unified strategic analysis from the allocator's perspective. Brunnermeier, James, and Landau (2019) identify the coordination failure but propose institutional solutions rather than analyzing the equilibrium outcome. This paper addresses the gap by formalizing the allocator's strategic problem directly.

We formalize the allocator's problem as a non-cooperative game—the "Exit Game"—and prove that the payoff advantage of Exit over Stay is strictly increasing in adoption. Every term in the payoff differential favors Exit under maintained monotonicity conditions. Each actor faces an adoption threshold that approaches zero when the capturable system delivers negative real returns—which, under structural debasement (Assumption 1), it does.

We also prove that no coalition can sustain Stay (Theorem 2) and that the resulting equilibrium is absorbing (Theorem 3). Bitcoin satisfies the seven necessary properties derived here—a result established by systematic elimination in Hash (2026b). A third paper (Hash, 2026c) extends the analysis to autonomous AI agents, for whom neutral settlement is not a preference but a structural requirement.

2. Related Work

2.1 Game Theory of Proof-of-Work Consensus

The foundational result is Biais, Bisiere, Bouvard, and Casamatta (2019): following the longest chain constitutes a Markov Perfect Equilibrium in a stochastic game among miners, and subsequent work largely builds on this framework. Eyal and Sirer (2014) complicate this picture by identifying conditions under which selfish mining is profitable—but their threshold (a miner controlling more than one-third of hashrate) has never been approached in practice, which says more about the incentive structure than about the theoretical vulnerability. Budish (2018) argues that attack cost scales with network value. This analysis does not fully account for the fee market: a network processing $1 billion in daily settlement generates security budget independent of token price, a gap that Chen (2025) partially fills with the RESUNE framework.

2.2 Monetary Competition and Coordination

Brunnermeier, James, and Landau (2019) is the essential reference on digital currency competition in a multipolar world. Their core argument— that competing monetary systems create coordination failures traditional institutions cannot resolve—is precisely our starting point. Where we diverge: Brunnermeier et al. treat the coordination failure as a problem to be solved by better institutional design. We treat it as a permanent structural feature (Assumption 1) that selects for a non-institutional solution. Huberman, Leshno, and Moallemi (2021) contribute the memorable characterization of Bitcoin as a "monopoly without a monopolist," directly supporting the neutrality property we formalize below.

2.3 Monotone Comparative Statics

The monotonicity conditions (M1)–(M5) that drive our main result belong to the framework of monotone comparative statics developed by Milgrom and Shannon (1994) and Topkis (1998). In their language, the payoff differential $\Delta_i(p)$ exhibits increasing differences in $(p, \text{adoption decision})$ under our maintained conditions—precisely the supermodularity property that guarantees monotone best responses. Our contribution is applying this machinery to the monetary coordination problem rather than to the firm and market settings where it originated.

2.4 Network Effects and Focal Points

Katz and Shapiro (1985, 1986) provide the theoretical machinery for our network effect assumption. What matters for our purposes is a specific implication they do not emphasize: in monetary networks, switching costs compound because the network's value proposition *is* its adoption. Schelling (1960) provides the concept of focal points—though applying focal point theory to money requires more care than the literature typically exercises. We argue that the focal point is determined not by precedence but by structural resistance to the "who benefits?" objection.

We note the distinction between our "Exit" concept and Hirschman's (1970) "exit" from organizations. In Hirschman's framework, exit is one of three responses to organizational decline (alongside voice and loyalty). Our Exit is a strategic action in a coordination game—adopting neutral settlement— and does not require leaving an institution, only diversifying settlement infrastructure.

3. Model

3.1 Maintained Assumptions

The analysis rests on four empirical axioms. Each is independently assessable. All subsequent claims derive from these assumptions; rejecting any axiom invalidates the specific claims that depend on it.

Assumption 1 (Multipolarity)**.** No single entity permanently governs global economic activity. Power distributes across competing centers with superlinear coordination costs.

Falsification: A single coordinator achieves permanent, stable control over global monetary policy with sublinear coordination costs (F1).

Assumption 2 (Rational Self-Interest)**.** Actors optimize for self-interest. When defection from cooperative agreements is unpunished or unpunishable, actors defect.

Falsification: Stable cooperation persists indefinitely among self-interested actors without enforcement mechanisms (F3).

Assumption 3 (Computational Hardness)**.** Certain mathematical problems remain computationally intractable. Digital scarcity and cryptographic custody are implementable.

Assumption 4 (Network Effect Persistence)**.** Past critical mass, switching costs exceed marginal gains of alternatives. Incumbency compounds (Katz and Shapiro, 1985).

3.2 Definitions

Definition 1 (Neutral Settlement)**.** An asset S is a *neutral settlement asset* if and only if S is immune to seizure, immune to debasement, and immune to political capture by any single actor or coordinated coalition.

Definition 2 (Exit)**.** The action of moving capital from a capturable system to a neutral settlement asset.

Definition 3 (Capture Surface)**.** For a monetary system with parameter space Θ, the capture surface for actor A is $CS_A = \{\theta \in \Theta : A \text{ can unilaterally change } \theta\}$. A system is neutral if $|CS_A| = 0$ for all A. The terminology adapts "attack surface" from security engineering to institutional capture (cf. Stigler, 1971 on regulatory capture).

Definition 4 (Absorbing State)**.** A state p^* is absorbing if $P(p_{t+1} = p^* \mid p_t = p^*) = 1$. In the monetary coordination context, the transition to $p^* = 1$ destroys the trust conditions required for reversal, making the state absorbing in the standard Markov chain sense (cf. Arthur, 1989 on lock-in).

3.3 Necessary Properties

From Definition 1, we derive seven necessary properties for any neutral settlement asset. Each blocks a specific attack class:

D1 Requirement	Attack If Missing	Required Property
Immune to seizure	Physical confiscation	P6: Informational security
Immune to seizure	Transaction censorship	P3: Permissionless access
Immune to debasement	Supply inflation	P5: Absolute scarcity
Immune to debasement	Protocol rule change	P1: Protocol security
Immune to capture	Governance takeover	P2: Neutrality
Functions as settlement	Prohibitive cost/delay	P4: Cheap finality
Survives future threats	Obsolescence	P7: Adaptive resilience

Proposition 1 (Necessity)**.** *P1–P7 are individually necessary. Removing any single property enables the corresponding attack, violating Definition 1.*

Proof. By construction from the derivation table. □

Bitcoin is the unique asset satisfying P1–P7—the systematic elimination across seven asset classes appears in Hash (2026b).

3.4 Game Structure

We model monetary coordination as a non-cooperative game:

Game $G = (N, S, u)$ where:

- $N = \{1, 2, \ldots, n\}$ is the set of capital allocators.
- $S_i = \{\text{Exit}, \text{Stay}\}$ is the strategy set for player i.
- $u_i : S \to \mathbb{R}$ is the utility function for player i.

Let $p \in [0, 1]$ denote the fraction of total capital that has chosen Exit. We assume R_B, σ_B, K_A, R_A, and K_N are continuous functions of p, bounded on $[0, 1]$. The utility functions are:

$$u_i(\text{Exit}, s_{-i}) = R_B(p) - \lambda_i \cdot \sigma_B(p) - \kappa_i \cdot K_A(p) - \rho_i \cdot R_A(p)$$

$$u_i(\text{Stay}, s_{-i}) = R_F - \lambda_i \cdot \sigma_F - K_N(p)$$

Here $K_A(p)$ is the adoption penalty—the switching cost of moving to neutral settlement, analogous to technology adoption costs in network economics (Katz and Shapiro, 1985; Farrell and Saloner, 1985). $K_N(p)$ is the non-adoption penalty—the opportunity cost imposed on non-adopters as the competing network grows, a form of negative network externality.

Maintained monotonicity conditions:

(M1) $R'_B(p) > 0$ — Network effects increase return (Assumption 4)

(M2) $\sigma'_B(p) < 0$ — Deeper markets reduce volatility

(M3) $K'_A(p) < 0$ — Adoption penalty falls with adoption

(M4) $R'_A(p) < 0$ — Regulatory penalty decreases as adoption normalizes compliance

(M5) $K'_N(p) > 0$ — Non-adoption penalty rises as competitors exit

4. Main Results

4.1 Exit Dominance

Theorem 1 (Monotone Exit Advantage)**.** *Under Assumptions 1–4 and maintained conditions (M1)–(M5):*

(i) The payoff differential $\Delta_i(p)$ is strictly increasing in p for all player types.

(ii) For each player i, there exists a threshold $p_i^* \in [0, 1]$ such that Exit is the unique best response for $p > p_i^*$.

(iii) Under Assumption 1 ($R_F < 0$), thresholds cluster near zero for actors with typical risk preferences.

Proof. Define the payoff differential:

$$\Delta_i(p) = u_i(\text{Exit}) - u_i(\text{Stay})$$
$$= [R_B(p) - R_F] - \lambda_i[\sigma_B(p) - \sigma_F] - \kappa_i \cdot K_A(p) - \rho_i \cdot R_A(p) + K_N(p)$$

Taking the derivative with respect to p:

$$\frac{d\Delta_i}{dp} = R'_B(p) - \lambda_i \cdot \sigma'_B(p) - \kappa_i \cdot K'_A(p) - \rho_i \cdot R'_A(p) + K'_N(p)$$

Under the maintained monotonicity conditions (M1)–(M5), each term is positive:

Term	Sign	Reason
$R'_B(p)$	> 0	(M1)
$-\lambda_i \cdot \sigma'_B(p)$	> 0	$\lambda_i > 0$, $\sigma'_B(p) < 0$ by (M2)
$-\kappa_i \cdot K'_A(p)$	> 0	$\kappa_i \geq 0$, $K'_A(p) < 0$ by (M3)
$-\rho_i \cdot R'_A(p)$	> 0	$\rho_i \geq 0$, $R'_A(p) < 0$ by (M4)
$K'_N(p)$	> 0	(M5)

Therefore $d\Delta_i/dp > 0$: the advantage of Exit is strictly increasing in p.

For each player i, there exists a threshold p_i^* satisfying $\Delta_i(p_i^*) = 0$ (by the intermediate value theorem, since Δ_i is continuous). For $p > p_i^*$, $\Delta_i(p) > 0$ and Exit strictly dominates.

Under Assumption 1, $R_F < 0$ in real terms (sovereign debasement is structural—no nation has repaid debt exceeding 120% of GDP through taxation alone; IMF, 2024). This ensures that even at $p = 0$, the first term $[R_B(0) - R_F]$ provides positive contribution, lowering p_i^* toward zero for many actor types.

For the marginal allocation argument: even for actors with $p < p_i^*$, a small allocation $\varepsilon > 0$ provides $u_i(w_i = \varepsilon) > u_i(w_i = 0)$ because the first-order gain from network participation exceeds the first-order volatility cost for sufficiently small ε. This holds for a wide range of parameter values. □

Remark. The game matrix at the individual level:

	Others: Stay	Others: Exit
Player i: Stay	$R_F < 0$	$R_F - K_N(p)$
Player i: Exit	$R_B(0) - K_A(0)$	$R_B(p)$

Under Assumption 1 ($R_F < 0$), Exit yields higher payoff in both columns for actors whose threshold p_i^* is at or below the current adoption level.

4.2 Coordination Failure

Theorem 2 (Coalition Instability)**.** *No coalition can sustain coordinated Stay when Exit is available and unpunishable.*

Proof. Consider a coalition $C \subseteq N$ maintaining Stay. For this coalition to be self-enforcing in the sense of Bernheim, Peleg, and Whinston (1987), no member $j \in C$ can profit from unilateral defection.

But by Theorem 1, $u_j(\text{Exit}) > u_j(\text{Stay})$ for j when p exceeds p_j^*. Property P3 (permissionless access) ensures the defection cost is approximately zero—no institutional cooperation is required to initiate settlement.

The coalition faces a standard free-rider problem: the benefit of collective Stay is a public good (maintained value of capturable system), but the benefit of individual defection is a private good (captured network effects, avoidance of debasement). Under Assumption 2, the private good dominates.

The dynamics compound: the first actor to defect captures fleeing capital, raising $K_N(p)$ for remaining coalition members and accelerating further defection.

Formally:

$$u_j(\text{defect from } C) > u_j(\text{remain in } C)$$

for all $j \in C$, making the all-Stay profile not coalition-proof. □

Remark. The first-mover dynamic is empirically observable: El Salvador adopted Bitcoin as legal tender in 2021 (Bukele, 2021); Singapore issued 13 digital payment token licenses in 2024 (MAS, 2024). Each sovereign defection raises $K_N(p)$ for remaining coalition members.

4.3 Absorbing State

Theorem 3 (Absorption)**.** *The state* $(\textit{Exit}, \textit{Exit})$ *is absorbing: once entered, there exists no feasible sequence of individual actions leading back to* $(\textit{Stay}, \textit{Stay})$*.*

Proof. Define the aggregate adoption process $\{p_t\}_{t\geq 0}$ on state space $[0, 1]$.

For $p_t \geq p_c$ (critical mass threshold), we show that $p_{t+1} \geq p_t$ with probability 1.

Step 1 (No Exit-to-Stay switching). For any actor currently at Exit: switching to Stay incurs infrastructure lock-in costs (custody solutions, regulatory frameworks, accounting standards) and $\Delta_i(p_t) > 0$ by Theorem 1. No rational actor switches from Exit to Stay.

Step 2 (Positive drift). For actors currently at Stay: $\Delta_i(p_t)$ is increasing in p_t (shown in Theorem 1). As p_t increases, more actors cross their thresholds p_i^*, increasing p_{t+1}.

Step 3 (Convergence). The process $\{p_t\}$ is monotone non-decreasing and bounded above by 1. By the monotone convergence theorem (see Fudenberg and Tirole, 1991, Ch. 12), $p_t \to p^*$ for some $p^* \leq 1$.

If $p^* < 1$, then—assuming the distribution of thresholds $\{p_i^*\}$ has continuous support on $[0, 1]$ (i.e., no gap in actor types)—there exists an actor j at Stay with $\Delta_j(p^*) > 0$. This contradicts stationarity at p^*. Therefore $p^* = 1$.

The state $p = 1$ is absorbing: $P(p_{t+1} = 1 \mid p_t = 1) = 1$, since return coordination requires re-establishing trust in a capturable system, but exiting *is* the revelation that the system was capturable—and revelations do not un-reveal. A return cartel faces the same coordination failure as a stay cartel (Theorem 2). □

4.4 Cascade Dynamics

Proposition 2 (Adoption Cascades)**.** *Given Theorems 1–3, adoption occurs in bursts (cf. Bikhchandani, Hirshleifer, and Welch, 1992 on informational cascades; our mechanism is payoff-driven rather than information-driven).*

The cascade operates through the interaction of $K_A(p)$ and $K_N(p)$: as p increases, the adoption penalty falls and the non-adoption penalty rises, pushing additional actors past their thresholds p_i^*. This creates punctuated equilibria: periods of stability followed by rapid adoption bursts when a critical mass of actors simultaneously crosses their thresholds.

Comparative statics. By the implicit function theorem applied to $\Delta_i(p_i^*) = 0$:

$$\frac{dp_i^*}{d\lambda_i} = -\frac{\partial \Delta_i / \partial \lambda_i}{d\Delta_i / dp} > 0$$

More risk-averse actors have higher thresholds (they wait longer). But they still eventually exit because $d\Delta_i/dp > 0$ pushes the advantage past any finite risk aversion as $p \to 1$.

5. Falsification

Note what does *not* falsify the model: price declines, regulatory actions against specific intermediaries, developer controversies, or energy consumption debates. These affect adoption

ID	Condition	Breaks
F1	Global coordination cost becomes sublinear	Assumption 1
F2	An alternative asset satisfies P1–P7 (tested in Hash, 2026b)	Uniqueness
F3	(Stay, Stay) is stable equilibrium when Exit exists	Theorem 1
F4	Stable cartel prevents Exit indefinitely	Theorem 2
F5	Quantum computing breaks cryptographic primitives before migration	Assumption 3
F6	AI agents gain enforceable legal personhood (tested in Hash, 2026c)	Limiting case

timing, not equilibrium structure.

6. Limitations

The binary Exit/Stay model is a simplification. Real actors face continuous allocation decisions—a pension fund does not choose between 0% and 100% but between 0%, 1%, 5%. The marginal allocation argument (Section 4.1) partially addresses this, but a proper continuous-strategy extension with portfolio optimization would strengthen the formalism. This is the most natural direction for future work.

A deeper concern: the monotonicity conditions (M1)–(M5) are empirically motivated but not derived from first principles. Each could be violated in specific regimes (e.g., $R'_B(p)$ could be negative during a network congestion crisis). The model's robustness depends on these conditions holding *on average and in the long run*, not at every instant. Formalizing this distinction—perhaps through time-averaged conditions— would strengthen the theoretical foundation.

7. Conclusion

The argument rests on four axioms: multipolarity is persistent, actors are self-interested, cryptography works, and network effects compound. From these we derive a monotone increasing payoff advantage, coalition instability, and absorbing dynamics. Bitcoin is the unique asset serving as the Exit destination (Hash, 2026b); autonomous agents face the strongest version of this convergence pressure (Hash, 2026c). Six falsification conditions across the series identify when and how these claims fail.

References

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *Economic Journal*, 99(394), 116–131.

Bernheim, B. D., Peleg, B., and Whinston, M. D. (1987). Coalition-proof Nash equilibria I. Concepts. *Journal of Economic Theory*, 42(1), 1–12.

Biais, B., Bisiere, C., Bouvard, M., and Casamatta, C. (2019). The blockchain folk theorem. *Review of Financial Studies*, 32(5), 1662–1715.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5), 992–1026.

Brunnermeier, M. K., James, H., and Landau, J.-P. (2019). *The Digitalization of Money*. NBER Working Paper 26300.

Budish, E. (2018). The economic limits of Bitcoin and the blockchain. *NBER Working Paper* No. 24717.

Bukele, N. (2021). Bitcoin Law. Legislative Decree No. 57, Republic of El Salvador.

Chen, Y. (2025). A game-theoretic foundation for Bitcoin's price. Working paper.

Eyal, I. and Sirer, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. *Communications of the ACM*, 61(7), 95–102.

Farrell, J. and Saloner, G. (1985). Standardization, compatibility, and innovation. *RAND Journal of Economics*, 16(1), 70–83.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. MIT Press.

Hash, S. (2026b). Bitcoin as unique neutral settlement: A seven-property elimination. Working paper, bitcoingametheory.com.

Hash, S. (2026c). Settlement at zero trust: Bitcoin and autonomous economic agents. Working paper, bitcoingametheory.com.

Hirschman, A. O. (1970). *Exit, Voice, and Loyalty*. Harvard University Press.

Huberman, G., Leshno, J. D., and Moallemi, C. (2021). Monopoly without a monopolist. *Review of Financial Studies*, 34(7), 3266–3302.

International Monetary Fund (2024). World Economic Outlook Database.

Katz, M. L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, 75(3), 424–440.

Katz, M. L. and Shapiro, C. (1986). Technology adoption in the presence of network externalities. *Journal of Political Economy*, 94(4), 822–841.

Milgrom, P. and Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1), 157–180.

Monetary Authority of Singapore (2024). Digital Payment Token Licensing Statistics.

Pagnotta, E. S. (2022). Decentralizing money: Bitcoin prices and blockchain security. *Review of Financial Studies*, 35(2), 866–907.

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

Stigler, G. J. (1971). The theory of economic regulation. *Bell Journal of Economics and Management Science*, 2(1), 3–21.

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.

Notation

Symbol	Definition
G	Settlement game (N, S, u)
N	Set of capital allocators
S_i	Strategy set: $\{\text{Exit}, \text{Stay}\}$
u_i	Utility function
p	Fraction of capital at Exit
p_i^*	Threshold for player i
$\Delta_i(p)$	Payoff differential: $u_i(\text{Exit}) - u_i(\text{Stay})$
$R_B(p)$	Expected real return on neutral settlement asset
R_F	Expected real return on capturable assets
$\sigma_B(p)$	Volatility of neutral settlement asset
σ_F	Volatility of capturable assets
$K_A(p)$	Adoption penalty
$R_A(p)$	Regulatory penalty (decreases as regulatory clarity improves)
$K_N(p)$	Non-adoption penalty
$\lambda_i, \kappa_i, \rho_i$	Risk aversion, career penalty, regulatory penalty
CS_A	Capture surface for actor A: $\{\theta \in \Theta : A \text{ can unilaterally change } \theta\}$
(M1)–(M5)	Monotonicity conditions