The Shape of Possible States[*]

Robert J. Carroll[†]

December 31, 2025

Abstract

Many theories of international relations assume that states are "like units"—but what, precisely, does it mean to say that states are the same? I argue that sameness must be grounded in the properties of the class of all possible states. Modeling the state as a producer of force, I show that this class forms a single, unified component with trivial structural properties: in a precise mathematical sense, all states are qualitatively alike. I then develop a meta-theory of state modeling, propose a formal criterion for adequate representation, and demonstrate that a well-behaved subclass preserves the structure of the full class while exhibiting stronger quantitative regularities. The analysis derives as a theorem the foundational premise that states are alike, opening the door to a new mode of theorizing about state behavior.

[*]Version `2.1` (December 2025). Previous versions of this paper—under working title "The State as a Producer of Force"—have benefitted from presentations at the University of Illinois at Urbana-Champaign's Student-Faculty Seminar and Princeton University's Political Economy Seminar, not to mention feedback from students in UIUC's graduate seminar in international political economy (P.S. 582): Ina Azifa, Do Young Gong, Ruth Holloway, Seung-Uk Huh, Jaeyoung Lee, Ikromjon Tuhtasunov, and Paul Un. I am further grateful to Ekrem Baser, Jake Bowers, Don Casler, Tyson Chatagnier, Germán Gieczewski, Nick Grossman, Matias Iaryczower, Gleason Judd, Brenton Kenkel, Avital Livny, Bob Pahre, Gino Pauselli, Alyssa Prorok, Kris Ramsay, Joe Ruggiero, Matt Winters, and Elaine Yao for feedback and helpful conversations. All errors are my own.

[†]University of Illinois at Urbana-Champaign. Email: `rjc@illinois.edu`

Contents

A system consists of interacting units whose collective behavior cannot be understood through isolated study. Nearly all systemic theories work from the premise that these units are similar to some degree and in some respect. In its strong form, the premise holds that units are identical in kind; only their interactions—and not their intrinsic characteristics—matter for explaining outcomes. Such theories help us understand firms in markets, birds in flocks, and even neurons in the brains that devise systemic theories in the first place.

So too 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 how they interact, 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 article begins from that puzzle. My aim is to clarify what it means to say that states are the same. 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, then, 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, the premise of "like units" is not merely a simplifying assumption; it is a structural truth about the domain of theory.

To approach this formally, I model the state as a producer of force: an agent that transforms resources into coercive capacity, subject to constraints imposed by a technology and a cost function. This is not an arbitrary starting point. From Weber (1922) and Hintze (1975) to Tilly (1975; 1990), the bellicist tradition has held that the state's defining activity lies in the organization of coercion. As Tilly's slogan has it, "war made the state, and the state made war" (Tilly, 1975, p. 42). Even if war no longer makes states, the production and management of coercive power remains a necessary, if not sufficient, feature of statehood. To model the state in this way is to begin from the one dimension of its activity that all states share. It is 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. The resulting policy map $\pi_{\tau,\kappa} : M \to X$ specifies, for each m, how the state mobilizes. The collection of all such maps, across all admissible technologies and costs, constitutes the *set of possible states*. Studying the structure of this set—its topology and geometry—is a way of studying the logic of the state itself. It also lets us ask, in rigorous terms, what sameness among states means.

The analysis proceeds in two parts. First, the topological. Under minimal assumptions, the space of states $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is *contractible*. Contractibility is a strong form of connectedness: every point can be continuously deformed into every other without leaving the space. In effect, the class of all possible states is topologically trivial—any member can represent any other. In systemic terms, qualitative variation is second-order. Defined by force production alone, there is only one kind of state. Indeed, there is a ladder of sameness: as we move from mere connectedness to contractibility, states are alike in increasingly strong senses.

Second, the modeling. If the space of possible states is contractible, what kinds of models can represent it adequately? I introduce a subclass of functions—log-linear technologies and linear costs—that is both tractable and faithful to the general form. These *tame states* preserve the topological simplicity of the full space but exhibit a richer geometry: they form a convex set. Convexity implies not only unity but also linearity; it permits interpolation, optimization, and equilibrium analysis. In this sense, the tame model is not just a simplification but an illumination: a representation that makes the underlying geometry of statehood visible.

Foundations. This analysis builds on three intellectual traditions.

1. *The ontological status of the state.* Erik Ringmar (1996) distinguishes between realist and pluralist views: realists treat the state as a pre-given actor, pluralists as an emergent bundle of sub-units. Ringmar suggests that we adjudicate metaphorically—by asking what the state is *like.* He analogizes the state to a person, maintaining a longstanding tradition; here, I analogize it to a firm.[1] Both metaphors are useful, but metaphor should not be the end of the ontological exercise. Following Quine's (1948) dictum that "to be is to be the value of a variable," I treat ontology as a kind of set theory. As Lowe (2005) puts it, an ontology is "the set of things whose existence is acknowledged by a particular theory." To posit the state, on this view, is to posit a *set of states*; the defining properties of that set become the defining properties of the state. The difficulty of ontologizing the state is part of what makes it unavoidable. As Bartelson (2001) argues, even those who seek to reject the state as a concept often find it creeping back in. Recent work continues to take up these questions directly (Epstein, 2013; Hay, 2014; Jessop, 2014; Knio, 2023).

[1] "Metaphorizing" and "modeling" are not the same. In modeling, I formalize force production as the action of an agent, as in Alexander Wendt's *Social Theory of International Politics* (1999).

2. *The state as a firm.* Frederic Lane (1979) envisioned the state as a firm providing services—among them protection—in exchange for revenue. Similar ideas appear in Douglass North (1981, Ch. 3), Margaret Levi (1988), David Lake (1992), and Elizabeth Kier (2021). These accounts differ in detail but 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: a device mapping inputs (resources) to outputs (force), subject to technological and cost constraints. This shares the structure of a firm's problem in microeconomic theory (Mas-Colell, Whinston and Green, 1995; Chambers, 1988). The point is not to reduce the state to a firm, but to use the firm-as-producer template as a clean starting point. Metaphor becomes model becomes ontology—but only along one axis of what the state might be.

3. *The bellicist tradition.* From Weber and Hintze through Tilly, a longstanding line of thought holds that states are born in war. In this view, the coercive apparatus emerges from, and is shaped by, violent conflict. Here I take no stand on whether warmaking is constitutive of statehood or 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 encompass others. One could, for instance, use this structure to model variation in force production as a function of territory or infrastructural power.

Contributions. The paper contributes in four ways.

1. **It clarifies the "like unit" premise.** Structural realism and kindred approaches treat states as functionally identical, but this claim has been under-theorized. I provide a formal framework to interrogate the assumption on its own terms and determine when, and in what sense, states can be treated as the same. The claim becomes a theorem rather than a heuristic.

2. **It constructs a general yet spare model of the state.** Modeling the state as a producer of force isolates a core feature of systemic behavior while remaining agnostic about institutional detail. It provides enough structure to support geometric and topological analysis without loading in unnecessary assumptions.

3. **It introduces global methods.** The technical contribution lies less in novel mathematics than in a shift of focus. Using elementary topology and convex analysis, I study *global* properties of the state space: not how it looks near any particular state, but how it looks as a whole.

4. **It provides a theory of adequate representation.** Recognizing that models simplify, I ask when a model of states can represent the broader space. I show that the tame states are a strong deformation retract of the general model and that they form a convex set. This explains how simpler models preserve structural insights while gaining analytic tractability; it also cautions that not all model properties reflect properties of the space, urging humility even when representation is adequate.

Two limits are worth highlighting. First, I have taken only one slice of what states do. They tax, build, regulate, legitimate; they make and maintain order. But if the state has many faces, this is the one it cannot do without, and it furnishes a clean foundation for formal analysis. Second, what follows is a study of a function in isolation. A state is more than a function; its relations matter. Later work will reconstruct state identity from its pattern of relations with others—from the transformations it sends and receives. For now, the claim is modest: if we look only at the production of force, the class of all possible states is unified; and within that unity, a tractable, faithful model is available.

Roadmap. The paper proceeds in four sections. Section 1 introduces the model of state force production and defines the set of possible states. Section 2 analyzes the topology of that set, showing that it is contractible. Section 3 introduces the tame states and shows that they form a convex subset that adequately represents the full space. Section 4 discusses the implications of these results for systemic theorizing about the state. Appendix A contains proofs and technical details.

1 The State as a Producer of Force

We now turn to the formal model of state militarization. Our goal is to construct a family of models that characterize how states mobilize resources to achieve desired force levels. To do so, we will define the primitives of the model, then formulate the state's production problem. This will yield a general framework for analyzing state militarization decisions.

1.1 Motivating Game

To orient the problem, however, we begin *in media res.* Consider the following game drawn from the strategic arming literature.

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:

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

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

This game is a stylized representation of a militarization contest between two states. Each state chooses a militarization level, which determines both its probability of winning a prize and the costs it incurs. The probability of winning is determined by a contest function, which depends on the relative militarization levels and the parameter α. The costs of militarization are linear in the sum of the militarization levels, scaled by the parameter k. The game has a unique Nash equilibrium in pure strategies, which can be found by solving the first-order conditions of the expected utility functions.[2]

As was noted in the introduction, this game is representative of a large class of models in the international relations literature. Authors in this tradition use a wide vareity of terms to describe the choice variable m_i. For example, Carmen Beviá and Luis C. Corchón (2010) and Jack Hirshleifer (1991) simply call it "war efforts," a concept one might measure in francs, battalions, or barrels

[2]A caveat is in order: here in the main text, we have ignored the corner possibility where both players choose zero militarization. It is shown in the appendix that this corner solution is never an equilibrium when the contest takes its usual form at zero.

of oil. Michelle R. Garfinkel (1990) and Robert Powell (1993) both call it a "good," suggesting an output of spending that does not accumulate across periods. Adam Meirowitz and Anne E. Sartori (2008, fn. 8, p. 333) 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."[3] "Arms" is also used by Stergios Skaperdas (1992) and Andrew J. Coe and Jane Vaynman (2020). Roland Hodler and Hadi Yektaş (2012) go so far as to call it "power," whereas James D. Fearon (2018) lands on what will be our preferred term: "force level." None of the results that follow—nor those in the literature cited above—depend on the specific terminology used, but to the degree that we wish to imbue the model with state-centric meaning, "force level" seems most appropriate.

Irrespective of the terminology, however, one question remains unaddressed in these models.

---

2 Question

How is the force level m_i^* produced by State i in models like Game 1, and what does this tell us about the structure of states?

We seek to answer this question by constructing a model of state militarization. To do so, we will define the primitives of the model: the state's desired force level, its technology for converting resources into force, and its cost of mobilizing resources. With these in hand, we will formulate the state's production problem, which will yield a general framework for analyzing state militarization decisions. The remainder of this section is devoted to this task.

1.2 A Program for Modeling State Militarization

In light of Question 2, we propose the following program.

---

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] Meirowitz and Sartori study a far more reduced—and elegant—version of the interaction where the military investment happens first, is privately known, and then is input into a very general function; their focus is less on the terms of battle and more on how uncertainty influences bargaining under incomplete information.

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

Call such a map the opportunity cost of militarization.

A baseline task is to convince ourselves that this program points us toward a reasonable (albeit minimalistic) characterization of state behavior. To that end, we now define the primitives of the model.

Force. The first primitive of the model is the state's desired force level. This is the choice variable in Game 1, and it represents the quantity of military power the state wishes to field. We can think of a player in Game 1 as first choosing a desired force level, then coming to us to ask how to produce it. Since that player most commonly chooses over non-negative real numbers, we let $M := \mathbb{R}_{\geqslant 0}$ denote the set of all possible desired force levels.

Resources. Force is made from stuff, and we call that stuff *resources*. We suppose the resources arrive in different types, which we call *commodities*. We index the set of all commodities by L, where we suppose L is nonempty and finite. (Abusively, L will sometimes refer to the cardinality of the set of commodities, and this should introduce no confusion.) Each commodity $\ell \in L$ is a good or service that the state can mobilize—*e.g.,* steel, oil, labor, or fresh-cut flowers. In common interpretations of models like this one, these commodities may be differentiated not just by their physical properties but also their time of delivery, location of delivery, or state of the world in which they are delivered. We will not consider these interpretational complications here, but they are often important in practice. In case $L = 1$, the model collapses to one without across-commodity trade-offs.

Mobilization plans. The state's decision is to decide how much steel, oil, labor, and so on to mobilize to achieve its desired force. This is a decision about investment, in the sense that the state is choosing to allocate resources to a particular end. We refer to a particular decision about how to allocate resources as a *mobilization plan*, encoded as a vector $x \in X := \mathbb{R}^L_{\geqslant 0}$. We refer to the ℓth element of x as x_ℓ, which is the amount of commodity ℓ that the state mobilizes.

Invisible parameters. Naturally, the state does not make its decisions in a vacuum; it must consider the world around it. In the interest of keeping things simple, we will ignore several features likely to be relevant to the state's decision-making process. Of course, one cannot make an exhaustive list of all factors one has ignored in a given model, but we can at least mention a few likely candidates.

1. We will not include market conditions, most notably the prices of commodities. It seems obvious that the cost of mobilization depends in part on these prices—indeed, a textbook approach might be something like

$$\text{cost of mobilization at } x = \sum_{\ell} \text{price of commodity } \ell \times x_{\ell}.$$

   The influence of such prices on costs will be left implicit.

2. We will not include the state's endowment of resources, nor the territory over which it has control (which influences those endowments). It is likely the case that a state's militarization technology depends on its endowment of resources; for example, a state with a large endowment of coal may be more competent at converting coal into force. The same may go for how the state experiences costs when mobilizing resources. Again, the influences of territory and endowments will be left implicit.

One can and should think of the influence of such factors, but for now, we will think of these as *invisible parameters* influencing the two functions we define next.

1.3 The Technology of Militarization

We now turn to the state's technology for converting resources into force. In words, the state's technology is a rule telling us how much force the state can produce using a given mobilization plan. It is the machine that turns stone into hatchets, bronze into shields, steel into tanks, and labor into soldiers. This is a fundamental aspect of the state's decision-making process, as it is the mechanism by which the state converts resources into force, an important precursor to higher-order concerns like power or security.

The relevance of such technologies predates the state system, as the ability to convert resources into force is a fundamental aspect of human society once we move past the hunter-gatherer stage. Historian Ian Morris (2014, p. 7) observes that Stone Age societies were tiny and that violence among people was small in scale. (Nevertheless, some 10–20% of all people who lived in Stone Age societies died at the hands of other people.) We still see this in the anthropological record, where the study of violence in contemporary hunter-gatherer societies has long been a topic of interest. Ethnic groups vary in their propensity for violence; anthropologist Ruth Benedict (1934) classified societies as either "Apollonian" (authoritarian and warlike) or "Dionysian" (egalitarian and peaceful) based on this propensity. As societies transitioned to the Neolithic era, they developed agriculture, which allowed them to support larger populations. With larger populations came more complex social structures, including the emergence of states. States developed technologies for mobilizing resources and converting them into force, which allowed them to field larger armies and engage in more complex forms of warfare. Thus, the technology of militarization has been a fundamental aspect of human society for millennia.

The important role of technology, and particularly military technology, in shaping the behavior of states is well-recognized in military history. In a series of influential works, military historian Martin van Creveld focuses on how technological innovations have shaped the conduct of war (1991), with deeper institutional ramifications thanks to advances in supply chains (2004) and command (1985). Military historian Geoffrey Parker (1996) argues that technological innovations, such as the development of gunpowder and the printing press, played a crucial role in shaping the behavior of states during the early modern period. These technologies allowed states to field larger armies and engage in more complex forms of warfare, which in turn influenced their strategic behavior. Similarly, historian John Keegan (1978) emphasizes the importance of technology in shaping the conduct of war, arguing that technological innovations have often been the decisive factor in determining the outcome of battles and wars. Political scientist Brian Downing (1993) goes so far as to argue that the military revolution of the early modern period was a key driver not only in the development of the state as we know it, but in democracy as well. And perhaps most audaciously of all, historian Priya Satia (2018) argues that the military revolution was itself a driver of the industrial revolution, rather than the other way around as is commonly supposed.

These handful of works are far from exhaustive, but they illustrate the centrality of technology in shaping the behavior of states. To capture this in our model, we define the state's technology as a function, imposing several palatable properties on its shape.

---

4 Definition

The state's militarization technology *is a function*

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

We assume τ possesses the following properties:

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 $\left(\widetilde{\mathfrak{M}}_\tau\right)$: *$\tau$ is weakly increasing in all commodities; and*

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

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

   is concave.[4]

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

The state's technology is a simple machine. Mathematically speaking, it does little more than input a mobilization plan and output a scalar quantity of force. We impose four requirements on how the machine performs this task:

1. small changes in the mobilization plan must yield small changes in the force level;

2. any force level must be achievable by some mobilization plan;

3. no good hinders the production of force, and any force level may be augmented by adding more of *some* good; and

4. the force level does not experience returns to scale too quickly.

The set of functions $\mathcal{T}$—a *function space*—gathers all technologies that satisfy these properties. It is a space of possibilities, constrained solely by its domain, codomain, and the structural properties we have imposed.

The function space $\mathcal{T}$ contains many familiar forms of production functions. For example, the well-known Cobb-Douglas production function

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

where $A > 0$ and $\beta_\ell \in (0, 1)$ with $\sum_{\ell \in L} \beta_\ell \leqslant 1$, satisfies all four properties and is therefore an element of $\mathcal{T}$. Similarly, the Constant Elasticity of Substitution (CES) production function

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

[4] We use the term "log-concavity" here in a nonstandard way. Ordinarily *log-concavity* refers to functions f such that $\log(f(x))$ is concave. Here, we use $\log(1 + \tau(x))$ to ensure that the function is well-defined at $\tau(x) = 0$. Many a regression-runner has been burned by the logarithm's misbehavior at zero, and nearly all of them remedy this 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.

where $A > 0$, $\gamma_\ell > 0$, $\sigma \in (0,1]$, and $\rho \leqslant 1$ with $\rho \neq 0$, also satisfies all four properties and is an element of $\mathcal{T}$. We take as canonical a function reminiscent of the logarithmic utility function common in economics:

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

where $\beta_\ell > 0$ and $\sum_{\ell \in L} \beta_\ell = 1$. This function is particularly convenient for quantitative work, as we will see later. But what matters is not the specific form of any one technology, but rather the family of all technologies satisfying our structural properties.

1.4 The Cost of Militarization

The second major component of the state's decision-making environment is the cost the state experiences when mobilizing resources. This is no less fundamental than the state's technology, as the state must weigh the benefits of mobilization against the costs. Despite its massive importance, the cost of militarization is far less studied than the technology of militarization. In a tremendous contribution to an understudied problem, Rosella Cappella Zielinksi (2016) studies how states finance their militarization, finding that states rely on a variety of methods, including taxation, borrowing, and printing money. Her work highlights the complexity of the cost of militarization, which likely depends on a variety of factors, including the state's fiscal capacity, the structure of its economy, and the political environment in which it operates. As such, one useful way to think about the cost of militarization is as a suite of channels sending resources in X to costs in $\mathbb{R}$. Certain commodities hit different channels more heavily than others; for example, conscripted labor may impose political costs, while oil may impose economic costs, while printing money may impose inflationary costs. These costs may be felt more at home, as Cappella Zielinski emphasizes, or abroad—say, in the form of international indebtedness as studied by Jennifer Siegel (2014).

But the notion of cost ought to transcend mere dollars and cents, which is difficult when the current order of magnitude for war costs is in the trillions of American dollars (Stiglitz and Bilmes, 2008). Militarization imposes diverse costs—economic, political, social, and environmental—that extend beyond immediate fiscal burdens. Conscription may generate political and social strain through inequality and contestation (Leander, 2004; Kriner and Shen, 2016; Horowitz and Levendusky, 2011; Asal, Conrad and Toronto, 2017; Levi, 1996). Debt and other war-financing mechanisms can create long-term fiscal and political distortions (Cappella Zielinski and Poast, 2024; Flores-Macías and Kreps, 2015; Slantchev, 2012; Mosley and Rosendorff, 2023; Siegel, 2014). Beyond these material costs, militarization produces ecological and social externalities: the environmental damage and carbon emissions associated with military activity (Crawford, 2022), and the gendered restructuring of labor, identity, and care work that accompany war and militarism (Enloe, 2000). Of course, these channels are neither exhaustive nor mutually exclusive, and despite their diversity, they all contribute to the overall cost of militarization.

Taken together, these literatures suggest that militarization imposes a multidimensional structure of costs. Each channel—fiscal, political, social, ecological—translates mobilization into a different kind of strain on the state. To capture this idea abstractly, we represent the state's experience of mobilization as a cost function: a mapping from a mobilization plan in X to a scalar measure of cost in $\mathbb{R}$. This formalization does not privilege any single source of cost but instead provides a general framework within which particular mechanisms—taxation, conscription, borrowing, or social disruption—can be modeled as components of a unified cost surface.

---

5 Definition

The state's cost function *is a function*

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

We assume κ possesses the following properties:

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

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

3. Coerciveness ($\mathfrak{O}_\kappa$): $\kappa(x) \to \infty$ *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 $\mathcal{K}$.

The state's cost function is another simple machine. Mathematically speaking, it inputs a mobilization plan and outputs a scalar quantity of cost. We impose five requirements on how the machine performs this task:

1. small changes in the mobilization plan must yield small changes in the cost;
2. mobilizing nothing incurs no cost;
3. mobilizing more and more resources eventually becomes prohibitively expensive;
4. mobilizing more of any good always increases the cost; and
5. the cost does not experience returns to scale too quickly.

The set of functions $\mathcal{K}$—another function space—gathers all cost functions that satisfy these properties. It is a space of possibilities, constrained solely by its domain, codomain, and the structural properties we have imposed.

Once again, the function space $\mathcal{K}$ contains many familiar forms of cost functions. Canonical among these is the familiar linear cost function

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

where $q \in \mathbb{R}^L_{>0}$ is a vector of prices for each commodity. This function satisfies all five properties and is therefore an element of $\mathcal{K}$. Another familiar form is 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. This function also satisfies all five properties and is an element of $\mathcal{K}$. It's worth noting that there exist strictly concave cost functions as well—they need only be "less concave than log" to satisfy the exp-convexity property. For example, the function

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

where $\beta_\ell > 0$ and $\sum_{\ell \in L} \beta_\ell = 1$, satisfies all five properties and is an element of $\mathcal{K}$. Once again, what matters is not the specific form of any one cost function, but rather the family of all cost functions satisfying our structural properties.

1.5 The State's Production Problem

We have now defined the two major components of the state's decision-making environment: the state's technology for converting resources into force, and the state's cost of mobilizing resources. These represent the primary data that the state must consider when deciding how to mobilize resources to achieve its desired force. Though they are functions housed in function spaces, they are also the parameters of the state's decision-making environment.

But how do these data cohere? Naturally, we consider the organization of the production of force as a decision problem, where the state must choose a mobilization plan that minimizes the costs of mobilization while achieving its desired force level. This is the state's *production problem*, which we now define.

---

6 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*

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

$\mathsf{SPP}\,(m, \tau, \kappa)$ is the star of the show, the choice that makes a state a state in this most primitive sense of the word. This humble minimization problem is the most basic expression of the foundational decisions in organizing the production of force. It points toward the answer to the "how" question of militarization: given a desired force level like m_i^* from Game 1, how does State i mobilize resources to produce it? The state chooses a mobilization plan $x \in X$ that minimizes its cost of mobilization $\kappa(x)$ while achieving the desired force level $\tau(x) = m$.

Because we have been quite broad in our definitions of technologies and costs, the state's production problem is a general object that can be applied to a wide variety of contexts. τ might be good at converting stone, steel, or enriched uranium into force, and κ might demonstrate extreme sensitivity to the price of labor, or horses, or oil. Large or small, capitalist or socialist, ancient or modern, democratic or authoritarian, rich or poor—*all* states must solve the same basic problem: how to mobilize resources to achieve a desired force level at the lowest possible cost, where we recall that this cost is rather broadly construed.

1.6 The Currying Trick

But solving $\mathsf{SPP}\,(m, \tau, \kappa)$ is insufficient for purposes of characterizing the state, given the way the story began in Game 1. We need to know how the state will mobilize resources given a demanded force level as derived from strategic considerations. We therefore do not merely want to solve $\mathsf{SPP}\,(m, \tau, \kappa)$ for fixed m, τ, and κ; rather, we seek a function

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

that tells us how the state will mobilize resources given any desired force level. This is a *policy function*, as it prescribes the state's mobilization plan for any desired force level. To construct this function, we employ a standard technique from functional analysis called *currying*.[5] Currying is a method for transforming a function that takes multiple arguments into a sequence of functions that each take a single argument. In our case, we can curry the state's production problem to obtain the desired policy function.

---

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 by $\mathcal{P}_{\mathcal{T}\times\mathcal{K}} := \{\pi_{\tau,\kappa} : \tau \in \mathcal{T}, \kappa \in \mathcal{K}\}$.

The policy function $\pi_{\tau,\kappa}$ is the solution to the state's production problem for all desired force levels $m \in M$. It tells us how the state will mobilize resources given any desired force level, given its technology τ and cost function κ. The set $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is the family of all such policy functions, generated by all possible combinations of technologies and cost functions. This family is the main object of interest in this paper, as it tells us exactly how a given force level is produced given characteristics encoded in τ and κ. It is, for our purposes, the set of all states.

Needless to say, states do thing other than solve SPP (m, τ, κ). They govern, tax, legislate, police, and adjudicate; they build roads, schools, and hospitals; they provide public goods and services; they regulate markets and economies; they interact with other states diplomatically and militarily. All of these activities are important, but they are not the focus of this paper. Our focus is on the state's production of force, and the family of policy functions $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ captures this aspect of state behavior in a general and flexible way. Whether from above—as in structural IR theories like Waltz's—or from below—as in bellicist conceptions of the state like Tilly's—the state's production of force is a foundational aspect

[5] Currying is named for logician Haskell Curry, who formalized the technique in the 20th century. But the idea is older, dating to Frege and Church in the late 19th and early 20th centuries.

of its existence. In studying the family of policy functions $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$, we study this foundational aspect in a general and flexible way. We do so in a way, it is hoped, that will be useful in the study of these other crucial state functions as well.

2 The Structure of States

In the previous section, we introduced a simple cost minimization model designed to represent the basic problem of militarization. The modeling process culminated in Definition 6, which defined the state's production problem $\mathsf{SPP}(m, \tau, \kappa)$ in terms of a desired force level $m \in M$, a militarization technology $\tau \in \mathcal{T}$, and a cost function $\kappa \in \mathcal{K}$. In this section, we turn our attention to the solutions to $\mathsf{SPP}(m, \tau, \kappa)$.

2.1 Three Basic Questions

When faced with an optimization problem like $\mathsf{SPP}(m, \tau, \kappa)$, three questions naturally arise. The first two relate to whether the problem is well-posed:

1. *Existence:* does it admit at least one solution? If $\mathsf{SPP}(m, \tau, \kappa)$ does not admit a solution, then there is no way for the state to achieve its desired force level. This would be a serious problem, as it would imply that the state is unable to achieve its most basic goal.

2. *Uniqueness:* if it admits a solution, is this solution unique? If there are multiple solutions, then the state faces a second choice that, by definition, *cannot* be determined by the optimization problem itself. This multiplicity represents the limit of the model's precision, the modeler's control.

These questions are important, but in the present context they are more means than ends. Let us dispatch them without dawdling.

---

8 Lemma

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

Thus, the two concerns just raised are resolved.

The third question that arises when faced with an optimization problem is both more substantive and more capacious:

3. *Stability:* how does the solution change as the parameters of the problem change? Do changes in the desired force level, the militarization technology, or the cost function lead to changes in the solution? If so, are these changes mild or drastic? Do any patterns emerge in relating the data to the solution?

Here begin the questions of comparative statics. But in the background lurks a subtle and fascinating problem: two of the parameters in SPP (m, τ, κ) are themselves functions. What does this subtlety mean for the stability of the solution? How can we meaningfully speak of the solution as a function of the data, when the data is itself functional and infinite-dimensional?

Put differently: what sort of information do these independent variables carry? The desired force level $m \in M \coloneqq \mathbb{R}_{\geqslant 0}$ is a scalar, a simple non-negative number. It is easy to formulate questions like

> Will the amount of steel needed to produce a given force level increase or decrease as the force level increases? If so, by how much?

This is because $\mathbb{R}_{\geqslant 0}$ includes all sorts of structure that we take for granted, both theoretically and empirically. For example, it possesses an order structure, so the word "increase" makes sense; similarly, it possesses a metric structure, so the words "by how much" make sense. But what about the militarization technology $\tau \in \mathcal{T}$ and the cost function $\kappa \in \mathcal{K}$? What sort of structure do they possess? How can we formulate questions about their behavior?

The most basic question is: do the solutions to Problem SPP (m, τ, κ) change continuously as the data changes? We therefore need to equip the function spaces $\mathcal{T}$ and $\mathcal{K}$ with enough structure that we can make sense of continuity. This is a *topological* problem, topology being the branch of mathematics concerned with continuity and convergence. In the present context, we are interested in the topology of function spaces, which is a rich and fascinating subject in its own right. While the formal machinery is developed in the appendix, the core intuition is simple—and happily, it is spatial in nature. We equip the set of technologies $\mathcal{T}$ with a function that tells us how "close" two technologies are. This function is called a *metric,* and it works just like the familiar distance function on $\mathbb{R}$. Under the metric we define here, two technologies are close if they send all mobilization plans to force levels that are close. For example, the functions

$$\tau_0(x) = \sum_{\ell \in L} \log(1 + x_\ell) \quad \text{and} \quad \tau_1(x) = \sum_{\ell \in L} \log(1 + (1 \pm \varepsilon)x_\ell)$$

are close if ε is small, because they send similar mobilization plans to similar force levels. We can think of this creating an ε-ball around τ_0 that contains τ_1, just as we do with points in $\mathbb{R}$ when we speak of open intervals like $(-\varepsilon, \varepsilon)$. The cost functions $\mathcal{K}$ are equipped with a similar metric, which tells us how close two cost functions are. For example, the functions

$$\kappa_0(x) = q \cdot x \quad \text{and} \quad \kappa_1(x) = (1 \pm \varepsilon)q \cdot x, \quad \text{where} \quad q \in \mathbb{R}^L_{>0},$$

are close if ε is small, because they charge similar amounts for similar mobilization plans. To be sure, the problem at hand creates some complications—the interested reader is referred to Appendix A—but the basic idea is clear. Remarkably, this simple intuition allows us to ask questions like: does the solution to SPP (m, τ, κ) change continuously as force levels, technologies, and costs change? Can we, in fact, obtain well-defined comparative statics in this more general setting?

It turns out that we can.

---

9 Lemma

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

This result is the first step in understanding the structure of the state's production problem, as it suggests that we can learn about the solution by studying the data. We just noted that topology is the study of continuity and convergence. But it is also the study of structure, namely of the properties that are preserved under continuous transformations. We refer to such properties as *invariants*, and they are the key to understanding the structure of the state's production problem—and its solutions.

Moving toward comparative statics, we next observe that the policy function $\pi_{\tau,\kappa}$ defined in Definition 7—that is, the curried function mapping technologies and cost functions to mobilization plans—is itself continuous as a function of technologies and cost functions.

---

10 Corollary

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

Thus, not only does the solution to SPP (m, τ, κ) vary continuously with the data, but the entire policy function does as well. It is worth noting that these data—functions themselves—encode infinite-dimensional information. This makes the continuity results all the more powerful, as they tell us that even in this infinite-dimensional setting, the solutions behave nicely as the data change.

Finally, we report three key structural results about the policy function $\pi_{\tau,\kappa}$.

---

11 Lemma

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

1. Centeredness (o_π): *we have*

$$\pi_{\tau,\kappa}(0) = 0;$$

2. Coerciveness ($\mathfrak{O}_\pi$): *we have*

$$\lim_{m\to\infty} \|\pi_{\tau,\kappa}(m)\| = \infty; \text{ and}$$

3. Weak Monotonicity $\left(\widetilde{\mathfrak{M}}_\pi\right)$: *we have*

$$m_1 \leqslant m_2 \Longrightarrow \pi_{\tau,\kappa}(m_1) \leqslant \pi_{\tau,\kappa}(m_2),$$

where the inequality on the right-hand side is taken component-wise. [*Proof*.]

These properties are important because they tell us that the policy function behaves in ways that are both economically sensible and mathematically tractable. Without them, it would be difficult to draw meaningful conclusions about the structure of the state's production problem. With them in hand, we may proceed to the main event: understanding the structure of the data and the solutions they generate.

2.2 The Structure of the Data and the Solutions

$\mathcal{T}$ and $\mathcal{K}$ are sets of functions, which naturally raises the question of how to compare them. In the spirit of systemic theory as discussed in the introduction, we seek to understand whether they reflect any essential sameness—a question of *equivalence.* Two objects are equivalent if they are the same in some sense, and the study of equivalence relations is a powerful tool for understanding structure. Identical objects are always equivalent, but the converse is not true: equivalent objects are not always identical. For example, the number $1/2$ is both identical and equivalent to itself, but it is only equivalent to the number $2/4$. They are not identical, but they are equivalent because they represent the same quantity. In much the same way, two militarization technologies may yield different numerical outputs for the same mobilization plan, yet still represent structurally equivalent approaches to force production. In the present context, we are interested in understanding whether the militarization technologies and cost functions are equivalent in some sense and, if they are, just how strong this sense of equivalence is.

In the previous section, we spent considerable time discussing just how little has been assumed in defining $\mathcal{T}$ and $\mathcal{K}$. These are large and diverse sets, and it is not immediately obvious how much sameness we can expect to find among their elements. Rather than focusing on local properties or on the tedious details of equivalence relations for particular sub-classifications of functions, we take a more global perspective. We are interested in the *structure* of the data—literally the shape of the sets $\mathcal{T}$ and $\mathcal{K}$. This is a question of *homotopy,* which is the study of continuous transformations of spaces. The reader may be familiar with the old quip that a topologist cannot distinguish a coffee cup from a doughnut, because both of them possess a single hole. That single hole is the invariant that allows the topologist to say that the two objects are the same: they are not identical, but they are equivalent in the sense of homotopy. This sense of equivalence binds them together but distinguishes them from a ball, which has no holes, and from a pretzel, which has more than one.

We are interested in the homotopy of the sets $\mathcal{T}$ and $\mathcal{K}$ because it tells us how much sameness we can expect to find among their elements. If, for example, $\mathcal{T}$ is homotopy equivalent to two disconnected spaces, then we can expect to find two fundamentally different types of militarization technology: one that is structurally similar to the first space and one that is structurally similar to the second. If $\mathcal{K}$ is homotopy equivalent to a single connected space, then we can expect to find a single type of cost function that is structurally similar to all others. Moreover, the shape of these individual spaces may provide even deeper insights into the structure of the data.

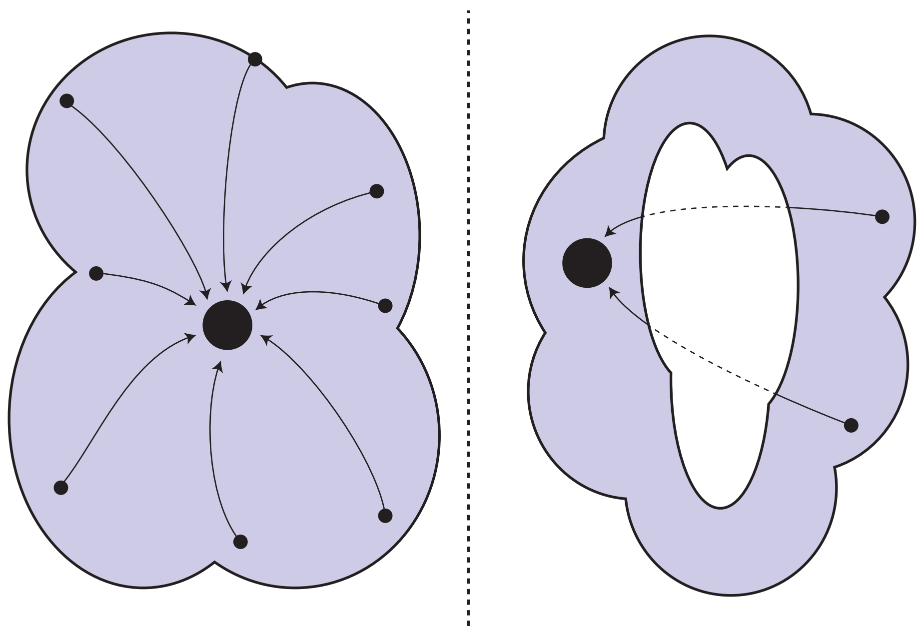

Figure 1: *Contractible (left) versus non-contractible (right) spaces.*

In terms of homotopy, the most extreme form of sameness is *contractibility.* Intuitively speaking, a space is contractible if it can be shrunk to a point without tearing or gluing. For example, consider the set $[-1, 1]$, a simple compact interval centered around zero. Now imagine that we pinched the interval from both ends and pulled the ends together until they met at the center point 0. This pinching and pulling is a continuous deformation, and it shrinks the entire interval to the single point 0; hence, this is a contractible space. In contrast, no rubber band could possibly deform the doughnut just mentioned into a single point, so the doughnut is not contractible. Figure 1 illustrates the difference between a contractible space (left) and a non-contractible space (right). At left, we see a blob-like shape with arrows pointing inward toward a central point. Any point in this shape can be continuously deformed to the central point, so the entire shape can be shrunk to that point. And, as it turns out, if a space can be contracted to such a point, it can be contracted to *any* point. Conversely, the shape at right is an annulus, a ring-like object with a hole in the middle. No matter how we try to deform this shape, the hole remains; there is no way to shrink the entire shape to a single point. Put differently, we can use a point to represent the left shape, but the smallest representation of the right shape must include the hole—*i.e.,* it must be a ring, not a point.

And this is the question we now ask about the function spaces $\mathcal{T}$ and $\mathcal{K}$: are they contractible? Can they be represented by a single point, or do they possess holes or other more complicated structure? It turns out that the classical assumptions given in Definitions 4 and 5 are strong enough to guarantee that both function spaces are contractible.

---

12 Lemma

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

Thus, the militarization technologies and cost functions are structurally equivalent in the strongest possible sense. Not doughnuts, nor pretzels, nor even simple rings; mere points. This obtains despite the fact that neither log-concavity nor exp-convexity are closed under addition or scalar multiplication, so neither $\mathcal{T}$ nor $\mathcal{K}$ is a vector space.[6]

We therefore have shown that the data are structurally simple to the point of triviality. As promised, this simplicity has deep implications for the structure $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$, the set of solutions to the family of problems those data generate. In the main result of this section, we learn that the set of states is also contractible.

---

[6] Since both of these classes relate closely to quasiconcavity and quasiconvexity, it is worth noting that analogous families defined by those properties are likewise contractible, even though their governing properties fail to preserve addition. The homotopies constructed in the *Proof* of Lemma 12 follow more general, non-linear paths that do not rely on vector structure.

13 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. [*Proof*.]

In the name of concreteness, we provide an explicit basepoint for the contractibility of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$, though this is not strictly necessary. If a space can be contracted to some point, then it can be contracted to any point; truly, any point is the center of the universe of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$. The basepoint given happens to be a particularly easy-to-derive solution to a particularly easy-to-solve instance of the state's production problem. It also happens to be aesthetically pleasing, performing the task of sending force demands to resource investments in a particularly simple way. It sends the zero force demand to the zero resource investment and divvies exponentially-increasing resource investments symmetrically across all commodities as the force demand increases. Highly unrealistic from an empirical perspective, it nevertheless serves as a worthy representative of the set of states. Happily, it does not matter which point we choose as the basepoint for the contractibility of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$; once we know the space is contractible, we know it can be shrunk to any point. We record this fact in the following corollary.

---

14 Corollary

$\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ can be strongly deformation retracted onto any point.

Thus, the set of all states is so structurally simple that it can be represented by a single point—and indeed, by *any* point.

Proposition 13 is the main structural result of this section, and it has deep implications for our understanding of states. But, its full significance becomes clear only when we consider the various ways in which states can be considered the same. In the next subsection, we work our way through several corollaries of Proposition 13, each of which reveals a different aspect of the structure of states. These corollaries are in ascending order of strength, each one building on the last. As such, they form a ladder of sameness, each rung revealing a deeper layer of equivalence among states.

2.3 The Ladder of Sameness

This subsection presents four corollaries of Proposition 13, each of which reveals a different aspect of the structure of states. Before considering the rungs in detail, we quickly summarize the ladder. Proposition 13 implies that the set of all states ignores variation in:

1. *Sorts:* qualitative distinctions among states are necessarily second-order to their sameness as producers of force (Corollary 15);
2. *Transformations:* any two states can be transformed into one another through a continuous deformation (Corollary 16);
3. *Histories:* any two paths between the same pair of states are continuously deformable into one another (Corollary 17); and
4. *Information:* any function defined on the set of states is homotopic to a constant function (Corollary 21).

Thus, when we say that all states are the same, we mean this in four increasingly strong senses. Let us now climb the ladder.

Qualitative distinctions among states are necessarily second-order to their sameness as producers of force. The first corollary is that the set of states is *connected*, which we state like so:

---

15 Corollary

$\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ cannot be written as the union of two disjoint non-empty open sets.

Visually, this means that the set of states is a single blob-like object without holes or disconnected pieces. Figure 2 illustrates the difference between a connected space (left) and a disconnected space (right). At left, we see a single blob-like shape; no matter how we try to slice it, we cannot separate it into two pieces without tearing it apart. At right, we see two separate blob-like shapes; no matter how we try to connect them, they remain separate. This is the difference between connected and disconnected spaces. In the event that $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ were disconnected, we would have two fundamentally different types of states, each of which we might call a sort of state. This would be a first-order distinction, a primary way of dividing the set of states in response to their basic task of producing force. But Corollary 15 tells us otherwise: there is only one sort of state. Any distinctions we draw must therefore be second-order. This result does not deny variation,

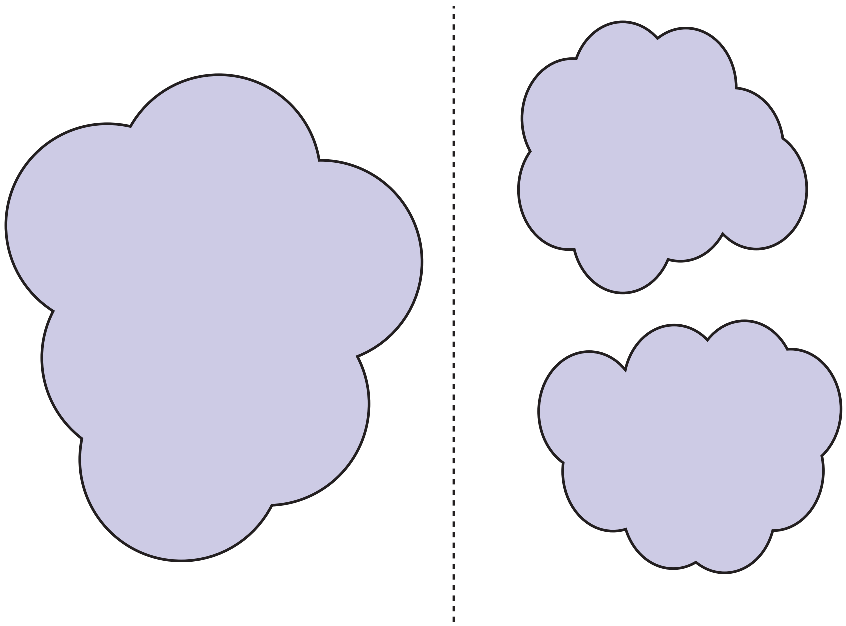

Figure 2: *Connected (left) versus disconnected (right) spaces.*

but it does demand that variation be interpreted through the lens of a shared structural core. This is the first layer of sameness that the contractibility of states reveals.

Of course, qualitative distinctions can and should be drawn among states. Democracies and autocracies, capitalist and socialist states, and so on—these are all distinctions that are both meaningful and important. But these distinctions are necessarily second-order to the sameness of states as producers of force. Put differently, if one has a classification structure P that divides states into categories, then P must be attached to a partition of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ into connected components. More formally 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 $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$. Figure 3 illustrates such a classification structure: four different subclasses of states are drawn within a single connected component. This is the only way to draw meaningful qualitative distinctions among states without contradicting Corollary 15. Thus, distinctions among states either need to live outside of this component of the ontology of states or they need to provide further information about how those distinctions manifest within the connected component.

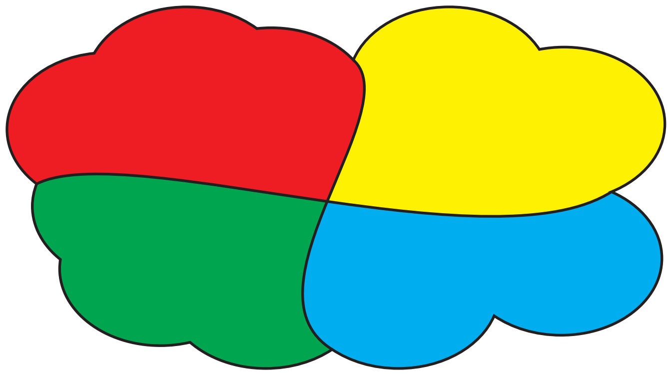

Figure 3: *Illustration of a classification structure P dividing states into categories within connected components.*

Remarkably, this partitioning logic helps us see the structuralist ideal as a limiting case of a more general framework. Classification structures, being partitions, can be refined or coarsened, and this allows us to put them into a hierarchy. One classification scheme might divide states into democracies and autocracies, while a finer scheme might divide them into presidential democracies, parliamentary democracies, military autocracies, and one-party autocracies. As we refine the classification structure, we approach the structuralist ideal: a classification structure that divides states into singleton sets. In this limiting case, each state is its own category, and the classification structure provides no further information beyond the identity of the states themselves. In the other direction, we can coarsen the classification structure until it divides states into a single category: $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ itself. This is the structuralist ideal in its purest form: all states are the same, and no distinctions are drawn among them. Functions sending classification schemes to outcomes of interest can then be assessed in terms like monotonicity or continuity with respect to refinements and coarsenings of the classification structure. Metatheoretically, this provides a way to situate structuralist theorizing within a broader framework of classification structures. And since $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is connected, any such classification structure must be second-order to the sameness of states as producers of force.

Any two states can be transformed into one another through a continuous deformation. The second corollary, deeply related to the first, is that the set of states is *path-connected*, which we state like so:

16 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 a somewhat stronger way of saying that the set of states is connected, but now it tells us the positive side of the story. Not only are there no fundamental distinctions among states, but any two states can be transformed into one another through a continuous deformation. Figure 4 illustrates the difference between a

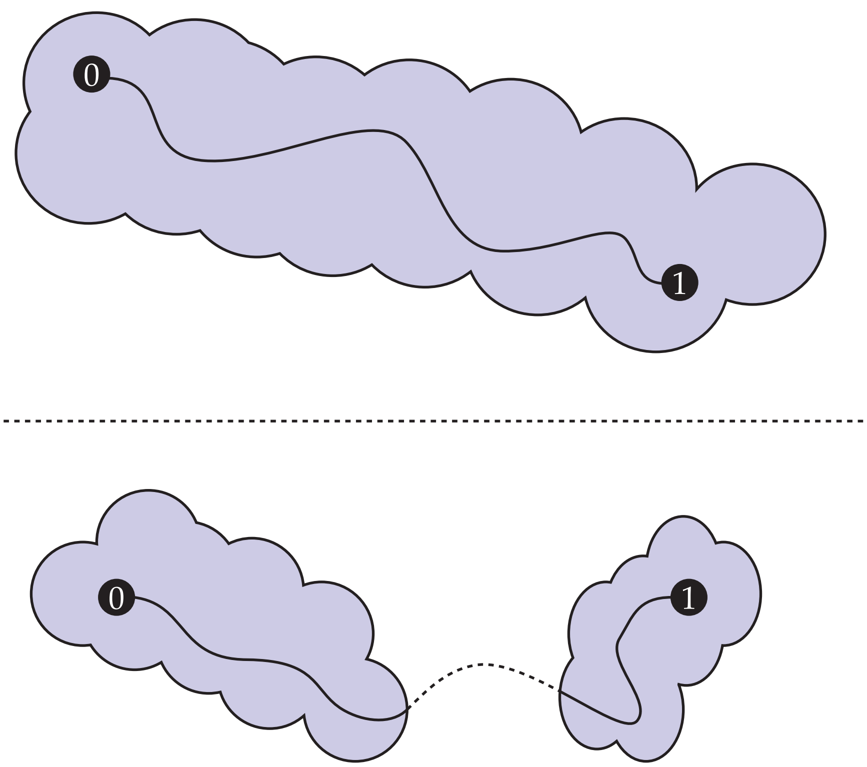

Figure 4: *Path-connected (top) versus non-path-connected (bottom) spaces.*

path-connected space (top) and a non-path-connected space (bottom). At top, we see a single blob-like shape; no matter which two points we pick, we can draw a continuous path between them without leaving the shape. Conversely, at bottom, we see two separate blob-like shapes; if we pick one point from each shape, there is no continuous path between them that remains within the shape.

This might seem like an arcane and abstract result if we think about how to turn Djibouti into Denmark, but it is much more powerful when we think about both the evolution of states over time and the emulation of states by other states. In terms of the former, it tells us that the transformation of one state into another is a continuous process; we might imagine the United States in 1776 as one state and the United States in 2025 as another and γ as the continuous transformation that takes us from one to the other as time passes. In terms of the latter, Corollary 16 tells us that states can learn from one another in a continuous way; we might imagine a state attempting to emulate the militarization style of another state, much as South American states did with Prussian militarization in the 19th century (Resende-Santos, 2007); again, this emulation can be seen as a continuous process. This is the second layer of sameness that the contractibility of states reveals: the sameness liberating the process of continual evolution and emulation.

Again, discontinuous transformations are possible and meaningful, but they are not woven into the fabric of the state system. If flashpoints and tipping points are to be found, they must be sought in the second-order distinctions that divide the set of states. Formally, a discontinuous transformation of states might be represented as a flashpoint 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 of states that is not continuous, characterized by a moment of fundamental change encoded in the parameter ζ. Naturally, more complicated structures can be imagined—multiple flashpoints, wrinkly evolution functions, and so on—but the basic point remains: one can envision discontinuous transformations, but they must be included in some second-order structure.

The space of states does not necessitate path dependency. The third corollary is that the set of states is *simply connected*, which we state like so:

---

17 Corollary

For all states $\pi_0, \pi_1 \in \mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ and all continuous paths $\gamma_\alpha, \gamma_\beta : [0,1] \to \mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ such that $\gamma_\alpha(0) = \gamma_\beta(0) = \pi_0$ 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 are all states the same in that all can be linked via a continuous path, but the paths themselves can be continuously deformed into one another. Figure 5

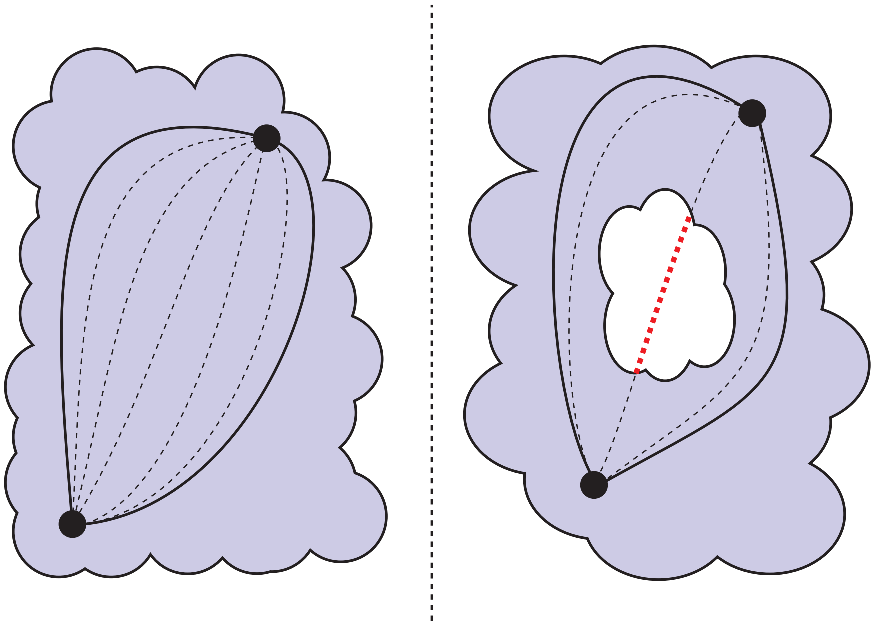

Figure 5: *Simply connected (left) versus non-simply connected (right) spaces.*

illustrates the difference between a simply connected space (left) and a non-simply connected space (right). At left, we see a single blob-like shape; no matter which two paths we pick between the same two points, we can continuously deform one path into the other without leaving the shape. Conversely, at right, we see a ring-like shape; if we pick one path that goes around the hole and another path that goes the other way, there is no continuous deformation between them that remains within the shape.

This might seem like an even more arcane and abstract result than the last, but it too has powerful implications when we think about both the evolution of states over time and the emulation of states by other states. In particular, it gives us a new perspective on path dependency. Path dependency is often understood as the idea that the history of a state matters for its present and future. One way of saying this might be:

> Two paths that link the same start and end points may nevertheless lead to different outcomes.

But Corollary 17 tells us that this is not quite right: if two paths link the same start and end points, then they are homotopic relative to their endpoints, and thus we think of them as essentially the same. Conversely, in the figure at right, paths "to the left" of the hole are fundamentally different from paths "to the right" of the hole; they cannot be continuously deformed into one another, and thus we think of them as essentially different. This is what path dependency looks like in a non-simply connected space.

We can think of this a little more formally using the language of differential forms. Suppose $\gamma : [0,1] \to \mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is a path—representing, for example, the temporal evolution of a state or the emulation of one state by another—and let ω be a piecewise-smooth 1-form on $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ that encodes infinitesimal changes in payoff or cost.[7] For two paths with common endpoints, define the difference in their line integrals as

$$\Delta(\gamma_0, \gamma_1; \omega) = \int_{\gamma_0} \omega - \int_{\gamma_1} \omega = \oint_{\gamma_0 * \overline{\gamma_1}} \omega,$$

where $\overline{\gamma_1}$ is γ_1 traversed in reverse and $*$ denotes concatenation. Thus, "path dependence" here means that there exists a closed loop with nonzero circulation. This leads us to the following lemma, which provides a useful test for path dependence.

---

18 Lemma

For a piecewise-smooth 1*-form* ω *on* $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$*, the following are equivalent:*

1. $\Delta(\gamma_0, \gamma_1; \omega) = 0$ *for all paths* γ_0, γ_1 *with common endpoints; and*
2. $\oint_\ell \omega = 0$ *for every piecewise-smooth loop* ℓ *in* $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$.

This is simply a matter of applying definitions, so we omit the proof. The lemma tells us that to check for path dependence, we need only check for nonzero circulation around loops. But now we can bring in Corollary 17, which tells us that all loops in $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ can be continuously contracted to a point. This has powerful implications when combined with *Poincaré's lemma*, which states that closed forms are exact on simply connected domains.[8] Poincaré's lemma gives us the following proposition:

---

[7] Explicitly, line integrals will be understood for C^1 paths with additivity under concatenation and sign reversal under reparameterization.

[8] By "exact," we mean that there exists a function Φ such that $\omega = d\Phi$. Such a function is called a *potential function* for the form ω. It provides a scalar field whose gradient is given by ω, and thus is the concept to which conservative vector fields correspond. A classic example is the gravitational field, which is the gradient of the gravitational potential.

19 Proposition

If $d\omega = 0$ on $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ and $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is simply connected, then there exists $\Phi : \mathcal{P}_{\mathcal{T}\times\mathcal{K}} \to \mathbb{R}$ such that

$$\omega = d\Phi \quad \Longrightarrow \quad \int_\gamma \omega = \Phi(\gamma(1)) - \Phi(\gamma(0)) \textit{ for all paths } \gamma.$$

Hence any closed field is path independent on $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$.

In words, this means that if the 1-form ω is closed (*i.e.,* $d\omega = 0$) and the space of states is simply connected, then the line integral of ω along any path depends only on the endpoints of that path. Thus, observed path-dependence can only arise if ω is not closed somewhere. This contrapositive form is useful in that it helps us to fully appreciate just how "outside" the model we must go to find path dependence.

---

20 Corollary

If

$$\exists \gamma_0, \gamma_1 \textit{ such that } \Delta(\gamma_0, \gamma_1; \omega) \neq 0,$$

then $d\omega \neq 0$ on some subset of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$—i.e., *$\omega$ is not closed.*

This is the third layer of sameness that the contractibility of states reveals: the sameness that confines path dependence to arise only from second-order distinctions among states. Put differently, if we observe path dependence in the evolution or emulation of states, it must reflect differences that are derivative of—rather than constitutive of—their shared function as producers of force. Formally, such dependence appears only when the differential field ω defined on the space of states is non-closed. To the extent that the statistics we compute to characterize states—be they measures of governance, economic structure, or social cohesion—can be represented as closed 1-forms on $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$, we can be sure that they do not entail path dependence.

Future work should focus on applying cohomology theory to classify the types of non-closed forms that might arise on $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$. Some initial details on the subject will be relegated to a footnote.[9]

[9] Formally, our simple-connectedness results corresponds to a trivial first de Rham cohomology group, $H^1_{\mathrm{dR}}(\mathcal{P}_{\mathcal{T}\times\mathcal{K}}) = 0$. Every closed 1-form on this space is therefore exact, and no topological degrees of freedom exist for storing historical information. The "memory" of the system, if any, must live in the field ω itself—in the particular political or institutional processes that make it non-closed. In this sense, topology constrains the geometry of possible histories: the state system cannot generate path dependence by shape alone.

$\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ **carries no information beyond the fact that its elements solve Problem** SPP (m, τ, κ)**.** The fourth and final corollary is a particular definition of contractibility, which we state like so:

---

21 Corollary

Let Y be any nonempty space, and let $f_0, f_1 : \mathcal{P}_{\mathcal{T}\times\mathcal{K}} \to Y$ be continuous maps. Then there exists a continuous map $F : \mathcal{P}_{\mathcal{T}\times\mathcal{K}} \times [0,1] \to Y$ such that $F(\cdot, 0) = f_0$ and $F(\cdot, 1) = f_1$.

In particular, any function $f : \mathcal{P}_{\mathcal{T}\times\mathcal{K}} \to Y$ is homotopic to a constant function.

This is the contractibility of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ seen *from the outside*: not as an internal deformation of the space, but as a statement about what any continuous observable can do with it. If any two maps f_0, f_1 can be deformed into one another, then no continuous observable can stably separate points of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$; any measured difference can be washed out by a homotopy. In the limit, every observable collapses to constancy.

Along these lines, we have the following corollary, which makes the point more explicit.

---

22 Corollary

For any nonempty Y, the homotopy set $[\mathcal{P}_{\mathcal{T}\times\mathcal{K}}, Y]$ contains a single element. Equivalently, every continuous invariant $I : \mathcal{P}_{\mathcal{T}\times\mathcal{K}} \to Y$ is, up to homotopy, a constant.

This places the discussion in informational terms. Suppose we define f to assign each state a number, say the soldiers used to field a force set to $m = 1$. Then Corollary 21 and Corollary 22 together imply that f is homotopic to a constant. Neither quantitative nor qualitative structure can be extracted from $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ alone beyond the fact that its elements solve SPP (m, τ, κ). This is the final layer of sameness revealed here: a domain that admits motion but no contrast, measurement but no differentiation, a genuinely blank slate awaiting further structure.

Assessment. We have built a logic of sameness. The global properties of $\mathcal{T}$ and $\mathcal{K}$—and the solutions they generate—yield a space of states that is connected, path-connected, simply connected, and finally null-homotopic as a source for observables. The result is both liberating and constraining. It is liberating in permitting continual evolution and emulation: any state can be deformed into any other without tearing the fabric. It is constraining in fixing the terms on which distinctions may be drawn: they cannot be sourced in the topology of the state space. If differences are to appear, they must do so as second-order features—through the fields we place upon the surface or through additional structures not yet introduced.

Interpretive scope and bridge. It bears emphasis that these claims are formal and conditional. They concern the topology of the model's state space, not the empirical world it seeks to illuminate. To say that $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is contractible is to say that, within this architecture, the space itself supplies no intrinsic coordinates of difference or memory. Empirical heterogeneity, historical inertia, and institutional specificity can—and will—enter, but only through additional geometric or dynamical structure. Accordingly, we now turn from topology to geometry. The next section asks what a *good model* of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ should look like: what properties it should preserve, what distortions it may justifiably introduce, and what it might buy us in return. There we restrict attention to a tractable class of technologies and costs, show that their solutions adequately represent the ambient space, and uncover a stronger property than contractibility—*convexity*—that equips the model with a richer, more usable shape.

3 Models of States

We just saw that the set of states $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is contractible, a proposition loaded with important implications for the structure of the bucket of states. As a topological property, contractibility conveys deep structural information about the set, from how connected it is to how much information it contains. However, it has little to offer about the *geometry* of the set, save for the fact that the geometry is, in some sense, unified and homogeneous. 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.

We would be well within our rights to begin modeling the set of states $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ as it stands. However, now is a good time to think about what additional structure we might be willing to tolerate in exchange for a clearer view of the geometry of the set. One natural step is to impose additional regularity conditions on the technologies and costs. To see why such conditions might be helpful, consider what Game 1 might look like if we incorporated the primitives we have defined so far.

---

23 Game

Two states, $i \in \{1, 2\}$, simultaneously choose resource investments $x_i \in X = \mathbb{R}^L_+$. Their *payoffs are given by von Neumann-Morgenstern expected utility functions:*

$$U_1(x_1, x_2) = \frac{\lambda\,(\tau_1(x_1))^{\alpha}}{\lambda\,(\tau_1(x_1))^{\alpha} + (\tau_2(x_2))^{\alpha}} \left(V - k\left(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\right)\right),$$

$$U_2(x_1, x_2) = \frac{(\tau_2(x_2))^{\alpha}}{\lambda\,(\tau_1(x_1))^{\alpha} + (\tau_2(x_2))^{\alpha}} \left(V - k\left(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\right)\right),$$

where:[10]

1. *λ, α, V, and k are as in Game 1;*
2. *$\tau_i \in \mathcal{T}$ is State i's technology; and*
3. *$\kappa_i \in \mathcal{K}$ is State i's cost function.*

In terms of the richness of the underlying political economy, Game 23 is an upgrade over Game 1, as it incorporates the technologies and costs we have defined so far, and it makes the choice variable something observable—-the resource investment vector x_i—rather than the abstract military capability m_i. The contest success function runs on inputs that have been sent to a force-like output through the technology, and the costs subtracted from the prize have been sent to a value-like output through the cost function. Of course, this richness comes at the expense of parsimony and clarity, as the game includes more moving parts and the choice variables are more complex.

It also comes at the expense of tractability, as the additional complexity makes it more difficult to analyze the game. We can at least provide general existence results for equilibria—for example:

---

24 Proposition

Game 23 has at least one pure-strategy Nash equilibrium. [*Proof*.]

But beyond this, it is difficult to say much more without imposing additional structure on what each τ and κ actually looks like. 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.

[10] The same disclaimer applies here as in Game 1: we set the contest outcomes to $\lambda/\lambda+1$ and $1/\lambda+1$ for the two States in case $\tau_1(x_1) = \tau_2(x_2) = 0$. The subsequent discontinuity in the utility functions makes the proof of Proposition 24 a little trickier than usual.

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.

Topologically speaking, Proposition 13 has already given an extreme answer: $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ is contractible, and thus representable by a single point. But while such a representation is formally faithful, it is geometrically and substantively vacuous. A single point cannot vary, and without variation there can be no explanation: no difference in how useful one resource is under a given technology or how dear another resource is in a given cost. The purpose of modeling is to make sense of such variation—to describe how differences in technology and cost structure shape behavior—and a single point can do none of that. To recover explanatory power, we must look for a representation that preserves not just connectedness but *shape* and variety.

This section addresses these questions by modeling the process of modeling itself. We seek a structured subset of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$ that is both tractable and topologically faithful. This suggests two conditions for a successful model:

1. *Tractability* concerns, particular in light of the intended use case, suggest that we need at least one degree of differentiability, else we will not be able to study games like Game 23 in the usual manner; and

2. *Adequacy* concerns, given the structure we have already established, suggest that we need the model to be homotopy equivalent to $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$, else we risk losing the essential topological features of the ambient space.

Put differently, we need to balance quantitative considerations about differentiability with qualitative considerations about topological structure.

To preview the results: we will indeed see that such a model exists. The particular model we will construct is one in which the technologies and costs take on especially simple functional forms, which we will call *tame* technologies and costs. These tame functions will be shown to adequately represent the general functions by being homotopy equivalent to them. Moreover, we will arrive at the tame states by using first-order information about regularized versions of the general functions, which will allow us to characterize the geometry of the tame states in traditional fashion. This means we will address both tractability and adequacy, arriving at a model that is both analyzable and faithful.

3.1 The Model

It will be far easier to introduce the modeling process in stages. For any writer with literary aspirations, there is a temptation to save the information meant for the climax for the end; however, some of the steps will not make sense if we wait until the end to introduce them. Let us therefore introduce the target model first, and then work backward to see how we might arrive at it.

---

25 Definition

We say that a technology $\tau \in \mathcal{T}$ is tame $(\mathfrak{T}_\tau)$ *if it takes the form*

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

where $A_\tau > 0$ is a "scale" parameter and the vector of "input elasticities,"

$$\beta \in \Delta_L := \left\{ b \in \mathbb{R}^L_{\geqslant 0} \;\middle|\; \sum_{\ell \in L} b_\ell = 1 \right\},$$

witnesses τ's tameness.

Similarly, we say that a cost $\kappa \in \mathcal{K}$ is tame $(\mathfrak{T}_\kappa)$ *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 the vector of "input prices,"

$$q \in \Delta_L := \left\{ p \in \mathbb{R}^L_{\geqslant 0} \;\middle|\; \sum_{\ell \in L} p_\ell = 1 \right\},$$

witnesses κ's tameness.

We denote the sets of all tame technologies and costs by $\mathcal{T}^{[\mathfrak{T}]}$ and $\mathcal{K}^{[\mathfrak{T}]}$, respectively.

The tame technologies and costs are both familiar and extraordinarily simple. They have many desirable properties, including differentiability (indeed, smoothness), concavity/convexity, and monotonicity. Their first and second derivatives are extraordinarily easy to compute, work with, and interpret. They are easy to bring to data and easy to use in models. In other words, they more than live up to their name as tame functions. It would be wonderful news indeed if we could show that these tame functions adequately represent the general functions we have defined so far.

The question of this section therefore becomes: do $\mathcal{T}^{[\mathfrak{T}]}$ and $\mathcal{K}^{[\mathfrak{T}]}$ adequately represent $\mathcal{T}$ and $\mathcal{K}$, respectively? For starters, straightforward checking shows that the tame functions are indeed members of their respective ambient spaces:

26 Lemma

The tame technologies and costs are elements of $\mathcal{T}$ and $\mathcal{K}$, respectively.

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

But there are many properties such subsets might have or lack, and we just spilled much ink about the important topological properties of the general functions. The one we cared the most about was contractability, which implied all sorts of interesting notions of sameness among the states. Just to keep everything above board, we should check whether the tame functions retain this property. This is straightforward to do:

27 Lemma

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

To see this, consider the straightforward homotopies

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

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

which send a given tame technology or cost to a particular "central" tame technology or cost. The central technology and cost have equal weights on all commodities and unit scale. For each commodity ℓ, the ℓth component of the homotopy takes the form

$$
\begin{aligned}
&(1-t)A_\tau\beta_\ell \log(1+x_\ell) + t\frac{1}{L}\log(1+x_\ell), \\
&\qquad\qquad\qquad\qquad (1-t)A_\kappa q_\ell x_\ell + t\frac{1}{L}x_\ell,
\end{aligned}
$$

so that the homotopies remain within the tame function spaces for all $t \in [0,1]$—their scale terms become

$$
\begin{aligned}
A_\tau(t) &:= (1-t)A_\tau + t, \\
A_\kappa(t) &:= (1-t)A_\kappa + t,
\end{aligned}
$$

and their weight terms become

$$
\begin{aligned}
\beta_\ell(t) &:= \frac{(1-t)A_\tau\beta_\ell + t\frac{1}{L}}{(1-t)A_\tau + t}, \\
q_\ell(t) &:= \frac{(1-t)A_\kappa q_\ell + t\frac{1}{L}}{(1-t)A_\kappa + t},
\end{aligned}
$$

It is not hard to show that these terms remain non-negative and sum to one for all $t \in [0,1]$. Thus, the tame functions retain contractibility.

So, we have demonstrated that the tame functions are subsets of the general functions and that they retain contractibility. We could be done, 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 $(\tau^{[\mathfrak{T}]}, \kappa^{[\mathfrak{T}]}) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$?

And now that we have set up this problem in a way that makes the modeling action clear, we may turn to answering it. The next step is to define a process that takes a general technology-cost pair and produces a differentiable approximation; we call this process *regularization*.

3.2 Regularization

Observe that neither Definition 4 nor Definition 5 imposed any differentiability assumptions; we only required 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. As a result, constructions like Problem SPP (m, τ, κ) and Game 23 cannot yet be treated by calculus. Before we can talk about optimal responses or marginal adjustments, we must pass through a stage of *regularization*: a systematic smoothing of rough functions into differentiable ones.

Regularization should be thought of as a gentle lens: it blurs the small irregularities of τ and κ while leaving their large-scale shape intact. Formally, we seek a continuous operator

$$\mathfrak{D} : \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T} \times \mathcal{K},$$

that replaces each pair (τ, κ) by a smoothed pair $(\mathfrak{D}_\tau(\tau), \mathfrak{D}_\kappa(\kappa))$ whose members are *smooth*—that is, infinitely differentiable on X.[11] Several features of this operator are essential: it must be continuous as a map, preserve the structural properties of $\mathcal{T}$ and $\mathcal{K}$ (such as monotonicity, convexity, and concavity), and—most importantly—fix the "tame" functions that already behave well:

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

Smooth regularization is not an exotic device; 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. Hence the existence of a continuous smoothing map is not surprising; it is a canonical way to make explicit what this density already implies.[12]

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. The following proposition records the analytic fact that such an operator exists and behaves as required.

---

28 Proposition

There exists a continuous regularization operator

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

[11] We write $\mathcal{T}^{[\infty]} := \mathcal{T} \cap C^\infty(X, M)$ and $\mathcal{K}^{[\infty]} := \mathcal{K} \cap C^\infty(X, \mathbb{R}_{\geq 0})$.

[12] There are many constructions that achieve it: convolution with a smooth kernel, Moreau envelopes, and spline regularization are all standard. Among them, 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.

such that

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

in other words, $\mathfrak{D}$ fixes the tame functions. [*Proof*.]

The proof is constructive, providing an explicit formula for $\mathfrak{D}$ based on causal convolution with a smooth kernel.

The first step of our modeling process is now complete. We have defined a regularization operator $\mathfrak{D}$ that smooths arbitrary technologies and costs into differentiable ones while leaving tame functions unchanged. That smoothing process takes the form

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

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.

3.3 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. Let us normalize this gradient to obtain 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.[13] 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, we define the *tame representation* of τ by

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

Because τ is increasing, $\nabla\tau(0) \in \mathbb{R}^L_{\geqslant 0}$, so $\beta \in \Delta_L$ and $A_\tau > 0$, ensuring that $\mathfrak{T}(\tau)$ is indeed a well-defined tame technology.

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

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

and let

$$\mathfrak{T}(\kappa)(x) := A_\kappa \sum_{\ell\in L} q_\ell x_\ell.$$

Here, 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

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

[13] The reader might worry that the denominator is zero, but this is precluded by ray surjectivity of τ; if there exists a direction such that τ strictly increases along the direction, then there exists at least one partial derivative that is positive.

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. Remarkably, the tamification of technologies and costs renders non-identical objects equivalent. Put differently, $\mathfrak{T}$ is not a one-to-one function from $\mathcal{T}^{[\infty]}$ to $\mathcal{T}^{[\mathfrak{T}]}$ and from $\mathcal{K}^{[\infty]}$ to $\mathcal{K}^{[\mathfrak{T}]}$, but rather a many-to-one function. Consider two distinct technologies τ_0 and τ_1 in $\mathcal{T}^{[\infty]}$. If it happens to be that $\nabla\tau_0(0) = \nabla\tau_1(0)$, then $\mathfrak{T}(\tau_0) = \mathfrak{T}(\tau_1)$, even though the two functions are distinct. This is a powerful result, as it 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.[14]

Crucially, tamification leaves tame functions unchanged, as we record in the following lemma:

---

29 Lemma

For all $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, *we have* $\mathfrak{T}(\tau, \kappa) = (\tau, \kappa)$. [*Proof*.]

This is a simple consequence of the definition of tamification, but it is an important one. It states that the tame functions are *fixed points* of the tamification process. This is both a powerful and philosophically appealing result. 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. 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.

This is the second step of our modeling process. We have defined a tamification operator $\mathfrak{T}$ that extracts from each regularized technology-cost pair a canonical tame representative. That representative captures the first-order structure of the original pair at the origin and leaves the tame functions unchanged. The tamification process takes the form

$$
\begin{aligned}
&t \longmapsto (1 + \mathfrak{D}_\tau(\tau))^{1-t} \cdot (1 + (\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau))^t - 1, \\
&t \longmapsto \log\left((1-t)\exp(\mathfrak{D}_\kappa(\kappa)) + t\exp((\mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)(\kappa))\right),
\end{aligned}
$$

for $t \in [1/2, 1]$, 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 $t \in [0, 1]$, the tame functions are fixed points of the entire deformation.

[14] One could imagine quotienting out 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*, which seem a promising avenue for enriching the present construction with more data.

3.4 Adequacy of the Tame Representation

We have now introduced a full theory of representation through regularization and tamification. Regularization smooths arbitrary technologies and costs into differentiable ones, while tamification extracts from those smooth functions a canonical tame representative. To show that this representation is *adequate*—that the tame functions faithfully reflect the topological structure of the general ones—we combine the two steps 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

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

and for the costs by

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

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

---

30 Proposition
$\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ is a strong deformation retract of $\mathcal{T} \times \mathcal{K}$. [*Proof*.]

The existence of this retraction completes the circle: every general technology-cost pair can be continuously deformed into its canonical tame representative without leaving the ambient space, and every tame pair remains invariant along the way. The adequacy of the tame representation is therefore not merely heuristic but topological.

In which "adequacy" is finally explained. I have been remiss in not defining what I mean by "adequacy" until now; it seemed easier to provide the definition after the construction. Adequacy, in this context, means that the tame representation preserves the essential topological structure of the ambient space. A strong deformation retraction provides an exact and conceptually disciplined sense in which this is true. It ensures that the tame space $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ is not only contained within the ambient space $\mathcal{T} \times \mathcal{K}$ but is, up to homotopy, equivalent to it. Every general technology-cost pair can be continuously deformed into its tame counterpart through a path that remains entirely within the original functional space, and every tame pair remains fixed throughout the deformation. This means that no information about the global topological structure of $\mathcal{T} \times \mathcal{K}$ is lost in passing to the tame subspace: all homotopy invariants—connectedness, contractibility, and higher homotopy groups—are preserved. In geometric terms, one might think of the ambient space as a possibly irregular cloud enclosing a smooth inner region. The deformation retraction defines a continuous flow from the cloud to its core, collapsing extraneous irregularities while leaving the essential shape untouched. Adequacy, in this sense, is neither mere approximation nor abstraction: it is the existence of a continuous correspondence that preserves topological identity while improving analytic tractability. The tame functions do not merely approximate the general ones; they constitute a canonical, homotopically faithful image of them, sufficient for any analysis that depends on global qualitative structure rather than local idiosyncrasy.

Having now established the adequacy of the tame representation in the parameter space, we may turn to the states themselves.

3.5 The Geometry of Tame States

We just saw that the tame technologies and costs adequately represent their ambient spaces. What about the states they generate? Recall from Section 1 that each technology-cost pair $(\tau, \kappa) \in \mathcal{T} \times \mathcal{K}$ induces a state $\pi_{\tau,\kappa} : M \to X$ by solving the production problem $\mathsf{SPP}\,(m, \tau, \kappa)$ for each output level $m \in M$. The collection of all such states forms the state space $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$. We may now consider the subset of states generated by tame technologies and costs:

---

31 Definition

We define the set of tame states *as*

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

Because the solutions to the production problem $\mathsf{SPP}\,(m, \tau, \kappa)$ vary continuously with the technologies and costs, the tame states are themselves a strong deformation retract of $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$; they are just as adequate at representing their ambient space as are the tame technologies and costs.[15]

As promised, we proceed in traditional style. Now that we have smooth functions, we may characterize the solutions to the production problem SPP (m, τ, κ) by relating marginal products to marginal costs.

---

32 Lemma

For all $(m, \tau, \kappa) \in M \times \mathcal{T}^{[\infty]} \times \mathcal{K}^{[\infty]}$, $\pi_{\tau,\kappa}(m)$ solves SPP (m, τ, κ) *if and only if there exists a Lagrangian multiplier $\lambda_{m,\tau,\kappa} \in \mathbb{R}_{>0}$ and a vector of Karush-Kuhn-Tucker multipliers $\eta_{m,\tau,\kappa} \in \mathbb{R}^L_{\geqslant 0}$ such that:*

$$\begin{aligned} \frac{\partial \kappa}{\partial x_\ell}\left(\pi_{\tau,\kappa}(m)\right) - \lambda_{m,\tau,\kappa}\frac{\partial \tau}{\partial x_\ell}\left(\pi_{\tau,\kappa}(m)\right) - \eta_{m,\tau,\kappa,\ell} &= 0 \quad \textit{for all} \quad \ell \in L, \\ m - \tau\left(\pi_{\tau,\kappa}(m)\right) &= 0, \\ \eta_{m,\tau,\kappa,\ell} \times \pi_{\tau,\kappa,\ell}(m) &= 0 \quad \textit{for all} \quad \ell \in L, \end{aligned} \tag{FOC (m,τ,κ)}$$

In case $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, FOC (m, τ, κ) *takes the simpler form*

$$\begin{aligned} A_\kappa q_\ell - \lambda_{m,\tau,\kappa}\frac{A_\tau \beta_\ell}{1 + \pi_{\tau,\kappa,\ell}(m)} - \eta_{m,\tau,\kappa,\ell} &= 0 \quad \textit{for all} \quad \ell \in L, \\ m - A_\tau \sum_{\ell\in L} \beta_\ell \log\left(1 + \pi_{\tau,\kappa,\ell}(m)\right) &= 0, \\ \eta_{m,\tau,\kappa,\ell} \times \pi_{\tau,\kappa,\ell}(m) &= 0 \quad \textit{for all} \quad \ell \in L, \end{aligned}$$

where A_τ, β, A_κ, and q are the scale and weight parameters witnessing tameness of τ and κ respectively.

We state the lemma without proof; the reader is referred to any introductory optimization textbook (*e.g.*, Sundaram, 1996) for further details.

The first-order conditions FOC (m, τ, κ) are a function, the zeroes of which are precisely the solutions to the production problem SPP (m, τ, κ). 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 we have imposed a shape condition on the cost functions. Essentially, the quantitative functions

[15] In the proof of Proposition 13, we defined a canonical lift that sent arbitrary $\pi_{\tau,\kappa}$ to a particular (τ, κ) combination, and this lift is continuous in π. The lift is also of use here; to retract arbitrary π onto tame π, we first lift to a canonical technology-cost pair, then work through the homotopy as constructed in the previous subsection.

have been defined to unlock the door for this sort of analysis; they are the functions for which first-order analysis is straightforward and appropriate. They are a particularly appealing choice for modeling the state in this style, as their instantiations of FOC (m, τ, κ) are particularly simple. Owing to this simplicity, it is straightforward matter to arrive at the climax of Section 3.

---

33 Proposition

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

We have therefore shown that the set of tame states is a convex set, which is 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?

1. For starters, 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. If nothing else, this means that the tame states are a more informative model of the general states.

2. But more than this, convexity is an extraordinarily useful property for a set to have. It means that the set is "nice" in a way that contractibility does not. For example, 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, and it is easy to see that the 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/quasiconvexity, which 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.

3. Convexity is indeed a geometric property, rather than a topological one; it refers to shape, not merely connectivity. For example, consider a disk

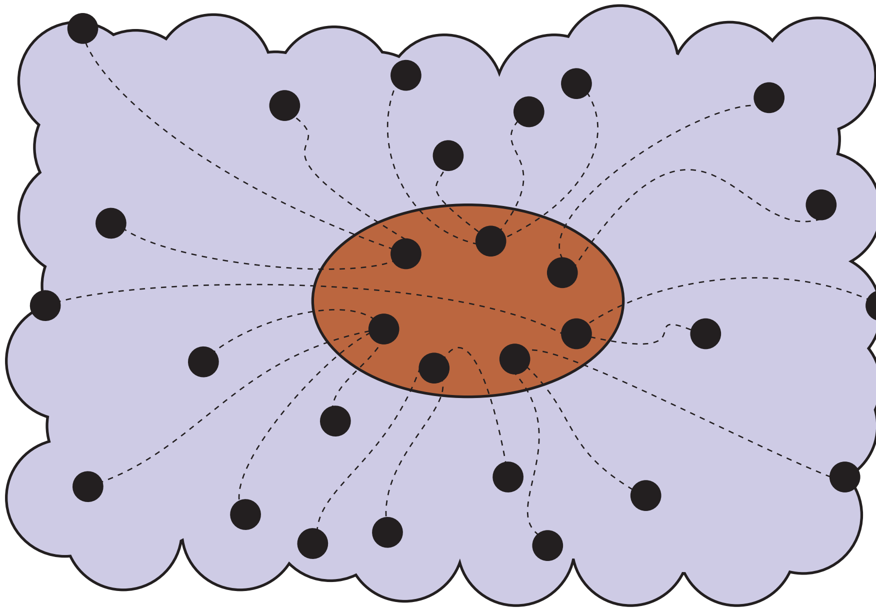

Figure 6: *Illustration of convexity representation.*

embedded in $\mathbb{R}^2$. The disk is contractible, as it can be continuously shrunk to a point. But it is also convex, as any two points in the disk can be connected by a straight line that lies entirely within the disk. Now consider the fact that the boundary of the circle maintains curvature with constant sign: it always bends in the same direction. This is a property of the shape of the disk, not the disk's connectivity. Thus, we are able to learn more about the nitty-gritty details of the set from the convexity of the tame states than we could from the contractibility of the general states.

4. Finally, 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. Far from a convenient mathematical trick, they provide deep insight into both the structure of states and the act of representation itself.

We are left with a model that is smaller than the original, but richer in structure.

This is the paradox of tamification: by simplifying, we reveal. The tame states are not the full story, but they are the story told clearly, with lines drawn straight and connections left intact.

3.6 Closing the Book

We opened this section by transferring the basic logic of Game 1 to its "enriched" counterpart, Game 23. Having found ourselves at an impasse, we begged for tractability by introducing tame functions. After thinking hard about the adequacy of that representation, we arrived at a convex set of tame states. We may now return to Game 23 and test whether these tame states yield a more explicit characterization of 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. Proposition 34 shows that each player's best response takes on a clear analytical structure: the equilibrium allocation is determined by a single scalar parameter that balances marginal productivity and marginal cost across all inputs. The entire strategic problem collapses to a one–dimensional fixed point, and the equilibrium emerges in a "water–filling" shape: continuous, ordered, and interpretable.

---

34 Proposition

Suppose the game in Game 23 is tame*, that is, for each player $i \in \{1,2\}$ we have*

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

with $A_{\tau,i}, A_{\kappa,i} > 0$, $\beta_i, q_i \in \Delta_L$ and all entries strictly positive. Let p_i denote the contest success probability and set

$$\begin{aligned} W(x_1,x_2) &= V - k\big(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\big),\\ p_1(x_1,x_2) &= \frac{\lambda\, \tau_1(x_1)^\alpha}{\lambda\, \tau_1(x_1)^\alpha + \tau_2(x_2)^\alpha}, \qquad p_2(x_1,x_2) = 1 - p_1(x_1,x_2). \end{aligned}$$

Then any Nash equilibrium (x_1^*, x_2^*) with strictly positive allocations on a (possibly player-specific) active set has the water–filling form*: for each player i there exists a scalar $c_i > 0$ such that*

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

and the scalar c_i solves the one-dimensional fixed point

$$c_i = \frac{(1-p_i)\,\alpha\,A_{\tau,i}}{k\,A_{\kappa,i}} \frac{W(x_1^*, x_2^*)}{e^{\kappa_i(x_i^*)}\,\tau_i(x_i^*)}. \tag{2}$$

In particular, if player i's equilibrium has all inputs active, then

$$\tau_i(x_i^*) = A_{\tau,i}\Big(\log c_i + \sum_{\ell\in L} \beta_{i,\ell} \log \frac{\beta_{i,\ell}}{q_{i,\ell}}\Big), \qquad \kappa_i(x_i^*) = A_{\kappa,i}\,(c_i - 1), \tag{3}$$

so that (2) becomes a scalar equation in c_i.

The proof follows directly from the first-order conditions in Lemma 32 together with the structure of the contest success functions. 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 thus reduces to a scalar equilibrium condition, a geometric equilibrium of pressures.

Finally, in the symmetric case, the simplification is complete.

---

35 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

$$x_\ell^* = \max\Big\{0, c\,\frac{\beta_\ell}{q_\ell} - 1\Big\} \qquad \textit{for all } \ell \in L, \tag{4}$$

where $c > 0$ solves

$$c = \frac{\alpha\,A_\tau}{2\,k\,A_\kappa} \frac{V - 2k\big(e^{A_\kappa(c-1)} - 1\big)}{e^{A_\kappa(c-1)}\,A_\tau\big(\log c + \sum_{\ell\in L} \beta_\ell \log \frac{\beta_\ell}{q_\ell}\big)}. \tag{5}$$

If $c \geq \max_{\ell\in L} \frac{q_\ell}{\beta_\ell}$ then all inputs are active and (4) 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. It is the strategic shadow price of efficiency, the market-clearing fulcrum against which every actor measures marginal gain, the residue of competition distilled to its purest form. The equilibrium surface is perfectly smooth, its contours given by the ratios ${}^{\beta_\ell}\!/_{q_\ell}$: a literal map of efficiency. The tame world does not merely approximate strategic interaction—it makes it visible.

If the aim of Section 3 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.

4 Conclusion

Let us recall our formal goals.

---

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.*

Call such a map the opportunity cost of militarization.

As promised, we have constructed a map, $\pi_{\tau,\kappa} : M \to X$, that specifies which resources $x \in X$ the state mobilizes given a desired level of force $m \in M$, a militarization technology $\tau \in \mathcal{T}$, and a cost function $\kappa \in \mathcal{K}$. This map is not arbitrary: it is the solution to the optimization problem SPP (m, τ, κ). We have, in effect, taken seriously a familiar metaphor—the state as a kind of firm—and given it mathematical substance. The analogy is not merely rhetorical. It reflects a lineage of thought in which the state's most fundamental activity is the organized production of coercive capacity. As Tilly put it, states make war and war makes states. If the state does many things, this is among the first.

And yet, what we have done here is austere. We have not modeled diplomacy, legitimacy, or social order. We have studied a single function. We have treated the state as a mapping from desired power to resource allocation, stripped of history, culture, and contingency. This is what a state can be in isolation: an operator defined by the logic of its production. It is the thin silhouette of the state when all that remains are its necessary conditions for force. That silhouette is enough to tell us something fundamental. It shows us that even when reduced to its most skeletal form, the state's structure obeys a logic both simple and revealing.

Because the state's production problem is continuous, its set of solutions inherits the structure of the functions that define it. We showed that $\mathcal{P}_{\mathcal{T}\times\mathcal{K}}$—the set of all such "states"—is contractible under very general assumptions. This means that, at a deep structural level, all states belong to a single connected manifold. There are no categorical fractures among them. Diversity, in this sense, is variation within a unified topology. To the degree that there are "types" of states, they occupy continuous regions within the same space rather than distinct kinds. This formalizes what structural realists have long argued: that the variety of states reflects the play of structure, not the eruption of essence. As in Waltz, the system constrains before it differentiates.

The argument could have ended there—with a statement about the state's topological unity. But topology tells us only that something holds together, not what shape it takes. We therefore passed from the topological to the geometric, introducing tame functions to represent technologies and costs. This was not a turn to realism, but to adequacy. The tame classes preserve the structure of the general classes up to deformation: they are simpler, but not simpler than the truth. They let us see the shape of the state's possibility space without inventing properties that were not already implicit in the general formulation. In Quine's sense, this was a maneuver of economy rather than ontology. We have not multiplied entities, only clarified our language.

The reward was geometric. Where the general set of states was contractible, the tame set was convex. Convexity is not merely a mathematical convenience—it is a statement about structure. It means that mixtures of states are still states, that intermediate configurations are coherent. It grants a linear geometry to the space of statehood. In that geometry, we can speak of interpolation and equilibrium; we can connect two points by a straight path rather than by a contorted one. The tame representation transforms the state system from a loose topological fabric into a smooth, navigable surface.

Once both technology and cost were tame, the strategic problem of contestation—the game of mobilization—became solvable in closed form. The equilibrium took on a shape: a "water-filling" pattern, smooth and ordered, governed by a single scalar parameter balancing productivity and cost. The state's strategic behavior, once tangled in many dimensions, reduced to an intelligible surface. This was not a trick of algebra, but the visible reward of tamification: a clarity earned by structure rather than assumed by fiat. It shows that when we discipline our representations, the phenomena we study sometimes discipline themselves.

At this point, we can see that the project has always been double. Formally, it has been about the state's production of force; philosophically, it has been about what it means to model. We have constructed not a picture of the state, but a lens for seeing what makes pictures possible. The "class of all states" is not a metaphysical claim about what exists, but a linguistic construction that lets us talk about what can be represented. Our results—contractibility, convexity, equilibrium—are properties of that representation, not of the world itself. They are, in Quinean spirit, the residuum of what must be true if our talk of states is to cohere. They tell us what holds once we have disciplined our speech about the state.

And yet, even in this spare, functional portrait, we glimpse something of the real state. Force production is not the whole of statehood, but it is never far from its core. It is the part of the state that can be most cleanly formalized because it is the part that must, in the end, work. To model the state through this function is to study the minimal conditions under which coercive capacity can exist at all. In that sense, we have isolated one strand of a much larger braid. The state does not merely produce force; it also allocates attention, defines boundaries, and sustains recognition. Those processes depend on relations—with other states, with societies, with environments. We have bracketed those relations here not because they are unimportant, but because understanding them requires first knowing what the isolated state looks like. A relation presupposes relata; we have studied one of them.

But the next step, inevitably, is to reintroduce relation. If the state can be defined by the way it converts resources into force, it can also be known by the ways it maps into, and is mapped by, others. Its identity lies not only in what it does alone, but in how it acts upon and is acted upon—how it transforms, and is transformed by, the networks of which it is a part. There is a sense in which to know a state is to know the family of mappings that express its relations. The future task, then, is to reconstruct the state not as a solitary object but as a structure of correspondences: a system that is determined, not by its contents, but by its position in a web of transformations. If the present paper has shown what a state can be in isolation, the next must show what a state becomes in relation.

For now, it is enough to recognize what has been achieved. We have taken the simplest and most severe abstraction of the state—the act of force production—and treated it as a mathematical object. We have found in that austerity a topology of unity, a geometry of convexity, and a structure of equilibrium. We have shown that even when pared down to a single function, the logic of the state yields an intelligible form. That is a modest claim, but it carries an unexpected grace. To study the state in this way is not to reify it, but to remind ourselves that beneath the flux of politics lies a disciplined structure of reasoning—a space where things can, at last, be seen clearly.

A Proofs

This section contains the proofs of all the results in the main text. However, since many of the technical details are not relevant to the argument presented in the main text, the section will have to be broken up into several parts. In particular, we need to study the structure of the spaces of militarization technologies and cost functions; happily, these are similar enterprises, and many of the results we prove for one space will carry over to the other.

A.1 The Motiving Game

Here we derive the unique Nash equilibrium of Game 1.

---

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:

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

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

Proof. We first show that there cannot exist a Nash equilibrium where $m_1 = 0 = m_2$, which involves specifying contest probabilities for this case. We simply set

$$p_1(0,0) = \frac{\lambda}{\lambda+1} \quad \text{and} \quad p_2(0,0) = \frac{1}{\lambda+1}, \tag{6}$$

which reflect the same relative effectiveness ratio as in the positive-effort case. The expected utilities for both players are given by:

$$\begin{aligned} U_1(0,0) &= \frac{\lambda}{\lambda+1} \times V, \\ U_2(0,0) &= \frac{1}{\lambda+1} \times V. \end{aligned} \tag{7}$$

Without loss of generality, suppose Player 1 deviated to $m_1 = \varepsilon$ for some $\varepsilon > 0$. Then Player 1's expected utility is given by:

$$\begin{aligned} U_1(\varepsilon,0) &= \frac{\lambda\varepsilon^{\alpha}}{\lambda\varepsilon^{\alpha} + 0^{\alpha}} \times (V - k(\varepsilon + 0)), \\ &= \frac{\lambda\varepsilon^{\alpha}}{\lambda\varepsilon^{\alpha}} \times (V - k\varepsilon), \\ &= V - k\varepsilon. \end{aligned} \tag{8}$$

Note that

$$\begin{aligned} U_1(\varepsilon,0) - U_1(0,0) &= V - k\varepsilon - \frac{\lambda}{\lambda+1}V, \\ &= \left(1 - \frac{\lambda}{\lambda+1}\right)V - k\varepsilon, \\ &= \frac{V}{\lambda+1} - k\varepsilon. \end{aligned} \tag{9}$$

Therefore, we may choose $\varepsilon < \frac{V}{k(\lambda+1)}$, and this ensures that $U_1(\varepsilon,0) > U_1(0,0)$. A similar argument shows that Player 2 will not choose an effort level of 0 against 0. We conclude that there cannot exist a Nash equilibrium where $m_1 = 0 = m_2$.

Now we show that there cannot exist a Nash equilibrium where $m_1 = 0$ and $m_2 > 0$. The expected utilities for both players are given by:

$$\begin{aligned} U_1(0,m_2) &= \frac{\lambda 0^{\alpha}}{\lambda 0^{\alpha} + m_2^{\alpha}} \times (V - k(0 + m_2)), \\ &= 0, \\ U_2(0,m_2) &= \frac{m_2^{\alpha}}{\lambda 0^{\alpha} + m_2^{\alpha}} \times (V - k(0 + m_2)), \\ &= V - km_2. \end{aligned} \tag{10}$$

Holding $m_1 = 0$ fixed, consider a deviation by Player 2 to ζm_2 for $\zeta \in (0, 1)$. Then Player 2's expected utility is given by:

$$\begin{aligned} U_2(0, \zeta m_2) &= \frac{(\zeta m_2)^\alpha}{\lambda 0^\alpha + (\zeta m_2)^\alpha} \times (V - k(0 + \zeta m_2)), \\ &= \frac{(\zeta m_2)^\alpha}{(\zeta m_2)^\alpha} \times (V - k\zeta m_2), \\ &= V - k\zeta m_2. \end{aligned} \tag{11}$$

This is a strict improvement over $U_2(0, m_2)$, so we conclude that there cannot exist a Nash equilibrium where $m_1 = 0$ and $m_2 > 0$. A similar argument shows that there cannot exist a Nash equilibrium where $m_1 > 0$ and $m_2 = 0$. We therefore study only the interior case where $m_1, m_2 > 0$.

Since any pure-strategy Nash equilibrium, if it exists, must be interior, we proceed by studying the first-order conditions of the expected utility functions. These conditions are

$$\frac{\alpha m_2^\alpha V}{\lambda m_1^{1+\alpha} + m_2^\alpha \left((1+\alpha)m_1 + \alpha m_2\right)} = k, \tag{12}$$

$$\frac{\lambda \alpha m_1^\alpha V}{m_2^{1+\alpha} + \lambda m_1^\alpha \left(\alpha m_1 + (1+\alpha) m_2\right)} = k. \tag{13}$$

Define $\rho = \frac{m_1}{m_2}$, so that $m_1 = \rho m_2$. Substituting this into (12) and (13) gives:

$$\frac{\alpha V}{m_2 \left(\alpha + (1+\alpha)\rho + \lambda \rho^{1+\alpha}\right)} = k, \tag{14}$$

$$\frac{\alpha \lambda \left(\rho m_2\right)^\alpha V}{m_2 \left(m_2^\alpha + \lambda \left(\rho m_2\right)^\alpha \left(1 + \alpha(1+\rho)\right)\right)} = k. \tag{15}$$

Equating the left-hand sides of (14) and (15) gives:

$$\frac{\alpha V}{m_2 \left(\alpha + (1+\alpha)\rho + \lambda \rho^{1+\alpha}\right)} = \frac{\alpha \lambda \left(\rho m_2\right)^\alpha V}{m_2 \left(m_2^\alpha + \lambda \left(\rho m_2\right)^\alpha \left(1 + \alpha(1+\rho)\right)\right)}. \tag{16}$$

Cross-multiplying and simplifying gives

$$1 + \frac{1}{\lambda \rho^\alpha} = \rho + \lambda \rho^{1+\alpha}. \tag{17}$$

Let $x \coloneqq \lambda\rho^{1+\alpha} > 0$. Since $\lambda\rho^{\alpha} = {}^{x}/_{\rho}$, (17) becomes

$$1 + \frac{\rho}{x} = \rho + x. \tag{18}$$

Multiplying by x and rearranging yields

$$x^2 + (\rho - 1)x - \rho = 0, \tag{19}$$

whose roots are $x \in \{-\rho, 1\}$. Because $x > 0$, we must have $x = 1$, so

$$\lambda\rho^{1+\alpha} = 1 \qquad \Longrightarrow \qquad \tilde{\rho} = \lambda^{-\frac{1}{1+\alpha}}. \tag{20}$$

Substituting $\tilde{\rho}$ into (14) and using $\lambda\tilde{\rho}^{1+\alpha} = 1$ gives

$$\alpha + (1+\alpha)\tilde{\rho} + \lambda\tilde{\rho}^{1+\alpha} = \alpha + (1+\alpha)\tilde{\rho} + 1 = (1+\alpha)\big(1+\tilde{\rho}\big). \tag{21}$$

Hence the equilibrium efforts are

$$m_2^* = \frac{\alpha V}{k(1+\alpha)\big(1+\tilde{\rho}\big)} = \frac{\alpha}{1+\alpha} \cdot \frac{V}{k} \cdot \frac{1}{1+\lambda^{-\frac{1}{1+\alpha}}}, \tag{22}$$

$$m_1^* = \tilde{\rho}\, m_2^* = \frac{\alpha}{1+\alpha} \cdot \frac{V}{k} \cdot \frac{\lambda^{-\frac{1}{1+\alpha}}}{1+\lambda^{-\frac{1}{1+\alpha}}}. \tag{23}$$

In particular, the equilibrium ratio is uniquely pinned down by

$$\frac{m_1^*}{m_2^*} = \tilde{\rho} = \lambda^{-\frac{1}{1+\alpha}}, \tag{24}$$

and the scale is uniquely determined by the level first-order condition above, so there is a unique pair $\left(m_1^*, m_2^*\right)$ satisfying the first-order conditions.

We must verify that the solution to the first-order conditions corresponds to a maximum. First, it helps to show the following. Using the first-order condition

$$k = \frac{\alpha V}{m_2\left(\alpha + (1+\alpha)\rho + \lambda\rho^{1+\alpha}\right)}, \tag{25}$$

we obtain

$$m_2 = \frac{\alpha V}{k\left(\alpha + (1+\alpha)\rho + \lambda\rho^{1+\alpha}\right)}, \qquad m_1 = \rho m_2. \tag{26}$$

Hence

$$\frac{k(m_1 + m_2)}{V} = \frac{\alpha(1 + \rho)}{\alpha + (1 + \alpha)\rho + \lambda\rho^{1+\alpha}}. \tag{27}$$

From the ratio identity implied by the first-order conditions,

$$1 + \frac{1}{\lambda\rho^{\alpha}} = \rho + \lambda\rho^{1+\alpha}, \tag{28}$$

we rewrite the denominator as

$$\begin{aligned} \alpha + (1 + \alpha)\rho + \lambda\rho^{1+\alpha} &= \alpha(1 + \rho) + \left(\rho + \lambda\rho^{1+\alpha}\right) \\ &= \alpha(1 + \rho) + 1 + \frac{1}{\lambda\rho^{\alpha}}. \end{aligned} \tag{29}$$

Therefore

$$\frac{k(m_1 + m_2)}{V} = \frac{\alpha(1 + \rho)}{\alpha(1 + \rho) + 1 + \frac{1}{\lambda\rho^{\alpha}}} < 1, \tag{30}$$

which implies

$$V - k(m_1 + m_2) > 0. \tag{31}$$

Now, let us consider the second-order behavior at the proposed solution. Let $h_i(m_1, m_2) = (\log \circ U_i)(m_1, m_2)$. For Player 1,

$$h_1(m_1, m_2) = \log\left(\lambda m_1^{\alpha}\right) - \log\left(\lambda m_1^{\alpha} + m_2^{\alpha}\right) + \log\left(V - k(m_1 + m_2)\right). \tag{32}$$

Differentiating with respect to m_1,

$$\begin{aligned} \frac{\partial h_1}{\partial m_1} &= \frac{\alpha}{m_1} - \frac{\lambda\alpha m_1^{\alpha-1}}{\lambda m_1^{\alpha} + m_2^{\alpha}} - \frac{k}{V - k(m_1 + m_2)}, \\ &= \frac{\alpha}{m_1(1 + \lambda\rho^{\alpha})} - \frac{k}{V - k(m_1 + m_2)}, \end{aligned} \tag{33}$$

where again $\rho = \frac{m_1}{m_2}$. The second derivative is

$$\frac{\partial^2 h_1}{\partial m_1^2} = -\frac{\alpha m_2^{\alpha}\left((1 + \alpha)\lambda m_1^{\alpha} + m_2^{\alpha}\right)}{m_1^2\left(\lambda m_1^{\alpha} + m_2^{\alpha}\right)^2} - \frac{k^2}{\left(V - k(m_1 + m_2)\right)^2} < 0, \tag{34}$$

where the inequality uses $V - k(m_1 + m_2) > 0$ derived above. Thus h_1 is strictly concave in m_1 for any fixed m_2, so $U_1 = \exp(h_1)$ is log-concave, hence quasiconcave, in its own action. The same calculation applies to Player 2, implying each best response is uniquely pinned down by its first-order condition, and the Nash equilibrium constructed from the FOCs is unique.

Finally, that the equilibrium levels vary smoothly in the parameters is obvious from their functional forms as given above. [*Back to the text*.] ∎

A.2 Preliminaries on Technologies

Let us re-state the assumptions we make about the technologies in our model.

---

4 Definition

The state's militarization technology *is a function*

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

We assume τ *possesses the following properties:*

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 $\left(\widetilde{\mathfrak{M}}_\tau\right)$: τ *is weakly increasing in all commodities; and*
4. Log-Concavity $\left(\tilde{\mathfrak{L}}_\tau\right)$: *the map*

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

   is concave.[16]

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

[16] We use the term "log-concavity" here in a nonstandard way. Ordinarily *log-concavity* refers to functions f such that $\log(f(x))$ is concave. Here, we use $\log(1 + \tau(x))$ to ensure that the function is well-defined at $\tau(x) = 0$. Many a regression-runner has been burned by the logarithm's misbehavior at zero, and nearly all of them remedy this 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.

Let us demonstrate that log-concavity implies quasiconcavity.

36 Lemma

If τ possesses $\widetilde{\mathfrak{L}}_\tau$, then τ possesses Weak Quasiconcavity $\left(\widetilde{\mathfrak{Q}}_\tau\right)$: for all $x_0, x_1 \in X$ and all $\lambda \in (0,1)$,

$$x_0 \neq x_1 \Longrightarrow \tau\left(\lambda x_0 + (1-\lambda)x_1\right) \geqslant \min\left\{\tau(x_0), \tau(x_1)\right\}.$$

Proof. Choose any τ possessing $\widetilde{\mathfrak{L}}_\tau$ and any $x_0, x_1 \in X$ such that $x_0 \neq x_1$. Choose any $\lambda \in (0,1)$, and define $x_\lambda := \lambda x_0 + (1-\lambda)x_1$. We need to show that $\tau(x_\lambda) > \min\{\tau(x_0), \tau(x_1)\}$. Without loss of generality, we may assume that $\tau(x_0) \leqslant \tau(x_1)$, so we need to show that $\tau(x_\lambda) > \tau(x_0)$.

Since τ possesses $\widetilde{\mathfrak{L}}_\tau$, we have

$$\log\left(1+\tau(x_\lambda)\right) \geqslant \lambda \log\left(1+\tau(x_0)\right) + (1-\lambda)\log\left(1+\tau(x_1)\right). \tag{35}$$

Exponentiating both sides, we have

$$1+\tau(x_\lambda) \geqslant \left(1+\tau(x_0)\right)^\lambda \left(1+\tau(x_1)\right)^{1-\lambda}. \tag{36}$$

Since $\tau(x_0) \leqslant \tau(x_1)$, we have $1+\tau(x_0) \leqslant 1+\tau(x_1)$, and thus

$$1+\tau(x_\lambda) \geqslant \left(1+\tau(x_0)\right)^\lambda \left(1+\tau(x_0)\right)^{1-\lambda} = 1+\tau(x_0). \tag{37}$$

Rearranging, we have $\tau\left(x_\lambda\right) \geqslant \tau(x_0)$, as desired. ∎

We define the following metric for the space of technologies.

37 Definition

For technologies $\tau_0, \tau_1 \in \mathcal{T}$, we define the distance[17]

$$d\left(\tau_0, \tau_1\right) = \sum_{n\in\mathbb{N}} \frac{1}{2^n} \times \frac{\max_{x\in[0,n]^L} |\tau_0(x) - \tau_1(x)|}{1 + \max_{x\in[0,n]^L} |\tau_0(x) - \tau_1(x)|}$$

Let us confirm that d is a metric.

[17] Since $[0,n]^L$ is compact for all $n \in \mathbb{N}$ and the map $x \mapsto |\tau_0(x) - \tau_1(x)|$ is continuous, we have taken the liberty of writing "max" in place of "sup" for d.

38 Lemma

d is a metric on $\mathcal{T}$.

Proof. We need to show that d satisfies the properties of a metric.

1. *Codomain:* we need to show that for all $\tau_0, \tau_1 \in \mathcal{T}$, we have $d(\tau_0, \tau_1) \in \mathbb{R}_{\geqslant 0}$. As we are taking maxima of absolute values, non-negativity is immediate. For finiteness, we observe that

$$\begin{aligned} d(\tau_0, \tau_1) &= \sum_{n \in \mathbb{N}} \frac{1}{2^n} \times \frac{\max_{x \in [0,n]^L} |\tau_0(x) - \tau_1(x)|}{1 + \max_{x \in [0,n]^L} |\tau_0(x) - \tau_1(x)|} \\ &< \sum_{n \in \mathbb{N}} \frac{1}{2^n} \times 1 = 1 < \infty. \end{aligned} \tag{38}$$

Thus, $d(\tau_0, \tau_1) \in \mathbb{R}_{\geqslant 0}$, and we officially write $d : \mathcal{T} \times \mathcal{T} \to [0, 1)$.

2. *Identity of Indiscernibles:* we need to show that for all $\tau_0, \tau_1 \in \mathcal{T}$, we have $d(\tau_0, \tau_1) = 0$ if and only if $\tau_0 = \tau_1$. For the first direction, suppose that $d(\tau_0, \tau_1) = 0$. Since $X = \bigcup_{n \in \mathbb{N}} [0, n]^L$, $d(\tau_0, \tau_1) = 0$ implies $\tau_0(x) = \tau_1(x)$ for all $x \in X$, implying $\tau_0 = \tau_1$. The other direction is immediate.

3. *Symmetry:* we need to show that for all $\tau_0, \tau_1 \in \mathcal{T}$, we have $d(\tau_0, \tau_1) = d(\tau_1, \tau_0)$. This is immediate from the symmetry of d_0 and d_1.

4. *Triangle Inequality:* we need to show that for all $\tau_0, \tau_1, \tau_2 \in \mathcal{T}$, we have $d(\tau_0, \tau_2) \leqslant d(\tau_0, \tau_1) + d(\tau_1, \tau_2)$. Consider any fixed $n \in \mathbb{N}$, and define the functions

$$\begin{aligned} \psi_n(\tau_0, \tau_1) &= \max_{x \in [0,n]^L} |\tau_0(x) - \tau_1(x)|, \text{ and} \\ \xi(\psi) &= \frac{\psi}{1 + \psi}, \end{aligned} \tag{39}$$

where $\xi : \mathbb{R}_{\geqslant 0} \to [0, 1)$. The nth component of the sum defining d is proportional to $\xi(\psi_n(\tau_0, \tau_1))$. Let us show that ψ_n is subadditive; choose and fix any $\tau_0, \tau_1, \tau_2 \in \mathcal{T}$, and let $\overline{x}_n \in [0, n]^L$ be a maximizer of $\psi_n(\tau_0, \tau_2)$. Then, we have

$$\begin{aligned} \psi_n(\tau_0, \tau_2) &= |\tau_0(\overline{x}_n) - \tau_2(\overline{x}_n)| \\ &= |\tau_0(\overline{x}_n) - \tau_1(\overline{x}_n) + \tau_1(\overline{x}_n) - \tau_2(\overline{x}_n)| \\ &\leqslant |\tau_0(\overline{x}_n) - \tau_1(\overline{x}_n)| + |\tau_1(\overline{x}_n) - \tau_2(\overline{x}_n)| \\ &\leqslant \psi_n(\tau_0, \tau_1) + \psi_n(\tau_1, \tau_2). \end{aligned} \tag{40}$$

The zeroth step is because $\overline{x}_n$ is a maximizer of $\psi_n(\tau_0, \tau_2)$; the first step simply substracts and adds $\tau_1(\overline{x}_n)$; the second step is because the absolute value function is subadditive; and the third step is from the definition of ψ_n. We therefore have shown that ψ_n is subadditive.

Now, consider ξ. Because ξ is increasing, (40) implies

$$\xi\left(\psi_n(\tau_0, \tau_2)\right) \leqslant \xi\left(\psi_n(\tau_0, \tau_1) + \psi_n(\tau_1, \tau_2)\right). \tag{41}$$

Because ξ is itself subadditive for non-negative arguments, (41) implies

$$\xi\left(\psi_n(\tau_0, \tau_2)\right) \leqslant \xi\left(\psi_n(\tau_0, \tau_1)\right) + \xi\left(\psi_n(\tau_1, \tau_2)\right). \tag{42}$$

We have shown that the nth component of the sum defining d is subadditive. Since n was arbitrary, we have shown that $d(\tau_0, \tau_2) \leqslant d(\tau_0, \tau_1) + d(\tau_1, \tau_2)$.

We have shown that d is a metric on $\mathcal{T}$. ■

Naturally, we use this metric to topologize the space of technologies.

---

39 Definition

The topology on $\mathcal{T}$ *is the topology induced by the metric d.*

Thus, the open sets in $\mathcal{T}$ are the unions of open balls of the form

$$B_\epsilon(\tau) = \left\{\tau' \in \mathcal{T} \,\middle|\, d\left(\tau, \tau'\right) < \epsilon\right\}, \tag{43}$$

for all $\tau \in \mathcal{T}$ and $\epsilon > 0$.

As a matter of course, we now define convergence in the space of technologies, which is standard uniform convergence on compact subsets of X.

---

40 Definition

Let $\{\tau_n\}_{n\in\mathbb{N}}$ be a sequence of technologies in $\mathcal{T}$, and let $\tau \in \mathcal{T}$ be a technology. We say that $\{\tau_n\}_{n\in\mathbb{N}}$ converges to τ under d *just in case for all $n \in \mathbb{N}$ and $K \in \mathbb{K}(X)$, we have*

$$\sup_{x\in K} |\tau_n(x) - \tau(x)| \to 0 \textit{ as } n \to \infty.$$

Because the domain $X = \mathbb{R}^L_{\geqslant 0}$ is covered by the ascending sequence of compact boxes $[0, K]^L$, convergence on all compact subsets of X is equivalent to convergence on each such box.

The next result plays a key role early in the proof of the proposition around which Section 2 is built.

41 Lemma

There exists a continuous function

$$\tilde{x} : M \times \mathcal{T} \longrightarrow X$$

such that $\tau(\tilde{x}(m, \tau)) = m$ *for all* $m \in M$ *and all* $\tau \in \mathcal{T}$.

Proof. We will prove this lemma in two steps. The first involves constructing a continuous selection from the upper contour set of τ. The second scales this selector to ensure that $\tau(\xi(m, \tau, \kappa)) = m$, not just $\tau(\xi(m, \tau, \kappa)) \geqslant m$.

Step 1: there exists a continuous selection from the upper contour set of τ. We will appeal to the Michael selection theorem (Aliprantis and Border, 2006, Theorem 17.66, pp. 589–590) for the map

$$\begin{aligned} \psi : M \times \mathcal{T} &\rightrightarrows X, \\ (m, \tau) &\mapsto \{x \in X \mid \tau(x) \geqslant m\}. \end{aligned} \tag{44}$$

This introduces a few requirements.

1. *Requirement 1:* the domain $M \times \mathcal{T}$ must be paracompact. It is well-known that the product of two paracompact spaces need not be paracompact, so we need a stronger condition for at least one of the spaces. Morita (1963, Theorem 1) showed that the product $X \times Y$ is normal and paracompact if X is normal, paracompact, and σ-locally compact and Y is normal and paracompact.

   (a) *M is normal, paracompact, and σ-locally compact.* Being a subspace of the metrizable space $\mathbb{R}$, $M \coloneqq \mathbb{R}_{\geqslant 0}$ is metrizable; since any metrizable space is perfectly normal (Aliprantis and Border, 2006, Corollary 3.21, p. 81) and paracompact (Aliprantis and Border, 2006, Theorem 3.22, pp. 81–83), we conclude that M is normal and paracompact. For σ-local compactness, we observe that

$$M \coloneqq \mathbb{R}_{\geqslant 0} = \bigcup_{n \in \mathbb{N}} [0, n], \tag{45}$$

which is a countable union of locally compact spaces. We conclude that M is σ-locally compact. M, then, may serve as the "X" in Morita's theorem.

(b) *$\mathcal{T}$ is normal and paracompact.* We have already constructed a metric d_τ on $\mathcal{T}$, so $\mathcal{T}$ equipped with the metric topology is metrizable; once again, we conclude that $\mathcal{T}$ is normal and paracompact. $\mathcal{T}$, then, may serve as the "Y" in Morita's theorem.

Thus, Morita's theorem ensures that the product $M \times \mathcal{T}$ is normal and paracompact, so it satisfies Requirement 1.

2. *Requirement 2:* the codomain $X = \mathbb{R}^L_{\geqslant 0}$ must be a Fréchet space—this means it must be completely metrizable and locally convex. Equip X with the Euclidean metric, and observe that this metric is complete on X because X is a closed subspace of the complete space $\mathbb{R}^L$. Being globally convex, X is locally convex. We conclude that the codomain X satisfies Requirement 2.

3. *Requirement 3:* the map ψ must take nonempty, closed, and convex values. Choose any $(m, \tau) \in M \times \mathcal{T}$. We must show that the set

$$\{x \in X \mid \tau(x) \geqslant m\}$$

is nonempty, closed, and convex. Nonemptiness follows because τ has $\mathfrak{S}_\tau$. Closedness follows because τ has $\mathfrak{C}_\tau$ and this set is the preimage of the closed set $[m, \infty)$ under τ. Convexity follows because τ has $\widetilde{\mathfrak{Q}}_\tau$. We conclude that the map ψ satisfies Requirement 3.

4. *Requirement 4:* the map ψ must be lower hemicontinuous. Choose any open set $V \subseteq X$, and consider the preimage $\psi^{-1}(V) \subseteq M \times \mathcal{T}$. Let $(m_0, \tau_0) \in \psi^{-1}(V)$, meaning there exists $x_0 \in V$ such that $x_0 \in \psi(m_0, \tau_0)$—that is, $\tau_0(x_0) \geqslant m_0$.

We seek a neighborhood of (m_0, τ_0) such that for all (m, τ) in this neighborhood, we have $\psi(m, \tau) \cap V \neq \varnothing$. We do this by showing that x_0 remains in the upper contour set of τ at m across that neighborhood. Note that the map $(m, \tau) \mapsto \tau(x_0) - m$ is continuous. Since $\tau_0(x_0) - m_0 \geqslant 0$, there exists $\varepsilon > 0$ and a neighborhood U of (m_0, τ_0) such that for all $(m, \tau) \in U$, we have $\tau(x_0) - m > -\varepsilon$. Choosing ε small enough ensures $\tau(x_0) \geqslant m$ throughout U. Thus, for all $(m, \tau) \in U$, we have $x_0 \in \psi(m, \tau) \cap V \neq \varnothing$, and so $(m, \tau) \in \psi^{-1}(V)$. This shows that $\psi^{-1}(V)$ is open, and we conclude that ψ is lower hemicontinuous. We therefore conclude that the map ψ satisfies Requirement 4.

These are all the requirements of the Michael selection theorem, so we conclude that there exists a continuous selector $\xi : M \times \mathcal{T} \to X$ such that $\tau\left(\xi(m,\tau)\right) \geqslant m$ for all $m \in M$ and $\tau \in \mathcal{T}$.

Step 2: we scale the selector to ensure that $\tau\left(\tilde{x}(m,\tau)\right) = m$. For any $m \in M$ and $\tau \in \mathcal{T}$, consider the function

$$\begin{aligned} \eta_{m,\tau} : [0,1] &\longrightarrow M, \\ t &\longmapsto \tau\left(t\xi(m,\tau)\right). \end{aligned} \tag{46}$$

We observe that $\eta_{m,\tau}$ is continuous, that $\eta_{m,\tau}(0) = 0$, and that $\eta_{m,\tau}(1) = \tau\left(\xi(m,\tau)\right) \geqslant m$. By the intermediate value theorem, there exists some $t^*_{m,\tau} \in [0,1]$ such that $\eta_{m,\tau}(t^*_{m,\tau}) = m$. Moreover, since τ possesses $\widetilde{\mathfrak{M}}_\tau$, $\eta_{m,\tau}$ is strictly increasing in t. This implies that $t^*_{m,\tau}$ is uniquely defined in $[0,1]$.

We now argue that the map $(m,\tau) \mapsto t^*_{m,\tau}$ is continuous. To do so, observe that the function

$$\begin{aligned} [0,1] \times M \times \mathcal{T} &\longrightarrow M \\ (t,m,\tau) &\longmapsto \tau\left(t\xi(m,\tau)\right) - m \end{aligned} \tag{47}$$

is jointly continuous in all arguments, as τ is continuous, scalar multiplication is continuous, and the selection function ξ was constructed to be continuous in (m,τ). Moreover, for each fixed (m,τ), the map $t \mapsto \tau(t\xi(m,\tau))$ is strictly increasing on $[0,1]$, so the zero set of this function is a singleton. Thus, the zero set of the function $(t,m,\tau) \mapsto \tau(t\xi(m,\tau)) - m$ is a continuous function of (m,τ), and we conclude that the map $(m,\tau) \mapsto t^*_{m,\tau}$ is continuous.

Conclusion. Finally, we define the continuous selector

$$\tilde{x}(m,\tau) \coloneqq t^*_{m,\tau}\xi(m,\tau). \tag{48}$$

We observe that $\tilde{x}(m,\tau)$ is continuous in (m,τ), and that

$$\tau\left(\tilde{x}(m,\tau)\right) = \tau\left(t^*_{m,\tau}\xi(m,\tau)\right) = \eta_{m,\tau}\left(t^*_{m,\tau}\right) = m. \tag{49}$$

This completes the proof of the lemma. ■

A.3 Preliminaries on Cost Functions

Again, we restate our assumptions about the cost function.

---

5 Definition

The state's cost function *is a function*

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

We assume κ possesses the following properties:

1. Continuity ($\mathfrak{C}_\kappa$): *κ is continuous;*
2. Centeredness ($\mathfrak{o}_\kappa$): *$\kappa(0) = 0$;*
3. Coerciveness ($\mathfrak{O}_\kappa$): *$\kappa(x) \to \infty$ 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*

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

   is strictly convex.

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

As with technologies, we observe that the shape condition entails quasiconvexity.

---

42 Lemma

If κ possesses $\mathfrak{L}_\kappa$, then it also possesses Strict Quasiconvexity ($\mathfrak{Q}_\kappa$): for all $x_0, x_1 \in X$ and all $\lambda \in (0,1)$,

$$x_0 \neq x_1 \Longrightarrow \kappa(\lambda x_0 + (1-\lambda)x_1) < \max\{\kappa(x_0), \kappa(x_1)\}.$$

Proof. Let κ satisfy $\mathfrak{L}_\kappa$, and let $x_0, x_1 \in X$ with $x_0 \neq x_1$. Choose any $\lambda \in (0,1)$, and define $x_\lambda \coloneqq \lambda x_0 + (1-\lambda)x_1$.

Let us assume without loss of generality that $\kappa(x_0) \geqslant \kappa(x_1)$. Then

$$\begin{aligned} \exp(\kappa(x_\lambda)) &< (1-\lambda)\exp(\kappa(x_0)) + \lambda\exp(\kappa(x_1)) \\ &< (1-\lambda)\exp(\kappa(x_0)) + \lambda\exp(\kappa(x_0)) = \exp(\kappa(x_0)), \end{aligned} \tag{50}$$

where the first inequality follows from strict convexity of $\exp \circ \kappa$, and the second from the assumption $\kappa(x_1) < \kappa(x_0) \Rightarrow \exp(\kappa(x_1)) < \exp(\kappa(x_0))$. Taking logarithms (which preserves strict inequality because log is strictly increasing), we obtain:

$$\kappa(x_\lambda) < \kappa(x_0) = \max\{\kappa(x_0), \kappa(x_1)\}, \tag{51}$$

which is what we wanted to show. We conclude that κ possesses $\mathfrak{Q}_\kappa$. ∎

We impose the same metric on the space of cost functions.

---

43 Definition

For cost functions $\kappa_0, \kappa_1 \in \mathcal{K}$, we define the distance[18]

$$d(\kappa_0, \kappa_1) = \sum_{n \in \mathbb{N}} \frac{1}{2^n} \times \frac{\max_{x \in [0,1]^L} |\kappa_0(x) - \kappa_1(x)|}{1 + \max_{x \in [0,1]^L} |\kappa_0(x) - \kappa_1(x)|}.$$

And again, we topologize the space of cost functions with the metric d_κ.

---

44 Definition

The topology on $\mathcal{K}$ *is the topology induced by the metric d_κ.*

A.4 For Section 2

Let us recall that the state's production problem is to choose a resource investment $x \in X$ that minimizes the cost of production $\kappa(x)$ while satisfying the desired force level m given the militarization technology τ.

---

6 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*

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

We take on the traditional questions, attempting to show that:

1. Problem SPP (m, τ, κ) has a solution for all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$;
2. this solution is unique for all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$; and
3. this solution varies continuously with m, τ, and κ.

The following is a useful start to this endeavor.

---

[18] We will now use d_τ to denote the metric on $\mathcal{T}$ and d_κ to denote the metric on $\mathcal{K}$. But notice that they work exactly the same way, and indeed they're even well defined if we attempted to measure the distance between a technology and a cost function. But we won't do that.

45 Lemma

There exists a continuous function

$$x^{\max} : M \times \mathcal{T} \times \mathcal{K} \longrightarrow X$$

such that, for all $(m, \tau, \kappa) \in M \times \mathcal{T} \times \mathcal{K}$, we have

$$x \text{ solves } \mathsf{SPP}\,(m, \tau, \kappa) \quad \Longrightarrow \quad x \in \underbrace{\prod_{\ell \in L} \left[0, x_\ell^{\max}(m, \tau, \kappa)\right]}_{:= X^{\max}(m,\tau,\kappa)}.$$

Moreover, for all $(m, \tau, \kappa) \in M \times \mathcal{T} \times \mathcal{K}$, we have $\tilde{x}(m, \tau) \in X^{\max}(m, \tau, \kappa)$, where $\tilde{x}(m, \tau)$ is the continuous selector from Lemma 41.

Proof. Choose any $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$. Lemma 41 guarantees the existence of a continuous selector $\tilde{x}(m, \tau)$ such that $\tau\,(\tilde{x}(m, \tau)) = m$, and we choose such a selector. We do not eliminate any minimizers by adding the requirement that $\kappa\,(x) \leqslant \kappa\,(\tilde{x}(m, \tau))$, as this is satisfied by all x that solve $\mathsf{SPP}\,(m, \tau, \kappa)$.

Now, for all $\ell \in L$ and all $\lambda \in \mathbb{R}_{\geqslant 0}$, we define the mobilization vector

$$\lambda_\ell = \lambda e_\ell \in \mathbb{R}^L_{\geqslant 0}, \tag{52}$$

where e_ℓ is the ℓth unit vector in $\mathbb{R}^L$. We observe that

$$\kappa\,(0_\ell) = 0 \quad \text{and} \quad \left[\lambda \to \infty \Rightarrow \|\lambda_\ell\| \to \infty \Rightarrow \lim_{\lambda \to \infty} \kappa\,(\lambda_\ell) = \infty\right], \tag{53}$$

where the second implication is because κ possesses $\mathfrak{O}_\kappa$. Since κ possesses $\mathfrak{C}_\kappa$, we may appeal to the intermediate value theorem to conclude that there exists some $\lambda_\ell^{\max} \in \mathbb{R}_{\geqslant 0}$ such that

$$\kappa\left(\lambda_\ell^{\max} e_\ell\right) = \kappa\,(\tilde{x}(m, \tau))\,. \tag{54}$$

And since κ possesses $\mathfrak{M}_\kappa$, this $\lambda_\ell^{\max}$ is unique.

It remains to show that $\lambda_\ell^{\max}$ is continuous in (m, τ, κ). Define

$$F\,(\lambda; m, \tau, \kappa) = \kappa\,(\lambda e_\ell) - \kappa\,(\tilde{x}\,(m, \tau))\,, \tag{55}$$

which is jointly continuous in λ and (m, τ, κ). Observe that $\lambda_\ell^{\max}$ is the unique solution to $F\,(\lambda; m, \tau, \kappa) = 0$ and that F is strictly increasing in λ. Thus, the root $\lambda_\ell^{\max}$ varies continuously in (m, τ, κ).

Finally, we define the continuous function

$$x_\ell^{\max}(m,\tau,\kappa) := \lambda_\ell^{\max} \in \mathbb{R}_{\geqslant 0}. \tag{56}$$

The bounding box $X^{\max}(m,\tau,\kappa) := \prod_{\ell\in L}\left[0, x_\ell^{\max}(m,\tau,\kappa)\right]$ is a compact subset of X guaranteed to contain all solutions to SPP (m,τ,κ).

For the final claim, we recall that $\lambda_\ell^{\max}$ is the unique scalar satisfying

$$\kappa\left(\lambda_\ell^{\max} e_\ell\right) = \kappa\left(\tilde{x}(m,\tau)\right). \tag{57}$$

Since κ possesses $\mathfrak{M}_\kappa$, it follows that

$$\tilde{x}^\ell(m,\tau) \leqslant \lambda_\ell^{\max} \quad \text{for all } \ell \in L; \tag{58}$$

were such not the case, then $\tilde{x}(m,\tau)$ would cost strictly more than $\lambda_\ell^{\max} e_\ell$ does in the ℓth commodity and at least as much as $\lambda_\ell^{\max} e_\ell$ does in all other commodities, meaning their costs could not be equal. Thus, we have shown that $\tilde{x}(m,\tau) \in X^{\max}(m,\tau,\kappa)$, as claimed. This completes the proof of the lemma. ∎

Thus, we have shown that there exists a compact bounding box $X^{\max}(m,\tau,\kappa)$ that contains all solutions to the state's production problem SPP (m,τ,κ) for all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$. Moreover, the bounds of this bounding box vary continuously with (m,τ,κ). As a result, we can use this bounding box to construct a continuous constraint set for the state's production problem.

---

46 Lemma

The correspondence

$$X^{\max} : M \times \mathcal{T} \times \mathcal{K} \rightrightarrows X$$

is upper and lower hemicontinuous.

Proof. We address upper and lower hemicontinuity in turn.

Upper hemicontinuity. Let $\mathcal{V} \subseteq X$ be open such that $X^{\max}(m^*,\tau^*,\kappa^*) \subseteq \mathcal{V}$ for some $(m^*,\tau^*,\kappa^*) \in M \times \mathcal{T} \times \mathcal{K}$. We need to show that there exists a neighborhood $\mathcal{U}$ of (m^*,τ^*,κ^*) such that for all $(m,\tau,\kappa) \in \mathcal{U}$, we have

$$X^{\max}(m,\tau,\kappa) \subseteq \mathcal{V}. \tag{59}$$

Since $\mathcal{V}$ is open and $X^{\max}(m^*,\tau^*,\kappa^*) \subseteq \mathcal{V}$, there exists $\varepsilon > 0$ such that the open box $B_\varepsilon(X^{\max}(m^*,\tau^*,\kappa^*)) \subseteq \mathcal{V}$. Now note: for each $\ell \in L$, the continuity of

$x_\ell^{\max}$ implies the existence of an open neighborhood $\mathcal{U}_\ell$ of (m^*, τ^*, κ^*) such that for all $(m, \tau, \kappa) \in \mathcal{U}_\ell$, we have

$$x_\ell^{\max}(m, \tau, \kappa) < x_\ell^{\max}(m^*, \tau^*, \kappa^*) + \varepsilon. \tag{60}$$

Define $\mathcal{U} := \bigcap_{\ell \in L} \mathcal{U}_\ell$, which is open and contains (m^*, τ^*, κ^*). Then for every $(m, \tau, \kappa) \in \mathcal{U}$, we have $X^{\max}(m, \tau, \kappa) \subseteq B_\varepsilon(X^{\max}(m^*, \tau^*, \kappa^*)) \subseteq \mathcal{V}$. Thus, $X^{\max}$ is upper hemicontinuous.

Lower hemicontinuity. Let $\mathcal{V} \subseteq X$ be open, and suppose $(m^*, \tau^*, \kappa^*) \in M \times \mathcal{T} \times \mathcal{K}$ and $x^* \in X^{\max}(m^*, \tau^*, \kappa^*) \cap \mathcal{V}$. We need to show that there exists a neighborhood $\mathcal{U}$ of (m^*, τ^*, κ^*) such that for all $(m, \tau, \kappa) \in \mathcal{U}$, we have

$$X^{\max}(m, \tau, \kappa) \cap \mathcal{V} \neq \varnothing. \tag{61}$$

Since $\mathcal{V}$ is open and $x^* \in \mathcal{V}$, there exists $\varepsilon > 0$ such that the open box $B_\varepsilon(x^*) \subseteq \mathcal{V}$. Now note: for each $\ell \in L$, the continuity of $x_\ell^{\max}$ implies the existence of an open neighborhood $\mathcal{U}_\ell$ of (m^*, τ^*, κ^*) such that for all $(m, \tau, \kappa) \in \mathcal{U}_\ell$, we have

$$x_\ell^{\max}(m, \tau, \kappa) > x_\ell^* - \varepsilon. \tag{62}$$

Define $\mathcal{U} := \bigcap_{\ell \in L} \mathcal{U}_\ell$, which is open and contains (m^*, τ^*, κ^*).

Then for every $(m, \tau, \kappa) \in \mathcal{U}$, we have $X^{\max}(m, \tau, \kappa) \cap B_\varepsilon(x^*) \neq \varnothing$. Since $B_\varepsilon(x^*) \subseteq \mathcal{V}$, we conclude that $X^{\max}(m, \tau, \kappa) \cap \mathcal{V} \neq \varnothing$. Thus, $X^{\max}$ is lower hemicontinuous.

This completes the proof that $X^{\max}$ is continuous. ∎

We now use this bounding box in tandem with the level set at m to construct a continuous constraint set for the state's production problem.

---

47 Lemma

The correspondence

$$\begin{aligned} &\mathcal{X} : M \times \mathcal{T} \times \mathcal{K} \rightrightarrows X, \\ &\qquad (m, \tau, \kappa) \mapsto X^{\max}(m, \tau, \kappa) \cap \tau^{-1}(\{m\}), \end{aligned}$$

is nonempty, compact-valued, and continuous.

Proof. Choose any $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$. We will address nonemptiness, compactness, and continuity in turn.

Nonemptiness. From Lemma 45, we know that the bounding box $X^{\max}(m, \tau, \kappa)$ contains $\tilde{x}(m, \tau)$, which is a continuous selector such that $\tau(\tilde{x}(m, \tau)) = m$. Thus, $\tilde{x}(m, \tau) \in \tau^{-1}(\{m\})$, and we conclude that $\mathcal{X}(m, \tau, \kappa)$ is nonempty.

Compactness. The bounding box $X^{\max}(m, \tau, \kappa)$ is compact because it is a finite product of compact intervals in $\mathbb{R}_{\geqslant 0}$, and the level set $\tau^{-1}(\{m\})$ is closed because τ possesses $\mathfrak{C}_\tau$. The intersection of a compact set with a closed set is compact, so we conclude that $\mathcal{X}(m, \tau, \kappa)$ is compact.

Upper hemicontinuity. Aliprantis and Border (2006, Theorem 17.25, pp. 567–568) demonstrate that the intersection of an upper hemicontinuous, compact-valued correspondence with a closed-valued correspondence is upper hemicontinuous. We observe that $X^{\max}$ is upper hemicontinuous (by Lemma 46) and compact-valued (by construction), and that $\tau^{-1}(\{m\})$ is closed-valued because τ possesses $\mathfrak{C}_\tau$. Thus, we conclude that $\mathcal{X}$ is upper hemicontinuous.

Lower hemicontinuity. Choose $(m, \tau, \kappa) \in M \times \mathcal{T} \times \mathcal{K}$, and suppose $V \subseteq X$ is an open set satisfying $V \cap \mathcal{X}(m, \tau, \kappa) \neq \varnothing$. Then there exists some $x \in V$ such that $x \in \mathcal{X}(m, \tau, \kappa)$, meaning that $x \in X^{\max}(m, \tau, \kappa)$ and $\tau(x) = m$. We need to identify a neighborhood $U \subseteq M \times \mathcal{T} \times \mathcal{K}$ of (m, τ, κ) such that for all $\left(m', \tau', \kappa'\right) \in U$, we have $V \cap \mathcal{X}\left(m', \tau', \kappa'\right) \neq \varnothing$. By the continuity of $X^{\max}$ (from Lemma 46), we can choose a neighborhood $U_1 \subseteq M \times \mathcal{T} \times \mathcal{K}$ of (m, τ, κ) such that for all $\left(m', \tau', \kappa'\right) \in U_1$, we have

$$x \in X^{\max}\left(m', \tau', \kappa'\right). \tag{63}$$

Since τ possesses $\mathfrak{C}_\tau$, we can also choose a neighborhood $U_2 \subseteq M \times \mathcal{T}$ of (m, τ) such that for all $\left(m', \tau'\right) \in U_2$, we have

$$\tau(x) = m \quad \Longrightarrow \quad x \in \tau^{-1}\left(\left\{m'\right\}\right). \tag{64}$$

Now define

$$\mathcal{U} := U_1 \cap (U_2 \times \mathcal{K}) \subseteq M \times \mathcal{T} \times \mathcal{K}. \tag{65}$$

Being a finite intersection of open sets, $\mathcal{U}$ is open and contains (m, τ, κ). We claim that for all $\left(m', \tau', \kappa'\right) \in \mathcal{U}$, we have

$$V \cap \mathcal{X}\left(m', \tau', \kappa'\right) \neq \varnothing. \tag{66}$$

To see this, observe that for all $\left(m', \tau', \kappa'\right) \in \mathcal{U}$, we have

$$x \in X^{\max}\left(m', \tau', \kappa'\right) \quad \text{and} \quad x \in \tau^{-1}\left(\left\{m'\right\}\right). \tag{67}$$

Thus, we conclude that $x \in \mathcal{X}\left(m', \tau', \kappa'\right)$, meaning that $V \cap \mathcal{X}\left(m', \tau', \kappa'\right) \neq \varnothing$, as required. This completes the proof of the lemma. ■

Having done all the heavy lifting, we can now move on to the main results of the section. First, we demonstrate that the state's production problem has a solution for all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$.

---

48 Lemma

For all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$, SPP (m, τ, κ) has a solution.

Proof. This follows immediately from Lemma 45: any solution to SPP (m, τ, κ) must lie in the compact set $X^{\max}(m, \tau, \kappa)$, which is nonempty by Lemma 46. The cost function κ possesses $\mathfrak{C}_\kappa$, so SPP (m, τ, κ) is a continuous optimization problem over a compact set, and thus it has a solution—this is from the Weierstrass extreme value theorem. ■

Next, we show that the solution to the state's production problem is unique for all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$, a simple consequence of our convexity assumptions about technologies and cost functions.

---

49 Lemma

For all $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$, SPP (m, τ, κ) has a unique solution.

Proof. Choose any $m \in M$, $\tau \in \mathcal{T}$, and $\kappa \in \mathcal{K}$. For sake of contradiction, suppose that there exist two distinct solutions $x_0, x_1 \in X$ to SPP (m, τ, κ). Thus, our assumption implies that

$$\kappa\left(x_0\right) = \kappa\left(x_1\right) = \min_{x \in X} \kappa(x) \quad \text{and} \tag{68}$$

$$\tau\left(x_0\right) = \tau\left(x_1\right) = m. \tag{69}$$

Since X is convex, we may define

$$x_{1/2} := \frac{1}{2}x_0 + \frac{1}{2}x_1 \in X. \tag{70}$$

Since κ possesses $\mathfrak{Q}_\kappa$, $x_0 \neq x_1$ implies

$$\kappa\left(x_{1/2}\right) < \min\left\{\kappa\left(x_0\right), \kappa\left(x_1\right)\right\} = \kappa\left(x_0\right) = \kappa\left(x_1\right).$$

Since τ possesses $\widetilde{\mathfrak{Q}}_\tau$, $x_0 \neq x_1$ implies

$$\tau\left(x_{1/2}\right) \geqslant \tau\left(x_0\right) = \tau\left(x_1\right) = m. \tag{71}$$

In case $\tau\left(x_{1/2}\right) = m$, we have found a contradiction to the optimality of x_0 and x_1 because

$$\kappa\left(x_{1/2}\right) < \min\left\{\kappa\left(x_0\right), \kappa\left(x_1\right)\right\} = \kappa\left(x_0\right) = \kappa\left(x_1\right). \tag{72}$$

Thus, we may suppose without loss of generality that $\tau\left(x_{1/2}\right) > m$. Define the function

$$\begin{aligned} \eta : [0,1] &\longrightarrow M \\ t &\longmapsto \tau\left(t x_{1/2}\right) \end{aligned} \tag{73}$$

and observe that $\eta(0) = \tau(0) = 0$ and $\eta(1) = \tau\left(x_{1/2}\right) > m$. Moreover, η is continuous because τ possesses $\mathfrak{C}_\tau$ and $x_{1/2}$ does not depend on t. By the intermediate value theorem, there exists some $t^* \in (0,1)$ such that $\eta(t^*) = m$. But then, the fact that κ possesses $\mathfrak{M}_\kappa$ implies that

$$\kappa\left(t^* x_{1/2}\right) < \min\left\{\kappa\left(x_0\right), \kappa\left(x_1\right)\right\} = \kappa\left(x_0\right) = \kappa\left(x_1\right), \tag{74}$$

which contradicts the fact that x_0 and x_1 are both solutions to SPP (m, τ, κ). Thus, we conclude that the solution to SPP (m, τ, κ) is unique. ■

We pause to record the main-text statement of the previous two lemmas.

---

8 Lemma

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

Proof. This is a restatement of Lemmas 48 and 49, which the reader may find immediately above. [*Back to the text*.] ■

Finally, we show that the solution to the state's production problem varies continuously with m, τ, and κ.

---

9 Lemma

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

Proof. We will appeal to Berge's theorem (Aliprantis and Border, 2006, Theorem 17.31, pp. 570–571). In Lemma 47, we showed that the correspondence $\mathcal{X}$ is nonempty, compact-valued, and continuous; this is just what is required on constraint sets. As for the objective function, note that the map

$$(x;\kappa) \longmapsto \kappa(x) \tag{75}$$

is continuous in x and κ because κ possesses $\mathfrak{C}_\kappa$. This is just what is required on objective functions. Thus, we apply Berge's theorem to conclude that the solution to $\mathsf{SPP}\,(m,\tau,\kappa)$ varies continuously with m, τ, and κ. [*Back to the text*.] ∎

Having obtained the main results about $\mathsf{SPP}\,(m,\tau,\kappa)$, we now turn to naming the solutions. We first consider the *raw* solution to the state's production problem, which we denote by $x^*(m,\tau,\kappa)$.

50 Definition

For all $(m,\tau,\kappa) \in M \times \mathcal{T} \times \mathcal{K}$, *let* $x^*(m,\tau,\kappa)$ *be the* raw solution *to the state's production problem* $\mathsf{SPP}\,(m,\tau,\kappa)$. *The set of all raw solutions is denoted by*

$$\mathsf{Raw} \coloneqq \left\{x^*(m,\tau,\kappa) \in X \;\middle|\; (m,\tau,\kappa) \in M \times \mathcal{T} \times \mathcal{K}\right\}.$$

Our next theoretical maneuver is to curry the raw solution to the state's production problem with respect to the militarization technology τ and the cost function κ.

51 Definition

Define the curried solution function

$$\begin{aligned} \pi_{\tau,\kappa} : M &\longrightarrow X \\ m &\longmapsto x^*(m,\tau,\kappa). \end{aligned}$$

The set of all curried solutions is denoted by

$$\mathcal{P}_{\mathcal{T}\times\mathcal{K}} \coloneqq \left\{\pi_{\tau,\kappa} : M \to X \;\middle|\; (\tau,\kappa) \in \mathcal{T} \times \mathcal{K}\right\}.$$

In other words, $\pi_{\tau,\kappa}$ *is the function that maps each desired force level* $m \in M$ *to the raw solution* $x^*(m,\tau,\kappa)$ *to the state's production problem* $\mathsf{SPP}\,(m,\tau,\kappa)$ *for the given militarization technology* τ *and cost function* κ.

We now have a complete definition of the state's production problem, its solution, and the curried solution function. We need to show that the curried solution function is continuous with respect to the militarization technology τ and the cost function κ.

52 Lemma

The curried solution function $\pi_{\tau,\kappa}$ is continuous with respect to the militarization technology τ and the cost function κ when the function space X^M is endowed with the compact-open topology. In other words, for all $(\tau^*, \kappa^*) \in \mathcal{T} \times \mathcal{K}$, the map

$$\begin{aligned} \mathcal{M} : \mathcal{T} \times \mathcal{K} &\longrightarrow X^M, \\ (\tau, \kappa) &\longmapsto \pi_{\tau,\kappa} \end{aligned}$$

is continuous at (τ^*, κ^*) in the compact-open topology. Even more explicitly, for all $\varepsilon > 0$, there exists a neighborhood $\mathcal{U}$ of (τ^*, κ^*) such that for all $(\tau, \kappa) \in \mathcal{U}$, we have

$$\|\pi_{\tau,\kappa} - \pi_{\tau^*,\kappa^*}\| < \varepsilon,$$

where $\|\cdot\|$ is the supremum norm on X^M.

Proof. Choose any $(\tau^*, \kappa^*) \in \mathcal{T} \times \mathcal{K}$ and any $\varepsilon > 0$. Let $K \subseteq M$ be any compact subset. By Lemma 9, the raw solution map $(m, \tau, \kappa) \mapsto x^*(m, \tau, \kappa)$ is jointly continuous, and thus it is uniformly continuous on the compact set $K \times \{\tau^*\} \times \{\kappa^*\}$. Thus, there exists a neighborhood $\mathcal{U}$ of (τ^*, κ^*) such that for all $(\tau, \kappa) \in \mathcal{U}$, we have

$$\sup_{m \in K} \|x^*(m, \tau, \kappa) - x^*(m, \tau^*, \kappa^*)\| < \varepsilon. \tag{76}$$

Since $\pi_{\tau,\kappa}(m) = x^*(m, \tau, \kappa)$, this implies

$$\sup_{m \in K} \|\pi_{\tau,\kappa}(m) - \pi_{\tau^*,\kappa^*}(m)\| < \varepsilon, \tag{77}$$

which shows that the map $(\tau, \kappa) \mapsto \pi_{\tau,\kappa}$ is continuous at (τ^*, κ^*) in the compact-open topology on X^M. This completes the proof of the lemma. ∎

We now give the main text statement of the previous lemma.

10 Corollary

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

Proof. This is a restatement of Lemma 52, which the reader may find immediately above. [*Back to the text*.] ∎

Our final lemma for Section 2 demonstrates that the function spaces $\mathcal{T}$ and $\mathcal{K}$ are both contractible with respect to the compact-open topology.

---

12 Lemma

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

Proof. We will construct homotopies for both $\mathcal{T}$ and $\mathcal{K}$.

Homotopy for $\mathcal{T}$. We define the target function

$$\tau_0(x) = \sum_{\ell \in L} \log(1 + x_\ell). \tag{78}$$

Let us confirm that $\tau_0 \in \mathcal{T}$:

1. Ray-Surjectivity $(\mathfrak{R}_\tau)$. We need to show that there exists a point $v \in X$ such that the map

$$t \longmapsto \tau_0(tv) : \mathbb{R}_{\geqslant 0} \longrightarrow M \tag{79}$$

is continuous, strictly increasing, and surjective. Take $v = \mathbf{1} = (1)_{\ell \in L}$. Then for all $t \in \mathbb{R}_{\geqslant 0}$, we have

$$\tau_0(tv) = \sum_{\ell \in L} \log(1 + t) = |L| \log(1 + t). \tag{80}$$

The map $t \mapsto |L| \log(1 + t)$ is continuous and strictly increasing because the logarithm is continuous and strictly increasing on $\mathbb{R}_{>0}$. At $t = 0$, we have $|L| \log(1 + 0) = 0$, so the map attains the minimum of $M = \mathbb{R}_{\geqslant 0}$. Moreover, $\lim_{t \to \infty} |L| \log(1 + t) = \infty$; an appeal to the intermediate value theorem shows that the map attains every value in M. Thus, the map $t \mapsto \tau_0(tv)$ is continuous, strictly increasing, and surjective, as required.

2. Continuity $(\mathfrak{C}_\tau)$. This is immediate because τ_0 is a finite sum of continuous functions.

3. Weak Monotonicity $\left(\widetilde{\mathfrak{M}}_\tau\right)$. Since τ_0 is smooth, we may compute the gradient:

$$D\tau_0\left(x\right) = \left(\frac{1}{1+x_\ell}\right)_{\ell\in L}. \tag{81}$$

Since $x_\ell \geqslant 0$ for all $\ell \in L$, we have $1 + x_\ell > 0$ for all $\ell \in L$, and thus $D\tau_0\left(x\right) > 0$ for all $x \in X$.

4. Log-Concavity $\left(\tilde{\mathfrak{L}}_\tau\right)$. We need to show that

$$h(x) = \log\left(1+\tau_0(x)\right) = \log\left(1 + \sum_{\ell\in L}\log\left(1+x_\ell\right)\right) \tag{82}$$

is concave in x. For fun, let us take the scenic route and compute the first and second derivatives of this function. From the chain rule, the first derivative is given by

$$\begin{aligned} Dh(x) &= \frac{1}{1+\tau_0(x)} D\tau_0(x), \\ &= \left(\frac{1}{1+\sum_{\ell\in L}\log\left(1+x_\ell\right)}\right)\left(\frac{1}{1+x_\ell}\right)_{\ell\in L}. \end{aligned} \tag{83}$$

Then by the product rule, the elements of the Hessian matrix are given by

$$\begin{aligned} D^2_{\ell\ell}h(x) &= -\frac{1}{\left(1+\tau_0(x)\right)\left(1+x_\ell\right)^2} - \frac{1}{\left(1+\tau_0(x)\right)^2\left(1+x_\ell\right)^2}, \\ D^2_{\ell k}h(x) &= -\frac{1}{\left(1+\tau_0(x)\right)^2\left(1+x_\ell\right)\left(1+x_k\right)} \quad \text{for } \ell \neq k. \end{aligned} \tag{84}$$

We write the Hessian with the form

$$\begin{aligned} D^2h(x) &= -\left(\frac{1}{1+\tau_0(x)}\right)A - \left(\frac{1}{\left(1+\tau_0(x)\right)^2}\right)\left(A + uu^\top\right), \\ A &\coloneqq \operatorname{diag}\left(\frac{1}{\left(1+x_\ell\right)^2}\right)_{\ell\in L}, \\ u &\coloneqq \left(\frac{1}{1+x_\ell}\right)_{\ell\in L}. \end{aligned} \tag{85}$$

We observe that A, being a diagonal matrix with strictly positive entries, is positive definite. Now choose any $z \in \mathbb{R}^L \setminus \{0\}$ and compute

$$\begin{aligned} z^\top D^2 h(x) z = &-\left(\frac{1}{1+\tau_0(x)}\right) z^\top A z - \left(\frac{1}{(1+\tau_0(x))^2}\right) z^\top A z \\ &- \left(\frac{1}{(1+\tau_0(x))^2}\right) z^\top u u^\top z. \end{aligned} \tag{86}$$

The first two terms are negative because A is positive definite and $1 + \tau_0(x) > 0$. For the third term, observe that

$$z^\top u u^\top z = \left(z^\top u\right)^2 \geq 0,$$

Equation (86) therefore implies that $z^\top D^2 h(x) z < 0$ for all $z \in \mathbb{R}^L \setminus \{0\}$, which shows that h is concave (in fact, strictly concave) in x.

We conclude that τ_0 satisfies all four properties, and thus $\tau_0 \in \mathcal{T}$.

Now we define the homotopy

$$\begin{aligned} H : \mathcal{T} \times [0,1] &\longrightarrow \mathcal{T}, \\ (\tau, t) &\longmapsto (1+\tau)^{1-t} \times (1+\tau_0)^t - 1. \end{aligned} \tag{87}$$

We need to show that H is continuous in (τ, t). Since $\mathcal{T}$ is metrized by d_τ, it suffices to show that for any $\epsilon > 0$, there exists a neighborhood of (τ, t) such that for all $\left(\tau', t'\right)$ sufficiently close to (τ, t), we have $d\left(H\left(\tau', t'\right), H(\tau, t)\right) < \epsilon$. Fix any $n \in \mathbb{N}$. Over the compact set $[0, n]^L$, the map

$$(x; \tau, t) \mapsto H(\tau, t)(x) = (1+\tau(x))^{1-t} \times (1+\tau_0(x))^t - 1 \tag{88}$$

is jointly continuous in (τ, t), since $(\tau, x) \mapsto \tau(x)$ is continuous under d, and the arithmetic operations are smooth. Therefore, $d_\tau\left(H(\tau, t), H\left(\tau', t'\right)\right)$ is small for small perturbations in (τ, t), and H is continuous in d.

We also need to show that $H(\tau, t) \in \mathcal{T}$ for all $\tau \in \mathcal{T}$ and $t \in [0,1]$. Choose and fix any such $\tau \in \mathcal{T}$ and $t \in [0,1]$.

1. Ray-Surjectivity $(\mathfrak{R}_\tau)$. Since τ satisfies $(\mathfrak{R}_\tau)$, there exists a point $v \in X$ such that the map

$$a(s) := 1 + \tau(sv) \tag{89}$$

is continuous, strictly increasing, and unbounded as $s \to \infty$. For $\tau_0(x) = \sum_{\ell \in L} \log(1 + x_\ell)$, we also have, along the same ray,

$$b(s) \coloneqq 1 + \tau_0(sv) = 1 + \sum_{\ell \in L} \log(1 + sv_\ell), \tag{90}$$

which is continuous, strictly increasing, and unbounded since at least one $v_\ell > 0$.

Now, for any fixed $t \in [0, 1]$, define

$$\begin{aligned} F_t(s) &\coloneqq 1 + H(\tau, t)(sv), \\ &= a(s)^{1-t}\, b(s)^t. \end{aligned} \tag{91}$$

Being a product of continuous functions, $F_t(s)$ is continuous in s. Moreover, since both $a(s)$ and $b(s)$ are strictly increasing in s, and the map $z \mapsto z^c$ is strictly increasing for any $c > 0$, it follows that $F_t(s)$ is strictly increasing in s. Finally, since both $a(s)$ and $b(s)$ are unbounded as $s \to \infty$, it follows that $F_t(s)$ is unbounded as $s \to \infty$. Thus, we conclude that $H(\tau, t)$ satisfies $(\mathfrak{R}_\tau)$.

2. Continuity $(\mathfrak{C}_\tau)$. This is immediate because $H(\tau, t)$ is a finite product of continuous functions.

3. Weak Monotonicity $\left(\widetilde{\mathfrak{M}}_\tau\right)$. Define

$$\phi(x) \coloneqq (1 + \tau(x))^{1-t}, \qquad \psi(x) \coloneqq (1 + \tau_0(x))^t. \tag{92}$$

Since both τ and τ_0 are weakly increasing and nonnegative, and the map $z \mapsto (1 + z)^a$ is strictly increasing for any $a > 0$, it follows that $\phi(x)$ and $\psi(x)$ are each weakly increasing in x. Moreover, since τ is strictly increasing in at least one coordinate at each point, and τ_0 is strictly increasing in all coordinates, we conclude that the product $\phi(x) \times \psi(x)$ is strictly increasing in at least one coordinate at each point.

4. Log-Concavity $\left(\tilde{\mathfrak{L}}_\tau\right)$. We need to show that

$$\begin{aligned} \log\left(1 + H(\tau, t)(x)\right) = (1 - t)\log\left(1 + \tau(x)\right) \\ + t\log\left(1 + \tau_0(x)\right) \end{aligned} \tag{93}$$

is strictly concave in x. Since τ and τ_0 both possess $\tilde{\mathfrak{L}}_\tau$, this is a sum of two concave functions, and thus it is concave in x.

We conclude that $H(\tau, t) \in \mathcal{T}$ for all $\tau \in \mathcal{T}$ and $t \in [0, 1]$.

Finally, as a matter of course, we confirm that H is a strong deformation retraction from $\mathcal{T}$ onto τ_0:

1. At $t = 0$, we have

$$H(\tau, 0) = (1 + \tau)^{1-0} \times (1 + \tau_0)^{0} - 1 = \tau, \tag{94}$$

which shows that $H(\tau, 0) = \tau$ for all $\tau \in \mathcal{T}$, as required.

2. At $t = 1$, we have

$$H(\tau, 1) = (1 + \tau)^{1-1} \times (1 + \tau_0)^{1} - 1 = \tau_0, \tag{95}$$

which shows that $H(\tau, 1) = \tau_0$ for all $\tau \in \mathcal{T}$, as required.

3. For all $t \in [0, 1]$, we have

$$H(\tau_0, t) = (1 + \tau_0)^{1-t} \times (1 + \tau_0)^{t} - 1 = (1 + \tau_0) - 1 = \tau_0, \tag{96}$$

which shows that $H(\tau_0, t) = \tau_0$ for all $t \in [0, 1]$, as required.

We conclude that H is a strong deformation retraction from $\mathcal{T}$ onto τ_0, and thus $\mathcal{T}$ is contractible with respect to the compact-open topology.

Homotopy for $\mathcal{K}$. We define the target function

$$\kappa_0(x) = \sum_{\ell \in L} x_\ell. \tag{97}$$

Though this one is a bit more straightforward, let us confirm that $\kappa_0 \in \mathcal{K}$:

1. Continuity $(\mathfrak{C}_\kappa)$. This is immediate because κ_0 is a finite sum of continuous functions.

2. Centeredness $(\mathfrak{o}_\kappa)$. Evidently, $\kappa_0(0) = 0$.

3. Coerciveness $(\mathfrak{O}_\kappa)$. The coordinatewise limit of $\kappa_0(x)$ is ∞; *a fortiori*, the norm limit is ∞.

4. Strict Monotonicity $(\mathfrak{M}_\kappa)$. Again, this is immediate because κ_0 is a finite sum of strictly increasing functions.

5. Strict Exp-Convexity ($\mathfrak{L}_\kappa$). We need to show that the map

$$x \longmapsto \exp\left(\sum_{\ell \in L} x_\ell\right)$$

is strictly convex. κ_0 is linear, and the exponential function is strictly convex and increasing, so their composition is strictly convex.

So, we can move on to defining the homotopy, which we set as

$$\begin{aligned} H : \mathcal{K} \times [0,1] &\longrightarrow \mathcal{K}, \\ (\kappa, t) &\longmapsto \log\left((1-t)\exp\kappa + t\exp\kappa_0\right). \end{aligned}$$

We need to show that H is continuous in (κ, t). Since $\mathcal{K}$ is metrized by d_κ, it suffices to show that for any $\epsilon > 0$, there exists a neighborhood of (κ, t) such that for all $\left(\kappa', t'\right)$ sufficiently close to (κ, t), we have $d\left(H\left(\kappa', t'\right), H(\kappa, t)\right) < \epsilon$. Fix any $n \in \mathbb{N}$. Over the compact set $[0, n]^L$, the map

$$(x; \kappa, t) \mapsto H(\kappa, t)(x) = \log\left((1-t)\exp\kappa(x) + t\exp\kappa_0(x)\right) \tag{98}$$

is jointly continuous in (κ, t), since $(\kappa, x) \mapsto \kappa(x)$ is continuous under d, and the arithmetic operations are smooth. Therefore, $d_\kappa\left(H(\kappa, t), H\left(\kappa', t'\right)\right)$ is small for small perturbations in (κ, t), and H is continuous in d.

We also need to show that $H(\kappa, t) \in \mathcal{K}$ for all $\kappa \in \mathcal{K}$ and $t \in [0,1]$. Choose and fix any such $\kappa \in \mathcal{K}$ and $t \in [0,1]$.

1. Continuity ($\mathfrak{C}_\kappa$). This is immediate because $H(\kappa, t)$ is a composition of continuous functions.

2. Centeredness ($\mathfrak{o}_\kappa$). The weighted average inside the logarithm evaluates to 1 at $x = 0$, so $H(\kappa, t)(0) = \log(1) = 0$.

3. Coerciveness ($\mathfrak{O}_\kappa$). As $\|x\| \to \infty$, at least one of $\kappa(x)$ or $\kappa_0(x)$ goes to ∞, so the weighted average inside the logarithm goes to ∞, and thus $H(\kappa, t)(x)$ goes to ∞.

4. Strict Monotonicity ($\mathfrak{M}_\kappa$). Since both κ and κ_0 are strictly increasing, and the map $z \mapsto \log(z)$ is strictly increasing for $z > 0$, it follows that $H(\kappa, t)$ is strictly increasing.

5. Strict Exp-Convexity ($\mathfrak{L}_\kappa$). We need to show that the map

$$x \longmapsto \exp\left(H(\kappa, t)(x)\right) = (1-t)\exp\kappa(x) + t\exp\kappa_0(x) \tag{99}$$

is strictly convex. Since κ and κ_0 both possess $\mathfrak{L}_\kappa$, this is a positive weighted sum of two strictly convex functions, and thus it is strictly convex in x.

We conclude that $H(\kappa, t) \in \mathcal{K}$ for all $\kappa \in \mathcal{K}$ and $t \in [0, 1]$.

Finally, as a matter of course, we confirm that H is a strong deformation retraction from $\mathcal{K}$ onto κ_0:

1. At $t = 0$, we have

$$H(\kappa, 0) = \log((1-0)\exp\kappa + 0\exp\kappa_0) = \kappa, \tag{100}$$

which shows that $H(\kappa, 0) = \kappa$ for all $\kappa \in \mathcal{K}$, as required.

2. At $t = 1$, we have

$$H(\kappa, 1) = \log((1-1)\exp\kappa + 1\exp\kappa_0) = \kappa_0, \tag{101}$$

which shows that $H(\kappa, 1) = \kappa_0$ for all $\kappa \in \mathcal{K}$, as required.

3. For all $t \in [0, 1]$, we have

$$H(\kappa_0, t) = \log((1-t)\exp\kappa_0 + t\exp\kappa_0) = \log(\exp\kappa_0) = \kappa_0, \tag{102}$$

which shows that $H(\kappa_0, t) = \kappa_0$ for all $t \in [0, 1]$, as required.

We conclude that $H(\kappa, t) \in \mathcal{K}$ for all $\kappa \in \mathcal{K}$ and $t \in [0, 1]$.

Thus, H is a strong deformation retraction from $\mathcal{K}$ onto κ_0, and thus $\mathcal{K}$ is contractible with respect to the compact-open topology. [*Back to the text*.] ∎

Next we prove three important structural properties of policies.

---

11 Lemma

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

1. Centeredness $(\mathfrak{o}_\pi)$: *we have*

$$\pi_{\tau,\kappa}(0) = 0;$$

2. Coerciveness $(\mathfrak{O}_\pi)$: *we have*

$$\lim_{m\to\infty} \|\pi_{\tau,\kappa}(m)\| = \infty; \text{ and}$$

3. Weak Monotonicity $\left(\widetilde{\mathfrak{M}}_\pi\right)$: *we have*

$$m_1 \leqslant m_2 \implies \pi_{\tau,\kappa}(m_1) \leqslant \pi_{\tau,\kappa}(m_2),$$

where the inequality on the right-hand side is taken component-wise. [*Proof*.]

Proof. We address each claim in turn.

Coerciveness. Fix any norm $\|\cdot\|$ on $\mathbb{R}^L$ and suppose, for sake of contradiction, that there exists some sequence $\{m_n\}_{n\in\mathbb{N}}$ such that $m_n \to \infty$ but that $\|\pi_{\tau,\kappa}(m_n)\| \leqslant R_L := (R,\ldots,R)$ with $0 < R < \infty$. The box $[0,R]^L \subsetneq \mathbb{R}^L_+$ is evidently a compact subset of $\mathbb{R}^L_+$. Since τ is continuous, the restriction $\tau|_B$ attains its maximum value $\overline{m}_R < \infty$. Hence, for all $n \in \mathbb{N}$, we have

$$m_n \leqslant \tau\left(\pi_{\tau,\kappa}(m_n)\right) \leqslant \overline{m}_R, \tag{103}$$

contradicting $m_n \to \infty$. We conclude that $\|\pi_{\tau,\kappa}(m)\| \to \infty$ whenever $m \to \infty$.

Weak monotonicity. Choose and fix $(\tau,\kappa) \in \mathcal{T} \times \mathcal{K}$ and define, for each $m \geqslant 0$,

$$\begin{aligned} F(m) &:= \left\{x \in \mathbb{R}^L_+ : \tau(x) \geqslant m\right\}, \\ &= \left\{x \in \mathbb{R}^L_+ : \log(1+\tau(x)) \geqslant \log(1+m)\right\}. \end{aligned}$$

We will use the notation $G = \log(1+\tau)$ for the remainder of this proof.

A few remarks on F are in order:

1. F is nonempty, because τ possesses $\mathfrak{R}_\tau$;
2. F is convex, because τ possesses $\widetilde{\mathfrak{Q}}_\tau$;
3. F is closed, because it is the superlevel set of the continuous function G;
4. F has the upper set property that
$$\left(x \in F(m) \text{ and } y \geqslant x\right) \Longrightarrow y \in F(m), \tag{104}$$
where this is also due to $\widetilde{\mathfrak{M}}_\tau$; and
5. F has the antitone property that
$$m_1 < m_2 \Longrightarrow F(m_1) \supset F(m_2), \tag{105}$$
where again this follows from $\widetilde{\mathfrak{M}}_\tau$.

From here, we proceed in three steps:

1. *For any $m \in M$, there exists a unique minimal element of $F(m)$.* Choose and fix any $m \in M$; since $F(m)$ is nonempty, we may also choose and fix some $x^0 \in F(m)$. Consider the order interval
$$[0,x^0] := \left\{x \in \mathbb{R}^L_+ : 0 \leqslant x \leqslant x^0\right\}. \tag{106}$$

This set is compact, so the intersection $F(m) \cap [0, x^0]$ is also compact (being the intersection of a closed set and a compact set). Finally, we know the intersection is nonempty, as it contains x^0. Define the linear functional

$$L(x) := \mathbf{1}^\top x. \tag{107}$$

By the Weierstrass Theorem, L attains a minimum over $F(m) \cap [0, x^0]$. Let $u(m)$ denote any such minimizer. For uniqueness, suppose for sake of contradiction that there exist $x_1, x_2 \in F(m)$ with $x_1 \neq x_2$ with both minimal. Since G is strictly concave, we have

$$G\left((1-\lambda)x_1 + \lambda x_2\right) > (1-\lambda)G(x_1) + \lambda G(x_2) \geqslant \log(1+m), \tag{108}$$

where the second part is because $x_1, x_2 \in F(m)$. We therefore have

$$z := (1-\lambda)x_1 + \lambda x_2 \in \operatorname{int} F(m).$$

Then there exists some $\varepsilon > 0$ small enough that we may define

$$w := z - \varepsilon \mathbf{1}, \tag{109}$$

such that $w \in F(m)$ and $w \leqslant \min\{x_1, x_2\}$. Contradiction; we conclude that the minimizer is unique.

2. *The optimizer equals the minimal element.* Recall that $\pi_{\tau,\kappa}(m)$ is defined as the unique minimizer of κ over the feasible set $F(m)$. Since κ possesses $\mathfrak{M}_\kappa$, it increases strictly in every coordinate. Hence, for any upper set $U \subseteq \mathbb{R}^L_+$ with minimal element u, we have

$$\kappa(x) > \kappa(u) \qquad \forall x \in U \setminus \{u\}. \tag{110}$$

Because $F(m)$ is an upper set with unique minimal element $u(m)$, this property implies

$$\pi_{\tau,\kappa}(m) = \operatorname*{argmin}_{x \in F(m)} \kappa(x) = u(m). \tag{111}$$

Thus, the optimizer of κ over $F(m)$ coincides with its minimal element.

3. *Monotonicity.* Let $m_1 < m_2$. By the antitone property of F, we have

$$F(m_2) \subset F(m_1). \tag{112}$$

Since $u(m_1)$ is the unique minimal element of $F(m_1)$ and $u(m_2) \in F(m_2) \subset F(m_1)$, it follows that

$$u(m_1) \leqslant u(m_2), \tag{113}$$

where the inequality is understood coordinatewise. Therefore,

$$\pi_{\tau,\kappa}(m_1) = u(m_1) \leqslant u(m_2) = \pi_{\tau,\kappa}(m_2), \tag{114}$$

establishing that $\pi_{\tau,\kappa}$ is coordinatewise nondecreasing in m.

We have thus established both claims. [*Back to the text*.] ∎

Finally, we may prove the main result of Section 2.

---

13 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. [*Proof*.]

Proof. We proceed in two steps. First, we give the intuitive argument that motivates the construction. Second, we use this fact to construct a canonical lift from the policy space $\mathcal{P}$ to the parameter space $\mathcal{T} \times \mathcal{K}$, which allows us to complete the homotopy.

Step 1: Intuition. By Lemma 12, there exist strong deformation retractions

$$\begin{aligned} H_\kappa : \mathcal{K} \times [0,1] &\longrightarrow \mathcal{K}, \\ (\kappa, t) &\longmapsto H_\kappa(\kappa, t), \\ H_\tau : \mathcal{T} \times [0,1] &\longrightarrow \mathcal{T}, \\ (\tau, t) &\longmapsto H_\tau(\tau, t), \end{aligned} \tag{115}$$

onto the functions

$$\kappa_0(x) = \sum_{\ell \in L} x_\ell, \qquad \text{and} \qquad \tau_0(x) = \sum_{\ell \in L} \log(1 + x_\ell). \tag{116}$$

For any $(\tau, \kappa) \in \mathcal{T} \times \mathcal{K}$, define the path $(\tau_t, \kappa_t) := (H_\tau(\tau, t), H_\kappa(\kappa, t))$, and let $\pi_t := \pi_{\tau_t, \kappa_t}$ denote the corresponding curried solutions of

$$\min_{x \geqslant 0} \kappa_t(x) \quad \text{s.t.} \quad \tau_t(x) \geqslant m. \tag{117}$$

Per Lemma 52, the solution map $S : (\tau, \kappa) \mapsto \pi_{\tau,\kappa}$ is continuous. Thus, $t \mapsto \pi_t$ is a continuous path in the policy space $\mathcal{P}$ from $\pi_0 = \pi_{\tau,\kappa}$ to $\pi_1 = \pi_{\tau_0,\kappa_0}$. This establishes path connectedness of $\mathcal{P}$. 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 policy π. This is the content of Step 2.

Step 2: Canonical lift. We seek a continuous choice of (τ_π, κ_π) for each policy $\pi \in \mathcal{P}$. The harder step will be on τ_π; once τ_π is defined, we can choose an appropriate κ_π that is constant in π and satisfies the necessary properties.

Our strategy is to define τ_π so that its hypograph is the set of all feasible (x, m) pairs for the policy π. Then, we will verify that τ_π possesses all the required properties to be in $\mathcal{T}$. Finally, we will show that τ_π varies continuously in π under the compact-open topology.

Given a policy $\pi \in \mathcal{P}$, we define the function

$$\tau_\pi(x) := \sup \{m \in M : \pi(m) \leqslant x\}.$$

We now verify that $\tau_\pi \in \mathcal{T}$:

1. Continuity $(\mathfrak{C}_\tau)$: Fix $\pi \in \mathcal{P}$ and $x^0 \in X$. Write $\tau = \tau_\pi$ and set $m^0 = \tau(x^0) = \max\{m \in M : \pi(m) \leqslant x^0\}$.

   Upper semicontinuity at x^0. Pick any $m^+ \in M$ with $m^+ > m^0$. Since m^0 is maximal, $\pi(m^+) \nleqslant x^0$, so there exists an index j with $\pi_j(m^+) > x_j^0$. Let $\delta_1 = \frac{1}{2}(\pi_j(m^+) - x_j^0) > 0$. If $\|x - x^0\| < \delta_1$ then $x_j < \pi_j(m^+)$, hence $\pi(m^+) \nleqslant x$ and therefore $\tau(x) < m^+$. Since $m^+ > m^0$ was arbitrary, this implies $\limsup_{x \to x^0} \tau(x) \leqslant m^0$.

   Concavity of $g = \log(1 + \tau)$. For each $m \in M$, the set $\{x : \tau(x) \geqslant m\} = \{x : \pi(m) \leqslant x\} = \pi(m) + \mathbb{R}_+^L$ is convex. Thus all superlevel sets of g are convex, so g is concave.

   Continuity on the interior. g is finite and concave on the convex set $X^\circ = (0, \infty)^L$, hence g is continuous on X°. Because $t \mapsto e^t - 1$ is continuous and strictly increasing, τ is continuous on X°.

Lower semicontinuity at x^0. If $x^0 \in X^\circ$, then τ is continuous at x^0 by the previous step, so $\liminf_{x \to x^0} \tau(x) \geqslant \tau(x^0) = m^0$.

If x^0 lies on the boundary of X, define $x^k = x^0 + k^{-1}\mathbf{1} \in X^\circ$ and $y^k = (x^0 - k^{-1}\mathbf{1})_+ \in X$ for $k \in \mathbb{N}$. Then $y^k \leqslant x^0 \leqslant x^k$, $x^k \downarrow x^0$, $y^k \uparrow x^0$, and by monotonicity of τ, $\tau(y^k) \leqslant \tau(x) \leqslant \tau(x^k)$ whenever $y^k \leqslant x \leqslant x^k$. By continuity of τ on X° and upper semicontinuity at x^0, we have $\tau(x^k) \downarrow m^0$ and $\tau(y^k) \uparrow m^0$. Hence, for any $\epsilon > 0$, there exists k_ϵ such that for all $k \geqslant k_\epsilon$ and all x with $y^k \leqslant x \leqslant x^k$, $m^0 - \epsilon \leqslant \tau(y^k) \leqslant \tau(x) \leqslant \tau(x^k) \leqslant m^0 + \epsilon$. This yields $\liminf_{x \to x^0} \tau(x) \geqslant m^0$.

Combining upper semicontinuity and lower semicontinuity, τ is continuous at x^0. Since x^0 was arbitrary, τ_π is continuous on X.

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

$$s \longmapsto \tau_\pi(sv) : \mathbb{R}_{\geqslant 0} \longrightarrow M \tag{118}$$

is continuous, strictly increasing, and unbounded.

Fix $\pi \in \mathcal{P}$ and define, for any $v \gg 0$ and $s \geqslant 0$,

$$s_m(v) := \max_{\ell \in L} \frac{\pi_\ell(m)}{v_\ell}. \tag{119}$$

By definition of τ_π, we have

$$\begin{aligned} \tau_\pi(sv) &= \sup\{\, m \in M : \pi(m) \leqslant sv \,\}, \\ &= \sup\{\, m \in M : s \geqslant s_m(v) \,\}. \end{aligned} \tag{120}$$

Hence the map $s \mapsto \tau_\pi(sv)$ is the (right-continuous) generalized inverse of $m \mapsto s_m(v)$.

Unboundedness. Since π is weakly increasing in m, $\pi(m) \to \infty$ componentwise as $m \to \infty$. For every fixed $m \in M$, we can choose $s \geqslant s_m(v)$ so that $sv \geqslant \pi(m)$, which implies $\tau_\pi(sv) \geqslant m$. Letting $m \to \infty$ yields $\tau_\pi(sv) \to \infty$ as $s \to \infty$.

Monotonicity. For $s_2 > s_1$, we have

$$\{\, m : s_1 \geqslant s_m(v) \,\} \subseteq \{\, m : s_2 \geqslant s_m(v) \,\}, \tag{121}$$

so $\tau_\pi(s_2 v) \geqslant \tau_\pi(s_1 v)$. Thus $s \mapsto \tau_\pi(sv)$ is weakly increasing.

Strict increase for generic rays. Fix rational $m_1 < m_2$ and define

$$E_{m_1,m_2} := \{\, v \gg 0 : s_{m_2}(v) = s_{m_1}(v) \,\}. \tag{122}$$

If $\pi(m_2) = \pi(m_1)$, the set E_{m_1,m_2} is empty. Otherwise, equality requires

$$\max_{\ell} \frac{\pi_\ell(m_2)}{v_\ell} = \max_{k} \frac{\pi_k(m_1)}{v_k}, \tag{123}$$

which defines a finite union of smooth hypersurfaces of codimension one in the positive cone $\{v \gg 0\}$. Each E_{m_1,m_2} is therefore closed and nowhere dense. Define the residual set

$$\mathcal{V} := \{\, v \gg 0 : s_{m_2}(v) > s_{m_1}(v) \text{ for all rationals } m_1 < m_2 \,\}. \tag{124}$$

For any $v \in \mathcal{V}$, the map $m \mapsto s_m(v)$ is strictly increasing on $\mathbb{R}_+$, and hence its inverse $s \mapsto \tau_\pi(sv)$ is strictly increasing.

Continuity. From $(\mathfrak{C}_\tau)$, τ_π is continuous on X. Thus $s \mapsto \tau_\pi(sv)$ is continuous for each v.

Combining these properties, we find that for any generic $v \in \mathcal{V}$, the map $s \mapsto \tau_\pi(sv)$ is continuous, strictly increasing, and unbounded. Hence τ_π satisfies $(\mathfrak{R}_\tau)$.

3. Weak Monotonicity $(\widetilde{\mathfrak{M}}_\tau)$: Fix $x, y \in X$ with $x \leqslant y$ coordinatewise. By definition, $\tau_\pi(x) = \sup\{m \in M : \pi(m) \leqslant x\}$ and $\tau_\pi(y) = \sup\{m \in M : \pi(m) \leqslant y\}$. Since $x \leqslant y$, we have $\{m : \pi(m) \leqslant x\} \subseteq \{m : \pi(m) \leqslant y\}$, hence $\tau_\pi(x) \leqslant \tau_\pi(y)$. Therefore τ_π is weakly increasing in each coordinate.

4. Log-Concavity $(\widetilde{\mathfrak{L}}_\tau)$: Let $g(x) = \log(1 + \tau_\pi(x)) = \log\big(1 + \sup\{m \in M : \pi(m) \leqslant x\}\big)$. To show g is concave, it suffices to verify that for all $m_1, m_2 \in M$ and $\lambda \in (0,1)$,

$$\lambda U_\pi(m_1) + (1-\lambda) U_\pi(m_2) \subseteq U_\pi(m_\lambda), \tag{125}$$

where $U_\pi(m) = \{x : \tau_\pi(x) \geqslant m\} = \{x : x \geqslant \pi(m)\}$ and m_λ satisfies

$$\log(1 + m_\lambda) = \lambda \log(1 + m_1) + (1-\lambda) \log(1 + m_2), \tag{126}$$

that is,

$$m_\lambda = (1 + m_1)^\lambda (1 + m_2)^{1-\lambda} - 1. \tag{127}$$

We compute

$$\lambda U_\pi(m_1) + (1-\lambda)U_\pi(m_2) = \lambda\pi(m_1) + (1-\lambda)\pi(m_2) + \mathbb{R}^L_+. \tag{128}$$

Hence the desired inclusion is equivalent to

$$\lambda\pi(m_1) + (1-\lambda)\pi(m_2) \geqslant \pi(m_\lambda) \quad \text{(coordinatewise)}. \tag{129}$$

Since $\pi \in \mathcal{P}$, there exist $\tau \in \mathcal{T}$ and $\kappa \in \mathcal{K}$ such that $\pi = \pi_{\tau,\kappa}$, with $\log(1+\tau)$ concave and κ strictly increasing. Let $x_i = \pi(m_i)$ for $i = 1, 2$. Concavity of $\log(1+\tau)$ gives

$$\begin{aligned} \log\big(1 + \tau(\lambda x_1 + (1-\lambda)x_2)\big) \geqslant{}& \lambda \log(1+\tau(x_1)) \\ &+ (1-\lambda)\log(1+\tau(x_2)) \\ ={}& \lambda\log(1+m_1) \\ &+ (1-\lambda)\log(1+m_2), \end{aligned} \tag{130}$$

so $\tau(\lambda x_1 + (1-\lambda)x_2) \geqslant m_\lambda$. Thus $\lambda x_1 + (1-\lambda)x_2$ is feasible at level m_λ. Because κ is strictly increasing and $\pi(m_\lambda)$ is the unique κ-minimizer among points x with $\tau(x) \geqslant m_\lambda$, we obtain

$$\pi(m_\lambda) \leqslant \lambda x_1 + (1-\lambda)x_2 = \lambda\pi(m_1) + (1-\lambda)\pi(m_2), \tag{131}$$

where the inequality is coordinatewise. This proves the inclusion above, and hence $x \mapsto \log(1+\tau_\pi(x))$ is concave.

We therefore have $\tau_\pi \in \mathcal{T}$ for all $\pi \in \mathcal{P}$.

Now we verify that the map $\pi \mapsto \tau_\pi$ is continuous from $\mathcal{P}$ to $\mathcal{T}$ under the compact-open topologies. Fix any $\pi_0 \in \mathcal{P}$ and let $\tau_0 = \tau_{\pi_0}$. Let $K \subseteq X$ be compact and $\epsilon > 0$. We must find $\delta > 0$ such that $d_{\mathcal{P}}(\pi, \pi_0) < \delta$ implies

$$\max_{x\in K} |\tau_\pi(x) - \tau_0(x)| < \epsilon. \tag{132}$$

Set $\overline{m}_K = \max_{x\in K} \tau_0(x)$ and choose $m_K \in M$ with $m_K > \overline{m}_K + 1$. For $(m, x) \in [0, m_K] \times K$ define

$$f(m,x) = \min_{1\leqslant i\leqslant L} \big(x_i - \pi_{0,i}(m)\big). \tag{133}$$

Note that f is continuous, nonincreasing in m for each fixed x, and

$$\tau_0(x) = \sup\{m \in [0, m_K] : f(m,x) \geqslant 0\}. \tag{134}$$

Let $\pi \in \mathcal{P}$ and define $f_\pi(m, x) = \min_i \big(x_i - \pi_i(m)\big)$. Note that if

$$\sup_{m \in [0, m_K]} \|\pi(m) - \pi_0(m)\| < \eta, \tag{135}$$

then

$$\sup_{(m,x) \in [0, m_K] \times K} |f_\pi(m, x) - f(m, x)| \leq \eta. \tag{136}$$

Upper bound. For each $x \in K$ set $m^+(x) = \tau_0(x) + \epsilon$. Then $f(m^+(x), x) < 0$ because $m^+(x) > \tau_0(x)$. By continuity of f there exist $\eta_x > 0$ and an open neighborhood $V_x \subset K$ such that

$$f(m^+(x), x') \leq -2\eta_x \quad \text{for all } x' \in V_x. \tag{137}$$

If $\eta \leq \eta_x$, then for all $x' \in V_x$,

$$f_\pi(m^+(x), x') \leq -\eta_x < 0, \tag{138}$$

hence $\tau_\pi(x') < m^+(x) = \tau_0(x) + \epsilon$.

Lower bound. For each $x \in K$ either $f(\tau_0(x), x) > 0$, in which case set $m^-(x) = \tau_0(x)$, or $f(\tau_0(x), x) = 0$, in which case by continuity in m choose $m^-(x) \in [0, \tau_0(x))$ with

$$f(m^-(x), x) \geq 2\eta'_x \tag{139}$$

for some $\eta'_x > 0$. By continuity in x, shrinking if necessary, there is an open neighborhood $W_x \subset K$ of x such that

$$f(m^-(x), x') \geq 2\eta'_x \quad \text{for all } x' \in W_x. \tag{140}$$

If $\eta \leq \eta'_x$, then for all $x' \in W_x$,

$$f_\pi(m^-(x), x') \geq \eta'_x > 0, \tag{141}$$

so $\tau_\pi(x') \geq m^-(x) \geq \tau_0(x) - \epsilon$.

By compactness of K, select $x^1, \ldots, x^p \in K$ so that $U_\ell := V_{x^\ell} \cap W_{x^\ell}$ cover K. Set

$$\overline{\eta} = \min_{1 \leq \ell \leq p} \{\eta_{x^\ell}, \eta'_{x^\ell}\} > 0. \tag{142}$$

Choose $\delta > 0$ so that $d_{\mathcal{P}}(\pi, \pi_0) < \delta$ implies

$$\sup_{m \in [0, m_K]} \|\pi(m) - \pi_0(m)\| < \overline{\eta}. \tag{143}$$

Then for any $x \in K$ there exists ℓ with $x \in U_\ell$, and we have

$$\tau_0(x^\ell) - \epsilon \leq \tau_\pi(x) \leq \tau_0(x^\ell) + \epsilon. \tag{144}$$

Finally, continuity of τ_0 on K allows us (shrinking U_ℓ if needed) to ensure

$$|\tau_0(x) - \tau_0(x^\ell)| < \epsilon \quad \text{whenever } x \in U_\ell, \tag{145}$$

so that for all $x \in K$,

$$|\tau_\pi(x) - \tau_0(x)| \leq |\tau_\pi(x) - \tau_0(x^\ell)| + |\tau_0(x^\ell) - \tau_0(x)| < 2\epsilon. \tag{146}$$

As $\epsilon > 0$ was arbitrary, the claim follows. We therefore have shown that the map $\pi \mapsto \tau_\pi$ is continuous from $\mathcal{P}$ to $\mathcal{T}$. To complete the lift, we need to choose κ_π for each π.

Defining κ_π. Fix $\alpha \in \mathbb{R}^L_{++}$ and $\varepsilon > 0$ and define

$$\kappa_\pi(x) = \alpha \cdot x + \varepsilon \|x\|^2, \quad x \in X. \tag{147}$$

$\kappa_\pi \in \mathcal{K}$: Continuity and $\kappa_\pi(0) = 0$ are immediate. Coerciveness holds since $\|x\| \to \infty$ implies $\kappa_\pi(x) \to \infty$. Strict coordinatewise monotonicity holds because $\alpha_i > 0$ for all i and the quadratic term is nondecreasing in each coordinate when others are fixed. Strict convexity of κ_π is standard, hence $x \mapsto \exp(\kappa_\pi(x))$ is strictly convex.

Compatibility with τ_π: For any $m \in M$, the feasible set for level m is

$$\{x \in X : \tau_\pi(x) \geq m\} = \{x \in X : x \geq \pi(m)\} = \pi(m) + \mathbb{R}^L_+. \tag{148}$$

On such an upper set, any strictly increasing function in each coordinate achieves its unique minimum at the minimal element. Therefore,

$$\operatorname{argmin}\{\kappa_\pi(x) : \tau_\pi(x) \geq m\} = \{\pi(m)\}. \tag{149}$$

Hence, with this choice of κ_π (which does not depend on π), we have $\pi = \pi_{\tau_\pi, \kappa_\pi}$ for every $\pi \in \mathcal{P}$.

Completing the homotopy. We may now complete the homotopy on $\mathcal{P}$. For any $\pi \in \mathcal{P}$, define the path $(\tau_t, \kappa_t) = (H_\tau(\tau_\pi, t), H_\kappa(\kappa_\pi, t))$ for $t \in [0, 1]$. Let $\pi_t = \pi_{\tau_t,\kappa_t}$ denote the corresponding curried solutions. By continuity of the solution map $S : (\tau, \kappa) \mapsto \pi_{\tau,\kappa}$, the map $t \mapsto \pi_t$ is a continuous path in $\mathcal{P}$ from $\pi_0 = \pi_{\tau_\pi,\kappa_\pi} = \pi$ to $\pi_1 = \pi_{\tau_0,\kappa_0}$. Since this construction works for any $\pi \in \mathcal{P}$, we have established a homotopy from the identity map on $\mathcal{P}$ to the constant map with value π_{τ_0,κ_0}. This shows that $\mathcal{P}$ is contractible.

Identifying the basepoint. In the name of completeness, let us identify π_{τ_0,κ_0}. Recall that τ_0 and κ_0 are defined as

$$\begin{aligned} \tau_0(x) &= \sum_{\ell \in L} \log(1 + x_\ell), \\ \kappa_0(x) &= \sum_{\ell \in L} x_\ell. \end{aligned} \tag{150}$$

The solution to Problem (SPP (m, τ, κ)) is characterized by the first-order condition

$$1 - \frac{\lambda}{1 + x_\ell} = 0 \quad \text{for all } \ell \in L, \qquad m - \sum_{\ell \in L} \log(1 + x_\ell) = 0, \tag{151}$$

where λ is the Lagrange multiplier associated with the production constraint. Clearly, the first condition implies that $x_\ell = x^* = \lambda - 1$ for all $\ell \in L$. Plugging this into the second condition, we have

$$m - L \log\left(1 + x^*\right) = 0, \tag{152}$$

which implies that

$$x^* = \exp\left(\frac{m}{L}\right) - 1. \tag{153}$$

Thus, the curried solution function π_{τ_0,κ_0} is given by

$$\pi_{\tau_0,\kappa_0}(m) = \left(\exp\left(\frac{m}{L}\right) - 1\right)\mathbf{1}. \tag{154}$$

We conclude that for any $(\tau, \kappa) \in \mathcal{T} \times \mathcal{K}$, there exists a homotopy from $\pi_{\tau,\kappa}$ to the function π_{τ_0,κ_0} defined above. [*Back to the text*.] ■

A.5 For Section 3

First, we handle the existence proof for Nash equilibrium in the disaggregated game.

---

24 Proposition

Game 23 has at least one pure-strategy Nash equilibrium. [*Proof*.]

Proof. The main sticking point here is the discontinuity of the contest success function when both States choose zero investment; this precludes us from appealing to standard existence results that require continuity of the payoff functions. We will appeal to a well-known result due to Reny (1999), which states that a game has a pure-strategy Nash equilibrium if:

1. each strategy set X_i is a nonempty, compact, and convex subset of a topological vector space;
2. each payoff function is bounded and quasiconcave in its owner's inputs; and
3. the game satisfies a condition called *better-reply secureness*, to be defined at the relevant part of the proof.

We address each condition in turn.

Strategy sets. Since each κ_i is coercive $\left(\mathfrak{O}_{\kappa_i}\right)$, each State's cost goes to infinity as the norm of their investment vector goes to infinity. So, there exists some compact box

$$\overline{X} = \prod_{\ell \in L} [0, \overline{x}_\ell]$$

such that any equilibrium must live in $\overline{X}$. We focus on this box without loss of any generality, and we observe that it is nonempty, compact, and convex.

Payoff functions. We address boundedness and quasiconcavity separately.

Boundedness. On the compact box $\overline{X}$, each τ_i is bounded and each e^{κ_i} is bounded below by 1. Define

$$p_i(x) = \frac{\lambda_i \, \tau_i(x_i)^\alpha}{\lambda_1 \, \tau_1(x_1)^\alpha + \lambda_2 \, \tau_2(x_2)^\alpha}. \tag{155}$$

Then $p_i(x) \in [0,1]$, and for every $x \in \overline{X}^2$,

$$0 \le U_i(x) = p_i(x)\left(V - k\left(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\right)\right) \\ \le V. \tag{156}$$

Thus U_i is bounded on $\overline{X}^2$.

Quasi-concavity. Fix the opponent's action $x_{-i} \in \overline{X}$ and consider $x_i \mapsto U_i(x_i, x_{-i})$. Write

$$\log U_i = \log p_i + \log A, \tag{157}$$

where

$$A(x) = V - k\left(e^{\kappa_1(x_1)} + e^{\kappa_2(x_2)} - 2\right). \tag{158}$$

(i) Concavity of $\log p_i$ *in* x_i. Let $w_i(x_i) = \log \lambda_i + \alpha \log \tau_i(x_i)$. Since $g_i = \log(1 + \tau_i)$ is concave, one obtains on $\{\tau_i > 0\}$ the matrix inequality

$$\nabla^2 \log \tau_i \preceq -\frac{1}{\tau_i^2(1+\tau_i)} \nabla\tau_i \nabla\tau_i^\top \preceq 0, \tag{159}$$

so w_i is concave. With the opponent fixed, set $C = e^{w_{-i}(x_{-i})} > 0$ and note that

$$\log p_i = w_i - \log(e^{w_i} + C). \tag{160}$$

Since $w \mapsto w - \log(e^w + C)$ is increasing and strictly concave, $\log p_i(\cdot, x_{-i})$ is concave on $\{\tau_i > 0\}$. Along $\{\tau_i = 0\}$ the upper contour sets extend by closure (axis values are 0 against a positive opponent and are set by convention on the joint zero set), so $\log p_i$ has convex upper contours on all of $\overline{X}$.

(ii) Concavity of $\log A$ *in* x_i. Let $B_i = e^{\kappa_i}$, which is strictly convex by assumption. Then, for fixed x_{-i},

$$\nabla^2_{x_i} \log A = -\frac{k}{A} \nabla^2 B_i - \frac{k^2}{A^2} \nabla B_i \nabla B_i^\top \preceq 0, \tag{161}$$

so $\log A(\cdot, x_{-i})$ is concave on $\overline{X}$.

(iii) Quasi-concavity of U_i. On $\{\tau_i > 0\}$, $\log U_i = \log p_i + \log A$ is a sum of concave functions, hence concave; therefore U_i is log-concave and thus quasi-concave there. By the boundary argument in (i), the upper contour sets extend by closure to all of $\overline{X}$ and remain convex. Therefore $U_i(\cdot, x_{-i})$ is quasi-concave on $\overline{X}$.

Better-reply secureness. We begin by recalling the definition of *better-reply secureness.* Let $G = (X_i, U_i)_{i\in\{1,2\}}$ denote the game, with $X = X_1 \times X_2$ and $U = (U_1, U_2)$. A game is said to be *better-reply secure* if, for every pair $(x, u) \in X \times \mathbb{R}^2$ satisfying

$$U(x^n) \longrightarrow u \quad \text{and} \quad x^n \longrightarrow x, \tag{162}$$

where $U_i(x^n) \le u_i$ for all n and each state i, there exists at least one state i and an action $\widehat{x}_i \in X_i$ such that

$$\begin{aligned} &U_i(\widehat{x}_i, x_{-i}) > u_i, \\ &U_i(\widehat{x}_i, y_{-i}) \ge u_i \quad \text{for all } y_{-i} \text{ near } x_{-i}. \end{aligned} \tag{163}$$

Intuitively, even at any limit point of a sequence of approximate play, at least one state can profitably and *securely* deviate—that is, choose a nearby action that guarantees a payoff strictly exceeding the limit level u_i against all sufficiently small perturbations of the opponent's action.

We now verify that the present game is better-reply secure. Let $S_i = \{x_i \in \overline{X} : \tau_i(x_i) = 0\}$ denote the zero-technology set, and define

$$D = S_1 \times S_2. \tag{164}$$

The payoff functions U_i are continuous on $\overline{X}^2 \setminus D$, so only points in D require attention.

Fix $(x_1^*, x_2^*) \in D$, a state i, and a small $\epsilon \in (0, 1)$. We will construct a deviation $\widehat{x}_i$ and a neighborhood V_{-i} of x_{-i}^* such that

$$U_i(\widehat{x}_i, y_{-i}) \ge (1-\epsilon)\left(V - \frac{\epsilon V}{2}\right) \quad \text{for all } y_{-i} \in V_{-i}. \tag{165}$$

Step 1: Small capability, small cost. Because τ_i is onto and continuous, we may choose $\widehat{x}_i \in \overline{X}$ with $\tau_i(\widehat{x}_i) = \overline{t} > 0$, as small as desired. Because $e^{\kappa_i(\widehat{x}_i)} \downarrow 1$ as $\widehat{x}_i \to S_i$ and A is continuous on $\overline{X}^2$, we may shrink $\overline{t}$ such that

$$k\left(e^{\kappa_i(\widehat{x}_i)} - 1\right) \le \frac{\epsilon V}{4}. \tag{166}$$

Step 2: Controlling the opponent near x_{-i}^*. Set

$$\delta_\tau = \left(\frac{\epsilon}{1-\epsilon}\lambda_i\right)^{1/\alpha} \overline{t}, \tag{167}$$

where $\lambda_1 = \lambda$ and $\lambda_2 = 1$. Since $\tau_{-i}(x^*_{-i}) = 0$ and τ_{-i} is continuous, there exists a neighborhood $V^{(1)}_{-i}$ of x^*_{-i} such that

$$\tau_{-i}(y_{-i}) \leq \delta_\tau \quad \text{for all } y_{-i} \in V^{(1)}_{-i}. \tag{168}$$

Then

$$p_i(\widehat{x}_i, y_{-i}) = \frac{\lambda_i \bar{t}^\alpha}{\lambda_i \bar{t}^\alpha + \tau_{-i}(y_{-i})^\alpha} \geq 1 - \epsilon. \tag{169}$$

Because $e^{\kappa_{-i}(x^*_{-i})} = 1$ and $e^{\kappa_{-i}}$ is continuous, there exists a neighborhood $V^{(2)}_{-i}$ of x^*_{-i} such that

$$k\left(e^{\kappa_{-i}(y_{-i})} - 1\right) \leq \frac{\epsilon V}{4} \quad \text{for all } y_{-i} \in V^{(2)}_{-i}. \tag{170}$$

Let $V_{-i} = V^{(1)}_{-i} \cap V^{(2)}_{-i}$.

Step 3: Uniform security bound. For every $y_{-i} \in V_{-i}$,

$$\begin{aligned} U_i(\widehat{x}_i, y_{-i}) &= p_i(\widehat{x}_i, y_{-i})\left[V - k\left(e^{\kappa_i(\widehat{x}_i)} + e^{\kappa_{-i}(y_{-i})} - 2\right)\right] \\ &\geq (1-\epsilon)\left[V - k\left((e^{\kappa_i(\widehat{x}_i)} - 1) + (e^{\kappa_{-i}(y_{-i})} - 1)\right)\right] \\ &\geq (1-\epsilon)\left(V - \frac{\epsilon V}{2}\right). \end{aligned} \tag{171}$$

Since $U_i \leq V$ everywhere, any sequence approaching (x^*_1, x^*_2) has lim sup $\leq V$. Hence State i can secure within $\frac{3}{2}\epsilon V$ of the maximal limit payoff uniformly over V_{-i}. Therefore, the game is better-reply secure at (x^*_1, x^*_2).

Conclusion. Because $(x^*_1, x^*_2) \in D$ and i were arbitrary, the game is better-reply secure. By Reny (1999), the existence of a pure-strategy Nash equilibrium follows. [*Back to the text*.] ∎

Next, we introduce the operators $\mathfrak{D}$ that regularize technologies and costs to be differentiable.

---

53 Definition

The regularization operator *for technologies is a map*

$$\mathfrak{D} : \mathcal{T} \longrightarrow \mathcal{T}.$$

For each $\tau \in \mathcal{T}$, the function $\mathfrak{D}(\tau)$ is defined in three steps:

1. Curvature chart. *Define*

$$\phi(x) = \log(1 + \tau(x)),$$

which, by the defining properties of $\mathcal{T}$*, is continuous, concave, and weakly increasing in every coordinate.*

2. Causal convolution smoothing. *Extend* ϕ *to a concave, coordinatewise nondecreasing function on all of* $\mathbb{R}^L$ *by*

$$\tilde{\phi}(x) = \inf_{y \in \mathbb{R}^L_+} \inf_{g \in \partial\phi(y)} \{\phi(y) + \langle g, x - y\rangle\},$$

where $\partial\phi(y)$ *is the superdifferential of* ϕ *at* y.[19] *For each* $\varepsilon > 0$*, choose a smooth nonnegative kernel* η_ε *supported in* $[0, c\varepsilon]^L$ *satisfying* $\int_{\mathbb{R}^L} \eta_\varepsilon(u)\, du = 1$ *and define the causal convolution*

$$\tilde{\phi}_\varepsilon(x) = \int_{\mathbb{R}^L} \tilde{\phi}(x - u)\, \eta_\varepsilon(u)\, du,$$

with the convention $\tilde{\phi}_0 = \tilde{\phi}$*. The restriction of* $\tilde{\phi}_\varepsilon$ *to* $X = \mathbb{R}^L_+$ *is*

$$\phi_\varepsilon = \tilde{\phi}_\varepsilon|_X,$$

which is smooth, concave, and coordinatewise nondecreasing on X*.*

3. Gauge and return to the original chart. *Fix a nonempty compact set* $K \subset X$ *with nonempty interior (for example,* $K = [0, 2]^L$*) and a reference point* $x^* \in \mathrm{int}(K)$ *(for example,* $x^* = (e - 1)\mathbf{1}$*). For* $\phi = \log(1 + \tau)$*, define the normalized tame template*

$$T_\beta(x) := A \frac{\sum_{\ell \in L} \beta_\ell \log(1 + x_\ell)}{\sum_{\ell \in L} \beta_\ell \log(1 + x^*_\ell)}, \qquad A > 0, \beta \in \Delta_L.$$

Set the projection residual

$$R_{\mathcal{T}}(\tau) := \inf_{A > 0, \beta \in \Delta_L} \sup_{x \in K} \left|\phi(x) - \phi(x^*)\, T_\beta(x)\right|.$$

[19] Because ϕ is concave and finite on the open convex set $\mathbb{R}^L_{++}$, the superdifferential $\partial\phi(y)$ is nonempty for every $y \in \mathbb{R}^L_{++}$. As for the boundary points $y \in \partial\mathbb{R}^L_+$, we may take limits of supergradients at interior points approaching y to see that $\partial\phi(y)$ is nonempty there as well.

Choose a continuous strictly increasing gauge $\Theta : [0, \infty) \to [0, \overline{\varepsilon}]$ *with* $\Theta(0) = 0$, *and define*

$$\varepsilon(\tau) \coloneqq \Theta\big(R_{\mathcal{T}}(\tau)\big).$$

The regularized technology is then

$$\mathfrak{D}(\tau)(x) = \exp\left(\phi_{\varepsilon(\tau)}(x)\right) - 1.$$

We can now show that the operator works as intended.

---

54 Lemma

Fix the operator $\mathfrak{D} : \mathcal{T} \to \mathcal{T}$ *as in Definition 53. Then for every* $\tau \in \mathcal{T}$*:*

1. Class preservation. $\mathfrak{D}(\tau) \in \mathcal{T}$. *Equivalently,*
   (a) $\mathfrak{D}(\tau)$ *is continuous on* X*;*
   (b) *there exists* $v \in X$ *such that* $t \mapsto \mathfrak{D}(\tau)(tv)$ *is continuous, strictly increasing, and unbounded on* $[0, \infty)$*;*
   (c) $\mathfrak{D}(\tau)$ *is weakly increasing in every coordinate; and*
   (d) $\log(1 + \mathfrak{D}(\tau))$ *is concave on* X.
2. Smoothness. $\mathfrak{D}(\tau) \in C^{\infty}(X)$.
3. Fixed points on tame technologies. *If* $\tau \in \mathcal{T}^{[\mathfrak{T}]}$, *then* $\mathfrak{D}(\tau) = \tau$.
4. Continuity in τ. *The map* $\mathfrak{D} : (\mathcal{T}, d) \to (\mathcal{T}, d)$ *is continuous with respect to the compact-open metric* d.

Proof. Write $\phi = \log(1 + \tau)$ on X. Let $\tilde{\phi}$ be the monotone-concave extension to $\mathbb{R}^L$, and let

$$\phi_\varepsilon = (\tilde{\phi} * \eta_\varepsilon)|_X \tag{172}$$

as in Definition 53. Set $\widehat{\tau} = \exp(\phi_{\varepsilon(\tau)}) - 1$.

Remark on compact bounds on A. On the compact set $\Delta_L \times K$, the map $(\beta, x) \mapsto T_\beta(x)$ is continuous and strictly positive. Hence

$$c_{\min} := \min_{(\beta,x)\in\Delta_L\times K} T_\beta(x) > 0, \qquad c_{\max} := \max_{(\beta,x)\in\Delta_L\times K} T_\beta(x) < \infty. \tag{173}$$

Let

$$m_\phi = \min_{x\in K} \phi(x), \qquad M_\phi = \max_{x\in K} \phi(x). \tag{174}$$

Any minimizer (A^*, β^*) in the definition of $R_{\mathcal{T}}(\tau)$ satisfies

$$\frac{m_\phi}{c_{\max}} \le A^* \le \frac{M_\phi}{c_{\min}}. \tag{175}$$

Consequently, the search over A may be restricted to the compact interval

$$\mathcal{A}_\phi = \left[\, m_\phi / c_{\max},\ M_\phi / c_{\min} \,\right]. \tag{176}$$

Proof sketch. If $A < m_\phi/c_{\max}$, then for all $x \in K$, $A\, T_\beta(x) \le A\, c_{\max} < m_\phi \le \phi(x)$, so the sup error at the point achieving m_ϕ exceeds $m_\phi - A\, c_{\max}$ and can be strictly reduced by increasing A. If $A > M_\phi/c_{\min}$, then for all $x \in K$, $A\, T_\beta(x) \ge A\, c_{\min} > M_\phi \ge \phi(x)$, so the sup error at the point achieving M_ϕ exceeds $A\, c_{\min} - M_\phi$ and can be strictly reduced by decreasing A.

Class preservation. We address each axiom in turn.

1. *Continuity.* The extension $\tilde{\phi}$ is concave on $\mathbb{R}^L$, hence continuous on the interior and upper semicontinuous everywhere. Convolution with a smooth kernel η_ε produces a C^∞ function $\tilde{\phi}_\varepsilon = \tilde{\phi} * \eta_\varepsilon$ on $\mathbb{R}^L$. Restricting to X preserves continuity, and composition with $x \mapsto \exp(x) - 1$ preserves continuity. Therefore $\hat{\tau}$ is continuous.

2. *Ray surjectivity.* By the defining axiom for $\mathcal{T}$, there exists $v \in X$ such that $t \mapsto \tau(tv)$ is continuous, strictly increasing, and unbounded on $[0, \infty)$. Then

$$\phi(tv) = \log(1 + \tau(tv)) \tag{177}$$

is continuous, nondecreasing, and unbounded. The extension satisfies $\tilde{\phi}(x) = \phi(x)$ for $x \in X$, so along the ray $t \mapsto tv$ we have $\tilde{\phi}(tv) = \phi(tv)$. The convolution

$$t \mapsto \tilde{\phi}_\varepsilon(tv) \tag{178}$$

is a positive average of translates of $\tilde{\phi}$ along the same ray; it inherits concavity and nondecreasingness on $[0, \infty)$. Because the kernel has bounded support, a bounded-window average of an unbounded function remains unbounded. A concave, nondecreasing, unbounded function on $[0, \infty)$ is strictly increasing. Therefore $t \mapsto \phi_{\varepsilon(\tau)}(tv)$ is continuous, strictly increasing, and unbounded, and the same holds for

$$t \mapsto \hat{\tau}(tv) = \exp(\phi_{\varepsilon(\tau)}(tv)) - 1. \tag{179}$$

3. *Weak monotonicity.* By construction, $\tilde{\phi}$ is coordinatewise nondecreasing on $\mathbb{R}^L$: it is the infimum of affine majorants whose slopes $g \in \partial\phi(y)$ satisfy $g \geq 0$ componentwise. Let $x, y \in \mathbb{R}^L$ with $x \leq y$ coordinatewise, and let u lie in the support of η_ε, which is contained in the positive orthant. Then $x - u \leq y - u$, so $\tilde{\phi}(x - u) \leq \tilde{\phi}(y - u)$. Integrating against η_ε gives $\tilde{\phi}_\varepsilon(x) \leq \tilde{\phi}_\varepsilon(y)$. Restricting to X and composing with the increasing function $x \mapsto \exp(x) - 1$ yields that $\hat{\tau}$ is weakly increasing in each coordinate.

4. *Concavity of* $\log(1 + \hat{\tau})$. Translates of a concave function are concave, and positive averages of concave functions are concave. Therefore $\tilde{\phi}_\varepsilon$ is concave on $\mathbb{R}^L$, and ϕ_ε is concave on X. By definition,

$$\log(1 + \hat{\tau}) = \phi_{\varepsilon(\tau)}, \tag{180}$$

so $\log(1 + \hat{\tau})$ is concave.

Smoothness. The convolution $\tilde{\phi}_\varepsilon = \tilde{\phi} * \eta_\varepsilon$ is C^∞ on $\mathbb{R}^L$ because η_ε is a smooth kernel. Hence $\phi_\varepsilon = \tilde{\phi}_\varepsilon|_X$ is C^∞ on X. Composition with $x \mapsto \exp(x) - 1$ preserves smoothness, so $\hat{\tau}$ is C^∞ on X, and therefore in particular $\hat{\tau} \in C^{1,1}(X)$.

Fixed points on tame technologies. If $\tau \in \mathcal{T}^{[\mathfrak{T}]}$, then by construction $R_{\mathcal{T}}(\tau) = 0$ and the gauge satisfies $\varepsilon(\tau) = \Theta(0) = 0$. By the convention in Definition 53, $\tilde{\phi}_0 = \tilde{\phi}$ and $\phi_0 = \tilde{\phi}|_X = \phi$ on X. Therefore

$$\begin{aligned} \hat{\tau}(x) &= \exp(\phi_{\varepsilon(\tau)}(x)) - 1 \\ &= \exp(\phi(x)) - 1 \\ &= \tau(x), \end{aligned} \tag{181}$$

so $\mathfrak{D}$ fixes every tame technology.

Continuity in τ. Let (τ_n) converge to τ in the compact-open metric d. Then $\phi_n = \log(1+\tau_n)$ converges to $\phi = \log(1+\tau)$ uniformly on every compact subset of X. The monotone-concave extension operator $\phi \mapsto \tilde{\phi}$ is continuous for uniform convergence on compacts.[20] Hence $\tilde{\phi}_n \to \tilde{\phi}$ uniformly on compacts in $\mathbb{R}^L$. The convolution map $(f, \varepsilon) \mapsto f * \eta_\varepsilon$ is continuous for uniform convergence on compacts and continuous in ε. The gauge $\varepsilon(\cdot) = \Theta\big(R_{\mathcal{T}}(\cdot)\big)$ is continuous on $(\mathcal{T}, d)$. Therefore, we have uniform convergence on compact subsets of $\mathbb{R}^L$:

$$\tilde{\phi}_n * \eta_{\varepsilon(\tau_n)} \longrightarrow \tilde{\phi} * \eta_{\varepsilon(\tau)}. \tag{182}$$

Restricting to X gives uniform convergence on compact subsets of X:

$$\phi_{n,\varepsilon(\tau_n)} \longrightarrow \phi_{\varepsilon(\tau)}. \tag{183}$$

Finally, applying the smooth chart map $h \mapsto \exp(h) - 1$ yields

$$\mathfrak{D}(\tau_n)(x) = \exp(\phi_{n,\varepsilon(\tau_n)}(x)) - 1 \longrightarrow \exp(\phi_{\varepsilon(\tau)}(x)) - 1 = \mathfrak{D}(\tau)(x), \tag{184}$$

uniformly on compact subsets of X. Hence $\mathfrak{D} : (\mathcal{T}, d) \to (\mathcal{T}, d)$ is continuous.

Conclusion. Having addressed all four points, the proof is complete. ■

---

55 Definition

The regularization operator *for costs is a map*

$$\mathfrak{D} : \mathcal{K} \longrightarrow \mathcal{K}.$$

For each $\kappa \in \mathcal{K}$, the function $\mathfrak{D}(\kappa)$ is defined in three steps:

1. Convexity chart. *Define*

$$\psi(x) = \exp(\kappa(x)),$$

which, by the defining properties of $\mathcal{K}$, is continuous, strictly convex, and strictly increasing in every coordinate.

[20] For a concave, coordinatewise nondecreasing ϕ, every compact $K \subset \mathrm{int}(X)$ admits $M < \infty$ such that $\|g\| \le M$ for all $g \in \partial\phi(y)$ and $y \in K$. Hence the family of supporting hyperplanes $\{\phi(y) + \langle g, x - y\rangle\}$ is equicontinuous on compacts. Uniform convergence $\phi_n \to \phi$ then implies $\tilde{\phi}_n \to \tilde{\phi}$ uniformly on compacts.

2. Causal convolution smoothing. *Extend ψ to a convex, coordinatewise strictly increasing function on all of $\mathbb{R}^L$ by*

$$\tilde{\psi}(x) = \sup_{y\in\mathbb{R}^L_+} \sup_{g\in\partial\psi(y)} \{\psi(y) + \langle g, x - y\rangle\},$$

where $\partial\psi(y)$ is the subdifferential of ψ at y.[21] *For each $\varepsilon > 0$, choose a smooth nonnegative kernel η_ε supported in $[0, c\varepsilon]^L$ satisfying $\int_{\mathbb{R}^L} \eta_\varepsilon(u)\, du = 1$, and define the causal convolution*

$$\tilde{\psi}_\varepsilon(x) = \int_{\mathbb{R}^L} \tilde{\psi}(x - u)\, \eta_\varepsilon(u)\, du,$$

with the convention $\tilde{\psi}_0 = \tilde{\psi}$. The restriction of $\tilde{\psi}_\varepsilon$ to $X = \mathbb{R}^L_+$ is

$$\psi_\varepsilon = \tilde{\psi}_\varepsilon|_X,$$

which is smooth, convex, and coordinatewise strictly increasing on X.

3. Gauge and return to the original chart. *For $\psi = \exp(\kappa)$, define the normalized tame template*

$$C_q(x) := A \frac{\sum_{\ell\in L} q_\ell x_\ell}{\sum_{\ell\in L} q_\ell x^*_\ell}, \qquad A > 0, q \in \Delta_L,$$

and the residual

$$R_{\mathcal{K}}(\kappa) := \inf_{A>0, q\in\Delta_L} \sup_{x\in K} \left|\psi(x) - \psi(x^*)\, C_q(x)\right|.$$

With the same gauge Θ as in Definition 53, define

$$\varepsilon(\kappa) := \Theta\big(R_{\mathcal{K}}(\kappa)\big).$$

The regularized cost is then

$$\hat{\kappa}(x) = \log(\psi_{\varepsilon(\kappa)}(x)) - \log(\psi_{\varepsilon(\kappa)}(0)).$$

We can again show that the operator works as intended.

---

[21] Because ψ is convex and finite on the open convex set $\mathbb{R}^L_{++}$, the subdifferential $\partial\psi(y)$ is nonempty for every $y \in \mathbb{R}^L_{++}$, and remains nonempty on the boundary by closure under limits of interior subgradients.

56 Lemma

Fix the operator $\mathfrak{D} : \mathcal{K} \to \mathcal{K}$ as in Definition 55. Then for every $\kappa \in \mathcal{K}$:

1. Class preservation. $\widehat{\kappa} := \mathfrak{D}(\kappa)$ *belongs to* $\mathcal{K}$. *Equivalently:*
   (a) $\widehat{\kappa}$ *is continuous on* X;
   (b) $\widehat{\kappa}(0) = 0$;
   (c) $\widehat{\kappa}(x) \to \infty$ *as* $\|x\| \to \infty$;
   (d) $\widehat{\kappa}$ *is strictly increasing in every coordinate; and*
   (e) $\exp(\widehat{\kappa})$ *is strictly convex on* X.
2. Smoothness. $\widehat{\kappa} \in C^{\infty}(X)$.
3. Fixed points on tame costs. *If* $\kappa \in \mathcal{K}^{[\mathfrak{T}]}$, *then* $\mathfrak{D}(\kappa) = \kappa$.
4. Continuity in κ. *The map* $\mathfrak{D} : (\mathcal{K}, d) \to (\mathcal{K}, d)$ *is continuous with respect to the compact-open metric* d.

Proof. Write $\psi = \exp(\kappa)$ on X. Let $\tilde{\psi}$ be the convex, coordinatewise nondecreasing extension to $\mathbb{R}^L$ from Definition 55, and set

$$\begin{aligned} \psi_\varepsilon &= (\tilde{\psi} * \eta_\varepsilon)\big|_X, \\ \widehat{\kappa}(x) &= \log(\psi_{\varepsilon(\kappa)}(x)) - \log(\psi_{\varepsilon(\kappa)}(0)). \end{aligned} \tag{185}$$

We verify the axioms in order after another brief remark.

Remark on compact bounds on A. On the compact set $\Delta_L \times K$, the map $(q, x) \mapsto C_q(x)$ is continuous and strictly positive. Hence

$$c^{\text{cost}}_{\min} := \min_{(q,x)\in\Delta_L\times K} C_q(x) > 0, \qquad c^{\text{cost}}_{\max} := \max_{(q,x)\in\Delta_L\times K} C_q(x) < \infty. \tag{186}$$

Let

$$m_\psi = \min_{x\in K} \psi(x), \qquad M_\psi = \max_{x\in K} \psi(x). \tag{187}$$

Any minimizer (A^*, q^*) in $R_{\mathcal{K}}(\kappa)$ satisfies

$$\frac{m_\psi}{c^{\text{cost}}_{\max}} \le A^* \le \frac{M_\psi}{c^{\text{cost}}_{\min}}. \tag{188}$$

Thus the search over A may be restricted to the compact interval

$$\mathcal{A}_\psi = \left[\, m_\psi / c^{\text{cost}}_{\max},\; M_\psi / c^{\text{cost}}_{\min} \,\right]. \tag{189}$$

The proof is identical to the technology case, replacing T_β and ϕ by C_q and ψ.

Continuity and centeredness. Convolution with a smooth kernel yields $\tilde{\psi}_\varepsilon \in C^\infty(\mathbb{R}^L)$, hence $\psi_\varepsilon \in C^\infty(X)$ and continuous. The centering term ensures $\widehat{\kappa}(0) = 0$.

Coerciveness. Since κ is coercive on X, $\psi = \exp(\kappa)$ is coercive on X. The extension satisfies $\tilde{\psi} = \psi$ on X. Convolution with a compactly supported positive kernel preserves coerciveness: for $\|x\|$ large, all points $x - u$ in the kernel window remain large in X, so $\tilde{\psi}(x-u)$ is large and hence so is $\psi_\varepsilon(x)$. Therefore $\widehat{\kappa}(x) \to \infty$ as $\|x\| \to \infty$.

Strict monotonicity. Fix $x, y \in X$ with $x \le y$ coordinatewise and $x \neq y$. Since ψ is strictly increasing in every coordinate, $\psi(y) > \psi(x)$. The extension agrees with ψ on X, so $\tilde{\psi}(y) > \tilde{\psi}(x)$. By continuity, there exists a neighborhood U of 0 in $[0, c\varepsilon]^L$ such that $\tilde{\psi}(y-u) > \tilde{\psi}(x-u)$ for all $u \in U \cap X$. The kernel η_ε assigns strictly positive mass to U, hence

$$\psi_\varepsilon(y) - \psi_\varepsilon(x) = \int \big(\tilde{\psi}(y-u) - \tilde{\psi}(x-u)\big)\eta_\varepsilon(u)\, du > 0. \tag{190}$$

Therefore ψ_ε is strictly increasing in every coordinate, and so is $\widehat{\kappa} = \log(\psi_{\varepsilon(\kappa)}) - \log(\psi_{\varepsilon(\kappa)}(0))$.

Strict exp-convexity. Each translate $z \mapsto \tilde{\psi}(z-u)$ is strictly convex on X because ψ is strictly convex and $\tilde{\psi} = \psi$ on X. A positive average of strictly convex functions is strictly convex; therefore ψ_ε is strictly convex on X. Hence $\exp(\widehat{\kappa}) = \psi_{\varepsilon(\kappa)}$ is strictly convex.

Smoothness. Since $\psi_\varepsilon \in C^\infty(X)$ and log is smooth on $(0, \infty)$, it follows that $\widehat{\kappa} \in C^\infty(X)$.

Fixed points on tame costs. If $\kappa \in \mathcal{K}^{[\mathfrak{T}]}$, then $R_{\mathcal{K}}(\kappa) = 0$ and the gauge satisfies $\varepsilon(\kappa) = \Theta(0) = 0$. Thus $\psi_{\varepsilon(\kappa)} = \psi$, and

$$\widehat{\kappa}(x) = \log(\psi(x)) - \log(\psi(0)) = \kappa(x) - \kappa(0) = \kappa(x), \tag{191}$$

using centeredness of κ.

Continuity in κ. If $\kappa_n \to \kappa$ in the compact-open metric, then $\psi_n = \exp(\kappa_n) \to \psi$ uniformly on compacts. The convex, coordinatewise monotone extension operator $\psi \mapsto \tilde{\psi}$ is continuous for uniform convergence on compacts.[22] The convolution map $(f, \varepsilon) \mapsto f * \eta_\varepsilon$ is continuous in both arguments for uniform convergence on compacts, and the gauge $\varepsilon(\cdot) = \Theta(R_{\mathcal{K}}(\cdot))$ is continuous on $(\mathcal{K}, d)$.

Therefore, we have uniform convergence on compact subsets of $\mathbb{R}^L$:

$$\tilde{\psi}_n * \eta_{\varepsilon(\kappa_n)} \longrightarrow \tilde{\psi} * \eta_{\varepsilon(\kappa)}. \tag{192}$$

Restricting to X gives uniform convergence on compact subsets of X:

$$\psi_{n,\varepsilon(\kappa_n)} \longrightarrow \psi_{\varepsilon(\kappa)}. \tag{193}$$

Finally, applying the smooth chart map $h \mapsto \log(h) - \log(h(0))$ yields

$$\mathfrak{D}(\kappa_n)(x) = \log(\psi_{n,\varepsilon(\kappa_n)}(x)) - \log(\psi_{n,\varepsilon(\kappa_n)}(0)), \tag{194}$$

which converges to

$$\log(\psi_{\varepsilon(\kappa)}(x)) - \log(\psi_{\varepsilon(\kappa)}(0)) = \mathfrak{D}(\kappa)(x) \tag{195}$$

uniformly on compact subsets of X. Hence $\mathfrak{D} : (\mathcal{K}, d) \to (\mathcal{K}, d)$ is continuous.

Conclusion. Having addressed all four points, the proof is complete. ■

Now we are justified in writing out the proposition from the main text.

---

28 Proposition
There exists a continuous regularization operator

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

such that

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

in other words, $\mathfrak{D}$ fixes the tame functions. [*Proof*.]

[22] For convex, coordinatewise nondecreasing ψ, every compact $K \subset \operatorname{int}(X)$ admits $M < \infty$ with $\|g\| \le M$ for all $g \in \partial\psi(y)$ and $y \in K$. Hence the supporting hyperplanes are equicontinuous on compacts, and uniform convergence $\psi_n \to \psi$ implies $\tilde{\psi}_n \to \tilde{\psi}$ uniformly on compacts.

Proof. This is a direct consequence of Lemmas 54 and 56; the reader may find the actual constructions in Definitions 53 and 55. The reader may find all of these immediately preceding this proof. [*Back to the text*.] ■

The next lemma confirms that the derivatives of the regularized technologies and costs depend continuously on the input technology or cost, respectively.

57 Lemma

Let $\mathfrak{D}_\tau : \mathcal{T} \to \mathcal{T}$ be the technology regularization operator from Definition 53. Then the derivative of the regularized technology depends continuously on the input under the compact–open topology. Specifically, for every compact set $K \subset X$,

$$\tau_n \underset{d}{\to} \tau \quad \Longrightarrow \quad \sup_{x \in K} \|\nabla \mathfrak{D}_\tau(\tau_n)(x) - \nabla \mathfrak{D}_\tau(\tau)(x)\| \longrightarrow 0.$$

Equivalently, the map

$$\mathcal{T} \longrightarrow C\left(X; \mathbb{R}^L\right), \qquad\qquad \tau \longmapsto \nabla \mathfrak{D}_\tau(\tau),$$

is continuous when both spaces are endowed with the compact–open topology.

The analogous result holds for the cost regularization operator $\mathfrak{D}_\kappa : \mathcal{K} \to \mathcal{K}$; that is,

$$\kappa_n \underset{d}{\to} \kappa \quad \Longrightarrow \quad \sup_{x \in K} \|\nabla \mathfrak{D}_\kappa(\kappa_n)(x) - \nabla \mathfrak{D}_\kappa(\kappa)(x)\| \longrightarrow 0.$$

Proof. Fix a compact set $K \subset X = \mathbb{R}^L_+$ and let $(\tau_n)_{n \in \mathbb{N}} \subset \mathcal{T}$ with $\tau_n \underset{d}{\to} \tau$. Write

$$\phi_n = \log(1 + \tau_n), \qquad\qquad \phi = \log(1 + \tau), \tag{196}$$

and let $\tilde{\phi}_n, \tilde{\phi}$ denote their concave, coordinatewise nondecreasing extensions to $\mathbb{R}^L$ from Definition 53. For $\varepsilon > 0$, let η_ε be the causal kernel, and set

$$\phi_{n,\varepsilon} = (\tilde{\phi}_n * \eta_\varepsilon)\big|_X, \qquad\qquad \phi_\varepsilon = (\tilde{\phi} * \eta_\varepsilon)\big|_X. \tag{197}$$

Finally, define the bandwidths

$$\varepsilon_n = \Theta\big(R_{\mathcal{T}}(\tau_n)\big), \qquad\qquad \varepsilon = \Theta\big(R_{\mathcal{T}}(\tau)\big), \tag{198}$$

and recall $\mathfrak{D}_\tau(\tau_n) = \exp(\phi_{n,\varepsilon_n}) - 1$.

Step 1 (Chart continuity). Since $t \mapsto \log(1+t)$ is smooth and uniformly continuous on bounded sets, $\tau_n \to \tau$ uniformly on every compact implies $\phi_n \to \phi$ uniformly on every compact subset of X. Hence,

$$\forall K' \Subset X : \quad |\phi_n - \phi|_{\infty;K'} \to 0. \tag{199}$$

Step 2 (Stability of the concave extension). Define the extension operator

$$E[\varphi](x) = \inf_{y \in \mathbb{R}^L_+} \inf_{g \in \partial\varphi(y)} \{\varphi(y) + \langle g, x-y \rangle\}. \tag{200}$$

Let $K'' \Subset \mathbb{R}^L$ be compact. If $|\varphi_1 - \varphi_2|_{\infty;B} \le \delta$ on a compact $B \supset K'' \cap \mathbb{R}^L_+$, then

$$|E[\varphi_1] - E[\varphi_2]|_{\infty;K''} \le \delta. \tag{201}$$

Reason: for any $x \in K''$ and (y,g) feasible for $E[\varphi_2]$, shifting the intercept of the affine minorant for φ_1 changes the value by at most δ, and taking infima preserves this bound. Hence E is 1-Lipschitz locally in the sup norm. Applying this with $\varphi_1 = \phi_n$, $\varphi_2 = \phi$ yields

$$|\tilde{\phi}_n - \tilde{\phi}|_{\infty;K''} \to 0. \tag{202}$$

Step 3 (Convolution continuity and derivatives at fixed bandwidth). For fixed $\varepsilon > 0$, convolution with η_ε is continuous from the compact–open topology on $\mathbb{R}^L$ to $C^\infty(\mathbb{R}^L)$ and

$$\nabla(\tilde{\phi} * \eta_\varepsilon) = \tilde{\phi} * (\nabla\eta_\varepsilon). \tag{203}$$

For every compact $K'' \Subset \mathbb{R}^L$,

$$|\nabla(\tilde{\phi}_n * \eta_\varepsilon) - \nabla(\tilde{\phi} * \eta_\varepsilon)|_{\infty;K''} \le |\tilde{\phi}_n - \tilde{\phi}|_{\infty;K''+[0,c\varepsilon]^L} \, \|\nabla\eta_\varepsilon\|_{L^1} \to 0. \tag{204}$$

Restricting to X gives, for every compact $K \Subset X$,

$$|\nabla\phi_{n,\varepsilon} - \nabla\phi_\varepsilon|_{\infty;K} \to 0. \tag{205}$$

Step 4 (Continuity of the bandwidth selection). The residual

$$R_{\mathcal{T}}(\tau) = \inf_{\beta \in \Delta_L} \sup_{x \in K_0} \left|\phi(x) - \phi(x^*) T_\beta(x)\right| \tag{206}$$

is continuous in τ because the map $(\phi, \beta, x) \mapsto \phi(x) - \phi(x^*) T_\beta(x)$ is continuous on the compact set $K_0 \times \Delta_L$, and infima and suprema over compact sets preserve continuity. As Θ is continuous and strictly increasing, we obtain $\varepsilon_n \to \varepsilon$.

Step 5 (Joint continuity in function and bandwidth). Let $K \Subset X$ be fixed. For $\varepsilon', \varepsilon > 0$,

$$\nabla(\tilde{\phi} * \eta_{\varepsilon'})(x) - \nabla(\tilde{\phi} * \eta_\varepsilon)(x) = \tilde{\phi} * \big(\nabla\eta_{\varepsilon'} - \nabla\eta_\varepsilon\big)(x). \tag{207}$$

Since $\varepsilon' \mapsto \nabla\eta_{\varepsilon'}$ is continuous in L^1 and $\tilde{\phi}$ is locally bounded, we obtain

$$|\nabla\phi_{\varepsilon'} - \nabla\phi_\varepsilon|_{\infty;K} \xrightarrow[\varepsilon' \to \varepsilon]{} 0. \tag{208}$$

Combining with Step 3 and $\varepsilon_n \to \varepsilon$ yields joint continuity:

$$|\nabla\phi_{n,\varepsilon_n} - \nabla\phi_\varepsilon|_{\infty;K} \to 0. \tag{209}$$

Step 6 (Return to original chart and product rule). Set $u_n = \phi_{n,\varepsilon_n}$ and $u = \phi_\varepsilon$. Then

$$\nabla\mathfrak{D}_\tau(\tau_n) = \nabla\big(\exp(u_n) - 1\big) = \exp(u_n)\nabla u_n, \tag{210}$$
$$\nabla\mathfrak{D}_\tau(\tau) = \nabla\big(\exp(u) - 1\big) = \exp(u)\nabla u. \tag{211}$$

From Steps 3–5 we have $u_n \to u$ and $\nabla u_n \to \nabla u$ uniformly on K. Since $z \mapsto \exp(z)$ and $(a, b) \mapsto ab$ are continuous and u_n are uniformly bounded on K, it follows that

$$\sup_{x \in K} \big|\nabla\mathfrak{D}_\tau(\tau_n)(x) - \nabla\mathfrak{D}_\tau(\tau)(x)\big| \longrightarrow 0. \tag{212}$$

As $K \Subset X$ was arbitrary, the claimed continuity in the compact–open topology follows.

Costs. For $\kappa_n \underset{d}{\to} \kappa$, repeat the argument with the convex chart $\psi_n = \exp(\kappa_n)$, the convex extension

$$\tilde{\psi}(x) = \sup_{y \in \mathbb{R}^L_+} \sup_{g \in \partial\psi(y)} \big\{\psi(y) + \langle g, x - y\rangle\big\}, \tag{213}$$

the same causal kernel family, and the normalization

$$\hat{\kappa}_n = \log(\psi_{n,\varepsilon_n}) - \log(\psi_{n,\varepsilon_n}(0)). \tag{214}$$

Step 2 is replaced by the analogous 1-Lipschitz stability for convex extensions. The normalization term depends continuously on (ψ, ε) by Step 5 (evaluate at $x = 0$). The same product and chain rule argument yields

$$\sup_{x \in K} \left| \nabla \mathfrak{D}_\kappa(\kappa_n)(x) - \nabla \mathfrak{D}_\kappa(\kappa)(x) \right| \to 0. \tag{215}$$

This completes the proof. ∎

---

58 Lemma

The tamification maps are continuous in the compact–open topology on $\mathcal{T}^{[\infty]}$ and $\mathcal{K}^{[\infty]}$ when defined via the regularized derivatives at 0*:*

$$\mathfrak{T}_\tau : \mathcal{T}^{[\infty]} \longrightarrow \mathcal{T}^{[\mathfrak{T}]}, \qquad \mathfrak{T}_\kappa : \mathcal{K}^{[\infty]} \longrightarrow \mathcal{K}^{[\mathfrak{T}]},$$

where

$$\begin{aligned} A_\tau &:= \sum_{\ell \in L} \partial_\ell\big(\mathfrak{D}_\tau(\tau)\big)(0), \\ \beta_\ell &:= \frac{\partial_\ell\big(\mathfrak{D}_\tau(\tau)\big)(0)}{A_\tau}, \\ \mathfrak{T}_\tau(\tau)(x) &:= A_\tau \sum_{\ell \in L} \beta_\ell \log(1 + x_\ell), \end{aligned}$$

and

$$\begin{aligned} A_\kappa &:= \sum_{\ell \in L} \partial_\ell\big(\mathfrak{D}_\kappa(\kappa)\big)(0), \\ q_\ell &:= \frac{\partial_\ell\big(\mathfrak{D}_\kappa(\kappa)\big)(0)}{A_\kappa}, \\ \mathfrak{T}_\kappa(\kappa)(x) &:= A_\kappa \sum_{\ell \in L} q_\ell x_\ell. \end{aligned}$$

Proof. We do technologies; costs are analogous.

By Lemma 57, the map

$$\mathcal{T} \longrightarrow C\big(X; \mathfrak{R}^L\big), \qquad \tau \longmapsto \nabla \mathfrak{D}_\tau(\tau) \tag{216}$$

is continuous for the compact–open topology. Evaluating at 0 is continuous on $C\big(X; \mathbb{R}^L\big)$, hence

$$\tau \longmapsto D\big(\mathfrak{D}_\tau(\tau)\big)(0) \tag{217}$$

is continuous from $\mathcal{T}^{[\infty]}$ (with the compact–open topology) into $\mathbb{R}^L$.

On the open cone

$$C := \{a \in \mathbb{R}^L_{\geq 0} \, : \, \sum_{\ell} a_\ell > 0\}, \tag{218}$$

the normalization map

$$N : C \longrightarrow (0, \infty) \times \Delta_L, \qquad a \longmapsto \Big(A = \sum_{\ell} a_\ell, \; \beta_\ell = a_\ell / A\Big) \tag{219}$$

is continuous. For $\tau \in \mathcal{T}$, coordinatewise monotonicity of $\mathfrak{D}_\tau(\tau)$ implies $D\big(\mathfrak{D}_\tau(\tau)\big)(0) \in C$.

Finally, the assembly map

$$(A, \beta) \longmapsto \Big[\, x \mapsto A \sum_{\ell \in L} \beta_\ell \log(1 + x_\ell) \,\Big] \tag{220}$$

is continuous from $(0, \infty) \times \Delta_L$ into $\mathcal{T}^{[\mathfrak{T}]}$ endowed with the compact–open topology. Composing the three continuous maps

$$\tau \longmapsto \nabla\big(\mathfrak{D}_\tau(\tau)\big)(0) \longmapsto (A_\tau, \beta) \longmapsto \mathfrak{T}_\tau(\tau) \tag{221}$$

yields continuity of $\mathfrak{T}_\tau$.

For costs, replace $\log(1 + x_\ell)$ with x_ℓ and use Lemma 56 to obtain continuity of

$$\kappa \longmapsto \nabla\big(\mathfrak{D}_\kappa(\kappa)\big)(0). \tag{222}$$

Normalization and assembly are the same, giving continuity of $\mathfrak{T}_\kappa$. ∎

---

29 Lemma

For all $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, we have $\mathfrak{T}(\tau, \kappa) = (\tau, \kappa)$. [*Proof*.]

Proof. Let $(\tau, \kappa) \in \mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$, and let

$$\tau(x) = A_\tau \sum_{\ell \in L} \beta_\ell \log(1 + x_\ell), \qquad \kappa(x) = A_\kappa \sum_{\ell \in L} q_\ell x_\ell, \tag{223}$$

where $A_\tau, A_\kappa > 0$ and $\beta, q \in \Delta_L$ witness their tameness.

By direct differentiation,

$$\frac{\partial \tau}{\partial x_\ell}(0) = A_\tau \beta_\ell, \qquad\qquad \frac{\partial \kappa}{\partial x_\ell}(0) = A_\kappa q_\ell. \tag{224}$$

Therefore,

$$\sum_{j\in L} \frac{\partial \tau}{\partial x_j}(0) = A_\tau, \qquad\qquad \sum_{j\in L} \frac{\partial \kappa}{\partial x_j}(0) = A_\kappa, \tag{225}$$

and hence the normalized weights computed by $\mathfrak{T}$ are

$$\begin{aligned} \beta'_\ell &:= \frac{\partial_\ell \tau(0)}{\sum_{j\in L} \partial_j \tau(0)} = \frac{A_\tau \beta_\ell}{A_\tau} = \beta_\ell, \\ q'_\ell &:= \frac{\partial_\ell \kappa(0)}{\sum_{j\in L} \partial_j \kappa(0)} = \frac{A_\kappa q_\ell}{A_\kappa} = q_\ell. \end{aligned} \tag{226}$$

Similarly, the scales recovered by $\mathfrak{T}$ are

$$A'_\tau = \sum_{j\in L} \partial_j \tau(0) = A_\tau, \qquad A'_\kappa = \sum_{j\in L} \partial_j \kappa(0) = A_\kappa. \tag{227}$$

Substituting these values into the definition of $\mathfrak{T}$ yields

$$\begin{aligned} \mathfrak{T}(\tau,\kappa) &= \Big(A'_\tau \sum_{\ell\in L} \beta'_\ell \log(1+x_\ell),\ A'_\kappa \sum_{\ell\in L} q'_\ell x_\ell\Big) \\ &= \Big(A_\tau \sum_{\ell\in L} \beta_\ell \log(1+x_\ell),\ A_\kappa \sum_{\ell\in L} q_\ell x_\ell\Big) = (\tau,\kappa). \end{aligned} \tag{228}$$

Hence $\mathfrak{T}$ acts as the identity on $\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$. [*Back to the text*.] ■

---

30 Proposition
$\mathcal{T}^{[\mathfrak{T}]} \times \mathcal{K}^{[\mathfrak{T}]}$ is a strong deformation retract of $\mathcal{T} \times \mathcal{K}$. [*Proof*.]

Proof. We handled most of the details in the main text, but we can be more explicit here. Recall that the homotopies

$$H : [0,1] \times \mathcal{T} \times \mathcal{K} \longrightarrow \mathcal{T} \times \mathcal{K}, \tag{229}$$

are defined for the technologies by

$$\begin{cases} \left((1+\tau)^{1-2t} \cdot (1+\mathfrak{D}_\tau(\tau))^{2t} - 1\right), & t \in [0, 1/2], \\ \left((1+\mathfrak{D}_\tau(\tau))^{2-2t} \cdot (1+(\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau))^{2t-1} - 1\right), & t \in [1/2, 1], \end{cases} \tag{230}$$

and for the costs by

$$\begin{cases} \log((1-2t)\exp\kappa + 2t\exp\mathfrak{D}_\kappa(\kappa)), & t \in [0, 1/2], \\ \log((2-2t)\exp\mathfrak{D}_\kappa(\kappa) + (2t-1)\exp(\mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa)(\kappa)), & t \in [1/2, 1]. \end{cases} \tag{231}$$

Since the $\mathfrak{D}$ operators are continuous by Lemmas 54 and 56 and the $\mathfrak{T}$ operators are continuous by Lemma 58, it follows that H is continuous. Moreover, at $t = 0$, we have $H(0, \tau, \kappa) = (\tau, \kappa)$, and at $t = 1$, we have $H(1, \tau, \kappa) = (\mathfrak{T}_\tau \circ \mathfrak{D}_\tau(\tau), \mathfrak{T}_\kappa \circ \mathfrak{D}_\kappa(\kappa))$; by construction, these are tame. Since both $\mathfrak{D}$ and $\mathfrak{T}$ fix tame technologies and costs by Lemmas 29, 54 and 56, it follows that $H(t, \tau, \kappa) = (\tau, \kappa)$ for all t whenever (τ, κ) is tame.

Therefore, the only thing we really need to show is that the homotopy remains in $\mathcal{T} \times \mathcal{K}$ for all $t \in [0, 1]$. We will keep this brief; write $\phi_t = \log(1 + \tau_t)$ and $\psi_t = \exp(\kappa_t)$.

1. **Technologies.** For each $t \in [0, 1]$, the function τ_t defined above satisfies:
   (a) *Continuity:* Each τ_t is continuous since the defining operations (addition, multiplication, exponentiation, and logarithm) are continuous and the composing functions τ, $\mathfrak{D}_\tau(\tau)$, and $(\mathfrak{T}_\tau \circ \mathfrak{D}_\tau)(\tau)$ are continuous.
   (b) *Monotonicity:* Each of $\phi_0 = \log(1+\tau)$, $\phi_{\mathfrak{D}}$, and $\phi_{\mathfrak{T}\circ\mathfrak{D}}$ is coordinatewise nondecreasing. Since ϕ_t is a convex combination of these functions on each interval, it is also coordinatewise nondecreasing. Hence $\tau_t = \exp(\phi_t) - 1$ is coordinatewise nondecreasing.
   (c) *Log-Concavity:* Concavity of ϕ_t follows because a convex combination of concave functions is concave. Therefore $\log(1+\tau_t) = \phi_t$ is concave.
   (d) *Ray surjectivity:* Let $v_0, v_{\mathfrak{D}} \in X$ witness ray surjectivity for τ and $\mathfrak{D}_\tau(\tau)$ respectively. Set $v_t = v_0 + v_{\mathfrak{D}}$. Because each ϕ_i is coordinatewise nondecreasing,

$$\phi_i(sv_t) \geq \phi_i(sv_i) \tag{232}$$

   for all $s \geq 0$, and the right-hand sides diverge to $+\infty$. Thus $\phi_t(sv_t)$ is unbounded and strictly increasing in s, implying that $\tau_t(sv_t) =$

$\exp(\phi_t(sv_t)) - 1$ is continuous, strictly increasing, and unbounded along that ray. The same reasoning applies for $t \in [\frac{1}{2}, 1]$ using $v_{\mathfrak{D}}$ and $v_{\mathfrak{T} \circ \mathfrak{D}}$. Hence τ_t is ray surjective.

2. **Costs.** For each $t \in [0, 1]$, the function κ_t defined above satisfies:
   (a) *Continuity:* Continuity of κ_t follows from the same reasoning as for τ_t.
   (b) *Monotonicity:* Each $\psi_i = \exp(\kappa_i)$ is coordinatewise strictly increasing, and convex combinations preserve coordinatewise increase. Hence ψ_t is strictly increasing and so is $\kappa_t = \log(\psi_t)$.
   (c) *Convexity:* Convex combinations of convex functions are convex, so ψ_t is convex and $\exp(\kappa_t) = \psi_t$ remains convex.
   (d) *Coerceiveness:* This is closed under convex combinations, so κ_t remains coercive.
   (e) *Centering:* $\kappa_t(0) = 0$ for all t, and κ_t remains finite on compact sets.

We therefore conclude that $\kappa_t \in \mathcal{K}$ for all t; this is the final piece we needed before concluding that H is a homotopy in $\mathcal{T} \times \mathcal{K}$, and thus $\mathcal{T}_{[\mathfrak{T}]} \times \mathcal{K}_{[\mathfrak{T}]}$ is a strong deformation retract of $\mathcal{T} \times \mathcal{K}$. [*Back to the text*.] ∎

---

33 Proposition

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

Proof. Choose and fix any $m \in M$, which we will suppress from notation whenever possible. In case $m = 0$, the result is trivial, as the only solution to the production problem SPP (m, τ, κ) is $\pi_{\tau,\kappa}(0) = 0$. We therefore suppose $m > 0$. Let π_0 and π_1 be the solutions to the production problem SPP (m, τ, κ) for two tame pairs (τ_0, κ_0) and (τ_1, κ_1), with parameters $(A_{\tau,0}, \beta_0, A_{\kappa,0}, q_0)$ and $(A_{\tau,1}, \beta_1, A_{\kappa,1}, q_1)$, respectively. Define the interpolated state

$$\pi_t := (1 - t)\pi_0 + t\pi_1,$$

for $t \in [0, 1]$. We will construct tame parameters $(A_\tau(t), \beta_t, A_\kappa(t), q_t)$ and multipliers (λ_t, η_t) such that π_t satisfies the first-order conditions FOC (m, τ, κ) for (τ_t, κ_t).

Step 1: Choosing q_t. Pick any $q_t \in \Delta_L$ with strictly positive coordinates; for definiteness, take $q_t^\ell = 1/L$.

Step 2: Defining β_t. Set

$$c_t := \left(\sum_{j\in L}(1+\pi_t^j)q_t^j\right)^{-1}, \qquad \beta_t^\ell := c_t(1+\pi_t^\ell)q_t^\ell \quad \text{for all } \ell \in L. \tag{233}$$

Then $\beta_t \in \Delta_L$, $\beta_t^\ell \geqslant 0$ for all ℓ, and this choice ensures that a common multiplier λ_t can satisfy all stationarity conditions simultaneously.

Step 3: Setting the scale parameters. Let $A_\kappa(t) > 0$ be arbitrary (for simplicity, take $A_\kappa(t) = 1$) and define

$$S_t := \sum_{\ell\in L}\beta_t^\ell \log(1+\pi_t^\ell), \qquad A_\tau(t) := \frac{m}{S_t}. \tag{234}$$

Then the feasibility condition $m - A_\tau(t)\sum_{\ell\in L}\beta_t^\ell \log(1+\pi_t^\ell) = 0$ holds by construction.

Step 4: Determining the multipliers. Define

$$\lambda_t := \frac{A_\kappa(t)}{A_\tau(t)\, c_t}, \qquad \eta_t^\ell := A_\kappa(t)q_t^\ell - \lambda_t\frac{A_\tau(t)\beta_t^\ell}{1+\pi_t^\ell}. \tag{235}$$

By the definitions of β_t, c_t, and λ_t, it follows that $\eta_t^\ell = 0$ for all $\ell \in L$, so the complementarity conditions $\eta_t^\ell \pi_t^\ell = 0$ are satisfied automatically.

Step 5: Verification. For every $\ell \in L$,

$$A_\kappa(t)q_t^\ell - \lambda_t\frac{A_\tau(t)\beta_t^\ell}{1+\pi_t^\ell} - \eta_t^\ell = 0, \tag{236}$$

and the feasibility condition holds by Step 3. Thus, the full system $\mathsf{FOC}(m, \tau_t, \kappa_t)$ is satisfied. Because π_t is a convex combination of π_0 and π_1, the mapping $t \mapsto \pi_t$ is continuous, and the construction above produces a corresponding tame pair (τ_t, κ_t) for each $t \in [0, 1]$.

Conclusion. The interpolated state π_t therefore solves the production problem for some tame pair (τ_t, κ_t) for every $t \in [0, 1]$. Hence the set of tame states $\mathcal{P}_{\mathcal{T}^{[\mathfrak{T}]}\times\mathcal{K}^{[\mathfrak{T}]}}$ is convex. [*Back to the text.*] ∎

References

Aliprantis, Charalambos D. and Kim C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Third ed. Berlin: Springer.

Asal, Victor, Justin Conrad and Nathan Toronto. 2017. "I Want You! The Determinants of Military Conscription." *Journal of Conflict Resolution* 61(7):1456–1481.

Bartelson, Jens. 2001. *The Critique of the State*. Cambridge: Cambridge University Press.

Benedict, Ruth. 1934. *Patterns of Culture*. Mariner books (2005) ed. Boston: Houghton Mifflin.

Beviá, Carmen and Luis C. Corchón. 2010. "Peace Agreements Without Commitment." *Games and Economic Behavior* 68(2):469–487.

Cappella Zielinski, Rosella. 2016. *How States Pay for Wars*. Ithaca, NY: Cornell University Press.

Cappella Zielinski, Rosella and Paul Poast. 2024. War Financing and Foreign Debt. In *What Do We Know about War?*, ed. Sara McLaughlin Mitchell and John A. Vasquez. Bloomsbury Publishing USA pp. 260–271.

Chambers, Robert G. 1988. *Applied Production Analysis: A Dual Approach*. Cambridge, UK: Cambridge 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.

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, Charlotte. 2013. "Theorizing Agency in Hobbes's Wake: The Rational Actor, the Self, or the Speaking Subject?" *International Organization* 67(2):287–316.

Fearon, James D. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* 72(3):523–559.

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.

Garfinkel, Michelle R. 1990. "Arming as a Strategic Investment in a Cooperative Equilibrium." *American Economic Review* 80(1):50–68.

Hay, Colin. 2014. "Neither Real nor Fictitious but 'as If Real'? A Political Ontology of the State." *British Journal of Sociology* 65(3):459–480.

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.

Jessop, Bob. 2014. "Towards a Political Ontology of State Power: A Comment on Colin Hay's Article." *British Journal of Sociology* 65(3):481–486.

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.

Knio, Karim. 2023. "Why and How Ontology Matters: A Cartography of Neoliberalism(s) and Neoliberalization(s)." *Journal for the Theory of Social Behaviour* 53(2):160–182.

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.

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.

Morita, Kiiti. 1963. "On the Product of Paracompact Spaces." *Proceedings of the Japan Academy, Series A, Mathematical Sciences* 39(8):559–563.

Morris, Ian. 2014. *War: Who Is It Good For? The Role of Conflict in Civilisation, from Primates to Robots*. Paperback (2015) ed. London: Profile Books.

Mosley, Layna and B. Peter Rosendorff. 2023. "Government Choices of Debt Instruments." *International Studies Quarterly* 67(2).

North, Douglass C. 1981. *Structure and Change in Economic History*. New York: Norton.

Parker, Geoffrey. 1996. *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800*. Second ed. Cambridge, UK: Cambridge University Press.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1):115–132.

Quine, Willard V. 1948. "On What There Is." *Review of Metaphysics* 2(5):21–38.

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.

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.

Stiglitz, Joseph E. and Linda J. Bilmes. 2008. *The Three Trillion Dollar War: The True Cost of the Iraq Conflict*. First ed. New York: W.W. Norton.

Sundaram, Rangarajan K. 1996. *A First Course in Optimization Theory*. Cambridge, UK: Cambridge University Press.

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990–1992*. Revised (1992) paperback ed. Cambridge, MA: Blackwell.

Tilly, Charles, ed. 1975. *The Formation Of National States In Western Europe*. Number 8 *in* "Studies in Political Development" Princeton, NJ: Princeton University Press.

Van Creveld, Martin. 1985. *Command in War*. Cambridge, MA: Harvard University Press.

Van Creveld, Martin. 1991. *Technology and War: From 2000 B.C. to the Present*. Revised and expanded ed. New York: Free Press.

Van Creveld, Martin. 2004. *Supplying War: Logistics from Wallenstein to Patton*. Second ed. Cambridge, UK: Cambridge 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.

Wendt, Alexander. 1999. *Social Theory of International Politics*. Number 67 *in* "Cambridge Studies in International Relations" Cambridge, UK: Cambridge University Press.