

Marx, Revolution, and the Credibility Problem

Pablo Paniagua

Universidad del Desarrollo (UDD) - King's College London (KCL)

Abstract

Why have Marxist revolutions consistently produced authoritarian regimes rather than the liberated society Marx envisioned? This paper argues that revolutionary failure is not contingent but structurally determined by commitment problems inherent to the Marxist theoretical and philosophical program. Drawing on the self-enforcing constitutions literature and the time-consistency problem, I identify three credibility challenges that jointly render the revolutionary project unachievable: (i) a horizontal commitment problem among workers, (ii) a vertical commitment problem between revolutionary leadership and followers, and (iii) an inter-class commitment problem between revolutionaries and property owners. Using backward induction, I show that these problems interact to produce a self-defeating structural logic: anticipating that post-revolutionary commitments will fail, negotiations collapse; anticipating violent confrontation, revolutionary movements adopt organizational forms—militarization, hierarchy, ideological rigidity—that make authoritarian outcomes more likely. The Marxist program is inherently fragile because its demands—abolition of private productive property, withering of the state—eliminate the institutional conditions for credible commitment: bright-line rules, ongoing bargaining leverage, and constraints on authority. The same analytical tools that explain why liberal constitutional settlements succeed explain why revolutionary transformation cannot achieve the outcomes it promises. While this paper remains agnostic on Marx's diagnosis of capitalism, it demonstrates the institutional impossibility of his cure.

Keywords: *Karl Marx; communist revolution; credible commitment; time consistency; game theory; self-enforcing constitutions; backward induction*

JEL Codes: D72, D74, D82, P16, P26

1. Introduction

Karl Marx's theoretical system contains a striking asymmetry. His diagnosis of capitalism—the analysis of exploitation, alienation, the tendency of the rate of profit to fall, the concentration and centralization of capital—has generated an enormous and enduring scholarly literature, much of it sympathetic. His proposed cure—revolutionary seizure of productive property, dictatorship of the proletariat, eventual withering of the state into a communist society of free association—has been tested repeatedly and has failed with remarkable consistency. The Soviet Union, Maoist China, Cambodia, Cuba, Ethiopia, and numerous other revolutionary regimes did not produce the liberated society Marx envisioned; they produced authoritarian states that, on Marx's own account of exploitation, merely substituted one ruling class for another (Kolakowski 2005).

The standard explanations for this failure divide into two camps. Marxists and sympathetic critics attribute the outcome to contingent factors: the backwardness of Russia and China (revolution occurred in the wrong countries), imperialist encirclement, the personal pathologies of Stalin or Mao, or deviations from authentic Marxist theory (Anderson 1976; Blackburn 1991). Critics from the liberal and conservative traditions attribute the failure to deeper problems: the impossibility of economic calculation without prices (Mises 1920; Hayek 1945), the inherent tendency of concentrated political power toward tyranny (Popper 1945; Talmon 1952), or the incompatibility of Marxist utopianism with human nature.

This paper proposes a different and more precisely specified explanation. I argue that the failure of Marxist revolution is neither contingent nor attributable to any single factor but is structurally determined by *commitment problems* inherent to the Marxist program itself. Using the analytical tools of the self-enforcing constitutions literature (Weingast 1997; North and Weingast 1989) and the theory of time-consistent policy (Kydland and Prescott 1977), I identify three distinct credibility problems that jointly render the revolutionary project unachievable on its own terms.

The first is a *horizontal commitment problem*: revolutionary mobilization requires coordination among workers, but collective action faces free-riding that can only be overcome through vanguard organization—which immediately creates new agency problems that substitute for rather than resolve the original exploitation. The second is a *vertical commitment problem*: revolutionary leaders cannot credibly commit to followers that post-revolutionary authority will serve proletarian

interests rather than the vanguard's own interests, because the revolutionary program eliminates the institutional mechanisms through which such commitments are typically enforced. The third is an *inter-class commitment problem*: revolutionaries cannot credibly commit to property owners regarding post-transition treatment, making negotiated transformation impossible and violent confrontation structurally inevitable.

These three problems interact through backward induction to produce a self-defeating logic. Anticipating that post-revolutionary governance will produce exploitation rather than liberation (Stage 3), property owners refuse negotiated transition (Stage 2); anticipating violent confrontation, revolutionary movements adopt organizational forms—centralized command, ideological discipline, suppression of internal dissent—that are optimized for military success but guarantee authoritarian governance (Stage 1). The very features that enable revolutionary success ensure revolutionary failure in achieving its stated aims.

The paper's central claim is that Marxist revolution fails because of a *scope problem* that is distinctive to the Marxist program. Credible commitments in political life depend on three institutional conditions: bright-line rules that define transgression (Weingast 1997), ongoing bargaining leverage that gives affected parties power to punish violations (North and Weingast 1989), and institutional constraints on authority that make defection costly (Acemoglu and Robinson 2006). The Marxist program—abolition of private productive property, dismantling of the bourgeois state apparatus, concentration of economic and political power in revolutionary hands pending the state's eventual withering—systematically destroys all three conditions. There are no bright-line rules because “abolition of private property” grants revolutionary authority unbounded discretion over economic life. There is no ongoing bargaining leverage because the working class, having surrendered independent organizational capacity to the vanguard, possesses no mechanism for disciplining revolutionary leaders. There are no institutional constraints because the revolutionary program explicitly demands the dismantling of existing state institutions and their replacement with revolutionary authority that is, by design, unconstrained.

The argument is not that revolution is “hard” or that it faces practical difficulties. It is that the *specific features of the Marxist program* generate commitment problems that are logically irresolvable within the program's own terms. The same analytical tools that explain why the Glorious Revolution's constitutional settlement *could* sustain credible commitments explain why

Marxist revolutionary transformation *cannot*. The difference is not in the quality of actors or the contingencies of history but in the institutional structure of the commitments being attempted.

I remain agnostic on Marx's diagnosis of capitalism. Whether capitalism exploits workers as Marx argued is a separate question that this paper does not address. The paper's contribution is to demonstrate that even if the diagnosis is correct, the proposed cure is institutionally impossible—that the revolutionary program contains within itself the seeds of its own failure, not as a contingent historical pattern but as a structural necessity derivable from the logic of credible commitment.

The argument builds on and extends several intellectual traditions. From constitutional political economy (Buchanan 1975; Buchanan and Tullock 1962), it takes the insight that institutional rules must be analysed for their incentive-compatibility, not merely their normative desirability. From the self-enforcing constitutions literature (Weingast 1997; Mittal and Weingast 2013), it takes the framework of coordination games and focal points that explains when political commitments are credible. From the rationalist theory of war (Fearon 1995), it takes the demonstration that bargaining failures—not irrationality or misperception—explain violent conflict when commitment problems preclude negotiated settlement. From transaction cost economics (Williamson 1985), it takes the concept of relationship-specific investment and the hold-up problem. These are not novel analytical tools; what is novel is their systematic application to the Marxist revolutionary program, which reveals that the program's failure is not contingent but structurally determined.

The paper proceeds as follows. Section 2 develops the three commitment problems verbally, showing how each arises from specific features of the Marxist program and illustrating each with historical examples. Section 3 presents a formal three-stage game and solves by backward induction, establishing the core impossibility result. Section 4 extends the baseline model through a selection analysis of revolutionary leadership and a continuous enforcement probability formulation that strengthens the result. Section 5 engages with the most important objections, including the analytical Marxist response associated with G.A. Cohen and Jon Elster and the political economy framework of Acemoglu and Robinson. Section 6 develops the totality argument explaining why the analysis is specific to the Marxist program. Section 7 concludes.

2. Three Commitment Problems

This section develops the three credibility challenges that jointly render the Marxist revolutionary program unachievable. Each problem arises from specific features of the Marxist theoretical framework and is not merely a practical difficulty but a structural impossibility given the program's own demands.

2.1 The Horizontal Commitment Problem: Coordination and Its Costs

Revolutionary mobilization is a collective action problem of extreme severity. Marx recognized this: the transition from a “class in itself” to a “class for itself”—from workers who share objective interests to workers who act collectively on those interests—was central to his political theory (Marx [1847] 1995; Marx and Engels [1848] 1998). The difficulty is that revolutionary action resembles a stag hunt game with catastrophic failure payoffs: each individual worker benefits from successful collective revolution but faces severe punishment (imprisonment, death, loss of livelihood) if the revolution fails and they have participated. The individually rational strategy is to free-ride on others' revolutionary efforts, which means that absent a coordination mechanism, the equilibrium is universal acquiescence (Olson 1965; Tullock 1971).

Marx's own writings oscillate between two solutions to this problem. The first, developed primarily in the earlier writings, relies on the spontaneous development of class consciousness through the objective experience of exploitation and immiseration: capitalism itself produces the conditions for its overthrow by creating an ever-larger, ever-more-miserable proletariat whose collective interest in revolution becomes self-evident (Marx and Engels [1848] 1998). The second, developed most explicitly by Lenin (1902) but implicit in Marx's own organizational practice, relies on a revolutionary vanguard: a disciplined party of professional revolutionaries who solve the coordination problem by providing organizational infrastructure, ideological leadership, and strategic direction.

The first solution faces the standard objections to spontaneous collective action (Olson 1965): shared interest is insufficient to generate coordination when individual participation is costly and individual contribution is negligible. Even under conditions of severe immiseration, the rational worker free-rides. But the second solution, while potentially solving the coordination problem, immediately generates a new and arguably more severe problem: the agency relationship between

the vanguard and the working class. By creating a centralized organization with the capacity to coordinate collective action, the revolutionary movement creates an agent with enormous power and no institutional mechanism for accountability. The vanguard is simultaneously the solution to the horizontal commitment problem and the source of the vertical commitment problem analyzed below.

This is not a contingent feature of Leninist organizational theory; it follows from the structure of the coordination problem itself. Any organization capable of overcoming the collective action barriers to revolution—capable of imposing discipline, punishing defection, maintaining operational security against state repression—must possess concentrated authority, hierarchical command, and the capacity for coercion. These features are not organizational pathologies; they are functional requirements of the revolutionary task. But they are also precisely the features that make post-revolutionary accountability impossible. The organizational solution to the coordination problem is structurally isomorphic to the organizational form of authoritarian governance.

The historical record illustrates this paradox with uncomfortable consistency. Lenin's (1902) prescription in *What Is to Be Done?* for a disciplined vanguard of professional revolutionaries was not an arbitrary organizational choice but a rational response to the coordination problem under conditions of tsarist repression. The Bolshevik party's centralized structure, strict discipline, and intolerance of factional dissent—the very features that enabled it to seize power in October 1917—were also the organizational template for the post-revolutionary state. Similarly, Mao's organizational innovations during the Long March and the Yan'an period—guerrilla discipline, ideological rectification campaigns, ruthless elimination of internal opposition—solved the coordination problem for revolutionary mobilization but created the institutional DNA of the post-revolutionary regime. The Cultural Revolution was not a departure from revolutionary organizational logic but its extension: the same tools used to discipline the revolutionary movement were turned against the population (Dikotter 2010).

The horizontal commitment problem thus creates a dilemma that has no resolution within the Marxist framework. Without a vanguard, coordination fails and revolution does not occur. With a vanguard, coordination succeeds but the organizational form guarantees authoritarian governance. The vanguard is the solution to the first problem and the source of the second; there is no

organizational form that solves both simultaneously. This is not a claim about the psychology of revolutionary leaders but about the structural requirements of the revolutionary task.

2.2 The Vertical Commitment Problem: Quis Custodiet Ipsos Custodes?

Suppose the coordination problem is solved—whether through vanguard organization, spontaneous mass mobilization during a crisis, or some combination. Revolutionary leadership now holds state power. Can it credibly commit to implementing the communist program rather than exploiting its position?

The answer, I argue, is no—and the reason is not that revolutionary leaders are necessarily venal but that the Marxist program eliminates the institutional conditions under which any political authority can credibly commit to self-restraint. The self-enforcing constitutions literature identifies three conditions for credible political commitment (Weingast 1997; North and Weingast 1989; Mittal and Weingast 2013). First, there must be *bright-line rules* that define transgression—observable criteria that enable citizens to distinguish compliance from violation. Second, there must be *ongoing bargaining leverage*—mechanisms through which citizens can impose costs on authority that violates its commitments. Third, there must be *institutional constraints*—formal and informal rules that limit authority’s discretion and make defection costly.

The Marxist revolutionary program systematically destroys all three conditions. Consider each in turn.

Bright-line rules. Liberal constitutional settlements succeed because they define specific, observable rights: the Crown shall not levy taxes without parliamentary consent; private property shall not be seized without due process; individuals shall not be imprisoned without charge. These are focal points around which citizens can coordinate punishment of transgressions (Weingast 1997). The Marxist program offers no comparable focal points. “Abolition of private productive property” is not a bright-line rule but a mandate for wholesale economic transformation that grants revolutionary authority unbounded discretion over production, distribution, and exchange. “Withering of the state” is not an observable commitment but an eschatological promise with no specified timeline, no measurable criteria, and no mechanism for enforcement. When does the revolutionary authority’s failure to wither constitute a violation of its commitment? The question

has no answer, which means citizens cannot coordinate to punish non-compliance because there is no identifiable moment of transgression.

Ongoing bargaining leverage. Constitutional commitments are sustained because the parties to the bargain retain independent sources of power: Parliament controls taxation, the judiciary controls legal interpretation, citizens can exit or resist. This mutual dependency creates self-enforcing equilibria: each party's compliance is sustained by the other parties' credible threat of retaliation (North and Weingast 1989). The Marxist program concentrates all productive property—and therefore all economic power—in revolutionary hands. Workers who depend on the revolutionary state for employment, housing, food, and physical security have no independent resource base from which to challenge authority. The abolition of private property does not merely redistribute wealth; it eliminates the material foundation for independent political action. As Trotsky (1937) recognized with bitter clarity: in a state where the government is the sole employer, opposition means slow starvation.

Institutional constraints. Credible commitment requires institutions that constrain authority's future choices: independent courts, legislative vetoes, constitutional entrenchment, federal division of powers. The Marxist program explicitly demands the dismantling of such institutions as bourgeois instruments of class domination (Marx [1871] 1970; Lenin [1917] 1992). The revolutionary state is to be governed by the proletariat through direct democratic mechanisms—but as the horizontal commitment analysis shows, the “proletariat” in practice means the vanguard, and the vanguard faces no institutional constraints because it has dismantled them by revolutionary design. The dictatorship of the proletariat is institutionally unconstrained not by accident but by programmatic intention.

The vertical commitment problem is therefore not a contingent failure of particular revolutionary leaderships but a structural feature of the Marxist program. Any revolutionary authority that faithfully implements the program's demands—abolishing private property, dismantling the bourgeois state, concentrating economic and political power—simultaneously destroys the institutional conditions under which its own commitments to the working class could be credible. This is the deepest paradox of the revolutionary project: the more thoroughly the program is implemented, the less credible any remaining commitments become.

The historical evidence is stark. The Soviet Union under Stalin, China under Mao, Cambodia under Pol Pot, Cuba under Castro, and Ethiopia under the Derg all followed structurally similar trajectories despite enormous differences in culture, economic development, and geopolitical context. In each case, revolutionary leadership consolidated power, eliminated independent centres of authority, collectivized productive property, and then governed without effective institutional constraint. The variation among these cases is real—the degree of repression, the pace of economic transformation, the severity of humanitarian consequences—but the structural pattern is uniform: revolutionary authority, unconstrained by the institutional mechanisms the program itself demanded be dismantled, exploited its position. This uniformity across wildly different contexts is precisely what a structural explanation predicts and what a contingency explanation cannot account for.

Trotsky's (1937) critique of Stalin from within the Marxist tradition is particularly illuminating. Trotsky diagnosed the problem correctly—the revolutionary state had become a new instrument of exploitation—but could not identify a solution because the solution required the very institutional constraints (independent judiciary, private economic power, competitive political parties) that Marxist theory demanded be abolished. Trotsky's proposed remedy—a “political revolution” to restore workers' democracy within the framework of state ownership—itself faced the identical commitment problem: who would constrain the leaders of the political revolution? The regress is infinite because the program eliminates the conditions for its own termination.

2.3 The Inter-Class Commitment Problem: The Impossibility of Negotiated Revolution

The third commitment problem concerns the relationship between revolutionaries and the class they seek to dispossess. Even if revolutionary leaders sincerely intend to implement a humane transition—offering amnesty, compensation, or guaranteed civil rights to former property owners—they cannot make such promises credible.

The reason is straightforward. Any promise made by revolutionary leadership regarding post-transition treatment of property owners is subject to the same time-consistency problem that afflicts Stage 3 governance: once the transition is complete and revolutionary authority is consolidated, the enforcement mechanism for pre-transition promises has been dismantled. Property owners, reasoning backward, recognize that promises of amnesty or compensation are

not credible because the revolutionary authority will have both the incentive and the capacity to renege. The logic is identical to Fearon's (1995) rationalist theory of war: even when negotiated settlements exist that both parties would prefer to violent confrontation, the inability to commit to honoring the settlement makes war rational. The structure here echoes Fearon's (1995) rationalist theory of war precisely: even when peaceful transition would be Pareto-superior to violent confrontation, the inability to commit renders negotiation impossible.

This problem is distinctive to the Marxist case because of the *scope* of revolutionary demands. In ordinary political bargaining—including many revolutionary situations—the parties can reach agreements because each retains some bargaining leverage post-settlement. The Glorious Revolution of 1688 succeeded as a negotiated settlement because it reallocated authority within the propertied classes: the Crown surrendered fiscal prerogative but retained executive functions; Parliament gained control over taxation but depended on the Crown for governance. This asymmetry distinguishes the Marxist case from constitutional settlements: in 1688, the relevant parties shared an interest in property protection and the bargain reallocated authority rather than abolishing the counterparty's fundamental interest (North and Weingast 1989). Each party retained ongoing leverage, which made commitments self-enforcing. In the Marxist case, the demand is not reallocation but *abolition*: the complete expropriation of productive property. There is no intermediate position that satisfies revolutionary demands while preserving property owners' bargaining leverage. The bargaining range is empty.

The consequence is that violent confrontation is not a regrettable but avoidable feature of Marxist revolution; it is a structural necessity. Since negotiated transition is impossible (no credible commitments can be offered), and since property owners rationally prefer violent resistance to uncompensated expropriation (even at high cost), revolution must proceed through force. But violent revolutionary success requires exactly the organizational features—military discipline, centralized command, suppression of dissent—that guarantee authoritarian post-revolutionary governance. The inter-class commitment problem thus feeds back into the horizontal and vertical problems, creating a self-reinforcing cycle.

This feedback mechanism deserves emphasis because it is the source of the model's distinctive predictive power. The three commitment problems are not independent failures that might be addressed separately; they interact to produce an equilibrium that is globally stable and resistant

to reform. The horizontal problem (vanguard organization is necessary for coordination) creates the vertical problem (the vanguard cannot be constrained post-revolution). The vertical problem (revolutionary leadership will exploit) creates the inter-class problem (property owners cannot accept negotiated transition). The inter-class problem (violent confrontation is inevitable) reinforces the horizontal problem (military organization requires even more centralized, hierarchical command). Each problem exacerbates the others, producing a structural trap from which there is no exit within the program's own terms.

The contrast with the Glorious Revolution is instructive. In 1688, William of Orange and the parliamentary coalition could negotiate because the demands were limited (reallocation of fiscal authority, not abolition of monarchy) and both parties retained ongoing leverage (the Crown retained executive power; Parliament retained fiscal control). The bargaining range was non-empty because neither party demanded the other's total dispossession. North and Weingast (1989) show that the resulting institutional settlement sustained credible commitments precisely because both parties remained powerful enough to punish violations. In the Marxist case, the demand for complete expropriation of productive property eliminates the bargaining range and destroys the conditions for any settlement that both parties would prefer to conflict.

3. A Formal Model of Revolutionary Credibility Failure

This section formalizes the three commitment problems as a sequential game and solves by backward induction, establishing that no equilibrium path leads to the communist outcome.

3.1 Players and Setup

Consider a polity with three actors: a continuum of Workers (W), indexed by $i \in [0,1]$; a Capitalist class (C), treated as a unitary actor for simplicity; and a Revolutionary Leadership (R), the vanguard that coordinates revolutionary action and governs post-revolution. The game proceeds in three stages: Mobilization, Transition, and Governance. I solve by backward induction, beginning at the terminal stage.

3.2 Stage 3: Post-Revolutionary Governance

Suppose revolution has succeeded and R holds state power. R chooses policy $g \in \{\text{Implement, Exploit}\}$. Under Implement, R pursues the genuine communist program; workers receive payoff $u_W^I > 0$. Under Exploit, R extracts rents from its position; workers receive $u_W^E < u_W^I$, while R receives $B > 0$.

Workers can attempt to punish exploitation through collective resistance. Let $\sigma \in [0,1]$ denote the fraction of workers who resist, at individual cost $c > 0$. Resistance succeeds (R is deposed) if $\sigma \geq \bar{\sigma}(\kappa)$, where $\bar{\sigma}$ is the coordination threshold and κ measures the coercive capacity of the ruling authority. The critical structural feature is that $\bar{\sigma}(\kappa)$ is increasing in κ : the greater the authority's coercive capacity, the higher the threshold for successful collective resistance.

Under revolutionary governance, κ takes the value κ_R , which exceeds the pre-revolutionary value κ_C for three reasons. First, the vanguard that solved the pre-revolutionary coordination problem now actively suppresses coordination against itself. Second, the abolition of private property eliminates independent economic bases from which opposition could be organized and financed. Third, the dismantling of civil society organizations, opposition parties, and independent media—undertaken as part of the revolutionary program—removes the organizational infrastructure for collective resistance. Therefore $\bar{\sigma}(\kappa_R) > \bar{\sigma}(\kappa_C)$.

In the continuum model with atomistic workers, each individual's contribution to resistance is negligible. The unique Nash equilibrium is $\sigma = 0$: no worker resists. Even relaxing the continuum assumption, the coordination threshold $\bar{\sigma}(\kappa_R)$ is prohibitively high because the revolutionary state has systematically destroyed the coordination mechanisms (independent organizations, private economic power, free communication) that would be needed to reach it.

Proposition 1 (Vertical Commitment Failure). In any subgame perfect equilibrium of Stage 3, revolutionary leadership chooses Exploit. Workers, unable to coordinate resistance, do not resist. The communist program is not implemented.

Note that this result does not depend on R 's type. Even a leadership composed partially of sincere idealists faces $B > 0$ and $p_E \approx 0$; exploitation is the dominant strategy. (An independent argument reinforces this conclusion: even if commitment problems were resolved, implementation of collective ownership would face the economic calculation problem identified by Mises (1920) and Hayek (1945). Abolishing private productive property eliminates the price system, which is the

only mechanism capable of aggregating the dispersed knowledge required for rational economic coordination (Lavoie 1985). The commitment and knowledge problems are logically independent; either alone renders the Marxist program unachievable.)

3.3 Stage 2: Revolutionary Transition

Step back one stage. Revolution has succeeded militarily; R and C bargain over the transition. Let $K > 0$ denote the value of capitalist property. R makes an offer to C : a transfer $\tau \in [0, K]$ that C retains in exchange for peaceful surrender of the remaining property ($K - \tau$). If C accepts, peaceful transition occurs. If C rejects, armed conflict ensues: R wins with probability $p \in (0,1)$, both sides pay conflict cost $\delta > 0$.

The commitment problem is immediate. If C accepts R 's offer, the game proceeds to Stage 3. By Proposition 1, R will exploit its position—which includes expropriating any property C was promised under the transition agreement. R 's promise of τ is not credible because no enforcement mechanism exists in Stage 3.

C 's reasoning is therefore: acceptance yields payoff 0 (since τ will be expropriated), or worse if R pursues additional punishment of class enemies. Rejection yields $(1 - p) \cdot K - \delta$ from armed resistance. For any $p < 1$ and sufficiently large K , C strictly prefers violent resistance.

Proposition 2 (Inter-Class Commitment Failure). No peaceful transition is achievable in equilibrium. For any offer τ , C rejects and violent confrontation occurs.

This result holds even when Pareto-improving negotiated settlements exist. If R could credibly commit to honoring τ , both parties would prefer peaceful transition (avoiding conflict cost δ). But the absence of post-revolutionary enforcement means that τ cannot be made credible, and the surplus from peaceful transition is destroyed by the commitment problem. This is a pure commitment-failure inefficiency: both parties are made worse off by R 's inability to bind itself.

The inefficiency is worth emphasizing because it reveals the distinctive character of the inter-class commitment failure. In Coasean bargaining, parties can negotiate efficient outcomes regardless of initial endowments provided transaction costs are low and property rights are well defined. The Marxist revolutionary situation violates both conditions: transaction costs are enormous (the stakes are total, the information asymmetries severe, and the time horizon compressed into a

revolutionary moment) and the entire dispute concerns the definition and allocation of property rights. The Coase theorem, which guarantees efficient bargaining under its assumptions, is precisely inapplicable to the revolutionary situation because the conditions the theorem requires are the conditions the revolution destroys.

Furthermore, the inter-class commitment problem intensifies as the revolutionary movement gains strength. A weak revolutionary threat can be addressed through limited concessions (higher wages, improved working conditions, political representation)—concessions that remain within the framework of private property and are therefore credible. A revolutionary movement that demands the *abolition* of private property cannot be satisfied by concessions within the system it seeks to destroy. The stronger the Marxist revolutionary threat, the less room there is for negotiated settlement, and the more certain violent confrontation becomes. This is the opposite of the democratic transition dynamic analyzed by Acemoglu and Robinson (2006), where a sufficiently strong revolutionary threat *enables* negotiated transition to democracy.

3.4 Stage 1: Revolutionary Mobilization

Now consider the initial coordination problem. Each worker i chooses $a_i \in \{\text{Revolt}, \text{Acquiesce}\}$. Let m denote the fraction who revolt. Revolution succeeds if $m \geq \bar{m} \in (0,1)$.

If revolution fails ($m < \bar{m}$): revolting workers receive $-\lambda$ (punishment: imprisonment, death, loss of livelihood); acquiescing workers receive w (status quo wage). If revolution succeeds: all workers proceed through Stages 2 and 3. By Propositions 1 and 2, the outcome is violent confrontation followed by vanguard exploitation. Workers' expected payoff from successful revolution is therefore: $p \cdot u_w^E + (1 - p) \cdot w - d$, where $d > 0$ is the cost of violent confrontation borne by workers.

A worker revolts if and only if the expected payoff from revolt, conditional on beliefs about others' actions, exceeds w . The crucial observation is that the expected payoff from successful revolution is *not* the communist utopia u_w^I but the exploitation payoff u_w^E discounted by the probability of revolutionary military victory and the costs of violent confrontation. Revolution is rational for workers only if:

$$p \cdot u_w^E + (1 - p) \cdot w - d > w$$

which simplifies to $p \cdot (u_w^E - w) > d$. Revolution is individually rational only when the probability-weighted improvement from exploitation under the vanguard over wages under capitalism exceeds the costs of violent confrontation. If $u_w^E \leq w$ (vanguard exploitation is no better than capitalism), revolution is never rational.

Proposition 3 (Mobilization Failure). If $u_w^E \leq w$, no worker revolts in equilibrium. If $u_w^E > w$, revolution may occur but does not produce communism—it produces a change of masters.

3.5 The Backward Induction Result

Solving the complete game by backward induction yields the paper's central result.

Theorem 1 (Revolutionary Impossibility). In any subgame perfect equilibrium of the three-stage revolutionary game, no equilibrium path leads to the communist outcome. Specifically: (i) revolutionary leadership exploits its position in Stage 3 because no enforcement mechanism constrains it; (ii) peaceful transition is impossible in Stage 2 because revolutionary commitments are not credible; (iii) revolution in Stage 1, if it occurs, produces at best a substitution of exploiters, not liberation.

The theorem establishes that the Marxist program fails *on its own terms*. Even if capitalism exploits workers as Marx claimed, even if immiseration generates revolutionary consciousness, even if a revolutionary movement successfully seizes state power—the program's own structural features ensure that the promised outcome (communist society) cannot be achieved. The program is self-defeating: the institutional conditions it demands (abolition of private property, concentration of power, dismantling of constraints) are precisely the conditions that make its commitments incredible.

It is worth noting what the theorem does and does not establish. It establishes that no equilibrium path leads to the *communist* outcome—the society of free association, collective ownership, and eventual statelessness that Marx envisioned. It does not establish that revolution is necessarily irrational for workers. If $u_w^E > w$ —if exploitation under the vanguard is preferable to wages under capitalism—then revolution may improve workers' material conditions. But this would represent the failure, not the success, of the Marxist program: workers would remain exploited, merely by different masters. The revolution would have succeeded as regime change and failed as liberation.

Marx's own theory predicts and condemns precisely this outcome: the substitution of one form of exploitation for another is not the communist emancipation Marx promised but a new form of class society, which Marxist theory would classify as exploitative (Roemer 1982). The program thus fails by its own evaluative criteria.

3.6 The Role of Expectations and Self-Fulfilling Dynamics

The backward induction argument reveals a feature of the revolutionary problem that deserves independent emphasis: the *self-fulfilling* character of commitment failure. Even a revolutionary movement that begins with sincere intentions faces an expectations trap. Property owners, reasoning backward from Stage 3, refuse negotiation in Stage 2. This refusal forces violent confrontation, which in turn selects for militarized organizational forms in Stage 1. The militarized organization then governs post-revolution, and its hierarchical, coercive character makes exploitation the dominant strategy in Stage 3—confirming the expectations that drove the original refusal.

The expectations are rational and self-confirming. This distinguishes the revolutionary commitment failure from cases where pessimistic expectations are mistaken. In the Marxist case, the expectations are *correct*: the structural features of the program really do prevent credible commitment, militarized organization really does produce authoritarian governance, and post-revolutionary exploitation really does follow from the absence of institutional constraints. The tragedy is not that the parties are trapped by unfounded pessimism but that the pessimism is warranted by the program's own institutional logic.

This self-fulfilling dynamic also explains why attempts to reform the revolutionary process from within have consistently failed. Gorbachev's attempt to introduce *glasnost* and *perestroika* within the Soviet system illustrates the problem: creating space for political contestation and economic decentralization undermined the institutional basis of the revolutionary state without creating alternative mechanisms for credible commitment. The result was not reformed communism but system collapse—because the revolutionary institutional framework could not accommodate partial reform. The all-or-nothing character of the Marxist program, which is the source of its commitment problems, is also the reason it cannot be reformed incrementally.

The self-fulfilling logic also illuminates the peculiar resilience of pre-revolutionary acquiescence—the Hobbesian trap in the revolutionary context. Workers who correctly anticipate the backward induction result—that successful revolution leads to vanguard exploitation rather than liberation—may rationally prefer the status quo even under severe capitalist exploitation, if the expected outcome of revolution is no better or is worse. This creates an equilibrium of frustrated acquiescence that is stable even when workers share a collective interest in transformation, because each worker recognizes that the institutional vehicle for transformation (revolutionary seizure) cannot deliver the promised outcome. The equilibrium is not one of false consciousness—workers may be perfectly clear about their exploitation under capitalism—but of rational pessimism about the available alternative. Marx’s theory attributes the absence of revolution to ideological mystification; the commitment analysis suggests that rational assessment of revolutionary prospects may be the more parsimonious explanation.

4. Extensions: Selection Effects and Probabilistic Enforcement

4.1 Revolutionary Selection and Leadership Type

The baseline model treats revolutionary leadership as a unitary actor choosing between implementation and exploitation. In reality, leadership emerges through a competitive process with selection pressures that systematically favor exploitative types.

Model potential revolutionary leaders as varying in type $\theta \in \{\text{Idealist}, \text{Opportunist}\}$. Idealists receive high utility from implementing communism; Opportunists receive high utility from exploitation. Let π denote the prior fraction of Idealists among potential leaders. Revolutionary success requires capabilities that I denote C : willingness to use violence, organizational ruthlessness, strategic deception capacity, and tolerance for casualties. Revolutionary competition is a tournament: leaders with higher C prevail in factional struggles, armed conflict, and organizational competition.

The critical structural feature is that the capabilities required for revolutionary success are *negatively correlated* with Idealist type. Idealists face moral constraints on violence and deception; commitment to communist principles creates strategic rigidity (refusing tactically advantageous

compromises); unwillingness to sacrifice followers limits tactical options; honesty is a competitive disadvantage in factional politics. Formally, $\text{Corr}(C, \theta = \text{Idealist}) < 0$.

Proposition 4 (Revolutionary Selection). Revolutionary competition selects for leadership types whose characteristics (willingness to use violence, organizational ruthlessness, strategic flexibility) are negatively correlated with post-revolutionary commitment to the communist program. The probability that revolutionary leadership will exploit exceeds the base rate of opportunism in the movement: $P(\theta^* = \text{Opportunist}) > 1 - \pi$.

This result is not merely theoretical. Lenin outmaneuvered the more idealistic Mensheviks through superior organizational ruthlessness. Stalin defeated Trotsky, Bukharin, and Kamenev through factional cunning rather than ideological purity. Mao's rise involved the systematic elimination of rivals; the Cultural Revolution subsequently targeted remaining idealists. The pattern documented by Brinton (1965) and others—that “moderate” revolutionary phases consistently give way to more ruthless factions (Girondins to Jacobins, Kerensky to Bolsheviks)—is precisely what Proposition 4 predicts.

The selection mechanism operates through multiple channels. Within revolutionary movements, factional competition rewards leaders who can mobilize followers, eliminate internal rivals, and maintain organizational cohesion under extreme pressure. These tasks require precisely the qualities—strategic ruthlessness, willingness to deceive, tolerance for violence against former comrades—that predict post-revolutionary exploitation. Between revolutionary movements competing for the same constituency, the more militant and disciplined organization tends to prevail: the Bolsheviks over the Mensheviks and Socialist Revolutionaries, the CCP over the various warlord and democratic factions, the Khmer Rouge over less extreme Cambodian left movements. And within the armed struggle against the existing state, military effectiveness selects for leaders whose command structure prefigures authoritarian governance.

The selection effect interacts powerfully with the vertical commitment problem. Even if the initial pool of potential leaders contains a substantial fraction of sincere idealists (π is large), the competitive dynamics of revolution systematically filter them out. The leader who emerges victorious is, with high probability, the type least likely to implement the communist program. Stalin was not an accident; he was the equilibrium outcome of a selection process that rewarded

organizational ruthlessness over ideological commitment. The structural prediction is clear: revolutionary competition produces Stalins, not Luxemburgs.

4.2 Continuous Enforcement and Comparative Institutional Analysis

A potential objection to the baseline model is that it treats enforcement as binary: either post-revolutionary authority faces constraints or it does not. In reality, post-revolutionary regimes face *some* constraints—factional competition within the vanguard, military loyalty limits, international pressure, threat of popular unrest. A more defensible formulation replaces the binary structure with a continuous enforcement probability.

Let $p_E \in [0,1]$ denote the probability that post-revolutionary authority faces effective constraints on exploitation. R 's expected payoff from exploitation becomes $(1 - p_E) \cdot B$; the payoff from implementation remains $\varepsilon > 0$. R exploits whenever $(1 - p_E) \cdot B > \varepsilon$, which obtains for any $p_E < 1 - \varepsilon/B$. Given $B \gg \varepsilon$, exploitation is the dominant strategy for any p_E substantially below 1.

The comparative institutional claim is then not that $p_E^R = 0$ but that $p_E^R \ll p_E^C$, where the superscripts denote revolutionary and constitutional governance respectively. Under constitutional settlements, multiple independent power centres with separate resource bases, explicit procedural rules enabling coordinated punishment, repeated interactions among parties with ongoing stakes, and clear focal points for identifying transgressions jointly produce a high p_E^C . Under revolutionary governance, a single power centre controlling the coercive apparatus, no explicit and enforceable constraints, one-shot transformation eliminating repeated-game incentives, and no bright-line rules defining transgression jointly produce a low p_E^R .

Theorem 2 (Comparative Institutional Result). Revolutionary transformation produces exploitation with probability approaching $1 - p_E^R$. Since $p_E^R \ll p_E^C$, Marxist revolution produces authoritarian outcomes with systematically higher probability than constitutional settlements—not with logical certainty, but as a strong structural tendency.

This probabilistic formulation is both more defensible and more analytically powerful than the binary version. It does not invite counterexample-hunting (“but what about Cuba’s healthcare?”) because it does not claim logical impossibility. It claims systematic structural bias—which is

exactly what the historical record of nearly universal revolutionary failure, with rare partial exceptions, displays.

5. Objections and Responses

5.1 *The Analytical Marxist Response: Cohen and Elster*

The most sophisticated philosophical defence of Marxism comes from the analytical Marxist tradition, particularly G.A. Cohen and Jon Elster. Cohen (2000) argues that the normative core of Marxism—the commitment to equality, community, and human flourishing—is independent of, and survives, the failure of Marx’s specific institutional proposals. Elster (1985) provides the most rigorous rational-choice reconstruction of Marx’s theoretical system, identifying both its genuine insights and its logical failures. Engaging with both is essential.

Cohen (2000) might respond that the commitment problems identified here are features of *Leninist* revolutionary strategy, not of Marx’s theory itself. Marx’s vision of revolution, Cohen could argue, was democratic and spontaneous—the self-emancipation of the working class—not the vanguard-led seizure of power that produced Bolshevism. The commitment problems I identify arise from the organizational form (vanguardism), not from the program (communism).

This objection fails for two reasons. First, as the horizontal commitment analysis demonstrates, spontaneous democratic revolution faces even more severe coordination problems than vanguard-led revolution. Without organizational infrastructure, the free-rider problem is insurmountable. If Cohen’s response is that material conditions eventually make coordination spontaneous, he must explain the mechanism by which shared immiseration translates into coordinated action without organizational mediation—and the entire collective action literature since Olson (1965) demonstrates that shared interest alone is insufficient. The spontaneous coordination story also faces an empirical challenge: every successful Marxist revolution—Russia, China, Cuba, Vietnam, Cambodia—has been organized through a vanguard party, not through spontaneous mass action. The only large-scale spontaneous revolutionary mobilizations that have occurred (the Paris Commune, the German Revolution of 1918–19, the Hungarian Revolution of 1956) were either crushed quickly or failed to establish durable revolutionary governance. The empirical record provides no support for the spontaneous coordination mechanism.

Second, and more fundamentally, even if spontaneous revolution were possible, the vertical commitment problem remains in full force. Post-revolutionary governance requires *some* institutional form. Whether authority is exercised by a vanguard party, a workers' council, a revolutionary committee, or a direct-democratic assembly, the problem is identical: the body that holds power after revolution cannot credibly commit to the communist program because the program itself eliminates the conditions for credible commitment. Workers' councils face the same collective action problems as any democratic body—capture by organized minorities, free-riding on monitoring, inability to coordinate punishment of leadership—compounded by the absence of institutional constraints, independent media, and alternative power centres. The problem is in the *program*, not the *party*. Cohen's attempt to separate Marxist normative commitments from Leninist organizational strategy fails because the commitment problems identified here are generated by the normative commitments themselves (abolition of private property, withering of the state), not by the organizational form chosen to pursue them.

Elster (1985) provides a different line of challenge. His rational-choice analysis of Marx identifies “functional explanation” as a pervasive methodological error in Marx's work: Marx explains institutions by their consequences without specifying the mechanisms by which those consequences are produced. Elster's critique is compatible with and complementary to mine. Where Elster identifies methodological problems in Marx's explanatory framework, I identify structural problems in Marx's prescriptive program. Elster shows that Marx's *theory* contains logical gaps; I show that Marx's *program* contains institutional impossibilities. The two critiques operate at different levels but converge: the Marxist system fails both as explanation (Elster) and as prescription (this paper).

Roemer (1982) offers a further analytical Marxist refinement, reconstructing exploitation theory using general equilibrium models and arguing that exploitation can be defined without reference to the labour theory of value. Roemer's contribution is important for the diagnosis of capitalism but is irrelevant to the prescriptive question this paper addresses. Even if Roemer's exploitation concept is correct—even if capitalism is exploitative in a formally rigorous sense—the commitment problems of revolutionary transformation remain. The diagnosis may be valid; the cure remains impossible.

5.2 Acemoglu and Robinson: Revolution, Commitment, and Democratic Transition

Acemoglu and Robinson (2006) provide the most important contemporary formal analysis of revolution, democratization, and institutional change. Their framework models revolution as a credible threat that forces elite concessions: when the revolutionary threat is sufficiently severe, elites extend the franchise to prevent revolution, because democracy commits future governments to redistribution in a way that temporary concessions cannot. Their analysis addresses commitment problems extensively and at several points overlaps with the argument developed here.

My paper extends and sharpens their framework in a specific direction. Acemoglu and Robinson's model explains when elites concede democratic transitions: when the cost of revolution exceeds the cost of redistribution under democracy. Critically, their model works because the demands being negotiated are *incremental*: the franchise, progressive taxation, social insurance. These demands are compatible with the continued existence of private property and therefore with the preservation of ongoing bargaining leverage. Elites under democracy retain economic power, organizational capacity, and exit options; democratic commitments to redistribution are credible precisely because they are limited.

The Marxist case is structurally different. The demand is not redistribution within capitalism but abolition of capitalism. There is no intermediate concession that satisfies revolutionary demands while preserving the institutional conditions for credible commitment. In Acemoglu and Robinson's framework, revolution is a threat that enables bargaining; in the Marxist case, revolution is an end that precludes bargaining. Their model explains why democratic transitions can be negotiated; mine explains why communist transitions cannot. The two analyses are complementary: Acemoglu and Robinson identify the conditions under which political transformation succeeds; I identify why the Marxist form of transformation is uniquely resistant to those conditions.

This distinction also illuminates why the historical record is asymmetric. Democratic transitions through negotiation have occurred successfully in many countries—Spain, Portugal, South Korea, Chile, much of Eastern Europe after 1989. Communist transitions through negotiation have never occurred. The asymmetry is not accidental; it reflects the difference between demands that preserve the conditions for credible commitment (democratic transition) and demands that destroy them (communist revolution).

Acemoglu and Robinson's framework also helps clarify a point sometimes raised in defence of the Marxist program: that the revolutionary threat is valuable even if revolution itself is unachievable, because it forces elites to make concessions. This is correct within Acemoglu and Robinson's model, where the threat of revolution motivates democratic transition and redistribution. But note what the concessions involve: expanded franchise, progressive taxation, social insurance—all of which are compatible with preserved property rights and ongoing bargaining leverage. The revolutionary threat is productive *precisely to the extent that it generates outcomes short of revolution*. The full Marxist program—actual revolutionary seizure and abolition of private property—remains unachievable for the reasons this paper identifies. The practical implication, though I note it only in passing since this paper remains agnostic on the diagnosis, is that Marxist theory may be most useful when it is not implemented: as a threat that generates reformist concessions rather than as a program that can achieve its own objectives.

5.3 “You Assume What You Prove”

A structural objection holds that the model's conclusions are built into its assumptions: by assuming no post-revolutionary enforcement mechanism, the paper simply asserts the conclusion it claims to derive. This objection conflates assumption and derivation and misidentifies the source of the model's assumptions.

The model's key assumptions are not arbitrary stipulations but descriptions of the Marxist program itself, taken from Marx's own writings and from the programmatic demands of actually existing Marxist movements. The assumption that post-revolutionary authority faces no institutional constraints follows directly from Marx's demand to dismantle the bourgeois state apparatus (Marx [1871] 1970; Lenin [1917] 1992). The assumption that workers lack independent bargaining leverage follows from the abolition of private productive property, which is the program's central demand (Marx and Engels [1848] 1998). The assumption that no bright-line rules define transgression follows from the program's demand for wholesale economic transformation rather than incremental reform. These are not external assumptions imposed on the model; they are the Marxist program's own demands, taken seriously and traced to their logical implications.

The paper does not assume that commitment fails; it shows *why* it fails, given the specific institutional features that the Marxist program demands. The argument is: if you implement the

Marxist program (premise), then institutional constraints, bargaining leverage, and bright-line rules are destroyed (derivation from the program's content), then credible commitment is impossible (derivation from institutional theory), then authoritarian exploitation is the equilibrium outcome (derivation from game theory). Each step follows from the previous; the conclusion is derived, not assumed. The only "assumption" is that the Marxist program means what it says—an assumption that Marxists themselves would endorse.

5.4 Material Conditions and Historical Contingency

Marxists frequently argue that actually existing socialism failed because revolutions occurred in the "wrong" countries—economically backward Russia and China rather than the industrialized West. If revolution had occurred where Marx predicted (in advanced capitalist societies with a mature proletariat), the outcome would have been different.

This objection does not survive the commitment analysis. The horizontal, vertical, and inter-class commitment problems are structural features of the Marxist program, not contingent features of Russian or Chinese conditions. A revolution in advanced Germany or Britain in 1920 would have faced identical problems: the vanguard would still have lacked institutional constraints; the working class would still have surrendered independent bargaining leverage upon the abolition of private property; negotiated transition would still have been impossible because revolutionary promises could not be made credible. Material conditions affect *whether* revolution occurs; they do not affect *whether post-revolutionary commitments can be sustained*. The commitment problem is structural, not contextual.

Indeed, the commitment problem may be *worse* in advanced economies, because the value of assets subject to revolutionary expropriation (K in the model) is larger, intensifying the inter-class commitment problem and making violent resistance by property owners more determined. Furthermore, the density of institutional relationships in advanced economies—complex financial systems, international trade dependencies, global capital mobility—means that revolutionary disruption would be more costly and the bargaining problem more severe. The very features that Marx expected would facilitate revolution (advanced productive forces, concentrated capital, a mature proletariat) actually intensify the commitment problems that prevent its success.

6. The Totality Argument: Why Marx Is a Special Case

The preceding analysis might seem to demonstrate merely that revolution is difficult—a conclusion neither novel nor distinctive to Marxism. This section argues that the commitment problems identified here are not generic difficulties of political transformation but specific consequences of the *scope and character* of Marxist demands. The argument is that the Marxist program is uniquely vulnerable to commitment failure because of three features that are individually unusual and jointly distinctive.

First, the *scope of demanded change*. Political transformations succeed when they meet three institutional conditions: defined boundaries on change, preserved bargaining leverage for affected parties, and mechanisms for monitoring and enforcing compliance. The Glorious Revolution, the American constitutional founding, and post-war democratic transitions all met these conditions. Each involved the reallocation of authority among parties who retained independent power, within frameworks that defined specific rights and obligations and created mechanisms for mutual monitoring. The Marxist demand for abolition of private productive property has no defined boundary: it encompasses all means of production, distribution, and exchange, and ultimately the entire structure of economic life. Where a constitutional settlement says “the Crown shall not tax without Parliament’s consent,” the Marxist program says “all productive property shall be collectively owned.” The first is a bright-line rule; the second is an unbounded mandate.

Second, the *elimination of counterparty leverage*. Successful political settlements preserve the capacity of all parties to punish violations. Parliamentary sovereignty works because the Crown depends on Parliament for revenue; democratic redistribution works because the wealthy retain economic power that constrains the state (Acemoglu and Robinson 2006). The Marxist program demands the elimination of all independent economic power: once productive property is collectivized, no group retains the material capacity to discipline the state. The program does not merely weaken counterparty leverage; it eliminates it by design.

Third, the *destruction of institutional monitoring capacity*. Weingast (1997) shows that self-enforcing constitutions require citizens to detect violations and coordinate punishment. Detection requires observable criteria; coordination requires organizational capacity. The Marxist program destroys both: “abolition of private property” provides no observable criterion for detecting when

revolutionary authority has exceeded its mandate (since the mandate is unbounded), and the elimination of independent organizations (opposition parties, independent media, autonomous trade unions, private firms) destroys the organizational infrastructure for coordinated punishment.

Each feature alone would create severe commitment problems; together they produce a situation in which credible commitment is, in the precise sense established by the formal model, impossible. This is why the analysis is specific to the Marxist program rather than a generic critique of political transformation. Revolutions that pursue limited objectives—the franchise, religious toleration, national independence, even substantial redistribution—can succeed because their demands are compatible with the institutional conditions for credible commitment. The Marxist program is uniquely self-defeating because its demands *constitute* the destruction of those conditions. The cure is impossible because implementing it requires eliminating the institutional conditions under which any political cure could be sustained.

A useful way to see this point is through comparison with the Lockean social contract tradition. Locke’s constitutional framework succeeds as a commitment device because it *limits* what government may do: specified rights, defined procedures, enumerated powers. These limits are the bright-line rules around which citizens coordinate. Marx’s revolutionary program fails as a commitment device because it *removes all limits*: the revolutionary state is to control all productive property, restructure all economic relations, and reshape social life in its entirety. There is no bright line because the mandate is total. The Lockean state is credibly constrained because its power is bounded; the Marxist state is incredible because its power is, by programmatic design, unbounded.

7. Conclusion

This paper has argued that the consistent failure of Marxist revolutions to achieve their stated aims is neither contingent nor accidental but structurally determined by commitment problems inherent to the Marxist program. Three distinct credibility problems—horizontal (coordination among workers generates agency costs), vertical (revolutionary authority cannot credibly commit to implementation over exploitation), and inter-class (revolutionaries cannot credibly commit to property owners, precluding negotiated transition)—interact through backward induction to produce a self-defeating logic in which the organizational requirements for revolutionary success guarantee the failure of revolutionary aims.

The paper's central insight is that the Marxist program is uniquely self-defeating because its demands—abolition of private productive property, dismantling of existing state institutions, concentration of economic and political power in revolutionary hands—systematically destroy the three institutional conditions on which credible political commitment depends: bright-line rules, ongoing bargaining leverage, and institutional constraints on authority. The same analytical tools from the self-enforcing constitutions literature that explain why liberal constitutional settlements *can* sustain credible commitments explain why Marxist revolutionary transformation *cannot*.

The analysis has proceeded at the level of institutional structure rather than historical narrative. I have not argued that revolutionary leaders are uniquely venal, that workers are irrational, or that capitalism is just. I have remained agnostic on Marx's diagnosis. The claim is narrower and stronger: even granting the diagnosis, the prescribed cure is institutionally impossible. The commitment problems identified here are features of the program, not of the people who attempted to implement it, and they would obtain regardless of the material conditions, cultural context, or personal integrity of the actors involved.

The selection extension (Proposition 4) strengthens this conclusion by showing that revolutionary competition systematically selects for the leadership types least likely to implement the communist program, and the probabilistic formulation (Theorem 2) demonstrates that the result holds as a strong structural tendency rather than a binary impossibility—matching the empirical pattern of nearly universal revolutionary failure with rare and partial exceptions.

The paper's conclusions are compatible with, and complementary to, the analytical Marxist tradition's own internal critique. Elster (1985) showed that Marx's explanatory theory contains logical gaps; Cohen (2000) argued that Marx's normative commitments must be separated from his institutional proposals; Roemer (1982) demonstrated that exploitation theory can be reconstructed without the labour theory of value. This paper adds a further dimension: even if the normative commitments are valid and the diagnosis is sound, the institutional vehicle proposed to realize them—revolutionary seizure and transformation of productive property—cannot work, for reasons derivable from the logic of credible commitment.

Acemoglu and Robinson (2006) have shown that democratic transitions can be negotiated because the demands at stake—the franchise, redistribution—are compatible with preserving the

institutional conditions for credible commitment. The Marxist case fails because its demands are not. This asymmetry explains the historical pattern: democratic transitions through negotiation are common; communist transitions through negotiation are nonexistent.

I conclude with a note on the paper's limitations and a broader reflection. The argument does not address whether capitalism is exploitative, whether Marx's normative vision is desirable, or whether some alternative institutional vehicle—short of revolutionary abolition of private property—might realize aspects of Marx's aspirations. It addresses only the specific claim that revolutionary seizure and transformation can produce the communist society Marx envisioned. That claim, I have argued, is refuted by the logic of credible commitment.

The broader reflection concerns the relationship between normative aspiration and institutional feasibility. Marx's vision of a society free from exploitation, alienation, and domination retains its moral force regardless of this paper's conclusions. The question is not whether the vision is attractive but whether the proposed institutional vehicle can realize it. The analysis presented here suggests that it cannot—not because the aspiration is wrong but because the mechanism is self-defeating. The institutional conditions required for credible commitment to non-exploitation are precisely the conditions the revolutionary program demands be destroyed. This is a tragedy in the strict sense: not a story of foolishness or villainy but of a program whose internal logic undermines its own objectives.

The covenants of revolutionary promise, lacking a sword that is not itself corrupted by power, remain but words.

References

- Acemoglu, Daron, and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.
- Anderson, Perry. 1976. *Considerations on Western Marxism*. London: New Left Books.
- Blackburn, Robin. 1991. *After the Fall: The Failure of Communism and the Future of Socialism*. London: Verso.
- Brinton, Crane. 1965. *The Anatomy of Revolution*. Rev. ed. New York: Vintage.
- Buchanan, James M. 1975. *The Limits of Liberty: Between Anarchy and Leviathan*. Chicago: University of Chicago Press.
- Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.
- Cohen, G.A. 2000. *If You're an Egalitarian, How Come You're So Rich?* Cambridge, MA: Harvard University Press.
- Dikotter, Frank. 2010. *Mao's Great Famine: The History of China's Most Devastating Catastrophe, 1958–1962*. London: Bloomsbury.
- Elster, Jon. 1985. *Making Sense of Marx*. Cambridge: Cambridge University Press.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414.
- Hamilton, Alexander, James Madison, and John Jay. [1788] 2003. *The Federalist Papers*. Edited by Clinton Rossiter. New York: Signet Classics.
- Hayek, F.A. 1945. "The Use of Knowledge in Society." *American Economic Review* 35 (4): 519–30.
- Hobbes, Thomas. [1651] 1991. *Leviathan*. Edited by Richard Tuck. Cambridge: Cambridge University Press.
- Kolakowski, Leszek. 2005. *Main Currents of Marxism*. New York: W.W. Norton.
- Kydland, Finn E., and Edward C. Prescott. 1977. "Rules Rather Than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85 (3): 473–1.
- Lavoie, Don. 1985. *Rivalry and Central Planning: The Socialist Calculation Debate Reconsidered*. Cambridge: Cambridge University Press.
- Lenin, V.I. [1902] 1961. *What Is to Be Done?* In *Collected Works*, vol. 5. Moscow: Progress Publishers.
- Lenin, V.I. [1917] 1992. *The State and Revolution*. London: Penguin.
- Marx, Karl. [1847] 1995. *The Poverty of Philosophy*. Amherst, NY: Prometheus Books.
- Marx, Karl. [1871] 1970. *The Civil War in France*. In *Marx/Engels Selected Works*, vol. 2. Moscow: Progress Publishers.
- Marx, Karl, and Friedrich Engels. [1848] 1998. *The Communist Manifesto*. London: Verso.
- Mises, Ludwig von. 1920. "Die Wirtschaftsrechnung im sozialistischen Gemeinwesen." *Archiv für Sozialwissenschaft und Sozialpolitik* 47: 86–121.

- Mittal, Sonia, and Barry R. Weingast. 2013. "Self-Enforcing Constitutions: With an Application to Democratic Stability in America's First Century." *Journal of Law, Economics, and Organization* 29 (2): 278–302.
- North, Douglass C., and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49 (4): 803–32.
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Paniagua, Pablo. Forthcoming. "Covenants, Swords, and Sunk Costs: A Signaling Theory of Constitutional Commitment."
- Popper, Karl R. 1945. *The Open Society and Its Enemies*. 2 vols. London: Routledge.
- Roemer, John E. 1982. *A General Theory of Exploitation and Class*. Cambridge, MA: Harvard University Press.
- Talmon, J.L. 1952. *The Origins of Totalitarian Democracy*. London: Secker and Warburg.
- Trotsky, Leon. 1937. *The Revolution Betrayed*. Garden City, NY: Doubleday.
- Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11 (1): 89–99.
- Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *American Political Science Review* 91 (2): 245–63.
- Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.