# The Binding Force of Law: Stare Decisis as Constraint Accumulation

**Abstract**

Why does law bind independent judges? We model horizontal stare decisis as constraint accumulation in a multidimensional parameter space. "The law" is not a single determinate rule but a feasible set of admissible decision rules consistent with past holdings. Each case generates a halfspace constraint on this set, and the feasible set shrinks monotonically over time. We establish six results. First, the penumbra—the set of cases where multiple outcomes are legally plausible—shrinks monotonically as holdings accumulate (Propositions 1–3). Second, judges choose their ideologically preferred outcome whenever the law permits, generating path-dependent doctrinal drift even though every individual decision is lawful (Propositions 4–5). Third, overruling occurs when ideological gain exceeds the cost of removing blocking constraints (Proposition 6). A VC-dimension corollary shows that higher-dimensional legal questions require proportionally more precedent to resolve, formalizing why multidimensional doctrine resists convergence. We illustrate the framework with Equal Protection jurisprudence, mapping landmark cases from *Brown* to *Students for Fair Admissions* to constraint-accumulation dynamics.

## 1 Introduction and Motivation

Why, and when, does law bind judges? In modern legal systems judges enjoy substantial independence, legal texts are linguistically open-ended, and factual situations are highly heterogeneous. In principle, nothing prevents judges from deciding cases purely in line with their own normative views. Yet in practice, judicial decisions display a remarkable degree of continuity, stability, and mutual constraint across judges and over time. This phenomenon—commonly described as the *binding force of law*—is a foundational premise of markets, rights, and democratic governance. At the same time, persistent disagreement, ideological voting in hard cases, and gradual doctrinal drift are equally familiar features of appellate judging. A convincing theory of law must explain how these two facts coexist.

Existing approaches fall short in opposite directions. Standard economic models of adjudication typically take the content of the law as given and assume that judges follow it (e.g., Gennaioli and Shleifer, 2007; Fernández and Ponzetto, 2012), thereby abstracting away from the core question of why legal rules constrain independent decision-makers at all. Attitudinal and behavioral models (e.g., Segal and Spaeth, 2002), by contrast, emphasize judicial preferences and ideological bias, but struggle to explain why legal argumentation, precedent, and reason-giving play such a central and persistent role in judicial practice. If judges simply maximize preferences, why does legal reasoning matter, and why are some decisions widely regarded as illegitimate even when they reflect sincere normative disagreement?

This paper develops a formal model of *horizontal stare decisis*—the mutual constraint judges exert on one another through precedent—that reconciles binding force with discretion. The starting point is two simple but often neglected observations. First, cases are inherently multidimensional (cf. Callander and Clark, 2017): each dispute involves many legally salient features, and the probability that a future case replicates past facts exactly is essentially zero. Second, judicial decisions nonetheless constrain future courts because they do more than announce outcomes; they articulate *holdings*—general statements about how legally relevant features relate to outcomes. Holdings are neither mere descriptions of facts nor complete rules. Instead, they impose partial, structured restrictions on how future cases may be justified.

We formalize this idea by modeling "the law" not as a single determinate rule, but as a set of admissible decision rules that remain consistent with past holdings (cf. Re, 2021). Each holding imposes a constraint on this set: at a minimum, every admissible rule must agree with the decided outcome at the case's facts. A decision is legally permissible if it can be justified by at least one rule within the current admissible set. Decisions that fall outside this set trigger sanctions from peer judges, capturing reputational and institutional consequences of appearing not to follow the law. In this way, the binding force of law arises endogenously from the need to remain within a commonly recognized space of plausible legal justifications.

This framework yields three core insights. First, law can bind strongly even when it is underdetermined. Judges are constrained because they must remain within the admissible set of rules, but they retain discretion because that set typically contains many elements. Second, ideology operates through law rather than against it: judges differ in their preferred legal rules and therefore select different outcomes and holdings when multiple justifications are legally plausible. Third, doctrine evolves gradually. Holdings restrict the admissible set over time, but because constraints are partial, multidimensional, and sometimes reversed at a cost, the law need not converge to a fixed point. Instead, it may drift in response to a sequence of individually plausible decisions, giving rise to familiar concerns about "slippery slopes" and doctrinal instability.

Even when judges have no discretion over the form of their holdings—when each decision mechanically generates the constraint "all admissible rules must agree with this outcome at these facts"—the geometry of constraint accumulation produces rich dynamics: law becomes more determinate over time, yet the feasible set need not converge to a point, and the direction of doctrinal drift depends on the sequence of cases. The base model isolates this pure geometric channel. An extension allows judges to also shape the *breadth* of their holdings, creating a tradeoff between influence on future doctrine and risk of overruling; this adds a strategic dimension to doctrine-making while preserving the geometric core. The result is a tractable, positive theory of horizontal stare decisis that explains how law simultaneously constrains, enables disagreement, and changes over time.

## 1.1 Related Literature

Our model connects to several literatures that have studied precedent, legal indeterminacy, and judicial decision-making from different angles.

**Formal models of precedent.** Within economics and political science, our closest neighbors are models of case-law evolution and judicial strategy. Gennaioli and Shleifer (2007) model judges choosing law with a cost of change, but treat law as a known single rule—we model it as an ambiguous feasible set. Niblett et al. (2010) analyze holding breadth and case-law evolution in one dimension; our framework is the multidimensional extension. Callander and Clark (2017) are the closest predecessor: they study judicial learning in a multi-dimensional fact space with a shrinking feasible set and path dependence, but do not formalize holdings as objects separate from outcomes and do not model the breadth tradeoff or overruling. Fox and Vanberg (2014) study narrow versus broad judicial decisions and their strategic implications. Cameron et al. (2019) explain why horizontal stare decisis is sustainable among polarized judges via implicit log-rolls across cases. Baker and Mezzetti (2012) endogenize analogical reasoning in a dynamic model of judge-made law. Our model builds on all of these but adds the holding-as-linear-constraint formalization that connects binding force, discretion, and drift in a unified framework.

**Legal indeterminacy and admissible sets.** The philosophical and legal-theoretic foundations for our approach lie in Hart's (1961) concept of open texture and Dworkin's (1986) argument that multiple principled interpretations can fit the same body of past decisions—our feasible set $\mathcal{F}_t$ operationalizes these ideas. Re (2021) reconceptualizes precedent as *permission*, defining an admissible set of rulings rather than primarily prohibiting departures, directly paralleling our framework. Kornhauser (1989) develops the coordination value of precedent as a focal point. Kaplow (1992) analyzes rules versus standards; our model endogenizes the choice, as judges sometimes convert open-ended standards into bright-line constraints. Volokh (2003) catalogs slippery-slope mechanisms—our drift result formalizes the precedent slippery slope he identifies.

**AI & Law case-space models.** An important tradition that economic models of courts have largely ignored represents cases as vectors in a factor or dimension space and reasons about precedential constraint via structured argumentation. Since the HYPO system (Ashley, 1990), this literature has separated outcome from rationale: Branting (1993) models ratio decidendi as a justification structure; Horty (2004) and Horty and Bench-Capon (2012) formalize when a precedent constrains future cases via factor orderings; Rigoni (2015, 2018) and Prakken (2021) refine these models for dimensions and hierarchies. Rissland and Collins (1986) explicitly connected doctrine to version-space learning. Our model imports the AI & Law insight that "what binds is the holding, not the outcome" into economics, while replacing discrete factor orderings with continuous linear constraints and adding strategic judicial behavior with ideology.

**Learning theory and empirical connections.** Our feasible set $\mathcal{F}_t$—a polytope in parameter space that shrinks via intersection with linear constraints—is formally a version space (Mitchell, 1982), and our constraint-accumulation dynamics are isomorphic to cutting-plane methods (James E. Kelley, 1960). This connection provides ready-made tools: convex body shrinkage results quantify binding force as volume reduction, and VC dimension (Vapnik, 1995) formalizes the expressiveness of the constraint language. Recent work applies learning theory to common-law institutions: Hartline et al. (2022) model the system's information-aggregation efficiency; Dutz et al. (2025) study strategic litigation as teaching. We model the judge-side constraint problem. Empirically, the model predicts that ideology matters more when the feasible set is large—consistent with Lindquist and Cross (2005), who find that precedent constrains ideology but that sufficiently rich case law can also supply ammunition for preferred outcomes—and with prediction-accuracy measures (Katz et al., 2017) that implicitly estimate feasible-set diameter.

The remainder of the paper formalizes this framework and explores its implications. We first present the model and define legal plausibility, holdings, and sanctions. We then analyze equilibrium behavior, showing how binding force, ideological disagreement, and doctrinal drift arise naturally in multidimensional legal environments. We conclude by discussing implications for legal doctrine, judicial design, and empirical research on judicial behavior.

## 2 Model

Time is discrete, $t = 1, 2, \ldots$. In each period a case arrives with fact vector

$$z_t \in \mathbb{R}^k,$$

drawn i.i.d. from a distribution $G$ on $\mathbb{R}^k$ (for simplicity, $G$ has a density). A judge $j_t$ is drawn i.i.d. from a finite set $J$.

### 2.1 Decision rules (the object of "law")

Let an affine decision rule be parameterized by $(w, c) \in \mathbb{R}^k \times \mathbb{R}$:

$$f_{w,c}(z) \equiv w^\top z - c.$$

The associated binary outcome is

$$d(z; w, c) \equiv \mathbf{1}\{w^\top z \geq c\} \in \{0, 1\}.$$

Interpret $d = 1$ as "plaintiff wins" / "liability".

**Why affine rules?** Affine decision boundaries are not a literal model of judicial cognition. Rather, they are a tractable reduced form for any setting in which outcomes depend on a weighted combination of legally salient factors—the standard assumption in balancing tests, multifactor standards, and totality-of-the-circumstances analyses. Just as linear utility in consumer theory approximates smooth preferences locally, affine rules approximate smooth decision boundaries via first-order Taylor expansion around any reference case. The affine class is also the simplest that is genuinely multidimensional ($k \geq 2$) and admits a complete geometric analysis via halfspace constraints. The monotone-shrinkage and version-space results generalize to any parametric hypothesis class for which holdings correspond to convex constraints in parameter space; affine rules are the simplest such class where holdings are halfspaces and the full geometric analysis admits closed-form solutions.

### 2.2 Judges' ideology (preferences over what the law should be)

Each judge $j \in J$ has an ideal affine rule $(w_j, c_j)$. A simple reduced-form utility over the outcome in case $z$ is

$$u_j(d, z) \equiv \alpha \cdot \mathbf{1}\{d = d(z; w_j, c_j)\},$$

where $\alpha > 0$ indexes the strength of outcome preferences. (Any monotone loss in the distance between the chosen outcome and the ideal outcome can be used instead.)

### 2.3 Holdings as Linear Constraints

In the most general form, a *linear-constraint holding* is a finite set of linear inequalities on $(w, c)$:

$$H \equiv \{a_m^\top (w, c) \leq b_m\}_{m=1}^M,$$

where each $a_m \in \mathbb{R}^{k+1}$ and $b_m \in \mathbb{R}$. Let $\mathcal{H}$ denote the set of all such finite systems. For a holding $H \in \mathcal{H}$, write

$$(w, c) \in H \iff a_m^\top (w, c) \leq b_m \text{ for all } m.$$

The public legal state at time $t$ is a feasible set $\mathcal{F}_t \subset \mathbb{R}^{k+1}$ of admissible affine rules ("plausible versions of law"). Initialize with a broad set, e.g.

$$\mathcal{F}_1 \equiv \{(w, c) : \|w\|_\infty \leq W,\ |c| \leq C_0\},$$

or any nonempty compact convex set.

If a new holding $H_t$ is issued and treated as binding, the law updates by intersection:

$$\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t.$$

Thus, over time, holdings restrict which affine decision boundaries remain admissible.

### Canonical holdings (base model)

We begin with the simplest disciplined holding language: each case generates exactly one halfspace, mechanically determined by the outcome and the case facts. Given outcome $d_t$ at case $z_t$, the *canonical holding* is

$$H_t \ \equiv\ \big\{(w, c) \in \mathbb{R}^{k+1} : (-1)^{1-d_t}(w^\top z_t - c) \geq 0\big\}.$$

If $d_t = 1$ (plaintiff wins), the constraint is $w^\top z_t \geq c$: every admissible rule must classify $z_t$ for the plaintiff. If $d_t = 0$ (defendant wins), the constraint is $w^\top z_t \leq c$: no admissible rule may strictly favor the plaintiff at $z_t$.[1] This is the minimal holding consistent with the outcome—it constrains the feasible set to agree with $d_t$ at $z_t$, and nothing more.

Geometrically, $H_t$ is a halfspace in $(w, c)$-space with normal determined by the augmented case vector $\tilde{z}_t \equiv (z_t, -1) \in \mathbb{R}^{k+1}$ (or its negative), so the feasible-set update $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t$ is a single halfspace cut. After $T$ cases, the feasible set is the initial set intersected with $T$ halfspaces—a convex polytope whose facets are determined by past case fact vectors.

## 2.4 Horizontal Stare Decisis: Plausibility and Sanctions

Given $(z_t, \mathcal{F}_t)$, an outcome $d_t \in \{0, 1\}$ is *existentially plausible* iff it can be generated by at least one admissible rule:

$$\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t) \iff \exists (w, c) \in \mathcal{F}_t \text{ such that } d_t = d(z_t; w, c).$$

This matches the idea that, with multidimensional facts, "the law" does not pin down a unique $f$, but rules out outcomes that no plausible $f$ can justify.

### Plausibility characterization

The plausibility of each outcome at case $z$ is determined by two functions:

$$m_t(z) \equiv \min_{(w,c)\in\mathcal{F}_t} (w^\top z - c), \qquad M_t(z) \equiv \max_{(w,c)\in\mathcal{F}_t} (w^\top z - c).$$

Since $\mathcal{F}_t$ is compact and the objective is linear, both extrema are attained. Outcome 1 is plausible iff $M_t(z) \geq 0$ (some admissible rule classifies $z$ for the plaintiff); outcome 0 is plausible iff $m_t(z) < 0$

[1] Under the convention $d = \mathbf{1}\{w^\top z \geq c\}$, a rule exactly on the boundary ($w^\top z_t = c$) classifies $z_t$ as $d = 1$. The closed-halfspace holding for $d_t = 0$ therefore includes this boundary rule. Since the boundary has measure zero in parameter space and cases on the boundary have probability zero under continuous $G$, this does not affect any volume or probability statement. All results hold identically under the alternative convention $d = \mathbf{1}\{w^\top z > c\}$.

(some admissible rule classifies $z$ for the defendant). Case $z$ lies in the *penumbra* when both outcomes are plausible—equivalently, $m_t(z) < 0 \le M_t(z)$.

The *penumbra probability* under the case distribution $G$ is

$$P_t \;\equiv\; \Pr_{z\sim G}\big(m_t(z) < 0 \le M_t(z)\big).$$

This measures how much discretion current law leaves: $P_t$ close to 1 means nearly all cases are undetermined; $P_t$ close to 0 means the law is nearly determinate.

A holding $H_t$ is *feasible* at $\mathcal{F}_t$ iff $\mathcal{F}_t \cap H_t \neq \emptyset$.

If a judge chooses an outcome that is not existentially plausible, or issues an infeasible holding, peers impose a reputational/institutional sanction of size $K > 0$ (think: "clearly not law-following").

Formally, define the period-$t$ sanction indicator

$$\mathbf{S}_t \equiv \mathbf{1}\{\neg\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)\} \;+\; \mathbf{1}\{\mathcal{F}_t \cap H_t = \emptyset\},$$

and the sanction cost is $K \cdot \min\{1, \mathbf{S}_t\}$. (Under canonical holdings, a plausible outcome always generates a feasible holding, so the sanction reduces to $K \cdot \mathbf{1}\{\neg\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)\}$.)

### 2.5 Overruling

In period $t$, the judge may refuse to treat some subset of past holdings as binding. Model this as removing a (finite) subset of constraints from the current feasible set at a fixed cost $C > 0$ per removed holding.

To keep notation light, represent the law as an intersection of surviving holdings:

$$\mathcal{F}_t = \mathcal{F}_1 \cap \bigcap_{\tau<t,\ \tau\notin\mathcal{R}_t} H_\tau,$$

where $\mathcal{R}_t$ is the set of indices of overruled holdings (possibly empty). Overruling cost in period $t$ is $C \cdot |\Delta\mathcal{R}_t|$ where $\Delta\mathcal{R}_t$ are the newly overruled holdings at $t$.

**Remark.** Even with a fixed per-holding cost, overruling an early holding can indirectly create incompatibilities with later holdings that relied on it, because those later constraints remain in the intersection unless separately overruled. This captures the idea that "many rely on it" increases the practical cost.

### 2.6 Entailment: Holdings Must Follow from the Case (No Dicta)

A holding must be logically justified by the case—a court may not impose constraints on the law that go beyond what the outcome requires. This is the formal analog of the common-law prohibition on *dicta*.

Given $(z_t, \mathcal{F}_t, d_t)$, define the subset of admissible rules that deliver the chosen outcome:

$$\mathcal{F}_t(d_t; z_t) \equiv \{(w, c) \in \mathcal{F}_t : d(z_t; w, c) = d_t\}.$$

Note that $\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)$ iff $\mathcal{F}_t(d_t; z_t) \neq \emptyset$.

A holding $H_t$ is *entailed by the case* if (i) it is consistent with the chosen outcome—$\mathcal{F}_t(d_t; z_t) \cap H_t \neq \emptyset$, so that the holding does not contradict the decision itself—and (ii) $H_t \subseteq \{(w, c) :$

$d(z_t; w, c) = d_t\}$, so that the holding only constrains the feasible set in ways implied by the outcome at $z_t$.

Under canonical holdings, entailment is automatic: the halfspace $H_t$ contains exactly the rules that agree with $d_t$ at $z_t$. The judge's only instrument is the outcome $d_t$, and the holding follows mechanically. This is a feature, not a limitation: it isolates the pure geometric effect of constraint accumulation.

**Remark (richer holding languages).** In practice, courts do more than announce outcomes—they articulate rationales that constrain future cases beyond the minimum. An extension (to be developed) allows the judge to select an intended rule $(\hat{w}_t, \hat{c}_t) \in \mathcal{F}_t$ and issue a holding centered on that rationale, with a breadth parameter controlling how tightly the holding constrains the feasible set around the chosen rule. Under such a richer language, the citation incentive $\gamma$ and risk cost $\rho$ defined below become operative, and holdings are a strategic instrument distinct from the outcome itself. The formal treatment of this extension is deferred to a future draft; the base model uses canonical holdings throughout.

## 2.7 Breadth

Assume $\mathcal{F}_t$ is compact and has positive volume in $\mathbb{R}^{k+1}$. Define the *breadth* of a holding $H_t$ at $\mathcal{F}_t$ as

$$B(H_t; \mathcal{F}_t) \equiv \log\left(\frac{\mathrm{Vol}(\mathcal{F}_t)}{\mathrm{Vol}(\mathcal{F}_t \cap H_t)}\right),$$

with the convention that $B(H_t; \mathcal{F}_t) = +\infty$ if $\mathrm{Vol}(\mathcal{F}_t \cap H_t) = 0$ (which would also trigger feasibility sanctions). Larger $B$ means the holding is more constraining—it eliminates a larger share of the feasible set.

Under canonical holdings, breadth is not a choice variable: given outcome $d_t$ and case $z_t$, the halfspace $H_t$ is determined, and $B(H_t; \mathcal{F}_t)$ is a function of the geometry—how the hyperplane $w^\top z_t = c$ intersects $\mathcal{F}_t$. Cases whose fact vector $z_t$ cuts through the "center" of $\mathcal{F}_t$ generate broader holdings (more volume removed); cases near the boundary of the plausibility region generate narrower ones. This connects breadth directly to the informativeness of the case.

**Citation incentives and risk (extension).** When holdings are a strategic instrument (richer holding language), the judge faces a tradeoff: broader holdings are more likely to be cited (they constrain more future cases) but also more likely to be contested or overruled. We model this as a citation payoff $\gamma\, B(H_t; \mathcal{F}_t)$ and a risk cost $\rho\, B(H_t; \mathcal{F}_t)$, with $\gamma, \rho \geq 0$. This tradeoff is inert under canonical holdings (no discretion in $H_t$) but becomes the key strategic margin under richer holding languages where the judge chooses breadth.

## 2.8 Timing

At each period $t$:

1. Nature draws $(j_t, z_t)$ and reveals them publicly; $j_t$ observes $\mathcal{F}_t$.
2. Judge chooses an overrule set $\Delta\mathcal{R}_t$ (possibly empty), paying $C|\Delta\mathcal{R}_t|$, which updates $\mathcal{F}_t$ accordingly.
3. Judge chooses outcome $d_t \in \{0, 1\}$.
4. The canonical holding $H_t$ is determined by $(d_t, z_t)$.

5. The law updates: $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t$.
6. Sanctions (if any) are realized.

### 2.9 Per-Period Payoff

Given choices $(\Delta\mathcal{R}_t, d_t)$ and state $(j_t, z_t, \mathcal{F}_t)$, the per-period payoff under canonical holdings is

$$U_t \;=\; u_{j_t}(d_t, z_t) \;-\; C|\Delta\mathcal{R}_t| \;-\; K \cdot \mathbf{1}\{\neg\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)\}.$$

The judge trades off ideological preferences (choosing $d_t$ close to the ideal outcome) against sanctions for implausible outcomes and the cost of overruling. Since the holding is determined by the outcome, the only strategic choices are $d_t$ and $\Delta\mathcal{R}_t$.

### 2.10 Geometry of Constraint Accumulation

The following results are immediate consequences of the canonical-holding model.

**Proposition 1** (Plausibility and the penumbra)**.** *Let $\mathcal{F}_t$ be compact and convex. Then:*

(i) *$m_t(z)$ is concave and $M_t(z)$ is convex in $z$.*

(ii) *Outcome 1 is plausible at $z$ iff $M_t(z) \geq 0$; outcome 0 is plausible iff $m_t(z) < 0$.*

(iii) *The penumbra is the set $\{z : m_t(z) < 0 \leq M_t(z)\}$.*

*Proof sketch.* $M_t(z) = \max_{(w,c)\in\mathcal{F}_t}(w^\top z - c)$ is the support function of $\mathcal{F}_t$ evaluated at the direction $(z, -1)$, hence convex; $m_t(z) = -\max_{(w,c)\in\mathcal{F}_t}(-(w^\top z - c))$, hence concave. Parts (ii) and (iii) are immediate from the definition of existential plausibility. □

**Proposition 2** (Monotone determinacy)**.** *$\mathcal{F}_{t+1} \subseteq \mathcal{F}_t$ implies $m_{t+1}(z) \geq m_t(z)$ and $M_{t+1}(z) \leq M_t(z)$ for all $z$. Therefore the penumbra probability $P_t$ is weakly decreasing in $t$.*

*Proof sketch.* Minimizing over a smaller set raises the minimum; maximizing over a smaller set lowers the maximum. Since $P_t = \Pr_G(m_t(z) < 0 \leq M_t(z))$ and the event $\{m_t(z) < 0 \leq M_t(z)\}$ shrinks as $m_t$ rises and $M_t$ falls, $P_t$ is weakly decreasing. □

**Proposition 3** (Breadth and penumbra shrinkage)**.** *Define $\Delta P_t \equiv P_t - P_{t+1}$, the penumbra reduction from holding $H_t$. Under canonical holdings, $\Delta P_t \geq 0$. Moreover, if two feasible holdings $H, H'$ satisfy $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$ (i.e., $H$ is weakly more constraining), then $B(H;\mathcal{F}_t) \geq B(H';\mathcal{F}_t)$ and $\Delta P(H) \geq \Delta P(H')$: more constraining holdings produce weakly more breadth and weakly more penumbra shrinkage.*

*Proof sketch.* $\Delta P_t \geq 0$ follows from Proposition 2. For monotonicity: $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$ implies $\mathrm{Vol}(\mathcal{F}_t \cap H) \leq \mathrm{Vol}(\mathcal{F}_t \cap H')$, giving $B(H;\mathcal{F}_t) \geq B(H';\mathcal{F}_t)$. Applying Proposition 2 to the inclusion, the penumbra under $\mathcal{F}_t \cap H$ is a subset of the penumbra under $\mathcal{F}_t \cap H'$, so $\Delta P(H) \geq \Delta P(H')$. Full proof in Appendix. □

(The set-inclusion ordering $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$ is strictly stronger than volume ordering: a holding that removes more volume from $\mathcal{F}_t$ need not produce a subset of the post-holding feasible set, and the penumbra-shrinkage monotonicity does not hold under volume ordering alone.)

These three results hold regardless of how holdings are chosen (canonical or richer language). They establish the geometric spine: stare decisis makes law more determinate over time, with broader holdings accelerating the process.

### 2.11 Equilibrium Behavior and Doctrinal Drift

We now characterize equilibrium behavior and its dynamic consequences. Throughout this subsection, assume $K > \alpha$ (sanctions exceed the ideological gain from any single outcome), so that judges never willingly choose implausible outcomes absent overruling.

**Proposition 4** (Myopic outcome choice)**.** *Under canonical holdings with $K > \alpha$ and no overruling:*

(i) *If $z_t$ is in the penumbra ($m_t(z_t) < 0 \leq M_t(z_t)$), judge $j_t$ chooses the ideal outcome: $d_t = d(z_t; w_{j_t}, c_{j_t})$.*

(ii) *If only one outcome is plausible, the judge chooses that outcome regardless of ideology.*

(iii) *Every decision is lawful ($\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)$ holds), and the canonical holding $H_t$ updates the feasible set to $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t$.*

*Proof sketch.* The per-period payoff is $U_t = \alpha \cdot \mathbf{1}\{d_t = d(z_t; w_j, c_j)\} - K \cdot \mathbf{1}\{\neg\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)\}$. Choosing an implausible outcome yields at most $\alpha - K < 0$; choosing a plausible outcome yields at least 0. In the penumbra both outcomes are plausible and sanction-free, so the judge maximizes ideology. □

The simple structure of Proposition 4 is what makes the drift result sharp: every decision is individually lawful, yet the cumulative effect on $\mathcal{F}_t$ depends on which judges happen to decide which cases.

**Proposition 5** (Drift and path dependence)**.** *Under canonical holdings with myopic ideal-outcome choice (Proposition 4):*

(i) ***Monotone shrinkage.*** *$\mathcal{F}_{t+1} \subseteq \mathcal{F}_t$ for all $t$. If $z_t$ lies in the open penumbra ($m_t(z_t) < 0 < M_t(z_t)$), $\mathrm{Vol}(\mathcal{F}_{t+1}) < \mathrm{Vol}(\mathcal{F}_t)$.*

(ii) ***Path dependence.*** *The feasible set $\mathcal{F}_T$ depends on the order in which cases arrive, not only on the multiset of cases and judges.*

(iii) ***Ideological drift.*** *If the distribution of judge ideal rules is asymmetric—e.g., plaintiff-favoring judges are more likely—and $\mathcal{F}_t$ is symmetric about the cutting hyperplane (as holds when $\mathcal{F}_1$ is centrally symmetric and $G$ is symmetric), then the expected barycenter of $\mathcal{F}_t$ drifts toward the dominant ideology.*

*Proof sketch. (i)* Each canonical holding is a halfspace containing $\mathcal{F}_t(d_t; z_t) \subsetneq \mathcal{F}_t$ (strict when both outcomes are plausible), so the intersection is strict.

*(ii)* Constructive example in $k = 1$. Let $\mathcal{F}_1 = \{(w, c) : |w| \leq 1,\ |c| \leq 1\}$, and consider two penumbra cases $z_1 = 0.5$, $z_2 = 0.8$ with judges $j_A$ (ideal $(1, 0)$, plaintiff-leaning) and $j_B$ (ideal $(-1, 0)$, defendant-leaning). Both cases lie in the penumbra of $\mathcal{F}_1$: at $z_1 = 0.5$, $M_1(0.5) = 1.5 > 0$ and $m_1(0.5) = -1.5 < 0$; similarly for $z_2$.

- *Order 1:* $j_A$ decides $z_1 = 0.5$ (plaintiff wins, $H_1 = \{0.5w \geq c\}$), then $j_B$ decides $z_2 = 0.8$ (defendant wins, $H_2 = \{0.8w \leq c\}$). Intersection: $\mathcal{F}_3 = \{(w, c) : |w| \leq 1,\ 0.8w \leq c \leq 0.5w\}$. The constraint $0.8w \leq 0.5w$ forces $w \leq 0$: only defendant-leaning rules survive.

- *Order 2:* $j_B$ decides $z_1 = 0.5$ (defendant wins, $H_1' = \{0.5w \leq c\}$), then $j_A$ decides $z_2 = 0.8$ (plaintiff wins, $H_2' = \{0.8w \geq c\}$). Intersection: $\mathcal{F}_3' = \{(w, c) : |w| \leq 1,\ 0.5w \leq c \leq 0.8w\}$. The constraint $0.5w \leq 0.8w$ forces $w \geq 0$: only plaintiff-leaning rules survive.

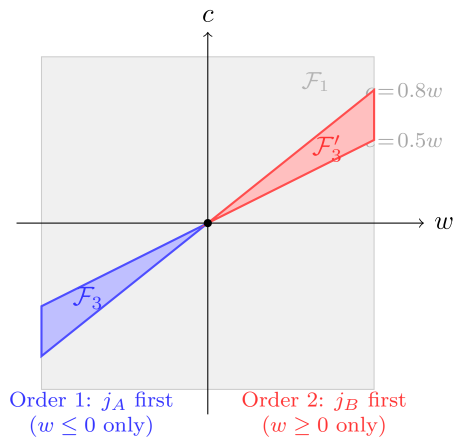


Figure 1: Path dependence in the constructive example of Proposition 5(ii). Both feasible sets start from $\mathcal{F}_1 = \{(w, c) : |w| \leq 1, |c| \leq 1\}$ (gray square). In Order 1, a plaintiff-leaning judge decides $z_1 = 0.5$ first, then a defendant-leaning judge decides $z_2 = 0.8$; the surviving set $\mathcal{F}_3$ (blue) contains only defendant-leaning rules ($w \leq 0$). In Order 2, the judges decide in reverse order; $\mathcal{F}_3'$ (red) contains only plaintiff-leaning rules ($w \geq 0$). The two sets are disjoint except at the origin: the order of cases determines the direction of the law.

Same cases, same judges, different order: the feasible sets $\mathcal{F}_3$ and $\mathcal{F}_3'$ are essentially disjoint (intersecting only at the single point $(0, 0)$). The direction of the law is entirely determined by the order of cases (Figure 1).

*(iii)* In penumbra cases, the outcome follows the judge's ideology (Proposition 4(i)). If plaintiff-favoring judges are more frequent, the halfspace cut $\{w^\top z_t \geq c\}$ is chosen more often than $\{w^\top z_t \leq c\}$, systematically removing defendant-leaning rules from $\mathcal{F}_t$. The expected barycenter shifts toward the dominant ideology at a rate proportional to the asymmetry in the judge distribution. □

**Remark 1** (Slippery slopes without lawlessness)**.** *Proposition 5 formalizes the "slippery slope" concern: a sequence of individually lawful decisions can cumulatively transform the feasible set so that outcomes previously considered implausible become the only plausible option. Crucially, no single judge violates the law at any step; the drift is an emergent property of the interaction between judicial ideology and the geometry of constraint accumulation.*

### 2.12 Overruling Threshold

**Proposition 6** (Overruling threshold)**.** *Suppose judge $j_t$'s ideal outcome $d_t^* \equiv d(z_t; w_{j_t}, c_{j_t})$ is not plausible under $\mathcal{F}_t$. Let*

$$W^* \equiv \min\Big\{|S| : S \subseteq \{1, \ldots, t-1\},\ \ \mathcal{F}_1 \cap \bigcap_{\tau \notin S} H_\tau \text{ contains a rule producing } d_t^* \text{ at } z_t\Big\}$$

*be the minimum number of past holdings whose removal restores plausibility. Then:*

(i) *The judge overrules iff $\alpha > C \cdot W^*$.*

(ii) *Under canonical holdings, $W^*$ equals the minimum number of past halfspace constraints that must be removed to admit a rule classifying $z_t$ as $d_t^*$.*

(iii) *For a fixed case $z$ and desired outcome $d^*$, $W^*$ is weakly increasing along histories without overruling: as holdings accumulate, overruling to restore a given outcome becomes weakly harder.*

*Proof sketch. (i)* Without overruling, the judge must choose the (unique) plausible outcome, receiving payoff 0 from ideology. Overruling $W^*$ holdings costs $C \cdot W^*$ and yields ideology payoff $\alpha$. The judge overrules iff $\alpha - C \cdot W^* > 0$.

*(ii)* Under canonical holdings each past holding is one halfspace, so removing a holding removes one facet of the polytope $\mathcal{F}_t$.

*(iii)* New holdings can only add halfspace constraints. If outcome $d^*$ at case $z$ is blocked at time $t$, any additional halfspace at $t+1$ that is inconsistent with $d^*$ at $z$ adds a new constraint that must also be removed, so $W^*$ can only increase. □

### 2.13 Precedent Complexity and the VC Dimension

The version-space structure of canonical holdings connects directly to statistical learning theory.

**Corollary 1** (Precedent complexity)**.** *Under canonical holdings, $\mathcal{F}_t$ is the version space of affine classifiers on $\mathbb{R}^k$ consistent with past case outcomes. The VC dimension of this class is $d_{\mathrm{VC}} = k+1$. Therefore:*

(i) ***Shattering.*** *There exist configurations of $k+1$ cases with fact vectors in general position such that any assignment of outcomes is consistent with some rule in $\mathcal{F}_1$. No configuration of $k+2$ cases has this property.*

(ii) ***Sample complexity.*** *If all outcomes were generated by a single fixed rule, then after $t = O\big(\frac{k+\log(1/\delta)}{\varepsilon}\big)$ i.i.d. cases, every surviving rule in $\mathcal{F}_t$ classifies at least a $(1-\varepsilon)$-fraction of future cases identically to the generating rule, with probability at least $1-\delta$.*

(iii) ***Multidimensionality.*** *Higher-dimensional fact spaces (larger $k$) require proportionally more precedent to achieve a given level of legal determinacy.*

*Proof sketch.* The VC dimension of affine (linear threshold) classifiers on $\mathbb{R}^k$ is $k+1$: the $k+1$ standard basis vectors in $\mathbb{R}^{k+1}$ (augmented fact space) can be shattered, while no $k+2$ points can be. Part (ii) follows from the fundamental theorem of statistical learning applied to the version space; since all labels are consistent with $\mathcal{F}_t$, any surviving hypothesis has generalization error at most $\varepsilon$ after $O(d_{\mathrm{VC}}/\varepsilon)$ samples. Part (iii) is immediate from the linear dependence of the sample complexity bound on $k$. □

**Interpretation.** In the language of the model, Corollary 1 says that Equal Protection doctrine ($k$ large—many legally salient dimensions) requires many more precedents to pin down the law than a simple threshold rule ($k = 1$). This formalizes the intuition that multidimensional legal questions resist convergence: not because judges are lawless, but because the constraint language is insufficiently expressive relative to the complexity of the fact space.

### 2.14 Comparative Statics

The propositions above yield several comparative statics that distinguish the model from alternatives.

**Dimension and determinacy.** Corollary 1 implies that higher-dimensional doctrinal areas (larger $k$) require proportionally more precedent to achieve legal determinacy. For fixed $t$, the penumbra probability $P_t$ is weakly larger when $k$ is larger: adding legally relevant dimensions expands the space of admissible rules without adding constraints, widening the gap between existing holdings and determinacy. Concretely, a one-dimensional rule ($k = 1$, e.g., a bright-line age threshold) can be pinned by two cases, while a five-factor balancing test ($k = 5$) requires at least six.

**Doctrinal entrenchment.** By Proposition 6(iii), the minimum hitting set $W^*$ for any fixed case and desired outcome is weakly increasing over time. Since overruling requires $\alpha > C \cdot W^*$ (Proposition 6(i)), the fraction of judges willing to overrule any given precedent decreases as holdings accumulate, making established doctrine progressively harder to reverse. This captures the common-law intuition that "old precedent binds more strongly" as a consequence of constraint geometry, without explicitly modeling reliance or reputation.

**Bench composition and drift direction.** Proposition 5(iii) implies that the direction of doctrinal drift depends on the composition of the bench. A symmetric bench (equal likelihood of plaintiff- and defendant-favoring judges for each penumbra case) produces zero expected drift in the barycenter of $\mathcal{F}_t$, while asymmetry produces drift proportional to the imbalance. This provides a mechanism for the empirical finding that changes in bench composition predict changes in legal doctrine, even when individual judges follow the law at every step.

**Sanctions and the binding-force threshold.** The assumption $K > \alpha$ ensures universal lawfulness (Proposition 4). As $K/\alpha$ decreases toward 1, the binding-force mechanism weakens: lawfulness holds by strict dominance when $K > \alpha$ but breaks down when $K \leq \alpha$, as judges become willing to absorb sanctions for ideological gain. This suggests that institutional features raising the salience or magnitude of horizontal sanctions—opinion publication, panel unanimity norms, appellate review—are necessary for the binding-force mechanism to operate.

### 2.15 Discussion

- **Binding force.** Existential plausibility plus large $K$ implies judges cannot select outcomes that no admissible $(w, c)$ can justify.
- **Discretion in the penumbra.** Because $\mathcal{F}_t$ is a *set* of admissible affine rules, many outcomes remain plausible for many $z_t$; judges can select among them in ways that reflect their ideal $(w_j, c_j)$.
- **No mechanical convergence.** In $k \geq 2$, intersecting $\mathcal{F}_t$ with halfspaces typically leaves a non-singleton polytope; the law becomes more structured without collapsing to a point absent very strong informational/expressiveness assumptions. Corollary 1 makes this precise: $O(k/\varepsilon)$ cases are needed to reduce disagreement among admissible rules below $\varepsilon$.
- **Drift and slippery slopes.** Each individually plausible decision shrinks and tilts $\mathcal{F}_{t+1}$, thereby changing which outcomes are plausible tomorrow (Proposition 5). The feasible set drifts even though every decision is lawful (Remark 1).
- **Overruling prevents inconsistency collapse.** Because halfspace constraints can accumulate to empty the feasible set in high dimensions, the ability to overrule at cost $C$ keeps $\mathcal{F}_t$ nonempty while preserving that overruling is exceptional (Proposition 6).

# 3 Equal Protection as Constraint Accumulation in a Multidimensional Legal Space

This section illustrates how the model maps onto Equal Protection jurisprudence. Equal Protection is a natural testing ground for the framework because it is explicitly multidimensional, relies heavily on precedent and reasoning rather than mechanical rules, and exhibits both strong claims of constraint and persistent ideological disagreement. We show how canonical Equal Protection cases can be understood as imposing linear constraints on a high-dimensional decision boundary, how those constraints bind future courts without fully determining outcomes, and how doctrinal drift emerges endogenously.

## 3.1 Equal Protection as a Multidimensional Decision Problem

Equal Protection cases rarely turn on a single factual dimension. Instead, courts consider a constellation of legally salient features, including (among others):

- the nature of the classification (race, sex, alienage, wealth, etc.),
- the governmental interest asserted,
- the strength and generality of that interest,
- the degree of tailoring or fit,
- historical patterns of discrimination,
- evidence of discriminatory purpose or animus,
- the institutional context (education, voting, criminal justice, etc.).

In the language of the model, each case is represented by a fact vector $z \in \mathbb{R}^k$, where different coordinates correspond to these features. The constitutional outcome (upholding or invalidating the challenged law) is determined by an underlying decision boundary $f(z) = w^\top z - c$. Crucially, the law does not specify $(w, c)$ uniquely. Instead, existing precedent defines a feasible set $\mathcal{F}_t$ of admissible $(w, c)$ pairs—that is, a set of plausible ways of weighing and trading off the relevant factors. With at least seven salient dimensions ($k \geq 7$), Corollary 1 implies that at least eight landmark cases would be needed to fully pin down the law—and many more to achieve near-determinacy across the full range of possible disputes.

## 3.2 Holdings as Constraints on Admissible Reasoning

In this framework, the binding force of Equal Protection precedent arises from holdings that restrict the admissible shape of $f$. A holding does not merely announce an outcome; it rules out certain ways of justifying outcomes in future cases. We illustrate this with a sequence of canonical cases.

**Brown v. Board of Education.** The holding of *Brown* that state-imposed racial segregation in public education violates Equal Protection is often described as categorical. Yet in the model it is best understood as imposing powerful but partial constraints. The holding rules out admissible decision rules that treat racial classification as neutral or benign in the educational context. Formally, it removes from $\mathcal{F}_t$ all $(w, c)$ for which the weight on race is weak or positive when applied to public education—a halfspace cut that shrinks the feasible set and reduces the penumbra (Proposition 2).

At the same time, it does not specify a complete theory of Equal Protection: it leaves open how race interacts with other dimensions, how far the principle extends beyond education, and how strong countervailing interests might matter in other contexts.

**Strict scrutiny and Korematsu.** Subsequent cases articulated the idea that racial classifications are "suspect" and subject to strict scrutiny. In the model, this does not amount to a decision tree that mechanically invalidates all race-based laws. Instead, it introduces conditional constraints on admissible $(w, c)$: when the race dimension is active, admissible decision rules must assign race a large negative weight and permit constitutionality only if other dimensions (such as the strength of the governmental interest and the degree of tailoring) receive sufficiently large positive weight. The controversial holding in *Korematsu* illustrates that even under strict scrutiny, the feasible set $\mathcal{F}_t$ was not empty: upholding the internment order was existentially plausible ($M_t(z) > 0$ in the notation of Proposition 1), though only at an extreme point of the admissible set—a move later courts would treat as erroneous or exceptional.

**Grutter v. Bollinger and narrow tailoring.** *Grutter* further refined the constraints associated with strict scrutiny, holding that diversity in higher education can constitute a compelling interest and specifying requirements of individualized consideration and limited use of race. In the model, this adds new linear constraints linking the race dimension to dimensions capturing governmental interest and tailoring. These constraints shrink $\mathcal{F}_t$ and reduce the penumbra (Proposition 3) but do not collapse it to a singleton. Different judges may still disagree about how demanding narrow tailoring must be, yet they are bound to reason within a constrained space shaped by prior holdings.

**Parents Involved and ideological disagreement.** Later cases such as *Parents Involved in Community Schools* demonstrate how ideological disagreement operates within this constrained space. The Justices largely accepted the existing constraint structure—that race is suspect and that compelling interest and tailoring matter—but disagreed sharply over how those constraints should be applied. In model terms, the case lay in the penumbra (Proposition 1(iii)), and the Justices selected different admissible $(w, c)$ within $\mathcal{F}_t$ according to their ideal rules (Proposition 4(i))—reflecting different preferred locations of the decision boundary rather than outright rejection of precedent.

**SFFA and regime shift through overruling.** The most recent major development, *Students for Fair Admissions v. Harvard* (2023), illustrates the model's overruling mechanism. The Court effectively rejected the permissive reading of strict scrutiny that had sustained race-conscious admissions from *Grutter* onward. In model terms, this corresponds to removing constraints from $\mathcal{F}_t$ at cost $C$: the holdings in *Grutter* and its progeny—that educational diversity constitutes a compelling interest and that holistic, individualized review can satisfy narrow tailoring—were treated as no longer binding. The decision did not simply add a new constraint to $\mathcal{F}_t$; it expanded the feasible set in the dimension of permissible governmental interests while simultaneously contracting it in the dimension of means-ends fit. This combination of constraint removal and constraint addition captures the distinctive character of overruling in the model: the feasible set is reshaped, not merely shrunk, and the institutional cost $C$ reflects the reputational and coordination consequences of departing from settled doctrine. In the language of Proposition 6, the majority's ideological gain $\alpha$ from eliminating race-conscious admissions exceeded $C \cdot W^*$, the cost of removing the $W^*$ holdings that had sustained *Grutter*'s framework.

### 3.3 Horizontal Stare Decisis and the Role of Holdings

These examples illustrate horizontal stare decisis as modeled here. A judge deciding a new Equal Protection case must choose an outcome that is existentially plausible given $\mathcal{F}_t$; the canonical holding then constrains all admissible rules to agree with that outcome at the case's facts. Distinguishing a precedent corresponds to a case whose facts lie near the boundary of the plausible region, so the resulting halfspace cut removes little from $\mathcal{F}_t$; extending a precedent corresponds to a case that cuts deeper into the feasible set, further constraining future doctrine. Refusing to follow a past holding requires explicitly removing constraints at a cost, capturing the exceptional nature of overruling.

### 3.4 Doctrinal Drift and Slippery Slopes

The model also explains gradual doctrinal drift in Equal Protection law (Proposition 5). Each new holding slightly tilts or shrinks $\mathcal{F}_t$, thereby changing which future outcomes are existentially plausible. Decisions that are plausible today can make outcomes plausible tomorrow that would not have been plausible earlier—and, critically, the resulting feasible set depends on the *order* in which cases arise, not just their content (Proposition 5(ii)). This mechanism captures familiar "slippery slope" concerns in Equal Protection debates, such as whether recognizing one compelling interest or one form of tailoring will inevitably expand or contract the scope of strict scrutiny over time.

Under a richer holding language where judges choose how broadly to formulate their rationales, this dynamic is reinforced: broader holdings exert greater influence on the future shape of $\mathcal{F}_t$ but also increase the risk of being challenged or overruled. Even under canonical holdings, however, the drift mechanism operates purely through the geometry of constraint accumulation, explaining why Equal Protection doctrine evolves through incremental, contested steps rather than through either complete stasis or wholesale revision.

### 3.5 Summary

Viewed through the lens of the model, Equal Protection jurisprudence is neither a collection of mechanical rules nor a sequence of unconstrained ideological choices. It is a process of constraint accumulation in a high-dimensional legal space. Holdings bind by ruling out entire classes of justificatory rules, yet they preserve discretion by leaving multiple admissible decision boundaries. This structure explains how Equal Protection law can simultaneously exhibit strong claims of legal constraint, deep ideological disagreement in hard cases, and gradual doctrinal drift over time.

## 4 Due Process as Open-Textured Constraint Accumulation

Due Process jurisprudence provides a complementary and, in many ways, even clearer illustration of the model. If Equal Protection highlights how law can bind through structured categories and tiers, Due Process shows how binding force can arise despite persistent doctrinal openness, contested tests, and explicit balancing. Few areas of constitutional law are as openly multidimensional, normatively charged, and resistant to convergence as Due Process, making it a particularly demanding setting for any theory of legal constraint.

### 4.1 Due Process as a High-Dimensional Legal Space

Both procedural and substantive Due Process cases involve a wide array of legally relevant considerations. For procedural due process, courts routinely consider:

- the nature and importance of the private interest at stake,
- the risk of erroneous deprivation,
- the value of additional procedural safeguards,
- the government's interest, including fiscal and administrative burdens,
- the institutional context in which the deprivation occurs.

For substantive due process, relevant dimensions include:

- the nature of the asserted liberty interest,
- its historical grounding and traditional recognition,
- the level of generality at which the interest is described,
- the strength and type of governmental justification,
- the degree of intrusion on personal autonomy.

In the model, these considerations form the coordinates of a fact vector $z \in \mathbb{R}^k$. Constitutional validity is determined by an underlying decision boundary $f(z) = w^\top z - c$, but, as in Equal Protection, neither the text of the Due Process Clauses nor precedent specifies $(w, c)$ uniquely. Instead, the law consists of a feasible set $\mathcal{F}_t$ of admissible decision rules consistent with past holdings.

### 4.2 Holdings and the Structure of Procedural Due Process

The canonical example of procedural due process doctrine is the balancing test articulated in *Mathews v. Eldridge*. The Court identified three factors to be weighed: the private interest affected, the risk of erroneous deprivation and value of additional procedures, and the government's interest. In the language of the model, *Mathews* did not define a determinate rule for when procedures are constitutionally required. Rather, it imposed structural constraints on admissible decision boundaries.

Specifically, the holding restricts $\mathcal{F}_t$ by requiring that admissible $(w, c)$ assign positive weight to each of the three factors and permit tradeoffs among them. It rules out decision rules that ignore error risk entirely, that treat governmental convenience as dispositive, or that mechanically require trial-type procedures in all cases. At the same time, it leaves substantial freedom regarding the relative magnitudes of these weights, thereby preserving discretion in application.

Subsequent procedural due process cases can be understood as adding further linear constraints. For example, holdings emphasizing the importance of notice and opportunity to be heard restrict admissible rules that place insufficient weight on error risk. Decisions limiting the reach of due process in contexts such as prison administration or welfare benefits add constraints linking procedural requirements to institutional context. None of these holdings collapse $\mathcal{F}_t$ to a single rule; instead, they gradually shape the space of admissible balancing regimes.

### 4.3 Substantive Due Process and Persistent Underidentification

Substantive Due Process illustrates even more starkly why convergence is neither expected nor desirable in a multidimensional legal system. Cases involving economic regulation, family autonomy, bodily integrity, and sexual privacy raise fundamentally different combinations of factual and normative dimensions. Early twentieth-century cases associated with economic liberty imposed constraints emphasizing freedom of contract, while later repudiation of that line of cases removed or reversed those constraints at significant institutional cost.

Modern substantive due process cases impose different kinds of restrictions. Holdings that require a liberty interest to be "deeply rooted in this Nation's history and tradition" constrain admissible decision rules by linking the weight on liberty interests to historical evidence. Other cases emphasize personal autonomy or dignity, imposing different constraints on the shape of $f$. These constraints coexist uneasily, leaving $\mathcal{F}_t$ broad and contested. Judges who disagree about how Due Process should operate often disagree not about outcomes alone, but about which dimensions should carry decisive weight and how abstractly liberty interests should be characterized.

### 4.4 Horizontal Stare Decisis and Doctrinal Contestation

Within this framework, horizontal stare decisis in Due Process operates through constraints on reasoning rather than through fixed rules. A judge deciding a new case must select an outcome that is existentially plausible given $\mathcal{F}_t$; the resulting canonical holding then constrains all admissible rules to agree with that outcome at the case's facts. Claims that a prior case was "wrongly decided," that a test was "unworkable," or that a line of cases should be limited to its facts correspond to explicit or implicit removal of constraints at a cost.

Because Due Process doctrine is highly multidimensional, conflicts among constraints are common. Allowing for overruling or limiting past holdings prevents the feasible set from becoming empty while preserving the idea that such moves are exceptional and costly. The model thus captures why Due Process doctrine is simultaneously stable enough to constrain judges and fluid enough to accommodate deep and persistent disagreement.

### 4.5 Drift, Slippery Slopes, and the Evolution of Due Process

The model naturally accounts for the gradual evolution of Due Process jurisprudence. A decision that expands recognition of a liberty interest, or that relaxes the historical grounding requirement, can be represented as a holding that tilts or shrinks $\mathcal{F}_t$. That change may render future outcomes existentially plausible that previously were not, enabling further doctrinal movement without any single decision appearing lawless. Conversely, decisions emphasizing restraint or institutional deference can shift the feasible set in the opposite direction.

This dynamic explains recurrent concerns about slippery slopes in Due Process cases. Each step may be defensible within existing constraints, yet the cumulative effect can substantially alter the space of admissible decision rules. The risk of sanction or overruling discourages overly aggressive moves, while the geometry of constraint accumulation ensures that even individually cautious decisions can cumulate into significant doctrinal shifts. The resulting dynamics feature incremental change, contested boundaries, and enduring underidentification of "the law."

### 4.6 Summary

Due Process jurisprudence exemplifies how law can bind without converging. Through holdings that impose partial, multidimensional constraints on admissible decision rules, courts coordinate on a shared space of plausible legal reasoning. That space constrains outcomes and justifications, yet remains broad enough to permit ideological disagreement and gradual evolution. The model thus explains why Due Process law appears at once indeterminate, contentious, and meaningfully binding—a combination that has long puzzled both legal theorists and social scientists.

# 5 Next Steps: Closing the Model, Empirical Tests, and Policy Implications

The framework developed in this paper is intentionally parsimonious and foundational. Its purpose is to isolate a small number of mechanisms—constraint accumulation through holdings, existential plausibility, and horizontal sanctions—that together generate binding force, discretion, and drift. Several natural next steps follow. This section outlines a research agenda for closing the model, deriving testable predictions, confronting those predictions with data, and drawing policy implications.

## 5.1 Closing the Model

The current framework leaves several components deliberately open in order to maintain generality. A first step is to "close" the model by specifying these components more tightly in particular institutional settings.

One direction is to enrich the holding language beyond canonical holdings. In the base model, each case generates a single halfspace constraint determined by the outcome and the case facts. A natural extension allows judges to select an intended rationale and choose how broadly to formulate the holding around it, introducing a strategic breadth dimension. Under such a richer language, the constraint language could be limited to specific doctrinal forms—such as sign restrictions, relative-weight constraints, or threshold conditions—that correspond closely to actual judicial reasoning, allowing sharper characterizations of equilibrium holdings and the evolution of the feasible set.

A second direction is to endogenize overruling and sanctioning more fully. In the present model, sanctions are direct and automatic, and overruling carries a fixed cost. A richer model could make sanctions the outcome of coordination among peers, panels, or higher courts, allowing the cost of overruling to depend on factors such as the age of a precedent, the number of subsequent citations, or the degree of reliance by other holdings. This would bring the model closer to observed judicial practice while preserving its core logic.

## 5.2 Testable Predictions

Even in its current form, the model generates a number of testable predictions that distinguish it from both purely attitudinal models and models that treat law as fully determinate.

First, the model predicts that ideological effects on judicial decisions should be strongest in cases where the feasible set $\mathcal{F}_t$ is large—that is, where doctrine leaves many admissible decision boundaries (Proposition 4). Conversely, ideology should matter less in doctrinally settled areas where accumulated holdings have substantially shrunk $\mathcal{F}_t$ (Proposition 2). This suggests that measures of doctrinal clarity or constraint should moderate the relationship between judicial ideology and outcomes.

Second, the model predicts systematic variation in overruling. Proposition 6 implies that overruling is more likely when (a) the judge's ideological stake $\alpha$ is large relative to the overruling cost $C$, and (b) the minimum hitting set $W^*$ is small—that is, when few past holdings block the desired outcome. Overruling should therefore cluster in doctrinal areas where a small number of pivotal precedents constrain the feasible set, rather than in areas with many overlapping holdings.

Third, the model predicts path dependence and drift (Proposition 5). Early holdings in a line of cases should have outsized effects on the future shape of doctrine, even when later judges disagree with them. Doctrinal change should often occur through a sequence of individually plausible steps rather than through abrupt reversals. The VC dimension result (Corollary 1) further predicts that higher-dimensional doctrinal areas require proportionally more cases to achieve the same level of determinacy.

## 5.3 Empirical Strategies

Testing these predictions requires moving beyond simple outcome-based analyses of judicial voting. Because the model emphasizes holdings and reasoning, not just results, empirical work should focus on the structure and content of opinions.

One promising approach is to measure the constraining effect of holdings using text-based methods. Advances in natural language processing make it feasible to identify what doctrinal constraints an opinion articulates and to track how those constraints are cited and applied in subsequent cases. Such measures can serve as empirical proxies for the model's feasible-set shrinkage (Proposition 3) and penumbra reduction.

Another approach is to exploit quasi-experimental variation in judicial assignment. Random assignment of judges to panels or cases can be used to study how differences in judicial ideology affect the evolution of doctrine, not only case outcomes. The model predicts that ideological differences should manifest most strongly in outcomes when doctrine is unsettled—that is, when $\mathcal{F}_t$ is large and many outcomes remain plausible (Proposition 4).

Finally, longitudinal analyses of specific doctrinal areas—such as Equal Protection or Due Process—can trace how the space of admissible reasoning expands, contracts, or tilts over time. The model provides a conceptual map for interpreting such patterns as changes in the feasible set of decision rules rather than as shifts in a single underlying "law."

## 5.4 Policy and Institutional Implications

The framework also has implications for institutional design and legal policy. If the binding force of law depends on the accumulation and enforcement of constraints on reasoning, then institutions that affect clarity, observability, and sanctioning will influence how strongly law binds.

For example, practices that promote clear articulation of holdings—such as reason-giving requirements, publication norms, and structured opinions—may strengthen coordination among judges but also accelerate doctrinal drift by making broader constraints easier to articulate and cite. Conversely, mechanisms that raise the cost of overruling or narrowing precedent may stabilize doctrine at the expense of adaptability.

The model also suggests that concerns about judicial ideology cannot be addressed solely by focusing on outcomes. Institutional reforms that affect how holdings are written, aggregated, and enforced may be at least as important for shaping the evolution of the law. More broadly, the framework cautions against viewing legal indeterminacy as a flaw to be eliminated. Some degree of underidentification may be essential for maintaining both the binding force and the adaptability of legal systems.

### 5.5 Conclusion

This paper offers a first step toward a positive theory of the binding force of law that takes multidimensionality, precedent, and judicial reasoning seriously. Closing the model, testing its predictions, and exploring its policy implications are natural next steps. Together, these avenues promise to deepen our understanding of how law constrains judges, how it changes over time, and how legal institutions can be designed to balance stability with principled evolution.

# A Proofs

This appendix provides complete proofs for all propositions and the corollary stated in the main text. Throughout, write $\tilde{z} = (z, -1) \in \mathbb{R}^{k+1}$ for the augmented case vector, so that $w^\top z - c = (w, c)^\top \tilde{z}$ for any $(w, c) \in \mathbb{R}^{k+1}$.

*Proof of Proposition 1. (i)* $M_t(z) = \max_{(w,c)\in\mathcal{F}_t}(w, c)^\top \tilde{z}$ is the support function $\sigma_{\mathcal{F}_t}(\tilde{z})$ of the compact convex set $\mathcal{F}_t$. The support function of any convex set is convex (Rockafellar, 1970, Thm. 13.2). Since $z \mapsto \tilde{z} = (z, -1)$ is affine, $M_t(z) = \sigma_{\mathcal{F}_t}(z, -1)$ is convex in $z$ as the composition of a convex function with an affine mapping.

For $m_t(z)$: $m_t(z) = \min_{\mathcal{F}_t}(w, c)^\top \tilde{z} = -\max_{\mathcal{F}_t}(w, c)^\top(-\tilde{z}) = -\sigma_{\mathcal{F}_t}(-z, 1)$. Since $z \mapsto (-z, 1)$ is affine and $\sigma_{\mathcal{F}_t}$ is convex, $\sigma_{\mathcal{F}_t}(-z, 1)$ is convex in $z$, so $m_t(z) = -\sigma_{\mathcal{F}_t}(-z, 1)$ is concave.

*(ii)* Outcome 1 is plausible iff $\exists\,(w, c) \in \mathcal{F}_t$ with $w^\top z - c \geq 0$, iff $\max_{\mathcal{F}_t}(w^\top z - c) \geq 0$, iff $M_t(z) \geq 0$. Outcome 0 is plausible iff $\exists\,(w, c) \in \mathcal{F}_t$ with $w^\top z - c < 0$, iff $\min_{\mathcal{F}_t}(w^\top z - c) < 0$, iff $m_t(z) < 0$.

*(iii)* Both outcomes are plausible iff $M_t(z) \geq 0$ and $m_t(z) < 0$, giving the penumbra $\{z : m_t(z) < 0 \leq M_t(z)\}$. □

*Proof of Proposition 2.* Let $\mathcal{F}_{t+1} \subseteq \mathcal{F}_t$. Minimizing a linear objective over a subset yields a weakly larger value:

$$m_{t+1}(z) = \min_{x\in\mathcal{F}_{t+1}} x^\top \tilde{z} \;\geq\; \min_{x\in\mathcal{F}_t} x^\top \tilde{z} = m_t(z).$$

Maximizing over a subset yields a weakly smaller value: $M_{t+1}(z) \leq M_t(z)$.

For $P_t$: suppose $z$ is in the penumbra at $t{+}1$, i.e., $m_{t+1}(z) < 0 \leq M_{t+1}(z)$. Then $m_t(z) \leq m_{t+1}(z) < 0$ and $M_t(z) \geq M_{t+1}(z) \geq 0$, so $z$ is in the penumbra at $t$. The time-$(t{+}1)$ penumbra is thus a subset of the time-$t$ penumbra, and $P_{t+1} \leq P_t$. □

*Proof of Proposition 3.* That $\Delta P_t \geq 0$ follows immediately from Proposition 2, since $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t \subseteq \mathcal{F}_t$ implies $P_{t+1} \leq P_t$.

For the monotonicity claim, we prove a precise version under the partial order of set inclusion on post-holding feasible sets. Let $H, H'$ be two feasible holdings at $\mathcal{F}_t$ with $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$.

*Breadth:* $\mathrm{Vol}(\mathcal{F}_t \cap H) \leq \mathrm{Vol}(\mathcal{F}_t \cap H')$, so

$$B(H; \mathcal{F}_t) = \log \frac{\mathrm{Vol}(\mathcal{F}_t)}{\mathrm{Vol}(\mathcal{F}_t \cap H)} \;\geq\; \log \frac{\mathrm{Vol}(\mathcal{F}_t)}{\mathrm{Vol}(\mathcal{F}_t \cap H')} = B(H'; \mathcal{F}_t).$$

*Penumbra shrinkage:* Apply Proposition 2 to the inclusion $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$: the penumbra under $\mathcal{F}_t \cap H$ is a subset of the penumbra under $\mathcal{F}_t \cap H'$. Therefore $P(\mathcal{F}_t \cap H) \leq P(\mathcal{F}_t \cap H')$, and

$$\Delta P(H) = P_t - P(\mathcal{F}_t \cap H) \;\geq\; P_t - P(\mathcal{F}_t \cap H') = \Delta P(H').$$

Note: the set-inclusion ordering is stronger than volume ordering—$\mathrm{Vol}(\mathcal{F}_t \cap H) \leq \mathrm{Vol}(\mathcal{F}_t \cap H')$ does not in general imply $\mathcal{F}_t \cap H \subseteq \mathcal{F}_t \cap H'$—so the penumbra-shrinkage monotonicity requires the set-inclusion condition. □

*Proof of Proposition 4.* The per-period payoff under canonical holdings with no overruling is $U_t = \alpha \cdot \mathbf{1}\{d_t = d(z_t; w_{j_t}, c_{j_t})\} - K \cdot \mathbf{1}\{\neg \mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)\}$.

*(i)* If $z_t$ is in the penumbra, both outcomes are plausible (no sanction either way). Choosing the ideal outcome $d_t = d(z_t; w_{j_t}, c_{j_t})$ yields $U_t = \alpha > 0$. Choosing the other outcome yields $U_t = 0$. The ideal outcome is strictly preferred.

*(ii)* If only one outcome $d^*$ is plausible, choosing $d^*$ yields $U_t \geq 0$ (either $\alpha$ if $d^*$ matches the ideal, or 0 if not). Choosing the implausible outcome yields at most $\alpha - K < 0$ (since $K > \alpha$). The plausible outcome is strictly preferred regardless of ideology.

*(iii)* In both cases the judge chooses a plausible outcome, so $\mathsf{Plaus}(d_t \mid z_t, \mathcal{F}_t)$ holds. Under canonical holdings, a plausible outcome generates a feasible holding: $\mathcal{F}_t \cap H_t \supseteq \mathcal{F}_t(d_t; z_t) \neq \emptyset$. Therefore $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t$ is well defined and nonempty. □

*Proof of Proposition 5. (i)* By Proposition 4(iii), every decision is plausible and every canonical holding is feasible, so $\mathcal{F}_{t+1} = \mathcal{F}_t \cap H_t \subseteq \mathcal{F}_t$.

For strict volume decrease when $z_t$ is in the open penumbra ($m_t(z_t) < 0 < M_t(z_t)$): both outcomes are plausible, so $\mathcal{F}_t(0; z_t) \neq \emptyset$ and $\mathcal{F}_t(1; z_t) \neq \emptyset$. The judge chooses outcome $d_t$, and the canonical holding $H_t$ is the closed halfspace $\{(w, c) : (-1)^{1-d_t}(w^\top z_t - c) \geq 0\}$. The complementary open halfspace contains $\mathcal{F}_t(1-d_t; z_t) \neq \emptyset$, so the removed set $\mathcal{F}_t \setminus H_t$ is nonempty.

Since $\mathcal{F}_t$ has positive volume and the holding hyperplane $\{(w, c) : w^\top z_t - c = 0\}$ has points of $\mathcal{F}_t$ strictly on both sides (by the penumbra condition), the hyperplane intersects the interior of $\mathcal{F}_t$. Both $\mathcal{F}_t \cap H_t$ and $\mathcal{F}_t \setminus H_t$ therefore have positive $(k+1)$-dimensional volume, so $\mathrm{Vol}(\mathcal{F}_{t+1}) < \mathrm{Vol}(\mathcal{F}_t)$.

*(ii)* We exhibit a constructive example with $k = 1$. Set $\mathcal{F}_1 = \{(w, c) \in \mathbb{R}^2 : |w| \leq 1, |c| \leq 1\}$ and consider two penumbra cases $z_1 = 0.5$, $z_2 = 0.8$ with judges $j_A$ (ideal $(1, 0)$, plaintiff-leaning) and $j_B$ (ideal $(-1, 0)$, defendant-leaning). Verify the penumbra condition: at $z = 0.5$, $M_1(0.5) = \max_{|w| \leq 1, |c| \leq 1}(0.5w - c) = 0.5 + 1 = 1.5 > 0$ and $m_1(0.5) = \min_{|w| \leq 1, |c| \leq 1}(0.5w - c) = -0.5 - 1 = -1.5 < 0$; similarly for $z = 0.8$.

*Order 1:* $j_A$ decides $z_1 = 0.5$: outcome $d_1 = 1$ (since $1 \cdot 0.5 \geq 0$), canonical holding $H_1 = \{0.5w \geq c\}$. Then $j_B$ decides $z_2 = 0.8$: outcome $d_2 = 0$ (since $(-1) \cdot 0.8 < 0$), canonical holding $H_2 = \{0.8w \leq c\}$. The feasible set becomes $\mathcal{F}_3 = \{(w, c) : |w| \leq 1,\ 0.8w \leq c \leq 0.5w\}$. The constraint $0.8w \leq 0.5w$ requires $0.3w \leq 0$, forcing $w \leq 0$: only defendant-leaning rules survive.

*Order 2:* $j_B$ decides $z_1 = 0.5$ first: $d_1 = 0$, $H_1' = \{0.5w \leq c\}$. Then $j_A$ decides $z_2 = 0.8$: $d_2 = 1$, $H_2' = \{0.8w \geq c\}$. Now $\mathcal{F}_3' = \{(w, c) : |w| \leq 1,\ 0.5w \leq c \leq 0.8w\}$, which requires $w \geq 0$: only plaintiff-leaning rules survive.

The sets $\mathcal{F}_3$ (with $w \leq 0$) and $\mathcal{F}_3'$ (with $w \geq 0$) intersect only at $(0, 0)$. Same cases, same judges, different order: the resulting law is qualitatively different.

*(iii)* We show the one-step expected drift of the barycenter under the additional assumption that $\mathcal{F}_t$ is symmetric about the cutting hyperplane $\{(w, c) : w^\top z_t = c\}$. Write $\bar{x}(S) = \frac{1}{\mathrm{Vol}(S)} \int_S x\, dx$ for the centroid of a body $S$. Fix a penumbra case $z_t$ and let $\mathcal{F}_t^+ = \mathcal{F}_t \cap \{w^\top z_t \geq c\}$ and $\mathcal{F}_t^- = \mathcal{F}_t \cap \{w^\top z_t \leq c\}$. Both have positive volume (penumbra condition).

*Claim:* $\tilde{z}_t^\top \bar{x}(\mathcal{F}_t^+) > 0 > \tilde{z}_t^\top \bar{x}(\mathcal{F}_t^-)$.

All points in $\mathcal{F}_t^+$ (excluding the measure-zero boundary) satisfy $\tilde{z}_t^\top x = w^\top z_t - c > 0$, so $\tilde{z}_t^\top \bar{x}(\mathcal{F}_t^+) > 0$. Similarly, $\tilde{z}_t^\top \bar{x}(\mathcal{F}_t^-) < 0$.

Under Proposition 4, a plaintiff-favoring judge ($d_t = 1$) sets $\mathcal{F}_{t+1} = \mathcal{F}_t^+$ and a defendant-favoring judge ($d_t = 0$) sets $\mathcal{F}_{t+1} = \mathcal{F}_t^-$. If the judge distribution is asymmetric with $\Pr(d_t = 1 \mid z_t \text{ in penumbra}) = p > \frac{1}{2}$, the expected value $\mathbb{E}[\tilde{z}_t^\top \bar{x}(\mathcal{F}_{t+1})] = p \cdot \tilde{z}_t^\top \bar{x}(\mathcal{F}_t^+) + (1-p) \cdot \tilde{z}_t^\top \bar{x}(\mathcal{F}_t^-)$ is a $p$-weighted average of a positive and a negative term. Since $p > \frac{1}{2}$ and the symmetry assumption ensures $|\tilde{z}_t^\top \bar{x}(\mathcal{F}_t^+)| = |\tilde{z}_t^\top \bar{x}(\mathcal{F}_t^-)|$ (the two halves are reflections across the cutting hyperplane), the expectation is positive: the barycenter drifts in the $\tilde{z}_t$-direction, toward plaintiff-favoring rules. Taking expectations over $z_t \sim G$, the expected barycenter drifts toward the dominant ideology. $\square$

*Proof of Proposition 6. (i)* Without overruling, only one outcome is plausible at $z_t$ (since $d_t^*$ is not plausible by assumption), and the judge must choose it by Proposition 4(ii), receiving ideology payoff 0 (since the forced outcome differs from $d_t^*$). With optimal overruling, the judge removes the minimum set of $W^*$ holdings that restores plausibility of $d_t^*$, paying $C \cdot W^*$, and then chooses $d_t^*$, receiving ideology payoff $\alpha$. Net payoff from overruling: $\alpha - C \cdot W^*$.

Could overruling fewer than $W^*$ holdings be beneficial? No: by definition of $W^*$, fewer removals do not restore plausibility, so the judge still cannot choose $d_t^*$ and gains nothing. Could overruling more than $W^*$ be beneficial? No: once plausibility is restored, additional removals only add cost. Therefore the judge overrules iff $\alpha - C \cdot W^* > 0$.

*(ii)* Under canonical holdings, each past holding $H_\tau$ is a single halfspace. Removing holding $\tau$ from the intersection removes one halfspace constraint from the system defining $\mathcal{F}_t$. Therefore $W^*$ equals the minimum number of facets whose removal admits a rule producing $d_t^*$ at $z_t$.

*(iii)* We show $W_t^* \leq W_{t+1}^*$ along a history with no overruling between $t$ and $t+1$, where subscripts index the time at which the minimum hitting set is computed and the target $(z, d^*)$ is fixed.

First, if $d^*$ is not plausible at time $t$, it remains not plausible at $t+1$: since $\mathcal{F}_{t+1} \subseteq \mathcal{F}_t$, Proposition 2 implies $M_{t+1}(z) \leq M_t(z)$ and $m_{t+1}(z) \geq m_t(z)$, so the plausibility condition can only become harder to satisfy. Hence $W_{t+1}^*$ is well defined.

Let $S^* \subseteq \{1, \ldots, t\}$ be an optimal removal set at time $t+1$ with $|S^*| = W_{t+1}^*$. We construct a valid removal set at time $t$ of size $\leq |S^*|$.

*Case 1:* $t \notin S^*$. Then $S^* \subseteq \{1, \ldots, t-1\}$. The feasible set at $t+1$ after removing $S^*$ is $\mathcal{F}_1 \cap \bigcap_{\tau \leq t,\, \tau \notin S^*} H_\tau$, which includes $H_t$. The feasible set at time $t$ after removing $S^*$ is $\mathcal{F}_1 \cap \bigcap_{\tau \leq t-1,\, \tau \notin S^*} H_\tau \supseteq \mathcal{F}_1 \cap \bigcap_{\tau \leq t,\, \tau \notin S^*} H_\tau$ (the former omits $H_t$). Since the latter admits a rule producing $d^*$ at $z$, so does the former. Thus $S^*$ is valid at time $t$ with $|S^*| = W_{t+1}^*$.

*Case 2:* $t \in S^*$. Set $S' = S^* \setminus \{t\}$, so $|S'| = W_{t+1}^* - 1$. The feasible set at $t+1$ after removing $S^*$ is $\mathcal{F}_1 \cap \bigcap_{\tau \leq t,\, \tau \notin S^*} H_\tau$. Since $t \in S^*$, this equals $\mathcal{F}_1 \cap \bigcap_{\tau \leq t-1,\, \tau \notin S'} H_\tau$, which is exactly the feasible set at time $t$ after removing $S'$. So $S'$ is valid at time $t$ with $|S'| = W_{t+1}^* - 1$.

In both cases, $W_t^* \leq W_{t+1}^*$. $\square$

*Proof of Corollary 1.* Under canonical holdings, each case outcome $(z_\tau, d_\tau)$ constrains $\mathcal{F}_t$ to the halfspace $\{(w, c) : (-1)^{1-d_\tau}(w^\top z_\tau - c) \geq 0\}$. The surviving set $\mathcal{F}_t$ is the set of all affine classifiers in $\mathcal{F}_1$ consistent with every past outcome—precisely the version space of affine (linear threshold) classifiers on $\mathbb{R}^k$.

$d_{\text{VC}} = k+1$. For the lower bound, consider the $k+1$ points $e_1, \ldots, e_k, \mathbf{0}$ in $\mathbb{R}^k$, where $e_i$ denotes the $i$-th standard basis vector. Lift to the augmented space: $\tilde{e}_i = (e_i, -1) \in \mathbb{R}^{k+1}$ and $\tilde{\mathbf{0}} = (\mathbf{0}, -1)$. These $k+1$ vectors are linearly independent in $\mathbb{R}^{k+1}$. For any labeling $y_1, \ldots, y_{k+1} \in \{+1, -1\}$, linear independence guarantees the existence of $\tilde{w} \in \mathbb{R}^{k+1}$ with $\tilde{w}^\top \tilde{e}_i = y_i$ for $i = 1, \ldots, k$ and $\tilde{w}^\top \tilde{\mathbf{0}} = y_{k+1}$. Writing $\tilde{w} = (w, c)$, the affine classifier with parameters $(w, c)$ achieves the desired labeling with strict separation. Hence these $k+1$ points are shattered.

For the upper bound: by Radon's theorem, any set of $k+2$ points in $\mathbb{R}^k$ can be partitioned into two nonempty sets $A, B$ with $\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \emptyset$. No hyperplane can separate $A$ from $B$, so the labeling assigning $+1$ to $A$ and $-1$ to $B$ cannot be realized. Hence no $k+2$ points can be shattered, and $d_{\text{VC}} = k+1$.

*(ii)* The sample complexity bound follows from the fundamental theorem of statistical learning in the realizable case (Vapnik, 1995, Ch. 6). Since all past case outcomes are consistent with some

rule in the version space, every surviving hypothesis has generalization error at most $\varepsilon$ after

$$t = O\left(\frac{d_{\mathrm{VC}} + \log(1/\delta)}{\varepsilon}\right) = O\left(\frac{k + \log(1/\delta)}{\varepsilon}\right)$$

i.i.d. samples, with probability at least $1 - \delta$.

*(iii)* The bound is linear in $k = d_{\mathrm{VC}} - 1$, so higher-dimensional fact spaces require proportionally more cases for any fixed $(\varepsilon, \delta)$. □

## References


**Ashley, Kevin D.**, "Modeling Legal Argument: Reasoning with Cases and Hypotheticals," in "PhD Thesis, University of Massachusetts" 1990. Published by MIT Press.

**Baker, Scott and Claudio Mezzetti**, "A Theory of Rational Jurisprudence," *Journal of Political Economy*, 2012, *120* (3), 513–551.

**Branting, L. Karl**, "A Computational Model of Ratio Decidendi," *Artificial Intelligence and Law*, 1993, *2* (1), 1–31.

**Callander, Steven and Tom S. Clark**, "Precedent and Doctrine in a Complicated World," *American Political Science Review*, 2017, *111* (1), 184–203.

**Cameron, Charles M., Lewis A. Kornhauser, and Giri Parameswaran**, "Stare Decisis and Judicial Log-Rolls: A Gains-from-Trade Model," *RAND Journal of Economics*, 2019, *50* (3), 505–531.

**Dutz, Denizalp, Ira Shao, Avrim Blum, and Peter Cohen**, "A Machine Learning Theory Perspective on Strategic Litigation," 2025. Working paper, arXiv.

**Dworkin, Ronald**, *Law's Empire*, Harvard University Press, 1986.

**Fernández, Raquel and Giacomo A. M. Ponzetto**, "Stare Decisis: Rhetoric and Substance," *Journal of Law, Economics, and Organization*, 2012, *28* (2), 313–346.

**Fox, Justin and Georg Vanberg**, "Narrow versus Broad Judicial Decisions," *Journal of Theoretical Politics*, 2014, *26* (3), 355–383.

**Gennaioli, Nicola and Andrei Shleifer**, "The Evolution of Common Law," *Journal of Political Economy*, 2007, *115* (1), 43–68.

**Hart, H. L. A.**, *The Concept of Law*, Oxford University Press, 1961.

**Hartline, Jason, Madeline Linna, Liren Shan, and Ao Tang**, "Algorithmic Learning Foundations for Common Law," 2022. Working paper, arXiv:2206.11903.

**Horty, John F.**, "The Result Model of Precedent," *Legal Theory*, 2004, *10* (1), 19–31.

_ **and Trevor J. M. Bench-Capon**, "A Factor-Based Definition of Precedential Constraint," *Artificial Intelligence and Law*, 2012, *20* (2), 181–214.

**Kaplow, Louis**, "Rules Versus Standards: An Economic Analysis," *Duke Law Journal*, 1992, *42* (3), 557–629.

**Katz, Daniel Martin, II Michael J. Bommarito, and Josh Blackman**, "A General Approach for Predicting the Behavior of the Supreme Court of the United States," *PLOS ONE*, 2017, *12* (4), e0174698.

**Kelley, Jr. James E.**, "The Cutting-Plane Method for Solving Convex Programs," *Journal of the Society for Industrial and Applied Mathematics*, 1960, *8* (4), 703–712.

**Kornhauser, Lewis A.**, "An Economic Perspective on Stare Decisis," *Chicago-Kent Law Review*, 1989, *65* (1), 63–92.

**Lindquist, Stefanie A. and Frank B. Cross**, "Empirically Testing Dworkin's Chain Novel Theory: Studying the Path of Precedent," *New York University Law Review*, 2005, *80* (4), 1156–1206.

**Mitchell, Tom M.**, "Generalization as Search," *Artificial Intelligence*, 1982, *18* (2), 203–226.

**Niblett, Anthony, Richard A. Posner, and Andrei Shleifer**, "The Evolution of a Legal Rule," *Journal of Legal Studies*, 2010, *39* (2), 325–358.

**Prakken, Henry**, "A Formal Analysis of Some Factor- and Precedent-Based Accounts of Precedential Constraint," *Artificial Intelligence and Law*, 2021, *29* (4), 559–585.

**Re, Richard M.**, "Precedent as Permission," *Texas Law Review*, 2021, *99* (5), 907–963.

**Rigoni, Adam**, "An Improved Factor-Based Approach to Precedential Constraint," *Artificial Intelligence and Law*, 2015, *23* (1), 1–30.

_ , "Representing Dimensions within the Reason Model of Precedent," *Artificial Intelligence and Law*, 2018, *26* (1), 1–22.

**Rissland, Edwina L. and Robert D. Collins**, "The Law as a Learning System," *Proceedings of the First International Conference on Artificial Intelligence and Law*, 1986, pp. 1–8.

**Rockafellar, R. Tyrrell**, *Convex Analysis*, Princeton University Press, 1970.

**Segal, Jeffrey A. and Harold J. Spaeth**, *The Supreme Court and the Attitudinal Model Revisited*, Cambridge University Press, 2002.

**Vapnik, Vladimir N.**, *The Nature of Statistical Learning Theory*, Springer, 1995.

**Volokh, Eugene**, "The Mechanisms of the Slippery Slope," *Harvard Law Review*, 2003, *116* (4), 1026–1137.