The Shape of Possible States

Robert J. Carroll

University of Illinois at Urbana-Champaign

Draft — March 12, 2026

Contents

Preface

This book asks what it means for states to be alike, and it answers the question by studying the topology of the space of possible states—where a "state" is modeled as a producer of coercive capacity. The main result is that, under regularity conditions on the production technology and cost structure, this space is contractible: topologically trivial, admitting no intrinsic divisions. The real contribution, however, is not the contractibility theorem but the admissibility conditions that make it possible and the phase boundary they define between the classical regime (where the theorem holds) and the wild regime (where it does not).

The book is part of a larger project, but it is designed to be read on its own. Two companion studies address questions that this book deliberately brackets. *A State Is Its Relations* (2026) develops the relational side of the problem: using category theory and the Yoneda lemma, it shows that a state's identity within any relational theory is constituted by its complete relational profile—the totality of relationships that every other state bears to it. Where this book establishes the structural uniformity of the possibility space, *A State Is Its Relations* establishes the relational apparatus by which states are individuated within it. *A Theory of Force* (2025) develops the compositional structure of military capability: what force is, how it decomposes, and how components compose into operational effect. Where this book takes the force target m as given, *A Theory of Force* characterizes the product that target indexes.

None of the three studies requires the others. But taken together, the whole might be more than the sum of the parts. It's a nice thought.

Chapter 1

Introduction

> *War made the state, and the state made war.*
>
> —————————————————————
>
> Charles Tilly

A system consists of interacting units whose collective behavior cannot be understood by studying those units in isolation. Nearly all systemic theories rest on the premise that the units are similar in some respect. In its strong form the premise holds that units are identical in kind: only their interactions—not their intrinsic characteristics—matter for explaining outcomes. Such theories illuminate firms in markets, birds in flocks, and neurons in brains.

So too, in principle, with states in the international system. The most influential systemic theory in international relations, structural realism, developed by Kenneth Waltz (1979), treats states as functionally identical. Like profit-maximizing firms in microeconomics, states are reduced to security-maximizing actors under anarchy; system behavior depends only on their interaction, not on their interiors. This is a strong claim, often relaxed in practice. After all, states differ in size, regime type, ideology, and much else. How, then, can they be treated as the same? And what, exactly, would sameness mean in a theory that insists on it?

This book begins from that puzzle. My aim is not to defend or attack structural realism, but to clarify what it means—and what it would take—to say that states are alike. I argue that any adequate notion of sameness must be grounded not in how states happen to be distributed, but in the structure of the *class of all possible states*. A truly systemic theory cannot rest on the empirical variation of particular cases; it must be about the space of possible ones. The

question is not whether some states are more similar than others, but whether the space of possible states admits meaningful divisions at all. If it does not, "like units" is not a simplifying assumption—it is a structural truth about the domain of theory.

1.1 The Approach

To approach this formally, I model the state as a *producer of force*: an agent that transforms resources into coercive capacity, subject to a technology and a cost structure. This is not an arbitrary starting point. From Weber (1922) and Hintze (1975) to Tilly (1990), the bellicist tradition has held that the state's defining activity is the organization and deployment of coercion. Even if war no longer makes states, producing and managing coercive power remains necessary, if not sufficient, for statehood. To model the state as a force producer is to begin from the one dimension of its activity that all states share—a minimalist ontology, but a defensible one.

For any technology τ and cost function κ, the state's problem is to choose the resource allocation x that most efficiently produces a desired level of force m. This is the *State Production Problem* (SPP):

$$\min_x \ \kappa(x) \quad \text{subject to} \quad \tau(x) = m.$$

Solving the SPP for every possible m yields a *policy function* $\pi_{\tau,\kappa} \colon M \to X$, specifying how the state mobilizes for any force target. The collection of all such policy functions—across all admissible pairs (τ, κ)—is the *space of possible states*, $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$. Studying the topology and geometry of this space is a way of studying the logic of the state itself.

1.2 The Main Results

The analysis proceeds in three movements, each asking a different question about the space of possible states.

The classical topology. Under regularity conditions on τ and κ—the *admissibility conditions*—the space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is *contractible*. Contractibility means the entire space can be continuously shrunk to a single point: every member can be deformed into every other without leaving the space. The class of all admissible

states is topologically trivial—any member can represent any other. Defined by force production alone, there is only one kind of state.

More precisely, there is a *ladder of sameness*: as we strengthen the notion of structural equivalence, states are alike in increasingly strong senses.

1. *Connected*: no first-order qualitative divisions among states.

2. *Path-connected*: any state can be continuously deformed into any other.

3. *Simply connected*: any two paths between states are homotopic; path dependence cannot arise from the topology alone.

These three properties follow from contractibility of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, which is itself derived from contractibility of the technology and cost spaces individually. "Like units" becomes a theorem, not a heuristic. This is the clean result, and it vindicates Waltz—within the classical regime.

Modeling the classical. If the space of possible states is contractible, what kinds of models can represent it faithfully? This is not merely a technical question. Following van Fraassen (1980), I argue that the criterion for representational adequacy must be structural: a submodel is adequate not when it approximates the ambient space pointwise, but when it can be *retracted* from it via a continuous, admissibility-respecting deformation. I introduce a subclass—*tame states*, with log-linear technologies and linear costs—and show that tamification defines a strong deformation retract of the full classical space. No topological information is lost.

Within the tame subspace, richer structure emerges. The set of tame states is not merely contractible but *convex*: paths are straight, interpolations are states, and the geometry carries no intrinsic memory. Convexity deepens each rung of the ladder of sameness and opens the door to differential structure—where the calculus of 1-forms and exterior derivatives can diagnose exactly where, and how, path dependence enters the system from the dynamics imposed upon it. The tame model is not a simplification but an illumination: a structurally adequate representation that makes the underlying geometry of statehood visible.

What the classical models. The admissibility conditions are not innocent. They purchase the contractibility theorem by restricting attention to states whose production problems have unique solutions. When they fail—when the production

problem admits multiple locally optimal modes of force production—the topology is no longer trivial. The space of *wild* states admits disconnected components, topological path dependence, and branch structure that the classical case rules out.

The natural question is whether the classical model is adequate for these wild states in the same sense that the tame model is adequate for the classical ones. It is not. The shadow map that compresses wild states into their classical silhouettes fails the structural adequacy criterion: the deformation path passes through inadmissible territory, violating path-fidelity. The classical whiteboard is a faithful model of its tame interior but a lossy projection of the wild room it claims to represent. This asymmetry—faithful inward, lossy outward—is as much a result of the book as the contractibility theorem itself.

1.3   Why the Shape Matters

The results above are stated in the language of topology, but their consequences are practical.

First, the admissibility conditions provide a *testable criterion for when comparison is warranted*. Comparativists routinely place states on continuous scales—strong to weak, high capacity to low—without asking whether the underlying space supports such comparison. The contractibility theorem says it does, but only when both states satisfy the admissibility conditions. If one does not, comparison through the classical framework is lossy: it preserves the cost envelope but discards the branch structure that makes the state's situation structurally distinct. The admissibility conditions tell you when the comparison is legitimate and when it is lossy, and they tell you what is lost.

Second, the topology provides a *diagnostic for path dependence*. In the classical regime, the space is simply connected: path dependence cannot arise from the topology of the production problem. If one observes what looks like path dependence among admissible states, the exterior derivative developed in Chapter 5 (within the tame subspace where differential structure is available) locates the source in the dynamics imposed on the space, not in the space itself. In the wild regime, by contrast, path dependence is topological—monodromy and branch structure make it a feature of the space that no change in dynamics can undo. The distinction matters: dynamic path dependence can, in principle, be overcome by changing the policy that drives transitions; topological path dependence cannot.

Third, the wild regime offers a *formal language for structural traps*. A state whose production problem violates admissibility may occupy a locally optimal branch of its production correspondence—not because it lacks resources but because transitioning to a different branch would require passing through a catastrophe. The state is not weak; it is trapped. This reframes questions about persistent military dysfunction, warlord politics, and coup-proofing: these may be structurally distinct modes of force production, locally stable under their own logic, rather than failures of a single mode.

These consequences are developed at length in the conclusion. They are mentioned here to make the stakes clear: the topology of the space of possible states is not an abstract curiosity but a diagnostic tool, and the admissibility conditions are not technical assumptions but substantive claims about the political economy of force production whose truth or falsehood determines what kind of theory the state requires.

1.4 What This Book Is Not

Two limits are worth stating plainly.

First, I take only one slice of what states do. They tax, legitimate, regulate, and build. But if the state has many faces, force production is the one it cannot do without, and it furnishes a clean foundation for formal analysis. The fiscal state, the regulatory state, and the protective state can each be modeled as problem-solvers with their own production problems; the method generalizes. This book develops the force-production instance because the international relations literature specifically delivers a force target as the relevant variable, and because the bellicist tradition identifies force production as the one activity all states share.

Second, this is a study of a function in isolation. A state is more than a policy function; its relations to other states matter. What individuates states—what makes France different from Germany even if both solve classically admissible production problems—is a relational question that lies outside the scope of this book. The claim here is the structural one: if we look at force production alone, the classical regime is unified, the tame model faithfully represents it, and the wild regime marks the boundary where that unity breaks down.

1.5 Intellectual Lineage

The argument draws on four intellectual traditions.

The ontological status of the state. Ringmar (1996) distinguishes realist views (the state as pre-given unit) from pluralist views (the state as an emergent bundle of sub-units), and argues that neither resolves the question of what the state *is*. Bartelson (2001) deepens this: sovereignty is *parergonal*—it constitutes the inside/outside divide without belonging to either side, making it resistant to direct theorization. Chapter 2 develops these points at length, arguing that the force-production model is licensed precisely because it operates *below* the parergonal level: it characterizes the content of the frame without trying to theorize the frame itself.

The state as a firm. Lane (1979) envisioned the state as a firm selling protection in exchange for revenue. Similar ideas appear in North (1981), Levi (1988), Lake (1992), and Kier (2021). These accounts converge on a basic insight: the state governs, in part, by producing and distributing coercive capacity. The approach here is more abstract and more literal. I treat the state as a production function mapping inputs to outputs, in the spirit of the cost-minimizing firm in microeconomic theory (Mas-Colell, Whinston and Green, 1995). The point is not to reduce the state to a firm, but to use the firm-as-producer template as a clean and tractable starting point.

The bellicist tradition. From Weber and Hintze through Tilly, a longstanding line of thought holds that states are born in organized violence. Here I take no stand on whether war-making is constitutive of statehood or merely instrumental to it. I begin from a stylized fact: states possess and produce force. That premise is consistent with the bellicist tradition, but the framework is broad enough to accommodate others.

Philosophy of scientific representation. The book's central methodological question—when does a submodel faithfully represent its ambient space?—draws on a tradition in philosophy of science that includes Hertz's (1894) triad of permissibility, correctness, and appropriateness; van Fraassen's (1980) structural empiricism; and Weisberg's (2013) account of representational ideals. Chapter 5 argues that our setting selects a *topological completeness* ideal: the objects being

represented are function spaces whose identity is constituted by their topological structure, so a model must preserve that structure to count as a model *of that space* rather than merely a convenient subset. The formal criterion—structural adequacy via strong deformation retraction—is where this tradition meets the mathematics.

1.6 Plan of the Book

Chapter 2: What States Are. Before the formal model, a philosophical question demands attention: what kind of thing is a state, and what would it mean for a formal model to characterize it correctly? This chapter develops the book's ontological position. It argues that states are *functional unities*—defined by the role they play in force production, not by any particular physical or institutional substrate. Drawing on Putnam's multiple realizability, Fodor's philosophy of special sciences, and Epstein's grounding-anchoring framework, it defends the view that "state" is a legitimate special-science predicate that carves nature at joints neither physics nor sociology recognizes. It also addresses objections from Cartwright's simulacrum account, Friedman's instrumentalism, and Hacking's interactive kinds.

Chapter 3: The State as a Producer of Force. This chapter introduces the model. It motivates the SPP through a stylized two-state arming game, defines the three primitives (technology τ, cost κ, target m), and derives the policy function $\pi_{\tau,\kappa}$ as the SPP's solution. Lemmas in Chapter 4 establish existence and uniqueness of the solution and its continuity in all parameters.

Chapter 4: The Classical Topology. This chapter proves the main topological results for the classical regime—the states whose production problems satisfy the admissibility conditions. It shows that the technology and cost spaces are individually contractible, and that contractibility propagates to the full policy function space $\mathcal{P}$ via the currying construction. The ladder of sameness (connected $\rightarrow$ path-connected $\rightarrow$ simply connected) is derived as a sequence of corollaries. Within the classical regime, Waltz is right: states are like units as a topological theorem.

Chapter 5: Modeling the Classical. This chapter develops the book's modeling theory: when does a submodel faithfully represent the space it lives in?

Drawing on van Fraassen (1980), Hertz (1894), and Weisberg (2013), it argues that representational adequacy is a matter of structural retractability, not pointwise approximation. It introduces tame states, shows that tamification defines a strong deformation retract of the classical space, and proves that the tame subspace is convex. Convexity opens the door to differential structure: the exterior derivative becomes a diagnostic for path dependence, locating historical memory not in the geometry of the arena but in the non-closedness of the fields that drive transitions across it. The chapter closes by returning to the motivating arming game and deriving the water-filling equilibrium in closed form.

Chapter 6: What the Classical Models. The preceding chapters build the classical whiteboard and show that the tame model faithfully represents it. This chapter asks what happens when we step outside the whiteboard. Wild states—those whose production problems violate the admissibility conditions—admit multiple locally optimal modes of force production, disconnected components, and topological path dependence that the classical case rules out. The shadow map that compresses wild states into their classical silhouettes fails the structural adequacy criterion established in Chapter 5: the deformation path passes through inadmissible territory. The classical model is faithful inward but lossy outward, and the chapter confronts this asymmetry directly.

Chapter 7: Conclusion. The conclusion executes a reversal. The contractibility theorem is, in a precise sense, the book's least interesting finding—it follows from the admissibility conditions by standard arguments. The real contribution is the admissibility conditions themselves and the phase boundary they define between the classical and wild regimes. The chapter traces the consequences of the faithful-inward, lossy-outward asymmetry for existing scholarship on state fragility and military organization, confronts objections from realist, constructivist, and empiricist perspectives, and sketches how the method generalizes beyond force production to the fiscal and regulatory state.

Appendix A: Mathematical Primer. For readers less familiar with the topological and optimization machinery, a self-contained primer collects the definitions and standard results invoked in the main text.

Appendix B: Proofs. All proofs deferred from the main chapters are collected here. Results in the text are stated with references to their proofs; the appendix

is self-contained for readers who want the technical details.

Chapter 2

What States Are

To be is to be the value of a variable.

———————————————————

W. V. O. Quine

Every formal theory of international politics makes an ontological commitment before it makes a single empirical claim. The commitment is to what kind of thing the state is. This commitment is unavoidable, and it is consequential: the structure of the formal analysis depends on it. What follows examines that commitment directly.

The argument has four movements. Sections 2.1–2.2 show that existing accounts are circular or collapse into what they were meant to explain. Section 2.3 introduces the central move: descend below the sovereignty question to the force production problem. Sections 2.4–2.6 develop the positive ontological position and connect it to the contractibility theorem. Section 2.7 addresses four objections.

2.1 The Question No One Can Avoid

2.1.1 Like units, unlike inquiry

Every IR theory that uses "the state" as a unit of analysis has already made an ontological decision. Structural realism makes it by fiat: states are "like units," sovereign actors under anarchy (Waltz, 1979). The democratic peace makes it by attribution: states have regime types that generate preferences and constrain behavior. Constructivism makes it by constitution: states are bearers

10

of identities whose stability is an ongoing social achievement (Ringmar, 1996). None of these moves is ontologically neutral.

The standard response is to bracket the question: treat the state as whatever the theory needs it to be and proceed. But bracketing is itself a commitment—the commitment that the ontological question will not bite back. It bites back whenever a theory is pushed to explain why its units have the properties it attributes to them. Why are Waltz's units functionally like? Why do democratic states have different preferences than autocracies? Why do identities persist? Each question demands an answer that the theory cannot give without descending into the ontology it chose to bracket. Wight (2006, pp. 19–21) identifies the failure as systematic: the academic study of international relations has never unpacked the deeply embedded and often implicit ontologies underlying its theoretical commitments.

2.1.2   Three stances, none sufficient

Three major positions structure the debate, and none produces a derivable structure for the space of possible states.

The *realist* position treats the state as a pre-given unit: sovereign, transcendental, the starting point of analysis rather than its product. Waltz is the extreme case. By making the state a primitive, he purchases parsimony at the cost of accessibility: the unit of analysis cannot itself be analyzed. The ontological question is foreclosed by assumption.

The *pluralist* position dissolves the state into its components: domestic coalitions, bureaucratic actors, interest groups. There is no state above and beyond the interactions of the parts. This purchase comes at the opposite cost: once you look inside, only domestic actors remain, and the systemic question—what kind of thing is doing the interacting at the international level?—never gets answered.

The *constructivist* position makes the state a relational product: states are constituted by practices of recognition, intersubjective understandings, and the mutual attribution of sovereignty (Wendt, 1999). The insight is genuine but the result is indeterminate: if the state is constituted by recognition, then its properties are as varied as the practices that constitute them, and there is nothing to say about the structure of the space of possible states as such.

2.1.3 What this chapter does

This chapter does not resolve the metaphysics of sovereignty. It descends below that question and shows that one well-motivated model—the state as force producer—yields a substantive structural result derivable without presupposing what the model is supposed to establish. The argument is constructive: it produces a positive ontological position (the state as functional unity), a scope condition (classical admissibility), and a derived result (contractibility of the policy space). Sovereignty questions are not answered; they are bracketed at a principled level for principled reasons.

The payoff is a theorem—and a question the theorem opens. The space of all classically admissible policy functions is contractible: every possible classical state can be continuously deformed into every other, with no topological barriers separating them. "Like units" follows as a consequence of the production problem's structure, not as an assumption about what states happen to be. But whether the classical whiteboard is an adequate model of the full room— whether the well-behaved states faithfully represent the space that includes the wild ones—is a question the theorem frames but cannot answer on its own.

2.2 Two Diagnoses, Both Correct

2.2.1 Ringmar: the two-way vanishing trick

Ringmar (1996) identifies a symmetrical problem. Realists make the state transcendentally given and thus inaccessible to inquiry: it is the unit of analysis, so it cannot be analyzed. Pluralists dissolve the state into sub-units and domestic coalitions, losing the state in the process of theorizing it. The state "vanishes in both directions," as Ringmar puts it—upward into transcendence, downward into reduction.

Ringmar's diagnosis is precise. The state's unity is simultaneously assumed and unexplained. Realism presupposes it to generate predictions; pluralism denies it to achieve micro-foundations. Neither tradition asks what kind of thing would have to exist for the presupposition to be true or the denial to be interesting.

Ringmar's proposed solution is narrative ontology: a state becomes a state by telling a story about itself in terms recognizable to the community of states. The state acquires identity through narrative performance in a field of mutual recognition (pp. 452–455). This is a rich and philosophically attractive answer.

But it is not available to formal theory. Narrative identity is not mathematically tractable; it does not yield a space of possible states whose topology can be studied. The diagnosis stands; the solution is not the one we need.

2.2.2 Bartelson: the parergon

Bartelson (1995) develops the problem more radically. Sovereignty is *parergonal*: it constitutes the domestic inside and the international outside as distinct domains without itself belonging to either. A parergon is a frame—the frame of a painting, the marginal gloss of a text—that makes the work possible without being part of the work. Sovereignty is the frame of international politics: it divides inside from outside, domestic from international, and thereby makes both domains available for theorization. But it cannot be theorized from either side it constitutes.

The consequence for structural realism is damaging. Waltz's circularity runs as follows (Bartelson, 1995, pp. 34–36): states are "like units" because they are sovereign; they are sovereign because anarchy (which presupposes like units) produces them. The like-units claim is not derived from anything more fundamental—it is definitionally linked to sovereignty, which is definitionally linked to anarchy, which presupposes like units. The argument is a closed loop.

Bartelson's response is genealogy: trace the historical constitution of the sovereignty concept and show that its present meaning is a contingent sedimentation, not a necessary truth. Genealogy is effective as critique but unavailable as foundation. It dissolves the claim rather than grounding it.

2.2.3 The gap both open

Together, Ringmar and Bartelson clear the ground. Ringmar shows that the state's unity is a real question that existing theories evade. Bartelson shows that sovereignty-based answers are circular and that the circularity is structural, not accidental. Neither offers a way to ground the like-units claim non-circularly while remaining amenable to formal analysis.

That is this chapter's entry point. What is needed is an account that operates below the parergonal level—that characterizes the content of the frame rather than the frame itself—and that generates derivable structural results. The argument of this book provides that account.

2.3 Below the Parergon

2.3.1 Weber's license

The move below the parergon is licensed by a commitment already present in the bellicist tradition.

Weber (1922) insists that the state cannot be defined by the ends it pursues—welfare, justice, order—because states pursue all manner of ends. The state is defined, instead, by its *specific means*: the legitimate use of physical force within a territorial domain (§§ 12, 17). Not its purposes but its instruments are definitive.

This has an immediate consequence. If force is the state's defining means, then modeling the state as a producer of force is not a simplification—it is a definition. The descent below the parergon is licensed by the one ontological commitment the bellicist tradition has always made: the state is whatever monopolizes force in a territory. We are not replacing the Weberian state with a technical construct; we are formalizing the one feature Weber identifies as essential.

Tilly (1990) deepens the insight. War-making and state-making are not merely correlated; they are structurally linked through the logic of resource extraction and force deployment. The state's coercive apparatus emerges from, and is continuously shaped by, the demands of violent competition. Even in a world where interstate war has become less frequent, the structural relationship between resource mobilization and force capacity persists. The force production problem is not an artifact of a particular historical period; it is the core organizational challenge of statehood as such.

2.3.2 The inside/outside architecture of the SPP

The State Production Problem (SPP) instantiates the parergonal inside/outside divide without presupposing any particular theory of sovereignty.

The force target m^* arrives from *without*. It is determined by the systemic contest—the distribution of capabilities, the geometry of threats, the logic of the arming game described in Chapter 3. This is the Waltzian black box: the systemic level delivers a demand to the state without specifying how that demand is met.

The resource allocation x^* is produced from *within*. The state mobilizes personnel, materiel, and organizational capacity subject to its technology τ and cost structure κ. This is Tilly's content: the state builds its coercive apparatus through the extraction and organization of domestic resources.

The policy function $\pi_{\tau,\kappa} \colon M \to X$ is the formal object that bridges the two sides. It maps the external demand (force target) to the internal response (resource allocation). $\pi_{\tau,\kappa}$ is what the state *is* in the model: the function that transforms systemic pressure into domestic organization.

The parergon is not resolved by this architecture—it is presupposed. The SPP does not explain why there is an inside and an outside, or what makes one side domestic and the other international. It accepts the existence of the frame and characterizes the content of the frame: what is the structure of the function that solves the state's production problem? That is the question the book addresses.

2.3.3 What descent purchases

The force production problem has mathematical structure that sovereignty questions do not. The technology function $\tau \colon X \to M$ and cost function $\kappa \colon X \to \mathbb{R}$ live in function spaces with well-defined topology. The space of admissible pairs (τ, κ) can be given a natural metric and its properties studied. Contractibility, convexity, and the strong deformation retract relationship are properties of these mathematical objects—properties that can be proved or disproved, not merely asserted.

The parergon's intractability is not inherited by what lies below it. Whether the frame of sovereignty is philosophically resolvable is a question this book deliberately brackets. Whether the solution space of the SPP is contractible is a question with a definite answer, derived in Chapter 4. Working below the parergon means working at a level where mathematical methods are available.

2.4 The State as Functional Unity

2.4.1 The president, the guns, the election

The force production problem is a mathematical object. But the entity that solves it—the state—requires ontological characterization before mathematical analysis can be brought to bear. What kind of thing is a state, and what makes its parts a unity?

Consider the parts of a state: the president, the armed forces, the electoral apparatus, the revenue service, the bureaucracy. Each part is real in a straightforward sense—it consists of people, buildings, equipment, legal instruments.

The president is a person. The army is a collection of persons, weapons, and organizations. The election is an event with physical outcomes.

But the thing that makes all of these components of a single whole—the unity that constitutes them as parts of one state rather than parts of many separate institutions—is not real in the same sense. The state's unity is not an additional physical object alongside the president and the guns. It is a normative and relational fact: a web of recognitions, legal forms, collective intentionality, and practices through which disparate elements are organized into a functional pattern.

2.4.2 Social kinds and model-relative ontology

Social kinds—states, firms, elections, contracts—are constituted, not discovered. Unlike natural kinds (gold, water, electrons), social kinds do not exist independently of the practices and representations through which they are organized. This does not mean they are not real. But it does mean that to ask a meaningful existence question about a social kind, we must already be using a linguistic framework that introduces the relevant terms with determinate rules of use.

Carnap (1950) draws the fundamental distinction between *internal* and *external* existence questions.[1] An internal question is one asked *using* the terms of a linguistic framework, governed by the rules of use that introduce them into the language. An external question either attempts to ask about the framework's entities from outside any framework (yielding a pseudo-question—a term without rules of use has no determinate meaning) or is charitably read as a pragmatic question about whether to adopt the framework.

The SPP framework introduces terms—"state," "technology function," "cost function," "admissible policy function"—with explicit rules of use given by the model's definitions. Once those rules are in place, two distinct questions arise.

The framework question (external): should we model states as force producers? This is a pragmatic question, but not an arbitrary one. Carnap's point is that framework adoption "can be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended" (1950/1956, 214). For the SPP framework, the case is strong: the bellicist tradition converges on

---

[1]Thomasson (2015, Ch. 1) gives the cleanest reconstruction of Carnap's position: internal questions are those that *use* the terms of a framework, governed by the rules of use that introduce them; external questions either attempt to use those terms while severing them from their rules (yielding pseudo-questions) or, charitably, *mention* the terms while asking whether we should adopt the framework at all.

force production as the state's defining activity; the framework supports formal analysis unavailable without it; and the results are non-trivial. The framework is adopted not by ontological discovery but by reasoned practical judgment.

The constitutive alternative is worth naming. Badiou (2005) argues that set theory is not a framework chosen because it illuminates politics—it is the structure of being itself, of which politics is an instance. There is no framework question for Badiou, no moment of asking whether the mathematics is an expedient tool; set theory just is what ontology is. This is an impressive program, and it establishes that formal mathematics can be brought to bear on political ontology without trivializing either. But it is not the stance taken here. The contractibility of $\mathcal{P}$ is a result within a well-motivated framework, not a theorem about the fabric of political reality as such; and the mathematics that matters for force production—continuous maps, homotopy, deformation retracts—is not the combinatorics of inclusion and membership that set theory provides.[2]

The contractibility question (internal): is $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ contractible? This is an internal question, asked using the terms introduced by the SPP framework, and answerable by the mathematical methods of Chapters 4–5. As Carnap notes, internal questions can be answered "by purely logical methods or by empirical methods, depending upon whether the framework is a logical or a factual one" (1950/1956, 206). The SPP framework is mathematical; the contractibility question is answered by mathematical proof. The answer is simply true or false—not "true within the framework" in any relativist sense. When we prove that $\mathcal{P}$ is contractible, we establish a genuine structural fact, not a fiction or a stipulation. The position differs from instrumentalism, which treats any model as a useful fiction evaluable only by predictive success. Predictive success is the wrong criterion here: the topology of $\mathcal{P}$ characterizes what kind of thing the state is, not what states will do under specified conditions. The difference from natural kinds is not that social kinds are less real, but that their reality is accessible only via a framework—and for social kinds, that is what realism amounts to.

This brings us back to the epigraph. Quine's (1948) criterion holds that a theory's ontological commitments are determined by what its bound variables range over: to be is to be the value of a variable. As Lowe (2005) puts it, an ontology is "the set of things whose existence is acknowledged by a particular theory." The convergence is not accidental. When a computer scientist announces an ontology—a formal specification of the entities, types, and relations that a

---

[2]The set-theoretic alternative, and Badiou's program for grounding ontology in it, require a categorical response that lies outside the scope of this book.

system quantifies over—the operative meaning is the same: to declare an ontology is to declare a domain of quantification.[3] If systemic theory quantifies over states, then its ontological commitment is to the *space* of entities that "state" ranges over—and the structure of that space is the structure of the commitment. The project of this book is to study that structure directly. To posit the state is to posit a set of possible states; the topological and geometric properties of that set become the defining properties of the kind. Contractibility is not a property of any particular state but of the space of all admissible ones—a Quinean finding about the domain of quantification itself.

The position is intermediate between Quine and Carnap. It accepts the Carnapian point that "state" acquires determinate meaning only within a framework—the SPP framework introduces the term with explicit rules of use. But it sides with Quine in treating the structural facts about the domain as genuine ontological commitments, not mere artifacts of framework choice.

Quine's direct response to Carnap makes the case sharply. In "On Carnap's Views on Ontology," Quine argues that if the distinction between analytic and synthetic cannot be maintained, then "no basis at all remains for the contrast which Carnap urges between ontological statements and empirical statements of existence" (Quine, 1951, 71). Ontological questions, on this view, "end up on a par with questions of natural science"—and the contractibility of $\mathcal{P}$ is precisely such a question: a structural finding about a domain of quantification, answerable by mathematical proof, whose truth does not depend on whether we call the framework analytic or synthetic.[4]

When we prove that $\mathcal{P}$ is contractible, we do not merely report a fact internal to the SPP language; we discover a structural constraint on any theory that individuates states by force production. The framework is chosen; the topology is found.

---

[3]The canonical formulation is Gruber's (1993): "an ontology is an explicit specification of a conceptualization"—that is, of the objects, relations, and constraints that a representational vocabulary is committed to. The usage has drifted from its philosophical origins, but the Quinean core—to declare what exists is to declare what the variables range over—remains intact.

[4]The Quine–Carnap exchange has generated a substantial secondary literature. Bird (1995) argues that Quine's criticisms leave Carnap's central points untouched; Alspector-Kelly (2001) presses the case further. The position taken here does not require adjudicating the full dispute— only the narrower claim that structural facts about domains of quantification are genuine findings, not artifacts of framework adoption.

2.4.3 Functional individuation: Putnam's license

How should a kind like "state" be individuated—by physical substrate, by institutional form, or by functional role?

Putnam (1975) argues, in the context of mental states, that individuation by functional role is both possible and correct. Pain is pain not in virtue of any particular neural substrate but in virtue of the causal-functional role it plays: tissue damage causes it, avoidance behavior results from it. The same role can be realized in mammalian nervous systems, in Martian biochemistry, or in silicon—multiple realizability is not a defect of functional individuation but a consequence of getting the level of analysis right. What makes something a pain is what it *does*, not what it is made of.

The same argument applies to political institutions. "The state" is a functional unity defined by what it does: it produces force. More precisely, it solves the force production problem—it maintains a functioning policy function $\pi_{\tau,\kappa}$ mapping force targets to resource allocations. A presidential republic with an all-volunteer military, a constitutional monarchy with a conscript army, and a one-party state with a professionalized security apparatus can all realize the same $\pi_{\tau,\kappa}$ and thus be the same *kind* of state in the sense relevant to a systemic theory. Their institutional details differ; their structural role is the same.

This is multiple realizability at the institutional level. The force production function is the functional role; different institutional forms are its realizers. Abstracting from institutional particulars is not a methodological shortcut—it is the consequence of individuating states by the right criterion. Functional individuation is the correct criterion for the kind of entity that participates in systemic international politics.

2.4.4 The special science of states: Fodor's argument

Functional individuation implies that "state" is a special science predicate. Fodor (1974) argues that the special sciences—economics, psychology, biology—are legitimate even though their predicates are not reducible to the predicates of more fundamental sciences. Special science predicates are multiply realizable: the neurological predicate that realizes "pain" in one species may differ from the predicate that realizes it in another, but the economic, psychological, or behavioral regularity stated in terms of "pain" holds nonetheless. Reduction would destroy the regularity: no predicate at the level of physics or neuroscience picks out the same natural kind.

"State" is a special science predicate in exactly this sense. Multiple physical and institutional realizers satisfy the functional description; no description at the level of sociology, organizational theory, or physics picks out the same class of entities. The regularities that hold for states as force producers—including the contractibility of their policy space—are regularities at the level of the special science, not at a more fundamental level. $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is the special science's taxonomy: it is the right level of description at which structural regularities can be identified and derived.

The admissibility conditions play a specific role in this picture: they define the *ceteris paribus* scope of the special science law. Fodor's regularities hold only when the background conditions are right. For states, the background condition is classical admissibility—the requirement that the production problem has a unique, continuously varying solution. When this condition is satisfied, the structural results of Chapters 4–5 hold: the state space is contractible and the ladder of sameness applies. When it fails, the ceteris paribus clause is violated and the structural results change qualitatively—not because the theory has nothing to say, but because the wild regime has its own topology, developed in Chapter 6. Disunity of science is a feature, not a bug: political science does not need to reduce to sociology or physics to be legitimate.

2.4.5 Grounding versus anchoring: Epstein's distinction

The book makes a *grounding* claim, not an *anchoring* claim. These are distinct and must not be conflated.

A grounding claim specifies what makes it true, right now, that some entity has a given property. An anchoring claim specifies the historical, intentional, or social processes that brought it about that the grounding relation holds. Epstein (2015) insists on this distinction because conflating it generates systematic confusion in social ontology. The conditions that make something true today are not the same as the processes that produced those conditions.

The grounding claim of this book is: a concrete polity is "a state" in the relevant sense if and only if it realizes some admissible policy function $\pi_{\tau, \kappa}$. What makes it true that France is a state is that France has a functioning force production apparatus that maps military requirements to resource allocations continuously and uniquely. This is a structural ground—it specifies what constitutes France's statehood at the relevant level of description.

The anchoring stories are separate. Lane (1979) tells one: the state emerges when a protection entrepreneur achieves the scale economies needed to stabilize

and tax a territory. Levi (1988) tells another: rulers extract resources through bargaining constrained by their relative power and transaction costs. Tilly (1990) tells a third: war-making and state-making coevolve through a process of competitive selection. These are all anchoring stories—accounts of how the grounding relation came to hold. The book does not adjudicate among them because it does not need to. The grounding question and the anchoring question have different answers, and the formal analysis addresses only the grounding question.

2.5 Classical and Wild States

2.5.1 What the admissibility conditions require

The contractibility result of Chapter 4 holds under conditions of *classical admissibility*: log-concavity of the technology function τ and exp-convexity of the cost function κ. These are the conditions under which the SPP has a unique, continuously varying solution and the policy function $\pi_{\tau,\kappa}$ is well-defined.

Log-concavity of τ is a diminishing-returns condition on force production: each additional unit of input yields a strictly smaller increment to force capacity than the last, in the specific sense that $\log \tau$ is concave. This is satisfied by any force production technology whose marginal returns decay smoothly—the organizational pattern characteristic of bureaucratically coordinated militaries with complementary inputs and no hard indivisibilities.

Exp-convexity of κ governs cost structure: the map $x \mapsto \exp(\kappa(x))$ is convex. This is weaker than requiring κ itself to be convex—it permits cost functions that are concave, exhibiting economies of scale, provided they are not too concave (Section 3.5.2). The condition says that extraction costs accelerate fast enough to prevent the state from exploiting scale economies indefinitely: the political resistance, fiscal strain, and social disruption that accompany mass mobilization eventually dominate.

Together, the admissibility conditions characterize the class of *classical states*: those whose production problems are well-posed in the sense of microeconomic theory. Uniqueness and continuity of the solution are not assumed—they follow from the structure of the problem.

2.5.2 Waltz and Tilly were always theorizing classical states

The admissibility conditions do not restrict the class of states being theorized; they make explicit what the bellicist and realist traditions were always implicitly assuming.

Tilly's war-making state is a bureaucratically organized institution that extracts resources from a territory and deploys them as coercive capacity. It is not an entity with indivisible resources, combinatorial logistics, or multiple organizational equilibria; it is a continuous extraction and deployment machine, shaped by the competitive pressures of the state system. Tilly's state satisfies the admissibility conditions.

Levi (1988) provides a micro-foundation for why organized states tend to develop exactly this organizational form: rulers who can reduce transaction costs and increase their bargaining power achieve the coherence that makes continuous resource extraction possible. The classical state is not merely an idealization; it is the organizational form that emerges from competitive selection among coercive actors.

Waltz's structural realism requires that the arming game have stable equilibria. If each state's production problem had multiple solutions, the arming game's equilibrium analysis would collapse—there would be no unique response function to compute best responses against. Structural realism presupposes unique, continuous policy functions. The admissibility conditions make this presupposition explicit.

The leading formal extension of structural realism inherits the same presupposition. Fearon (1995) models war as a bargaining failure between states whose probability of victory p and cost of conflict c are taken as given primitives. Fearon himself draws the connection: "what I refer to here as 'rationalist explanations for war' could just as well be called 'neorealist explanations'" (Fearon, 1995, 380). But if rationalist bargaining theory just is neorealism formalized, then it inherits neorealism's foundational gap: p and c encode everything the SPP is designed to derive—p depends on the relative force each state can produce, c on the cost at which it can be produced—yet both are taken as given. The SPP is precisely that missing account. It derives p and c from τ and κ, bridging the "without" question (how do states relate in the international system?) and the "within" question (how are states constituted as the units that enter that system?). Without the bridge, the reduction of states to like units is an assumption imported by the equilibrium, not a theorem about what states are.

There is a further point, developed fully in the Conclusion. The admissibil-

ity conditions are not merely scope conditions that delimit where the theorem applies. They encode specific claims about the political economy of force production: that technologies exhibit diminishing returns (Section 3.4.4), that extraction costs accelerate with scale (Section 3.5.4). These are institutional achievements, not mathematical conveniences. A state earns its way into the classical regime by building the organizational capacity and fiscal infrastructure that make its production problem well-posed. The "like units" result is therefore not free: its price is admissibility, and admissibility is where the political science of force production lives.

2.5.3 Wild states

When classical conditions fail, the SPP may have multiple solutions: the policy function becomes a *correspondence* $\pi\colon M \rightrightarrows X$, assigning to each force target a set of possible resource allocations. The topology of a space of correspondences is potentially much richer than the topology of a space of functions.

A state whose force production technology exhibits increasing returns to particular combinations of resources, hard indivisibilities in force inputs, or multiple organizational equilibria corresponding to different command structures may violate the admissibility conditions. Whether any particular state's production problem is wild is an empirical question, not a classification to be read off regime type or institutional form. But where the conditions do fail, the SPP has multiple solutions and the policy function cannot vary continuously.

Chapter 6 takes up the wild case directly. When log-concavity fails, the production problem forces complete specialization in a single input; the solution jumps discontinuously at bifurcation points in the cost structure; and the resulting state space has nontrivial π_0 (disconnected components) and, in higher dimension, nontrivial π_1 (topological path dependence via monodromy around cusp singularities). Each rung of the ladder of sameness that the classical case satisfies, the wild case can violate. The wild state is not outside the theory; it is the regime in which the topology becomes genuinely interesting—where the shape of the state space is no longer trivial, and where the political processes that navigate that space leave traces that the topology cannot erase.

The formal relationship between the two regimes is mediated by a *classical shadow map*: every wild state projects onto a classical state that preserves its fiscal silhouette (the cost of producing force at each level) while discarding the branch structure that makes the wild topology nontrivial. Berge's Maximum Theorem guarantees that the shadow is well-behaved—continuous and single-

valued—even when the underlying correspondence is not. The shadow is a genuine approximation, not a faithful representation: it is lossy in precisely the information that the wild topology encodes.

2.6 What the Contractibility Result Shows

2.6.1 The result stated non-technically

The main topological result of Chapter 4 is:

2.1 Slogan

The space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ of all classically admissible policy functions is contractible.

Contractibility is the strongest form of topological simplicity. A space is contractible if there is a continuous deformation (a homotopy) that carries the entire space to a single point. This means: every possible classical state can be continuously deformed into every other possible classical state without leaving the space of classical states. There are no topological barriers between states; no first-order qualitative divisions that no continuous path can cross; no regions of the state space that are structurally inaccessible from others.

This is not a claim about any particular state. It is a claim about the *space* of possible states—the topology of the class of all functional unities that solve the force production problem classically. Waltz's "like units" becomes a structural theorem about this space, not an empirical generalization about states that happen to exist, and not a methodological assumption entered for analytical convenience. It is a consequence of the production problem's structure.

2.6.2 The ladder of sameness

Contractibility implies a ladder of successively stronger structural properties, each carrying its own ontological interpretation.

1. *Connected*: the space $\mathcal{P}$ cannot be partitioned into two disjoint nonempty open subsets. *Ontological reading:* there are no first-order qualitative divisions among classical states. No property of the production problem marks a fundamental boundary between two kinds of state.

2. *Path-connected*: any two policy functions can be connected by a continuous path within $\mathcal{P}$. *Ontological reading:* any classical state can be continuously transformed into any other. The most "different" states—in technology, cost structure, or force profile—are connected by a continuous transformation that stays within the class of admissible states.

3. *Simply connected*: every closed path in $\mathcal{P}$ can be continuously contracted to a point. *Ontological reading:* path dependence cannot arise from the topology of the state space alone. If historical or institutional path dependence is real, its source is in the field structure (non-closed differential forms on the state space) or in initial conditions, not in topological obstructions. The topology imposes no memory.

Together these three properties constitute a formal ladder of sameness. As the notion of structural equivalence strengthens from connectedness through contractibility, states are alike in increasingly demanding senses. The topological results cash out what "like units" means across four levels of structural commitment.

2.6.3   Why derivation is necessary: the exogeneity argument

The contractibility result is *derived*, not assumed. This matters more than it might appear.

Suppose instead that policy functions were posited exogenously: given as primitive elements of the state space, described only by the requirement that they be continuous maps from a force-target space M to a resource-allocation space X. The resulting function space $C(M, X)$ has rich and complicated topology. In general, function spaces are not contractible: they may have non-trivial fundamental groups, higher homotopy groups, or cohomological structure that encodes genuine topological information. Positing a policy function exogenously puts you in $C(M, X)$, where contractibility is a special property that must be assumed, not a consequence that can be proved.

The SPP changes this. The admissibility conditions on τ and κ place $\pi_{\tau,\kappa}$ in a highly structured subclass of $C(M, X)$—the class of solutions to optimization problems with log-concave technologies and exp-convex costs. This subclass's topology is determined by the optimization structure, not by generic properties of function spaces. It is this structure that makes contractibility provable. Berge's maximum theorem (Chapter 4, Lemma 4.2) plays a central role: it establishes

that the policy function varies continuously with the parameters precisely because the production problem is well-posed. Contractibility follows from the continuity properties that the optimization structure guarantees.

The analogy to structural versus reduced-form econometrics is apt. A reduced-form model can approximate the behavioral relationship between inputs and outputs; a structural model derives it from first principles. Only the structural approach tells you what properties the estimated object *must* have by construction. Here: only the structural approach, which derives the policy function from the SPP, tells you that the policy function space must be contractible. Exogenous postulation gives you a policy function; the SPP gives you a policy function with structure.

2.6.4   Breaking Bartelson's circle

The contractibility result breaks Bartelson's circle without contesting his diagnosis.

Bartelson shows that sovereignty-based derivations of "like units" are circular: the like-units claim requires sovereignty, sovereignty requires anarchy, anarchy presupposes like units. Any attempt to ground "like units" via the sovereignty concept inherits this circularity.

The contractibility result is not sovereignty-based. It is derived from mathematical properties of the technology and cost functions τ and κ—from conditions on the state's interior production structure, not from any definitional link between sovereignty and anarchy. The like-units claim is a theorem about the space of classical policy functions. Sovereignty does not appear in the proof.

This *complements* Bartelson. Accept his diagnosis in full: sovereignty is parergonal, philosophically intractable, and not available as a foundation for structural theorizing. The contractibility result stands regardless, because it operates below the parergon. The frame of sovereignty is presupposed (inside and outside must exist for the SPP to be defined), but the content of the frame (the topology of the policy space) is derived independently of how the frame is constituted. Bartelson clears the ground; the SPP builds below it.

But the circle is broken only for the classical regime. The wild states—those whose production problems violate the admissibility conditions—are not reached by the contractibility theorem. Whether the classical model is structurally adequate for the full space of states, wild included, is a question that reopens a version of the problem at a higher level. Chapter 6 confronts it directly.

2.7 Objections

2.7.1 Cartwright: the simulacrum objection

Cartwright (1983) argues that models in physics are simulacra: they do not describe reality as it is but construct a "prepared description" of how reality would be if it conformed to the idealized assumptions of the model. The ideal pendulum is not a claim about actual pendulums—it is a fictional object whose behavior is governed by the simple pendulum equation. On this view, the contractibility of $\mathcal{P}$ is a property of the model, not of reality. States don't have to lie in $\mathcal{P}$ any more than actual pendulums have to satisfy the ideal pendulum equation.

The response turns on the disanalogy between natural and social kinds. For natural kinds—pendulums, fluids, gases—there is a mind-independent physical phenomenon that the model approximates. The ideal pendulum departs from actual pendulums because actual pendulums have friction, air resistance, and non-point masses; the model is a simplification of a prior physical reality. For social kinds, there is no mind-independent physical phenomenon underneath the social kind that the model could be approximating. The unity that makes the president, the guns, and the election into a single state is a normative and relational fact. It is not a physical quantity that the model simplifies.

In the social case, model-relative is all that realism amounts to. The admissibility conditions define the scope of the claim: they specify not an empirical universal (every actual state satisfies these conditions) but the boundary of the class being characterized (a state is classically admissible if its production problem has these properties). The contractibility result is a structural truth about the class of classical states—a truth that follows from the definition of the class plus the mathematics of optimization. Cartwright's simulacrum account applies to models that approximate a prior physical reality; it does not apply to models that constitute a social kind.

2.7.2 Friedman: the instrumentalist challenge

Friedman (1953) argues that economic models should be judged by the accuracy of their predictions, not by the realism of their assumptions. Unrealistic assumptions are not a defect if the model generates good predictions; the methodology of positive economics is predictive, not constitutive. On this view, the topology of $\mathcal{P}$ is neither here nor there: it makes no predictions about state behavior,

so it cannot be evaluated by Friedman's criterion. Why take it as more than a convenient analytical device?

The response is direct: the book makes constitutive claims, not predictive ones. The topology of $\mathcal{P}$ characterizes what kind of thing the state is—the structure of the space of possible states—not what states will do under specified conditions. Constitutive analysis and predictive analysis are different enterprises, governed by different criteria. Friedman's criterion is the right criterion for behavioral models that make empirical predictions. It is not the right criterion for ontological analysis that characterizes the structure of a domain. The methodological divide is prior to Friedman's framework; his criterion cannot assess a claim that does not compete with it.

2.7.3   Hacking: the interactive kinds objection

Hacking (1999) distinguishes indifferent from interactive kinds. Quarks are indifferent: they do not know they are classified as quarks and their behavior is not altered by the classification. Social kinds are typically interactive: when people are classified as members of a kind, they may respond to the classification, changing their behavior and thereby changing the properties of the kind. The looping effect—classification induces behavior change induces reclassification—is endemic to social science.

Applied to the present argument: "state" may be an interactive kind. Rulers read Tilly and reorganize state-building strategies. Statesmen respond to being labeled "failed states" or "predatory states." International institutions create incentives for states to conform to the SPP ideal. If the formalization of $\pi_{\tau,\kappa}$ changes how governments organize force production, which changes π, which changes $\mathcal{P}$, then the space of possible states is itself on the move.

The response operates at two levels.

1. First, the SPP is defined below the level at which looping operates. The technology function τ and cost function κ are constrained by physical, economic, and organizational facts: the physics of force production, the economics of procurement, the logistical constraints of deploying and sustaining coercive capacity. These facts change slowly and are not significantly altered by academic classification. A government does not reorganize its infantry logistics because a political science paper characterized force production as a constrained optimization. The problem structure—$\min_x \kappa(x)$ subject to $\tau(x) = m$—is indifferent in Hacking's

sense even if the state-as-institution is mildly interactive at the level of legitimating discourses and organizational self-descriptions.

2. Second, contractibility provides topological robustness against small perturbations. A contractible space remains contractible under small continuous deformations. If looping effects produce small perturbations in τ and κ— shifting where in $\mathcal{T} \times \mathcal{K}$ a particular state operates—but do not push states outside the admissibility conditions, contractibility of $\mathcal{P}$ is preserved. The looping worry is real for classification schemes that operate at the level of institutional labels and policy discourse; it is attenuated, and topologically manageable, at the structural level of force production technology and cost.

The interactive kinds objection has force against classification schemes that operate at the level of discourse and self-description. It has less force against a formalization whose primitives are constrained by physics, logistics, and organizational capacity—and whose central topological claim is stable under exactly the small perturbations that looping would produce.

2.7.4 The warlord and the non-state actor

Weber's definition of the state requires a *claim to monopoly* of legitimate force within a territory—a relational, normative property that the model abstracts away. The model characterizes all classical force producers: any agent with a log-concave force production technology and exp-convex costs, regardless of whether it claims legitimacy, commands a territory, or is recognized as a state by others. A warlord who can solve the SPP classically would appear in $\mathcal{P}$ alongside a constitutional republic.

Two defensible responses are available, and they can be combined.

1. The first is to *accept the implication* and adjust the interpretation. The contractibility result holds for all classical force producers. Statehood, in the full Weberian sense, requires additional normative conditions— legitimacy, territorial control, international recognition— that the book deliberately brackets. This is a correct division of labor: the book establishes a structural property of the class of classical force producers, of which states are a subclass distinguished by normative properties that require a different analysis.

2. The second is to *argue that the admissibility conditions are implicitly political*. If a force producer's technology exhibits hard indivisibilities, increasing returns to particular input combinations, or organizational instabilities that prevent the production problem from having a unique, continuously varying solution, then it belongs to the wild regime—not by definitional exclusion but because its production problem fails the admissibility conditions. Whether any particular warlord or non-state actor satisfies or violates those conditions is an empirical question about its production structure, not a consequence of its institutional label. But the admissibility conditions are not politically neutral: the organizational achievements that make a production problem well-behaved—smooth diminishing returns, accelerating extraction costs, absence of hard indivisibilities—are precisely the achievements that Levi (1988) and Tilly (1990) associate with the emergence of the bureaucratically organized territorial state.

Both responses leave the contractibility result standing. The second is more interesting: it suggests that classical admissibility is not merely a mathematical condition but a mark of institutional achievement—the organizational form that emerges when competitive pressures drive force producers toward the smooth, well-posed production problems that the admissibility conditions encode.

Coda: Open Questions

This chapter establishes what classical states share as force producers and provides a philosophical warrant for studying the topology of that shared structure. Two questions it leaves open are worth naming.[5]

The first is the phase boundary question. The admissibility conditions define a boundary between the classical regime (contractible, comparable, systematically tractable) and the wild regime (disconnected components, catastrophes, monodromy, topological path dependence). Chapter 6 shows that the classical model is faithful inward—the tame subspace is structurally adequate for the classical space—but lossy outward: the shadow map that projects wild states onto their classical silhouettes destroys the branch structure that makes the wild

---

[5]What force itself *is*, and how individual states are distinguished within the shared structure, are questions this book deliberately brackets. The first is a question about the compositional structure of military capability; the second is a relational question requiring apparatus beyond the topological methods developed here.

topology nontrivial. The asymmetry is structural, not repairable by refinement. What a faithful model of the wild regime would look like—one that preserves its catastrophes, its monodromy, its individual topology—remains genuinely open.

The second is the scope question: how far does the force-production framing extend? The method generalizes to any state function that can be modeled as a constrained optimization with appropriate regularity conditions. The fiscal state (Levi, 1988), the regulatory state, the protection firm (Lane, 1979) all admit analogous production problems. Whether those policy spaces are contractible under their own admissibility conditions, and where their phase boundaries fall, are open questions—and interesting ones, because the shape of possible states is a product of all the dimensions along which states produce.

Chapter 3

The State as a Producer of Force

War and preparation for war involved rulers in extracting the means of war from others who held the essential resources—men, arms, supplies, or money to buy them—and who were reluctant to surrender them without strong pressure or compensation.

Charles Tilly

The previous chapter argued that the state is best understood as a functional unity—defined not by what it is made of but by the problem it solves. This chapter says what that problem is and shows that its parameters encode the deepest political facts about the state.

The model is a constrained optimization: given a force target set by the strategic environment, the state chooses the resource allocation that meets that target at minimum cost. The model has two structural parameters—a technology function τ and a cost function κ—and the central claim of this chapter is that these are not thin mathematical abstractions. The technology encodes the state's organizational capacity for violence: not just what equipment it possesses but how effectively it can convert resources into fighting power—a capacity shaped by regime type, civil-military relations, and doctrinal culture (Biddle, 2004; Talmadge, 2015). The cost function encodes the political economy of extraction: not just fiscal burdens but the institutional infrastructure, political cohesion, and coercion-capital balance that determine what the state can extract and at what price (Besley and Persson, 2011; Tilly, 1990). The math reveals the structural skeleton; the politics clothes it.

The chapter proceeds in six sections. Section 3.1 opens the black box that the arming literature leaves closed. Section 3.2 derives the production problem from

its intellectual sources. Section 3.3 defines the primitives. Sections 3.4 and 3.5 develop the technology and cost functions, justifying the admissibility conditions on material, institutional, and political grounds. Section 3.6 assembles the policy function and defines the space of possible states that the rest of the book will study.

3.1 The Black Box

3.1.1 A motivating game

The international relations literature has a well-developed theory of how states compete militarily—and a conspicuous silence about what military competition actually requires of the state internally. To see the silence, consider a standard model from the strategic arming literature.

3.1 Game

Two states, $i \in \{1, 2\}$, simultaneously choose a force level $m_i \in \mathbb{R}_+$. Their payoffs are given by von Neumann–Morgenstern expected utility functions:

$$U_1(m_1, m_2) = \frac{\lambda m_1^\alpha}{\lambda m_1^\alpha + m_2^\alpha} \times (V - k(m_1 + m_2)),$$

$$U_2(m_1, m_2) = \frac{m_2^\alpha}{\lambda m_1^\alpha + m_2^\alpha} \times (V - k(m_1 + m_2)),$$

where:

1. *$\lambda \in \mathbb{R}_{>0}$ captures the relative effectiveness of the forces;*

2. *$\alpha \in (0, 1]$ captures the decisiveness of superior force;*

3. *$V \in \mathbb{R}_{>0}$ captures the value of the prize; and*

4. *$k \in (0, 1]$ captures the inverse-recuperability of militarization costs.*

The game has a unique Nash equilibrium, given by:

$$\left(m_1^*, m_2^*\right) = \left(\frac{\alpha}{1 + \alpha} \cdot \frac{V}{k} \cdot \frac{\lambda^{-\frac{1}{1+\alpha}}}{1 + \lambda^{-\frac{1}{1+\alpha}}}, \ \frac{\alpha}{1 + \alpha} \cdot \frac{V}{k} \cdot \frac{1}{1 + \lambda^{-\frac{1}{1+\alpha}}} \right),$$

and this solution is continuous in all parameters. [*Proof*.]

Game 3.1 is a Tullock contest with linear costs: each state chooses how much force to field, the probability of winning depends on relative force, and the costs of fielding force are proportional to total militarization. The game has a unique equilibrium in pure strategies, and that equilibrium varies continuously with the parameters of the contest. As a model of strategic arming, it is clean and well-understood.

But notice what the game takes for granted. Each state chooses a force level m_i directly, as if force were a commodity available for purchase. It is not. Force is produced—from steel, labor, oil, training, logistics, institutional capacity—through processes that differ across states and across centuries. The game says nothing about how m_i^* comes into being. It treats the state as a black box: systemic pressure goes in, a force level comes out, and what happens inside is someone else's problem.

3.1.2 Someone else's problem

This silence is not peculiar to Game 3.1. It runs through the strategic arming literature. Authors in this tradition use a wide variety of terms to describe the choice variable m_i, and the variety is itself revealing. Beviá and Corchón (2010) and Hirshleifer (1991) call it "war efforts"—a concept one might measure in francs, battalions, or barrels of oil. Garfinkel (1990) and Powell (1993) call it a "good," suggesting an output of spending that does not accumulate across periods. Meirowitz and Sartori (2008) specifically mention their synonymous use of "capacity" and "arms," saying that what they hope to capture is "any factors that make a state more likely to win a war but are costly to accumulate— for example, a new technology or military strategy."[1] "Arms" is also the term of choice for Skaperdas (1992) and Coe and Vaynman (2020). Hodler and Yektaş (2012) call it "power," while Fearon (2018) settles on "force level," which will be our preferred term.

The proliferation of labels is a symptom. If the literature cannot agree on what m_i *is*, it is because no one has asked how m_i is *made*. None of the models cited above—nor the many others in their lineage—says anything about the internal process by which a state converts resources into coercive capacity. The

---

[1]Meirowitz and Sartori study a more reduced version of the interaction in which military investment happens first, is privately known, and then enters a very general contest function; their focus is on how uncertainty shapes bargaining under incomplete information.

force level appears as a scalar, chosen from a menu, as if producing a hundred tanks were the same kind of decision as choosing a number.

The omission is not merely technical. What lies inside the black box is not an engineering problem but a political one: the organizational capacity to convert resources into force, the extraction bargains that make those resources available, the regime constraints that shape what technologies are feasible and what costs are bearable. The military effectiveness literature has shown that identical equipment produces radically different outcomes depending on the organizational and political context in which it is employed (Biddle, 2004). The state capacity literature has shown that the ability to extract resources is itself a political achievement, endogenous to institutions and regime type (Besley and Persson, 2011). Opening the black box means confronting these political facts directly.

3.2 Question

How is the force level m_i^* produced by State i in models like Game 3.1, and what does the production process tell us about the structure of states?

This chapter answers the first half of the question by constructing a model of state force production. The remainder of the book answers the second half.

3.1.3 The program

We seek a formal object that opens the black box.

3.3 Program

Construct and investigate a map asserting which resources the state will mobilize given:

1. *some specified force level;*

2. *the state's technology for converting resources into force; and*

3. *the state's cost of mobilizing resources.*

The map described in Program 3.3 will be the *policy function*—a complete prescription of how the state mobilizes for any force target, given its technology and cost structure. Before defining it, we need to say where it comes from. The next section derives the production problem from the intellectual traditions that motivate it; the sections that follow define its components precisely.

3.2 From Waltz to the Production Problem

The black box of Game 3.1 exists because the arming literature inherits a division of labor from structural realism. Waltz provides the systemic frame; Tilly provides the internal content; the state-as-firm literature provides the formal template. The production problem that opens the black box is already implicit in these three traditions. Making it explicit is the work of this section.

3.2.1 Waltz's firm analogy

In Chapter 5 of *Theory of International Politics*, Waltz (1979) builds his theory of international structure by analogy with microeconomics. States are to international politics what firms are to markets:

> [T]he parallel with market theory is exact. Both firms and states are like units. Through all of their variations in form, firms share certain qualities: They are self-regarding units that, within governmentally imposed limits, decide for themselves how to cope with their environment and just how to work for their ends. Variation of structure is introduced, not through differences in the character and function of units, but only through distinctions made among them according to their capabilities. (Waltz, 1979, p. 98)

The analogy is productive. It licenses the systemic move that makes structural realism possible: just as a market theorist need not know how each firm is organized internally, a systemic theorist need not know how each state is organized internally. Internal organization drops out of the structural theory.

But the analogy is also incomplete. Market theorists *do* know how firms are organized internally—or rather, they have a powerful theory of the firm's internal organization that they choose to bracket at the systemic level. The theory of the firm in microeconomics (Mas-Colell, Whinston and Green, 1995) begins from a production function (technology) and a cost function, derives the firm's optimal input choices as solutions to a constrained optimization problem, and then studies how those choices respond to prices and market conditions. The systemic theory of market competition can afford to treat the firm as a black box precisely because the theory of the firm has already characterized what is inside the box.

Waltz's structural realism has no such foundation. The "like units" claim is asserted at the systemic level without a corresponding theory of the state's in-

ternal production problem. States are like firms, but unlike firms, their interiors are left unexamined. The omission is substantive, not just formal. In bracketing the interior, Waltz brackets the politics of production—the organizational capacity that determines whether equipment becomes fighting power (Biddle, 2004), the extraction bargains that determine whether resources are available at all (Olson, 1993), and the regime constraints that shape which technologies are feasible (Talmadge, 2015). The present chapter takes the Waltz analogy literally: if states are like firms, model them *as* firms—as producers that transform inputs into outputs subject to technological and cost constraints. The result is a theory of the state's internal organization that can be bracketed at the systemic level, just as the theory of the firm is bracketed in market theory—but that, unlike Waltz's assertion, actually characterizes what "like units" means.

3.2.2   Tilly's content

If Waltz provides the systemic frame, Tilly provides what goes inside it—and what goes inside is politics all the way down.

In *Coercion, Capital, and European States*, Tilly (1990) argues that war drives state formation through a production logic:

> [T]he increasing scale of war and the knitting together of the European state system through commercial, military, and diplomatic interaction eventually gave the war-making advantage to those states that could field standing armies; states having access to a combination of large rural populations, capitalists, and relatively commercialized economies won out. They set the terms of war, and their form of state became the predominant one in Europe. Eventually European states converged on that form: the national state. (Tilly, 1990, p. 15)

Strip this to its formal skeleton and a constrained optimization emerges. The state has *resources*—"men, arms, supplies, or money to buy them"—that serve as inputs. It has a *technology*—the organizational and material apparatus by which those resources are converted into coercive capacity. And it faces *costs*—the "reluctance to surrender [resources] without strong pressure or compensation" that Tilly emphasizes throughout his account of extraction, resistance, and bargaining. The technology τ formalizes what Tilly calls the means of war. The cost function κ formalizes what Tilly calls the price of extraction. The

force target m is set by the systemic contest—the "increasing scale of war" that disciplines state behavior from without.

But Tilly's account does more than license the SPP. It tells us that the parameters of the production problem are *themselves* political objects.

Tilly identifies three paths of state formation, distinguished by the interaction between coercion and capital: *coercion-intensive* states (Russia, Brandenburg-Prussia) build force through direct extraction from subject populations, relying on large standing armies fed by conscription and taxation; *capital-intensive* states (Venice, the Dutch Republic) build force by purchasing it, leveraging commercial wealth to hire mercenaries, fund navies, and finance allies; *capitalized coercion* states (England, France) combine both, developing bureaucratic extraction apparatuses that tap commercial wealth while maintaining coercive capacity.

In the language of the SPP, these are three regions of the parameter space $\mathcal{T} \times \mathcal{K}$. The coercion-intensive state has a labor-heavy technology τ and a cost function κ that is steep in the extraction dimension—it can produce force, but the political price of extraction is high. The capital-intensive state has a capital-heavy technology and a cost function mediated by markets—κ is shaped by credit conditions, interest rates, and the willingness of capitalists to lend. The capitalized coercion state occupies an intermediate position: a technology that combines organizational capacity with material resources, and a cost structure that balances fiscal extraction with commercial intermediation.

Tilly's central empirical claim is that these paths *converge*. Competitive selection—the pressure of interstate war—drives states toward the capitalized coercion form, because states that cannot combine coercive and fiscal capacity are defeated or absorbed. The convergence is not on a single point but on a common organizational type: the bureaucratic national state with the institutional capacity to extract, organize, and deploy force at scale. This is an empirical claim about the topology of the state space: the surviving region of $\mathcal{T} \times \mathcal{K}$ is connected, perhaps contractible, because competitive selection eliminates the states at the periphery. Whether the formal structure of the production problem implies something similar—whether the admissibility conditions alone constrain the topology of the state space—is a question for Chapter 4.

3.2.3 The state-as-firm tradition

The Waltz–Tilly synthesis is not without precedent. A substantial literature has modeled the state as a firm-like entity that produces services from resources at a cost.

Lane (1979) envisioned the state as a firm selling protection in exchange for revenue—a "protection-producing enterprise" whose profits depend on the gap between the cost of providing protection and the tribute extracted for it. North (1981, Ch. 3) extended this logic, modeling the state as a revenue-maximizing organization constrained by transaction costs and the threat of rival providers. Levi (1988) developed the predatory theory of rule: rulers maximize revenue subject to their bargaining power relative to constituents, the transaction costs of taxation, and the discount rate governing their time horizons. Lake (1992) applied a similar production logic to the relationship between regime type and the willingness to use force, arguing that democratic accountability changes the effective cost of military mobilization. And Kier (2021) traced how labor institutions shape military organization, showing that the state's "technology" for producing force is not purely material but institutional—mediated by the political economy of labor relations.

The contemporary culmination of this tradition is Besley and Persson (2011), who model state capacity—both fiscal and legal—as an endogenous investment shaped by political institutions. Their central insight is that the *ability* to extract resources is not given; it is built, and the building is political. States invest in fiscal capacity when political cohesion between competing groups is high enough that both winners and losers benefit from a functioning extraction apparatus, and when external threats create shared incentives for collective investment. When cohesion is low and institutions weak, states do not invest; the result is a "weak state" with a fragile, steep cost structure and limited organizational capacity. In the language of this book, Besley and Persson's weak state is an *wild* state—one whose production problem may fail the admissibility conditions precisely because it has not invested in the institutional infrastructure that makes well-behaved extraction possible.

These accounts differ in what the state produces (protection, revenue, security, order), in what counts as a cost (transaction costs, political opposition, institutional constraints), and in which internal actors matter. But they share a common formal skeleton: the state takes in resources, transforms them through some organizational process, and the character of that process—its technology and its costs—determines what the state can do and at what price. The State Production Problem is their common formal ancestor. What this chapter adds is the recognition that the production problem's parameters are not thin mathematical abstractions: τ encodes the state's sociotechnical capacity for organized violence, and κ encodes the political economy of resource extraction. The math does not simplify away the politics; it reveals the structural skeleton that the

politics clothes.

3.2.4 The production problem

We are now in a position to state the problem precisely.

3.4 Definition

Given a desired force level $m \in M$, a militarization technology $\tau \in \mathcal{T}$, and a cost function $\kappa \in \mathcal{K}$, the state's production problem *is*

$$\min_{x \in X} \kappa(x) \quad \textit{subject to} \quad \tau(x) = m. \qquad\qquad \mathsf{SPP}(m, \tau, \kappa)$$

$\mathsf{SPP}(m, \tau, \kappa)$ says: given a force target m delivered by the systemic contest, find the mobilization plan x that achieves that target at minimum cost. The technology τ determines which mobilization plans can produce a given force level; the cost function κ determines what each plan costs; the state chooses the cheapest feasible plan.

This is formally identical to the cost-minimization problem of a firm in microeconomic theory (Mas-Colell, Whinston and Green, 1995, Ch. 5). The output is force rather than widgets; the inputs are military commodities rather than capital and labor; but the mathematical structure is the same. The Waltz analogy is now literal.

But the analogy should not be read as a reduction to economics. The SPP is not a simplification of what the state does; it is, as the preceding sections have argued, the *deepest* characterization of the state available to formal theory, precisely because τ and κ are politically constituted. The technology function encodes not just material conversion ratios but the organizational capacity that Biddle, Talmadge, and Rosen have shown to be the decisive determinant of military effectiveness. The cost function encodes not just fiscal burdens but the political economy of extraction that Olson, Besley and Persson, and Tilly have shown to be endogenous to institutions and regime type. The SPP is the formal skeleton of the state's defining activity, and its parameters are windows into the state's political character.

Large or small, capitalist or socialist, ancient or modern, democratic or authoritarian—all states that produce force face $\mathsf{SPP}(m, \tau, \kappa)$. They differ in their technologies, in their cost structures, and in the force targets they face; but the *form* of their problem is shared. Whether this shared form implies a

shared *structure* across the space of solutions is the question the rest of the book addresses.

The formulation is also robust to the direction of optimization. The dual problem—given a budget b, maximize $\tau(x)$ subject to $\kappa(x) \leqslant b$—is equivalent under the admissibility conditions developed below.[2] A state that receives a force target from the systemic contest and a state that receives a budget from its fiscal apparatus are solving the same problem, viewed from opposite sides: the systemic perspective (Waltz) delivers the target; the domestic perspective (Tilly) delivers the budget.

We have stated the production problem, but so far we have referred to its components—force, resources, technology, cost—only informally. The remainder of this chapter defines each one precisely and then assembles them into the policy function that is the book's central object of study.

3.3 The Primitives

The production problem involves three kinds of objects: a force level the state wishes to achieve, a set of resources from which force can be produced, and a mobilization plan specifying how much of each resource to deploy. These are the raw materials of the model. The substantive content enters later, through the technology and cost functions that act on them; for now, the task is to fix the domain and codomain.

3.3.1 Force

The first primitive is the state's desired force level. This is the choice variable in Game 3.1, reinterpreted as the *output* of a production process rather than a number chosen from a menu. We let $M := \mathbb{R}_{\geqslant 0}$ denote the set of all possible force levels.

Force is modeled as a scalar.[3]

---

[2]This is the standard duality between cost minimization and output maximization in the theory of the firm; see Mas-Colell, Whinston and Green (1995, Ch. 5).

[3]Carroll (2025) builds a discrete compositional model of force—atomic elements, molecules, configurations, organizational structures—and identifies the conditions under which the resulting space of possible forces reduces to a single numerically valued dimension, a *force scale*. The scalar case used here is the one delivered by that reduction and by the arming literature reviewed in Section 3.1.2, where the equilibrium concept produces a single force target. The

3.3.2 Resources and commodities

Force is made from stuff. We call that stuff *resources*, and we suppose they arrive in different types, which we call *commodities*. We index the set of all commodities by L, where L is nonempty and finite.[4] Each commodity $\ell \in L$ is a good or service that the state can mobilize: steel, oil, labor, uranium, or—should the need arise—fresh-cut flowers. In case $L = 1$, the model collapses to one without cross-commodity trade-offs. The interesting cases arise when $L \geq 2$: the state must decide not only *how much* to mobilize but *what mix* of resources to deploy.

3.3.3 Mobilization plans

The state's decision is how much of each commodity to mobilize. A particular decision is a *mobilization plan*, encoded as a vector $x \in X := \mathbb{R}_{\geq 0}^L$. The ℓth component x_ℓ is the amount of commodity ℓ that the state mobilizes.

One might object that several features likely relevant to this decision—commodity prices, resource endowments, territorial extent—are absent from the model. They are not absent; they are absorbed. A price vector q changes what the cost function looks like: if mobilizing commodity ℓ costs q_ℓ per unit, then the cost function includes a term $q_\ell x_\ell$, and the linear cost function $\kappa(x) = q \cdot x$ is precisely the price-weighted case. An endowment shapes what the technology can do: a state with abundant coal might have a higher marginal productivity of coal in its technology, reflected in larger coefficients on coal-related inputs. Territory influences both channels—through the resource base it makes available and the logistical costs it imposes on mobilization.

The point is that τ and κ are general enough to encode these factors; they are the *medium* through which prices, endowments, and the political economy of extraction enter the model. The examples in Sections 3.4.3 and 3.5.3 will make this concrete.

3.4 Technology

With the primitives in place, we turn to the first of the two functions that give the production problem its content. The technology is, in a sense, the easier of

---

topological results of Chapters 4–5 do not depend on the dimension of M.

[4]L will sometimes refer to the cardinality of the commodity set as well as the set itself; context will disambiguate.

the two to motivate: it says how resources become force. But the admissibility conditions we impose on it are doing more work than they might first appear. They are not merely technical convenience; they formalize substantive commitments about the character of force production, and they determine the scope of the results that follow. They also determine what is left out, and the reader should hold that cost in view throughout.

3.4.1 The definition

The state's technology is the machine that turns resources into force. It converts stone into hatchets, bronze into shields, steel into tanks, and labor into soldiers. Formally, it is a function from mobilization plans to force levels, subject to four structural conditions.

3.5 Definition

The state's militarization technology *is a function*

$$\tau : X \longrightarrow M.$$

We require τ to satisfy:

1. Continuity $(\mathfrak{C}_\tau)$: *τ is continuous;*

2. Ray Surjectivity $(\mathfrak{R}_\tau)$: *there exists a point $v \in X$ such that the map*

$$t \longmapsto \tau(tv) : \mathbb{R}_{\geqslant 0} \longrightarrow M$$

    is continuous, strictly increasing, and unbounded;

3. Weak Monotonicity $(\widetilde{\mathfrak{M}}_\tau)$: *$\tau$ is weakly increasing in all commodities; and*

4. Log-Concavity $(\widetilde{\mathfrak{L}}_\tau)$: *the map*

$$x \longmapsto \log\left(1 + \tau(x)\right)$$

    is concave.[5]

We denote the set of all such functions by $\mathcal{T}$.

---

[5]The term "log-concavity" is used here in a nonstandard way. Ordinarily it refers to functions f such that $\log(f(x))$ is concave. We use $\log(1 + \tau(x))$ to ensure well-definedness at $\tau(x) = 0$. Many a regression-runner has been burned by the logarithm's misbehavior at zero, and nearly all remedy the situation by adding one inside the logarithm—despite all the good statistical reasons not to. It is with a profound sense of solidarity that we follow suit.

3.4.2 What the conditions mean

The four conditions constrain the machine's behavior. Each is a substantive commitment about how force production works, and each can be read as a formalization of something the bellicist tradition already assumes.

Continuity. Small changes in the mobilization plan produce small changes in the force level. There are no magic thresholds—no point at which adding one more soldier causes a discontinuous jump in coercive capacity. This is a smoothness condition, weaker than differentiability, and it rules out production technologies with hard cutoffs or combinatorial phase transitions.

Ray surjectivity. Any desired force level is achievable if the state scales up enough along some resource direction. There exists at least one commodity bundle such that, by investing more and more along that ray, the state can reach arbitrarily high force levels. States are not fundamentally capped; there is always, in principle, a way to produce more force by mobilizing more resources. This does not mean that all rays work—some resource bundles may be useless—only that at least one path of scaling is available.

Weak monotonicity. More of any input does not decrease force. No commodity actively hinders the production of coercive capacity. This is a "do no harm" condition: additional steel cannot make an army weaker, though it may make it no stronger. The condition is weak in the sense that some inputs may have zero marginal product—they are not harmful, merely useless.

Log-concavity. This is the substantive condition, and it deserves the most attention. It says that force production exhibits diminishing returns, in a specific and controlled sense: the function $x \mapsto \log(1 + \tau(x))$ is concave. The hundredth tank adds less force than the tenth. Doubling inputs does not double output; the production technology becomes less efficient at the margin as the scale of mobilization grows.

Log-concavity is weaker than ordinary concavity—it permits some technologies that would violate standard concavity—but it is strong enough to rule out increasing returns, indivisibilities, and combinatorial complementarities. A technology that exhibits increasing returns to scale—where doubling inputs more than doubles output—violates log-concavity. So does a technology with

hard indivisibilities: if producing a tank requires exactly 10 tons of steel and no fewer, the resulting step function is not log-concave.

States whose production technologies violate log-concavity are what we call *wild*. Their production problems may have multiple solutions—distinct mobilization plans that achieve the same force level at the same cost—and the topology of the resulting state space is no longer trivial. Chapter 6 develops the wild case in detail: the fold and cusp catastrophes that arise when log-concavity fails, the rugged landscape of local optima, and the topological path dependence that the classical case rules out. The admissibility condition is not merely a scope limitation but a phase boundary: on one side, a smooth, contractible state space; on the other, a world in which the topology itself generates distinctions and memory.

As Section 3.4.4 will argue, the states that Waltz and Tilly had in mind—bureaucratically organized, territorially rooted, capable of sustained extraction and mobilization—are precisely the states for which log-concavity is a natural description of production.

3.4.3   Examples

The function space $\mathcal{T}$ contains many familiar production functions.

Cobb–Douglas.   The well-known Cobb–Douglas technology

$$\mathcal{T} \ni \tau(x) = A \prod_{\ell \in L} x_\ell^{\beta_\ell},$$

where $A > 0$ and $\beta_\ell \in (0, 1)$ with $\sum_\ell \beta_\ell \leq 1$, satisfies all four conditions on its domain. The exponents β_ℓ capture the relative importance of each commodity: a state whose technology places high weight on labor (large β_{labor}) is labor-intensive in its force production. A state rich in coal but poor in labor would not choose this technology—it would face a different τ—but the *form* of the function is available to any state whose production satisfies the structural conditions.

Constant Elasticity of Substitution (CES).   The CES production function

$$\mathcal{T} \ni \tau(x) = A \left(\sum_{\ell \in L} \gamma_\ell \, x_\ell^\rho \right)^{\sigma/\rho},$$

where $A > 0$, $\gamma_\ell > 0$, $\sigma \in (0, 1]$, and $\rho \leq 1$ with $\rho \neq 0$, is also an element of $\mathcal{T}$. The parameter ρ governs how easily the state can substitute between commodities: as $\rho \to 1$, inputs are perfect substitutes (oil and coal are interchangeable); as $\rho \to -\infty$, they are perfect complements (tanks are useless without fuel). A state that can flexibly substitute between resource types has a different τ from one that cannot, and this difference is encoded in the elasticity parameter.

The CES form also illustrates a subtler point about what the technology encodes. Consider Fisher and Churchill's campaign, beginning in 1904 and culminating in the years before the First World War, to convert the Royal Navy from coal to oil. The argument was not merely that oil was a better fuel. It was that coal-fired ships required enormous crews of stokers—men whose labor was consumed shoveling coal into furnaces rather than fighting. Fisher's case was, in essence, that a man is more useful firing a gun than feeding a boiler: the *cross-partial derivative* $\partial^2 \tau / \partial x_{\text{labor}} \, \partial x_{\text{fuel}}$ depends on the type of fuel. Under coal, labor and fuel are complements in a wasteful sense: more coal requires more stokers, diverting labor from combat. Under oil, the complementarity is broken: fuel is consumed mechanically, and labor is freed for its more productive military uses. The shift from coal to oil is a change in τ—a movement within the technology space $\mathcal{T}$—that altered the cross-partial structure of British force production.[6]

Log-linear (canonical form). We take as canonical the log-linear technology

$$\mathcal{T} \ni \tau(x) = A_\tau \sum_{\ell \in L} \beta_\ell \log\left(1 + x_\ell\right),$$

where $A_\tau > 0$, $\beta_\ell > 0$, and $\sum_\ell \beta_\ell = 1$. This function will play a distinguished role in Chapter 5, where it defines the subclass of *tame* technologies. Its appeal is partly analytical—it yields closed-form solutions to the production problem— and partly structural: it treats each commodity as contributing independently and with diminishing marginal returns, a natural baseline for force production.

---

[6] The conversion also illustrates how endowments shape the *politics* of technology choice. Britain had abundant domestic coal but no oil; the shift to oil created a strategic dependence on Persian supplies that Fisher and Churchill accepted as the price of a more efficient technology. The cost function κ absorbs this dependence: a state whose force production relies on imported inputs faces a different cost structure from one that relies on domestic resources.

Endowment-shaped technology.   Suppose a state's resource endowment is given by $\omega \in \mathbb{R}^L_{>0}$, and that abundant resources are more productive. A technology reflecting this might take the form

$$\tau(x) = \sum_{\ell \in L} \omega_\ell^{1/2} \log\left(1 + x_\ell\right).$$

A coal-rich state (ω_{coal} large) gets more force per unit of coal mobilized than a coal-poor state. The endowment is not a separate primitive; it shapes the technology τ directly. This is one sense in which "invisible parameters" are absorbed rather than omitted.

3.4.4   The political economy of technology

The admissibility conditions on τ are mathematical properties. But the technology function is not a fact about physics; it is a fact about the state's organizational capacity to convert resources into fighting power—and that capacity is politically constituted. This subsection develops the evidence for the conditions along three channels: material, organizational, and political. The channels are not independent—they interact in ways that determine the overall shape of τ—but distinguishing them clarifies what the admissibility conditions require and why classical states satisfy them.

Material capacity and logistics.   The most tangible channel runs through the physical infrastructure of force production: supply chains, training pipelines, arsenals, and the logistical apparatus that connects resources to fighting units. Martin van Creveld's studies of military logistics (2004) and command (1985) emphasize that the expansion of military capacity is a gradual process mediated by institutional routines. Armies do not spring into existence fully formed; they are built up through sustained investment in logistics, training, and material. The conversion of resources into force is continuous, not discontinuous—which is exactly what the continuity condition requires. Geoffrey Parker's thesis of the early modern "military revolution" (1996) reinforces the point from the supply side: technological innovations—gunpowder, fortification, naval power— opened new directions along which states could scale their coercive capacity. Each innovation created a new "ray" in resource space: a new combination of inputs (saltpeter, bronze, skilled labor) along which force could be increased without bound. Ray surjectivity formalizes this: there is always at least one

path of scaling available. And John Keegan's analyses of Agincourt, Waterloo, and the Somme (1978) illustrate the logic of accumulation that underwrites monotonicity: commanders accumulate material advantages—more men, more weapons, more supplies—on the premise that no available resource is actively counterproductive.

But material capacity is only the outer shell of τ. The same equipment in different hands produces radically different results, and the reason lies in the organizational channel.

Organizational capacity and force employment. Biddle (2004) demonstrates that force employment—the organizational ability to implement combined arms tactics, use cover and concealment, suppress enemy fire, and coordinate the movement of dispersed units—matters far more for battlefield outcomes than the quantity or quality of materiel. Iraqi and American forces in 1991 operated much of the same hardware; the difference in outcomes was a difference in τ. Biddle's "modern system" is not a description of what equipment a state possesses but of how effectively it can *use* that equipment—and effectiveness depends on training regimes, flexibility in execution, delegation of authority, and the organizational culture of the military. In the language of the SPP, force employment is the content of τ: the function's value at a given input bundle reflects not the engineering conversion of steel into tanks but the organizational conversion of resources into fighting power.

Posen (1984) identifies the organizational infrastructure through which this conversion is accomplished. Military organizations develop standard operating procedures and programs—routinized ways of combining inputs into combat power—that become institutionalized and resistant to change. A tank unit's hull-down firing procedure, the coordination protocols between infantry and armor, the logistics routines that move supplies to the front: these are the production routines that constitute τ at the operational level.[7] Posen's key insight is that these routines are not neutral engineering artifacts. Military organizations develop routines that serve organizational interests—uncertainty reduction, autonomy from civilian oversight, expansion of size and budget—and these routines persist

---

[7]The distinction between *doctrine* and *production routines* matters. In Carroll (2025), doctrine is a preference relation on available forces—it selects the force target m. The organizational routines that *execute* a given target are the content of τ, not of doctrine. Posen's account of SOPs and programs describes the latter: the operational machinery by which resources become force, not the strategic preference that determines what force to aim for.

long after they have outlived their usefulness. The technology function τ is not a frictionless conversion ratio; it is an organizational artifact, shaped by the political dynamics of the bureaucracy that operates it.

The organizational channel also explains why τ exhibits diminishing returns. Brian Downing's study of the military revolution's political consequences (1993) emphasizes that bureaucratic structures absorb resources with declining marginal efficiency: each additional layer of administration produces less additional military output. Priya Satia's account of the relationship between military and industrial production (2018) reinforces the point: even the most technically advanced states face diminishing returns in converting industrial output into coercive capacity. Log-concavity formalizes these organizational limits. A state whose production technology exhibited *increasing* returns—where doubling inputs more than doubled force—would be a state for which the standard institutional story of bureaucratic capacity and logistical constraint did not apply, and whose production problem would force complete specialization in a single input rather than the diversified mobilization that classical states pursue (Proposition 6.1). Such states are not merely edge cases; they are the regime against which the classical model must ultimately be tested.

Regime type and political constraints. The organizational channel establishes that τ encodes bureaucratic routines. The political channel shows that those routines are constrained by regime type and civil-military relations.

Talmadge (2015) identifies a specific mechanism. Authoritarian rulers who face the risk of military coups respond by *coup-proofing*: centralizing command, rotating officers to prevent the formation of loyal networks, politicizing promotion, and restricting the autonomy of subordinate commanders. These measures are rational from the perspective of regime survival, but they are devastating for military effectiveness. Coup-proofing prevents exactly the delegation, initiative, and coordination that Biddle's modern system requires. The result is a technology function τ that is *constrained by regime type*: a dictator who centralizes command to prevent coups accepts a less efficient τ as the price of staying in power.

Brooks (2008) generalizes the point elegantly. Where Talmadge identifies one mechanism—coup-proofing—by which regime type degrades τ, Brooks identifies the broader structure: civil-military relations shape the entire institutional environment in which the state assesses its own capabilities and coordinates strategy. Two variables—the intensity of preference divergence between political

and military leaders, and the balance of political-military power—interact to produce distinct configurations of information sharing, strategic coordination, and organizational competence. When preferences diverge and power is shared, competition between political and military leaders corrupts the assessment institutions themselves: information is compartmentalized, coordination collapses, and the state may not even perceive its own τ correctly, let alone optimize against it. The technology of force production is not exogenous to politics; it is endogenous to the domestic political environment—and so is the state's ability to perceive and act on its own production parameters.

Rosen (1991) extends the argument from static capacity to dynamic change. Military innovation—the reorganization of force structure, the development of new operational concepts, the adoption of new ways of combining inputs—is not an engineering project but a bureaucratic-political one. Rosen shows that successful innovation requires senior military leaders who can identify and promote junior officers committed to the new way of war, reorganizing career incentives within large, conservative organizations. In the language of the SPP, navigating the technology space $\mathcal{T}$—moving from one τ to another—is itself a political process, constrained by organizational politics and civil-military relations. Where Posen shows that organizational routines resist change from within, Rosen shows that overcoming that resistance requires political action from above.

Synthesis. The three channels are not additive; they interact. A state with weak logistical infrastructure (van Creveld) cannot sustain the organizational routines that Biddle's modern system requires (Posen), which limits the technologies that are feasible, which constrains what political leaders can demand of the military (Brooks), which shapes the innovation pathways that are available (Rosen). The admissibility conditions on τ—continuity, ray surjectivity, weak monotonicity, log-concavity—are conditions on *organized* force production: the kind that only bureaucratically coherent states, with functioning logistics, settled organizational routines, and manageable civil-military relations, can sustain. A state whose political circumstances prevent it from maintaining coherent organizational capacity is a state whose τ may fail these conditions—not because of an engineering failure but because of a political one. The consequence is not merely that the model ceases to apply but that the structure of the production problem changes qualitatively: multiple solutions, discontinuous parameter dependence, and a topology that is no longer contractible (Chapter 6). This

is worth being honest about: the admissibility conditions purchase the clean results of Chapters 4 and 5, but they purchase them by restricting attention to the well-behaved part of the landscape. Whether that restriction is structurally adequate—whether the classical whiteboard faithfully represents the room that includes the wild states—is a question the book must eventually confront.

3.5 Cost

The technology tells the state what it *can* produce. The cost function tells it what production *costs*—and cost, in this model, is construed broadly. It includes fiscal burdens, but also political strain, social disruption, and organizational friction. The diversity of what counts as a cost is precisely why the cost function is defined abstractly: it must be capacious enough to accommodate channels as different as tax revenue and conscription resistance within a single formal object.

3.5.1 The definition

Formally, the cost function maps mobilization plans to a scalar measure of cost.

3.6 Definition

The state's cost function *is a function*

$$\kappa : X \longrightarrow \mathbb{R}.$$

We require κ to satisfy:

1. Continuity $(\mathfrak{C}_\kappa)$: κ *is continuous;*

2. Centeredness $(\mathfrak{o}_\kappa)$: $\kappa(0) = 0;$

3. Coerciveness $(\mathfrak{D}_\kappa)$: $\kappa(x) \to \infty$ *as* $\|x\| \to \infty;$

4. Strict Monotonicity $(\mathfrak{M}_\kappa)$: κ *is strictly increasing in all commodities; and*

5. Strict Exp-Convexity $(\mathfrak{L}_\kappa)$: *the map*

$$x \longmapsto \exp\left(\kappa(x)\right)$$

   is strictly convex.

We denote the set of all such functions by $\mathcal{K}$.

3.5.2   What the conditions mean

The five conditions constrain how the state experiences the burden of mobilization. As with the technology conditions, each encodes a substantive commitment about the political economy of extraction.

Continuity.   Small changes in the mobilization plan produce small changes in cost. There is no threshold at which mobilizing one more barrel of oil triggers a discontinuous jump in the political, fiscal, or social price of extraction.

Centeredness.   Doing nothing costs nothing: a normalization that sets the zero point of the cost scale at the status quo ante.

Coerciveness.   Costs eventually become prohibitive. As the state attempts to mobilize more and more resources, the cost grows without bound. No state can mobilize infinite resources at finite cost. This is a boundary condition that ensures the production problem has well-behaved solutions: without it, the cost-minimizing mobilization plan might not exist.

Strict monotonicity.   Every additional unit of every resource costs something. There are no free inputs. Unlike the weak monotonicity of technology—where some inputs may be useless—the strict monotonicity of cost says that mobilizing *any* commodity, however marginal, increases the burden on the state. This is a strong condition, but a natural one: in a world of scarcity, extracting resources always has a price, even if that price is political rather than fiscal.

Exp-convexity.   This is the key condition on the cost side, and it deserves careful statement. Exp-convexity says that the function $x \mapsto \exp(\kappa(x))$ is strictly convex. This is *weaker* than requiring κ itself to be convex. It permits cost functions that are strictly concave—that exhibit economies of scale in a standard sense—provided they are not "too concave." The precise threshold is the logarithm: a cost function can be as concave as log but no more so.

What does this mean substantively? It means that costs accelerate, in the sense that the marginal cost of mobilization grows fast enough to prevent the state from exploiting scale economies indefinitely. A state that could mobilize resources at constant or decreasing marginal cost—that faced no political, fiscal, or organizational check on extraction—would violate exp-convexity. The condition

says: at some point, extracting more becomes harder. The political resistance that Tilly emphasizes, the fiscal strain that Cappella Zielinski (2016) documents, the social disruption that accompanies mass mobilization—all contribute to a cost structure that accelerates with scale.

3.5.3 Examples

Linear cost (price-weighted). The simplest cost function is the linear one:

$$\mathcal{K} \ni \kappa(x) = q \cdot x = \sum_{\ell \in L} q_\ell \, x_\ell,$$

where $q \in \mathbb{R}^L_{>0}$ is a vector of commodity prices. This function satisfies all five conditions. The linear case makes explicit how prices enter the model: they are not invisible parameters but components of the cost function. A state facing high oil prices (q_{oil} large) has a different cost function from one facing low oil prices, and this difference shapes its optimal mobilization plan.

Quadratic cost (interaction effects). The quadratic cost function

$$\mathcal{K} \ni \kappa(x) = x^\top Q \, x,$$

where $Q \in \mathbb{R}^{L \times L}$ is a positive definite matrix, captures interactions between commodities. The off-diagonal entries of Q encode cross-costs: mobilizing steel and labor simultaneously may be more expensive than mobilizing either alone if the same infrastructure (railways, ports) is strained by both. This function satisfies all five conditions.

Concave but exp-convex. The function

$$\mathcal{K} \ni \kappa(x) = \sum_{\ell \in L} \beta_\ell \left(1 - e^{-x_\ell} \right),$$

where $\beta_\ell > 0$, is strictly concave—it exhibits economies of scale in the standard sense. Yet it satisfies exp-convexity because $\exp(\kappa(x))$ is still strictly convex. This illustrates the permissiveness of the condition: costs can exhibit decreasing marginal cost at moderate scales, as long as they do not decrease too fast.

Price-augmented cost. Suppose a state's cost depends on both the quantity mobilized and the prevailing price level, with an additional convex friction term reflecting organizational strain:

$$\kappa(x) = \sum_{\ell \in L} p_\ell \, x_\ell + \gamma \|x\|^2,$$

where $p_\ell > 0$ is the price of commodity ℓ and $\gamma > 0$ captures the rate at which organizational costs accelerate. At small scales, costs are approximately linear in prices; at large scales, the quadratic term dominates and costs grow sharply. This is a natural model for a state whose fiscal costs of mobilization are price-determined but whose political and organizational costs are convex in scale. The prices are not omitted from the model; they are written into the cost function directly.

3.5.4 The political economy of cost

The admissibility conditions on κ are mathematical properties. But the cost function is not a fact about nature; it is a fact about the state's political economy. Costs are not merely experienced; they are *produced* by political institutions, shaped by regime type, and endogenous to the state's own history of institutional investment. This subsection develops the evidence for the conditions along three channels: fiscal, political, and organizational. The channels are not independent—they interact in ways that determine the overall shape of κ—but distinguishing them clarifies what the admissibility conditions require and why classical states satisfy them.

Fiscal costs and rising marginal extraction. The most visible channel runs through war finance. States pay for force through taxation, borrowing, and monetary expansion, and each channel imposes costs that rise with scale. Tax revenues face political limits; borrowing increases debt service and default risk; printing money generates inflation (Cappella Zielinski, 2016). International credit markets reinforce the pattern: Siegel (2014) documents how borrowing costs accelerated for belligerents in the early twentieth century as lenders priced in sovereign risk; Stasavage (2011) shows that representative institutions lowered borrowing costs by serving as credible commitment devices for repayment; and Flores-Macías and Kreps (2015) and Slantchev (2012) confirm that war finance imposes fiscal costs that grow with the scale of mobilization. The cost escalation

is not merely fiscal but macroeconomic. Ramey and Shapiro (1998) show that military buildups force costly reallocation of sector-specific capital, depressing consumption, wages, and manufacturing productivity; and Federle et al. (2026), analyzing 150 years of data across sixty countries, find that consumer prices rise approximately twenty percent in the war-site economy, interest rates climb, and output falls nearly ten percent—effects that compound the direct fiscal burden and make each additional unit of mobilization costlier than the last. These are empirical manifestations of the coerciveness and exp-convexity conditions: costs eventually become prohibitive, and they accelerate.

But the *ability* to extract at all is not given; it is built. Olson (1993) draws the fundamental distinction: a roving bandit extracts everything and invests nothing; a stationary bandit develops an "encompassing interest" in the population's productivity and invests in institutions—tax administration, property rights, public goods—that reduce the cost of extraction. Besley and Persson (2011) formalize this as a model of endogenous fiscal capacity, showing that investment in the extraction apparatus depends on political cohesion between competing groups (θ), external threats (ϕ), and resource rents (R). When cohesion is high, both groups benefit from investing in fiscal capacity, and the resulting κ is smooth, well-behaved, and satisfies the admissibility conditions. When institutions are weak and cohesion absent, the state does not invest; fiscal capacity remains underdeveloped, and κ is steep, fragile, and potentially non-classical. The boundary between classical and wild states tracks this political boundary: Besley and Persson's "common-interest" state is classical; their "weak state" is wild.

Tilly (1990) historicizes the same point. States with access to concentrated capital could finance force production through borrowing and taxation of commercial wealth, yielding a cost function mediated by credit markets. States in capital-poor regions relied on direct coercion—forced extraction of grain, labor, and conscripts—and the coercive cost function is steep, politically volatile, and organizationally crude. Tilly's convergence on the capitalized coercion form is, in cost function terms, a convergence on a κ that combines fiscal sophistication with coercive reach: costs that are manageable at scale because the extraction apparatus has been built through centuries of institutional investment.

Political costs of manpower. The conscription literature documents a specifically political channel of rising costs—and one that interacts with the technology in revealing ways. Conscription strains political community (Leander, 2004),

interacts with inequality and partisanship to generate costs that are convex in scale (Kriner and Shen, 2016), and shapes the willingness to mobilize through the anticipated political cost of casualties (Horowitz and Levendusky, 2011). Levi (1996) treats conscription itself as an institution whose costs depend on the bargaining power of the state relative to its constituents—a point that connects directly to the Olson–Besley-Persson logic of endogenous extraction capacity. The costs of manpower are not only political but structural: Broadberry and Harrison (2005) show that in World War I, mobilization capacity rose sharply with economic development, and that poor agrarian states faced catastrophic costs as conscription stripped labor from subsistence agriculture, collapsing food production and triggering the urban famines that brought down the Russian and Ottoman empires.

The sensitivity of κ to labor—the partial derivative $\partial\kappa/\partial x_{\text{labor}}$—is not merely a mathematical object. It encodes the political economy of manpower. A state with high $\partial\kappa/\partial x_{\text{labor}}$ is one for which the politics of manpower are acute, and its optimal mobilization plan will substitute away from labor toward material and capital wherever the technology permits. The interaction between τ and κ is where doctrine lives: changing the technology changes which inputs are needed, which changes the political costs of extraction, which feeds back into the technology that is feasible.[8]

Organizational and agency costs. Not all costs of force production flow through the fiscal system or the politics of extraction. Feaver (2003) identifies a channel specific to the production problem itself: civil-military relations generate agency costs. The political leadership delegates the use of force to the military, and the resulting monitoring, contracting, and occasional shirking impose costs that enter κ. A state in which the military resists oversight or the political leadership distrusts the officer corps faces a higher κ than one where the civil-military bargain is settled. Military activity also generates costs beyond the organizational: environmental costs (Crawford, 2022) and the restructuring of gendered divisions of labor and care (Enloe, 2000) accumulate with the scale of mobilization. The cost function aggregates all of these channels—fiscal, political, organizational, social—into a single scalar. As with the technology conditions,

---

[8]This is why the Fisher–Churchill conversion of the Royal Navy from coal to oil, discussed in Section 3.4.3, was simultaneously a technological and a political-economic decision. The new technology changed the cross-partial structure of τ, which changed the optimal mobilization plan, which changed the demands that force production made on the labor dimension of κ.

the admissibility conditions on κ are conditions on *organized* extraction—the kind that only states with functioning institutions and invested fiscal capacity can sustain. When these conditions fail, the production problem crosses the classical–wild boundary: the cost landscape becomes rugged, optimal mobilization plans bifurcate, and the state must navigate a topologically nontrivial space (Chapter 6).

3.6 The Policy Function and the Space of States

The technology and cost function define the state's environment. The production problem defines what the state does in that environment. What remains is to assemble these into the object we are ultimately after: a single function that encodes, for a given state, the complete relationship between what the system demands and how the state responds.

3.6.1 From problem to function

Solving SPP (m, τ, κ) for a single force target m tells us how the state mobilizes for that particular demand. But the systemic contest does not deliver a fixed m; it delivers a force target that depends on the strategic environment, which changes over time and across interactions. We therefore need to solve SPP (m, τ, κ) not for one m but for *all* $m \in M$. This yields a function.

3.7 Definition

Given a militarization technology $\tau \in \mathcal{T}$ and a cost function $\kappa \in \mathcal{K}$, the state's policy function *is the map*

$$\pi_{\tau,\kappa} : M \longrightarrow X$$

defined by

$$\pi_{\tau,\kappa}(m) \in \operatorname*{argmin}_{x \in X} \kappa(x) \quad \textit{subject to} \quad \tau(x) = m.$$

The set of all such functions is denoted $\mathcal{P}_{\mathcal{T} \times \mathcal{K}} := \{ \pi_{\tau,\kappa} : \tau \in \mathcal{T},\ \kappa \in \mathcal{K} \}$.

The move from a problem to a function deserves scrutiny, because it is doing more than it appears.

Start with the raw optimization. The solution to $\mathrm{SPP}\,(m, \tau, \kappa)$ is, in the first instance, a function of three arguments: the force target m, the technology τ, and the cost function κ. Call it $\Pi(m, \tau, \kappa)$. This is a single, large map from $M \times \mathcal{T} \times \mathcal{K}$ into X. It tells us, for any combination of demand, technology, and cost, how the state mobilizes. But it does not yet give us a *state*. It gives us an answer to a particular question about a particular state in particular circumstances.

The key step is to *curry* this map: to fix (τ, κ) and let m vary, producing a function $\pi_{\tau, \kappa} : M \to X$ that takes force targets as input and returns mobilization plans as output.[9] This is not merely a notational rearrangement. The uncurried map Π describes the solution to a parameterized family of optimization problems. The curried map $\pi_{\tau, \kappa}$ describes a *behavioral profile*: the complete specification of how one particular state—the state characterized by (τ, κ)—responds to every possible demand the system might place on it. The difference is the difference between knowing what a firm produces when facing a particular price and knowing the firm's entire supply curve.

In microeconomics, this is exactly the move that takes us from the firm's cost-minimization problem to its cost function $c(y) = \min_x w \cdot x$ subject to $f(x) \geq y$, where y is the output level and w the input prices.[10] Waltz said states are like firms. Here is where the analogy becomes literal: the policy function is the state's conditional factor demand, and the space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is the space of all possible supply-side characters a state can have.

The currying move also clarifies what the policy function *hides*. Once we have $\pi_{\tau, \kappa}$, the technology and cost structure are absorbed into the function; they are no longer visible as separate objects. Two states with very different technologies and cost structures could, in principle, produce the same policy function—if the differences happen to cancel in the optimization. This is a feature, not a defect. The policy function captures what the state *does*, not why it does it. It is behaviorally complete and mechanistically silent: given any force target, it tells us the mobilization plan, but it does not tell us whether the state achieves that plan through efficient technology or through low costs. In the language of Chapter 2, the policy function is the functional unity itself, not the parts that

---

[9]Currying is named for logician Haskell Curry, though the idea dates to Frege and Church. Formally, it is the canonical isomorphism $\mathrm{Hom}(A \times B, C) \cong \mathrm{Hom}(A, C^B)$ applied to function spaces.

[10]See Mas-Colell, Whinston and Green (1995, Ch. 5). The policy function $\pi_{\tau, \kappa}$ is the conditional factor demand; it plays the same role as the firm's cost-minimizing input bundle $x(w, y)$ viewed as a function of output y alone, with input prices absorbed into the cost function.

compose it.

Finally, currying is what makes the space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ possible as a mathematical object. Without it, we would have a parameterized family of solutions—one for each (m, τ, κ)—but no natural way to compare states, because no single object would represent a state. With it, each state is a *point* in a function space, and we can ask topological questions about the collection of such points: Is the space connected? Are there isolated clusters? Can every state be continuously deformed into every other? These are the questions the next chapter takes up.

3.6.2   A worked example

To see the policy function in action, consider a state with $L = 2$ commodities—call them steel and labor—equipped with the canonical log-linear technology and a linear cost function:

$$\tau(x) = \tfrac{2}{3} \log(1 + x_s) + \tfrac{1}{3} \log(1 + x_l), \qquad \kappa(x) = x_s + x_l.$$

Steel is twice as productive as labor ($\beta_s = \tfrac{2}{3}, \beta_l = \tfrac{1}{3}$), but both commodities cost the same per unit ($q_s = q_l = 1$). This is a material-intensive technology: steel contributes more to force production per unit mobilized than labor does.

To solve $\mathsf{SPP}\,(m, \tau, \kappa)$, form the Lagrangian and take first-order conditions. The interior solution satisfies

$$q_\ell = \frac{\mu \beta_\ell}{1 + x_\ell} \qquad \Longrightarrow \qquad x_\ell = \frac{\mu \beta_\ell}{q_\ell} - 1,$$

where μ is the Lagrange multiplier. A commodity ℓ activates—enters the mobilization plan—only when μ is large enough that $\mu \beta_\ell / q_\ell \geq 1$, i.e., when the force target is high enough that the state finds it worthwhile to pay the cost. The activation threshold is $\mu_\ell^* = q_\ell / \beta_\ell$: $\mu_s^* = \tfrac{3}{2}$ for steel, $\mu_l^* = 3$ for labor. Steel, being more productive per dollar, activates first.

The policy function therefore has two phases.

1. *Low demand* ($m \leq \tfrac{2}{3} \log 2$). Only steel is active. The constraint $\tfrac{2}{3} \log(1 + x_s) = m$ gives

   $$\pi(m) = \left(e^{3m/2} - 1,\ 0\right).$$

   The state mobilizes steel alone; labor's marginal product is too low relative to its cost.

2. *High demand* ($m > \frac{2}{3} \log 2$). Both commodities are active. Substituting the first-order conditions into the constraint and solving yields

$$\pi(m) = \left(2^{1/3} e^m - 1, \ 2^{-2/3} e^m - 1\right).$$

Both components grow exponentially in m, but the state mobilizes twice as much steel as labor: the ratio $x_s/x_l \to \beta_s/\beta_l = 2$ as $m \to \infty$. The asymptotic ratio reflects the productivity advantage—not the cost, which is equal here.

Several features of this example are general. First, the *water-filling structure*: commodities activate sequentially as the force target rises, ordered by their productivity-to-cost ratio β_ℓ/q_ℓ. What matters is not cost alone, nor productivity alone, but the ratio. A commodity that is expensive but highly productive may still activate before one that is cheap but marginal—and the activation order determines the qualitative character of the state's mobilization at low force levels. Second, *exponential scaling*: the interaction between diminishing returns in τ and linear costs in κ means that resource requirements grow exponentially in the force target.

Now change the technology. Suppose a technological shift—an investment in training, a doctrinal reform, a reorganization of logistics—raises the productivity of labor to match steel: $\beta_s = \beta_l = \frac{1}{2}$. The activation thresholds equalize, both commodities enter simultaneously, and the policy function becomes

$$\pi(m) = \left(e^m - 1, \ e^m - 1\right)$$

for all $m > 0$—the state mobilizes steel and labor in equal proportions. This is a *different state*: a different element of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, generated by a different technology but the same cost function. The shift in τ changed the state's character even though its cost structure remained the same.

Now change the cost function instead. Keep the equalized technology ($\beta_s = \beta_l = \frac{1}{2}$) but suppose the political cost of mobilizing labor triples: $q_s = 1$, $q_l = 3$. This is a state for which conscription is politically expensive—not because soldiers are paid well, but because extracting men from families, farms, and factories generates political resistance that the state must absorb. The activation thresholds are now $\mu_s^* = q_s/\beta_s = 2$ and $\mu_l^* = q_l/\beta_l = 6$: steel activates first again, but this time because of cost, not productivity.

The policy function acquires the same two-phase structure as before, but with different content:

1. *Low demand* ($m \leq \frac{1}{2} \log 3$). Only steel is active:

$$\pi(m) = \left(e^{2m} - 1, \; 0\right).$$

   The state avoids the expensive commodity entirely.

2. *High demand* ($m > \frac{1}{2} \log 3$). Both commodities are active. The first-order conditions and the constraint yield

$$\pi(m) = \left(\sqrt{3}\, e^{m} - 1, \; \tfrac{1}{\sqrt{3}}\, e^{m} - 1\right).$$

   The state mobilizes three times as much steel as labor asymptotically: $x_s/x_l \to q_l/q_s = 3$. The ratio now reflects the *cost* advantage, not the productivity advantage—the state substitutes away from the politically expensive input wherever the technology permits.

The three policy functions—the original material-intensive state, the post-reform equalized state, and the conscription-costly state—are three points in $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$. The first two differ in τ alone; the second and third differ in κ alone. The example illustrates concretely what the preceding sections argued abstractly: movements in technology space and movements in cost space both change the state's character, and the political content of the parameters—organizational capacity, extraction costs, regime constraints—is what gives those movements substantive meaning. Whether these three points are "close" to each other—and what closeness means in the space of all possible states—is a question about the topology of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$.

3.6.3 The space of possible states

The set $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is the *space of possible states*. Each element is a policy function generated by some admissible technology–cost pair; the collection of all such functions, across all of $\mathcal{T}$ and all of $\mathcal{K}$, is the main object of this book.

$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is, by construction, a very large set. It includes every possible way of converting resources into force that satisfies the admissibility conditions, paired with every possible cost structure that satisfies its own conditions. A 15th-century gunpowder state, a 19th-century industrial power, and a 21st-century technologically advanced military all correspond to elements of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, provided their production technologies and cost structures satisfy the structural conditions laid out above. The space is agnostic about regime type, ideology, geography,

or historical period; it cares only about the structure of force production. But "the structure of force production" is not a narrow concern. As the preceding sections have argued, each point in $\mathcal{T} \times \mathcal{K}$ encodes a political economy: a regime's organizational capacity for violence (shaped by civil-military relations, coup-proofing constraints, doctrinal culture), paired with its institutional capacity for extraction (shaped by fiscal investment, political cohesion, the coercion-capital balance). The space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is not an abstract mathematical object inhabited by dimensionless points; it is the space of all possible political economies of force production. The topology of this space is a claim about the structural relationships among those political economies.

States do many things other than solve SPP(m, τ, κ). They govern, tax, legislate, police, adjudicate, regulate, and build. But these other activities are not the focus of this book, and the space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ captures exactly the dimension of state behavior that the bellicist tradition and the strategic arming literature care about. Whether from above—as in Waltz's structural theory—or from below—as in Tilly's bellicist account—force production is the dimension along which the state's character is defined. In studying $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, we study that character in its full generality.

3.6.4 Connection to the ontological argument

It is worth pausing to connect the formal object we have constructed to the philosophical argument of Chapter 2. There, we argued that the state is not a physical object but a functional pattern: the organizing relation that maps external demand to internal response. The president, the guns, and the election are real; their unity—the fact that makes them components of a single entity—is a normative and relational fact. The policy function $\pi_{\tau, \kappa}$ is that fact, formalized. It does not model the president or the guns; it models the functional unity that makes them cohere as a force-producing organization.

This chapter has enriched that claim. The ontology chapter argued that $\pi_{\tau, \kappa}$ is a functional unity; this chapter has shown that it is a *politically constituted* functional unity. The technology and cost functions that generate the policy function are not thin mathematical parameters; they encode organizational capacity, regime constraints, extraction bargains, and civil-military relations. The function that maps external demand to internal response carries within it the political economy of the state's defining activity. To study the topology of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is not to study an abstract mathematical space; it is to study the structural relationships among all possible ways a state can organize its most

fundamental political task.

$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is the space of all possible such unities—all possible ways of being a state, in the sense of being a classical force producer. The space is now defined. We are in a position to ask the book's central question: what is its shape?

Chapter 4

The Classical Topology

The parallel with market theory is exact. Both firms and states are like units.

Kenneth Waltz

The previous chapter defined the space of possible states $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ and closed with a question: what is its shape? This chapter answers it. Under the admissibility conditions, the space is contractible—it can be continuously shrunk to a single point without tearing or puncturing. This is the strongest possible form of topological simplicity: all states are topologically the same, all paths between them are deformable into each other, and the space carries no intrinsic memory. The interpretive consequence is sharp: Waltz is right. States are like units—not as a metaphor, but as a topological theorem.

The chapter proceeds in five sections. Section 4.1 establishes that the classical production problem is well-posed: it has a unique solution that varies continuously with all parameters. Section 4.2 examines the shape of the parameter spaces $\mathcal{T}$ and $\mathcal{K}$ and finds that both are contractible—a fact whose proof reveals the paths along which institutional change can operate. Section 4.3 states the main result: $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is contractible. Section 4.4 develops the interpretive consequences through a *ladder of sameness*—three ascending rungs of equivalence among states, each revealing a deeper layer of structural uniformity. Section 4.5 assesses what the topology has established and identifies what it leaves open.

4.1 Well-Posedness

When faced with an optimization problem like SPP (m, τ, κ), three questions naturally arise. The first two concern whether the problem is well-posed: *existence*—does it admit at least one solution?—and *uniqueness*—if it does, is the solution unique? If the problem admits no solution, the state cannot achieve its force target. If it admits multiple solutions, the state faces a choice that the optimization problem itself cannot resolve: a limit on the model's precision.

Both concerns are resolved by the admissibility conditions.

4.1 Lemma

For all $(m, \tau, \kappa) \in M \times \mathcal{T} \times \mathcal{K}$, SPP (m, τ, κ) admits a unique solution. [*Proof*.]

The third question is more capacious: *stability*—how does the solution change as the parameters change? This is where the analysis becomes interesting, because two of the three parameters in SPP (m, τ, κ) are themselves functions. The force level $m \in M = \mathbb{R}_{\geq 0}$ is a scalar, and questions about its effect on the solution are standard comparative statics. But the technology $\tau \in \mathcal{T}$ and cost function $\kappa \in \mathcal{K}$ are elements of infinite-dimensional function spaces. To ask whether the solution changes continuously as these parameters change, we need to equip $\mathcal{T}$ and $\mathcal{K}$ with topological structure.

The natural choice is the compact-open topology, which measures how close two functions are by how similarly they behave on compact subsets of their domain.[1] Under this topology, the solution behaves well.

4.2 Lemma

The solution to SPP (m, τ, κ) varies continuously with m, τ, and κ. [*Proof*.]

This is the first step toward understanding the structure of the state's production problem: the solution responds continuously to changes in both the environment and the state's own parameters.

Moving toward comparative statics, we observe that the policy function $\pi_{\tau, \kappa}$ defined in Definition 3.7—the curried map from force targets to mobilization plans—is itself continuous as a function of the parameters.

---

[1]Under the compact-open topology, two technologies τ_0 and τ_1 are close if they send all mobilization plans within any bounded region to similar force levels. The formal construction and its properties are developed in the appendix.

4.3 Corollary

The policy function $\pi_{\tau,\kappa} : M \to X$ varies continuously with τ and κ. [*Proof*.]

Not only does a single solution vary continuously; the entire behavioral profile of the state—its response to every possible force demand—varies continuously as its technology and cost structure change.

Finally, the policy function inherits structural properties from the primitives.

4.4 Lemma

The policy function $\pi_{\tau,\kappa} : M \to X$ satisfies:

1. Centeredness $(\mathfrak{o}_\pi)$: $\pi_{\tau,\kappa}(0) = 0$;

2. Coerciveness $(\mathfrak{O}_\pi)$: $\lim_{m \to \infty} \|\pi_{\tau,\kappa}(m)\| = \infty$; *and*

3. Weak Monotonicity $(\widetilde{\mathfrak{M}}_\pi)$: $m_1 \leqslant m_2 \implies \pi_{\tau,\kappa}(m_1) \leqslant \pi_{\tau,\kappa}(m_2)$ *componentwise.*

[*Proof*.]

Zero demand elicits zero mobilization. Unbounded demand elicits unbounded mobilization. Higher demand never reduces mobilization in any commodity. These properties are economically natural and mathematically essential: they ensure that the policy function behaves in ways that make the subsequent topological analysis tractable.

With the well-posedness of the problem established, we turn to the structure of the spaces that generate it.

4.2 The Shape of the Parameter Spaces

$\mathcal{T}$ and $\mathcal{K}$ are large and diverse sets of functions. The technology space includes everything from Cobb–Douglas production functions to log-linear technologies to CES specifications with arbitrary substitution elasticities; the cost space includes everything from linear price-weighted costs to quadratic interaction models to concave-but-exp-convex cost structures. What, if anything, do the elements of these spaces have in common?

The answer is: more than one might expect. Despite their diversity, both spaces are *contractible*—they can each be continuously shrunk to a single point. In terms of their global topological structure, they are as simple as a space can be.

4.2.1 What contractibility means

Contractibility is the strongest form of topological triviality. A space is contractible if it can be continuously deformed to a single point without tearing, cutting, or gluing.

Consider a simple example: the interval $[-1, 1]$. Imagine pinching it from both ends, pulling the endpoints toward the center until they meet at 0. This is a continuous deformation that shrinks the entire interval to a single point; the interval is contractible. A circle, by contrast, is not contractible: any attempt to shrink it to a point must pass through the hole it encloses, and no continuous deformation can close a hole. The interval and the circle are topologically different, and the difference is precisely contractibility.
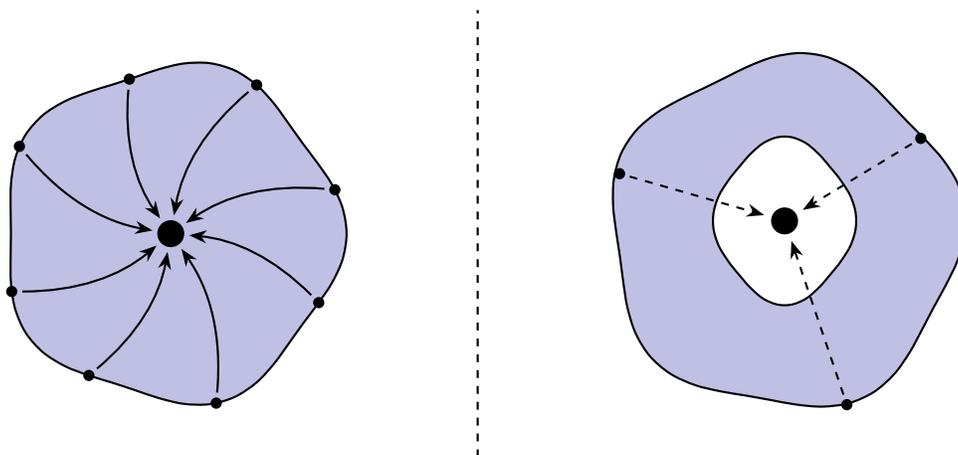
Figure 4.1: *Contractible (left) versus non-contractible (right) spaces. The blob at left can be continuously shrunk to any interior point. The annulus at right cannot: the hole is a topological obstruction.*

The question we now ask about $\mathcal{T}$ and $\mathcal{K}$ is: are they contractible? Can they be represented by a single point, or do they possess holes, disconnected components, or other structure that survives continuous deformation?

4.2.2 The homotopies and what they say

Both spaces are contractible, and the homotopies that prove it are interpretively rich.

4.5 Lemma

The function spaces $\mathcal{T}$ and $\mathcal{K}$ are contractible. [*Proof*.]

The proof constructs explicit homotopies—continuous paths that deform each space to a single point—and these homotopies deserve attention in their own right, because they reveal the structure of institutional change within each space.

Technology. The homotopy on $\mathcal{T}$ takes the form

$$H_\tau(\tau, t) = (1 + \tau)^{1-t} \cdot (1 + \tau_0)^t - 1,$$

where $\tau_0(x) = \sum_\ell \log(1 + x_\ell)$ is the symmetric log-linear technology. At $t = 0$, $H_\tau(\tau, 0) = \tau$: the original technology. At $t = 1$, $H_\tau(\tau, 1) = \tau_0$: the target. At each intermediate t, the function $H_\tau(\tau, t)$ is itself an admissible technology—it satisfies all four conditions in Definition 3.5. The key is that log-concavity is preserved under the geometric mean:

$$\log(1 + H_\tau(\tau, t)) = (1 - t) \log(1 + \tau) + t \log(1 + \tau_0),$$

which is a convex combination of concave functions and therefore concave.[2]

What does this homotopy *mean*? It says that any technology can be continuously deformed into any other through a path of admissible intermediate technologies. Each step along the path is itself a valid way of converting resources into force. A state whose technology is labor-intensive can be smoothly transformed into one whose technology is capital-intensive; a state with Cobb–Douglas complementarities can be deformed into one with CES substitutabilities. The deformation preserves admissibility at every step: no intermediate technology violates the structural conditions that define $\mathcal{T}$.

Cost. The homotopy on $\mathcal{K}$ takes the dual form

$$H_\kappa(\kappa, t) = \log((1 - t) \exp(\kappa) + t \exp(\kappa_0)),$$

---

[2]Neither $\mathcal{T}$ nor $\mathcal{K}$ is a vector space: log-concavity and exp-convexity are not closed under addition or scalar multiplication. The homotopies follow nonlinear paths—geometric means for technologies, log-sum-exp interpolation for costs—that respect the multiplicative and exponential structure of the admissibility conditions. This is why the contractibility of these spaces is not trivial.

where $\kappa_0(x) = \sum_\ell x_\ell$ is the symmetric linear cost function. Exp-convexity is preserved because $\exp(H_\kappa)$ is a convex combination of strictly convex functions. Every step along the path is an admissible cost function: a valid political economy of extraction. A state with steep, politically volatile extraction costs can be smoothly deformed into one with low, fiscally efficient costs, passing through a continuous family of intermediate cost structures, each satisfying the conditions in Definition 3.6.

The paths of institutional change. These homotopies are not merely proof devices. They formalize a claim about the structure of change: the admissibility conditions are exactly the conditions under which continuous institutional transformation is possible within each parameter space. Every convergence path that Tilly described—from coercion-intensive to capitalized coercion, from capital-intensive to bureaucratic national state—has a formal counterpart in the homotopy: a continuous curve in $\mathcal{T} \times \mathcal{K}$ along which every intermediate configuration is itself admissible.

At the systemic level, Gilpin (1981)'s differential growth rates are movements along these paths—one state's technology improving while another's cost structure stiffens—and the contractibility of the parameter spaces guarantees that such paths exist. Whether states follow them smoothly or through catastrophic transitions is a question about dynamics, not about the space of trajectories itself; Section 4.4 returns to this distinction.

4.3   The Contractibility of the State Space

The parameter spaces are contractible, and the solution map is continuous. These two facts together imply the main result.

4.6 Proposition

$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ *strongly deformation retracts onto the point*

$$\pi_0(m) = \left(\exp\left(\frac{m}{L}\right) - 1\right)\mathbf{1},$$

where $\mathbf{1} \in \mathbb{R}^L$ is the vector of ones. In particular, $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is contractible.   [*Proof*.]

The intuition is straightforward. Contract the technology space to its basepoint τ_0 via the geometric-mean homotopy; simultaneously contract the cost space to κ_0

via the log-sum-exp homotopy. At each instant $t \in [0,1]$, the intermediate parameters (τ_t, κ_t) are admissible, so the production problem has a unique solution, and that solution is the policy function π_{τ_t, κ_t}. By continuity of the solution map (Lemma 4.2 and Corollary 4.3), the resulting path $t \mapsto \pi_t$ is continuous in the policy space. At $t = 0$ the path starts at $\pi_{\tau, \kappa}$—the original state—and at $t = 1$ it arrives at $\pi_0 = \pi_{\tau_0, \kappa_0}$, the basepoint.

The technical subtlety lies in one step: distinct parameter pairs may generate the same policy function, so contracting the parameters does not immediately yield a well-defined homotopy on $\mathcal{P}$ itself. The proof resolves this by constructing a *canonical lift*—a continuous map $\pi \mapsto (\tau_\pi, \kappa_\pi)$ that recovers an admissible technology–cost pair from any policy function—and then composing this lift with the parameter-space homotopies. The canonical technology is defined as $\tau_\pi(x) = \sup\{m \in M : \pi(m) \leq x\}$, which reconstructs the technology from the policy's "inverse"; the canonical cost is chosen to be a fixed strictly increasing function independent of π. The details are in the appendix.

The basepoint π_0 sends zero demand to zero mobilization and distributes exponentially increasing resource requirements symmetrically across all commodities as demand grows. It is the policy function of a state that treats all commodities identically and faces no asymmetries in either technology or cost: a maximally symmetric, maximally simple state. This state is empirically unrealistic—no real state treats steel and labor identically—but topologically representative: if a space is contractible, it can be contracted to *any* point, so the choice of basepoint is a matter of convenience, not substance.

4.4 The Ladder of Sameness

Proposition 4.6 is the main structural result of this chapter. But its full significance becomes clear only when we unpack the various ways in which it implies that all states are "the same." Contractibility entails a hierarchy of topological properties, each one weaker than contractibility itself but with its own interpretive content. We present these as three rungs of a ladder, each revealing a deeper layer of equivalence among states.

In summary, Proposition 4.6 implies that the space of all states ignores variation in:

1. *Sorts:* qualitative distinctions among states are necessarily second-order to their sameness as producers of force (Corollary 4.7);

2. *Transformations:* two states can be transformed into one another through continuous deformation (Corollary 4.8); and

3. *Histories:* any two paths between the same pair of states are continuously deformable into one another (Corollary 4.9).

Let us climb the ladder.

4.4.1 Connectedness: no first-order distinctions

The first corollary is that the space of states is *connected.*

4.7 Corollary

$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ *cannot be written as the union of two disjoint non-empty open sets.*
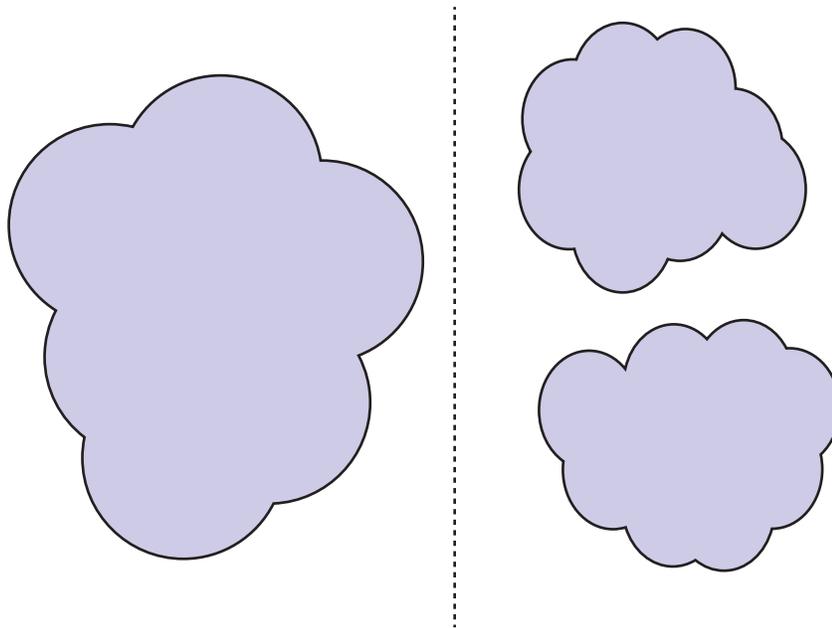
Figure 4.2: *Connected (left) versus disconnected (right) spaces. The connected space is a single piece; any attempt to divide it requires tearing. The disconnected space has two components with no continuous path between them.*

Connectedness means that the space of states is a single piece. No matter how we try to partition it, we cannot separate it into two disjoint groups without tearing

the space apart. If $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ were disconnected, we would have two fundamentally different *sorts* of state—two categories distinguished by the topology of the space itself, not by any external classification. This would be a first-order distinction: a primary division in how states produce force. Corollary 4.7 rules this out. There is only one sort of state.

This does not deny variation. Democracies and autocracies, maritime powers and continental powers, labor-intensive and capital-intensive states—these are all meaningful distinctions. But they are second-order: they divide a connected space into regions, rather than revealing a space that is already divided.

The point can be made precise. A classification of states is a map

$$\mathcal{P}_{\mathcal{T}\times\mathcal{K}} \ni \pi \longmapsto P(\pi) \in \Pi,$$

where Π is a partition of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ into categories. Figure 4.3 illustrates such
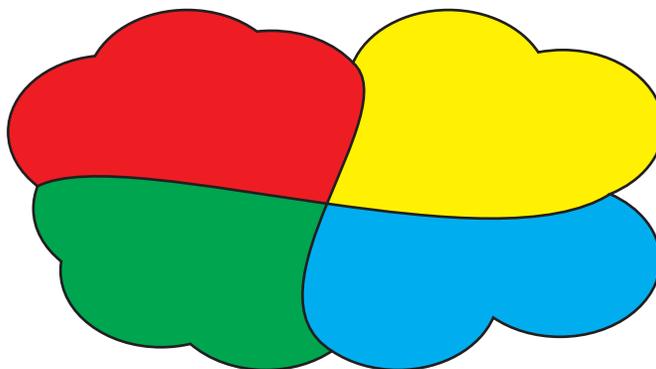
Figure 4.3: *A classification structure P dividing states into categories within a connected component.*

a classification: four subclasses of states, drawn within a single connected component. Since $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is connected, this is the only way to draw meaningful qualitative distinctions without contradicting Corollary 4.7.

Classification structures admit a natural hierarchy. One scheme might divide states into democracies and autocracies; a finer scheme might further distinguish presidential democracies, parliamentary democracies, military autocracies, and single-party regimes. Refinement approaches the limit of singleton categories— where each state is its own class and the classification conveys no information beyond identity. Coarsening approaches the opposite limit: a single category containing all states, the structuralist ideal in its purest form. Functions that send

classification schemes to outcomes of interest can then be assessed for properties like monotonicity or continuity with respect to refinement and coarsening. This provides a framework within which structuralist theorizing is a limiting case of a more general analysis of classification, and within which all such classifications are second-order to the connectedness of the space.

4.4.2 Path-connectedness: continuous transformation

The second corollary is that the space of states is *path-connected.*

4.8 Corollary

For all states $\pi_0, \pi_1 \in \mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, there exists a continuous map

$$\gamma : [0, 1] \longrightarrow \mathcal{P}_{\mathcal{T} \times \mathcal{K}}$$

such that $\gamma(0) = \pi_0$ and $\gamma(1) = \pi_1$.

This is stronger than connectedness. Not only are there no fundamental divisions among states, but any two states can be linked by a continuous path through the space. The path is a continuous family of intermediate states, each itself a valid policy function—a valid way of being a state.

The result has two natural readings. The first is temporal: the transformation of one state into another over time is a continuous process. The United States in 1790—a state with a minimal military, relying on militia and naval power, with a technology shaped by agrarian resources and a cost structure constrained by weak fiscal capacity—and the United States in 2025—a state with a technologically advanced, professionalized military, backed by the world's largest fiscal apparatus—are two points in $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, and the history connecting them is a continuous path γ.

The second reading is emulative: states can learn from one another in a continuous way. Resende-Santos (2007) documents how South American states emulated Prussian military organization in the late nineteenth century, adopting elements of Prussian training, doctrine, and force structure. This emulation is a path in $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$: a continuous deformation from the emulating state's original character toward the character of the model.

Gilpin (1981) elevates both readings to the systemic level. His central argument in *War and Change* is that differential growth rates shift the relative power of states, generating a continuous transformation of the international order.
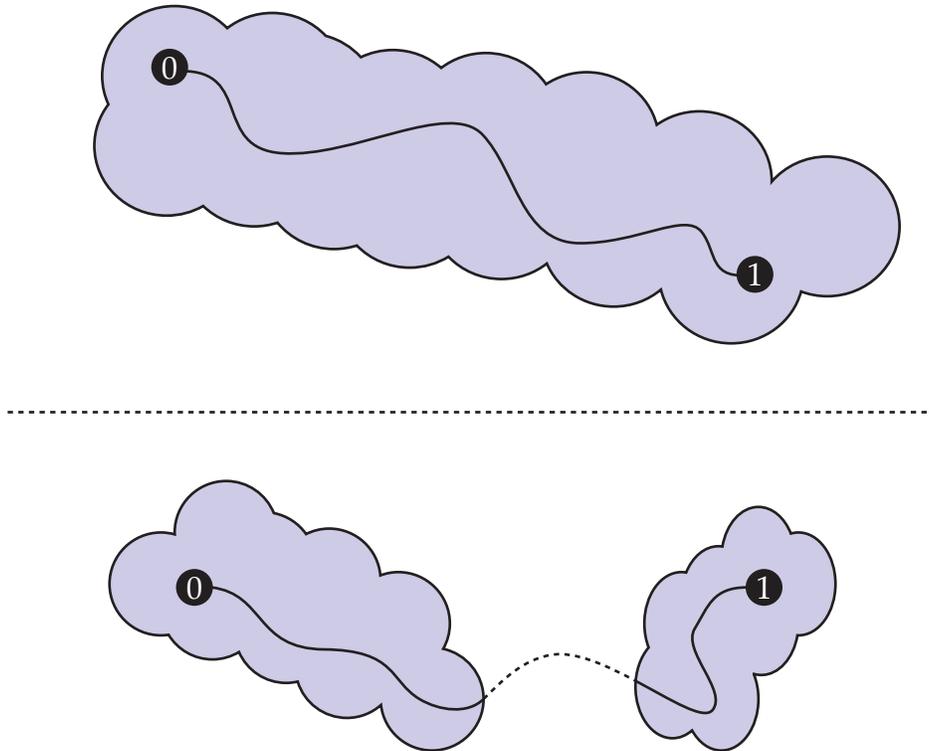
Figure 4.4: *Path-connected (top) versus non-path-connected (bottom) spaces. In the path-connected space, any two points can be joined by a continuous curve that stays within the space.*

The structure of the system persists even as its occupants change: hegemonic transitions are not ruptures in the fabric of international politics but movements along paths in the space of states. In the language of the model, Gilpin's differential growth rates are paths in $\mathcal{T} \times \mathcal{K}$—one state's technology improving while another's stagnates, one state's extraction capacity deepening while another's frays—and Corollary 4.8 guarantees that such paths exist and remain within the space.

Discontinuous transformations remain possible and meaningful, but they are not woven into the fabric of the state space. If flashpoints and tipping points are to be found, they must be sought in second-order structures. Formally, a discontinuous transformation might be represented as a moment $\zeta \in (0, 1)$ and

two evolution functions

$$\gamma_0 : [0, \zeta] \longrightarrow \mathcal{P}_{\mathcal{T} \times \mathcal{K}} \quad \text{and} \quad \gamma_1 : [\zeta, 1] \longrightarrow \mathcal{P}_{\mathcal{T} \times \mathcal{K}},$$

such that $\gamma_0(\zeta) \neq \gamma_1(\zeta)$. This is an evolution characterized by a moment of discontinuous change. Revolution, conquest, institutional collapse—these can all be modeled as flashpoints, but they must be imposed as additional structure on the space. The topology itself does not generate them.

4.4.3 Simple connectedness: the topology of path dependence

The third corollary is that the space of states is *simply connected.*

4.9 Corollary

For all states $\pi_0, \pi_1 \in \mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ and all continuous paths $\gamma_\alpha, \gamma_\beta : [0, 1] \to \mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ with $\gamma_\alpha(0) = \gamma_\beta(0) = \pi_0$ and $\gamma_\alpha(1) = \gamma_\beta(1) = \pi_1$, there exists a continuous homotopy

$$H : [0, 1] \times [0, 1] \longrightarrow \mathcal{P}_{\mathcal{T} \times \mathcal{K}}$$

such that $H(0, \cdot) = \gamma_\alpha$ and $H(1, \cdot) = \gamma_\beta$.

Not only can any two states be linked by a continuous path, but any two such paths between the same endpoints can be continuously deformed into each other. There are no holes in the space that would force paths to go different ways.

This rung of the ladder speaks directly to the theory of path dependence, and it deserves sustained attention.

Path dependence in the study of states. Path dependence is one of the central concepts in comparative politics and historical institutionalism. The idea, in its strongest form, is that the historical trajectory of a state matters for its present character: two states that arrive at the same structural position via different histories may nevertheless differ in ways that reflect those histories. Early institutional choices constrain later possibilities; critical junctures close off alternatives; the sequence of reforms matters, not just their content.

Simple connectedness places a sharp constraint on where such dependence can live. Consider two paths γ_α and γ_β in $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, both connecting the same initial state π_0 to the same final state π_1. Path γ_α might represent one historical
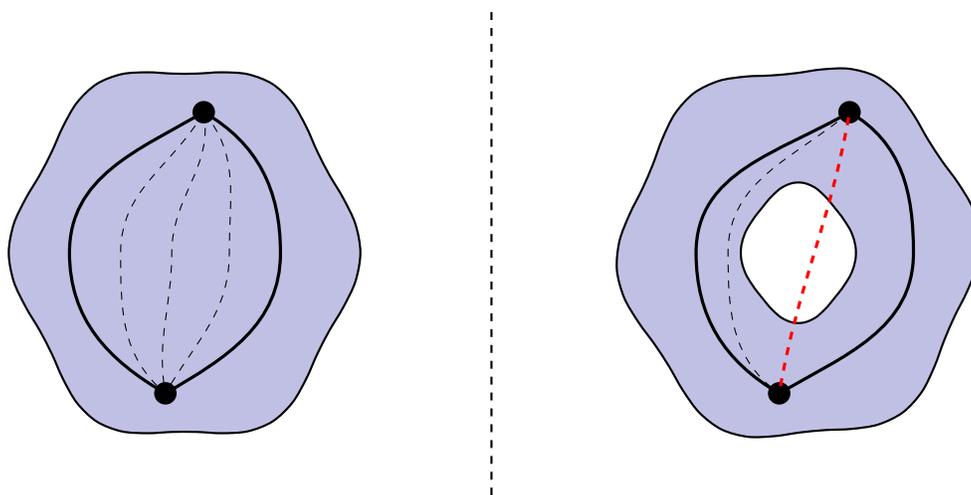
Figure 4.5: *Simply connected (left) versus non-simply connected (right) spaces. At left, any two paths between the same endpoints can be continuously deformed into each other. At right, the hole forces paths on opposite sides to be topologically distinct.*

trajectory: a state that industrialized early, built its extraction apparatus through fiscal innovation, and developed a capital-intensive military technology. Path γ_β might represent another: a state that industrialized late, relied on coercive extraction, and developed a labor-intensive military technology before eventually converging on the same technology–cost pair. Corollary 4.9 says that these two paths are homotopic: there exists a continuous deformation that transforms one into the other without leaving the space.

What does this *not* mean? It does not mean the paths are identical. It does not mean that the states traversed the same intermediate configurations. It does not mean that history is irrelevant. What it means is that there is no *topological obstruction* to deforming one path into the other—no hole in the space that would force the two trajectories to remain permanently distinct. The space of states has no topology-generating memory: no structural feature that, by its shape alone, makes one historical path fundamentally different from another.

Two sources of path dependence. This distinction is crucial, and it helps clarify what the result does and does not rule out. Path dependence can arise from two structurally different sources.

The first is *topological*: the space itself has holes or other obstructions that force different paths to be genuinely inequivalent. If $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ had a non-trivial

fundamental group—if there existed loops that could not be contracted to a point—then paths on opposite sides of a hole would be topologically distinct, and no continuous deformation could transform one into the other. Path dependence of this kind would be intrinsic to the structure of the state space. Corollary 4.9 rules this out entirely. The fundamental group of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is trivial; there are no holes, no obstructions, no topological memory.

The second source is *geometric*: the dynamics, fields, or forces placed on the space make different paths yield different outcomes, even though the paths are topologically equivalent. Two paths may be homotopic—continuously deformable into each other—and yet carry different costs, generate different sequences of intermediate states, or interact differently with external processes. The topology says the paths are equivalent; the geometry says they are not. Path dependence of this kind is not ruled out by Corollary 4.9. It is, however, constrained: it cannot arise from the shape of the space but only from additional structure imposed upon it.

The interpretive stakes. The distinction between topological and geometric path dependence is not merely technical. It tells us where to look for the mechanisms of historical persistence. If the space of states had holes—if it were not simply connected—then path dependence could be a structural feature of the state system, woven into the topology of possible configurations. One would need to know which side of a hole a state's history passed through, and this information would be irreducible. But the space is simply connected, so path dependence, to the extent that it exists, must arise from the particular dynamics that propel states along their trajectories: the political processes, institutional frictions, and power asymmetries that make one path costlier or more consequential than another. These are real and important, but they are second-order to the topology—they live in the geometry of the space, not in its shape.

Making this distinction precise—formalizing what it means for a field on the state space to generate path dependence, and identifying the conditions under which such fields are path-independent—requires geometric structure that the present chapter has not introduced. Contractibility provides the topological foundation; the geometric superstructure is the work of the next chapter, where restricting attention to a tractable subclass of states will equip the space with the richer structure needed to sharpen each rung of the ladder.

4.5   Assessment

Under the admissibility conditions, the space of states is contractible: connected, path-connected, simply connected, and continuously deformable to a point. The ladder of sameness unpacks what this means—no first-order distinctions, continuous transformation between any two states, no topological memory—and the interpretive consequence is sharp. Qualitative differences among states are second-order to their sameness as producers of force. Classifications, typologies, and regime labels are structures *imposed on* a topologically uniform space, not *read off* a space that is already divided.

But the admissibility conditions are not innocent. They were chosen—log-concavity of technology, exp-convexity of cost—because they deliver the well-posedness and continuity on which the entire analysis rests. Whether the space remains so well-behaved when these conditions fail is a question this chapter has deliberately deferred. Chapter 6 takes it up.

These results are topological. They concern the shape of the space, not the dynamics on it. Simple connectedness tells us that paths are homotopic, but it does not tell us what *fields* defined on the space can do—whether line integrals along different paths yield the same value, whether potential functions exist, whether the memory of the system is entirely captured by endpoint information. These are geometric questions, and they require the space to carry more structure than a bare topology.

The next chapter provides that structure. It restricts attention to a tractable subclass of technologies and costs—the *tame* states—shows that their policy functions faithfully represent the classical space, and equips the resulting subspace with a geometric structure richer than contractibility: *convexity.* That structure will deepen each rung of the ladder, sharpening what connectedness, path-connectedness, and simple connectedness mean when the space carries not just a topology but a geometry.

Chapter 5

Modeling the Classical

Mathematics is the art of giving the same name to different things.

Henri Poincaré

The previous chapter established that the space of possible states $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is contractible: connected, path-connected, simply connected, and continuously deformable to a single point. As a topological property, contractibility conveys deep structural information about the set, from how connected it is to how much qualitative information it carries. However, it has little to offer about the *geometry* of the set—its shape, its curvature, its capacity to support the calculations that modeling demands. Philosophically satisfying though it may be, contractibility is not especially helpful for modeling purposes. We therefore turn our attention from topological considerations to geometric ones.

The motivation is practical. Consider what Game 3.1, the motivating arming game of Chapter 3, would look like if we incorporated the primitives we have since defined. The choice variables become resource investment vectors $x_i \in X$; the contest success function runs on inputs that have been sent to a force-like output through the technology; the costs subtracted from the prize have been sent to a value-like output through the cost function. This richness comes at the expense of tractability, as the additional complexity makes the game difficult to analyze beyond general existence results. We can show that equilibria exist (Proposition 5.2), but without imposing additional structure on what each τ and κ actually looks like, we cannot say much about what those equilibria *are*.

Any disciplined restriction of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$—that is, any way of specifying explicit functional forms for τ and κ—constitutes a *model* of the broader class of states.

Introducing such structure is not a matter of convenience alone: it changes what can be said about equilibria, geometry, and even what it means for a "state" to be well-formed. This turns the familiar act of modeling into a deeper methodological question. Three issues immediately arise.

1. What intrinsic and relational properties should a good model of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ possess? How can we gain tractability without erasing the constraints that make states distinctive in the first place?

2. Which distortions are tolerable, and which would destroy the correspondence between the model and the ambient space? In other words, what geometric or topological features must any faithful model preserve?

3. What insight compensates for any loss of fidelity? If a model smooths or simplifies, what new perspective does that simplification reveal?

These questions frame the modeling problem: how to move from the abstract, intractable class of all possible states to a structured, analyzable family that still reflects its essence.

These are not merely technical questions; they are instances of a deep problem in the philosophy of scientific representation. What does it mean for a model to be *adequate* to what it represents—and adequate in what sense? The question is sharpened by the fact, established in Chapter 4, that both the ambient space and its natural subspaces are contractible: their homotopy invariants are all trivial, so invariant preservation cannot by itself distinguish a faithful model from a vacuous one. We will need a notion of adequacy—what we call *structural adequacy*—that locates the content of faithfulness not in which invariants are preserved but in *how* the passage from general to simple is achieved: continuously, within the admissible space, and with the simple functions held fixed throughout. Section 5.4 develops this criterion in detail, drawing on frameworks from van Fraassen (1980), Hertz (1894), and Weisberg (2013).

To make these ideas precise, we need to say what a "model" of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ actually is. A model of a space is a *chosen representative of an equivalence class*: a simpler subspace that stands in for the whole because it carries the same essential structure. In topology, the relevant notion of equivalence is *homotopy equivalence*, and the cleanest way to achieve it is through a *strong deformation retraction*—a continuous process that shrinks the ambient space onto a chosen subspace, leaving the subspace itself fixed throughout.

Hatcher (2002) opens his textbook with an example that makes the idea vivid. Consider the twenty-six letters of the alphabet written in thick block

form—solid regions bounded by curves, like the letters on a marquee. Each thick letter contains a thin one-dimensional "skeleton" inside it: the **A** contains an A made of line segments, the **B** contains a B, and so on. The thick letter can be continuously shrunk to the thin letter by sliding each point along a line segment toward the skeleton, with points already on the skeleton staying fixed. This is a deformation retraction: the thin letter is a subspace of the thick letter, the shrinking is continuous, and the result is a simpler object that is homotopy equivalent to the original. The thin letter is a *model* of the thick one: it discards the two-dimensional bulk but preserves the topology. A thick **A** and a thin A have the same fundamental group, the same number of holes, the same homotopy type. What the thick letter knows, the thin letter knows too.

The modeling problem for $\mathcal{T} \times \mathcal{K}$ is analogous. The ambient space of all admissible technologies and costs is the "thick letter"—rich, high-dimensional, and analytically intractable. We seek a thin skeleton inside it: a subspace of simple, explicit functional forms that is a deformation retract of the whole. If we can find such a subspace, then it is an *adequate model*—not merely a convenient simplification, but a homotopically faithful representative that preserves every topological invariant of the ambient space.

This framing reveals the modeling problem as a search for the right *quotient*. Many different technologies and costs should count as "the same" for modeling purposes—they should map to the same representative in the subspace. The modeling map therefore partitions the ambient space into equivalence classes, each class sharing a single canonical representative. The question is which equivalence relation to impose: which differences matter and which can be quotiented away without loss.[1]

In addition to adequacy, we need *tractability*: the representatives must be smooth enough to admit first-order analysis. The ambient space of Chapter 3 requires only continuity, which is too weak for calculus. A good model must therefore be both topologically faithful and analytically accessible.

---

[1] A partition of a space into equivalence classes, together with a choice of representative from each class, resembles a fiber bundle—and indeed a fiber bundle *could* serve as a model in this sense, with the base space as the space of representatives and the fibers as the equivalence classes. But a fiber bundle imposes additional structure that we do not require: local triviality (the fibers must all look alike) and smooth transition maps between local trivializations. The tamification map produces fibers that vary in shape—some equivalence classes are "thicker" than others—so the quotient is not locally trivial in general. What we need is weaker: a retraction onto a subspace that is a homotopy equivalence, which is exactly what a strong deformation retraction provides.

Topologically, Proposition 4.6 has already given an extreme answer: $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is contractible, and thus representable by a single point—the ultimate quotient, in which all states are identified. But while this representation is formally adequate (a point is homotopy equivalent to any contractible space), it is substantively vacuous. A single point cannot vary, and without variation there can be no explanation. The purpose of modeling is to make sense of variation—to describe how differences in technology and cost structure shape behavior—and a single point can do none of that. We need a quotient that is fine enough to preserve explanatory power yet coarse enough to be tractable.

To preview the results: we will construct such a model. The particular equivalence classes we define are organized by *first-order structure at the origin*: two technologies belong to the same class if and only if they have the same gradient at zero. The canonical representative of each class is a function of especially simple form—log-linear for technologies, linear for costs—which we call *tame*. We will show that the tame subspace is a strong deformation retract of the ambient space (adequate), that its functions are smooth (tractable), and that the resulting space of tame states is not merely contractible but *convex*—a geometric structure richer than anything contractibility alone can provide.

The chapter proceeds in six sections. Section 5.1 enriches the motivating arming game with the primitives developed in Chapter 3, shows that general existence results are possible but characterization is not, and introduces the tame technologies and costs that will serve as the target model. Sections 5.2 and 5.3 develop the two-stage modeling pipeline: regularization smooths arbitrary functions into differentiable ones, and tamification extracts their canonical first-order representatives. Section 5.4 confronts the question of what adequacy means for a subspace of a function space, defines a formal criterion of *structural adequacy*, and establishes that the tame subspace satisfies it. Section 5.5 develops the geometric payoff: tame states form a convex set, deepening each rung of the ladder of sameness from Chapter 4. Section 5.6 returns to the enriched game with tame functions and derives the water-filling equilibrium that repays the modeling investment.

5.1 The Model

To make the modeling question concrete, consider the following enrichment of Game 3.1.

5.1 Game

Two states, $i \in \{1, 2\}$, simultaneously choose resource investments $x_i \in X = \mathbb{R}_+^L$. Their payoffs are

$$U_i(x_1, x_2) = p_i(x_1, x_2) \cdot \left(V - k\left(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\right) \right),$$

where the contest success probabilities are

$$p_1(x_1, x_2) = \frac{\lambda \left(\tau_1(x_1)\right)^\alpha}{\lambda \left(\tau_1(x_1)\right)^\alpha + \left(\tau_2(x_2)\right)^\alpha}, \qquad p_2 = 1 - p_1,$$

with λ, α, V, and k as in Game 3.1, $\tau_i \in \mathcal{T}$ is State i's technology, and $\kappa_i \in \mathcal{K}$ is State i's cost function.[2]

Game 5.1 is an upgrade over Game 3.1 in that it incorporates the technologies and costs we have defined, making the choice variable something observable—the resource investment vector x_i—rather than the abstract military capability m_i. Where Game 3.1 treated military competition as a one-dimensional contest between abstract force levels, Game 5.1 reveals it as a multidimensional allocation problem: each state distributes resources across L commodities, transforming them through its technology and bearing costs through its cost function. The enrichment opens the black box of Chapter 3 inside a strategic interaction. We can at least provide general existence results:

5.2 Proposition

Game 5.1 has at least one pure-strategy Nash equilibrium.   [*Proof*.]

But existence is a thin reward. Compare what we know about Game 3.1—the original game has a unique equilibrium in closed form, with clean comparative statics in every parameter—with what we know about Game 5.1: equilibria exist, but we cannot characterize them. The difficulty is not technical carelessness but structural. Without knowing the functional form of each τ, we cannot compute marginal products; without knowing each κ, we cannot compute marginal costs; without both, first-order conditions yield no explicit solution. The game is well-posed but opaque. And the opacity is not merely a matter of computational difficulty: it reflects the fact that the admissibility conditions—designed for

---

[2]As in Game 3.1, we set the contest outcomes to $\lambda/\lambda+1$ and $1/\lambda+1$ when $\tau_1(x_1) = \tau_2(x_2) = 0$.

maximal generality in Chapter 3—admit an enormous variety of functional forms, each potentially yielding a different equilibrium structure.

We therefore need a modeling strategy: a way to restrict the class of technologies and costs to forms that are both analytically tractable and topologically faithful. It will be easier to introduce the target model first, and then work backward to see how we might arrive at it.

5.3 Definition (Tame technologies and costs)

A technology $\tau \in \mathcal{T}$ is tame *if it takes the form*

$$\tau(x) = A_\tau \sum_{\ell \in L} \beta_\ell \cdot \log\left(1 + x_\ell\right),$$

where $A_\tau > 0$ is a scale parameter and $\beta \in \Delta_L := \{b \in \mathbb{R}_{\geq 0}^L \mid \sum_\ell b_\ell = 1\}$ is a vector of input elasticities that witnesses τ's tameness.

A cost $\kappa \in \mathcal{K}$ is tame *if it takes the form*

$$\kappa(x) = A_\kappa \sum_{\ell \in L} q_\ell \cdot x_\ell,$$

where $A_\kappa > 0$ is a scale parameter and $q \in \Delta_L$ is a vector of input prices that witnesses κ's tameness.

The sets of all tame technologies and costs are denoted $\mathcal{T}^{[\mathfrak{T}]}$ and $\mathcal{K}^{[\mathfrak{T}]}$, respectively.

The tame technologies and costs are both familiar and extraordinarily simple. They are smooth, monotone, and have first and second derivatives that are easy to compute, work with, and interpret. The tame technology is the log-linear form previewed in Chapter 3: each commodity contributes independently with diminishing returns, and the elasticity vector β governs relative importance. The tame cost is linear: each unit of each commodity costs a fixed amount, and the price vector q governs relative expense. Together, they define a production problem whose solutions can be written in closed form.

Neither functional form is ad hoc. The log-linear technology belongs to the family of additively separable production functions that have been workhorses of economic theory since the study of production began in earnest (Mas-Colell, Whinston and Green, 1995, ch. 5). Its closest kin is the Cobb-Douglas function, which shares the property of independent, diminishing-returns contributions

from each input; the logarithmic form arises naturally when inputs enter multiplicatively in force production and are measured in log-units of effectiveness. The linear cost, meanwhile, is the cost structure of a price-taking state—one that faces fixed per-unit prices for each commodity. This is the simplest cost structure consistent with the admissibility conditions, and it corresponds to the standard competitive-market assumption in the state-as-firm literature (Lake, 1992): the state is a consumer of resources whose individual purchases do not move the market price. When the state is large enough to affect prices—when extraction itself generates political costs that increase at the margin—the cost function departs from linearity, and the general framework of Chapter 3 applies. The tame cost is the limiting case in which political friction is constant per unit of resource.

It would be wonderful news indeed if these tame functions could adequately represent the general functions. For starters, a straightforward check shows that the tame functions are members of their respective ambient spaces:

5.4 Lemma

$\mathcal{T}^{[\mathfrak{T}]} \subseteq \mathcal{T}$ and $\mathcal{K}^{[\mathfrak{T}]} \subseteq \mathcal{K}$.

We omit the proof, as it is a simple exercise in checking the definitions of Chapter 3. So the tame functions are at least subsets of the general functions.

But there are many properties such subsets might have or lack, and the previous chapter established that contractibility is the most important topological property of the general function spaces. The tame functions retain it:

5.5 Lemma

The tame function spaces $\mathcal{T}^{[\mathfrak{T}]}$ and $\mathcal{K}^{[\mathfrak{T}]}$ are contractible.

To see this, consider the homotopies

$$H_\tau(t, \tau) := (1 - t)\, \tau + t \left(\sum_{\ell \in L} \tfrac{1}{L} \cdot \log\left(1 + x_\ell\right) \right),$$

$$H_\kappa(t, \kappa) := (1 - t)\, \kappa + t \left(\sum_{\ell \in L} \tfrac{1}{L} \cdot x_\ell \right),$$

which send a given tame technology or cost to a "central" tame function with equal weights on all commodities and unit scale. For each commodity ℓ, the ℓth

component of the homotopy is a convex combination of the original weight and the uniform weight, so the homotopies remain within the tame function spaces for all $t \in [0, 1]$. In particular, their scale terms become $A_\tau(t) = (1 - t)A_\tau + t$ and $A_\kappa(t) = (1 - t)A_\kappa + t$, and their weight terms become

$$\beta_\ell(t) = \frac{(1 - t)A_\tau \beta_\ell + t/L}{(1 - t)A_\tau + t}, \qquad q_\ell(t) = \frac{(1 - t)A_\kappa q_\ell + t/L}{(1 - t)A_\kappa + t},$$

which remain non-negative and sum to one for all $t \in [0, 1]$.

So we have demonstrated that the tame functions are subsets of the general functions and that they retain contractibility. But the question remains: what is the modeling action here? The fact that the tame functions behave like the general functions is reassuring, but we have not yet introduced a notion of *representation*. What might be the nature of a mapping that sends a pair $(\tau, \kappa) \in \mathcal{T} \times \mathcal{K}$ to a tame pair in $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$?

5.2   Regularization

Recall that neither the technology conditions of Chapter 3 nor the cost conditions imposed any differentiability assumptions; we required only continuity. This was intentional: continuity captures responsiveness without presupposing smooth substitutability or differentiable marginal rates. It also gave us the largest possible ambient space in which to reason about technological and behavioral forms.

The drawback is analytic. A merely continuous technology τ or cost κ may have corners, flats, or kinks that block the use of gradients and first-order tools. Before we can extract the tame representative of an arbitrary function, we must pass through a stage of *regularization*: a systematic smoothing of rough functions into differentiable ones.

Like sandpaper, regularization smooths away the small irregularities of τ and κ—the corners, kinks, and rough patches—while leaving their large-scale shape intact. Formally, we seek a continuous operator

$$\mathfrak{D} : \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]},$$

where $\mathcal{T}^{[\infty]} := \mathcal{T} \cap C^\infty(X, M)$ and $\mathcal{K}^{[\infty]} := \mathcal{K} \cap C^\infty(X, \mathbb{R}_{\geq 0})$ are the smooth elements of the respective ambient spaces. Several features of this operator are

essential: it must be continuous as a map, preserve the structural properties of $\mathcal{T}$ and $\mathcal{K}$ (monotonicity, convexity, concavity), and—most importantly—fix the tame functions that already behave well:

$$\mathfrak{D}\big|_{\mathcal{T}^{[\mathfrak{T}]}\times\mathcal{K}^{[\mathfrak{T}]}} = \mathrm{id}_{\mathcal{T}^{[\mathfrak{T}]}\times\mathcal{K}^{[\mathfrak{T}]}}.$$

Intuitively, $\mathfrak{D}$ acts like a variable-bandwidth mollifier: it smooths aggressively where a technology or cost is rough, and not at all where the function is already tame. Smooth regularization is not exotic; it expresses a basic fact of functional analysis: smooth functions are dense in their continuous and convex-concave parents. On compact domains, every continuous function can be uniformly approximated by a smooth one, and the same holds under monotonicity, convexity, or concavity constraints.[3]

5.6 Proposition (Regularization)

There exists a continuous regularization operator

$$\mathfrak{D} : \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]},$$

such that $\mathfrak{D}\big|_{\mathcal{T}^{[\mathfrak{T}]}\times\mathcal{K}^{[\mathfrak{T}]}} = \mathrm{id}$; *in other words,* $\mathfrak{D}$ *fixes the tame functions.* [*Proof*.]

The regularization step takes the form of a homotopy

$$t \longmapsto (1 + \tau)^{1-t} \cdot (1 + \mathfrak{D}_\tau(\tau))^t - 1,$$
$$t \longmapsto \log\big((1 - t)\exp\kappa + t\exp\mathfrak{D}_\kappa(\kappa)\big),$$

for $t \in [0, 1/2]$, which continuously deforms any technology or cost into its regularized counterpart. Since $\mathfrak{D}$ fixes the tame functions, this homotopy remains within the tame function spaces when started there. This is the first half of our two-stage modeling pipeline.

---

[3]There are several constructions that achieve this: convolution with a smooth kernel, Moreau envelopes, and spline regularization are all standard. The *causal convolution* approach is the most natural here: it preserves coordinatewise monotonicity, respects the boundary of the nonnegative orthant, and can be tuned continuously through a gauge that measures distance to the tame subclass. Full details are given in Chapter B.

5.3   Tamification

What does it mean to represent a complex function by a simpler one? What, precisely, does simplification *do*—and how does it preserve what matters? Before claiming that the tame states adequately represent the general states, we must articulate what the act of representation consists of and what structure it must respect.

Consider a regularized technology $\tau \in \mathcal{T}^{[\infty]}$. Does it already come equipped with a tame representative? If so, there would exist a positive scale $A_\tau > 0$ and a weight vector $\beta \in \Delta_L$ such that

$$\tau(x) = A_\tau \sum_{\ell \in L} \beta_\ell \log(1 + x_\ell).$$

This is a highly restrictive condition, so we cannot expect an arbitrary τ to satisfy it exactly. Instead, we seek a systematic way to extract from τ a canonical pair (A_τ, β)—a "tame image" that captures its first-order structure.

A natural place to begin is the gradient of τ at the origin, $\nabla \tau(0)$. Because τ is smooth and strictly increasing, each partial derivative $\partial_\ell \tau(0)$ is positive, and the gradient encodes the initial marginal product of each input. Normalizing this gradient yields a probability vector

$$\beta_\ell := \frac{\partial_\ell \tau(0)}{\sum_{j \in L} \partial_j \tau(0)},$$

which measures the relative importance of each input at the start of production. The corresponding scale factor

$$A_\tau := \sum_{\ell \in L} \partial_\ell \tau(0)$$

records the total initial productivity of the technology. With these parameters, the *tame representation* of τ is

$$\mathfrak{T}(\tau)(x) := A_\tau \sum_{\ell \in L} \beta_\ell \log(1 + x_\ell).$$

The same reasoning applies to costs. For a regularized cost $\kappa \in \mathcal{K}^{[\infty]}$, define

$$q_\ell := \frac{\partial_\ell \kappa(0)}{\sum_{j \in L} \partial_j \kappa(0)}, \qquad A_\kappa := \sum_{\ell \in L} \partial_\ell \kappa(0),$$

and let $\mathfrak{T}(\kappa)(x) := A_\kappa \sum_{\ell \in L} q_\ell x_\ell$. The normalized vector $q \in \Delta_L$ expresses the initial marginal cost shares across inputs, while A_κ measures the overall cost scale. Since κ is strictly increasing, these quantities are positive, and $\mathfrak{T}(\kappa)$ is a well-defined tame cost.

Taken together, these constructions define a *tamification operator*

$$\mathfrak{T} : \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]} \longrightarrow \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]},$$

that sends each regularized technology-cost pair to its tame representation. This operator captures the first-order structure of technologies and costs at the origin, distilling them into their simplest functional forms. The contrast with Diewert's (1971) *flexible functional forms* is instructive. Diewert asks which functional forms provide second-order approximations to an arbitrary production function at a point—a local, metric criterion. Tamification asks which forms are the canonical first-order representatives of equivalence classes under gradient identity at the origin—a global, topological criterion. The two are complementary: Diewert tells us how well a specific form approximates locally; tamification tells us which form *represents the class* structurally.

Remarkably, tamification renders non-identical objects equivalent. Consider two distinct technologies τ_0 and τ_1 in $\mathcal{T}^{[\infty]}$. If $\nabla\tau_0(0) = \nabla\tau_1(0)$, then $\mathfrak{T}(\tau_0) = \mathfrak{T}(\tau_1)$, even though the two functions are distinct. This allows us to treat the tame functions as a *quotient*: each tame function represents an entire equivalence class of regularized functions that share the same first-order structure at the origin.[4]

Crucially, tamification leaves tame functions unchanged:

5.7 Lemma
For all $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, $\mathfrak{T}(\tau, \kappa) = (\tau, \kappa)$.

This is a simple consequence of the definition, but it is important. It states that the tame functions are *fixed points* of the tamification process. If we think of $\mathcal{T} \times \mathcal{K}$ as a peach, then the tame functions $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ are the pit at its center. We can squeeze the peach down to the pit, with the squishing of flesh and the dripping of juice representing the information lost in the tamification process.

---

[4]One could quotient by higher-order derivatives as well—functions with the same Hessian matrices, same third-order information, and so on. This brings us to the study of *jets*; see the discussion of k-tame subspaces in Section 5.4.

But the pit is unchanged by the squishing; it is invariant under the process. In naming it representative of its extrinsic abode, we do it no intrinsic injustice.

The tamification step takes the form of a second homotopy

$$t \longmapsto \left(1 + \mathfrak{D}_\tau(\tau)\right)^{1-t} \cdot \left(1 + (\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau)\right)^t - 1,$$

$$t \longmapsto \log\left((1-t)\exp(\mathfrak{D}_\kappa(\kappa)) + t\exp((\mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)(\kappa))\right),$$

for $t \in \left[1/2, 1\right]$, which continuously deforms the regularized functions into their tame representatives. Since $\mathfrak{T}$ fixes the tame functions, this homotopy remains within the tame function spaces when started there—across all $t \in \left[0, 1\right]$, the tame functions are fixed points of the entire deformation.

5.4   Structural Adequacy

We have now introduced a full theory of representation through regularization and tamification. Regularization smooths arbitrary technologies and costs into differentiable ones; tamification extracts from those smooth functions a canonical tame representative. But is the result *adequate*? Before constructing the retraction that combines these two steps, we must confront a prior question: what should "adequacy" mean in this setting?

5.4.1   The contractibility problem

The natural first answer—that a model is adequate if it preserves the topological invariants of the ambient space—turns out to be vacuous here. Both the ambient space $\mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]}$ and the tame subspace $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ are convex subsets of locally convex topological vector spaces, and convex subsets are contractible. Every homotopy invariant of both spaces is therefore trivial: $\pi_n = 0$ and $H_n = 0$ for all $n \geq 1$; both spaces are path-connected and simply connected. Any continuous map between them—including the constant map that sends every function to a single point—preserves all of these invariants. If invariant preservation were the whole content of adequacy, the claim that the tame subspace is adequate would say nothing at all.

This observation forces a sharper question. The tame subspace is not just any subspace; it was constructed by a specific process (regularization followed by tamification) that respects the admissibility conditions at every stage. If the content of adequacy does not lie in the invariants, it must lie in the *process*—in the nature of the map that takes general functions to their tame representatives.

5.4.2 A criterion of structural adequacy

We need a notion of adequacy suited to the relationship between a function space and a tractable subspace of it. The existing philosophical literature offers instructive frameworks but not quite the right one.

van Fraassen's (1980) *empirical adequacy* requires that a theory have a model whose empirical substructures are isomorphic to all observable appearances. This separates adequacy from truth and ties it to structural embedding—a valuable insight. But empirical adequacy governs the relationship between a theory and the world, not between a space and its subspace. It asks whether the model *contains* a faithful image of the observable; we need to ask whether the ambient space can be *continuously compressed* onto the subspace.

Hertz's (1894) Bild theory offers three desiderata: *permissibility* (mathematical consistency), *correctness* (empirical fit), and *appropriateness* (simplicity among correct representations). In our setting, tame functions are permissible—they satisfy the admissibility conditions. They are correct—they inhabit the ambient space. And they are appropriate—they are the simplest representatives of their equivalence classes, carrying only first-order structure where the general functions carry all orders. Hertz's triad captures the intuition but lacks the formal apparatus to say *how* the passage from general to simple is achieved.

Weisberg (2013) adds the observation that different *representational ideals* warrant different fidelity criteria: a model built for prediction may tolerate structural distortion that a model built for understanding may not. Our setting selects a *topological completeness* ideal, because the objects being represented are function spaces whose identity is constituted by their topological structure.

Drawing on these frameworks, we propose the following criterion.

5.8 Definition (Structural adequacy)

A subspace $M \subset X$ is structurally adequate *for X if there exists a strong deformation retraction $H: X \times [0,1] \to X$ such that $H(x,0) = x$, $H(x,1) \in M$, and $H(m,t) = m$ for all $m \in M$ and $t \in [0,1]$, with $H(x,t)$ satisfying the admissibility constraints of X for every t.*

The admissibility constraint is what distinguishes structural adequacy from bare retractability. It says: one can simplify without ever stepping outside the space of meaningful objects. The retraction is not a mathematical trick that routes through pathological intermediate functions; it is a genuine process of structural simplification, well-defined at every stage.

Structural adequacy differs from the principal alternatives. *Empirical adequacy* requires embeddability of observable substructures into a model; structural adequacy requires the *entire* ambient space to retract onto the submodel, a strictly stronger demand. *Metric approximation*, as in Diewert's (1971) flexible functional forms, requires small pointwise error in a neighborhood of a given point; structural adequacy says nothing about pointwise closeness but everything about the continuous, admissible structure of the passage from general to canonical. Weisberg's (2013) weighted feature-matching similarity is more general still; structural adequacy is a special case in which the relevant features are topological and the weights are binary—preserved or not.

The force of the criterion is threefold.

1. **Continuity.** Nearby general functions must map to nearby representatives. The parameterization is stable: small perturbations in $\mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]}$ produce small perturbations in the tame image. There is no discontinuous assignment or arbitrary choice.

2. **Path-fidelity.** The deformation path $H(f, t)$ must stay inside the admissible space at every intermediate step $t \in [0, 1]$. Every function encountered along the retraction satisfies the log-concavity and exp-convexity conditions of Chapter 3. We never leave the space of well-formed technologies and costs.

3. **Fixity.** Functions already in the subspace must be fixed: $H(f, t) = f$ for every $f \in M$ and every $t \in [0, 1]$. The retraction does not move what is already in canonical form.

Together, these properties ensure that the tame subspace is not merely a convenient subset but a *canonical core*: every general function has a unique tame representative, reachable by a continuous admissible path, with the representatives themselves held fixed throughout.

5.4.3   The retraction

We now show that the tame subspace satisfies this criterion. Combining regularization and tamification into a single continuous deformation

$$H : [0, 1] \times \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T} \times \mathcal{K},$$

defined for the technologies by

$$\begin{cases} (1 + \tau)^{1-2t} \cdot (1 + \mathfrak{D}_\tau(\tau))^{2t} - 1, & t \in [0, 1/2], \\ (1 + \mathfrak{D}_\tau(\tau))^{2-2t} \cdot (1 + (\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau))^{2t-1} - 1, & t \in [1/2, 1], \end{cases}$$

and for the costs by

$$\begin{cases} \log\left((1 - 2t)\exp\kappa + 2t\exp\mathfrak{D}_\kappa(\kappa)\right), & t \in [0, 1/2], \\ \log\left((2 - 2t)\exp\mathfrak{D}_\kappa(\kappa) + (2t - 1)\exp(\mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)(\kappa)\right), & t \in [1/2, 1]. \end{cases}$$

The two branches meet at $t = 1/2$, where $H(1/2, \tau, \kappa) = (\mathfrak{D}_\tau(\tau), \mathfrak{D}_\kappa(\kappa))$, ensuring continuity. At $t = 0$, H is the identity; at $t = 1$, it yields the tamified regularization $(\mathfrak{T}_\tau \circ \mathfrak{D}_\tau, \mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)$. Because $\mathfrak{D}$ and $\mathfrak{T}$ both act as the identity on the tame functions, these remain fixed throughout. Hence H defines a strong deformation retraction of $\mathcal{T} \times \mathcal{K}$ onto $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, and the admissibility conditions are satisfied at every stage of the deformation.

5.9 Proposition (Structural adequacy of the tame representation)
$\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ *is structurally adequate for* $\mathcal{T} \times \mathcal{K}$.   [*Proof*.]

The three questions posed at the opening of this chapter now have precise answers. What properties should a good model possess? Retractability from the ambient space via admissible paths. Which distortions are tolerable? Those that do not break admissibility along the deformation path. What insight compensates for any loss of fidelity? The tame functions, being finite-dimensional and smooth, admit gradient characterization—opening the door to the explicit geometric analysis that follows.

The tame model is not unique. Structural adequacy is a property of the retraction, not of a particular choice of subspace, and the tame functions are not the only structurally adequate model of the ambient space. To see this, consider what happens when we quotient by finer information. Tamification extracts the first-order jet of each function at the origin—the gradient $\nabla\tau(0)$—and discards all higher-order structure. But we could instead extract the full k-jet: the gradient together with the Hessian, the third-order tensor, and so on up to order k. The resulting k-*tame* subspace $\mathcal{T}^{[\mathfrak{T}_k]}$ would consist of functions determined by their first k derivatives at the origin, and the same regularize-then-quotient pipeline

would yield a strong deformation retraction satisfying admissibility at every stage. Each k-tame subspace is structurally adequate, and they form a nested sequence

$$\mathcal{T}^{[\mathfrak{T}]} = \mathcal{T}^{[\mathfrak{T}_1]} \subset \mathcal{T}^{[\mathfrak{T}_2]} \subset \cdots \subset \mathcal{T}^{[\infty]},$$

each finer than the last: a second-order tame representative distinguishes technologies that agree at the margin but differ in how fast diminishing returns set in, while a first-order representative collapses them.

The choice of jet order is therefore a modeling decision, not a mathematical necessity. But the tame model occupies a distinguished position in this hierarchy.

Recall that the ambient space of Chapter 3 required only continuity—no differentiability. Regularization is what purchases first-order analysis: it smooths the rough functions into objects on which gradients are defined. Having paid that cost, the natural question is: what is the *coarsest* equivalence relation on the smooth functions that does not destroy the first-order structure regularization has just provided? The answer is gradient identity at the origin. Any coarser relation would merge functions with distinct gradients, collapsing precisely the information that regularization was designed to make available. Any finer relation—matching Hessians, third derivatives, and so on—would preserve additional structure, but that structure plays no role in first-order analysis.

The tame model is therefore not just one quotient among many; it is the *coarsest* quotient of the smooth admissible functions that is compatible with the analytical purpose of regularization. The k-tame models for $k \geq 2$ are refinements: they make distinctions (in curvature, in higher-order substitution patterns) that the tame model deliberately collapses, because those distinctions are invisible to the gradient-based analysis that the rest of this chapter depends on.

Having established adequacy in the parameter space, we may turn to the states themselves. Because the solutions to the SPP vary continuously with the technologies and costs (Lemma 4.2 and Corollary 4.3), the deformation retraction of $\mathcal{T} \times \mathcal{K}$ onto its tame subspace propagates to the policy function space.

5.10 Definition

The set of tame states *is*

$$\mathcal{P}_{\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}} := \left\{ \pi_{\tau,\kappa} : M \to X \mid (\tau,\kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]} \right\}.$$

The tame states are structurally adequate for $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$: they stand in precisely the same relationship to the full space of states as the tame technologies and costs do to theirs.[5]

5.5 The Geometry of Tame States

Having established the adequacy of the tame representation, we may turn to its geometry. Now that we have smooth functions, we may characterize the solutions to the SPP by relating marginal products to marginal costs.

5.11 Lemma

For all $(m, \tau, \kappa) \in M \times \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]}$, $\pi_{\tau,\kappa}(m)$ solves the SPP if and only if there exist a Lagrangian multiplier $\lambda_{m,\tau,\kappa} > 0$ and Karush-Kuhn-Tucker multipliers $\eta_{m,\tau,\kappa} \in \mathbb{R}_{\geq 0}^{L}$ such that:

$$\frac{\partial \kappa}{\partial x_\ell}\left(\pi_{\tau,\kappa}(m)\right) - \lambda_{m,\tau,\kappa}\frac{\partial \tau}{\partial x_\ell}\left(\pi_{\tau,\kappa}(m)\right) - \eta_{m,\tau,\kappa,\ell} = 0 \quad \textit{for all } \ell \in L,$$

$$m - \tau\left(\pi_{\tau,\kappa}(m)\right) = 0,$$

$$\eta_{m,\tau,\kappa,\ell} \cdot \pi_{\tau,\kappa,\ell}(m) = 0 \quad \textit{for all } \ell \in L.$$

For tame $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, these reduce to

$$A_\kappa q_\ell - \lambda_{m,\tau,\kappa}\frac{A_\tau \beta_\ell}{1 + \pi_{\tau,\kappa,\ell}(m)} - \eta_{m,\tau,\kappa,\ell} = 0 \quad \textit{for all } \ell \in L,$$

$$m - A_\tau \sum_{\ell \in L} \beta_\ell \log\left(1 + \pi_{\tau,\kappa,\ell}(m)\right) = 0,$$

$$\eta_{m,\tau,\kappa,\ell} \cdot \pi_{\tau,\kappa,\ell}(m) = 0 \quad \textit{for all } \ell \in L.$$

The first-order conditions are the door that differentiability opens. We have access to them because we have introduced a differentiable structure on the set of states, and they are equivalent to a solution because the admissibility conditions

---

[5]In the proof of Proposition 4.6, we defined a canonical lift that sent arbitrary $\pi_{\tau,\kappa}$ to a particular (τ, κ) combination, and this lift is continuous in π. To retract arbitrary π onto tame π, we first lift to a canonical technology-cost pair, then work through the homotopy H as constructed above.

ensure the production problem is well-behaved. Owing to the simplicity of the tame forms, the first-order conditions yield explicit solutions, and it is a straightforward matter to arrive at the central geometric result of this chapter.

5.12 Proposition (Convexity of the tame state space)

$\mathcal{P}_{\mathcal{T}[\mathfrak{T}] \times \mathcal{K}[\mathfrak{T}]}$ *is a convex set.* [*Proof*.]

We have shown that the set of tame states is convex—a much stronger result than the contractibility of the general states. Not only does the peach pit adequately represent the peach without distortion, but it also has a richer structure than the peach itself.

What is convexity that contractibility is not?

Interpolation. Any convex set is automatically contractible, but not vice versa. Thus, there is more information in the convexity of the tame states than in the contractibility of the general states. Since the set of tame states is convex, we may define a *convex combination* of two states π_0 and π_1 as $\pi_\lambda := \lambda \pi_0 + (1 - \lambda)\pi_1$, where $\lambda \in [0, 1]$ sets the terms of the combination. This is a well-defined operation whose result is a state. Whereas we could only link two states in the general set through a continuous path, we can now link them through a straight line. Interpolations like this mean that we may define functions on $\mathcal{P}_{\mathcal{T}[\mathfrak{T}] \times \mathcal{K}[\mathfrak{T}]}$ that possess properties like concavity, convexity, or quasiconcavity—properties that are not well-defined on non-convex sets. As these properties are the bread and butter of economic analysis, this is a powerful result: one can imagine choosing an optimal state from a set of states, and then using the properties of the set to analyze the implications of that choice.

Geometry. Convexity is a geometric property, not merely a topological one; it refers to shape, not merely connectivity. The boundary of a convex set maintains curvature with constant sign—it always bends in the same direction. We are therefore able to learn more about the structure of the set from the convexity of the tame states than we could from the contractibility of the general states.

Discovery. Convexity in this setting is not assumed; it is *discovered.* We did not impose convexity on the production functions or the cost structure. Instead, we began with tame functions motivated by representational adequacy and computational accessibility, and convexity emerged from the internal geometry

of their solutions. The tame states are not just analytically convenient, nor just adequate for representation, nor just geometrically simple: they are, somehow, all of these things at once.

5.5.1 The ladder, deepened

The previous chapter established a ladder of sameness for the classical state space: connected, path-connected, simply connected. Each rung was a corollary of contractibility, and the interpretive content concerned what topological equivalence means for systemic theorizing. Convexity deepens each rung.

Connectedness. Contractibility tells us that $\mathcal{P}^{[\mathfrak{T}]}$ is a single piece. Convexity tells us it is a single *convex* piece: for any two states π_0, π_1, the midpoint $\frac{1}{2}\pi_0 + \frac{1}{2}\pi_1$ is also a state. No state is "far" from any other in the sense that the interpolating state is always reachable. Where connectedness forbids tearing, convexity forbids concavities—indentations in the boundary that would make some intermediate states unreachable by straight-line paths. For the classification structures of Section 4.4.1, this is a sharpening: not only do qualitative typologies live within a single connected component, but any two states that a typology places in different categories are connected by a straight line of intermediate states—each itself a valid state—that the typology must assign *somewhere*. The midpoint of a democracy and an autocracy, in the tame model, is not a logical impossibility but a well-defined policy function with intermediate characteristics.

Path-connectedness. In the general space, any two states can be connected by a continuous path, but that path may be tortuous. In the convex tame space, the straight-line path $t \mapsto (1-t)\pi_0 + t\pi_1$ always works. Every pair of states is connected by a geodesic that never detours. The "emulative" and "temporal" readings of path-connectedness from Chapter 4—gradual transformation, institutional learning—are sharpened: the transformation can always proceed by direct interpolation between the starting and ending configurations. Recall the emulative reading: Resende-Santos's South American states emulating Prussian military organization traced a continuous path in $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$. Convexity guarantees that the direct interpolation between the emulating and the emulated state—the "blend" that combines their elasticity vectors and price vectors at any ratio—is itself a valid state. Emulation need not follow a circuitous institutional route; a

straight line through the space of production characteristics is always available. Whether states *take* such direct paths is an empirical question, but the geometry ensures that no structural obstruction prevents it.

Simple connectedness. In the general space, any two paths between the same endpoints are homotopic: they can be continuously deformed into each other. In the convex tame space, there is a stronger fact: any two paths between the same endpoints can be linearly interpolated, and the interpolation itself stays within the space. The fundamental group is not merely trivial but trivially so. Moreover, convexity ensures that Poincaré's lemma applies in its strongest form: every closed 1-form on a convex domain is exact. This means that path dependence in the geometric sense—where line integrals along different paths yield different values—can arise *only* from non-closed fields, never from the topology or geometry of the space itself. The tame state space has no structural memory; any memory in the system must come from the dynamics imposed upon it, not from the arena in which those dynamics unfold. This completes the argument that Chapter 4 left open: there, simple connectedness ruled out *topological* path dependence but could not yet address geometric path dependence, because the space lacked the structure to define line integrals and potential functions. Convexity supplies that structure and delivers the sharp result: the only source of path dependence in the tame state space is the non-closedness of the field that drives the dynamics. Historical institutionalists who study how early choices constrain later possibilities must locate their mechanisms in the fields imposed on the space—the political processes, power asymmetries, and institutional frictions—not in the geometry of possible states itself.

Synthesis. Convexity does not merely confirm the ladder of sameness; it sharpens it. The space of tame states is not just contractible but *linearly* so: paths are straight, interpolations are states, and the geometry carries no intrinsic memory. This is the richest sense in which states are "alike"—not just topologically indistinguishable, but geometrically interchangeable.

5.5.2 Where path dependence lives

The ladder tells us where path dependence *cannot* live: not in the topology or the geometry of the state space itself. But differential structure does something

that topology alone cannot: it tells us where path dependence *does* live, and how to detect it.

To see this, we need to be precise about what "path dependence" means mathematically. Consider a 1-form ω on the tame state space $\mathcal{P}^{[\mathfrak{T}]}$—a field that assigns, at each state, a linear functional on the tangent space. Intuitively, ω encodes the "pressure" that some political or economic process exerts on states as they move through the space. A transition from state π_0 to state π_1 along a path γ accumulates a total effect $\int_\gamma \omega$. Path dependence, in the geometric sense, means that this integral depends on the choice of path: two trajectories from π_0 to π_1 yield different accumulated effects.

Stokes' theorem reduces the question to a local one. The line integral $\int_\gamma \omega$ is path-independent if and only if ω is *exact*—that is, $\omega = df$ for some scalar potential f. And Poincaré's lemma, which applies because the tame state space is convex, guarantees that ω is exact if and only if it is *closed*: $d\omega = 0$. So: compute the exterior derivative of the field. If $d\omega = 0$, the process is path-independent—there exists a potential function, and the accumulated effect depends only on the endpoints. If $d\omega \neq 0$, the process is genuinely path-dependent, and the exterior derivative tells you exactly where and how the path dependence arises.

This is the diagnostic that topology cannot provide. Simple connectedness tells us the space has no holes, so path dependence cannot arise from the *global* structure of the arena. But it says nothing about the *fields* on the arena. Differential structure gives us a calculus for those fields: a way to decompose political processes into their path-independent and path-dependent components, and to locate the sources of path dependence in the non-vanishing of $d\omega$.

Ertman (1997) provides a canonical illustration. Ertman argues that the character of European state formation depended not merely on the geopolitical and fiscal pressures that states faced, but on the *sequence* in which they faced them. States that confronted sustained geopolitical competition *before* developing bureaucratic infrastructure—the paradigmatic case is Prussia—built patrimonial absolutisms, concentrating military and fiscal power in the crown. States that bureaucratized *before* facing intense military pressure—England being the canonical example—developed constitutional regimes with representative institutions constraining executive power. The endpoint pressures were similar; the outcomes diverged because the paths differed.

In the language of this chapter, Ertman's argument is a claim about the non-closedness of a 1-form. The "institutional pressure" field ω that drives state transitions through $\mathcal{P}^{[\mathfrak{T}]}$—encoding, at each state, the direction and magnitude

of the political forces pushing toward military modernization, fiscal extraction, or bureaucratic rationalization—has $d\omega \neq 0$. The accumulated institutional residue of traversing the Prussian path differs from that of traversing the English path, even when the paths share their endpoints in the space of production characteristics. Topology cannot see this: both paths are homotopic, and the space is simply connected. But the exterior derivative can see it. The curl of the institutional field is nonzero, which means the field has no potential function, which means the integral is path-dependent.

This is where the differential structure earns its keep. Pierson (2000) identifies increasing returns as the mechanism of political path dependence: early institutional choices generate self-reinforcing dynamics that make alternatives progressively costlier. Translated into our setting, increasing returns manifest as the non-closedness of the driving field. A field with $d\omega \neq 0$ is precisely one where the "return" to a given transition depends on the transitions that preceded it—the hallmark of Pierson's increasing returns. The formalism does not merely restate the verbal argument; it sharpens it. It says: to determine whether a given political process exhibits path dependence, compute its exterior derivative. If the field is closed, the process has a potential function and the outcome is path-independent regardless of the trajectory. If the field is not closed, the specific location and magnitude of the non-closedness identify exactly which transitions, at which states, generate the path-dependent effects.

The contrast with Chapter 4 is instructive. Contractibility told us that the space itself carries no memory. Convexity told us that the space carries no geometric obstructions. But the differential structure tells us something new: it tells us how to *decompose* the forces acting on the space into a conservative component (the exact part of ω, which has a potential and is path-independent) and a non-conservative component (the non-exact part, which is the source of path dependence). This decomposition—Hodge-like in spirit, though the full Hodge theory requires a metric we have not yet introduced—is the kind of analytical leverage that topology alone cannot provide. The topologist can say "the space has no holes"; the geometer can say "here is the field, here is its curl, and here is where history matters."

5.6 The Water-Filling Equilibrium

We opened this chapter by enriching Game 3.1 into Game 5.1 and finding ourselves at an impasse: existence without characterization. Having introduced

tame functions and established their adequacy and convexity, we may now return to the enriched game and test whether tame states yield a more explicit equilibrium.

The reward for all this work is not a new assumption but a new kind of vision. Once both technologies and costs are tame, the game itself becomes geometrically simple. Where before the equilibrium was a tangle of implicit reactions, we now find a surface that can be described explicitly—and even elegantly.

5.13 Proposition (Water-filling equilibrium)

Suppose Game 5.1 is tame: *for each player $i \in \{1, 2\}$,*

$$\tau_i(x_i) = A_{\tau,i} \sum_{\ell \in L} \beta_{i,\ell} \log(1 + x_{i,\ell}), \qquad \kappa_i(x_i) = A_{\kappa,i} \sum_{\ell \in L} q_{i,\ell} \, x_{i,\ell},$$

with $A_{\tau,i}, A_{\kappa,i} > 0$, $\beta_i, q_i \in \Delta_L$, all entries strictly positive. Then any Nash equilibrium (x_1^*, x_2^*) with strictly positive allocations has the water-filling form: *for each player i there exists a scalar $c_i > 0$ such that*

$$x_{i,\ell}^* = \max \left\{ 0, \ c_i \frac{\beta_{i,\ell}}{q_{i,\ell}} - 1 \right\} \qquad \text{for all } \ell \in L, \tag{5.1}$$

and the scalar c_i solves the one-dimensional fixed point

$$c_i = \frac{(1 - p_i) \, \alpha \, A_{\tau,i}}{k \, A_{\kappa,i}} \, \frac{W(x_1^*, x_2^*)}{e^{\kappa_i(x_i^*)} \, \tau_i(x_i^*)}, \tag{5.2}$$

where p_i is the contest success probability and $W = V - k(e^{\kappa_1(x_1^*)} + e^{\kappa_2(x_2^*)} - 2)$. If all inputs are active, then

$$\tau_i(x_i^*) = A_{\tau,i} \left(\log c_i + \sum_{\ell \in L} \beta_{i,\ell} \log \frac{\beta_{i,\ell}}{q_{i,\ell}} \right), \qquad \kappa_i(x_i^*) = A_{\kappa,i} \, (c_i - 1),$$

so that (5.2) becomes a scalar equation in c_i alone.   [*Proof.*]

The resulting equilibrium allocations echo the "water-filling" solutions of information theory (Cover and Thomas, 1990): each player's resource distribution is governed by a single balancing constant c_i. Inputs receive positive allocations only when their productivity-to-cost ratio exceeds the threshold implied by c_i,

so that each player's resources quite literally *fill up* the most efficient channels first. The multidimensional strategic landscape collapses to a one-dimensional fixed point, and the equilibrium emerges in a water-filling shape: continuous, ordered, and interpretable.

The analogy with information theory is not merely nominal. In the classical water-filling problem of Cover and Thomas (1990), a transmitter allocates power across parallel communication channels to maximize total throughput: channels with better signal-to-noise ratios receive more power, and channels below a cutoff receive none. Here the state allocates resources across commodities to maximize military effectiveness at minimum cost: commodities with better productivity-to-cost ratios receive more investment, and commodities below the cutoff receive none. The ratio $\beta_{i,\ell}/q_{i,\ell}$ is an "efficiency ratio" that plays the role of the channel's signal-to-noise ratio: it measures how much military output a unit of expenditure on commodity ℓ buys, given the state's technology and cost structure. The scalar c_i is the water level: it rises until the total allocation exhausts the budget implied by the strategic equilibrium. States with many efficient channels—states whose technologies reward diversification across cheap inputs—spread their resources widely. States with few efficient channels concentrate their resources in their comparative advantages. The allocation pattern is a fingerprint of the state's production characteristics, rendered visible by the geometry of the tame model.

In the symmetric case, the simplification is complete:

5.14 Corollary

In the symmetric tame case ($A_{\tau,1} = A_{\tau,2} = A_\tau$, $A_{\kappa,1} = A_{\kappa,2} = A_\kappa$, $\beta_1 = \beta_2 = \beta$, $q_1 = q_2 = q$), there exists a symmetric equilibrium with $x_1^* = x_2^* =: x^*$ of the form

$$x_\ell^* = \max\left\{0, \, c\frac{\beta_\ell}{q_\ell} - 1\right\} \qquad \text{for all } \ell \in L, \tag{5.3}$$

where $c > 0$ solves

$$c = \frac{\alpha A_\tau}{2 k A_\kappa} \frac{V - 2k\left(e^{A_\kappa(c-1)} - 1\right)}{e^{A_\kappa(c-1)} A_\tau\left(\log c + \sum_{\ell \in L} \beta_\ell \log \frac{\beta_\ell}{q_\ell}\right)}. \tag{5.4}$$

If $c \geq \max_{\ell \in L} q_\ell/\beta_\ell$ then all inputs are active and (5.3) holds without truncation.

In symmetry, everything condenses to a single number. Both players face the same geometry of tradeoffs, and equilibrium is achieved when this common

scalar c balances their shared marginal returns. The equilibrium surface is perfectly smooth, its contours given by the ratios β_ℓ/q_ℓ: a literal map of efficiency.

The contrast with the wild regime is instructive. In the wild case, the production problem admits multiple solutions, and the enriched game would inherit this multiplicity: multiple equilibria corresponding to different branches of the argmin correspondence, with the selection among them determined by initial conditions and political contingency rather than optimization. The tame game has a clean scalar equilibrium precisely because it inherits the classical regime's uniqueness. The water-filling form is not an accident of functional form choice; it is a consequence of the structural simplicity that the admissibility conditions guarantee and that tamification preserves.

If the aim of this chapter was to find a tractable representation of complex strategic behavior, then we have succeeded. The tame states not only adequately represent the general states but endow the game with a geometry rich enough to admit explicit, interpretable equilibria. These equilibria are not accidents of simplification; they are the natural shapes that emerge when the model is seen clearly. Tameness, in the end, has not narrowed our view: it has brought the whole system into focus.

5.7   Assessment

This chapter has developed a modeling theory for the space of possible states. The pipeline—regularize, tamify, retract—takes an arbitrary technology-cost pair in $\mathcal{T} \times \mathcal{K}$ and produces a canonical tame representative whose policy function faithfully mirrors the original's topological properties. The tame subspace is not merely a convenient simplification but a strong deformation retract—and, as we have argued, a *structurally adequate* model in a precise sense that draws on but goes beyond van Fraassen's (1980) empirical adequacy. The content of adequacy here is not invariant preservation, which contractibility renders trivial, but the existence of a continuous, admissibility-respecting retraction: a process of structural simplification that is well-defined at every stage. This notion of structural adequacy—retractability via admissible paths—may have applications beyond the present setting, wherever one seeks a tractable submodel of a function space without sacrificing the structural relationship between model and ambient.

Within this tame subspace, we discovered convexity—a geometric property that contractibility alone cannot provide. Convexity deepened each rung of

the ladder of sameness established in Chapter 4: connectedness sharpens to convex connectedness, path-connectedness to geodesic connectedness, simple connectedness to the exactness of closed forms. The space of tame states is not just topologically trivial but geometrically interchangeable: paths are straight, interpolations are states, and the arena carries no intrinsic memory.

Together with the wild regime developed in Chapter 6, these results complete a three-level hierarchy—wild, classical, tame—that constitutes the formal architecture of the book. The shadow map of Chapter 6 compresses wild states into their classical silhouettes, discarding branch structure in the process. Tamification compresses classical states into their first-order representatives, preserving all topological information. Between the two compressions lies the full range of what the theory can and cannot model: the wild regime resists faithful representation, the classical regime admits it, and the tame subclass makes it geometrically rich.

Three directions remain genuinely open.

First, what would a faithful model of wild states look like? The tamification pipeline depends on uniqueness of the SPP solution, which the wild regime lacks. The obstacle is not merely technical: the wild regime's branch structure encodes genuinely political information—which mode of force production the state has settled into, and how it might transition between modes—and any faithful model must represent that information, not discard it. Jet spaces, which record not just the value but the derivatives of a correspondence at a point, or sheaf-theoretic constructions that track multiple local sections simultaneously, may offer a path forward; but the question of how to retract a multi-branched correspondence onto a canonical representative remains open.

Second, the tamification operator quotients by first-order structure at the origin. Higher-order tamification—using second derivatives (Hessian structure), or the full Taylor jet at the origin—would yield finer equivalence classes and richer models. A second-order tame representative could distinguish technologies that agree at the margin but differ in how fast diminishing returns set in—a distinction invisible to first-order tamification but potentially consequential for the geometry of allocation. The cost is complexity: each additional order of approximation introduces more parameters and weakens the geometric simplicity that makes the tame model tractable. Whether the additional structure is worth the cost is a question that applications must answer.

Third, the relational structure of states—what individuates France from Germany even when both solve admissible production problems—is a question that the present framework, which studies states in isolation, cannot address.

The space of possible states describes what states *are* as producers of force, but it says nothing about how states relate to one another in a system of mutual recognition and strategic interaction. Resolving it would require embedding each state in a web of relations to other states and recovering identity from relational position rather than intrinsic character—apparatus that lies beyond the topological methods developed here.

These questions remain open. The conclusion returns to them after the wild regime has been examined on its own terms.

Chapter 6

What the Classical Models

Whenever we are able to find rules to generate all the validities, we have the guarantee of having captured the essence of reasoning; i.e., its logic perfume in the most radical and barbarous sense.

María Manzano

The preceding two chapters told a clean story. The classical states—those satisfying the admissibility conditions of Chapter 3—inhabit a contractible space: topologically trivial, admitting no first-order distinctions, no obstructions, no structural memory. The tame states refine this further: a structurally adequate subspace, the coarsest quotient compatible with first-order analysis, carrying not just contractibility but convexity. Within this regime, Waltz is vindicated. States are like units. The admissibility conditions are the rules that, in Manzano's phrase, capture "the logic perfume" of statehood—its essence, distilled through the structure of force production (Manzano, 2005).

But not all states satisfy those rules. What lies beyond those conditions is not a featureless void but a richer landscape—one in which the topology itself generates distinctions, obstructions, and memory that the classical case lacks. We call these states *wild*: they resist the taming that Chapter 5 achieves for the classical regime, and they do so for structural reasons, not technical ones.

This chapter is an initial reconnaissance, not a complete survey. The wild regime is vast: it encompasses the full machinery of singularity theory, stratified Morse theory, and the classification of multi-valued correspondences—a literature that has been developed in mathematics for decades and that has barely been brought to bear on political phenomena. What follows is a first exploration, enough to establish that the structure exists, that it is rich, and that even a cursory

inspection reveals formal phenomena—catastrophes, monodromy, topological path dependence—that are invisible from the classical interior and that reward the effort of going native in the wild regime.

The chapter develops the wild landscape and then turns the structural adequacy criterion of Chapter 5 back on itself. The tame subspace is structurally adequate for the classical space—the retraction is faithful. Is the classical subspace structurally adequate for the full space including wild states? The answer is no. The shadow map that projects wild states into the classical regime is lossy: it preserves the fiscal silhouette while discarding the branch structure that makes the wild topology nontrivial. The asymmetry between the two directions—faithful inward, lossy outward—is the formal content of the tame-classical-wild hierarchy and the central result of this chapter.

6.1 The Production Correspondence

When log-concavity of τ or exp-convexity of κ fails, uniqueness of the SPP solution can fail with it. The production problem may admit two or more global minimizers—distinct mobilization plans that achieve the same force level at the same cost. When this happens, the policy "function" becomes a *correspondence*—a set-valued map

$$\Pi_{\tau,\kappa} : M \rightrightarrows X$$

that assigns to each force target not a single mobilization plan but a set of optimal plans.

More generally, the cost landscape on the level set $\{x : \tau(x) = m\}$ may have multiple *local* minima—mobilization plans that are optimal among nearby alternatives but not globally. Each local minimum is a mode of force production: a qualitatively distinct way to achieve the same force level, stable against small perturbations of the plan but potentially far from the global optimum.

The multiplicity is not a technical inconvenience; it is a political fact. When the production problem has a unique solution, the state's response to a given force target is determined by its technology and cost structure: politics fixes the parameters, and mathematics determines the output. When the problem has multiple solutions—or multiple local optima among which the state might settle—the parameters underdetermine the output. The state must choose, and the choice depends on who holds power, which constituencies bear costs, which

military traditions are entrenched, and which foreign models are emulated. The optimization provides a landscape; politics navigates it.

6.2 An Example: Increasing Returns to Specialization

Consider a state with $L = 2$ commodities and a technology exhibiting increasing returns to concentration:

$$\tau(x_1, x_2) = \left(x_1^2 + x_2^2\right)^{1/2}.$$

This is a CES technology with elasticity parameter $a = 2 > 1$. It satisfies continuity, ray surjectivity, and weak monotonicity, but it violates log-concavity: the level sets $\{x : \tau(x) = m\}$ are quarter-circles, which are concave curves rather than convex ones. Resources are *super-substitutable*—the technology rewards concentration over diversification. Allocating m units entirely to a single commodity yields force level m; splitting equally between two commodities yields only $m/\sqrt{2} \approx 0.71m$. To verify the failure of log-concavity directly: the midpoint of the chord from $(m, 0)$ to $(0, m)$ is $(m/2, m/2)$, where $\tau = m/\sqrt{2}$. So

$$\log\left(1 + \tau\left(\tfrac{m}{2}, \tfrac{m}{2}\right)\right) = \log\left(1 + \tfrac{m}{\sqrt{2}}\right) < \log(1 + m)$$
$$= \tfrac{1}{2}\log(1 + \tau(m, 0)) + \tfrac{1}{2}\log(1 + \tau(0, m)):$$

the function falls below the chord, violating concavity.[1]

Pair this technology with symmetric linear cost $\kappa(x) = x_1 + x_2$. The SPP asks: minimize total cost on the quarter-circle of radius m. Parametrize the level set as $x_1 = m\cos\theta$, $x_2 = m\sin\theta$ for $\theta \in [0, \pi/2]$. The cost becomes

$$\kappa(\theta) = m\left(\cos\theta + \sin\theta\right) = m\sqrt{2}\sin\left(\theta + \tfrac{\pi}{4}\right),$$

which is *maximized* at $\theta = \pi/4$ (equal allocation) and *minimized* at the endpoints $\theta = 0$ and $\theta = \pi/2$. The SPP has two global minima:

$$x^A = (m, 0), \qquad x^B = (0, m),$$

---

[1]More generally, for any $a > 1$, the function $x \mapsto \log(1 + \|x\|_a)$ is quasi-convex rather than quasi-concave: its super-level sets are complements of L^a balls, which are not convex. The failure of log-concavity is not a pathology of $a = 2$ but a generic feature of the $a > 1$ regime.

both with cost m. The state must choose between complete specialization in commodity 1 or complete specialization in commodity 2. The optimization is silent on which.
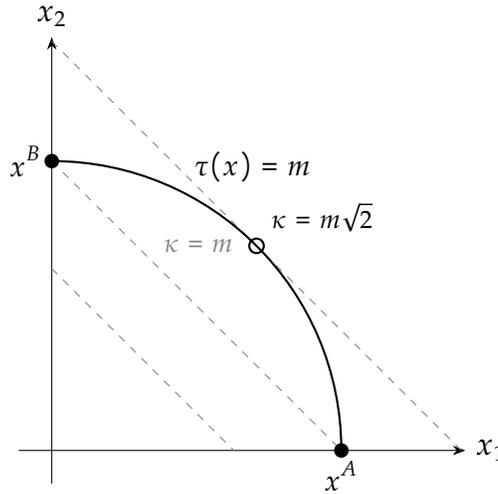
Figure 6.1: *The wild level set. Dashed lines are iso-cost contours ($\kappa = x_1 + x_2$). The quarter-circle (level set of τ) is concave rather than convex, so cost is* minimized *at the corners x^A and x^B (complete specialization) and maximized at the interior point (equal allocation).*

Now perturb the cost function. Let $\kappa_\epsilon(x) = x_1 + (1 + \epsilon) x_2$ for small ϵ. When $\epsilon > 0$, commodity 2 is more expensive, and the unique global optimum is $x^A = (m, 0)$. When $\epsilon < 0$, the unique global optimum is $x^B = (0, m)$. At $\epsilon = 0$, both solutions coexist. The solution jumps *discontinuously* as ϵ crosses zero: an infinitesimal change in the cost structure triggers a complete reorganization of the state's mobilization.

This is a *fold catastrophe*: two branches of solutions exchange global optimality at a critical parameter value. The phenomenon is generic—it persists under small perturbations of the technology—and it is entirely absent in the classical case, where convex level sets ensure a unique, continuously varying solution.

The $L = 2$ example generalizes.

6.1 Proposition

Let $a > 1$, $L \geq 2$, and consider the CES technology $\tau_a(x) = \left(\sum_\ell x_\ell^a \right)^{1/a}$ with linear cost $\kappa(x) = q \cdot x$, $q \in \mathbb{R}_{>0}^L$. Then τ_a satisfies $\mathfrak{C}_\tau$, $\mathfrak{R}_\tau$, and $\widetilde{\mathfrak{M}}_\tau$, but violates $\tilde{\mathfrak{L}}_\tau$. For
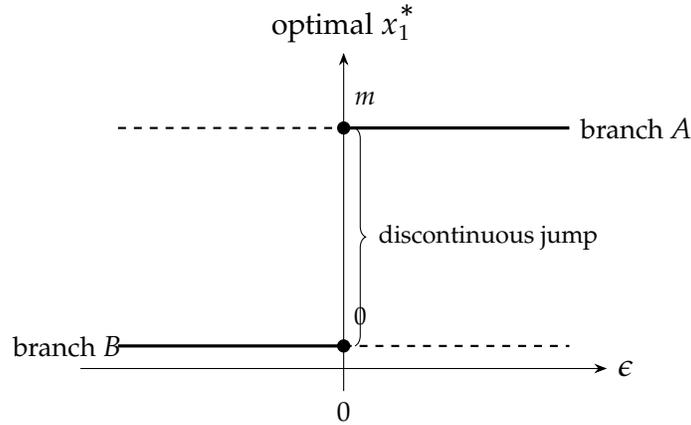
Figure 6.2: *The fold bifurcation in cost space. Solid lines indicate the global optimum; dashed lines indicate the suboptimal branch. At $\epsilon = 0$ both branches are equally optimal and the solution is discontinuous in the cost parameter.*

every $m > 0$:

1. *Every global minimum of the SPP is a vertex solution $x^{(\ell)} = m \cdot e_\ell$ (complete specialization in a single commodity);*

2. *The global minimum is unique if and only if the cheapest commodity is unique: $\left|\text{argmin}_\ell\, q_\ell\right| = 1$; and*

3. *The bifurcation set $\mathcal{B}_a = \{q \in \mathbb{R}^L_{>0} : \left|\text{argmin}_\ell\, q_\ell\right| > 1\}$ is a union of $\binom{L}{2}$ hyperplanes of codimension one.*

[*Proof*.]

The proof turns on a single computation. At the unique interior critical point of $q \cdot x$ on the L^a sphere, the cost is $m \cdot \|q\|_p$, where $p = a/(a-1)$ is the Hölder conjugate of a. Since $\|q\|_p > \min_\ell q_\ell$ for $L \geq 2$, the interior point is suboptimal; by induction on faces, the global minimum lies at a vertex.

The proposition reveals a sharp structural contrast with the classical case. When $a \leq 1$, the level sets of τ_a are convex, the SPP has a unique interior solution that spreads mobilization across commodities, and the solution varies continuously with q. When $a > 1$, the level sets are non-convex, the SPP forces complete specialization in the cheapest commodity, and the solution is discontinuous at the bifurcation set. The transition at $a = 1$ is the formal

content of the classical–wild boundary for this family: the moment at which diversification gives way to forced specialization, and continuous parameter dependence gives way to catastrophic switching.

6.3   The Landscape of Local Optima

The fold is the simplest wild phenomenon. In higher-dimensional commodity spaces ($L \geq 3$), the cost landscape on the level set $\{x : \tau(x) = m\}$ can be far richer. The level set is an $(L-1)$-dimensional manifold, and the cost function restricted to it is a scalar function on that manifold. With non-convex level sets, this function can have many local minima—each a distinct mode of force production, each locally stable, each representing a qualitatively different way to produce the same force level.

The resulting picture is a *rugged landscape*. Each local minimum is a basin of attraction: a region of the mobilization space from which small adjustments lead back to the local optimum. The basins are separated by saddle points— mobilization plans that are critical points of cost but not minima, marking the "mountain passes" between basins. The depth of a basin (the cost difference between the local minimum and the nearest saddle) measures the stability of the corresponding mode of force production. The height of the pass measures the cost of transition: the institutional effort required to move from one mode to another.

Politics, in this picture, operates on two levels. At the first level, it determines which basin the state occupies: initial conditions, historical contingency, and the distribution of power among constituencies select the mode of force production. A state whose military tradition is naval, whose commercial wealth is concentrated in shipbuilding, and whose officer corps is trained in maritime doctrine occupies one basin; a continental state with conscript traditions and land-based logistics occupies another. Both may achieve the same force level; they achieve it in qualitatively different ways.

At the second level, politics determines whether and when the state transitions between basins. A transition requires crossing a saddle point—an institutional transformation costly enough that the state must pass through a less efficient configuration before reaching a more efficient one. This is the formal content of military reform: not a smooth adjustment within a basin but a costly leap across a pass. The decision to undertake such a leap depends on the gap between the current basin's depth and the target basin's depth, on the height of

the pass, and on the political capacity to sustain costs during the crossing. Small perturbations do not induce transitions; only sufficiently large shocks—military defeat, fiscal crisis, regime change—push the state over the saddle.

Paths through the landscape are constrained by the topology. In the classical case, any two states can be connected by a direct interpolation: the homotopies of Section 4.2.2 construct explicit paths through admissible intermediaries. In the wild case, the direct interpolation may not exist—it may cross a bifurcation or leave the admissible set—and the state must find a route *around* the obstruction. Different routes between the same endpoints may cross different saddle points, pass through different intermediate basins, and incur different transition costs. The topology determines which routes are available; the geometry determines which among the available routes is optimal. Finding the optimal transition path amounts to constructing a geodesic in the wild landscape—a problem that requires the topological structure as scaffolding. This is where the topological framework earns its keep. The fact that a naive interpolation fails does not mean no path exists; it means that finding the path requires understanding the global structure of the landscape, and that the path one finds may be topologically constrained in ways that a classical state never encounters.

6.4   Topological Path Dependence

The wild landscape has topological features that the classical landscape lacks.

The most immediate is *disconnectedness*. Near a fold bifurcation, a state that selects branch A and an otherwise identical state that selects branch B are genuinely different elements of the policy space. The space of policy selections— single-valued functions chosen from the correspondence—can have multiple connected components near the bifurcation, one for each branch. This is already a sharp contrast with the classical case, where Corollary 4.7 guarantees a single component.

Richer structure emerges when two control parameters interact. The fold of Proposition 6.1 is a codimension-1 singularity: the bifurcation set $\mathcal{B}_a$ is a hyperplane in cost space, and crossing it exchanges one global optimum for another. The next entry in the classification of elementary catastrophes (Thom, 1975; Arnold, 2004) is the *cusp*: a codimension-2 singularity in which two solution branches coalesce at a degenerate critical point and re-separate with their identities exchanged. The CES family exhibits a cusp whose location and

consequences can be stated precisely.

6.2 Proposition (Cusp catastrophe in the CES family)

Let $L = 2$ and consider the CES technology τ_a of Proposition 6.1 with cost $\kappa_\epsilon(x) = x_1 + (1 + \epsilon)\, x_2$, parametrized by $(a, \epsilon) \in (0, \infty) \times \mathbb{R}$.

1. *For each $a > 1$, the bifurcation set is $\mathcal{B}_a = \{\epsilon = 0\}$, a single point in the ϵ-line. As $a \downarrow 1$, the two solution branches $x^A = (m, 0)$ and $x^B = (0, m)$ coalesce: the cost gap $\kappa(x^B) - \kappa(x^A) = |\epsilon|\, m$ closes, and at $a = 1$ the level sets become convex, the interior minimum becomes the unique global minimum, and the bifurcation vanishes.*

2. *The point $(a, \epsilon) = (1, 0)$ is a* cusp *point: the bifurcation locus $\{a > 1\} \times \{0\}$ terminates there, and the two-parameter family undergoes a cusp catastrophe in the sense of Arnold (2004, Ch. 9).*

3. *Let $\mathcal{U} = \{(a, \epsilon) : a > 1,\ \epsilon \neq 0\}$ be the complement of the bifurcation set in the wild regime. The fundamental group $\pi_1(\mathcal{U})$ is nontrivial: a loop in the (a, ϵ)-plane that encircles the cusp point induces a* monodromy—*continuous tracking of the global minimum around the loop returns a different branch to the starting parameters. The monodromy action is the transposition $x^A \leftrightarrow x^B$.*

[*Proof*.]

The proof of (i) is immediate from Proposition 6.1: the CES level sets are convex for $a \leq 1$ and non-convex for $a > 1$, so $a = 1$ marks the transition from unique interior solution to corner solutions. For (ii), the key observation is that the bifurcation set has a boundary: the half-line $\{a > 1, \epsilon = 0\}$ terminates at $(1, 0)$ because the fold disappears when the level sets become convex. This is the defining geometry of the cusp—a fold locus that terminates at a degenerate point where the two branches merge. For (iii), consider a loop $\gamma(t) = (1 + r \cos 2\pi t,\ r \sin 2\pi t)$ centered at $(1, 0)$ with radius r small enough that γ stays in the wild regime ($a > 1$) except at its leftmost point. Starting at $(1 + r, 0)$ on branch A, the global minimum varies continuously along γ as long as $\epsilon \neq 0$. When γ crosses the bifurcation locus at $\epsilon = 0$ (traversing from $\epsilon > 0$ to $\epsilon < 0$ via the arc through $a < 1$, where the unique minimum interpolates continuously between the branches), the tracking switches from x^A to x^B. Upon return, the state is on the opposite branch.
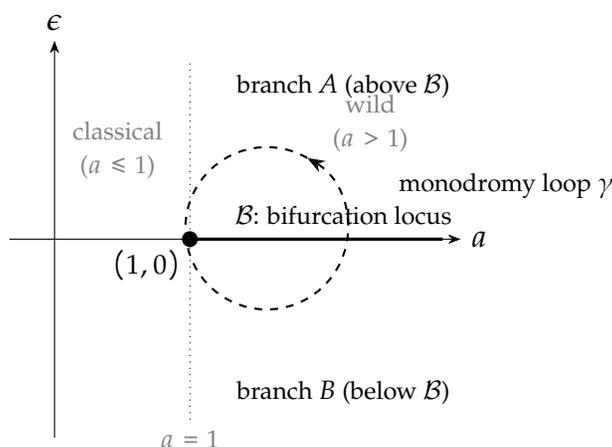
Figure 6.3: *The cusp catastrophe in the (a, ϵ) control plane. The bifurcation locus $\mathcal{B}$ (thick line) is the set of parameters at which the SPP has two equally optimal solutions; it terminates at the cusp point $(1, 0)$, where the wild regime gives way to the classical. A loop γ encircling the cusp point induces monodromy: continuous tracking of the global minimum returns the opposite branch to the starting parameters.*

What this means for a state is worth spelling out. Imagine a state that begins on branch *A*—specializing in commodity 1, with its military organized around that specialization. The state undertakes a sequence of institutional reforms: it adjusts its cost structure, modifies its technology, and then reverses both adjustments, returning to its original parameters. If the loop in parameter space encircles the cusp point, the state arrives back where it started in (τ, κ) space but finds itself on branch *B*—now specializing in commodity 2, with a fundamentally reorganized military. The reforms and counter-reforms have cancelled in parameter space but not in policy space. The state remembers the journey.

This is *topological path dependence* in the precise sense that Section 4.4.3 defined and ruled out for classical states. There exist cycles of institutional reform—loops in the parameter space—such that a state that traverses the cycle and returns to its original technology and cost structure finds itself on a different solution branch: a different mobilization strategy, a different way of being a state. The history of the cycle matters, and it matters irreducibly: no continuous deformation of the loop can eliminate the branch switch, because the cusp point is a topological obstruction. In the classical regime, simple connectedness guarantees that all loops are contractible and no cycle can produce a branch switch. In the wild

regime, the cusp singularity punctures the parameter space and creates loops that cannot be contracted away. The topology *generates* the path dependence; politics merely navigates it.

The connection to the historical-institutionalist literature is direct. Pierson (2000) identifies increasing returns as the mechanism of political path dependence: early institutional choices generate self-reinforcing dynamics that make reversal progressively costlier. The cusp sharpens this: the "increasing returns" are not a metaphor but a consequence of non-convex level sets ($a > 1$), and the irreversibility is not a matter of switching costs but of topology. A reform cycle that encircles the cusp point cannot be undone by retracing its steps, because the monodromy is a topological invariant of the loop class, not a function of the speed or cost of traversal. The cusp point also has a natural reading in Tilly's (1990) framework: it sits at $a = 1$, exactly where the wild regime gives way to the classical, where increasing returns to specialization dissolve into complementarities between inputs. Tilly's convergence on capitalized coercion—the merging of the coercion-intensive and capital-intensive paths—is, in this formalization, the passage through the cusp: the moment at which two branches cease to exist as distinct optima and the bifurcation resolves into a single basin.

Whether such cycles arise in practice—whether real states undergo institutional transformations that trace loops around cusp points in $\mathcal{T} \times \mathcal{K}$—is an empirical question. But the formal possibility is striking: the wild state space is a world in which topology generates memory, and in which the political processes of reform and counter-reform can leave permanent topological traces.

The full classification of wild topologies—which homotopy groups arise, what the generic bifurcation structure of the SPP looks like, whether the monodromy group is constrained by the form of the production problem—connects to deep results in singularity theory and is well beyond the scope of this chapter. But the structural contrast is clear:

	Classical	**Wild**
SPP solutions	unique	multiple
Policy object	function	correspondence
Continuity in parameters	yes	generically no
π_0 (connectedness)	trivial	potentially nontrivial
π_1 (path dependence)	trivial	potentially nontrivial
Landscape	smooth basin	rugged, multi-basin
Contractibility	yes	no

This table records what a first pass through the wild regime has turned up. The fold and cusp are the two simplest entries in Thom's (1975) classification of elementary catastrophes; higher entries—swallowtails, butterflies, umbilic singularities—arise with additional control parameters but are beyond the present scope.[2] What the present chapter establishes is that the state production problem, once stripped of the admissibility conditions that tame it, generates exactly the kinds of singularities that this mathematics was designed to study. The wild regime is not a pathological residual. It is a territory with its own formal structure, and the fold and cusp are its first landmarks.

6.5   The Classical–Wild Boundary

The admissibility conditions mark a topological phase transition in the structure of the state space.

Tilly's (1990) three paths of state formation acquire a new formal reading in this light. The coercion-intensive state and the capital-intensive state may be two branches of the same wild production problem—two locally optimal specializations, one in coercive extraction and manpower, the other in commercial wealth and purchased force. The "choice" between paths is not determined by the optimization; it is determined by initial conditions, political contingency, and the endowment structure that shapes the cost function. States that start in different basins follow different trajectories, even if their technologies and cost structures are similar.

Tilly's convergence on capitalized coercion then admits a striking reinterpretation. As states invest in institutional capacity—building fiscal apparatuses, developing bureaucratic routines, integrating coercive and commercial channels—their technologies become more log-concave: the increasing returns to specialization weaken, complementarities between inputs strengthen, and the production problem transitions from wild to classical. The bifurcation resolves. The multiple basins merge into a single basin—the mixed strategy of capitalized coercion—and the state space contracts from a rugged landscape with topologically nontrivial structure to a smooth, contractible one. Convergence, in this

---

[2] Arnold (2004, Ch. 12) gives the full classification of the seven elementary catastrophes and their normal forms. Each level of the hierarchy introduces richer monodromy groups and more complex stratifications of the bifurcation set. Whether the form of the state production problem constrains which catastrophes can arise generically is an open question that connects to the Thom–Boardman classification of singularities by corank and codimension.

reading, is not merely the triumph of one organizational form but the dissolution of the structural multiplicity that forced the choice in the first place.

The boundary between classical and wild tracks the political boundaries identified in Chapter 3. Besley and Persson's (2011) "common-interest" state—one with high cohesion and invested fiscal capacity—operates in the classical regime, where the production problem has a unique solution and the state space is contractible. Their "weak state"—low cohesion, underdeveloped institutions, fragile extraction—operates in the wild regime, where multiple locally optimal modes of force production coexist and the topology is nontrivial. Talmadge's (2015) coup-proofed dictator, who sacrifices military effectiveness for regime security, may be occupying a suboptimal basin in a wild landscape—a local minimum that is stable (the dictator stays in power) but globally inefficient (the military is degraded).

Formal touchpoint: parameterized cost families. The connection to Besley and Persson (2011) can be made precise. Their framework parameterizes the cost structure by a cohesion parameter $\theta \in [0, 1]$ (how much the winning group shares with the losing group) and an external threat probability $\varphi \in [0, 1]$. These parameters determine the state's investment in fiscal capacity, which in turn shapes the cost of mobilization. Fix a technology τ and let $(\theta, \varphi) \mapsto \kappa_{\theta,\varphi}$ be the continuous map from the political-parameter square into the cost space $\mathcal{K}$ induced by their model of fiscal-capacity investment. Composing with the SPP solution map gives a continuous map

$$(\theta, \varphi) \longmapsto (\tau, \kappa_{\theta,\varphi}) \longmapsto \pi_{\tau,\kappa_{\theta,\varphi}} \in \mathcal{P}.$$

Besley and Persson's three state types—common-interest (high θ, high φ), redistributive (intermediate θ), and weak (low θ, low φ)—correspond to connected regions of the (θ, φ) square. Each region maps, by continuity of the composite, to a connected subset of $\mathcal{P}$. The qualitative typology propagates: if two types occupy disjoint regions of the parameter square and the solution map is injective on their union, the corresponding policy sets are disjoint in $\mathcal{P}$.

Within the classical regime—the common-interest and redistributive regions, where $\kappa_{\theta,\varphi}$ satisfies exp-convexity—the solution map is single-valued and continuous. Small changes in cohesion produce small changes in policy, and the typology shades smoothly: the boundary between common-interest and redistributive is a threshold in (θ, φ) space, not a discontinuity in $\mathcal{P}$. Across the classical–wild boundary—into the weak-state region, where fiscal capacity is too

low to sustain convex costs—the solution map becomes a correspondence, and the propagation changes character. The weak-state region maps not to a single connected arc in $\mathcal{P}$ but to a set of branches among which the state must choose. The political typology, which varies smoothly through the classical regime, encounters a phase transition at the wild boundary: the same continuous parameterization that produces gentle variation above the boundary produces discontinuous branching below it.

The general lesson: the theory does not generate qualitative distinctions internally—contractibility rules that out. But it can *transmit* them. Any partition of the parameter space $\mathcal{T} \times \mathcal{K}$ motivated by political theory, comparative politics, or institutional analysis maps forward to $\mathcal{P}$ via the solution map, and the topology determines what survives the transmission: connectedness of each piece (always), disjointness (when the map is injective), and continuity of variation within each piece (in the classical regime only). The formalism is modular: it handles the propagation and is silent on which partitions matter. The substantive political argument supplies the partition.

Classical states inhabit a world in which all variation is second-order and all path dependence is geometric. Wild states inhabit a world in which the topology itself generates distinctions, obstructions, and memory—a world in which the shape of the state space is no longer trivial, and in which the political processes that navigate that space leave traces that the topology cannot erase.

6.6   The Classical Shadow

Can classical states *represent* wild ones? The classical regime has a single basin, continuous parameter dependence, and contractible topology. If a wild state could be faithfully encoded as a classical one, the wild phenomena of the preceding sections would be representational artifacts rather than structural features. Is there such an encoding?

There is, but it is lossy. Given a wild production pair (τ, κ)—one that violates the admissibility conditions—define the *cost function*

$$c_{\tau,\kappa}(m) = \min_x \kappa(x) \quad \text{s.t. } \tau(x) = m.$$

This is the value function of the SPP: it records, for each force target m, the minimum cost of achieving it. Even when the minimizer is non-unique—even when the production correspondence $\Pi_{\tau,\kappa}$ is set-valued—the cost function itself

is single-valued. It is an invariant of the pair (τ, κ) that survives the passage from correspondence to function.

The formal guarantee is Berge's (1963) Maximum Theorem. Under continuity of κ and appropriate continuity of the constraint correspondence $m \mapsto \{x : \tau(x) = m\}$, Berge gives two conclusions:

1. The value function $c(m)$ is continuous; and

2. The argmin correspondence $\Pi(m)$ is upper hemicontinuous (UHC) and compact-valued.

Both conclusions hold regardless of whether the production pair is classical or wild. Upper hemicontinuity means that solutions cannot escape: if a sequence of force targets $m_n \to m$ is paired with optimal allocations $x_n \in \Pi(m_n)$, every cluster point of $\{x_n\}$ lies in $\Pi(m)$. The solution set is well-behaved even when it is set-valued.

What Berge does *not* give, in general, is *lower* hemicontinuity (LHC) of the argmin. Lower hemicontinuity would mean that every solution $x \in \Pi(m)$ can be tracked continuously: for every sequence $m_n \to m$, there exist $x_n \in \Pi(m_n)$ with $x_n \to x$. In the classical case, convexity of the level sets ensures a unique minimizer, and a single-valued UHC correspondence is automatically continuous—so LHC holds trivially, the policy function is continuous, and the shadow adds nothing. In the wild case, non-convexity produces multiple minimizers, and LHC fails precisely at bifurcation points. Return to the example of Section 6.2: at $\epsilon = 0$, the argmin is $\Pi = \{x^A, x^B\}$, but for $\epsilon > 0$, $\Pi = \{x^A\}$ alone. The branch x^B belongs to the solution set at $\epsilon = 0$ but cannot be approximated by solutions at nearby $\epsilon > 0$—this is exactly the failure of LHC, and it is what makes continuous selection from the correspondence impossible. No single-valued function chosen from Π can be continuous at the bifurcation.

The failure of LHC is not incidental but structural: it is the mechanism by which non-convexity of the level sets propagates into the topology of the policy space. Berge guarantees that the *set* of solutions is well-behaved (UHC); convexity would further guarantee that any *selection* from that set is well-behaved (LHC, hence continuity). Without convexity, the set is tame but the selections are wild. The shadow map exists because Berge's first conclusion—continuity of the value function—is unconditional. It is lossy because Berge's second conclusion—UHC of the argmin—is weaker than the full continuity that convexity would provide.

With this in hand: $c = c_{\tau,\kappa}$ is a well-behaved cost function in its own right. Berge guarantees it is continuous; monotonicity of κ ensures it is increasing. The question is whether c is well-behaved *enough*—whether it satisfies the admissibility conditions of Definition 3.6 and therefore belongs to $\mathcal{K}$. If it does, the pair (id, c)—identity technology with cost c—is a classical production pair, and the wild state acquires a classical shadow.

The desideratum is therefore $c_{\tau,\kappa} \in \mathcal{K}$: the value function of the wild SPP is itself an admissible cost function. Continuity, monotonicity, and coerciveness come free from Berge and the properties of the original problem. The binding condition is exp-convexity: is the map $m \mapsto \exp(c(m))$ convex? This is not automatic. A sufficient condition is that κ itself be convex—not merely exp-convex, but ordinarily convex—since the infimum of a convex objective over constraint sets that are nested in the appropriate sense inherits convexity, and convexity implies exp-convexity.[3] The examples of this chapter all satisfy this sufficient condition: the CES technology is paired with linear cost, and the Besley–Persson parameterization produces convex cost structures. But convexity of κ is stronger than what the shadow construction actually requires. What is needed is $c \in \mathcal{K}$; convexity of κ is one route to that conclusion, not the only one. Whether weaker conditions on κ—exp-convexity alone, or some intermediate regularity—suffice to place c in $\mathcal{K}$ is an open question, and one whose answer would determine the full reach of the shadow construction.

For now, we restrict attention to the τ-*wild* regime: production pairs (τ, κ) whose technologies violate log-concavity but whose value functions $c_{\tau,\kappa}$ are admissible. Within this regime, the pair (id, c) is classical, and its unique policy function sends m to m itself, at cost $c(m)$. The *classical shadow* of the wild state (τ, κ) is the classical state $(\mathrm{id}, c_{\tau,\kappa})$.

6.3 Definition (Classical shadow map)

Let $\mathcal{W}_\tau \subset \mathcal{T} \times \mathcal{K}$ denote the τ-wild pairs for which $c_{\tau,\kappa} \in \mathcal{K}$. The shadow map *$\sigma : \mathcal{W}_\tau \to \mathcal{T}_{\mathrm{cl}} \times \mathcal{K}_{\mathrm{cl}}$ sends each such pair to its classical shadow: $\sigma(\tau, \kappa) = (\mathrm{id}, c_{\tau,\kappa})$.*

The shadow preserves fiscal behavior: it records exactly what the state can produce and what it costs. An external observer who sees only how much force a state can generate at each price point—who knows the cost function but not the

---

[3]Recall from Section 3.5.2 that exp-convexity is *weaker* than convexity: it permits cost functions that are concave, provided they are not too concave. Requiring κ to be convex is therefore *stronger* than what admissibility asks. The wild pathology lives in the *argmin*, not in the *value*.

mobilization plan—cannot distinguish the wild state from its classical shadow. In the language of the SPP, the shadow preserves the *what* (force-for-cost) while discarding the *how* (the allocation across commodities).

What is lost is the branch structure. The wild state (τ, κ) may achieve cost $c(m)$ by multiple qualitatively different allocations—the branches x^A, x^B of Section 6.2—each with its own political constituency, institutional history, and reform cost. The shadow collapses all of these to a single point. The monodromy of Section 6.4—the topological memory that tracks which branch a state occupies after a cycle of reform—vanishes in the shadow. Two states that differ only in their branch selection, and therefore inhabit different connected components of the wild policy space, map to the same classical shadow.

The shadow map σ is continuous but not injective, and the information it discards is precisely the information that makes the wild topology nontrivial. It cannot be a homotopy equivalence: it maps a space with nontrivial π_0 or π_1 into a contractible space, necessarily killing nontrivial homotopy classes. In categorical terms, the shadow is a *left adjoint to the inclusion* of classical states in the full state space—it is the best classical approximation, but it is a genuine approximation, not a faithful representation.

6.7 Structural Adequacy, Revisited

Chapter 5 established that the tame subspace is structurally adequate for the classical space: there exists a strong deformation retraction—continuous, admissibility-respecting at every stage, fixing the tame functions throughout—from the ambient classical space onto the tame subspace. The three desiderata of structural adequacy—continuity, path-fidelity, and fixity—are all satisfied.

We can now ask the same question one level up. Is the classical subspace structurally adequate for the full space of states, including the wild ones?

The shadow map provides a retraction of sorts: it sends every production pair to a classical one, and it fixes classical pairs (for which the shadow is the identity). But it fails the criterion of structural adequacy on the second desideratum: *path-fidelity*.

A structurally adequate retraction requires a continuous path $H(f, t)$ from each function to its representative, with every intermediate function satisfying the admissibility conditions. But the admissibility conditions are precisely what wild states violate. There is no admissible path from a wild state to its classical shadow, because the wild state itself is not admissible—its technology is not

log-concave, its cost is not exp-convex, and the SPP does not have a unique solution. The retraction cannot even begin within the admissible space, because the starting point lies outside it.

The tamification pipeline of Chapter 5—regularize, then quotient—worked because regularization smoothed rough-but-admissible functions into smooth-and-admissible ones. The key word is "admissible": the rough functions already satisfied the structural conditions; they merely lacked differentiability. Wild functions fail the structural conditions themselves. No amount of smoothing can make a non-log-concave technology log-concave; the pathology is not roughness but shape.

The result is the asymmetry that organizes the full hierarchy:

Level	Map	Section	Fidelity
Wild → Classical	Shadow map σ	Section 6.6	Lossy
Classical → Tame	$\mathfrak{T} \circ \mathfrak{D}$	Section 5.4	Faithful

Lossy means that the shadow map fails path-fidelity: there is no continuous, admissibility-respecting deformation from wild states to their classical representatives. The shadow preserves the fiscal silhouette—the cost function $c(m)$—but destroys the branch structure, the monodromy, and the topological memory that the wild regime carries. *Faithful* means that tamification satisfies all three desiderata of structural adequacy: the passage from classical to tame is a genuine structural simplification, well-defined at every stage, with no information lost.

The whiteboard is a perfect model of the circle drawn on it. But it is not a perfect model of the room it is in.

6.8   Contractibility as Attractor

The strongest reading of the wild landscape is this: the admissibility conditions of Chapter 3 are not modeling assumptions but the residue of a selection process. Early states inhabited the wild landscape—multiple locally optimal modes of force production, rugged topologies, path dependence woven into the structure of the problem. Competition among states selected for classicality: states that developed complementarities between inputs (strengthening log-concavity), that built fiscal institutions distributing costs smoothly (strengthening exp-convexity), that integrated coercive and commercial channels (resolving bifurcations)—these states accessed a smooth, single-basin landscape in

which the global optimum varies continuously with parameters. They became robust. States trapped in wild local optima—overspecialized, brittle, subject to catastrophic reorganization under small parameter shifts—were selected out or absorbed. On this reading, contractibility is not a theorem about a model but a theorem about what survives: the shape of the attractor toward which the historical process of state formation converges. The wild section describes the transient; the classical results describe the steady state.

This is a seductive argument, and there are at least three reasons to resist it.

The first is *survivorship*. We observe modern states that satisfy the admissibility conditions and infer that competition selected for those conditions. But this reads the selection mechanism off the outcome—a just-so story in which the model's assumptions are repackaged as evolutionary achievements. Much of the classicalization of states was not competitive selection but imperial imposition: European institutional forms were transplanted through colonialism, not discovered independently through trial and error. The convergence may be an artifact of hegemony, not fitness.

The second is the *Cartwright objection*, applied reflexively. The wild landscape is wild *relative to the model*. If the admissibility conditions are precisely those under which the SPP has a unique, continuously varying solution, then of course violating them produces pathology—that is true of any model's regularity conditions. The "instability" of wild states may be an artifact of the formalism rather than a feature of the world: the model breaks when its assumptions fail, and we are mistaking the model's limitations for the world's dynamics.

The third is the *persistence of the wild*. Multiple basins need not be unstable. Talmadge's (2015) coup-proofed dictator occupies a suboptimal basin indefinitely—the military is degraded, but the regime endures, because the dictator's objective function includes regime security as a component of κ that the analyst's welfare function does not. Besley and Persson's (2011) weak states are not transients on their way to classicality; they are equilibria of a political game with low cohesion. The wild landscape may have its own attractors, its own stable configurations, its own logic of persistence that the classical framework cannot see precisely because the shadow map collapses the branch structure that sustains them. If so, then "contractibility as attractor" is not a universal claim about state formation but a parochial claim about a particular historical trajectory—the European one—dressed in topological language.

The tension is productive and should not be resolved prematurely. The formal content of this chapter is unconditional: *if* the admissibility conditions fail, *then* the state space acquires nontrivial topology. The shadow construction

shows exactly what information is gained and lost in the passage between the two regimes. Whether the historical trajectory from wild to classical is driven by selection, imposition, contingency, or some combination is an empirical question that the topology frames but cannot answer. What the topology *can* say is that the two regimes are structurally different—that the passage between them is not a quantitative adjustment but a qualitative phase transition—and that any account of state formation that treats the admissibility conditions as obvious or natural is eliding the most interesting part of the story.

Chapter 7

Conclusion

Dropping either convexity or monotonicity from the picture leads to relatively impoverished theories.

Yves Balasko

7.1   The Price of Contractibility

The central result of this book—that the space of possible states is contractible—is, in a precise sense, its least interesting finding. Contractibility follows from the admissibility conditions by a standard argument: if the technology and cost spaces are individually contractible, and the policy map is continuous, the product structure propagates the deformation. The proof is clean, and unsurprising to anyone trained in algebraic topology. What is surprising is not the theorem but its premises.

The admissibility conditions—log-concavity of the technology function, exp-convexity of the cost function—are not technical conveniences adopted for tractability. As Sections 3.4.4 and 3.5.4 argued at length, they encode specific claims about the political economy of force production. Log-concavity says that force production exhibits diminishing returns: the hundredth tank adds less coercive capacity than the tenth, bureaucratic structures absorb resources with declining marginal efficiency, and the organizational capacity of the military saturates before the budget does. Exp-convexity says that extraction costs accelerate: each additional unit of mobilization is politically, fiscally, and macroeconomically costlier than the last, because the institutions of extraction

125

face limits that compound with scale. These are not mathematical hypotheses. They are claims about what it takes, institutionally and politically, for a state to produce force in a sustained, flexible, well-organized way.

In general equilibrium theory, the analogous admissibility conditions on consumer preferences—convexity, monotonicity, continuity—serve a specific purpose: they make agents comparable, so that equilibrium analysis can proceed across the system rather than one agent at a time. Balasko (2011) observes that dropping them impoverishes the theory. For states, the parallel is exact but the conclusion is inverted. Dropping the admissibility conditions does not impoverish the theory of states. It enriches it: the wild regime carries fold catastrophes, cusp singularities, monodromy, rugged energy landscapes, and topological path dependence that the classical regime rules out. Chapter 6 developed this structure in detail. But the enrichment is individual, not systemic. Each wild state has its own topology, its own branch structure, its own catastrophe set. There is no common framework for rendering them comparable. The admissibility conditions are not what make individual states interesting—the wild states are more interesting—but what make states comparable to one another. They are the price of systematic theory.

What "comparable" means here is precise. In the contractible space, any state can be continuously deformed into any other without leaving the space. One can interpolate between states, define neighborhoods, speak of proximity and gradation. The topological triviality of the classical regime is what makes it possible to treat states as a class: to place them on a common space, to ask how one differs from another by degree rather than by kind, and to build a theory in which variation is parametric rather than structural. In the wild regime, this collapses. Each wild state may have a different number of locally optimal branches, a different catastrophe set, a different monodromy group. One can study any one of them with great formal precision—catastrophe theory and singularity theory provide the tools—but one cannot place them on a shared manifold and compare. The very features that make them structurally interesting are the features that make them structurally incomparable.

This is not a deficiency to be repaired; it is a structural feature of the production problem. Richness and comparability are in tension—not because of a modeling choice but because of the mathematics. The classical model does not simplify by discarding interesting detail; it restricts attention to the regime where systematic comparison is possible at all. The wild model does not complicate by adding noise; it reveals structure that the requirement of comparability suppresses. The book set out to formalize Waltz's claim that states are like

units, and it did. But in doing so it discovered that the most interesting part of the theory is not the likeness—which, once the conditions are in place, is automatic—but the phase boundary where likeness breaks down and individual structure begins.

There is a further asymmetry. In general equilibrium, convexity and monotonicity are properties of consumer preferences, typically taken as given—part of the furniture of the model. In the state production problem, the analogous conditions are not given but *earned*. Log-concavity depends on the organizational capacity of the military: the ability to integrate diverse inputs, delegate command, and sustain operations without hard indivisibilities that would create step functions in the technology. Exp-convexity depends on the fiscal infrastructure of the state: the tax administration, credit institutions, and political mechanisms for distributing the burden of extraction that Besley and Persson (2011) and Tilly (1990) describe. A state earns its way into the comparable regime by building the institutions that make its production problem well-behaved. Comparability is not free; it is an institutional achievement, and the achievement is historically contingent, politically contested, and unevenly distributed.

Waltz argued that in any structural theory, units must be like units: functionally undifferentiated, distinguished only by their capabilities, different only in degree. This was not an empirical claim about states but a theoretical requirement: systematic comparison demands it. He was right—within the classical regime, states are topologically indistinguishable, and the requirement is met. But Waltz could state the requirement without being able to say what it demands of the units that satisfy it. The "what" matters, because without it the requirement is unfalsifiable: if likeness is an axiom, it cannot be tested, and the theory is immune to counterexample by construction. This book supplies the mechanism. States are like units because their force-production problems satisfy admissibility conditions that generate a contractible policy space. The likeness is not a modeling convenience but a topological consequence of the production problem's structure. And the structure is politically constituted: it is the institutional achievement of states that have built the fiscal infrastructure, organizational capacity, and complementary input channels required for smooth, sustained extraction and mobilization.

"Like units" is therefore not a starting point but an endpoint—the culmination of a historical and institutional process whose residue is encoded in the admissibility conditions. The question this book answers is not whether states are alike (they are, given admissibility) but what it costs them to be alike. And the answer is: admissibility. The conditions themselves are where the political

science lives.

7.2   Faithful Inward, Lossy Outward

The admissibility conditions do not merely partition the space of possible states. They define a phase boundary. On one side, the classical regime: contractible, path-connected, carrying no intrinsic topological memory. On the other, the wild regime: disconnected components, fold and cusp catastrophes, monodromy, and path dependence woven into the structure of the production problem itself. Chapter 6 established the formal content of this distinction. Here I draw out its consequences for how we study states.

The shadow map of Section 6.6 projects wild states onto their classical silhouettes. It preserves the cost function—the fiscal envelope of what the state spends to produce force—but it destroys the branch structure: which mode of force production the state occupies, whether it can transition between modes, and what the topological cost of reorganization would be. Any theory that operates within the classical framework implicitly applies this projection. The question is what is lost.

Consider the study of state fragility. When scholars model fragile states using the same conceptual apparatus as consolidated ones, they are working within the classical silhouette. The models capture the fiscal dimension: how much it costs the state to produce a given level of force, and how those costs compare across states. What the models cannot see is the branch structure of the production correspondence. If a state's production problem violates admissibility, it may occupy one of several locally optimal modes of force production—a militia-based mode, a patron-client mode, a warlord mode—and the cost of transitioning between modes may involve catastrophic reorganization that the classical framework rules out by construction. The shadow map flattens this landscape into a single cost curve. The state looks inefficient; it may actually be trapped.

Several literatures describe phenomena that are structurally consistent with what the topology predicts when admissibility fails. Herbst (2000) asks why certain states never consolidated in the European pattern, and answers in terms of the absence of the competitive pressures that drove institutional investment in fiscal and coercive capacity. The topology suggests a specific way to frame this: the question is whether the production problems of such states satisfy admissibility—whether their technologies exhibit diminishing returns and their

cost structures are exp-convex—or whether hard indivisibilities, increasing returns to coercion, or the absence of smooth fiscal infrastructure place them in a regime where the production problem has a different structure entirely. Whether any particular state's production problem violates admissibility is an empirical question, not a classification to be read off regime type or geography. But to the extent that the phenomena Herbst describes reflect violations of admissibility, placing these states on the same strong-to-weak continuum as consolidated ones applies the shadow map: it preserves the fiscal silhouette but discards the branch structure that makes their situations structurally distinct. Jackson (1990) identifies a related gap from the other direction. His "quasi-states" possess juridical sovereignty—the international system treats them as members of the common manifold—but the institutional conditions that admissibility encodes may not obtain. The classical silhouette is their legal status; the topology they actually inhabit may be quite different.

Talmadge's (2015) analysis of coup-proofing suggests what the wild regime might look like from the inside. If the coup-proofed military occupies a specific branch of a wild production correspondence—one in which command is centralized, initiative is suppressed, and effectiveness is deliberately degraded to prevent the military from threatening the regime—then the classical model sees only the aggregate cost: the state spends more for less force. What it cannot see is that the suboptimality may be deliberate, that the branch may be locally stable under the dictator's actual objective function, and that the cost of leaving it—decentralizing command, restoring meritocratic promotion—would pass through the fold catastrophe that the dictator is rationally avoiding. Reno's (1998) warlord states suggest the same structure from the other end of the coercive spectrum: warlord politics as a distinct *mode* of force production rather than a failure of it, occupying its own branch of the production correspondence, locally stable under its own logic of resource extraction and patronage. The persistence of apparently suboptimal military organizations need not be a puzzle to be explained by the classical framework. It may be a structural feature of the wild regime that the classical framework is constitutively unable to represent.

This is not a criticism of scholars who work within the classical framework. It is a diagnosis of the framework's structural limits—limits that are inherited, not chosen, and that become visible only when the topology of the production problem is taken seriously. The asymmetry established in Chapter 6—faithful inward, lossy outward—means that the classical model is reliable about the domain it claims, and structurally blind outside it. The blind spots are not random. They are precisely the states whose production problems violate

admissibility: the states with increasing returns, hard indivisibilities, multiple locally optimal modes, and the topological memory that branch structure carries. These are, by and large, the states that the international relations literature finds most difficult to theorize.

7.3 Objections

Three objections to the book's central result deserve direct answers.

"You proved what Waltz assumed." This is the structural realist's response, and it is half right. The contractibility theorem does vindicate Waltz's claim that states are like units. But the proof changes the claim's epistemic status. An assumption is unfalsifiable—it holds by fiat, and no evidence can dislodge it. A theorem has premises, and the premises can be tested. The admissibility conditions are empirically assessable claims about the structure of force production: that technologies exhibit diminishing returns, that extraction costs accelerate with scale. If these conditions fail for a given state, the theorem does not apply, and the book says exactly what replaces it: the wild regime, with its branch structure, catastrophes, and topological path dependence. Waltz's assumption immunized the theory from counterexample. The theorem exposes it to refutation. It also makes the counterexamples interesting: instead of saying "some states aren't like units" (which Waltz could absorb as noise), the framework identifies what structural feature of the production problem makes them unlike, and what topological consequences follow.

"Force production is only one slice—states do much more." This is the constructivist objection, and it comes in two forms. The first is that the model is too narrow: states tax, regulate, legitimate, and build, and force production is only one of many functions. This is correct, and Section 7.4 addresses it directly: the method generalizes, and each state function defines its own production problem with its own admissibility conditions. Force production is the instance developed here, not the limit of the approach.

The second form is deeper. The constructivist's real claim is that identity is relational—that what makes a state the state it is depends on its position in a system of interactions, not on any intrinsic property (Wendt, 1999). The book does not contest this. Contractibility is a statement about the space, not about individuation within it. France and Germany both solve admissible

production problems; they are topologically alike; they are not the same state. What distinguishes them is their position in a system of alliances, threats, trade dependencies, and mutual recognition—relational structure that the topology of the production problem does not encode. The individuation residual is the constructivist's point, restated in topological terms. The book does not dismiss it; it shows where, formally, that claim has to live: outside the production-problem topology, in a relational apparatus that this framework deliberately excludes.

"Topology cannot be tested." The admissibility conditions can. Log-concavity of τ is a claim about diminishing returns in force production—assessable, in principle, from production-function estimation. Exp-convexity of κ is a claim about the acceleration of extraction costs—assessable from fiscal data on the marginal cost of mobilization. The phase boundary between the classical and wild regimes is not a metaphysical line but an empirically locatable one: it is the locus in parameter space where the production problem's solution transitions from unique to non-unique. The wild regime makes specific predictions: multiple locally optimal modes of force production, hysteresis in military organization, catastrophic transitions between organizational forms. These are observable, at least in principle. The applications in Section 7.2—Herbst's unconsolidated states, Talmadge's coup-proofed armies, Reno's warlord polities—gesture at what the evidence might look like. The challenge is not that the theory is untestable but that the tests require production-function and cost-structure data at a level of disaggregation that existing datasets do not provide. That is a data problem, not a theory problem.

7.4   Beyond Force Production

This book develops the force-production instance of a general method. The method is: define a state function as a constrained optimization problem, impose regularity conditions on the technology and cost structure, and study the topology of the resulting policy space. Force production is a natural starting point—the bellicist tradition identifies it as the one activity all states share, and the international relations literature delivers a force target as the relevant variable—but it is not the only state function to which the method applies.

The fiscal state. A state that extracts revenue solves a production problem: it combines administrative capacity, compliance infrastructure, and political

capital to produce a target level of revenue, subject to a technology (the mapping from fiscal inputs to revenue extracted) and a cost structure (the political resistance, deadweight loss, and capital flight that extraction entails). What would admissibility look like? Log-concavity of the fiscal technology would mean diminishing returns to fiscal investment: the first tax collector extracts more marginal revenue than the hundredth; the first property registry yields more compliance than the second. This is plausible for consolidated states with functioning administrations—the kind of invested fiscal capacity that Besley and Persson (2011) describe. But it might fail where increasing returns dominate: network effects in tax enforcement (monitoring one firm reveals its trading partners), hard indivisibilities in institutional capacity (a functioning tax court, or none at all), or threshold effects in state legibility (Levi, 1988). If the fiscal technology violates log-concavity, the fiscal production problem enters its own wild regime—with multiple locally optimal modes of extraction, hysteresis in fiscal capacity, and a phase boundary that may fall in a different place than the force-production boundary.

The regulatory state. A state that enforces compliance solves a similar problem: inputs are information, enforcement capacity, and legislative authority; the output is a target level of compliance; the cost is the political, economic, and administrative burden of enforcement. Diminishing returns are plausible for regulatory effort (the first inspection catches the worst violators; the hundredth catches edge cases), but violation is equally plausible where regulatory capture creates increasing returns to incumbent enforcement or where threshold effects in institutional capacity make regulation all-or-nothing. The sketch is deliberately brief. The point is not to prove contractibility for the regulatory state but to show that the framework travels: the method defines the question, and the admissibility conditions determine the answer.

The product topology. The shape of possible states is not one-dimensional. A state that produces force, extracts revenue, and enforces regulation simultaneously occupies a point in the product of three policy spaces—one for each function. Each space has its own admissibility conditions, its own phase boundary, its own classical and wild regimes. If the fiscal space is contractible but the force space is not, or vice versa, the product carries structure that neither component reveals alone. A state can be classical along one dimension and wild along another: smooth and well-organized in its extraction but catastrophically

structured in its coercion, or the reverse. The full shape of possible states is the product of all such spaces, and each new dimension may introduce new phase boundaries, new regions where comparability holds or fails. This is an invitation. Whoever thinks they know what states do next—tax, borrow, protect, redistribute, build—can define the production problem and ask what it does to the shape.

What the framework enables. To make the invitation concrete, consider three questions that the framework makes answerable and that were not well-posed before it.

A comparativist studying state capacity can now ask: *is the comparison between these two states structurally warranted?* If both satisfy the admissibility conditions, they inhabit the same contractible space, continuous paths connect them, and variation between them is parametric—a matter of degree. If one violates admissibility, placing it on the same continuum as the other applies the shadow map. The comparison is not wrong, but it is lossy in a precise and diagnosable way: it preserves the cost envelope and discards the branch structure. The framework tells the analyst not only *that* something is lost but *what*.

A scholar of military organization can now ask: *is this organizational form a point on a shared continuum or a distinct branch of the production correspondence?* If a coup-proofed army, a warlord militia, and a professional military are branches of a wild production problem, they are not ranked by capacity but separated by catastrophe sets. The question "how much capacity does this state have?" gives way to "which branch does it occupy, and what is the topological cost of leaving it?" This reframes persistence: the organization persists not because the state lacks the resources to reform but because reform requires crossing a fold.

A scholar of state-building can now ask: *is this intervention moving a state along a branch or across one?* Parametric change within a branch is smooth—more funding, better training, incremental institutional investment. Change across branches is catastrophic—it passes through a discontinuity where the locally optimal mode of force production jumps. The phase boundary tells you which kind of change you are attempting, and the topology tells you whether the target branch is reachable from the current one without leaving the admissible regime. These are not metaphorical distinctions. They are structural features of the production problem, and they determine whether the tools of incremental policy reform are adequate or whether the intervention requires navigating a singularity.

7.5 The Shape of Possible States

The title of this book names its object. Not particular states—not France or Somalia, not Tilly's coercion-wielding early moderns or Waltz's functionally identical security-seekers—but the space of all states that *could* exist, given the structure of force production. The space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is a mathematical object: a set of functions, equipped with a topology, carrying a geometry. It has a shape. The question was what that shape is.

The shape is not flat. It has a classical interior: contractible, convex under tamification, carrying no intrinsic memory or division. Within this interior, any state can be continuously deformed into any other; paths between states are homotopic; the ladder of sameness—connected, path-connected, simply connected—holds at every rung. This is the regime where systematic theory is possible, where states can be compared, where "like units" is a theorem rather than an assumption.

The shape has a boundary: the admissibility conditions, where log-concavity and exp-convexity give way to increasing returns, hard indivisibilities, and the rugged landscapes of the wild regime. At the boundary, the topology changes—not gradually but through phase transition. Branches multiply, catastrophes appear, monodromy winds paths around singularities, and the space acquires the memory and division that the classical interior lacks.

And the shape has an exterior: the wild regime, structurally richer than the interior, carrying fold and cusp catastrophes, multiple basins, topological path dependence. The exterior is not formless—catastrophe theory and singularity theory give it precise structure—but the structure is individual, not systemic. Each wild configuration has its own topology. The exterior cannot be faithfully projected onto the interior; the shadow map that attempts this destroys exactly the features that constitute the exterior's distinctive character.

The shape, then, is stratified: a smooth, contractible core, surrounded by a structurally rich periphery, separated by a phase boundary whose location depends on the political economy of force production. The tame model sits inside the core as a convex representative—structurally adequate, losing no topological information. The classical regime sits inside the full space as a contractible representative—faithful inward, but unable to see past the boundary that defines it. The wild regime sits outside, casting lossy shadows inward, its structure visible only to theories willing to work outside the classical manifold.

To have mapped this shape is to have answered Waltz's question—but also to have replaced it. "Are states like units?" becomes: "What is the shape of the

space within which states are possible?" The answer is richer than the question. The space is not merely contractible or not; it is stratified, with a phase boundary whose location encodes the institutional conditions under which force can be produced with unique, continuously varying solutions. The contractible interior is the domain where systematic comparison works. The wild exterior is the domain where it does not—not because the theory fails, but because the topology of the space itself defeats the comparability that systematic theory requires.

The contribution of this book is the space itself. Not the contractibility theorem alone, not the tame model alone, not the wild regime alone, but the full object: the space of possible states, its shape, its stratification, and the phase boundary that organizes it. The shape is prior to the inhabitants. It determines what kinds of states can exist, how they can change, what transitions are available to them, and where the passage between configurations is smooth, catastrophic, or impossible. To know the shape of possible states is to know what statehood permits and what it forecloses.

Appendix A

Mathematical Preliminaries

This appendix collects the mathematical background needed to follow Chapters [4]–[6]. The reader is assumed to be comfortable with real analysis on $\mathbb{R}^n$—continuity, convergence, compactness—and with constrained optimization at the level of Sundaram ([1996]): Lagrangians, Karush–Kuhn–Tucker conditions, and the basics of upper and lower hemicontinuity.

What the reader is *not* expected to know is topology beyond $\mathbb{R}^n$: function spaces, homotopy, contractibility, fundamental groups. These are the tools that make the book's central results precise—that the space of policy functions is contractible (Proposition [4.6]), that the tame subspace is a strong deformation retract (Proposition [5.9]), and that the wild regime harbors monodromy (Proposition [6.2]). The primer introduces them from scratch, using the book's own objects as running examples wherever possible.

Throughout this appendix, $X = \mathbb{R}_{\geq 0}^L$ and $M = \mathbb{R}_{\geq 0}$ retain their meanings from Chapter [3]: the commodity space and the force level, respectively. $\mathcal{T}$, $\mathcal{K}$, and $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ are the technology space, cost space, and policy space defined there. When a letter like X appears in a general definition (e.g., "let (X, d) be a metric space"), the generic usage will be clear from context.

A.1   Metric Spaces Beyond $\mathbb{R}^n$

The commodity space $X = \mathbb{R}_{\geq 0}^L$ comes equipped with the Euclidean metric, and much of the reader's analytical toolkit rests on it: convergence is convergence in Euclidean distance, continuity is the ε–δ condition with respect to that distance, and compactness in $\mathbb{R}^n$ is characterized by the Heine–Borel theorem. The spaces

$\mathcal{T}$ and $\mathcal{K}$ that host the book's central arguments are *not* subsets of any $\mathbb{R}^n$: each point in $\mathcal{T}$ is an entire function $\tau : X \to M$, and the space is infinite-dimensional. The good news is that the same conceptual architecture—metric, convergence, continuity, compactness—extends to any set equipped with a distance function satisfying four axioms.

A.1 Definition (Metric space)

A metric space *is a pair* (S, d) *where S is a set and* $d : S \times S \to \mathbb{R}_{\geq 0}$ *is a function satisfying, for all* $x, y, z \in S$:

   (i) Identity of indiscernibles: $d(x, y) = 0$ *if and only if* $x = y$;

   (ii) Symmetry: $d(x, y) = d(y, x)$;

   (iii) Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

The function d is called a metric *or* distance function *on S.*

The Euclidean metric $d(x, y) = \sqrt{\sum_{\ell}(x_\ell - y_\ell)^2}$ on $\mathbb{R}^L$ is the familiar example, but the definition accommodates much more: discrete metrics on finite sets, metrics on spaces of probability measures, and—crucially for this book—metrics on spaces of functions.

Open balls and topology.   The *open ball* of radius $\varepsilon > 0$ around a point $x \in S$ is the set $B_\varepsilon(x) = \{y \in S : d(x, y) < \varepsilon\}$. A set $U \subseteq S$ is *open* if for every $x \in U$ there exists $\varepsilon > 0$ with $B_\varepsilon(x) \subseteq U$. The collection of all open sets is called the *topology* induced by d. This topology determines which subsets are "neighborhoods" of which points, and thereby determines the notions of convergence, continuity, and compactness that follow.

Convergence.   A sequence $\{x_n\}$ in (S, d) *converges* to $x \in S$ if $d(x_n, x) \to 0$ as $n \to \infty$. This is the same definition the reader knows from $\mathbb{R}^n$, with the Euclidean distance replaced by d. In $\mathcal{T}$, convergence means that a sequence of technologies $\{\tau_n\}$ converges to τ when the technologies become uniformly close on every bounded region of the commodity space—a notion that will be made precise in Section .

Continuity. A function $f : (S_1, d_1) \rightarrow (S_2, d_2)$ between metric spaces is *continuous* if it preserves convergence: whenever $x_n \rightarrow x$ in (S_1, d_1), we have $f(x_n) \rightarrow f(x)$ in (S_2, d_2). Equivalently, the preimage $f^{-1}(U)$ of every open set $U \subseteq S_2$ is open in S_1. The ε–δ formulation also generalizes verbatim: f is continuous at x if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_1(x, y) < \delta$ implies $d_2(f(x), f(y)) < \varepsilon$. These three characterizations are equivalent in metric spaces, and we will use whichever is most convenient.

Completeness. A sequence $\{x_n\}$ in (S, d) is *Cauchy* if for every $\varepsilon > 0$ there exists N such that $d(x_m, x_n) < \varepsilon$ whenever $m, n \geq N$. The space (S, d) is *complete* if every Cauchy sequence converges to a limit in S. Euclidean space is complete, and so are both $\mathcal{T}$ and $\mathcal{K}$ under their natural metrics: a Cauchy sequence of technologies converges to a technology, not to some pathological limit outside the space. Completeness is needed for several existence arguments in Chapters 4 and 5, and it holds for the function spaces because the admissibility conditions (Definitions 3.5 and 3.6) are preserved under uniform limits.

Compactness. A subset $A \subseteq S$ is *compact* if every sequence in A has a subsequence converging to a point in A.[1] In $\mathbb{R}^n$, the Heine–Borel theorem says: compact if and only if closed and bounded. In infinite-dimensional spaces like $\mathcal{T}$, this characterization fails spectacularly. A closed, bounded subset of a function space need not be compact: the unit ball in any infinite-dimensional normed space is not compact, because one can construct a sequence of functions with no convergent subsequence (e.g., functions with ever-more-oscillatory behavior). This failure is not a curiosity but a structural feature. It means that certain arguments that work effortlessly in $\mathbb{R}^n$—extracting a convergent subsequence from a bounded sequence—require more care in function spaces. The bounding box construction in Chapter 3 handles this by restricting the production problem to a compact subset of X, restoring the compactness needed for Weierstrass-type existence arguments.

A.2 Remark (Local compactness)

A metric space is locally compact *if every point has a neighborhood whose closure is compact. Euclidean space $\mathbb{R}^n$ is locally compact; $\mathcal{T}$ and $\mathcal{K}$ are not. This is why*

---

[1] This is *sequential* compactness. In metric spaces it coincides with the covering definition (every open cover has a finite subcover), so the distinction is harmless here. See Munkres (2000, Thm 28.2).

the book never invokes the Arzelà–Ascoli theorem in its usual form (which requires local compactness of the domain) and instead works with the Berge Maximum Theorem applied within the compact bounding box.

Product spaces. Given metric spaces (S_1, d_1) and (S_2, d_2), the *product space* $S_1 \times S_2$ carries a natural metric. Several choices are standard (sum metric, max metric, Euclidean combination); all induce the same topology, called the *product topology*. The product topology has the property that convergence in $S_1 \times S_2$ is equivalent to componentwise convergence: $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ in S_1 and $y_n \to y$ in S_2. This is the topology carried by $\mathcal{T} \times \mathcal{K}$ throughout the book: a sequence of institutional configurations (τ_n, κ_n) converges precisely when the technologies converge and the cost functions converge independently. See Munkres (2000, Ch. 2, §20) for the general construction.

A.2 Function Spaces

The single largest conceptual step in this book is the passage from finite-dimensional spaces like $\mathbb{R}^L$ to spaces whose elements are functions. A militarization technology $\tau : X \to M$ is, in the first instance, a function—a rule that assigns a force level to each mobilization plan. But when we write $\mathcal{T}$ for the set of all admissible technologies and ask questions like "is $\mathcal{T}$ contractible?" or "can any technology be continuously deformed into any other?", we are treating τ not as a function but as a *point* in a space. The set of all admissible technologies becomes a space, and the topological questions of Chapters 4–6 are questions about the shape of that space.

The analogy to keep in mind comes from econometrics. A demand curve $q : \mathbb{R}_+ \to \mathbb{R}_+$ is a function; the set of all possible demand curves for a good is a function space. Two demand curves are "close" if they assign similar quantities to similar prices. The metric we place on this set captures what "similar" means, and the resulting topology determines what "continuous transformation of demand" means. The same logic applies to $\mathcal{T}$ and $\mathcal{K}$: we need a metric that captures what it means for two technologies (or two cost functions) to be close to each other.

A.2.1 The compact-open topology

Let S and Y be metric spaces, and let $C(S, Y)$ denote the set of all continuous functions from S to Y. The question is: what is the right topology on $C(S, Y)$? "Right" here means: the topology under which convergence of functions captures what we intuitively mean by "a sequence of functions getting closer and closer to a limit function."

The standard answer is the *compact-open topology*, also called the *topology of compact convergence* or the *topology of uniform convergence on compact sets*.[2]

A.3 Definition (Compact convergence)

A sequence of continuous functions $\{f_n\}$ in $C(S, Y)$ converges compactly to $f \in C(S, Y)$ if, for every compact set $K \subseteq S$, the restrictions $f_n|_K$ converge uniformly to $f|_K$:

$$\sup_{x \in K} d_Y(f_n(x), f(x)) \to 0 \quad \text{as } n \to \infty.$$

The key phrase is *uniformly on compact sets*. On any bounded region of the commodity space, the technologies must converge uniformly: the approximation gets good *everywhere on the region at once*, not just pointwise. But the convergence need not be uniform globally. Two technologies can differ enormously in their behavior at very large mobilization plans while still being "close" in this topology, because what happens at large $\|x\|$ is screened off by the restriction to compact subsets.

This matches the modeling intent. The production problem in Chapter 3 restricts attention to a compact bounding box $\{x \in X : \|x\| \le R\}$; behavior outside the box is irrelevant to the solution. Two technologies that agree closely on every bounded region will generate nearly identical policy functions, regardless of how they differ at infinity. The compact-open topology formalizes precisely this indifference to the tails.

A.2.2 The Fréchet metric

The compact-open topology on $C(S, Y)$ can be realized by a metric when S is σ-*compact*: expressible as a countable union of compact sets. The commodity space

---

[2] These names describe the same topology on $C(S, Y)$ when S is locally compact and Hausdorff. For our function spaces, $S = X = \mathbb{R}_{\ge 0}^L$, which is locally compact, so the identification holds. See Munkres (2000, Thm 46.8).

$X = \mathbb{R}^L_{\geqslant 0}$ is σ-compact—it is the union of the boxes $[0, n]^L$ for $n = 1, 2, 3, \ldots$—so this construction applies directly.

A.4 Definition (Fréchet metric)

Let $\{K_n\}^{\infty}_{n=1}$ be an exhaustion of S by compact sets (i.e., $K_1 \subseteq K_2 \subseteq \cdots$ and $\bigcup_n K_n = S$). The Fréchet metric *on $C(S, Y)$ is*

$$d(f, g) = \sum_{n=1}^{\infty} 2^{-n} \frac{\sup_{x \in K_n} d_Y(f(x), g(x))}{1 + \sup_{x \in K_n} d_Y(f(x), g(x))}. \tag{A.1}$$

Each summand has the structure of a bounded sup-norm. The supremum $\sup_{x \in K_n} d_Y(f(x), g(x))$ measures the worst-case discrepancy between f and g on the nth compact set, and the fraction $r/(1+r)$ squashes this into the interval $[0, 1)$. The weights 2^{-n} ensure convergence of the series and discount the contribution of larger compact sets.

In the book's setting, the exhaustion is $K_n = [0, n]^L \subseteq X$, and the metric on $\mathcal{T}$ becomes

$$d_{\mathcal{T}}(\tau, \tau') = \sum_{n=1}^{\infty} 2^{-n} \frac{\sup_{x \in [0,n]^L} |\tau(x) - \tau'(x)|}{1 + \sup_{x \in [0,n]^L} |\tau(x) - \tau'(x)|}.$$

Two technologies are close in this metric when they assign similar force levels to similar mobilization plans across every bounded region, with discrepancies at large $\|x\|$ contributing exponentially less. The metric on $\mathcal{K}$ is defined identically with κ replacing τ.

A.5 Remark (Why not the sup-norm?)

The reader may wonder why we do not simply use the sup-norm $\|f - g\|_\infty = \sup_{x \in X} |f(x) - g(x)|$. The reason is that technologies and cost functions on the unbounded domain $X = \mathbb{R}^L_{\geqslant 0}$ can differ arbitrarily at large $\|x\|$: if τ is Cobb–Douglas and τ' is CES with elasticity of substitution greater than one, then $|\tau(x) - \tau'(x)| \to \infty$ as $\|x\| \to \infty$, and the sup-norm distance is infinite. The Fréchet metric handles this gracefully: the discrepancy on $[0, n]^L$ is bounded for each n, and the factor 2^{-n} ensures the sum converges. Two technologies are "close" when they agree closely on the regions that matter for bounded production problems, even if they diverge at extreme scales.

Convergence under the Fréchet metric. The following equivalence is the reason the Fréchet metric is the right choice:

$d(f_n, f) \to 0$ *if and only if* $f_n \to f$ *uniformly on every compact subset of S.*

That is, convergence in the Fréchet metric is exactly compact convergence (Definition A.3). The metric metrizes the compact-open topology. See Munkres (2000, Thm 46.8).

Key properties. Under the Fréchet metric, the function spaces $\mathcal{T}$ and $\mathcal{K}$ possess the following properties.

(i) *Completeness.* Both $\mathcal{T}$ and $\mathcal{K}$ are complete metric spaces. The admissibility conditions—log-concavity, exp-convexity, monotonicity—are preserved under compact convergence, so the limit of a Cauchy sequence of technologies is itself an admissible technology.

(ii) *Separability.* Both spaces are separable: each contains a countable dense subset. Separability follows from the separability of $C(K, \mathbb{R})$ under the sup-norm for compact K, which is a consequence of the Stone–Weierstrass theorem. This is a technical condition used in certain measurability arguments; the reader may safely note it and move on.

(iii) *Infinite-dimensionality.* Both $\mathcal{T}$ and $\mathcal{K}$ are infinite-dimensional: no finite set of "basis" technologies spans the space through linear (or any other finite) combination. The closed unit ball is not compact, and Heine–Borel fails.

(iv) *Joint continuity of evaluation.* The evaluation map $(\tau, x) \mapsto \tau(x)$ from $\mathcal{T} \times X$ to M is jointly continuous: if $\tau_n \to \tau$ in $\mathcal{T}$ and $x_n \to x$ in X, then $\tau_n(x_n) \to \tau(x)$. This property, which is automatic in the compact-open topology, is what makes the production problem depend continuously on the technology.

The proofs of (i)–(iv) are collected in Chapter B; the general theory is in Munkres (2000, §46) and Aliprantis and Border (2006, Ch. 3).

A.2.3 The product topology on $\mathcal{T} \times \mathcal{K}$

The space of institutional configurations is the product $\mathcal{T} \times \mathcal{K}$, equipped with the product topology from Section A.1. A natural product metric is

$$d_{\mathcal{T} \times \mathcal{K}}\big((\tau, \kappa), (\tau', \kappa')\big) \;=\; d_{\mathcal{T}}(\tau, \tau') + d_{\mathcal{K}}(\kappa, \kappa'),$$

and convergence $(\tau_n, \kappa_n) \to (\tau, \kappa)$ means precisely that $\tau_n \to \tau$ in $\mathcal{T}$ and $\kappa_n \to \kappa$ in $\mathcal{K}$. The product inherits completeness and separability from its factors.

This is the parameter space on which the solution map $(\tau, \kappa) \mapsto \pi_{\tau,\kappa}$ acts. All of the topological questions in Chapter 4—contractibility, path-connectedness, the existence of continuous deformations—are questions about the topology of $\mathcal{T} \times \mathcal{K}$, or of the policy space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ that it generates.

A.3 Connectedness and Path-Connectedness

The simplest topological property a space can have is *connectedness*: it is "all one piece." A disconnected space splits into components that cannot communicate with each other by any continuous process—no continuous path, no continuous deformation, no homotopy can cross the gap. The contractibility of $\mathcal{T} \times \mathcal{K}$ proved in Chapter 4 implies, as a corollary, that the parameter space is connected; but connectedness is an important stepping stone in its own right, and the stronger notion of *path-connectedness* will recur throughout.

A.3.1 Connected spaces

A.6 Definition (Connected space)

A metric space (S, d) is connected *if it cannot be written as a union $S = U \cup V$ of two nonempty, disjoint open sets.*

Equivalently, S is connected if the only subsets that are simultaneously open and closed are $\varnothing$ and S itself.

A.7 Example

$\mathbb{R}$ is connected. $\mathbb{R} \setminus \{0\}$ is not: the sets $(-\infty, 0)$ and $(0, \infty)$ are both open, nonempty, and disjoint, and their union is the whole space. Every convex subset of $\mathbb{R}^n$ is connected. The complement of the bifurcation set in the wild regime of Chapter 6—the set $\mathcal{U} = \{(a, \epsilon) : a > 1, \epsilon \neq 0\}$ in the cusp parameter plane—is not *connected: it splits into the half-planes $\epsilon > 0$ and $\epsilon < 0$. This disconnection is the topological source of the branch structure in the cusp catastrophe.*

Disconnection matters because it obstructs continuous variation. If the parameter space of a model is disconnected, then there exist pairs of parameter values that cannot be joined by any continuous path: the model has qualitatively distinct "regimes" with no smooth interpolation between them.

A.3.2   Path-connected spaces

A.8 Definition (Path-connected space)

A metric space (S, d) is path-connected *if for every pair of points $x, y \in S$ there exists a continuous map $\gamma : [0, 1] \to S$ with $\gamma(0) = x$ and $\gamma(1) = y$. The map γ is called a* path *from x to y.*

Path-connectedness is strictly stronger than connectedness, *i.e.,* every path-connected space is connected, but the converse fails in general (though counterexamples are exotic and will not arise here).

A.9 Proposition

If (S, d) is path-connected, then it is connected.

Proof. Suppose $S = U \cup V$ with U, V open, nonempty, and disjoint. Pick $x \in U$ and $y \in V$, and let $\gamma : [0, 1] \to S$ be a path from x to y. Then $\gamma^{-1}(U)$ and $\gamma^{-1}(V)$ are open subsets of $[0, 1]$ (by continuity of γ), nonempty (since $0 \in \gamma^{-1}(U)$ and $1 \in \gamma^{-1}(V)$), disjoint, and their union is $[0, 1]$. But $[0, 1]$ is connected—a contradiction. ∎

In the book's setting, the relevant path-connectedness results are:

(i) $\mathcal{T}$ is path-connected: any technology can be continuously deformed into any other through intermediate technologies. The explicit path is the geometric-mean homotopy of Chapter 4: given $\tau_0, \tau_1 \in \mathcal{T}$, the path

$$\gamma(t) \;=\; (1 + \tau_0)^{1-t} \cdot (1 + \tau_1)^t - 1, \qquad t \in [0, 1], \tag{A.2}$$

lies entirely in $\mathcal{T}$. A labor-intensive technology can be smoothly transformed into a capital-intensive one; a Cobb–Douglas technology can be deformed into a CES technology; every such path passes through admissible intermediate technologies.

(ii) $\mathcal{K}$ is path-connected: any cost function can be continuously deformed into any other. The explicit path is the log-sum-exp homotopy:

$$\gamma(t) \;=\; \log\!\big((1 - t)\exp(\kappa_0) + t\exp(\kappa_1)\big), \qquad t \in [0, 1]. \tag{A.3}$$

(iii) $\mathcal{T} \times \mathcal{K}$ is path-connected (apply both homotopies simultaneously). Any state can be continuously deformed into any other state: the space of institutional configurations is "all one piece," and the piece is traversable by smooth institutional reform.

(iv) $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ is path-connected, by composing the parameter-space path with the continuous solution map $(\tau, \kappa) \mapsto \pi_{\tau, \kappa}$.

These are in fact much stronger than path-connectedness—the spaces are contractible, as the next section will explain—but path-connectedness is the first dividend.

A.3.3 Components

A space that is not connected decomposes into maximal connected subsets called *connected components*. Each component is itself connected, the components are pairwise disjoint, and their union is the whole space. Two points lie in the same component if and only if they can be joined by a continuous path (when path-components and components coincide, which they do in the locally path-connected spaces we encounter).

The first corollary of Chapter 4's contractibility result is that $\mathcal{T} \times \mathcal{K}$, and therefore $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, has exactly *one* component: the entire space. Contrast this with the wild regime of Chapter 6, where the complement $\mathcal{U}$ of the bifurcation set has

two components (the half-planes $\epsilon > 0$ and $\epsilon < 0$). The topology changes: what was one piece becomes two, and the pieces carry distinct solution branches. The relationship between connectedness, contractibility, and the fundamental group is the subject of Sections A.4 and A.6.

A.4 Homotopy, Contractibility, and Deformation Retracts

Connectedness asks whether two points can be joined by a path. *Homotopy theory* asks a deeper question: can two *maps* be continuously deformed into each other? The answer organizes the central results of Chapters 4 to 6 into a single framework. When two maps are homotopic, they are "the same" for all topological purposes. When a space is *contractible*—homotopy equivalent to a point—it has the simplest possible topology: no holes, no obstructions, no path-dependence. This is the content of Proposition 4.6. When a space is *not* contractible—when it has a hole that prevents it from shrinking to a point—the obstructions are detected by the fundamental group, which is the subject of Section A.6.

A.4.1 Homotopy between maps

A.10 Definition (Homotopy)

Let X and Y be metric spaces. Two continuous maps $f, g : X \to Y$ are homotopic, *written $f \simeq g$, if there exists a continuous map $H : X \times [0,1] \to Y$ such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$. The map H is called a* homotopy *from f to g; the parameter $t \in [0,1]$ is called the* homotopy parameter.

Think of H as a movie: at time $t = 0$ we see the map f; at time $t = 1$ we see the map g; at each intermediate time t we see a map $H(\cdot, t)$ that is itself continuous, and the transition from one frame to the next is smooth. The homotopy is a continuous interpolation between f and g through intermediate maps.

A.11 Example (Straight-line homotopy)

If Y is a convex subset of $\mathbb{R}^n$ and $f, g : X \to Y$ are continuous, then

$$H(x,t) = (1-t) f(x) + t\, g(x)$$

is a homotopy from f to g: the convex combination of two points in a convex set is in the convex set, so $H(x,t) \in Y$ for all t, and H is continuous because addition and scalar multiplication are continuous. This is the simplest possible homotopy. It works whenever the codomain has no "holes" that the straight line might cross.

A.12 Example (The book's homotopies are NOT straight-line)

The technology space $\mathcal{T}$ is not a convex subset of any vector space—the pointwise convex combination $(1-t)\tau_0 + t\tau_1$ of two log-concave functions is not in general log-concave (Section A.7 gives a counterexample)—so the straight-line homotopy does not preserve admissibility. The geometric-mean homotopy (A.2) takes a nonlinear path that stays within $\mathcal{T}$. Section A.9 explains why this particular nonlinear path is, in a precise sense, the simplest one available: it is the straight-line homotopy, but in a different coordinate system.

Homotopy is an equivalence relation: $f \simeq f$ (take $H(x,t) = f(x)$); $f \simeq g$ implies $g \simeq f$ (reverse the parameter: $H'(x,t) = H(x,1-t)$); and $f \simeq g$ and $g \simeq h$ imply $f \simeq h$ (concatenate the two homotopies, reparametrizing so that the first runs on $[0, 1/2]$ and the second on $[1/2, 1]$).

Homotopy equivalence. Two spaces X and Y are *homotopy equivalent* if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$. This is a coarser relation than homeomorphism (topological equivalence): homotopy-equivalent spaces have the same "shape" for topological purposes—the same fundamental group, the same homology, the same obstructions to extending maps—even if they are not geometrically identical. A filled disk and a single point are homotopy equivalent (shrink the disk); a circle and a single point are not (the hole gets in the way).

A.4.2   Contractibility

A.13 Definition (Contractible space)

A metric space X is contractible *if it is homotopy equivalent to a point. Equivalently, X is contractible if there exists a point $x_0 \in X$ and a continuous map $H : X \times [0,1] \to X$ such that*

$$H(x,0) = x \quad \text{and} \quad H(x,1) = x_0 \quad \text{for all } x \in X.$$

The map H is called a contraction *of X to x_0.*

A contraction is a homotopy from the identity map id_X to the constant map $x \mapsto x_0$. It continuously deforms every point in the space to a single target, like deflating a balloon to a point. At each instant $t \in [0,1]$, the map $x \mapsto H(x,t)$ is a continuous "snapshot" of the deflation in progress.

Contractibility is the strongest "sameness" property a space can have. It implies:

- *Path-connectedness:* any two points can be joined by a path (connect each to the basepoint via the contraction, then compose).

- *Simply connected:* every loop can be shrunk to a point (Section A.6).

- *Trivial fundamental group:* $\pi_1(X) = \{e\}$—there is no topological path dependence (Section A.6).

- *All maps are homotopic:* any two continuous maps $f, g : Y \to X$ are homotopic, because both are homotopic to the constant map.

A.14 Example (Convexity implies contractibility)

Let $C \subseteq \mathbb{R}^n$ be a convex set. For any $x_0 \in C$, the straight-line homotopy $H(x,t) = (1-t)x + tx_0$ is a contraction of C to x_0: convex combinations of elements of C lie in C, so H maps $C \times [0,1]$ into C. The proof is three lines.

This is the familiar case: the reader's experience with convex feasible sets, convex constraint sets, and convex combinations of demand curves all involve contractible spaces. The novelty in Chapter 4 is that $\mathcal{T}$ is contractible without being convex—it requires the nonlinear homotopies of (A.2) and (A.3) to establish the contraction.

A.15 Example (The circle is not contractible)

The circle $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ is not *contractible: there is no continuous way to shrink the circle to a point while staying on the circle. The "hole" in the middle obstructs the contraction. This is the prototypical obstruction, and it is detected by the fundamental group: $\pi_1(S^1) = \mathbb{Z} \neq \{e\}$ (Section A.6). The monodromy phenomenon of Chapter 6 arises from precisely the same mechanism: the complement of the bifurcation set has a "hole" (the bifurcation locus itself), and loops encircling it cannot be contracted.*

A.16 Remark (Contractibility vs. convexity)

Convexity implies contractibility, but not vice versa. The technology space $\mathcal{T}$ is contractible but not convex; the set $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\} \setminus \{(0,0)\}$ (a disk with the origin removed) is contractible but not convex. The tame subspace $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, by contrast, is convex in a natural sense (Section A.7), which gives it extra geometric structure beyond what contractibility alone provides.

A.4.3 Deformation retracts

Contractibility says a space can be shrunk to a point. A *deformation retract* generalizes this: the space can be shrunk onto a subspace, not necessarily a point.

A.17 Definition (Retraction and deformation retract)

Let $A \subseteq X$. A retraction *of X onto A is a continuous map $r : X \to A$ such that $r(a) = a$ for all $a \in A$. The subspace A is a* deformation retract *of X if there exists a continuous map $H : X \times [0,1] \to X$ such that*

$$H(x,0) = x, \qquad H(x,1) \in A \quad \text{for all } x \in X.$$

The subspace A is a strong deformation retract *if, additionally,*

$$H(a,t) = a \quad \text{for all } a \in A \text{ and all } t \in [0,1].$$

The difference between a deformation retract and a *strong* deformation retract is the fixity condition: in a strong deformation retract, points already in A do not move at any time during the deformation. They are fixed from the start, not merely returned to their positions at the end.

This distinction matters for the book's modeling criterion. Chapter 5's structural adequacy (Definition 5.8) requires a *strong* deformation retraction of $\mathcal{T} \times \mathcal{K}$ onto the tame subspace, with the additional constraint that admissibility is preserved at every intermediate stage. If the retraction were merely a deformation retract, a tame technology could temporarily become non-tame during

the deformation before being returned to its tame form at $t = 1$—physically, this would mean that a first-order approximation is temporarily promoted to some wild intermediate form and then snapped back, which would undermine the claim that the simplification is canonical. Strong deformation retraction prevents this: tame functions never move.

A.18 Example (The main theorem as a strong deformation retract)

Proposition 4.6 states that $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ strongly deformation retracts onto a single point π_0. The deformation is the composite homotopy: contract the technologies via the geometric mean, simultaneously contract the costs via log-sum-exp, and compose with the solution map. At $t = 0$, each policy function is itself. At $t = 1$, every policy function has been deformed to π_0. At every intermediate t, the policy function at that stage is itself the solution to a legitimate production problem (with admissible intermediate parameters). A contractible space is one that strongly deformation retracts onto a point.

A.5 Structural Adequacy

The language of deformation retracts gives Chapter 5's central notion a precise home. Structural adequacy asks: is the tame subspace a strong deformation retract of the full parameter space, with the deformation preserving admissibility at every stage? This section unpacks the definition and explains why each of its components is load-bearing.

A.5.1 The definition

Recall from Definition 5.8 that a subspace $A \subset X$ is *structurally adequate* for X if there exists a strong deformation retraction $H : X \times [0, 1] \to X$ such that $H(x, 0) = x$, $H(x, 1) \in A$, $H(a, t) = a$ for all $a \in A$ and $t \in [0, 1]$, and $H(x, t)$ satisfies the admissibility constraints of X for every t.

The definition combines three requirements:

(i) **Continuity.** Nearby general functions must map to nearby tame representatives. The retraction is stable under perturbation: small changes in the input produce small changes in the simplified output.

(ii) **Path-fidelity.** The deformation path $t \mapsto H(x, t)$ must stay inside the admissible space at every intermediate step. Every function encountered along the retraction is itself a valid technology or cost function. The simplification never passes through a pathological intermediate.

(iii) **Fixity.** Functions already in the subspace are fixed throughout: $H(a, t) = a$ for all $a \in A$ and t. The retraction does not disturb what is already in canonical form.

A.5.2 Why "strong" matters for modeling

The force of the "strong" modifier is clearest by contrast. A bare deformation retract would allow the following scenario: a tame technology τ (a first-order approximation) is temporarily deformed into a complicated intermediate $H(\tau, t)$ for $0 < t < 1$, then returned to τ at $t = 1$. The deformation would achieve the topological goal—A is a retract—but the "simplification" would involve making simple things temporarily complex before returning them. This undermines the interpretation of the retraction as a canonical passage from general to simple.

Strong deformation retraction prevents this. If τ is already tame, it stays tame at every stage: H acts as the identity on A. The deformation is a one-way simplification: general functions move toward the tame core; tame functions stay put. This asymmetry is what justifies interpreting the retraction as a genuine simplification process rather than a topological trick.

A.5.3 The retraction of Chapter 5

The structural adequacy of the tame subspace (Proposition 5.9) is established by a two-stage retraction:

1. **Regularize** ($t \in [0, 1/2]$). Smooth the function via a causal convolution operator $\mathfrak{D}$, using the same chart interpolation as the contractibility homotopy: the geometric mean for technologies, log-sum-exp for costs.

2. **Tamify** ($t \in [1/2, 1]$). Project the regularized function to its first-order (affine or log-affine) approximation via $\mathfrak{T}$, again using chart interpolation to stay within the admissible space.

The two stages meet continuously at $t = 1/2$, and the composite $\mathfrak{T} \circ \mathfrak{D}$ acts as the identity on tame functions (they are already smooth and first-order), securing

the fixity condition. The full proof is in Chapter B; what matters here is the mechanism: the same chart-based homotopies that establish contractibility in Chapter 4 are reused to establish structural adequacy in Chapter 5.

A.6 The Fundamental Group π_1

Path-connectedness asks whether two points can be joined by a path. The fundamental group asks a subtler question: if a path starts and ends at the same point—forming a *loop*—can it be continuously shrunk to a point? The set of all "essentially different" loops at a basepoint forms a group, the *fundamental group π_1*, whose structure encodes the topological obstructions in a space. A trivial fundamental group ($\pi_1 = \{e\}$) means every loop unwinds—there are no holes, no irreducible circuits, no monodromy. A nontrivial fundamental group ($\pi_1 \neq \{e\}$) means the space has holes that certain loops must encircle, and the consequences of this for optimization are the subject of Chapter 6.

A.6.1 Loops and homotopy classes

A.19 Definition (Loop)

Let (X, d) be a metric space and $x_0 \in X$. A loop based at x_0 *is a continuous map $\gamma : [0, 1] \to X$ with $\gamma(0) = \gamma(1) = x_0$.*

A loop is a closed path: it starts and ends at the same point. In the book's setting, a loop in the parameter space $\mathcal{T} \times \mathcal{K}$ is a continuous sequence of institutional reforms—technology changes, cost structure adjustments—that returns the state to its original configuration.

A.20 Definition (Homotopy of loops)

Two loops γ_0, γ_1 based at x_0 are homotopic *(relative to the basepoint) if there exists a continuous map $H : [0, 1] \times [0, 1] \to X$ such that*

$$H(s, 0) = \gamma_0(s), \qquad H(s, 1) = \gamma_1(s),$$
$$H(0, t) = x_0, \qquad H(1, t) = x_0$$

for all $s, t \in [0, 1]$. The homotopy deforms one loop into the other while keeping the basepoint fixed.

If two loops are homotopic, they trace "essentially the same" circuit: one can be continuously deformed into the other without tearing, cutting, or moving the basepoint. The homotopy class of a loop is its equivalence class under this relation.

A.6.2 The group structure

A.21 Definition (Fundamental group)

The fundamental group $\pi_1(X, x_0)$ *is the set of homotopy classes of loops based at* x_0, *equipped with the following operations:*

- Product: $[\gamma_0] \cdot [\gamma_1]$ *is the class of the loop that first traverses* γ_0 *and then* γ_1 *(concatenation, reparametrized to* $[0, 1]$*).*

- Identity: *the class* $[e]$ *of the constant loop* $e(s) = x_0$ *for all* s.

- Inverse: $[\gamma]^{-1} = [\widehat{\gamma}]$, *where* $\widehat{\gamma}(s) = \gamma(1 - s)$ *is the reversed loop.*

These operations are well-defined on homotopy classes and satisfy the group axioms.

When X is path-connected, $\pi_1(X, x_0)$ is independent of the choice of basepoint up to isomorphism, so one writes simply $\pi_1(X)$.

A.6.3 Key computations

The reader need not compute fundamental groups from scratch; the following classical results suffice for all of the book's arguments.

(i) $\pi_1(\mathbb{R}^n) = \{e\}$. Euclidean space is *simply connected*: every loop can be shrunk to a point via the straight-line homotopy $H(s, t) = (1 - t)\gamma(s) + t x_0$.

(ii) $\pi_1(S^1) = \mathbb{Z}$. A loop on the circle is classified by its *winding number*: the integer k counting how many times (and in which direction) the loop winds around the circle. The loop with winding number k cannot be deformed into a loop with winding number $k' \neq k$. This is *the* fundamental example; see Hatcher (2002, Thm 1.7).

(iii) $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$. The punctured plane is homotopy equivalent to S^1 (retract radially onto the unit circle), so it has the same fundamental group. A loop encircling the origin cannot be shrunk to a point because it must cross the missing origin to do so.

(iv) If X is contractible, then $\pi_1(X) = \{e\}$. Every loop in a contractible space can be shrunk: compose the loop with the contraction to pull it to the basepoint.

These four facts, together with the computation for the cusp complement, are all the fundamental group theory the book requires.

A.6.4 Application to the book

The fundamental group is the device that makes precise the book's central contrast between Chapter 4 and Chapter 6.

Chapter 4: trivial π_1. $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is contractible, so $\pi_1(\mathcal{P}_{\mathcal{T}\times\mathcal{K}}) = \{e\}$. Every loop of institutional reform unwinds: any sequence of changes to a state's technology and cost structure that returns to the original parameters also returns to the original policy function. There is no topological path dependence. History does not matter—not because history is trivial, but because the space of possible states has no topological obstructions that could make it matter.

Chapter 6: nontrivial π_1. Consider the complement of the bifurcation set in the cusp parameter plane: $\mathcal{U} = \{(a, \epsilon) : a > 1, \epsilon \neq 0\}$. This is (homotopy equivalent to) a circle: the bifurcation half-line $\{a > 1, \epsilon = 0\}$ acts as a "barrier" that a loop must go around, and the cusp point $(1, 0)$ plays the role of the missing origin in the punctured plane. Therefore $\pi_1(\mathcal{U}) \cong \mathbb{Z}$.

A loop γ in $\mathcal{U}$ that encircles the cusp point cannot be shrunk to a point without crossing the bifurcation locus. Crucially, the global minimum of the production problem—the optimal solution branch—varies continuously along γ, but when the loop completes, the minimum has switched branches: $x^A \leftrightarrow x^B$. This is *monodromy*, the topological mechanism that converts a nontrivial π_1 into genuine path dependence. A state that traverses such a loop—adjusting its technology parameter a and its cost asymmetry ϵ in a circuit around the cusp— finds itself, upon return, on the opposite branch of specialization. Its institutions have undergone a topological phase transition with no local discontinuity.

The contrast is stark:

	Chapter 4	Chapter 6
Space	$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$	$\mathcal{U} \subseteq \{a > 1\} \times \mathbb{R}$
π_1	$\{e\}$	$\mathbb{Z}$
Loops	all unwind	some encircle cusp
Path dependence	none	monodromy

The reversal from trivial π_1 to nontrivial π_1 is the formal content of the book's shift from "states are alike" to "states can become trapped."

A.7 Convexity in Function Spaces

Convexity is the reader's most familiar structural property: a set is convex if the line segment between any two of its points lies entirely within the set, and convex sets are ubiquitous in optimization because they ensure that local optima are global optima. This section extends the concept to function spaces and explains a key asymmetry: the technology space $\mathcal{T}$ is *not* convex, but the tame subspace *is*.

A.7.1 Convex combinations of functions

In a vector space, the convex combination of elements f and g is $\lambda f + (1 - \lambda)g$ for $\lambda \in [0, 1]$. For function spaces embedded in a vector space, this is a *pointwise* operation:

$$(\lambda f + (1 - \lambda)g)(x) = \lambda f(x) + (1 - \lambda) g(x).$$

A subset $\mathcal{F}$ of a vector space of functions is *convex* if $\lambda f + (1 - \lambda)g \in \mathcal{F}$ whenever $f, g \in \mathcal{F}$ and $\lambda \in [0, 1]$.

A.22 Example (Concave functions form a convex set)

The set of all concave functions $f : \mathbb{R}^L \to \mathbb{R}$ is convex. If f and g are concave and $\lambda \in [0, 1]$, write $h = \lambda f + (1 - \lambda)g$. Since concavity is preserved under addition and nonnegative scaling, h is concave—that is, $h(\alpha x + (1 - \alpha)y) \geq \alpha h(x) + (1 - \alpha) h(y)$ for all x, y and $\alpha \in [0, 1]$. The same argument shows that convex functions form a convex set.

A.7.2 $\mathcal{T}$ and $\mathcal{K}$ are not convex

Despite the example above, the technology and cost spaces are *not* convex subsets of any ambient vector space. The reason is that log-concavity and exp-convexity are not preserved under pointwise addition.

A.23 Example (Failure of convexity in $\mathcal{T}$)

Consider L = 1 and two technologies in $\mathcal{T}$:

$$\tau_0(x) = x^2 \qquad and \qquad \tau_1(x) = e^x - 1.$$

Both satisfy log-concavity: $\log(1+x^2)$ is concave on $\mathbb{R}_{\geq 0}$ and $\log(e^x) = x$ is (trivially) concave. But their pointwise average $1/2(x^2 + e^x - 1)$ need not satisfy log-concavity: $\log(1 + 1/2(x^2 + e^x - 1))$ is not in general concave. The midpoint of the segment is outside $\mathcal{T}$.

This is why the contractibility of $\mathcal{T}$ is a *theorem* rather than a triviality. If $\mathcal{T}$ were convex, contractibility would follow from three lines (Example A.14). Because it is not convex, the contraction must take a nonlinear path—the geometric-mean homotopy—that avoids the "log-concavity barrier" separating admissible technologies from inadmissible ones.

A.7.3 The tame subspace is convex

While the full spaces $\mathcal{T}$ and $\mathcal{K}$ are not convex, the *tame* subspaces—the finite-dimensional families of first-order (log-affine and exp-affine) technologies and costs—are. At the level of policy functions, if π_0 and π_1 are tame policy functions, then $(1 - t)\pi_0 + t\pi_1$ is tame for every $t \in [0, 1]$.

The convexity of the tame subspace is a discovered structural property (Chapter 5), not assumed. It provides extra geometric structure beyond what contractibility alone delivers:

- Every local optimum over the tame subspace is a global optimum (convexity of the feasible set).

- The straight-line homotopy works within the tame subspace: one does not need the nonlinear chart-based interpolation.

- Statistical estimation of tame parameters inherits the nice properties of convex parameter spaces (unique MLEs, etc.).

A.7.4 Convexity implies contractibility, but not vice versa

The logical relationships are:

$$\text{convex} \implies \text{contractible} \implies \text{simply connected} \implies \pi_1 = \{e\}.$$

None of the converses holds. The full technology space $\mathcal{T}$ is contractible but not convex. The complement of the bifurcation set in the wild regime is not simply connected (and therefore not contractible).

The hierarchy matters because it locates the book's results with precision. Chapter 4 establishes contractibility of $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$, which gives trivial π_1 and therefore no path dependence. Chapter 5 establishes convexity of the tame subspace, which gives contractibility *and* the additional optimization guarantees that convexity provides. Chapter 6 shows that the wild regime fails even simply-connectedness, which means the failures are as bad as they can be: nontrivial π_1, monodromy, and genuine path dependence.

A.8 The Optimization Toolkit

The topological results of Chapters 4–6 rest on an optimization problem: the state's production problem ($\mathsf{SPP}\,(m, \tau, \kappa)$), which minimizes cost subject to a force-level constraint. The topological structure of the policy space $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ depends on how solutions to this problem vary as the technology and cost function change. This section restates the key optimization tools in the book's notation, emphasizing what is new relative to the finite-dimensional treatment in Sundaram (1996).

A.8.1 The production problem

The state's production problem, as defined in Chapter 3, is

$$\min_{x \in X} \ \kappa(x) \qquad \text{subject to} \qquad \tau(x) \geq m, \tag{SPP}$$

where $m \in M$ is the force target, $\tau \in \mathcal{T}$ is the technology, and $\kappa \in \mathcal{K}$ is the cost function. The solution is a mobilization plan $x^*(m, \tau, \kappa) \in X$ that achieves the required force level at minimum cost.

The problem is finite-dimensional in the decision variable x: x lives in $\mathbb{R}^L_{\geq 0}$, not in a function space. What is infinite-dimensional is the *parameter*: the pair (τ, κ) lives in $\mathcal{T} \times \mathcal{K}$, and the question is how the solution depends on this infinite-dimensional parameter.

A.8.2   Existence and uniqueness

Existence of a solution to (SPP) follows from the Weierstrass extreme value theorem, provided the feasible set is compact. The bounding box construction in Chapter 3 ensures this: the coerciveness condition ($\mathfrak{O}_\kappa$) and the boundedness of the force target guarantee that any minimizer lies within a compact box $\{x \in X : \|x\| \leq R(m)\}$, so the infimum is attained.

   Uniqueness follows from the strict convexity of the cost function on the level sets of the technology. The exp-convexity condition ($\mathfrak{L}_\kappa$) implies that κ is strictly quasiconvex, and a strictly quasiconvex function attains its minimum at most once on any sublevel set of a quasiconcave constraint. See Sundaram (1996, Ch. 6) for the finite-dimensional theory; the proof in Chapter B adapts the argument to the book's setting.

A.8.3   Berge's Maximum Theorem

The Maximum Theorem, due to Berge (1963), is the fundamental tool for establishing continuity of the solution map. It connects the topology of the parameter space to the continuity of the value function and the optimizer.

A.24 Theorem (Berge's Maximum Theorem)

Let Θ be a metric space, Y a metric space, $f : \Theta \times Y \to \mathbb{R}$ a continuous function, and $\Gamma : \Theta \rightrightarrows Y$ a continuous correspondence (both upper and lower hemicontinuous) with nonempty compact values. Define the value function $V(\theta) = \max_{y \in \Gamma(\theta)} f(\theta, y)$ *and the* argmax correspondence $Y^*(\theta) = \{y \in \Gamma(\theta) : f(\theta, y) = V(\theta)\}$. *Then:*

   (i) V is continuous;

   (ii) Y^* is upper hemicontinuous with nonempty compact values.

The reader has likely seen this theorem in the context of dynamic programming or consumer theory, where the parameter space Θ is finite-dimensional. In the book's application, $\Theta = M \times \mathcal{T} \times \mathcal{K}$, which is infinite-dimensional. The theorem still applies because its hypotheses are stated in terms of metric-space continuity, not dimensionality: what matters is that the objective is jointly continuous and the constraint correspondence is continuous.

   The joint continuity of the objective $-\kappa(x)$ (viewed as a function of (κ, x)) follows from the joint continuity of evaluation in the compact-open topology (Section A.2). The continuity of the constraint correspondence $\Gamma(m, \tau) = \{x \in$

$X : \tau(x) \geqslant m, \; \|x\| \leqslant R\}$ requires checking upper and lower hemicontinuity as τ varies in the Fréchet metric—this is the technical core of the argument, and the proof is in Chapter B.

See Sundaram (1996, Thm 9.14) for the standard statement; Aliprantis and Border (2006, Thm 17.31) for the general (not necessarily finite-dimensional) version.

A.8.4 Hemicontinuity

The Maximum Theorem requires the constraint correspondence to be *continuous*—both upper and lower hemicontinuous. These are the same definitions the reader knows from Sundaram (1996, Ch. 9); the only new ingredient is that "neighborhood" is now determined by the Fréchet metric on the parameter space.

A.25 Definition (Hemicontinuity)

Let $\Gamma : \Theta \rightrightarrows Y$ be a correspondence.

  (i) Γ *is* upper hemicontinuous *at θ_0 if, for every open set $V \supseteq \Gamma(\theta_0)$, there exists a neighborhood U of θ_0 such that $\Gamma(\theta) \subseteq V$ for all $\theta \in U$.*

  (ii) Γ *is* lower hemicontinuous *at θ_0 if, for every open set V with $V \cap \Gamma(\theta_0) \neq \varnothing$, there exists a neighborhood U of θ_0 such that $V \cap \Gamma(\theta) \neq \varnothing$ for all $\theta \in U$.*

Upper hemicontinuity says the correspondence does not "suddenly expand": the image cannot jump to include distant points. Lower hemicontinuity says it does not "suddenly shrink": every point currently in the image has nearby points in nearby images. Together, they ensure that the feasible set varies smoothly as the parameters change—which is exactly what is needed for the Maximum Theorem to deliver continuous dependence of the solution on the parameters.

A.9 Chart Homotopies: Why the Admissibility Conditions Have Their Form

This final section ties the primer together. The reader has seen three kinds of homotopy in the book: the geometric-mean homotopy on $\mathcal{T}$, the log-sum-exp homotopy on $\mathcal{K}$, and the straight-line homotopy on convex sets. The first two

look more complicated than the third. This section explains why they are, in fact, the *same thing*: all three are linear interpolation, just in different coordinate systems. The admissibility conditions turn out to be precisely the conditions that make such a coordinate change possible.

A.9.1   The problem with linear interpolation

As noted in Section A.7.2, the technology space $\mathcal{T}$ is not a convex subset of any vector space. The pointwise convex combination $(1-t)\tau_0 + t\tau_1$ of two admissible technologies is not in general admissible: log-concavity—the condition that $\log(1 + \tau(x))$ is concave—is not preserved under pointwise addition.

Dually, the cost space $\mathcal{K}$ is not convex either. The pointwise convex combination $(1 - t)\kappa_0 + t\kappa_1$ need not satisfy exp-convexity: $\exp(\kappa(x))$ being convex is not preserved under pointwise addition of κ.

If we cannot interpolate linearly in the "native" coordinates (τ, κ), we need a different coordinate system—one in which the admissibility conditions *become* convexity, and linear interpolation is therefore safe.

A.9.2   The chart idea

The technology chart.   Define the change of coordinates

$$\varphi \;=\; \log(1 + \tau). \tag{A.4}$$

Under this map, the log-concavity condition $\tilde{\mathfrak{L}}_\tau$—that $\log(1 + \tau(x))$ is concave—becomes simply: φ *is concave*. The image of $\mathcal{T}$ under this coordinate change is (a subset of) the cone of concave functions, which *is* a convex set (Example A.22).

Now perform linear interpolation in the φ-chart:

$$\varphi_t \;=\; (1 - t)\,\varphi_0 + t\,\varphi_1 \;=\; (1 - t)\,\log(1 + \tau_0) + t\,\log(1 + \tau_1).$$

This is a convex combination of concave functions, which is concave. So φ_t satisfies the admissibility condition for every t—the interpolation stays inside the admissible region.

Transform back to native coordinates:

$$\tau_t \;=\; \exp(\varphi_t) - 1 \;=\; (1 + \tau_0)^{1-t} \cdot (1 + \tau_1)^t - 1.$$

This is the geometric-mean homotopy of (A.2)—the formula from Chapter 4. *It was linear interpolation all along, in the φ-chart.*

The cost chart.   Define the change of coordinates

$$\psi \;=\; \exp(\kappa). \tag{A.5}$$

Under this map, the exp-convexity condition $\mathfrak{L}_\kappa$—that $\exp(\kappa(x))$ is convex—becomes: ψ *is convex*. Convex functions form a convex set, so linear interpolation in the ψ-chart preserves admissibility:

$$\psi_t \;=\; (1-t)\,\psi_0 + t\,\psi_1 \;=\; (1-t)\,\exp(\kappa_0) + t\,\exp(\kappa_1).$$

Transform back:

$$\kappa_t \;=\; \log(\psi_t) \;=\; \log\big((1-t)\exp(\kappa_0) + t\exp(\kappa_1)\big).$$

This is the log-sum-exp homotopy of (A.3)—linear interpolation in the ψ-chart.

A.9.3   What this means

The chart construction reveals a deep relationship between the admissibility conditions and the existence of homotopies.

The conditions are not arbitrary.   Log-concavity of τ and exp-convexity of κ are precisely the conditions under which a coordinate change linearizes the constraint. They are the conditions that make continuous deformation through admissible intermediates *possible*: without them, no chart exists, and the homotopy construction collapses.

Why these homotopies.   The geometric mean and log-sum-exp formulas look nonlinear, but they are the images of linear interpolation under the coordinate changes defined by the admissibility conditions. Other homotopies could establish contractibility—for instance, a retraction to a fixed basepoint need not interpolate between arbitrary pairs at all—but the chart-based construction has a specific virtue: every intermediate point is an explicit convex combination of the chart images of the endpoints, so the entire path is controlled by its boundary values and the deformation is transparent and verifiable.

When admissibility fails. When the admissibility conditions fail—when the technology is not log-concave, or the cost function is not exp-convex—the chart construction breaks down. The image of the inadmissible space under (A.4) is no longer contained in the cone of concave functions, so linear interpolation in the chart need not stay in the admissible region.

The failure of this particular construction does not, by itself, prove that the space is no longer contractible—some other homotopy might work. What establishes the topological change is a separate argument: Chapter 6 exhibits specific violations of admissibility where the topology is provably nontrivial. The CES technology $\tau_a(x) = (x_1^a + x_2^a)^{1/a}$ is log-concave for $a \leq 1$ and *not* log-concave for $a > 1$. When $a > 1$, the chart-based homotopy fails, but more importantly, the resulting space has $\pi_1 \neq \{e\}$: monodromy appears, and path dependence becomes a topological fact rather than a modeling assumption. No homotopy—chart-based or otherwise—can contract a space with nontrivial fundamental group.

The admissibility conditions are therefore *sufficient* for contractibility, and specific violations are shown to be *sufficient for non-contractibility*. Whether they are necessary—whether milder conditions might support contractibility via a different construction—remains open.

The phase boundary. The admissibility conditions define a *phase boundary* in the space of all possible technologies and cost functions. On one side (log-concave technologies, exp-convex costs), the chart homotopies work, the spaces are contractible, and Waltz's "like units" claim is valid. On the other side, specific violations produce nontrivial topology—monodromy, disconnected components, catastrophic divergence—and the structural similarity of states breaks down. This boundary is the formal content of the book's reversal: the very conditions that guarantee "states are alike" are the conditions whose violation produces "states are stuck."

Appendix B

Proofs

This appendix collects the proofs of all results stated in the main text. Each proof is preceded by a restatement of the result, and a back-link is provided for the reader's convenience. We freely cite the definitions and constructions developed in Chapter A; the notation and terminology established there are used throughout.

B.1 Properties of the Function Spaces

We begin with two foundational results that are promised in the primer and invoked in later proofs.

B.1.1 Metric properties

The Fréchet metric

$$d\left(f_0, f_1\right) = \sum_{n \in \mathbb{N}} \frac{1}{2^n} \cdot \frac{\max_{x \in [0,n]^L} |f_0(x) - f_1(x)|}{1 + \max_{x \in [0,n]^L} |f_0(x) - f_1(x)|}$$

is used to metrize both $\mathcal{T}$ and $\mathcal{K}$. We verify the four properties promised in Section A.2.2.

1. *d is a metric.* Non-negativity and symmetry are immediate. Identity of indiscernibles holds because $X = \bigcup_n [0, n]^L$. For the triangle inequality, define $\psi_n(f_0, f_1) = \max_{x \in [0,n]^L} |f_0(x) - f_1(x)|$ and $\xi(\psi) = \psi/(1 + \psi)$. The ordinary triangle inequality for the absolute value gives $\psi_n(f_0, f_2) \leq$

$\psi_n(f_0, f_1) + \psi_n(f_1, f_2)$; since ξ is increasing and subadditive on $\mathbb{R}_{\geqslant 0}$, the nth summand of d is subadditive, and summing over n yields $d(f_0, f_2) \leqslant d(f_0, f_1) + d(f_1, f_2)$.

2. *Completeness.* Convergence in d is equivalent to uniform convergence on every compact $[0, n]^L$. If (f_k) is Cauchy in d, then for each n, the restrictions $f_k|_{[0,n]^L}$ form a Cauchy sequence in the supremum norm and converge uniformly to some continuous limit $f^{(n)}$. Consistency gives a global limit $f \in C(X)$, and the admissibility axioms (log-concavity for $\mathcal{T}$, exp-convexity for $\mathcal{K}$) pass to pointwise limits.

3. *Separability.* Polynomial functions with rational coefficients are dense in $C([0, n]^L)$ under the sup-norm (Stone–Weierstrass), and a diagonal argument over n provides a countable dense subset of $(C(X), d)$.

4. *Joint continuity of evaluation.* The map $(f, x) \mapsto f(x)$ is continuous from $(\mathcal{T}, d) \times X \to \mathbb{R}$: $|f(x) - g(y)| \leqslant |f(x) - g(x)| + |g(x) - g(y)|$, where the first term is controlled by $d(f, g)$ on any compact set containing x, and the second by continuity of g.

B.1.2 Hemicontinuity of the constraint correspondence

The constraint correspondence for the state production problem is

$$\mathcal{X} : M \times \mathcal{T} \times \mathcal{K} \rightrightarrows X, \qquad (m, \tau, \kappa) \mapsto X^{\max}(m, \tau, \kappa) \cap \tau^{-1}(\{m\}),$$

where $X^{\max}$ is the bounding box defined below. We prove that $\mathcal{X}$ is nonempty, compact-valued, and continuous.

Step 1: Bounding box. By the Michael selection theorem (Aliprantis and Border 2006, Thm 17.66), the upper contour correspondence $\psi(m, \tau) = \{x \in X : \tau(x) \geqslant m\}$ admits a continuous selection $\tilde{x} : M \times \mathcal{T} \to X$ satisfying $\tau(\tilde{x}(m, \tau)) \geqslant m$. A scaling argument via the intermediate value theorem produces a modified selection with $\tau(\tilde{x}(m, \tau)) = m$ exactly. For each coordinate ℓ, the function $x_\ell^{\max}(m, \tau, \kappa)$ defined as the unique scalar satisfying $\kappa(x_\ell^{\max} e_\ell) = \kappa(\tilde{x}(m, \tau))$ is continuous and provides a compact box $X^{\max}(m, \tau, \kappa) = \prod_\ell [0, x_\ell^{\max}(m, \tau, \kappa)]$ containing all solutions to the SPP.

Step 2: Nonemptiness and compactness. The selection $\tilde{x}(m, \tau)$ lies in both $X^{\max}$ and $\tau^{-1}(\{m\})$, so $\mathcal{X}$ is nonempty. Since $X^{\max}$ is compact and $\tau^{-1}(\{m\})$ is closed (by continuity of τ), the intersection is compact.

Step 3: Upper hemicontinuity. By [Aliprantis and Border (2006, Thm 17.25)](#), the intersection of an upper hemicontinuous compact-valued correspondence with a closed-valued correspondence is upper hemicontinuous. The box $X^{\max}$ is uhc and compact-valued (its sides vary continuously); the level set $\tau^{-1}(\{m\})$ is closed-valued.

Step 4: Lower hemicontinuity. Suppose $x \in \mathcal{X}(m, \tau, \kappa) \cap V$ for some open $V \subseteq X$. Choose $\varepsilon > 0$ with $B_\varepsilon(x) \subseteq V$. By continuity of $X^{\max}$, there is a neighborhood U_1 of (m, τ, κ) such that $x \in X^{\max}(m', \tau', \kappa')$ for all $(m', \tau', \kappa') \in U_1$. Joint continuity of the evaluation map gives a neighborhood $U_2 \subseteq U_1$ with $|\tau'(x) - m| < \delta$ for all $(m', \tau') \in U_2$, where $\delta > 0$ is chosen below. Since τ is weakly monotone ($\widetilde{\mathfrak{M}}_\tau$), the map $s \mapsto \tau'(x + s\mathbf{1})$ is nondecreasing and continuous; by the intermediate value theorem, there exists $s' \in (-\varepsilon/\sqrt{L}, \varepsilon/\sqrt{L})$ with $\tau'(x + s'\mathbf{1}) = m'$, provided δ is small enough (choose δ so that the IVT window covers m'). Set $x' = x + s'\mathbf{1}$. Then $x' \in B_\varepsilon(x) \subseteq V$, $x' \in X^{\max}(m', \tau', \kappa')$, and $\tau'(x') = m'$, so $x' \in \mathcal{X}(m', \tau', \kappa') \cap V$.

B.2 Proofs for Chapter 3

3.1 Game

Two states, $i \in \{1, 2\}$, simultaneously choose a force level $m_i \in \mathbb{R}_+$. Their payoffs are given by von Neumann–Morgenstern expected utility functions:

$$U_1(m_1, m_2) = \frac{\lambda m_1^\alpha}{\lambda m_1^\alpha + m_2^\alpha} \times (V - k(m_1 + m_2)),$$

$$U_2(m_1, m_2) = \frac{m_2^\alpha}{\lambda m_1^\alpha + m_2^\alpha} \times (V - k(m_1 + m_2)),$$

where:

1. *$\lambda \in \mathbb{R}_{>0}$ captures the relative effectiveness of the forces;*

2. $\alpha \in (0, 1]$ *captures the decisiveness of superior force;*

3. $V \in \mathbb{R}_{>0}$ *captures the value of the prize; and*

4. $k \in (0, 1]$ *captures the inverse-recuperability of militarization costs.*

The game has a unique Nash equilibrium, given by:

$$\left(m_1^*, m_2^* \right) = \left(\frac{\alpha}{1+\alpha} \cdot \frac{V}{k} \cdot \frac{\lambda^{-\frac{1}{1+\alpha}}}{1 + \lambda^{-\frac{1}{1+\alpha}}}, \frac{\alpha}{1+\alpha} \cdot \frac{V}{k} \cdot \frac{1}{1 + \lambda^{-\frac{1}{1+\alpha}}} \right),$$

and this solution is continuous in all parameters. [*Proof*.]

Proof. We show that the unique Nash equilibrium is interior and derive its closed form.

Boundary cases. At $(m_1, m_2) = (0, 0)$, set $p_i(0, 0) = \lambda_i / (\lambda_1 + \lambda_2)$ with $\lambda_1 = \lambda$, $\lambda_2 = 1$. A unilateral deviation to $m_1 = \varepsilon$ yields $U_1(\varepsilon, 0) = V - k\varepsilon$, which exceeds $U_1(0, 0) = \lambda V / (\lambda + 1)$ for $\varepsilon < V / (k(\lambda + 1))$. Similarly, $(0, m_2)$ with $m_2 > 0$ is not an equilibrium: Player 2 can reduce m_2 and increase her payoff. By symmetry, the equilibrium must be interior.

First-order conditions. Define $\rho = m_1 / m_2$. The FOCs reduce to the ratio equation

$$1 + \frac{1}{\lambda \rho^\alpha} = \rho + \lambda \rho^{1+\alpha}. \tag{B.1}$$

Substituting $x = \lambda \rho^{1+\alpha}$ yields $x^2 + (\rho - 1)x - \rho = 0$, whose positive root is $x = 1$. Hence $\lambda \rho^{1+\alpha} = 1$ and $\tilde{\rho} = \lambda^{-1/(1+\alpha)}$. Substituting back into the level FOC gives

$$m_2^* = \frac{\alpha}{1+\alpha} \cdot \frac{V}{k} \cdot \frac{1}{1 + \lambda^{-\frac{1}{1+\alpha}}}, \qquad m_1^* = \tilde{\rho} \, m_2^*.$$

Second-order verification. Let $h_i = \log \circ U_i$. A direct computation shows

$$\frac{\partial^2 h_1}{\partial m_1^2} = -\frac{\alpha m_2^\alpha((1+\alpha)\lambda m_1^\alpha + m_2^\alpha)}{m_1^2(\lambda m_1^\alpha + m_2^\alpha)^2} - \frac{k^2}{(V - k(m_1 + m_2))^2} < 0,$$

using $V - k(m_1^* + m_2^*) > 0$ (which follows from $k(m_1 + m_2)/V = \alpha(1 + \rho)/(\alpha(1+\rho)+1+1/(\lambda\rho^\alpha)) < 1$). Thus U_1 is log-concave, hence quasiconcave, in m_1. The same holds for Player 2, so the FOC solution is the unique Nash equilibrium.

Parameter continuity. By inspection. [*Back to the text*.] ∎

B.3 Proofs for Chapter 4

4.1 Lemma

For all $(m, \tau, \kappa) \in M \times \mathcal{T} \times \mathcal{K}$, SPP (m, τ, κ) admits a unique solution. [*Proof*.]

Proof. We establish existence and uniqueness jointly.

Existence. By the hemicontinuity argument of Section B.1.2, the constraint set $\mathcal{X}(m, \tau, \kappa)$ is nonempty and compact. The cost function κ is continuous ($\mathfrak{C}_\kappa$), so by the Weierstrass extreme value theorem, κ attains its minimum on $\mathcal{X}$.

Uniqueness. Suppose $x_0 \neq x_1$ both solve the SPP. Then $\tau(x_0) = \tau(x_1) = m$ and $\kappa(x_0) = \kappa(x_1)$. Set $x_{1/2} = \frac{1}{2}x_0 + \frac{1}{2}x_1$. Since τ is quasiconcave (a consequence of log-concavity $\widetilde{\mathfrak{L}}_\tau$; see Section A.7.1), $\tau(x_{1/2}) \geqslant m$. Since κ is strictly quasiconvex (from exp-convexity $\mathfrak{L}_\kappa$), $\kappa(x_{1/2}) < \kappa(x_0)$. If $\tau(x_{1/2}) = m$, this contradicts optimality of x_0. If $\tau(x_{1/2}) > m$, then by ray monotonicity ($\widetilde{\mathfrak{M}}_\tau$) and continuity, there exists $t^* \in (0,1)$ with $\tau(t^* x_{1/2}) = m$ and $\kappa(t^* x_{1/2}) < \kappa(x_0)$ (using strict monotonicity $\mathfrak{M}_\kappa$), again a contradiction. [*Back to the text*.] ∎

4.2 Lemma

The solution to SPP (m, τ, κ) varies continuously with m, τ, and κ. [*Proof*.]

Proof. We will apply Berge's Maximum Theorem (Aliprantis and Border 2006, Thm 17.31). In Section B.1.2 we showed that the constraint correspondence $\mathcal{X}$ is nonempty, compact-valued, and continuous. The objective $(x; \kappa) \mapsto \kappa(x)$ is continuous. Berge's theorem then implies that the value function $v(m, \tau, \kappa) = \min_{x \in \mathcal{X}} \kappa(x)$ is continuous and that the argmin correspondence is upper hemicontinuous. Since the argmin is a singleton (by uniqueness, Lemma 4.1), upper hemicontinuity of a singleton-valued correspondence is equivalent to continuity of the selection. Hence $x^*(m, \tau, \kappa)$ is continuous in (m, τ, κ). [*Back to the text*.] ∎

4.3 Corollary

The policy function $\pi_{\tau, \kappa} : M \to X$ varies continuously with τ and κ.   [*Proof*.]

Proof. Fix $(\tau^*, \kappa^*) \in \mathcal{T} \times \mathcal{K}$ and a compact $K \subseteq M$. By Lemma 4.2, the raw solution $(m, \tau, \kappa) \mapsto x^*(m, \tau, \kappa)$ is jointly continuous, hence uniformly continuous on $K \times \{\tau^*\} \times \{\kappa^*\}$. For any $\varepsilon > 0$ there is a neighborhood $\mathcal{U}$ of (τ^*, κ^*) with $\sup_{m \in K} \|\pi_{\tau, \kappa}(m) - \pi_{\tau^*, \kappa^*}(m)\| < \varepsilon$ for all $(\tau, \kappa) \in \mathcal{U}$. This is continuity in the compact-open topology. [*Back to the text*.] ∎

4.4 Lemma

The policy function $\pi_{\tau, \kappa} : M \to X$ satisfies:

1. *Centeredness $(\mathfrak{o}_\pi)$: $\pi_{\tau, \kappa}(0) = 0$;*

2. *Coerciveness $(\mathfrak{O}_\pi)$: $\lim_{m \to \infty} \|\pi_{\tau, \kappa}(m)\| = \infty$; and*

3. *Weak Monotonicity $(\widetilde{\mathfrak{M}}_\pi)$: $m_1 \leq m_2 \implies \pi_{\tau, \kappa}(m_1) \leq \pi_{\tau, \kappa}(m_2)$ componentwise.*

[*Proof*.]

Proof. We address each property in turn.

Centeredness. $\pi_{\tau, \kappa}(0) = 0$ because the unique feasible point at $m = 0$ is $x = 0$ (from $\tau(0) = 0$ and $\kappa(0) = 0$).

Coerciveness. Suppose for contradiction that $m_n \to \infty$ but $\|\pi_{\tau,\kappa}(m_n)\| \leq R$ for all n. On the compact box $[0, R]^L$, continuity of τ gives a finite maximum $\overline{m}_R$. But $m_n = \tau(\pi_{\tau,\kappa}(m_n)) \leq \overline{m}_R$, contradicting $m_n \to \infty$.

Weak monotonicity. For $m \geq 0$, define the upper contour set $F(m) = \{x \in X : \tau(x) \geq m\}$. This set is nonempty (by $\mathfrak{R}_\tau$), convex (by $\widetilde{\mathfrak{Q}}_\tau$), closed (by $\mathfrak{C}_\tau$), and an upper set (by $\widetilde{\mathfrak{M}}_\tau$). We claim:

(i) $F(m)$ has a unique minimal element. Choose $x^0 \in F(m)$. On the compact set $F(m) \cap [0, x^0]$, the linear functional $\mathbf{1}^\top x$ attains a minimum at some $u(m)$. For uniqueness, if $x_1 \neq x_2$ are both minimal, then strict concavity of $\log(1 + \tau)$ implies $z = \frac{1}{2}x_1 + \frac{1}{2}x_2 \in \operatorname{int} F(m)$, so $z - \varepsilon\mathbf{1} \in F(m)$ for small $\varepsilon > 0$, contradicting minimality.

(ii) The optimizer is the minimal element. Since κ is strictly increasing in every coordinate ($\mathfrak{M}_\kappa$) and $F(m)$ is an upper set, κ is uniquely minimized at the minimal element: $\pi_{\tau,\kappa}(m) = u(m)$.

(iii) u is monotone. For $m_1 < m_2$, we have $F(m_2) \subset F(m_1)$. Since $u(m_2) \in F(m_1)$ and $u(m_1)$ is the minimal element of $F(m_1)$, we get $u(m_1) \leq u(m_2)$ coordinatewise. [*Back to the text*.] ∎

4.5 Lemma

The function spaces $\mathcal{T}$ and $\mathcal{K}$ are contractible. [*Proof*.]

Proof. We construct explicit strong deformation retractions for both spaces.

Homotopy for $\mathcal{T}$. Define the basepoint $\tau_0(x) = \sum_{\ell \in L} \log(1 + x_\ell)$. That $\tau_0 \in \mathcal{T}$ follows from the primer's discussion of the canonical technology: continuity and weak monotonicity are immediate; ray-surjectivity holds along $v = \mathbf{1}$ because $\tau_0(t\mathbf{1}) = L\log(1 + t) \to \infty$; and log-concavity holds because $\log(1 + \tau_0(x)) = \log(1 + \sum_\ell \log(1 + x_\ell))$ is concave (its Hessian is $-A/(1 + \tau_0) - (A + uu^\top)/(1 + \tau_0)^2 < 0$, where $A = \operatorname{diag}(1/(1 + x_\ell)^2)$ and $u = (1/(1 + x_\ell))_\ell$).

Define the homotopy

$$H_\tau(\tau, t)(x) = (1 + \tau(x))^{1-t} \cdot (1 + \tau_0(x))^t - 1. \tag{B.2}$$

We verify:

1. $H_\tau(\tau, t) \in \mathcal{T}$ for all $\tau \in \mathcal{T}, t \in [0, 1]$. Log-concavity: $\log(1 + H_\tau(\tau, t)) = (1 - t) \log(1 + \tau) + t \log(1 + \tau_0)$ is a convex combination of concave functions. Weak monotonicity: both factors $(1 + \tau)^{1-t}$ and $(1 + \tau_0)^t$ are nondecreasing compositions of nondecreasing functions. Continuity: immediate. Ray-surjectivity: along a witnessing ray v, $(1 + \tau(sv))^{1-t}(1 + \tau_0(sv))^t$ is a product of strictly increasing, unbounded functions, hence strictly increasing and unbounded.

2. H_τ *is continuous in* (τ, t). The map $(x; \tau, t) \mapsto (1 + \tau(x))^{1-t}(1 + \tau_0(x))^t - 1$ is jointly continuous on $[0, n]^L \times \mathcal{T} \times [0, 1]$ for each n, so H_τ is continuous in the Fréchet metric.

3. *Boundary conditions.* $H_\tau(\tau, 0) = \tau; H_\tau(\tau, 1) = \tau_0; H_\tau(\tau_0, t) = (1 + \tau_0)^{1-t}(1 + \tau_0)^t - 1 = \tau_0.$

Homotopy for $\mathcal{K}$. Define the basepoint $\kappa_0(x) = \sum_{\ell \in L} x_\ell$, which satisfies all axioms of $\mathcal{K}$ (exp-convexity: $\exp(\kappa_0)$ is the composition of an increasing strictly convex function with a linear function).

Define

$$H_\kappa(\kappa, t)(x) = \log\left((1 - t) \exp \kappa(x) + t \exp \kappa_0(x)\right). \tag{B.3}$$

Verification:

1. $H_\kappa(\kappa, t) \in \mathcal{K}$. Exp-convexity: $\exp(H_\kappa(\kappa, t)) = (1 - t) \exp \kappa + t \exp \kappa_0$ is a convex combination of strictly convex functions. Centeredness: at $x = 0$, $(1 - t) \cdot 1 + t \cdot 1 = 1$, so $H_\kappa(\kappa, t)(0) = 0$. Coerciveness, monotonicity, and continuity follow similarly.

2. H_κ *is continuous in* (κ, t). Same joint-continuity argument as for H_τ.

3. *Boundary conditions.* $H_\kappa(\kappa, 0) = \kappa; H_\kappa(\kappa, 1) = \kappa_0; H_\kappa(\kappa_0, t) = \log((1 - t + t) \exp \kappa_0) = \kappa_0.$

Both H_τ and H_κ are strong deformation retractions, so $\mathcal{T}$ and $\mathcal{K}$ are contractible. [*Back to the text.*] ∎

4.6 Proposition

$\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ *strongly deformation retracts onto the point*

$$\pi_0(m) = \left(\exp\left(\frac{m}{L} \right) - 1 \right) \mathbf{1},$$

where $\mathbf{1} \in \mathbb{R}^L$ *is the vector of ones. In particular,* $\mathcal{P}_{\mathcal{T} \times \mathcal{K}}$ *is contractible.* [*Proof*.]

Proof. We proceed in two steps: an intuitive argument via path connectedness, and then a canonical lift that upgrades this to a deformation retraction.

Step 1: Path connectedness. By Lemma 4.5, there exist strong deformation retractions H_τ and H_κ onto basepoints $\tau_0(x) = \sum_\ell \log(1 + x_\ell)$ and $\kappa_0(x) = \sum_\ell x_\ell$. For any $(\tau, \kappa) \in \mathcal{T} \times \mathcal{K}$, define $(\tau_t, \kappa_t) = (H_\tau(\tau, t), H_\kappa(\kappa, t))$ and let $\pi_t = \pi_{\tau_t, \kappa_t}$. By Corollary 4.3, the solution map $S : (\tau, \kappa) \mapsto \pi_{\tau, \kappa}$ is continuous, so $t \mapsto \pi_t$ is a continuous path in $\mathcal{P}$ from $\pi_{\tau, \kappa}$ to π_{τ_0, κ_0}.

However, since distinct parameter pairs may generate the same policy, this path can depend on the choice of representative. To obtain a homotopy on $\mathcal{P}$ itself, we must identify a continuous choice of (τ, κ) for each π.

Step 2: Canonical lift. Given $\pi \in \mathcal{P}$, define

$$\tau_\pi(x) = \sup\{m \in M : \pi(m) \leq x\}. \tag{B.4}$$

$\tau_\pi \in \mathcal{T}$: Define $g(x) = \log(1 + \tau_\pi(x))$ and the upper contour sets $U_\pi(m) = \{x : \tau_\pi(x) \geq m\} = \pi(m) + \mathbb{R}_+^L$. To show g is concave ($\widetilde{\mathfrak{L}}_\tau$), it suffices to verify the Minkowski-sum inclusion $\lambda U_\pi(m_1) + (1 - \lambda) U_\pi(m_2) \subseteq U_\pi(m_\lambda)$ for all $m_1, m_2 \in M$, $\lambda \in (0, 1)$, and $m_\lambda = (1 + m_1)^\lambda (1 + m_2)^{1-\lambda} - 1$. The left side equals $\lambda \pi(m_1) + (1 - \lambda) \pi(m_2) + \mathbb{R}_+^L$, so the inclusion reduces to showing $\pi(m_\lambda) \leq \lambda \pi(m_1) + (1 - \lambda) \pi(m_2)$ coordinatewise. Since $\pi \in \mathcal{P}$, there exist $\tau \in \mathcal{T}$ and $\kappa \in \mathcal{K}$ with $\pi = \pi_{\tau, \kappa}$. Let $x_i = \pi(m_i)$. Concavity of $\log(1 + \tau)$ gives $\tau(\lambda x_1 + (1 - \lambda) x_2) \geq m_\lambda$, so $\lambda x_1 + (1 - \lambda) x_2$ is feasible at level m_λ. By Lemma 4.4, $\pi(m_\lambda)$ is the coordinatewise-minimal element of $U(m_\lambda) = \{x : \tau(x) \geq m_\lambda\}$; hence $\pi(m_\lambda) \leq \lambda x_1 + (1 - \lambda) x_2$ and the inclusion holds.

Weak monotonicity ($\widetilde{\mathfrak{M}}_\tau$) follows from $x \leq y$ implying $\{m : \pi(m) \leq x\} \subseteq \{m : \pi(m) \leq y\}$. Continuity ($\mathfrak{C}_\tau$) follows from concavity on the interior and a squeeze argument on the boundary. Ray-surjectivity ($\mathfrak{R}_\tau$) holds along generic rays $v \gg 0$: the map $s \mapsto \tau_\pi(sv)$ is the generalized inverse of

$m \mapsto \max_\ell \pi_\ell(m)/v_\ell$, which is strictly increasing for generic v (the bad set is a countable union of hypersurfaces).

$\pi \mapsto \tau_\pi$ *is continuous:* Fix $\pi_0 \in \mathcal{P}$, a compact $K \subseteq X$, and $\epsilon > 0$. Set $\overline{m}_K = \max_{x \in K} \tau_{\pi_0}(x)$ and define $f(m, x) = \min_\ell(x_\ell - \pi_{0,\ell}(m))$. If we have $\sup_{m \in [0, \overline{m}_K+1]} \|\pi(m) - \pi_0(m)\| < \eta$, then $|f_\pi(m, x) - f(m, x)| \leq \eta$ uniformly on compacts. A covering argument on K (using upper and lower bounds from the zero sets of f and f_π) yields $\max_{x \in K} |\tau_\pi(x) - \tau_{\pi_0}(x)| < 2\epsilon$ for η small enough.

Defining κ_π: Fix $\alpha \in \mathbb{R}^L_{++}$ and set $\kappa_\pi(x) = \alpha \cdot x + \varepsilon \|x\|^2$. This does not depend on π, lies in $\mathcal{K}$ (strict exp-convexity follows from strict convexity), and is compatible with τ_π: the feasible set at level m is $\pi(m) + \mathbb{R}^L_+$, on which any strictly coordinate-increasing function is uniquely minimized at the minimal element $\pi(m)$.

Completing the homotopy. For $\pi \in \mathcal{P}$, define the homotopy $(\tau_t, \kappa_t) = (H_\tau(\tau_\pi, t), H_\kappa(\kappa_\pi, t))$ and $\pi_t = \pi_{\tau_t, \kappa_t}$. By continuity of the lift $\pi \mapsto (\tau_\pi, \kappa_\pi)$, the homotopies H_τ and H_κ, and the solution map S, the map $(\pi, t) \mapsto \pi_t$ is continuous. At $t = 0$, $\pi_0 = \pi_{\tau_\pi, \kappa_\pi} = \pi$. At $t = 1$, $\pi_1 = \pi_{\tau_0, \kappa_0}$ for all π. Hence $\mathcal{P}$ deformation-retracts to the single point π_{τ_0, κ_0}.

Identifying the basepoint. The SPP for (τ_0, κ_0) has the Lagrangian FOC $1 - \lambda/(1 + x_\ell) = 0$ for all ℓ, giving $x_\ell = \lambda - 1$ uniformly. The constraint $\sum_\ell \log(1 + x_\ell) = m$ yields $x_\ell = \exp(m/L) - 1$, so

$$\pi_{\tau_0, \kappa_0}(m) = (\exp(m/L) - 1)\mathbf{1}.$$

[*Back to the text*.] ∎

B.4 Proofs for Chapter 5

5.2 Proposition

Game 5.1 has at least one pure-strategy Nash equilibrium. [*Proof*.]

Proof. The contest success function is discontinuous at $(\tau_1(x_1), \tau_2(x_2)) = (0, 0)$, precluding standard existence theorems. We apply Reny (1999), whose sufficient conditions are: (a) strategy sets are nonempty, compact, and convex; (b) payoffs are bounded and quasiconcave in own action; (c) the game is better-reply secure.

Strategy sets. Since each κ_i is coercive, there exists a compact box $\overline{X}$ beyond which no equilibrium can lie.

Quasiconcavity. Fix the opponent's action and write $\log U_i = \log p_i + \log A$, where p_i is the contest probability and $A = V - k(\exp \kappa_1(x_1) + \exp \kappa_2(x_2) - 2)$ is the residual surplus.

Concavity of $\log p_i$: Let $w_i = \log \lambda_i + \alpha \log \tau_i(x_i)$. Since $\log(1 + \tau_i)$ is concave $(\tilde{\mathfrak{L}}_\tau)$, one obtains $\nabla^2 \log \tau_i \preceq -\nabla \tau_i \nabla \tau_i^\top / (\tau_i^2(1 + \tau_i)) \preceq 0$ on $\{\tau_i > 0\}$, so w_i is concave. With the opponent fixed at constant $C = e^{w_{-i}} > 0$, $\log p_i = w_i - \log(e^{w_i} + C)$; the map $w \mapsto w - \log(e^w + C)$ is increasing and strictly concave, so $\log p_i$ is concave.

Concavity of $\log A$: With $B_i = \exp \kappa_i$ strictly convex $(\mathfrak{L}_\kappa)$, $\nabla^2_{x_i} \log A = -(k/A)\nabla^2 B_i - (k/A)^2 \nabla B_i \nabla B_i^\top \preceq 0$.

Hence $\log U_i$ is a sum of concave functions, making U_i log-concave and therefore quasiconcave.

Better-reply secureness. Let $D = S_1 \times S_2$ where $S_i = \{x_i : \tau_i(x_i) = 0\}$. Payoffs are continuous off D, so only D requires attention. Fix $(x_1^*, x_2^*) \in D$ and $\epsilon > 0$. Choose $\hat{x}_i$ with $\tau_i(\hat{x}_i) = \bar{t} > 0$ small enough that $k(\exp \kappa_i(\hat{x}_i) - 1) \leq \epsilon V / 4$. Set $\delta_\tau = (\epsilon \lambda_i / (1 - \epsilon))^{1/\alpha} \bar{t}$. By continuity of τ_{-i} near $x_{-i}^* \in S_{-i}$, there is a neighborhood V_{-i} on which $\tau_{-i}(y_{-i}) \leq \delta_\tau$, giving $p_i(\hat{x}_i, y_{-i}) \geq 1 - \epsilon$. Combined with cost control near x_{-i}^*,

$$U_i(\hat{x}_i, y_{-i}) \geq (1 - \epsilon)(V - \epsilon V / 2)$$

for all $y_{-i} \in V_{-i}$. This is a secure improvement, and since $U_i \leq V$ everywhere, the game is better-reply secure. By Reny's theorem, a pure-strategy Nash equilibrium exists. [*Back to the text*.] ∎

5.6 Proposition (Regularization)

There exists a continuous regularization operator

$$\mathfrak{D} : \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]},$$

such that $\mathfrak{D}|_{\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}} = \mathrm{id}$; *in other words, $\mathfrak{D}$ fixes the tame functions.* [*Proof*.]

Proof. We construct $\mathfrak{D}$ via the three-step procedure described in Chapter 5 and verify the required properties. The argument is given for technologies; the cost case is parallel.

Construction. For $\tau \in \mathcal{T}$, set $\phi = \log(1 + \tau)$, which is continuous, concave, and coordinatewise nondecreasing on X.

1. *Curvature chart.* Extend ϕ to all of $\mathbb{R}^L$ via the monotone concave extension
$$\tilde{\phi}(x) = \inf_{y \in X, g \in \partial\phi(y)}\{\phi(y) + \langle g, x - y\rangle\}.$$

2. *Causal convolution.* For $\varepsilon > 0$, choose a smooth nonnegative kernel η_ε supported in $[0, c\varepsilon]^L$ with $\int \eta_\varepsilon = 1$, and set $\tilde{\phi}_\varepsilon = \tilde{\phi} * \eta_\varepsilon$; define $\phi_\varepsilon = \tilde{\phi}_\varepsilon|_X$.

3. *Gauge.* Fix a compact $K \subset X$ with nonempty interior and a reference $x^* \in \text{int}(K)$. Define the projection residual by the formula

$$R_{\mathcal{T}}(\tau) = \inf_{A>0, \beta\in\Delta_L} \sup_{x\in K} \left| \frac{\phi(x) - A\sum_\ell \beta_\ell \log(1 + x_\ell)}{\sum_\ell \beta_\ell \log(1 + x_\ell^*)} \right|,$$

choose a continuous strictly increasing gauge $\Theta : [0, \infty) \to [0, \bar{\varepsilon}]$ with $\Theta(0) = 0$, and set $\varepsilon(\tau) = \Theta(R_{\mathcal{T}}(\tau))$.

The regularized technology is $\mathfrak{D}(\tau) = \exp(\phi_{\varepsilon(\tau)}) - 1$.

Class preservation. Convolution of a concave function with a nonneg. kernel yields a concave function, so ϕ_ε is concave on X; hence $\log(1 + \mathfrak{D}(\tau)) = \phi_\varepsilon$ is concave. The causal (positive-orthant) support of η_ε ensures that convolution preserves coordinatewise monotonicity: if $x \leq y$, then $x - u \leq y - u$ for $u \geq 0$, so $\tilde{\phi}_\varepsilon(x) \leq \tilde{\phi}_\varepsilon(y)$. Ray-surjectivity is inherited: along a witnessing ray v, $\phi_\varepsilon(sv)$ is a bounded-window average of the unbounded function $\tilde{\phi}(sv - \cdot)$, hence unbounded; concavity then forces strict increase. Continuity of $\mathfrak{D}(\tau)$ follows from smoothness of convolution.

Smoothness. For $\varepsilon > 0$, $\tilde{\phi}_\varepsilon$ is C^∞ (convolution with a smooth kernel), so $\mathfrak{D}(\tau) = \exp(\phi_\varepsilon) - 1$ is C^∞ on X.

Fixed points. If $\tau \in \mathcal{T}^{[\mathfrak{T}]}$, then $\phi = A\sum_\ell \beta_\ell \log(1 + x_\ell)$ matches a tame template exactly, so $R_{\mathcal{T}}(\tau) = 0$, $\varepsilon(\tau) = 0$, $\phi_0 = \phi$, and $\mathfrak{D}(\tau) = \tau$.

Continuity in τ. If $\tau_n \to \tau$ in the Fréchet metric, then $\phi_n \to \phi$ uniformly on compacts; the monotone concave extension operator preserves this convergence (equicontinuity of supergradients on compacts); the convolution operator $(f, \varepsilon) \mapsto f * \eta_\varepsilon$ is jointly continuous; and $\varepsilon(\cdot)$ is continuous via the Berge-type argument on the compact parameter set $\mathcal{A}_\phi \times \Delta_L$. Composing, $\mathfrak{D}(\tau_n) \to \mathfrak{D}(\tau)$ in d.

Cost regularization. The cost operator $\mathfrak{D} : \mathcal{K} \to \mathcal{K}$ is constructed identically, replacing $\phi = \log(1 + \tau)$ with $\psi = \exp \kappa$, using the analogous extension, convolution, and gauge. Each translate $z \mapsto \widetilde{\psi}(z - u)$ is strictly convex on X (because $\widetilde{\psi} = \psi$ on X and $u \in [0, c\varepsilon]^L$). Since η_ε has full support in $[0, c\varepsilon]^L$, the integral $\psi_\varepsilon = \int \widetilde{\psi}(\cdot - u) \eta_\varepsilon(u) \, du$ is a positive-measure average of strictly convex functions, hence strictly convex. The fixed-point property on tame costs follows from $R_\mathcal{K}(\kappa) = 0$ when κ is tame. [*Back to the text*.] ∎

5.9 Proposition (Structural adequacy of the tame representation)
$\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ *is structurally adequate for* $\mathcal{T} \times \mathcal{K}$. [*Proof*.]

Proof. The homotopy from $\mathcal{T} \times \mathcal{K}$ to $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ proceeds in two phases.

Phase 1 ($t \in [0, 1/2]$): Regularize. For technologies, define

$$\tau_t = (1 + \tau)^{1-2t} \cdot (1 + \mathfrak{D}_\tau(\tau))^{2t} - 1.$$

For costs, define

$$\kappa_t = \log\big((1 - 2t) \exp \kappa + 2t \exp \mathfrak{D}_\kappa(\kappa)\big).$$

At $t = 0$: $(\tau_0, \kappa_0) = (\tau, \kappa)$. At $t = 1/2$: $(\tau_{1/2}, \kappa_{1/2}) = (\mathfrak{D}_\tau(\tau), \mathfrak{D}_\kappa(\kappa))$.

Phase 2 ($t \in [1/2, 1]$): Tamify. Replace τ with $\mathfrak{D}_\tau(\tau)$ and $\mathfrak{D}_\tau(\tau)$ with $(\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau)$ in the same homotopy formulas. At $t = 1$, we have $(\tau_1, \kappa_1) = \big((\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau), (\mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)(\kappa)\big) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$.

Admissibility along the path. At every t, $\log(1+\tau_t)$ is a convex combination of concave functions, hence concave ($\tilde{\mathfrak{L}}_\tau$). Weak monotonicity and ray-surjectivity transfer because each $\phi_i = \log(1 + \tau_i)$ is nondecreasing and the geometric-mean interpolation preserves these properties (the product of nondecreasing unbounded functions along a ray is nondecreasing and unbounded). Similarly, $\exp(\kappa_t)$ is a convex combination of strictly convex functions ($\mathfrak{L}_\kappa$); centeredness, coerciveness, and monotonicity pass through.

Continuity. H is continuous in (τ, κ, t) because $\mathfrak{D}$ and $\mathfrak{T}$ are continuous (the former by Proposition 5.6; the latter by continuity of the gradient-reading map at $x = 0$ on the smooth outputs of $\mathfrak{D}$).

Fixity. If (τ, κ) is tame, then $\mathfrak{D}$ and $\mathfrak{T}$ both act as the identity, so $H(t, \tau, \kappa) = (\tau, \kappa)$ for all t.

Hence H is a strong deformation retraction from $\mathcal{T} \times \mathcal{K}$ onto $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, establishing structural adequacy. [*Back to the text*.] ∎

5.12 Proposition (Convexity of the tame state space)

$\mathcal{P}_{\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}}$ *is a convex set.* [*Proof*.]

Proof. Fix $m > 0$ (the case $m = 0$ is trivial). Let π_0, π_1 be tame policies with parameters $(A_{\tau,0}, \beta_0, A_{\kappa,0}, q_0)$ and $(A_{\tau,1}, \beta_1, A_{\kappa,1}, q_1)$. Define $\pi_t = (1-t)\pi_0 + t\pi_1$ for $t \in [0, 1]$. We construct tame parameters $(A_\tau(t), \beta_t, A_\kappa(t), q_t)$ such that π_t satisfies the first-order conditions of the SPP.

Step 1: Choose q_t. Set $q_t = 1/L$ (any strictly positive vector works).

Step 2: Define β_t. Set $c_t = (\sum_j (1 + \pi_t^j) q_t^j)^{-1}$ and $\beta_t^\ell = c_t(1 + \pi_t^\ell) q_t^\ell$. Then $\beta_t \in \Delta_L$ with $\beta_t \gg 0$.

Step 3: Set scale parameters. Let $A_\kappa(t) = 1$ and $A_\tau(t) = m / \sum_\ell \beta_t^\ell \log(1+\pi_t^\ell)$. The feasibility condition $A_\tau(t) \sum_\ell \beta_t^\ell \log(1 + \pi_t^\ell) = m$ holds by construction.

Step 4: Determine multipliers. Fix $\lambda_t = A_\kappa(t)/(A_\tau(t)\,c_t)$ and fix $\eta_t^\ell = A_\kappa(t)\,q_t^\ell - \lambda_t A_\tau(t)\beta_t^\ell/(1 + \pi_t^\ell)$. By the definitions of β_t and c_t, $\eta_t^\ell = 0$ for all ℓ, so complementarity holds automatically.

Verification. The stationarity conditions $A_\kappa q_t^\ell - \lambda_t A_\tau(t)\beta_t^\ell/(1 + \pi_t^\ell) - \eta_t^\ell = 0$ hold for every ℓ and every $t \in [0, 1]$. Hence π_t solves the SPP for the tame pair (τ_t, κ_t), and $\pi_t \in \mathcal{P}_{\mathcal{T}[\mathfrak{T}] \times \mathcal{K}[\mathfrak{T}]}$. Since $t \mapsto \pi_t$ is affine, the tame state space is convex. [*Back to the text*.] ∎

5.13 Proposition (Water-filling equilibrium)

Suppose Game 5.1 is tame*: for each player $i \in \{1, 2\}$,*

$$\tau_i(x_i) = A_{\tau,i} \sum_{\ell \in L} \beta_{i,\ell} \log(1 + x_{i,\ell}), \qquad \kappa_i(x_i) = A_{\kappa,i} \sum_{\ell \in L} q_{i,\ell}\, x_{i,\ell},$$

with $A_{\tau,i}, A_{\kappa,i} > 0$, $\beta_i, q_i \in \Delta_L$, all entries strictly positive. Then any Nash equilibrium (x_1^*, x_2^*) with strictly positive allocations has the water-filling form*: for each player i there exists a scalar $c_i > 0$ such that*

$$x_{i,\ell}^* = \max\left\{0,\; c_i \frac{\beta_{i,\ell}}{q_{i,\ell}} - 1\right\} \qquad \text{for all } \ell \in L, \tag{5.1}$$

and the scalar c_i solves the one-dimensional fixed point

$$c_i = \frac{(1 - p_i)\,\alpha\,A_{\tau,i}}{k\,A_{\kappa,i}}\,\frac{W(x_1^*, x_2^*)}{e^{\kappa_i(x_i^*)}\,\tau_i(x_i^*)}, \tag{5.2}$$

where p_i is the contest success probability and $W = V - k\big(e^{\kappa_1(x_1^*)} + e^{\kappa_2(x_2^*)} - 2\big)$. If all inputs are active, then

$$\tau_i(x_i^*) = A_{\tau,i}\Big(\log c_i + \sum_{\ell \in L} \beta_{i,\ell} \log \frac{\beta_{i,\ell}}{q_{i,\ell}}\Big), \qquad \kappa_i(x_i^*) = A_{\kappa,i}\,(c_i - 1),$$

so that (5.2) becomes a scalar equation in c_i alone. [*Proof*.]

Proof. The tame technology and cost are $\tau_i(x_i) = A_{\tau,i} \sum_\ell \beta_{i,\ell} \log(1 + x_{i,\ell})$ and $\kappa_i(x_i) = A_{\kappa,i} \sum_\ell q_{i,\ell} x_{i,\ell}$. By Lemma 5.11, the interior first-order conditions for the SPP at level m_i are

$$A_{\kappa,i} q_{i,\ell} = \lambda_i \cdot \frac{A_{\tau,i} \beta_{i,\ell}}{1 + x_{i,\ell}}, \qquad \ell \in L. \tag{B.5}$$

Solving for $x_{i,\ell}$:

$$x_{i,\ell} = \frac{\lambda_i A_{\tau,i} \beta_{i,\ell}}{A_{\kappa,i} q_{i,\ell}} - 1.$$

This is nonneg. only if the right side exceeds zero, giving the water-filling structure:

$$x_{i,\ell} = \left[\frac{\lambda_i A_{\tau,i} \beta_{i,\ell}}{A_{\kappa,i} q_{i,\ell}} - 1 \right]^+ .$$

The multiplier λ_i is determined by the feasibility constraint $\tau_i(x_i) = m_i$, which becomes

$$A_{\tau,i} \sum_\ell \beta_{i,\ell} \log\left(1 + \left[\frac{\lambda_i A_{\tau,i} \beta_{i,\ell}}{A_{\kappa,i} q_{i,\ell}} - 1 \right]^+ \right) = m_i.$$

The left side is continuous and strictly increasing in λ_i (from 0 to ∞), so the intermediate value theorem gives a unique $\lambda_i(m_i)$.

When λ_i is large enough that all inputs are active ($x_{i,\ell} > 0$ for all ℓ), the constraint simplifies to

$$A_{\tau,i} \sum_\ell \beta_{i,\ell} \log\left(\frac{\lambda_i A_{\tau,i} \beta_{i,\ell}}{A_{\kappa,i} q_{i,\ell}} \right) = m_i,$$

a scalar equation in λ_i alone, consistent with the fixed-point equation stated in the proposition.

The cost at the optimum is

$$\kappa_i(x_i^*) = A_{\kappa,i} \sum_\ell q_{i,\ell} x_{i,\ell}^* = \lambda_i A_{\tau,i} \sum_\ell \beta_{i,\ell} - A_{\kappa,i} \sum_\ell q_{i,\ell} = \lambda_i A_{\tau,i} - A_{\kappa,i},$$

using $\sum_\ell \beta_{i,\ell} = 1 = \sum_\ell q_{i,\ell}$ (when all inputs are active and the sums are over Δ_L). This is the exponential cost expression stated in the proposition. [*Back to the text*.] ∎

B.5   Proofs for Chapter 6

6.1 Proposition

Let $a > 1$, $L \geq 2$, and consider the CES technology $\tau_a(x) = \left(\sum_\ell x_\ell^a \right)^{1/a}$ with linear cost $\kappa(x) = q \cdot x$, $q \in \mathbb{R}_{>0}^L$. Then τ_a satisfies $\mathfrak{C}_\tau$, $\mathfrak{R}_\tau$, and $\widetilde{\mathfrak{M}}_\tau$, but violates $\widetilde{\mathfrak{L}}_\tau$. For every $m > 0$:

1. *Every global minimum of the SPP is a vertex solution $x^{(\ell)} = m \cdot e_\ell$ (complete specialization in a single commodity);*

2. *The global minimum is unique if and only if the cheapest commodity is unique: $\left| \operatorname{argmin}_\ell q_\ell \right| = 1$; and*

3. *The bifurcation set $\mathcal{B}_a = \{ q \in \mathbb{R}_{>0}^L : \left| \operatorname{argmin}_\ell q_\ell \right| > 1 \}$ is a union of $\binom{L}{2}$ hyperplanes of codimension one.*

[*Proof.*]

Proof. Fix $a > 1$, $m > 0$, and $q \in \mathbb{R}_{>0}^L$. The SPP is to minimize $q \cdot x$ subject to $\left(\sum_\ell x_\ell^a \right)^{1/a} = m$ and $x \geq 0$.

Interior critical point. The Lagrangian on the interior $\{ x \gg 0 : \|x\|_a = m \}$ gives the stationarity condition $q_\ell = \lambda \cdot x_\ell^{a-1} / m^{a-1}$ for each ℓ. Solving: $x_\ell = (q_\ell / \lambda)^{1/(a-1)} m^{(a-1)/(a-1)} \propto q_\ell^{1/(a-1)}$. After normalization by the constraint $\|x\|_a = m$, the cost at the interior critical point is

$$q \cdot x^{\text{int}} = m \cdot \|q\|_p,$$

where $p = a/(a-1)$ is the Hölder conjugate of a.

Vertex cost. At the vertex $x^{(\ell)} = m\, e_\ell$, the cost is $q \cdot x^{(\ell)} = m\, q_\ell$. The cheapest vertex costs $m\, \min_\ell q_\ell$.

Comparison. For $L \geq 2$ and q not proportional to a unit vector, Hölder's inequality gives $\|q\|_p > \min_\ell q_\ell$ (strict because $p > 1$ and the q_ℓ are not all equal to their minimum). Hence $q \cdot x^{\text{int}} > m \min_\ell q_\ell$, and the interior critical point is suboptimal. By induction on faces (each lower-dimensional face is itself a CES problem with $a > 1$ on fewer commodities), the global minimum lies at a vertex.

Uniqueness and bifurcation. The global minimum is unique if and only if $|\operatorname{argmin}_\ell q_\ell| = 1$. The bifurcation set $\mathcal{B}_a = \{q : |\operatorname{argmin}_\ell q_\ell| > 1\} = \bigcup_{\ell < k} \{q : q_\ell = q_k\}$ is a union of $\binom{L}{2}$ hyperplanes. [*Back to the text.*] ∎

6.2 Proposition (Cusp catastrophe in the CES family)

Let $L = 2$ and consider the CES technology τ_a of Proposition 6.1 with cost $\kappa_\epsilon(x) = x_1 + (1 + \epsilon) x_2$, parametrized by $(a, \epsilon) \in (0, \infty) \times \mathbb{R}$.

1. *For each $a > 1$, the bifurcation set is $\mathcal{B}_a = \{\epsilon = 0\}$, a single point in the ϵ-line. As $a \downarrow 1$, the two solution branches $x^A = (m, 0)$ and $x^B = (0, m)$ coalesce: the cost gap $\kappa(x^B) - \kappa(x^A) = |\epsilon| m$ closes, and at $a = 1$ the level sets become convex, the interior minimum becomes the unique global minimum, and the bifurcation vanishes.*

2. *The point $(a, \epsilon) = (1, 0)$ is a* cusp *point: the bifurcation locus $\{a > 1\} \times \{0\}$ terminates there, and the two-parameter family undergoes a cusp catastrophe in the sense of* Arnold *(2004, Ch. 9).*

3. *Let $\mathcal{U} = \{(a, \epsilon) : a > 1, \epsilon \neq 0\}$ be the complement of the bifurcation set in the wild regime. The fundamental group $\pi_1(\mathcal{U})$ is nontrivial: a loop in the (a, ϵ)-plane that encircles the cusp point induces a* monodromy*—continuous tracking of the global minimum around the loop returns a different branch to the starting parameters. The monodromy action is the transposition $x^A \leftrightarrow x^B$.*

[*Proof.*]

Proof. We take $L = 2$ with $\tau_a(x) = (x_1^a + x_2^a)^{1/a}$ and $\kappa_\epsilon(x) = x_1 + (1 + \epsilon) x_2$, parametrized by (a, ϵ).

Part (i): Bifurcation structure. By Proposition 6.1, for $a > 1$ the global minima are vertices $x^A = (m, 0)$ with cost m and $x^B = (0, m)$ with cost $(1 + \epsilon)m$. Branch A is optimal for $\epsilon > 0$, branch B for $\epsilon < 0$, and both are optimal at $\epsilon = 0$: the bifurcation set is $\mathcal{B}_a = \{\epsilon = 0\}$. As $a \downarrow 1$, the level sets become convex, the interior minimum becomes the unique global minimum, and the cost gap $|\epsilon|m$ between branches closes.

Part (ii): Cusp geometry. The bifurcation locus in the (a, ϵ)-plane is the half-line $\{a > 1\} \times \{0\}$. It terminates at $(1, 0)$: for $a \leq 1$, the level sets of τ_a are convex and the SPP has a unique interior solution, so no bifurcation occurs. This is the defining geometry of a cusp catastrophe in the sense of Arnold (2004, Ch. 9): a fold locus (codimension 1) that terminates at a degenerate point (codimension 2) where two solution branches coalesce.

Part (iii): Monodromy. Let $\mathcal{U} = \{(a, \epsilon) : a > 1, \ \epsilon \neq 0\}$ and consider the loop $\gamma(t) = (1 + r\cos 2\pi t, \ r\sin 2\pi t)$ centered at $(1, 0)$ with radius r small enough that the loop passes through $a < 1$ at its leftmost point.

Start at $(1 + r, 0)$ on branch A (cost m). As t increases, $\epsilon = r\sin 2\pi t$ moves from 0^+ through positive values. Branch A remains the unique optimum as long as $\epsilon > 0$. Near $t = 1/2$, the loop crosses into $a < 1$ territory, where the level sets are convex and the solution is a unique interior minimum—the two branches have merged. The transition at $a = 1$ is continuous: at $a = 1$ with $\epsilon > 0$, the level set $\|x\|_1 = m$ is the simplex, on which $(1, 1+\epsilon) \cdot x$ is uniquely minimized at $(m, 0) = x^A$; for a slightly below 1 the interior minimizer converges to this vertex. Continuing, the loop returns to $a > 1$ with $\epsilon < 0$, where branch B is now the global optimum (with the same continuity at the return crossing, since the interior minimizer converges to $(0, m) = x^B$ when $\epsilon < 0$).

Upon return to $(1 + r, 0)$, continuous tracking of the global minimum has mapped $x^A \mapsto x^B$. A second traversal returns to x^A. Hence $\pi_1(\mathcal{U})$ acts on the branches via the transposition $x^A \leftrightarrow x^B$, and the monodromy is nontrivial. [*Back to the text.*] ∎

Bibliography

Aliprantis, Charalambos D. and Kim C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Third ed. Berlin: Springer.

Alspector-Kelly, Marc. 2001. "On Quine on Carnap on Ontology." *Philosophical Studies* 102(1):93–122.

Arnold, Vladimir I. 2004. *Catastrophe Theory*. Springer Science 3rd ed. Berlin: Springer.

Badiou, Alain. 2005. *Being and Event*. London: Continuum. Originally published as *L'Être et l'événement*, Paris: Seuil, 1988.

Balasko, Yves. 2011. *General Equilibrium Theory of Value*. Princeton, NJ: Princeton University Press.

Bartelson, Jens. 1995. *A Genealogy of Sovereignty*. Cambridge: Cambridge University Press.

Bartelson, Jens. 2001. *The Critique of the State*. Cambridge: Cambridge University Press.

Berge, Claude. 1963. *Topological Spaces*. Edinburgh: Oliver and Boyd. Translated by E. M. Patterson.

Besley, Timothy and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton, NJ: Princeton University Press.

Beviá, Carmen and Luis C. Corchón. 2010. "Peace Agreements Without Commitment." *Games and Economic Behavior* 68(2):469–487.

Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.

Bird, Graham H. 1995. "Carnap and Quine: Internal and External Questions." *Erkenntnis* 42(1):41–64.

Broadberry, Stephen and Mark Harrison. 2005. The Economics of World War I: An Overview. In *The Economics of World War I*, ed. Stephen Broadberry and Mark Harrison. Cambridge: Cambridge University Press.

Brooks, Risa A. 2008. *Shaping Strategy: The Civil-Military Politics of Strategic Assessment*. Princeton: Princeton University Press.

Cappella Zielinski, Rosella. 2016. *How States Pay for Wars*. Ithaca, NY: Cornell University Press.

Carnap, Rudolf. 1950. "Empiricism, Semantics, and Ontology." *Revue Internationale de Philosophie* 4(11):20–40.

Carroll, Robert J. 2025. "A Theory of Force." Unpublished manuscript.

Carroll, Robert J. 2026. "A State Is Its Relations: The Yoneda Lemma and Relational Identity in International Relations." Unpublished manuscript.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Coe, Andrew J. and Jane Vaynman. 2020. "Why Arms Control Is So Rare." *American Political Science Review* 114(2):342–355.

Cover, Thomas M. and Joy A. Thomas. 1990. *Elements of Information Theory*. Second (2006) ed. Hoboken, NJ: Wiley.

Crawford, Neta C. 2022. *The Pentagon, Climate Change, and War: Charting the Rise and Fall of U.S. Military Emissions*. Cambridge, MA: MIT Press.

Diewert, W. Erwin. 1971. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy* 79(3):481–507.

Downing, Brian M. 1993. *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. First paperback print ed. Princeton, NJ: Princeton University Press.

Enloe, Cynthia. 2000. *Maneuvers: The International Politics of Militarizing Women's Lives*. Berkeley, CA: University of California Press.

Epstein, Brian. 2015. *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford: Oxford University Press.

Ertman, Thomas. 1997. *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge: Cambridge University Press.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James D. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* 72(3):523–559.

Feaver, Peter D. 2003. *Armed Servants: Agency, Oversight, and Civil-Military Relations*. Cambridge, MA: Harvard University Press.

Federle, Jonathan, André Meier, Gernot J. Müller, Willi Mutschler and Moritz Schularick. 2026. "The Price of War." *American Economic Review* 116(3):791–827.

Flores-Macías, Gustavo A. and Sarah E. Kreps. 2015. "Borrowing Support for War: The Effect of War Finance on Public Attitudes toward Conflict." *Journal of Conflict Resolution* 61(5):997–1020.

Fodor, Jerry A. 1974. "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)." *Synthese* 28(2):97–115.

Friedman, Milton. 1953. The Methodology of Positive Economics. In *Essays in Positive Economics*. Chicago: University of Chicago Press pp. 3–43.

Garfinkel, Michelle R. 1990. "Arming as a Strategic Investment in a Cooperative Equilibrium." *American Economic Review* 80(1):50–68.

Gilpin, Robert. 1981. *War and Change in World Politics*. Cambridge: Cambridge University Press.

Gruber, Thomas R. 1993. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition* 5(2):199–220.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Hatcher, Allen. 2002. *Algebraic Topology*. Cambridge: Cambridge University Press.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, NJ: Princeton University Press.

Hertz, Heinrich. 1894. *Die Prinzipien der Mechanik, in neuem Zusammenhange dargestellt*. Leipzig: Johann Ambrosius Barth. English translation: *The Principles of Mechanics Presented in a New Form*, trans. D. E. Jones and J. T. Walley, Macmillan, 1899.

Hintze, Otto. 1975. Military Organization and the Organization of the State. In *The Historical Essays of Otto Hintze*, ed. Felix Gilbert. Oxford, UK: Oxford University Press pp. 175–215.

Hirshleifer, Jack. 1991. "The Paradox of Power." *Economics & Politics* 3(3):177–200.

Hodler, Roland and Hadi Yektaş. 2012. "All-Pay War." *Games and Economic Behavior* 74(2):526–540.

Horowitz, Michael C. and Matthew S. Levendusky. 2011. "Drafting Support for War: Conscription and Mass Support for Warfare." *Journal of Politics* 73(2):524–534.

Jackson, Robert H. 1990. *Quasi-States: Sovereignty, International Relations, and the Third World*. Cambridge: Cambridge University Press.

Keegan, John. 1978. *The Face of Battle*. Harmondsworth, UK: Penguin.

Kier, Elizabeth. 2021. *War and Democracy: Labor and the Politics of Peace*. Ithaca, NY: Cornell University Press.

Kriner, Douglas L. and Francis X. Shen. 2016. "Conscription, Inequality, and Partisan Support for War." *Journal of Conflict Resolution* 60(8):1419–1445.

Lake, David A. 1992. "Powerful Pacifists: Democratic States and War." *American Political Science Review* 86(1):24–37.

Lane, Frederic Chapin. 1979. *Profits from Power: Readings in Protection Rent and Violence-Controlling Enterprises*. Albany, NY: State University of New York Press.

Leander, Anna. 2004. "Drafting Community: Understanding the Fate of Conscription." *Armed Forces & Society* 30(4):571–599.

Levi, Margaret. 1988. *Of Rule and Revenue*. Number 9 *in* "California Series on Social Choice and Political Economy" first ed. Berkeley, CA: University of California Press.

Levi, Margaret. 1996. "The Institution of Conscription." *Social Science History* 20(1):133–167.

Lowe, E.J. 2005. Ontology. In *The Oxford Companion to Philosophy*, ed. Ted Honderich. Second ed. Oxford, UK: Oxford University Press.

Manzano, María. 2005. *Extensions of First Order Logic*. Number 19 *in* "Cambridge Tracts in Theoretical Computer Science" Cambridge: Cambridge University Press.

Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.

Meirowitz, Adam and Anne E. Sartori. 2008. "Strategic Uncertainty as a Cause of War." *Quarterly Journal of Political Science* 3(4):327–352.

Munkres, James R. 2000. *Topology*. 2nd ed. Upper Saddle River, NJ: Prentice Hall.

North, Douglass C. 1981. *Structure and Change in Economic History*. New York: Norton.

Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87(3):567–576.

Parker, Geoffrey. 1996. *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800*. Second ed. Cambridge, UK: Cambridge University Press.

Pierson, Paul. 2000. "Increasing Returns, Path Dependence, and the Study of Politics." *American Political Science Review* 94(2):251–267.

Posen, Barry R. 1984. *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars*. Cornell Studies in Security Affairs Ithaca: Cornell University Press.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1):115–132.

Putnam, Hilary. 1975. *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press.

Quine, Willard V. 1948. "On What There Is." *Review of Metaphysics* 2(5):21–38.

Quine, Willard V. 1951. "On Carnap's Views on Ontology." *Philosophical Studies* 2(5):65–72.

Ramey, Valerie A. and Matthew D. Shapiro. 1998. Costly Capital Reallocation and the Effects of Government Spending. Working Paper 6283 National Bureau of Economic Research.

Reno, William. 1998. *Warlord Politics and African States*. Boulder, CO: Lynne Rienner.

Reny, Philip J. 1999. "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games." *Econometrica* 67(5):1029–1056.

Resende-Santos, João. 2007. *Neorealism, States, and the Modern Mass Army*. Cambridge, MA: Cambridge University Press.

Ringmar, Erik. 1996. "On the Ontological Status of the State." *European Journal of International Relations* 2(4):439–466.

Rosen, Stephen Peter. 1991. *Winning the Next War: Innovation and the Modern Military*. Ithaca, NY: Cornell University Press.

Satia, Priya. 2018. *Empire of Guns: The Violent Making of the Industrial Revolution*. Stanford, CA: Stanford University Press.

Siegel, Jennifer. 2014. *For Peace and Money: French and British Finance in the Service of Tsars and Commissars*. Oxford, UK: Oxford University Press.

Skaperdas, Stergios. 1992. "Cooperation, Conflict, and Power in the Absence of Property Rights." *American Economic Review* 82(4):720–739.

Slantchev, Branislav L. 2012. "Borrowed Power: Debt Finance and the Resort to Arms." *American Political Science Review* 106(4):787–809.

Stasavage, David. 2011. *States of Credit: Size, Power, and the Development of European Polities*. Princeton: Princeton University Press.

Sundaram, Rangarajan K. 1996. *A First Course in Optimization Theory*. Cambridge, UK: Cambridge University Press.

Talmadge, Caitlin. 2015. *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes*. Ithaca, NY: Cornell University Press.

Thom, René. 1975. *Structural Stability and Morphogenesis*. Reading, MA: W. A. Benjamin. Translated by D. H. Fowler.

Thomasson, Amie L. 2015. *Ontology Made Easy*. Oxford: Oxford University Press.

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990–1992*. Revised (1992) paperback ed. Cambridge, MA: Blackwell.

Van Creveld, Martin. 1985. *Command in War*. Cambridge, MA: Harvard University Press.

Van Creveld, Martin. 2004. *Supplying War: Logistics from Wallenstein to Patton*. Second ed. Cambridge, UK: Cambridge University Press.

van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Boston, MA: McGraw-Hill.

Weber, Max. 1922. *Economy and Society*. Tribe-translated (2019) ed. Cambridge, Massachusetts: Harvard University Press.

Weisberg, Michael. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Wendt, Alexander. 1999. *Social Theory of International Politics*. Number 67 *in* "Cambridge Studies in International Relations" Cambridge, UK: Cambridge University Press.

Wight, Colin. 2006. *Agents, Structures and International Relations: Politics as Ontology*. Cambridge: Cambridge University Press.