MathGov Appendices Volume

(MathGov v5.0i, rev14.27)
Specification Annex

Scope Note

This volume contains the human-readable appendices referenced by the MathGov Foundation Paper. It provides definitions, reference values, registries, and explanatory material required to interpret and apply the normative methodology.

Runnable examples, scenario libraries, kernel edge registries, executable schemas, and deterministic replay artifacts are externalized to **MathGov ProofPack v1.0** and are referenced by cryptographic hash in the Proof-Carrying Case (PCC). This volume is non-runnable by design.

Appendix A: Symbol and Notation Reference

This appendix provides the complete symbol reference, organized by functional category. All symbols are defined with their domains, ranges, units where applicable, and first use location.

A.1 Union and Dimension Indices

Symbol	Name	Domain	Range	Units	First Defined
u	Union index	{1, 2, 3, 4, 5, 6, 7}	—	—	§2.2
d	Dimension index	{1, 2, 3, 4, 5, 6, 7}	—	—	§4.1
U	Operational union set	{Self, Household, Community, Organization, Polity, Humanity/CMIU, Biosphere}	—	—	§2.2

Symbol	Name	Domain	Range	Units	First Defined
*U*_meta	Meta-union set	{Cosmic, Universal}	—	—	§2.3
*U*_full	Complete union set	*U* ∪ *U*_meta	—	—	§2.3
D	Dimension set	{Material, Health, Social, Knowledge, Agency, Meaning, Environment}	—	—	§4.1
*U*_rights	Rights-bearing union set	Subset of *U*	—	—	§6.2
ϕ(*u*, *d*)	Cell-flattening map	*U* × *D*	{1, …, 49}	—	§8.3
I_{49}	Identity matrix	—	ℝ^(49×49)	—	§8.3

A.2 Actions and Option Sets

Symbol	Name	Domain	Range	Units	First Defined
a	Candidate action / option	*O*	—	—	§3.2
O	Option set	Finite set	—	—	§3.2

Symbol	Name	Domain	Range	Units	First Defined
A_NCRC	Set of NCRC-passing options	Subset of O	—	—	§3.2.3
A_adm	Set of admissible options	Subset of A_NCRC	—	—	§3.2.3
Admissible(a)	Admissibility predicate	O	{true, false}	—	§3.2.3
$\succ$	Preference relation	$A\_adm \times A\_adm$	—	—	§3.2.4

A.3 Applicability and Masking

Symbol	Name	Domain	Range	Units	First Defined
m_{u,d}	Applicability mask	$U \times D$	{0, 1}	—	§4.1.5
κ_{u,d}	Cell multiplier (per-cell scaling)	U × D	ℝ_{≥0}	—	§3.2.4 / Appendix AD

A.4 Impact Objects

Symbol	Name	Domain	Range	Units	First Defined
Ī_(u,d)^prop(a)	Propagated, post-saturation impact	$U \times D \times O$	[−1, +1]	Dimensionless	§4.1

Symbol	Name	Domain	Range	Units	First Defined
$I_{(u,d)}^{dir}(a)$	Direct impact after saturation	$U \times D \times O$	[−1, +1]	Dimensionless	§5.3
$\tilde{I}_{(u,d)}^{dir}(a)$	Direct impact aggregate (pre-saturation)	$U \times D \times O$	$\mathbb{R}$	Dimensionless	§5.2
$\tilde{I}_{(u,d)}^{prop}(a)$	Pre-saturation propagated impact	$U \times D \times O$	$\mathbb{R}$	Dimensionless	§8.3
$\bar{I}_{(u,d)}^{prop}(a \mid s)$	Scenario-conditioned propagated impact	$U \times D \times O \times S$	[−1, +1]	Dimensionless	§7.2.5
$\bar{I}_{(u,d)}^{rights}(a)$	Worst-off subgroup impact	$U \times D \times O$	[−1, +1]	Dimensionless	§3.2.8
$\mathbb{E}[\bar{I}_{(u,d)}^{prop}(a)]$	Expected propagated impact	$U \times D \times O$	[−1, +1]	Dimensionless	§11.1

Symbol	Name	Domain	Range	Units	First Defined
Help_(*u*,*d*)(*a*)	Magnitude of improvement	*U* × *D* × *O*	[0, +1]	Dimensionless	§3.2.2
Harm_(*u*,*d*)(*a*)	Magnitude of degradation	*U* × *D* × *O*	[0, +1]	Dimensionless	§3.2.2

A.5 Impact Instance Attributes

Symbol	Name	Domain	Range	Units	First Defined
k	Impact instance index	ℕ	—	—	§5.2
μ_k	Instance magnitude	—	[−1, +1]	Dimensionless	§5.2
r_k	Instance reach	—	[0, 1]	Proportion	§5.2
t_k	Instance time horizon	—	(0, ∞)	Years	§5.2
ℓ_k	Instance conditional likelihood	—	[0, 1]	Probability	§5.2

Symbol	Name	Domain	Range	Units	First Defined
c_k	Instance confidence	—	[0.1, 1]	Dimensionless	§5.2
e_k	Equity/resilience adjustment	—	[0.5, 2.0]	Dimensionless	§5.2
s_k	Sentience multiplier (SGP)	—	[0, 1]	Dimensionless	§5.2
T_ref	Reference horizon for temporal weighting	—	(0, ∞)	Years	§5.2
τ(t)	Temporal weighting function	(0, ∞)	Range: (0, ∞). τ(T_ref)=1; τ(t)>1 for t>T_ref. Optional cap only if declared in PCC (Tier 3 starter: no cap).	Dimensionless	§5.2
μ_phantom	Ignorance penalty magnitude	—	—	Dimensionless	§5.2

A.6 Saturation Parameters

Symbol	Name	Domain	Range	Units	First Defined
β	Direct saturation coefficient	—	(0, ∞)	Dimensionless	§5.3
β_prop	Post-propagation saturation coefficient	—	(0, ∞)	Dimensionless	§8.3.4

A.7 Vector Forms and Kernel

Symbol	Name	Domain	Range	Units	First Defined
I^dir	Flattened direct-impact vector	—	[−1, +1]^49	Dimensionless	§8.3
Ĩ^prop	Pre-saturation propagated impact vector	—	ℝ^49	Dimensionless	§8.3
Ī^prop	Post-propagation saturated impact vector	—	[−1, +1]^49	Dimensionless	§8.3.4

Symbol	Name	Domain	Range	Units	First Defined
K	Ripple kernel (propagation matrix)	—	ℝ^(49×49)	Dimensionless	§8.2
κ_(*i*,*j*)	Kernel entry	—	[−κ_max, +κ_max]	Dimensionless	§8.2
κ_max	Kernel entry bound	—	(0, 1)	Dimensionless	§8.4
ρ_max	Row-sum (ℓ_1-norm) bound	—	(0, 1)	Dimensionless	§8.4
ρ(*K*)	Spectral radius of kernel	—	[0, ∞)	Dimensionless	§8.3
KQS	Kernel Quality Score	—	[0, 1]	Dimensionless	§8.5
KOPS	Key Operational Pathways Set	—	Subset of kernel edges	—	§8.6

A.8 Rights Constraint Symbols

Symbol	Name	Domain	Range	Units	First Defined
R	Canonical rights set	{LIFE, BODY, LBTY, NEED, DIGN, PROC, INFO, ECOL}	—	—	§6.2
θ_r	Rights threshold for right *r*	*R*	[−1, 0)	Dimensionless	§6.2
C_r	Coverage set for right *r*	*R*	Subset of $U \times D$	—	§6.2
$v_r(a)$	Violation depth for right *r*	$O \times R$	[0, ∞)	Dimensionless	§6.3

A.9 Tail-Risk Constraint Symbols

Symbol	Name	Domain	Range	Units	First Defined
C_{cat}	Catastrophe cell set	—	Subset of $U \times D$	—	§7.2

Symbol	Name	Domain	Range	Units	First Defined
C_{cat}^{base}	Base catastrophe cell set	—	{(6,2), (6,7), (7,7)}	—	§7.2.2
s	Scenario index	S	—	—	§7.3
S	Scenario set	—	Finite set	—	§7.3
p_s	Scenario probability	S	[0, 1]	Probability	§7.3
$(x)_+$	Positive-part operator	ℝ	[0, ∞)	—	§3.2.2
$\omega_{(u,d)}$	Catastrophe-cell weight	C_{cat}	[0, 1]	Dimensionless	§7.2.4
ω_{min}	Minimum catastrophe-cell weight floor	—	[0, 1]	Dimensionless	§7.2.4
η	Catastrophe-weight floor slack factor	—	(0, 1]	Dimensionless	§7.2.4
$L(a, s)$	Scenario loss	$O \times S$	[0, 1]	Dimensionless	§7.2.6

Symbol	Name	Domain	Range	Units	First Defined
$L\_raw(a, s)$	Raw-indicator scenario loss	$O \times S$	[0, 1]	Dimensionless	§7.2.6
α	CVaR tail level	—	(0, 1)	Dimensionless	§7.3
CVaR_α	Conditional Value-at-Risk	—	[0, 1]	Dimensionless	§7.3
τ_TRC	Catastrophe corridor threshold	—	[0, 1]	Dimensionless	§7.3

A.10 Weighting and Scoring Symbols

Symbol	Name	Domain	Range	Units	First Defined
w_u	Union weight (HDW output)	U	[0, 1]	Dimensionless	§10.3
v_d	Dimension weight (HDW output)	D	[0, 1]	Dimensionless	§10.3
w_u^{floor}	Union weight floor	U	[0, 1]	Dimensionless	§10.2

Symbol	Name	Domain	Range	Units	First Defined
v_d^{floor}	Dimension weight floor	D	[0, 1]	Dimensionless	§10.2
λ_U	Union blend parameter	—	[0, 1]	Dimensionless	§10.3
λ_D	Dimension blend parameter	—	[0, 1]	Dimensionless	§10.3
RLS(a)	Ripple Logic Score	O	$\mathbb{R}$	Dimensionless	§11.1
$\sigma_{RLS}(a)$	RLS uncertainty estimate	O	[0, ∞)	Dimensionless	§11.2
$\sigma_{(u,d)}(a)$	Cell-level impact uncertainty	$U \times D \times O$	[0, ∞)	Dimensionless	§11.2
λ	Uncertainty aversion parameter	—	[0, ∞)	Dimensionless	§11.3
δ	Judgment Call discrimination threshold	—	(0, ∞)	Dimensionless	§11.4

A.11 Coherence and Containment Symbols

Symbol	Name	Domain	Range	Units	First Defined
UCI_u	Union Coherence Index for union *u*	*U*	[0, 1]	Dimensionless	§11.5
UCI	Overall Union Coherence Index	—	[0, 1]	Dimensionless	§11.5
$\Delta UCI_u(a)$	Predicted coherence shift for union *u*	$U \times O$	$\mathbb{R}$	Dimensionless	§3.4.2
HOI	Hollowing-Out Index	—	$\mathbb{R}$	Dimensionless	§11.5
τ_c	Containment tolerance threshold	—	[−1, 0]	Dimensionless	§3.4.2
θ_{pos}	Positive-impact threshold	—	[0, 1]	Dimensionless	§3.4.2
D_c	Containment depth limit	—	{1, 2, 3}	—	§3.4.2

Symbol	Name	Domain	Range	Units	First Defined
U_a^+	Set of positively impacted unions	—	Subset of U	—	§3.4.2
Anc(u, D_c)	Ancestor (containing) unions	$U \times \mathbb{N}$	Subset of U	—	§3.4.2
Containment(a)	Global containment predicate	O	{true, false}	—	§3.4.2
MRM(a)	Maximin Rights Margin	O	[0, 1]	Dimensionless	§11.6

A.12 Sentience Protocol Symbols

Symbol	Name	Domain	Range	Units	First Defined
SG(x)	Sentience Gradient score	Entity type	[0, 1]	Dimensionless	§9.2
SG_threshold	Rights plateau threshold	—	(0, 1)	Dimensionless	§9.4

Symbol	Name	Domain	Range	Units	First Defined
g_min	Sentience weighting floor	—	(0, 1)	Dimensionless	§9.5
ψ	Sentience weighting curvature	—	(0, 1]	Dimensionless	§9.5
α_*i*	UCI component weights	{1, 2, 3, 4}	[0, 1]	Dimensionless	§11.5
w_j	SGP component weights	{1, ..., 6}	[0, 1]	Dimensionless	§9.2

A.13 Operator Definitions

Symbol	Name	Definition	First Defined
clip(x, a, b)	Clipping operator	max(a, min(x, b))	§11.5
clamp_[0,1](x)	Unit clamp operator	max(0, min(1, x))	§9.2
(x)_+	Positive-part operator	max(x, 0)	§3.2.2
(x)_-	Negative-part operator	max(-x, 0)	§3.2.2
tanh(·)	Hyperbolic tangent	Saturation function	§5.3

Symbol	Name	Definition	First Defined
Smooth(·)	Smoothing operator	EMA with specified half-life	§11.5
PCC	Provenance & Compliance Certificate	Structured record artifact	§13

Appendix AB: Canonical Notation Charter

This appendix establishes the single authoritative notation system for MathGov. All mathematical notation in the framework must conform to this charter. No alternative notation systems are permitted.

AB.1.1 Narrative shorthand (normative).

In explanatory prose only, the documents may use compact shorthand (e.g., I_{u,d}(a) without stage tags) provided the surrounding sentence explicitly states which pipeline stage is being discussed. In all computational requirements (PCC fields, determinism rules, audit checks, registries, and schemas), the stage-specific canonical symbols defined in this appendix remain mandatory. If any ambiguity remains, auditors and implementers MUST treat the stage as undefined and hard-fail Tier ≥ 3.

AB.1 Primary Symbol Declarations

AB.1.1 Index Variables

Symbol	Name	Domain	First Defined
u	Union index	{1, 2, 3, 4, 5, 6, 7}	§2.2
d	Dimension index	{1, 2, 3, 4, 5, 6, 7}	§4.1
a	Candidate action	O	§3.2
s	Scenario index	S	§7.3

Symbol	Name	Domain	First Defined
r	Rights index	R	§6.2
k	Impact instance index	$\mathbb{N}$	§5.2
g	Subgroup index	$G_{(u,d)}$	§3.2.8
i, j	Cell indices (vectorized)	{1, …, 49}	§8.3

AB.1.2 Set Variables

Symbol	Name	Domain	First Defined
O	Option set	Finite set of actions	§3.2
S	Scenario set	Finite set of scenarios	§7.3
R	Rights set	{LIFE, BODY, LBTY, NEED, DIGN, PROC, INFO, ECOL}	§6.2
U	Operational union set	7 unions	§2.2
D	Dimension set	7 dimensions	§4.1
C_r	Coverage set for right r	Subset of $U \times D$	§6.2
C_{cat}	Catastrophe cell set	Subset of $U \times D$	§7.2

Symbol	Name	Domain	First Defined
$G_{(u,d)}$	Protected subgroups for cell	Finite set	§3.2.8
A_{NCRC}	NCRC-admissible options	Subset of O	§6.3
A_{adm}	Fully admissible options	Subset of A_{NCRC}	§3.2.3
U_a^+	Positively impacted unions	Subset of U	§3.4.2
$\mathrm{Anc}(u, D_c)$	Ancestor unions	Subset of U	§3.4.2

AB.2 Impact Objects

The impact pipeline uses distinct symbols for each transformation stage. This is the canonical specification:

Notation convenience (Normative). In some explanatory passages, stage-tags such as Ĩ^{dir,pre}_{u,d}(a) or I^{dir,post}_{u,d}(a) may appear. These are purely mnemonic labels. Treat them as aliases for the canonical stage objects Ĩ^{dir}_{u,d}(a) (direct, pre-saturation) and I^{dir}_{u,d}(a) (direct, post-saturation) respectively. No additional objects are introduced.

AB.2.1 Pre-Saturation Direct Impact Aggregate

Symbol: $\tilde{I}_{(u,d)}^{dir}(a)$

Definition: The raw aggregation of impact instances before normalization:

$$\tilde{I}_{(u,d)}^{dir}(a) = \Sigma_k I_{(u,d,k)}^{dir}(a)$$

Domain: Real-valued and unbounded ($\mathbb{R}$)

Interpretation: Represents cumulative contributions from all impact pathways before saturation.

AB.2.2 Post-Saturation Direct Impact

Symbol: $I\_(u,d)\text{^}dir(a)$

Definition: The saturated direct impact bounded to [−1, +1]:

$$I\_(u,d)\text{^}dir(a) \ = \ tanh(\beta \cdot \tilde{I}\_(u,d)\text{^}dir(a))$$

where the default saturation coefficient is β = 2.

Domain: [−1, +1]

Interpretation: Normalized direct impact ready for kernel propagation.

AB.2.3 Flattened Direct Impact Vector

Symbol: **I**^dir

Definition: The 49-element vector form of direct impacts:

$$I\text{^}dir(a) \ \in \ [-1,1]\text{^}49, \qquad [I\text{^}dir(a)]\_(\phi(u,d)) \ = \ I\_(u,d)\text{^}dir(a)$$

with components indexed by the canonical flattening map $\phi(u, d) = 7(u - 1) + d$.

Domain: [−1, +1]^49

AB.2.4 Pre-Saturation Propagated Impact

Symbol: $\tilde{I}\_(u,d)\text{^}prop(a)$ or $\tilde{\mathbf{I}}$^prop (vector form)

Definition: The output of kernel propagation before final saturation.

Quick mode:

$$\tilde{I}\text{^}prop(a) \ = \ I\text{^}dir(a) \ + \ K \cdot I\text{^}dir(a)$$

Full mode (requires ρ(K) < 1; prohibited for Tier 4 Pilot-Executable rev14.x):

$$\tilde{I}\text{^}prop(a) \ = \ (I \ - \ K)\text{^}(-1) \cdot I\text{^}dir(a), \qquad \rho(K) \ < \ 1$$

Domain: Real-valued (ℝ^49), potentially unbounded

AB.2.5 Post-Propagation Saturated Impact

Symbol: $\bar{I}\_(u,d)\text{^}prop(a)$ or $\bar{\mathbf{I}}$^prop (vector form)

Definition: The final impact value after post-propagation saturation:

$$I\_(u,d)\text{^}prop(a) \ = \ tanh(\beta\_prop \cdot \tilde{I}\_(u,d)\text{^}prop(a))$$

where the default post-propagation saturation coefficient is β_prop = 1.

Domain: [−1, +1]

Usage: This is the final impact value used in NCRC, TRC, and RLS calculations.

AB.2.6 Worst-Off Subgroup Impact

Symbol: Ī_(*u*,*d*)^rights(*a*)

Definition: The minimum impact across all protected subgroups:

$$\bar{I}\_(u, d)\hat{}rights(a) \; = \; min\_\{g \in G\_(u, d)\} \; I\_(u, d, g)\hat{}prop(a)$$

Domain: [−1, +1]

Usage: Used exclusively for NCRC rights checks.

AB.3 Deprecated Aliases

The following notations appeared in earlier drafts and must not be used:

Deprecated	Replacement	Reason
I(*a*) without superscript	Always use superscript (dir, prop, rights)	Ambiguity between pipeline stages
Î (circumflex)	Use tilde (Ĩ) for pre-saturation, overbar (Ī) for post-saturation	Consistency with saturation semantics
Impact(*u*,*d*,*a*)	Use subscript notation: *I*_(*u*,*d*)(*a*)	Compactness
*I*_cell	Use (*u*,*d*) subscript notation	Precision
CVaR(*L*) without subscript	Always use subscript: CVaR_α(*L*)	Tail level must be specified
Weight(*u*)	Use *wu for union weights, vd* for dimension weights	Standard notation

AB.4 First-Use-Defines Rule

Every symbol must be formally defined at its first occurrence. The definition must include:

- **Display equation format:** The symbol in a numbered, centered equation

- **Natural language description:** A clear explanation of what the symbol represents
- **Domain specification:** The set from which the symbol's values are drawn
- **Range specification:** The set of possible output values (for functions)
- **Default values:** Where applicable, the default parameter setting

Example of compliant first-use definition:

The saturation coefficient β controls the sensitivity of the saturation function. It is defined as:

$$\beta = 2$$

$\beta \in (0, \infty)$, default $\beta = 2$

Higher values of β produce sharper saturation near the boundaries, while lower values extend the approximately linear regime.

AB.5 Superscript and Subscript Conventions

AB.5.1 Superscript Usage

Superscripts indicate transformation stage or semantic role:

Superscript	Meaning	Example
dir	Direct (pre-propagation)	*I*^dir
prop	Propagated (post-kernel)	*I*^prop
rights	Worst-off subgroup	Ī^rights
floor	Constitutional floor	*w*^floor
base	Base configuration	*C*_cat^base

AB.5.2 Subscript Usage

Subscripts indicate indexing:

Tier-4 canonical cell identifier for registries and PCC snapshots is { "u": <1..7>, "d": <1..7> }. Labels and (u,d) tuple text are display-only.

Subscript	Meaning	Example
u	Union index	w_u
d	Dimension index	v_d
(u,d)	Cell specification	$I_{(u,d)}$
r	Right index	θ_r
s	Scenario index	p_s
k	Instance index	μ_k

AB.5.3 Combined Usage

When both superscript and subscript are needed, the subscript comes first, then the superscript:

Correct: $\bar{I}_{(u,d)}^{prop}(a)$

Incorrect: Ī^prop_(*u*,*d*)(*a*)

AB.6 Decoration Conventions

Decoration	Meaning	Example
Tilde (˜)	Pre-saturation / raw aggregate	$\tilde{I}$ (pre-saturation)
Overbar (¯)	Post-saturation / final	$\bar{I}$ (post-saturation)
Hat (ˆ)	NOT USED in MathGov	—
Bold	Vector or matrix	**I**, **K**
Blackboard bold	Sets	$\mathbb{R}$, $\mathbb{N}$

Decoration	Meaning	Example
Greek letters	Parameters and functions	β, τ, ρ

AB.7 Function Notation

Purpose. This section defines the canonical function symbols used throughout MathGov. These are not optional notations. They are the single allowed spellings and definitions for these functions under the Canonical Notation Charter. Any deviation must be explicitly declared in the PCC and treated as a nonstandard configuration requiring justification and sensitivity checks.

Conventions.

1. All functions are deterministic unless explicitly indexed by scenario $s$or an uncertainty object.

Negative part: (x)_- = max(-x, 0).

Clip: clip(x, a, b) = max(a, min(x, b)), with a <= b.

Unit clamp: clamp_[0,1](x) = max(0, min(1, x)).

Cell flattening (7x7): ϕ(u, d) = 7(u - 1) + d, for u, d in {1,...,7}.

Temporal weight: tau(t) = ln(1 + t) / ln(1 + T_ref), for t > 0, T_ref > 0.

EMA smoothing: EMA_lambda(x_t) = lambda x_t + (1 - lambda) EMA_lambda(x_{t-1}), lambda in (0,1], initialization declared in PCC.

Infinity norm: ||K||_infty = max_i sum_j |K_ij|.

Function (name)	Notation	Definition	Domain	Range	Defaults and notes
Saturation (hyperbolic tangent)	sat_β(x)	tanh(βx)	x ∈ ℝ, β>0	(−1, 1)	Canonical smooth saturation map. Larger increases boundary steepness; smaller

Function (name)	Notation	Definition	Domain	Range	Defaults and notes
					extends the near-linear regime around 0.
Positive part	(x)_+	max(x, 0)	x ∈ ℝ	ℝ_{≥0}	Used for nonnegative transforms (for example, “only penalize excess” constructions).
Negative part	(x)_-	max(-x, 0)	x ∈ ℝ	ℝ_{≥0}	Used for negative-part transforms (for example, tail-risk definitions).
Clipping (interval bound)	clip(x, a, b)	max(a, min(x, b))	x ∈ ℝ, a ≤ b	[a, b]	Canonical truncation to a closed interval. If a > b, the PCC/implementation is invalid.
Unit clamp	clamp_[0, 1](x)	max(0, min(x, 1))	x ∈ ℝ	[0, 1]	Canonical clamp into the unit interval. Use when a quantity is defined as a

Function (name)	Notation	Definition	Domain	Range	Defaults and notes
					bounded proportion, probability, or normalized score in .
Cell flattening (row-major)	ϕ(u,d)	7(u−1)+d	u,d ∈ {1,…,7}	{1,…,49}	Canonical flattening for 7×7 welfare matrix: ϕ(u,d)=7(u−1)+d with unions ordered {Self, Household, Community, Organization, Polity, Humanity/CMIU, Biosphere} and dimensions ordered {Material, Health, Social, Knowledge, Agency, Meaning, Environment}. Any deviation must be declared in PCC.

Function (name)	Notation	Definition	Domain	Range	Defaults and notes
Ancestor-union set	Anc(u, D_c)	Set of ancestor unions of u up the declared union ladder	u ∈ U	𝒫(U)	Canonical semantics: returns the set of ancestor unions above u at depth D_c on the canonical union ladder. Increasing D_c adds higher-level ancestor unions; decreasing D_c removes them.
Temporal weight (log scale)	τ(t)	ln(1+t) / ln(1+T_ref)	t>0, T_ref>0	(0, ∞)	Default T_ref = 25 years. τ(T_ref)=1. τ(t) increases with t with diminishing marginal gain. τ can exceed 1 when t > T_ref. Any cap τ_max is optional and must be declared in the PCC.
Exponential moving average	EMA_λ(x_t)	λx_t + (1−λ)EMA_λ(x_{t−1})	λ ∈ (0,1]	ℝ	Default initialization: unless a baseline

Function (name)	Notation	Definition	Domain	Range	Defaults and notes
smoothing (EMA)					initialization is declared in PCC. The PCC must report when Smooth is used (for HOI, monitoring, or trend reporting).
Infinity norm (row-sum norm)	$\|\|K\|\|_\infty$	$\max_i \Sigma_j \|K_{ij}\|$	$K \in \mathbb{R}^{49\times49}$	$\mathbb{R}_{\geq 0}$	

AB.7.2 Scope note

AB.7 defines general-purpose functions that recur across the framework (clipping, saturation, flattening, ancestor mapping, temporal weighting, smoothing, stability norms). Specialized operators (for example $\mathrm{CVaR}_\alpha(\cdot)$) are defined where introduced, but must obey the same canonical symbol discipline (same spelling, same parameters, same domains).

Appendix AD: Complete Default Parameter Registry

Tier-4 HDW appendix alignment note (Normative cross-reference)

This appendix may include illustrative embedded registry snippets for readability. For audit and replay, Tier-4 Pilot-Executable runs MUST treat the ProofPack registries and schemas as the authoritative sources, referenced and hash-bound in the PCC.

For Tier 4, hash-canonical artifacts (HDW ballots, HDW weights, registries, manifests, schemas, PCC override bundles) MUST use exact reduced rationals and MUST conform to the canonical JSON and NO_FLOATS rules. Computed quantities (impacts, saturation outputs, propagation outputs, losses, CVaR, RLS, gaps) MUST be computed under NDP_FIXEDPOINT_V1 and stored in the PCC as fixed-point integers only.

Tier-4 requires whole-ballot trimming and the canonical tie-break rule defined in the ProofPack: latest-wins by submitted_at, with sha256(canonical_json(ballot_obj)) hash-max on same-second ties.

Tier-4 timestamp format is strict UTC Z with second precision only: YYYY-MM-DDTHH:MM:SSZ.

This appendix consolidates all default parameter values for reference implementation.

Registry rule. A "default" applies only when the PCC does not override it. If a parameter has no default, it MUST be declared in the PCC.

Parameter	Default / Rule	Tier	Source anchor
β	2 (direct impact saturation: tanh(β·x))	All	§3.2.7; AB.7
β_prop	1 (post-propagation saturation: tanh(β_prop·x))	All	§3.2.7; AB.7
T_ref	25 years (reference horizon for τ)	All	B.1
τ(t)	ln(1+t) / ln(1+T_ref)	All	B.1; AB.7
w_u (union weights)	Uniform if not specified (Tier 1–2); declared fallback only (Tier 3). Tier 4 MUST be HDW-derived.	Tier 1–2; Tier 3 (declared fallback only); Tier 4 (HDW required)	§10.2; Appendix A (HDW)

v_d (dimension weights)	Uniform if not specified (Tier 1–2); declared fallback only (Tier 3). Tier 4 MUST be HDW-derived.	Tier 1–2; Tier 3 (declared fallback only); Tier 4 (HDW required)	§10.2; Appendix A (HDW)
μ_phantom	−0.10 (Tier 4 default ignorance penalty; tiered values may be declared)	Tier 4	§3.2.7; §5.2.2
α (CVaR tail level)	Context defaults per D.3.2 (informative). Default α = 0.95 for Organizational and Reversible policy contexts unless governance requires otherwise; PCC MUST declare context.	Tier 4	§7.3; MRE config
τ_TRC	Context defaults per D.3.2 (informative). Reversible policy τ_TRC = 0.15; Organizational τ_TRC = 0.20; other contexts per D.3.2. PCC MUST declare context.	Tier 4	§7.3; MRE config

p_floor (mandatory-tail probability floor)	0.02 per MTS category (categories may be split across scenarios)	Tier 4	§7.2.5
Kernel baseline	Tier 4 baseline: Starter-KOPS edges only; all other entries 0 (Quick mode) Note: subject to §8.5 KQS Policy; if KQS < 0.40 then set K = 0 and propagation_mode = NONE (recorded in PCC).	Tier 4	Appendix S; Appendix Q
Stability check	Require ρ(K) < 1 for full propagation; if violated, fall back to Quick mode or K=0 as declared	Tier 4	§8.3
Rights thresholds θ_r	Canonical defaults per Table C.2; any override requires governance justification	All	Appendix C.2; §6.3
Rights coverage sets C_r	Authoritative auditor mapping per Table C.3.7	All	Appendix C.3.7
γ (KOPS learning rate)	0.10 (default). Used only when adaptive kernel	Tier 4	Foundation §8.6.3

	learning is enabled.		
ε_KL (robust TRC KL radius)	Declared per PCC/run. Suggested ε_KL = 0.10 for Tier-3 exploratory runs; Tier-4 MUST declare if used.	Tier 3/4	Foundation §7.3.6
κ_{u,d} (cell multiplier)	1.0 (default). Used for per-cell scaling; if changed, PCC MUST declare.	Tier 4	Foundation §3.2.4
e_k (equity factor)	1.0 (default; no equity adjustment). If changed, PCC MUST declare + justify.	Tier 4	Foundation §5.2.1
γ_subgroup (subgroup multiplier)	1.5 (Tier-3 starter suggestion). If used in Tier-4, PCC MUST declare + sensitivity.	Tier 3	Foundation §3.2.8
UCI_ALPHA_H	1/4 (exact rational)	All (default)	Appendix E.7.3
UCI_ALPHA_F	1/4 (exact rational)	All (default)	Appendix E.7.3
UCI_ALPHA_R	1/4 (exact rational)	All (default)	Appendix E.7.3

UCI_ALPHA_E	1/4 (exact rational)	All (default)	Appendix E.7.3
UCI_WITHIN_COMPONENT_WEIGHT_RULE	Uniform: w_{C,k}=1/n_active (exact rational)	All (default)	Appendix E.7.3
SUBGROUP_N_MIN_DEFAULT	30 (or max(5, ceil(0.05*N)) if dataset smaller)	Tier ≥ 3	Foundation §3.2.9
HOI_EMA_INIT_RULE	EMA(x_0)=x_0; ΔRLS_0=0; ΔUCI_0=0	All	Appendix B.23
HDW_HASH_MAX_RULE	Lexicographic max over lowercase hex digest	Tier 4	Foundation Tier-4 HDW section

AD.1 Saturation Parameters

Parameter	Symbol	Default	Section
Direct saturation coefficient	**β**	**2**	**§5.3**
Post-propagation saturation coefficient	**β_prop**	**1**	**§8.3.4**

AD.2 Temporal Parameters

Parameter	Symbol	Default	Section
Reference time horizon	**T_ref**	**25 years**	**§5.2**
HOI smoothing half-life	**h**	**3 periods**	**§11.5**

AD.3 Rights Thresholds

Right	Symbol	Threshold	Section
Life	θ_LIFE	−0.90	§6.2
Bodily Integrity	θ_BODY	−0.70	§6.2
Liberty	θ_LBTY	−0.65	§6.2
Basic Needs	θ_NEED	−0.50	§6.2
Dignity	θ_DIGN	−0.55	§6.2
Due Process	θ_PROC	−0.45	§6.2
Information	θ_INFO	−0.40	§6.2
Ecological Integrity	θ_ECOL	−0.65	§6.2

AD.4 TRC Parameters

Parameter	Symbol	Default (Organizational)	Section
Tail level	α	0.95	§7.3.5
Corridor threshold	τ_TRC	0.20	§7.3.5
Catastrophe weight floor slack	η	0.5	§7.2.4
Minimum scenario probability	p_floor	0.02	§7.4.6

AD.5 Containment Parameters

Parameter	Symbol	Default	Section
Containment tolerance	**τ_c**	**−0.10**	**§3.4.2**
Positive-impact threshold	**θ_{pos}**	**0.05**	**§3.4.2**
Containment depth	**D_c**	**2**	**§11.7**

AD.6 Scoring Parameters

Parameter	Symbol	Default	Section
Discrimination threshold	**δ**	**2**	**§11.4**
Uncertainty aversion	**λ**	**0.5**	**§11.3**
UCI tie-break threshold	**δ_{UCI}**	**0.05**	**§11.4**

AD.7 Kernel Parameters

Parameter	Symbol	Default	Section
Entry bound	**κ_{max}**	**0.5**	**§8.4**
Row-sum bound	**ρ_{max}**	**0.9**	**§8.4**
Non-zero entry cap	**—**	**200**	**§8.4**

AD.8 Sentience Parameters

Parameter	Symbol	Default	Section
Rights plateau threshold	**$SG_{threshold}$**	**0.85**	**§9.4**
Sentience weighting floor	**g_{min}**	**0.05**	**§9.5**

Parameter	Symbol	Default	Section
Sentience weighting curvature	**ψ**	**0.5**	**§9.5**

AD.9 Ignorance Penalty by Tier

AD.10 Numeric determinism profile (Tier 4) (Normative)

numeric_profile_id: NDP_FIXEDPOINT_V1

S (fixed-point scale): 1000000000

SAT_LUT_ID: SAT_LUT_FP_V1 (hash-bound in ProofPack; referenced in PCC)

TEMPORAL_WEIGHT_REGISTRY_ID: REG_TEMPORAL_WEIGHTS_V1 (hash-bound in ProofPack; referenced in PCC)

Tier 4 rev14.x restriction: propagation_mode ∈ {NONE, QUICK}. Full propagation is prohibited for Tier 4 Pilot-Executable until a deterministic solver profile is hash-bound.

NDP_FIXEDPOINT_V1 edge cases (Normative). All fixed-point multiplications MUST use an intermediate wide integer (at least 128-bit or arbitrary precision) and MUST hard-fail on overflow before rescaling back to scale S. Fixed-point division MUST be implemented as integer division with round_half_even to the nearest representable value at scale S. Values whose absolute magnitude is < 0.5/S after rounding MUST resolve to 0 at scale S (underflow-to-zero). If any intermediate or final value exceeds the declared representable bounds, the run MUST hard-fail (no silent wrap).

Tier	μ_phantom	Section
1	−0.05	§5.2
2	−0.08	§5.2
3	−0.10	§5.2
4	−0.10	§5.2

Note (tier intent). μ_phantom is an ignorance penalty used only when a required instance is missing and the “phantom instance” rule is invoked (§5.2). Tier 4 keeps μ_phantom aligned to the canonical phantom defaults; more conservative values MAY be declared in the PCC for exceptional high-stakes contexts, but must be explicit and versioned.

AD.10 NCAR Triggers

Tier 3 starter NCAR triggers. The following conditions trigger Notice-Choose-Act-Reflect (NCAR) escalation and require explicit reflection notes in the PCC:

NCAR-1 Rights proximity: any rights-covered cell impact in the worst-off subgroup is within 0.05 of its rights threshold (sensitivity band).

NCAR-2 Tail-risk proximity: CVaR_alpha(L) is within 0.05 of the TRC threshold, or scenario feasibility constraints are tight (tail mass poorly resolved).

NCAR-3 Decision fragility: the top-ranked option changes under any declared sensitivity run (weights, kernel perturbation, or interval endpoint propagation).

NCAR-4 Low evidence kernel: KQS < 0.50 and any kernel propagation is used (Quick or Full).

NCAR-5 High uncertainty: sigma_RLS(a) >= 0.10 for any leading option, or the discrimination band overlaps among top candidates.

NCAR-6 Missing or weak data: more than 30% of asserted instances have c_k < 0.50, or key rights anchors are placeholders.

NCAR-7 Stakeholder dissent: recorded stakeholder objection from any union ballot above the Charter-set dissent threshold.

If any trigger fires, the PCC must include: (i) the trigger ID(s), (ii) what additional evidence was sought (if any), (iii) whether a safer alternative exists, and (iv) the final justification if proceeding.

AD.11 Equity Criteria Registry (Tier 4 starter set)

This registry constrains the equity/resilience adjustment e_k so it cannot be improvised ad hoc. The default rule is e_k = 1. Any e_k != 1 must cite exactly one criterion from this registry and must satisfy the evidence requirements below.

Allowed range (Tier 3 starter): e_k in [0.85, 1.15]. Values outside this range are disallowed in Tier 4 unless Charter-approved and explicitly documented in the PCC.

Tier 4 starter criteria (criterion ID, description, allowed e_k, evidence requirement):

ECR-1 Worst-off subgroup protection. Adjustments that reduce harm to the worst-off subgroup identified in the PCC subgroup analysis. Allowed e_k in [1.00, 1.15]. Evidence: subgroup identification, harm pathway, and counterfactual audit with e_k = 1.

ECR-2 Rights-adjacent harm mitigation (non-trigger). Harms that do not cross an NCRC threshold but are plausibly near-threshold in sensitivity analysis. Allowed e_k in [1.00, 1.10]. Evidence: threshold sensitivity run showing proximity.

ECR-3 Historic burden correction. Adjustments addressing historically accumulated disadvantage in the affected stakeholder set. Allowed e_k in [1.00, 1.10]. Evidence: documented disparity baseline and linkage to the instance.

ECR-4 Power asymmetry correction. Adjustments reflecting asymmetric ability to avoid harm or influence outcomes (voice deficit). Allowed e_k in [1.00, 1.10]. Evidence: documented asymmetry and PCC explanation.

ECR-5 Resilience protection for critical life support. Adjustments increasing weight on impacts that protect essential support systems for continued flourishing (health infrastructure, water access, etc.), short of TRC catastrophe triggers. Allowed e_k in [1.00, 1.10]. Evidence: system dependency mapping and justification.

ECR-6 Anti-gaming penalty. Down-weight instances that appear strategically framed to inflate welfare improvements without measurable support. Allowed e_k in [0.85, 1.00]. Evidence: audit rationale with specific flags and a sensitivity analysis.

Audit requirements (mandatory whenever e_k != 1). The PCC must include: (i) the criterion ID, (ii) evidence sources, (iii) a counterfactual recomputation with all e_k reset to 1, and (iv) a selection sensitivity statement indicating whether the final decision changes when e_k = 1.

Metric	Trigger Threshold	Section
Sign Accuracy	**< 0.60**	**§12.5**
Magnitude RMSE	**> 0.30**	**§12.5**
Rights Near-Miss Rate	**> 0.25**	**§12.5**
HOI Alert	**> 0.15 for 3+ periods**	**§11.5**

Appendix B: Key Equations

This appendix collects all canonical equations needed to implement MathGov. Notation is consistent across all sections: $\tilde{I}$ denotes pre-saturation values, $\bar{I}$ denotes post-saturation values, and all propagated impacts used for constraints and scoring are post-saturation bounded to [−1, +1].

B.1 Temporal Weighting (Log Scale)

Purpose. Temporal weighting maps an impact instance's time horizon into a dimensionless multiplier so that longer-lasting consequences receive greater weight without using exponential discounting that would systematically marginalize long-horizon and intergenerational effects. The temporal weight applies at the **impact-instance level** in the direct-aggregation step (Section 5.2) and can also be referenced in scenario-conditioned evaluations when instance horizons differ by scenario.

Tier 4 override (Normative). Tier 4 MUST NOT compute logarithms at runtime. Tier 4 temporal weights MUST be sourced from REG_TEMPORAL_WEIGHTS_V1 and applied under NDP_FIXEDPOINT_V1.

B.1.1 Definitions

Let **t** be the time horizon of an impact instance, expressed in years, with:

1. $t \in (0, \infty)$

Let **T_ref** be the reference horizon used to normalize the scale. Default:

1. T_ref = 25 years

Define the temporal weighting function τ(t) as:

Note. Because τ(t)=ln(1+t)/ln(1+T_ref), τ(t) is not bounded above by 1; τ(T_ref)=1 and τ(t)>1 for horizons longer than T_ref.

$$\tau(t) = \ln(1 + t) / \ln(1 + T\_ref)$$

This mapping satisfies:

1. $\tau(t) \geq 0$ for all $t > 0$
2. $\tau(T\_ref) = 1$
3. τ(t) increases with t but with diminishing marginal gain (logarithmic growth)

B.1.2 Illustrative values (default T_ref = 25)

Using T_ref = 25 years:

1. $\tau(1) = \ln(2) / \ln(26) \approx 0.21$
2. $\tau(5) = \ln(6) / \ln(26) \approx 0.56$
3. $\tau(10) = \ln(11) / \ln(26) \approx 0.75$
4. $\tau(25) = \ln(26) / \ln(26) = 1.00$
5. $\tau(50) = \ln(51) / \ln(26) \approx 1.22$

These values are illustrative; exact values are fully determined by the equation above.

B.1.3 Governance rules for overriding T_ref

T_ref may be overridden by governance (PCC-declared) in contexts where a different anchoring horizon is explicitly justified (e.g., some infrastructure or ecological policy contexts may use longer horizons). Any override must satisfy:

1. The PCC explicitly states the chosen T_ref and the rationale
2. All options in the decision are evaluated using the same T_ref
3. Sensitivity analysis is recommended when decisions are close (e.g., comparing T_ref = 25 vs. T_ref = 50)

B.1.4 Rationale: why logarithmic weighting instead of exponential discounting

MathGov adopts logarithmic temporal weighting rather than exponential discounting because:

1. **Exponential discounting can erase long-horizon impacts.** Even moderate annual discount rates drive far-future consequences toward near-zero weight, which is incompatible with the framework's emphasis on preserving conditions for continued flourishing and avoiding catastrophic tail risks.
2. **Log weighting preserves intergenerational salience without runaway dominance.** The log form increases weight with duration while avoiding the extremes of linear growth that can allow very long durations to dominate any finite near-term magnitude.
3. **Consistency with non-compensability.** Temporal weighting shapes the aggregation of welfare impacts, but it does not override the admissibility cascade (NCRC, TRC). Rights floors and catastrophe corridors remain hard constraints independent of $\tau(t)$.

B.1.5 Implementation note

The temporal weight τ(t) is applied multiplicatively inside the impact-instance aggregation equation in Section 5.2:

Containment Mode A: authoritative algorithm reference

Appendix B provides parameter names and predicate semantics. The authoritative step-by-step Mode A gate algorithm is Foundation §11.7A. In case of any mismatch, Foundation §11.7A governs.

- Each instance contribution is multiplied by τ(t_k) for its horizon t_k.

The PCC must document the horizon assignment method (e.g., duration until effect decays below a material threshold, policy sunset, ecological recovery window, etc.).

B.2 Direct Impact Aggregation (Pre-Saturation)

Purpose. This appendix section defines the **pre-saturation direct impact** for each union–dimension cell. It aggregates multiple predicted consequence instances into a single **unsaturated** scalar per cell, preserving sign (help vs harm) and weighting by reach, time horizon, likelihood, confidence, and (optionally) an equity/resilience factor. The output of B.2 is the canonical input to **B.3 Saturation**, which maps the unsaturated value into the bounded interval $[-1, +1]$.

B.2.1 Objects and indexing

Let:

1. $\mathcal{U}$be the set of unions (parameterized in the main text),
2. $\mathcal{D}$be the set of welfare dimensions,
3. $a \in \mathcal{A}$be an option under evaluation,
4. $(u, d) \in \mathcal{U} \times \mathcal{D}$be a union–dimension cell.

For each option $a$and cell (u, d), define the finite set of impact instances:

- $\mathcal{K}(u, d, a)$= set of impact instances asserted for cell (u, d)under option a

Each impact instance $k \in \mathcal{K}(u, d, a)$is a structured record with the attributes below.

B.2.2 Impact instance attributes (required fields)

Each instance $k$has:

1. **Magnitude** $\mu_k \in [-1, +1]$
   Signed direction and severity of the welfare change for the cell.

- $\mu_k > 0$indicates a benefit ("help")
- $\mu_k < 0$indicates a harm
- $\mu_k = 0$indicates no material effect

Magnitude is defined after calibration/anchoring (Section 5.4) and is a **dimensionless** score in the canonical scale.

- **Reach** $r_k \in [0,1]$
  Fraction of the relevant stakeholder population in union $u$meaningfully affected by the instance within the dimension d.
  Examples: $r_k = 0.10$means roughly 10% of the relevant union stakeholders are affected at non-trivial intensity.

- **Time horizon** $t_k \in (0, \infty)$years
  Approximate duration over which the effect persists at material relevance. Time horizons are weighted using the temporal weighting function $\tau(t)$in B.1.

- **Conditional likelihood** $\ell_k \in [0,1]$
  Probability that the instance occurs given the assumed context.

- If scenario evaluation is not used, $\ell_k$is conditional on the baseline forecast.

- If scenario evaluation is used, $\ell_k$may be conditional on a scenario s(this must be stated in the PCC).

- **Confidence** $c_k \in [0.1,1]$
  Analyst confidence in the parameterization of the instance (evidence quality, model support, measurement reliability, causal attribution).
  A lower bound of 0.1 prevents zeroing out impacts while still strongly penalizing weak claims.

- **Equity/resilience adjustment** $e_k \in [0.5,2.0]$(optional, default = 1)
  A governed multiplier that adjusts impact contribution when equity or resilience criteria apply. Any $e_k \neq 1$must be justified and audited in the PCC per the equity/resilience governance rules (Section 5.2.3).

B.2.3 Temporal weighting

Temporal weighting is applied using $\tau(t)$defined in B.1:

1. Default reference horizon: $T_{\text{ref}} = 25$years
2. $\tau(t) = \ln(1 + t)/\ln(1 + T_{\text{ref}})$

This weight is applied multiplicatively to each instance via $\tau(t_k)$.

B.2.4 Pre-saturation aggregation equation

Define the **pre-saturation direct impact** for option $a$in cell (u, d)as:

Ĩ^{dir,pre}_{u,d}(a) = sum_{k in K(u,d,a)} r_k * tau(t_k) * ell_k * e_k * c_k * s_k * mu_k

Canonical rule (Tier 4). Confidence c_k is multiplicative in direct aggregation. If c_k is omitted for any reason, the PCC must declare this deviation and justify it.

where mu_k is the signed impact magnitude for instance k (in [-1,+1]) prior to union aggregation, r_k is reach, tau(t_k) is temporal weight, ell_k is conditional likelihood, and e_k is the governed equity/resilience adjustment (default 1).

$$\tilde{I}_{u,d}^{dir}(a) \;=\; \sum_{k\in\mathcal{K}(u,d,a)} \mu_k \cdot r_k \cdot \tau(t_k) \cdot \ell_k \cdot c_k \cdot e_k$$

Where:

1. $\tilde{I}_{u,d}^{dir}(a) \in \mathbb{R}$is **unbounded** (it may exceed ± 1before saturation)
2. The sign of $\tilde{I}_{u,d}^{dir}(a)$indicates net help vs net harm for that cell, prior to saturation
3. The magnitude reflects cumulative contributions from multiple instances

This is the canonical "raw summed impact" that will be passed to B.3 for bounded mapping into $[-1, +1]$.

Terminology note. This is sometimes called the "multiplicative aggregation form" because each instance contribution is a product (r_k · τ(t_k) · ℓ_k · e_k · c_k · μ_k). The aggregation across instances is a sum. This term distinguishes the approach from alternatives where confidence is applied after aggregation or treated purely as a variance weight.

B.2.5 Help/harm decomposition (optional diagnostic)

For diagnostics and transparency, MathGov may decompose instance contributions into positive ("help") and negative ("harm") components.

Define the positive-part operator:

(x)_+ = max(x, 0)

(x)_- = max(-x, 0)

$$(x)^+ := \max(x, 0)$$

Then define:

H_{u,d}(a) = sum_{k in K(u,d,a)} (r_k * tau(t_k) * ell_k * e_k * c_k * s_k * mu_k)_+

$$Help_{u,d}^{dir}(a) := \sum_{k \in \mathcal{K}(u,d,a)} (\mu_k)^+ r_k \tau(t_k) \ell_k c_k e_k$$

$$Harm_{u,d}^{dir}(a) := \sum_{k \in \mathcal{K}(u,d,a)} (-\mu_k)^+ r_k \tau(t_k) \ell_k c_k e_k$$

D_{u,d}(a) = sum_{k in K(u,d,a)} (r_k * tau(t_k) * ell_k * e_k * c_k * s_k * mu_k)_-

so that Ĩ^{dir,pre}_{u,d}(a) = H_{u,d}(a) - D_{u,d}(a).

And:

$$\tilde{I}_{u,d}^{dir}(a) = Help_{u,d}^{dir}(a) - Harm_{u,d}^{dir}(a)$$

This decomposition is useful for explaining why a cell score is net-positive or net-negative and for auditing claims that an option has “many helps” but is still forbidden (via NCRC/TRC).

B.2.6 Missing-data rule (ignorance penalty)

To prevent score inflation by omission, if a cell $(u, d)$is active and materially relevant but no defensible instances can be specified for option a, the PCC must record an “unknown impact” and apply a phantom instance k_{ph}.

Canonical phantom instance parameters (Tier 4 default):

Define a phantom instance k_phi with parameters:

mu_phi = -0.10, r_phi = 1, t_phi = T_ref (25 years), ell_phi = 1, e_phi = 1, c_phi = 1.00. (So the default phantom contribution equals μ_phantom = -0.10.)

Then add its contribution r_phi * tau(t_phi) * ell_phi * e_phi * mu_phi (= -0.10 by default) to the pre-saturation sum for that cell whenever the cell is active but no empirical instances are asserted.

Tier 4 note (Normative). For Tier-4 Pilot-Executable runs, if m(u,d)=1 and the instance set K(u,d,a) is empty for an active cell, a phantom instance MUST be inserted to prevent unintended masking and forked implementations. The phantom instance MUST be recorded in the PCC with phantom_enabled: true and phantom_penalty_enabled: true (default). Sensitivity testing MAY be performed by rerunning with phantom_penalty_enabled: false as a diagnostic, but this diagnostic run MUST NOT be used to claim Tier-4 conformance unless explicitly declared as a downgrade.

The phantom instance is included in the sum exactly like any other instance and must be explicitly labeled in the PCC as an ignorance penalty.

B.2.7 PCC documentation requirements

For each cell (u, d) and option a, the PCC must record:

- the list of instances $\mathcal{K}(u, d, a)$(or the phantom instance if applicable)
- each instance's parameters $(\mu_k, r_k, t_k, \ell_k, c_k, e_k)$and evidence sources
- the method used to assign time horizons t_k
- whether $\ell_k$is baseline-conditional or scenario-conditional
- any equity/resilience justification and counterfactual audit when $e_k \neq 1$

B.2.8 Output and handoff to saturation

The output of B.2 is $\tilde{I}_{u,d}^{dir}(a)$for each active cell. The next step (B.3) maps this unsaturated value into the canonical bounded direct impact:

- $I_{u,d}^{dir}(a) \in [-1, +1]$

using the governed saturation function.

B.3 Saturation (Bounded Mapping to $-1, +1$)

Purpose. Pre-saturation aggregation (B.2) produces an unbounded real number $\tilde{I}_{u,d}^{dir}(a) \in \mathbb{R}$. Saturation maps this into the canonical bounded scale $[-1, +1]$while preserving sign and ensuring comparability across cells.

B.3.1 Saturation function

Tier 4 override (Normative). Tier 4 MUST NOT compute tanh(·) at runtime. Tier 4 saturation MUST use SAT_LUT_FP_V1 under NDP_FIXEDPOINT_V1.

MathGov uses smooth hyperbolic tangent saturation:

$$sat_\beta(x) := \tanh(\beta x)$$

Where:

1. $x \in \mathbb{R}$
2. $sat_\beta(x) \in (-1, +1)$
3. $\beta > 0$controls steepness

Default for **direct** (pre-propagation) saturation:

- $\beta = 2$

B.3.2 Direct impact after saturation

$$I_{u,d}^{dir}(a) := sat_\beta \left(\tilde{I}_{u,d}^{dir}(a)\right)$$

This produces a canonical, bounded direct impact score per cell, suitable for propagation (Appendix B.6 onward).

B.3.3 Rationale and governance

Saturation prevents:

2. runaway totals from many small instances,
3. disproportionate domination by one cell with extreme unbounded sum,
4. scale inconsistencies across dimensions.

Any override of βmust be PCC-declared and applied consistently across all options in the decision.

B.4 Union–Dimension Cell Set and Indexing

Purpose. Many MathGov computations operate on the full union–dimension grid. This section defines cell indexing and the standard flattening map used for vector operations.

Let:

- unions $u \in \{1, \dots, |\mathcal{U}|\}$
- dimensions $d \in \{1, \dots, |\mathcal{D}|\}$

In the standard 7×7 operational matrix (49 cells), $|\mathcal{U}| = 7$and $|\mathcal{D}| = 7$.

B.4.1 Canonical flattening map (row-major)

$$\phi(u,d) := |\mathcal{D}|(u-1)+d$$

For the 7×7 case:

$$\phi(u,d) = 7(u-1)+d$$

This maps each cell (u, d)into an index in $\{1,\ldots,49\}$.

B.5 Direct Impact Vector (Flattened Form)

Purpose. Define the flattened direct-impact vector used for kernel propagation.

B.5.1 Direct impact vector

$$I_a^{dir} \in [-1,1]^{|\mathcal{U}||\mathcal{D}|}$$

Component definition:

$$I_a^{dir}[\varphi(u,d)] := I_{u,d}^{dir}(a)$$

B.6 Kernel Propagation (Ripple Effects)

Purpose. Propagation maps direct impacts across cells using a governed kernel K. This captures second-order and structural spillovers.

Let:

$$K \in \mathbb{R}^{n\times n}, n = |\mathcal{U}||\mathcal{D}|$$

B.6.1 Propagated impact vector (pre-saturation). Let $I^{dir}(a) \in [-1,1]^{49}$be the flattened **saturated** direct impact vector and $K \in \mathbb{R}^{49\times 49}$the ripple kernel. Define the pre-saturation propagated vector $\tilde{I}^{prop}(a)$as:

Quick mode (first-order):

$$\tilde{I}^{prop}(a) := I^{dir}(a) + K\, I^{dir}(a).$$

Full mode (resummed; requires $\rho(K) < 1$):

$$\tilde{I}^{prop}(a) := (I_{49} - K)^{-1} I^{dir}(a).$$

B.6.2 Post-propagation saturation. Apply elementwise saturation with coefficient $\beta_{prop} > 0$(default $\beta_{prop} = 1$):

$$\bar{I}^{prop}(a) := \tanh (\beta_{prop} \, \tilde{I}^{prop}(a)),$$

yielding $\bar{I}^{prop}(a) \in [-1,1]^{49}$.

B.6.3 Reshaping back to cell form

$$\bar{I}^{prop}_{u,d}(a) := I^{prop}_{a}[\varphi(u,d)]$$

B.7 Worst-Off Subgroup Operator (Rights Evaluation)

Purpose. Rights constraints are evaluated on the **worst-off subgroup** in the relevant population, not on averages.

Let $g \in \mathcal{G}(u)$index subgroups within union u, and let $\bar{I}^{prop}_{u,d}(a; g)$be the propagated impact for subgroup g.

Define:

Tier-4 Subgroup Enumeration (Normative addendum)

For Tier-4 Pilot-Executable runs, subgroup sets are not optional. The PCC MUST (a) explicitly list G_{u,d} for every rights-covered cell, and (b) provide subgroup-conditioned impact totals per option. Tier-4 requires |G_{u,d}| ≥ 2 unless the union is single-entity for the decision; otherwise downgrade (audit_flag: SUBGROUP_ENUM_MIN_FAIL_TIER4).

In Tier-4 replay, subgroup aggregates are treated as declared inputs; implementers MUST NOT invent subgroup partitions not present in the PCC.

$$\bar{I}^{prop,worst}_{u,d}(a) := \min_{g \in \mathcal{G}(u)} \bar{I}^{prop}_{u,d}(a; g)$$

This operator is used for rights-floor evaluation in NCRC.

B.8 Rights Violation Depth and Rights Margin

Purpose. Convert rights-floor thresholds into a computable "violation depth" per right.

Let $r \in \mathcal{R}$index rights. Each right has:

5. a covered cell set $C_r \subseteq \mathcal{U} \times \mathcal{D}$
6. a rights floor threshold τ_r(canonical default often 0, but governance-chosen by right)

Define the **rights margin**:

$$M_r(a) := \min_{(u,d)\in C_r} \bar{I}_{u,d}^{prop,worst}(a) - \tau_r$$

Define the **violation depth** (nonnegative):

$$VD_r(a) := (-M_r(a))^+$$

Where $(x)^+ = \max(x, 0)$.

Interpretation:

- $M_r(a) \geq 0$ ⇒right $r$is satisfied (no violation)
- $M_r(a) < 0$ ⇒right $r$is violated with depth $VD_r(a) > 0$

B.9 NCRC Pass/Fail Predicate

Purpose. Determine whether an option passes the non-compensatory rights constraint.

$$NCRC(a) := \bigwedge_{r\in\mathcal{R}} (M_r(a) \geq 0)$$

Option passes NCRC if **all** rights margins are nonnegative under worst-off evaluation.

B.10 Catastrophe Cell Set and Catastrophe Weight Simplex

Purpose. TRC focuses only on a governance-chosen "catastrophe cell set" C_{cat} and aggregates tail risk using weights over that set.

Let:

$$C_{cat} \subseteq \mathcal{U} \times \mathcal{D}$$

Let weights $\omega_{u,d}$ be defined for $(u,d) \in C_{cat}$ with:

$$\omega_{u,d} \geq 0, \sum_{(u,d) \in C_{cat}} \omega_{u,d} = 1$$

Floors may be applied:

$$\omega_{u,d} \geq \omega_{u,d}^{floor}, \sum_{(u,d) \in C_{cat}} \omega_{u,d}^{floor} \leq 1$$

B.11 Scenario Set and Probability Normalization

Purpose. Tail risk is evaluated over scenarios $s \in S$ with probabilities p_s.

Let S be a finite scenario set. Define:

$$p_s \geq 0, \sum_{s \in S} p_s = 1$$

If robust probability bounds are used, see B.15.

B.12 Scenario-Conditioned Propagation

Purpose. When scenarios change impacts or kernel behavior, evaluate propagated impacts conditional on scenario s.

Let $I_a^{dir}(s)$ denote the direct-impact vector under scenario s. Then:

$$\tilde{I}_a^{prop}(s) := K(s)\, I_a^{dir}(s)$$

If kernel is not scenario-dependent, then $K(s) = K$.

Apply post-propagation saturation:

$$I_a^{prop}(s) := sat_{\beta_{prop}} \left(\tilde{I}_a^{prop}(s)\right)$$

And reshape:

$$\bar{I}_{u,d}^{prop}(a \mid s) := I_a^{prop}(s)[\varphi(u,d)]$$

B.13 Harm Function for Catastrophe Evaluation

Purpose. TRC uses harms (not benefits) on catastrophe cells.

Define scenario-conditioned harm:

$$Harm_{u,d}(a,s) := \max \left(-\bar{I}_{u,d}^{prop}(a \mid s), 0\right)$$

B.14 Scenario Loss Function (Bounded-impact form)

Purpose. Define a scenario loss $L(a,s)$for CVaR evaluation based on catastrophe harms.

Scenario loss (bounded-impact form): L(a, s) = Σ_{(u,d)∈C_cat} ω_{u,d} · h_{u,d}(a|s), where h_{u,d}(a|s) = max(0, −Ī^prop_{u,d}(a|s)) and Σ_{(u,d)∈C_cat} ω_{u,d} = 1.$L(a,s) := \sum_{(u,d)\in C_{cat}} \omega_{u,d}\ Harm_{u,d}(a,s)$

This loss is in $[0, 1]$when harms are computed from bounded impacts.

B.15 TRC Loss Mode and Raw-Indicator Loss (normative)

Define trc_mode ∈ { 'bounded_impact', 'raw_indicator' } and let tier = implementation_tier from the PCC.

Mode lock (tier rule). For tier ≤ 3, trc_mode MAY be 'bounded_impact' or 'raw_indicator'. For tier ≥ 4, trc_mode MUST be 'raw_indicator'. A Tier-4 PCC that sets trc_mode='bounded_impact' is invalid.

Admissibility lock (tier rule). When tier ≥ 4, bounded-impact loss L_impacts(a,s) MAY be computed for diagnostics only and MUST NOT be used to determine TRC admissibility. TRC admissibility MUST be computed using L_raw(a,s) derived from the AF-BASE registry (or AF-EXT only if the PCC declares a governed extension C_ext).

Raw-indicator mapping. For each catastrophe-bearing cell c and indicator $j \in AF(c)$, record $x_j(a,s)$ in its native units and map to bounded loss:

$$\ell_j(a,s) = \mathrm{clip}\left((x_j(a,s) - x_{onset,j}) / (x_{max,j} - x_{onset,j}) , 0 , 1 \right).$$

Raw-indicator aggregation. Aggregate within a cell by AF default (worst-case):

$$L_c(a,s) = \max_{j \in AF(c)} \ell_j(a,s).$$

Raw-indicator loss. Define:

$$L_{raw}(a,s) = \Sigma_{c \in C_{cat}} \omega_c \cdot L_c(a,s), \quad L_{raw}(a,s) \in [0,1].$$

Tier-4 requirement. For tier ≥ 4, AF-BASE MUST be non-placeholder (G.1A.4) and MUST include at least one indicator for each cell in C_{cat}^{base}. If any required indicator is missing or any mapping parameter is undefined, the run is invalid.

Then compute $CVaR_\alpha(L_{raw})$ and apply the TRC corridor threshold τ_{TRC} exactly as in B.13 and B.17–B.18, with α and τ_{TRC} taken from the Tier Requirements Matrix (§4.4.5) or a hash-bound parameter registry referenced by the PCC.

B.16 Value-at-Risk (VaR)

Purpose. Define VaR for scenario losses.

Let $L(a)$ be the random variable over scenarios with outcomes $L(a,s)$ and probabilities p_s. VaR at level $\alpha \in (0,1)$ is:

$$VaR_\alpha(L(a)) := \inf \{ z \in \mathbb{R} : \mathbb{P}(L(a) \le z) \ge \alpha \}$$

For discrete scenarios:

$$VaR_\alpha(L(a)) := \min \left\{ z : \sum_{s : L(a,s) \le z} p_s \ge \alpha \right\}$$

B.17 Conditional Value-at-Risk (CVaR)

Purpose. CVaR measures expected loss in the worst tail beyond VaR.

$$CVaR_\alpha(L(a)) := \mathbb{E}[L(a) \mid L(a) \ge VaR_\alpha(L(a))]$$

Equivalent optimization form (useful for implementation):

$$CVaR_{\alpha}(L(a)) = \min_{\eta \in \mathbb{R}} \left(\eta + \frac{1}{1-\alpha} \mathbb{E}[(L(a) - \eta)^{+}] \right)$$

Discrete scenario version:

$$CVaR_{\alpha}(L(a)) = \min_{\eta \in \mathbb{R}} \left(\eta + \frac{1}{1-\alpha} \sum_{s \in S} p_s \ (L(a,s) - \eta)^{+} \right)$$

B.18 TRC Corridor Constraint (Pass/Fail)

Purpose. TRC is a hard admissibility filter: options exceeding the corridor are forbidden.

Let τ_{TRC} be the governance-chosen corridor threshold. Then:

$$TRC(a) := (CVaR_{\alpha}(L(a)) \le \tau_{TRC})$$

B.18.1 Scenario-Aware RLS (When Scenario Evaluation Is Enabled)

Purpose. When scenario evaluation is used for welfare impacts, RLS uses scenario-expected propagated impacts.

Define scenario expectation of propagated impacts:

$$\mathbb{E}_S \left[\bar{I}_{u,d}^{prop}(a) \right] := \sum_{s \in S} p_s \ \bar{I}_{u,d}^{prop}(a \mid s)$$

Then scenario-aware RLS is:

$$RLS(a) = \sum_{u,d} w_u \ v_d \ m_{u,d} \ \mathbb{E}_S \left[\bar{I}_{u,d}^{prop}(a) \right]$$

This is identical to the baseline RLS with $\bar{I}^{prop}$ replaced by its scenario expectation.

B.19 Robust TRC (Worst-Case over Probability Set)

Purpose. When scenario probabilities are uncertain, apply a robust constraint over a set $\mathcal{P}$ of admissible distributions.

Let $\mathcal{P}$ be a governance-defined set of distributions over S. Then:

$$TRC_{rob}(a) := \sup_{p \in \mathcal{P}} CVaR_{\alpha,p}(L(a)) \leq \tau_{TRC}$$

The PCC must define $\mathcal{P}$ (bounds, floors, and justification).

B.20 Positive-Part, Clip, and Clamp Operators

Purpose. Standard operators used throughout the paper.

$$(x)^+ := \max(x, 0)$$

$$clip(x, L, U) := \max(L, \min(U, x))$$

$$clamp_0^1(x) := \max(0, \min(1, x))$$

B.21 Admissibility Predicate and Admissible Option Set

Purpose. Formal deontic separation: admissibility is not a weighted sum.

$$Admissible(a) := NCRC(a) \land TRC(a)$$

$$\mathcal{A}_{adm} := \{a \in \mathcal{A} \mid Admissible(a)\}$$

B.22 Containment "Positive Set" (Integrity Gate)

Purpose. Containment prevents sub-union gains from degrading higher unions beyond allowable bounds. Its exact parameters are governance-chosen; the PCC must declare them.

Let $\mathcal{U}^+$ denote unions whose protection has containment priority (e.g., Biosphere, Humanity/CMIU), with parameters in the PCC.

Define a containment pass/fail predicate generically:

$$Containment(a) := \text{PCC-defined test over designated unions and dimensions}$$

Then the containment-positive subset:

$$\mathcal{A}_{cont} := \{a \in \mathcal{A}_{adm} \mid Containment(a)\}$$

If containment is used as an escalation gate rather than a hard filter, the PCC must record override signatures and mitigation obligations.

B.23 Hollowing-Out Index (HOI) (Explicit EMA Form)

Purpose. HOI tracks divergence between welfare score improvements and coherence/health indicators, detecting "paper gains" that erode structural integrity.

Define exponential moving average:

Initialization: for any EMA_λ stream, set EMA_λ(x_0) := x_0 (first observation). For HOI in a new deployment with no prior period, set ΔRLS_0 := 0 and ΔUCI_0 := 0, compute HOI_0 using those zero deltas, and begin EMA from the first observed Δ values in the next period.

$$EMA(X)_t := \gamma X_t + (1-\gamma)EMA(X)_{t-1}$$

Default smoothing parameter from half-life $h$review periods:

$$\gamma := 1 - 2^{-1/h}$$

Default $h = 3$gives $\gamma \approx 0.206$.

Define HOI:

$$HOI_t := EMA(RLS)_t - EMA(UCI)_t$$

B.24 Maximin Rights Margin (Emergency-Mode Comparator)

Purpose. When all options violate at least one right, emergency mode selects options by minimizing rights-violation severity first.

Define the minimum rights margin for an option:

$$MRM(a) := \min_{r \in \mathcal{R}} M_r(a)$$

Higher is better. If all are negative, choose the option that makes $MRM(a)$least negative (closest to zero), with secondary criteria (e.g., CVaR) applied as defined in emergency procedures.

B.25 Summary of Appendix B Outputs

Appendix B provides the computable definitions for:

- instance aggregation $\tilde{I}_{u,d}^{dir}(a)$(B.2)
- saturation $I_{u,d}^{dir}(a)$(B.3)
- flattening/vectorization (B.4–B.5)
- propagation $\bar{I}_{u,d}^{prop}(a)$(B.6)
- worst-off evaluation (B.7)
- rights margins and NCRC pass/fail (B.8–B.9)
- catastrophe loss, VaR/CVaR, TRC corridor (B.10–B.18)
- scenario-aware RLS (B.18.1)
- robust TRC (B.19)
- admissibility sets, containment sets (B.21–B.22)
- HOI explicit computation (B.23)
- emergency comparator (B.24)
- shared operators (B.20)

Appendix C: Canonical Rights Specification (NCRC)

Authoritative rights coverage mapping (normative). Appendix C.3.7 is the single authoritative source for rights coverage sets C_r. Any other summary table must match C.3.7 exactly.

This appendix specifies the canonical rights used by the NCRC. These rights are implemented as lexicographic admissibility constraints: options that violate any canonical right are inadmissible except under explicitly declared Emergency Mode.

C.1 Scope and Coverage Semantics

Rights-bearing unions. Canonical rights apply primarily to unions that contain sentient moral patients and/or governance agents. Let:

The Biosphere union is protected under Ecological Integrity (ECOL) and via TRC rather than being treated as a generic "rights bearer" for all human-style rights.

Impacts used for NCRC. NCRC is evaluated on the worst-off subgroup post-propagation, post-saturation normalized impacts Ī^rights_(*u*,*d*)(*a*). Thresholds are stated on the [−1, +1] scale.

Canonical thresholds. Thresholds may be contextually tightened (more protective) but may not be loosened below canonical values without Charter-level revision.

C.2 Canonical Rights Table

Right	Code	Threshold	Coverage Set	Measurement Anchors
Life	LIFE	θ_LIFE = −0.90	Health cells for $u \in \{1,\ldots,6\}$	Mortality risk; excess deaths; acute lethality
Bodily Integrity	BODY	θ_BODY = −0.70	Health cells for $u \in \{1,\ldots,6\}$	Severe injury/disability; torture; trauma
Liberty	LBTY	θ_LBTY = −0.65	Agency cells for $u \in \{1,\ldots,6\}$	Arbitrary detention; forced labor; coercion
Basic Needs	NEED	θ_NEED = −0.50	Material cells for $u \in \{1,\ldots,6\}$	Food insecurity; homelessness; deprivation

Right	Code	Threshold	Coverage Set	Measurement Anchors
Dignity	DIGN	θ_DIGN = −0.55	{(u, Social): u ∈ {1,2,3,4,5,6}} ∪ {(u, Agency): u ∈ {1,2,3,4,5,6}}	Discrimination; humiliation; exclusion; coercive loss of voice/autonomy
Due Process	PROC	θ_PROC = −0.45	Agency cells for *u* ∈ {1,…,6}	Fair hearing denial; arbitrary enforcement
Information	INFO	θ_INFO = −0.40	Knowledge cells for *u* ∈ {1,…,6}	Censorship; misinformation; access denial
Ecological Integrity	ECOL	θ_ECOL = −0.65	Biosphere-Environment (required); Humanity-Environment (required)	Planetary boundaries; biodiversity

C.3 Rights coverage sets C_r: explicit union–dimension mapping and scope note

This appendix specifies the **canonical rights coverage sets** used by the Non-Compensatory Rights Constraint (NCRC). Each right r is evaluated on a declared set of union–dimension cells $C_r \subseteq U \times D$. These coverage sets define **where** a right applies in the 7×7 welfare matrix and make NCRC independently implementable.

C.3.1 Canonical unions and dimensions

Unions U are ordered:

$$U = \{1, \dots, 7\}$$
$$= \{\text{Self, Household, Community, Organization, Polity, Humanity/CMIU, Biosphere}\}.$$

Dimensions D are ordered:

$$D = \{1, \ldots, 7\} = \{\text{Material, Health, Social, Knowledge, Agency, Meaning, Environment}\}.$$

When this appendix refers to "rights-bearing unions," it means $U_{rb} = \{1,2,3,4,5,6\}$(Self through Humanity/CMIU). Biosphere ($u = 7$) is not treated as a rights-bearing moral patient in NCRC, but it can be protected as a **rights-relevant life-support system** through ECOL and through TRC and containment.

C.3.2 Scope note: what coverage sets mean and what they do not mean

Coverage sets are **not weights** and do not create compensatory tradeoffs. They specify:

1. which union–dimension cells are relevant to a given right, and
2. where the NCRC feasibility check must look for violations.

A cell being included in C_r means that harm in that cell can trigger a rights violation for right r if it crosses the governed rights threshold θ_r, under the scenario and subgroup rules specified in Section 6.

Coverage sets must be:

- **declared** (these defaults are canonical unless overridden),
- **stable across compared options**, and
- **auditable** (the PCC must state any deviation and justification).

C.3.3 Canonical rights set

MathGov uses the following canonical right labels:

$$R = \{\text{LIFE, BODY, LBTY, NEED, DIGN, PROC, INFO, ECOL}\}.$$

Where:

- LIFE = life preservation / avoid severe mortality harm
- BODY = bodily integrity / avoid severe physical harm, coercion, violence
- LBTY = liberty / freedom of movement and autonomy within rights floor
- NEED = basic needs / survival and essential access
- DIGN = dignity / non-degradation, non-discrimination, basic respect
- PROC = due process / fairness, transparency, contestability in coercive systems

- INFO = information integrity / consent-relevant truthfulness and privacy where needed
- ECOL = ecological life-support integrity as a rights-relevant constraint

C.3.4 Explicit coverage mapping table (canonical default)

Notation. (u, d)denotes union $u$and dimension d. “All rights-bearing unions” means $u \in \{1,2,3,4,5,6\}$.

Canonical mapping principle.

- Rights that protect direct sentient harms map primarily to **Health**, **Agency**, and **Social** cells across rights-bearing unions.
- Rights involving access and deprivation map to **Material** and **Health**.
- Rights involving institutional fairness map to **Social**, **Knowledge**, and **Agency** at governance-relevant unions (Organization, Polity, Humanity/CMIU).
- ECOL maps to Environment cells that represent life-support integrity affecting rights-bearing beings.

Table C.3-A. Rights coverage sets C_r(canonical default)

Right r	**Coverage set** C_r**(union–dimension pairs)**
LIFE	Defined in Appendix C.3.7 (authoritative auditor mapping table).
BODY	Defined in Appendix C.3.7 (authoritative auditor mapping table).
LBTY	Defined in Appendix C.3.7 (authoritative auditor mapping table).
NEED	Defined in Appendix C.3.7 (authoritative auditor mapping table).
DIGN	Defined in Appendix C.3.7 (authoritative auditor mapping table).
PROC	Defined in Appendix C.3.7 (authoritative auditor mapping table).
INFO	Defined in Appendix C.3.7 (authoritative auditor mapping table).
ECOL	Defined in Appendix C.3.7 (authoritative auditor mapping table).

C.3.5 Interpretation notes (to prevent misreadings)

1. **Why LIFE includes** $(6, \text{Environment})$**.** Life at scale depends on shared planetary and civilizational life-support functions. This inclusion prevents "death by environment collapse" from evading LIFE protections at humanity scale.
2. Why ECOL maps to Environment only. ECOL is a life-support integrity constraint anchored to environmental conditions. Downstream human health harms are protected by LIFE/BODY/NEED and by TRC, which remain separate admissibility corridors.
3. Non-duplication note. ECOL is not a second copy of LIFE/BODY. It constrains environmental life-support integrity directly; human mortality and bodily harms remain governed by LIFE and BODY, and catastrophic tails remain governed by TRC.

C.3.6 Governance and override rule

Any deviation from Table C.3-A must be declared in the PCC and must include:

- the modified set C_r for each affected right,
- the rationale (including any domain-specific rights constraints),
- and confirmation that the same C_r is applied across all options.

Appendix C.3.7: Rights coverage sets (C_r) — auditor mapping table

This table makes the NCRC independently implementable by specifying, for each right r, the coverage set C_r ⊆ U × D (union–dimension cells) to be checked against the rights threshold θ_r. Defaults apply unless overridden and documented in the PCC.

Canonical unions U (1–7): Self, Household, Community, Organization, Polity, Humanity/CMIU, Biosphere.

Canonical dimensions D: Material, Health, Social, Knowledge, Agency, Meaning, Environment.

Rights-bearing unions default: U_rb = {Self, Household, Community, Organization, Polity, Humanity/CMIU} (i.e., unions 1–6).

Right r	Coverage set C_r (union, dimension pairs)	Scope / audit note
LIFE	{(u, Health): u ∈ {1,2,3,4,5,6}} ∪ {(6, Environment)}	Protects severe mortality / life-ending harm across rights-bearing unions, plus

		humanity-scale life-support integrity.
BODY	{(u, Health): u ∈ {1,2,3,4,5,6}}	Protects bodily integrity: severe injury, violence, coercive harm with physical consequences.
LBTY	{(u, Agency): u ∈ {1,2,3,4,5,6}} ∪ {(u, Social): u ∈ {3,4,5,6}}	Protects liberty/autonomy primarily via Agency; adds Social for community–humanity where liberty is institutionally mediated.
NEED	{(u, Material): u ∈ {1,2,3,4,5,6}} ∪ {(u, Health): u ∈ {1,2,3,4,5,6}}	Protects basic needs and essential access; includes Health because deprivation often manifests as health harm.
DIGN	{(u, Social): u ∈ {1,2,3,4,5,6}} ∪ {(u, Agency): u ∈ {1,2,3,4,5,6}}	Protects non-degradation and non-discrimination; links to social treatment and autonomy/voice.
PROC	{(u, Agency): u ∈ {4,5,6}} ∪ {(u, Knowledge): u ∈ {4,5,6}} ∪ {(u, Social): u ∈ {4,5,6}}	Protects due process and contestability in coercive systems; applies where formal governance structures operate.
INFO	{(u, Knowledge): u ∈ {1,2,3,4,5,6}} ∪ {(u, Agency): u ∈ {1,2,3,4,5,6}}	Protects consent-relevant truthfulness and privacy (as applicable); ties to knowledge integrity and agency.
ECOL	{(6, Environment), (7, Environment)}	Protects ecological life-support integrity at humanity and biosphere scale; downstream health is covered by LIFE/BODY/NEED and TRC.

Governance rule: Any deviation from these defaults must be declared in the PCC, applied consistently across all options, and justified as preserving the intended protection scope. Coverage sets are not weights and do not create compensatory tradeoffs.

C.4 Rights-to-Cells Rationale Table

Right	Cells Covered	Rationale
LIFE	Health cells (U_1 through U_6)	Life protection requires health system integrity across all scales; mortality risk is fundamentally a health outcome
BODY	Health cells (U_1 through U_6)	Bodily integrity is a health dimension concern; severe injury and disability are health impacts
LBTY	Agency cells (U_1 through U_6)	Liberty is the capacity for autonomous action; coercion and detention restrict agency
NEED	Material cells (U_1 through U_6)	Basic needs are fundamentally material; food, shelter, and essential services are material resources
DIGN	Social and Agency cells	Dignity is operationalized as protection against social degradation/exclusion and against coercive loss of voice/autonomy; existential “Meaning” is scored in welfare (RLS) but is not used as a rights-floor trigger in the canonical mapping.
PROC	Agency cells (procedural)	Due process is a specific form of agency in institutional contexts; fair hearing enables effective action
INFO	Knowledge cells	Information access is the epistemic capacity foundation; censorship restricts knowledge acquisition
ECOL	Biosphere-Environment (required), Humanity-	Ecological integrity is environmental by definition; humanity-level extension captures civilization-scale environmental dependence

Right	Cells Covered	Rationale
	Environment (required)	

C.5 Threshold Calibration Protocol

Each rights threshold is calibrated through a three-step process:

C.5.1 Step 1: Normative Anchor Identification

Right	Harm Category	Normative Source
LIFE	Near-certain or highly probable death	UNHCR emergency mortality thresholds (doubling of baseline mortality)
BODY	Severe injury, disability, torture	Sphere Standards minimum thresholds
LBTY	Arbitrary detention, forced labor	Freedom House "partly free" threshold
NEED	Severe food insecurity, homelessness	FAO FIES severe threshold
DIGN	Systematic humiliation, dehumanization	UDHR dignity provisions
PROC	Denial of fair hearing	World Justice Project Rule of Law Index
INFO	Systematic censorship	Press freedom indices
ECOL	Planetary boundary transgression	Rockström et al. planetary boundaries framework

C.5.2 Step 2: Indicator Mapping

Example: LIFE threshold derivation

Reference indicator: Excess mortality rate (deaths per 1,000 population above baseline)

Reference class: WHO Global Health Observatory mortality data

Anchoring:

1. 5th percentile (best): 0 excess deaths per 1,000
2. 95th percentile (worst): 50+ excess deaths per 1,000 (severe epidemic/conflict)

The threshold $\theta_{LIFE} = -0.90$ corresponds to approximately 45 excess deaths per 1,000, representing near-certain widespread mortality if the decision proceeds.

C.5.3 Step 3: Philosophical Justification

Each threshold placement reflects convergent moral intuitions from:

1. Human rights jurisprudence (UDHR, ICCPR, ICESCR)
2. Humanitarian standards (Sphere, UNHCR)
3. Capability theory (Nussbaum's central capabilities)
4. Overlapping consensus across major ethical traditions

C.6 Threshold Sensitivity Analysis Requirement

Before adopting thresholds, conduct sensitivity analysis:

1. Vary each threshold by ±0.05
2. Apply to a test set of at least 20 decision scenarios
3. Document: How many decisions change admissibility status?
4. If >30% of decisions are sensitive to ±0.05 variation, provide additional justification for the chosen threshold

C.7 Threshold Revision Procedure

Thresholds may be revised only through charter-level governance:

1. Proposal with documented justification grounded in new evidence
2. Sensitivity analysis showing effects of proposed change
3. Supermajority vote in governance body (default: 2/3)
4. Independent review panel sign-off

5. Public disclosure and version increment

C.8 Emergency Mode Priority Order

Under NCRC Emergency Mode (§6.4), when no option satisfies all rights constraints, the system lexicographically minimizes violations in the following priority order:

1. LIFE (highest priority)
2. BODY
3. ECOL
4. LBTY
5. NEED
6. DIGN
7. PROC
8. INFO (lowest priority)

Normative status (authoritative). This priority order is the single authoritative ordering for lexicographic rights minimization in Emergency Mode. Any deviation requires Charter amendment and a new versioned registry hash. Any PCC using a non-canonical ordering MUST include audit_flag RIGHTS_PRIORITY_OVERRIDE.

This ordering reflects the relative severity and irreversibility of different rights violations.

Appendix D: TRC Parameter Defaults

This appendix provides default TRC parameters and the minimum scenario set required for a defensible tail-risk evaluation.

D.1 Catastrophe Cell Set (Base)

corresponding to Humanity/CMIU-Health, Humanity/CMIU-Environment, and Biosphere-Environment.

D.2 Catastrophe weights and minimum-weight floor (TRC aggregation weights)

This appendix provides the **canonical definition and construction** of catastrophe-cell weights used by the Tail-Risk Constraint (TRC). It is the appendix counterpart to §7.2.4 and must remain **identical in formula and meaning**.

D.2.1 Catastrophe cell set C_{cat}

Let $U = \{1, \dots, 7\}$be the unions and $D = \{1, \dots, 7\}$be the welfare dimensions. Define the catastrophe cell set:

$$C_{cat} \subseteq U \times D,$$

the set of union–dimension cells in which catastrophic failure is meaningful and must be evaluated under TRC.

Unless otherwise declared in the PCC, MathGov uses the canonical default:

$$C_{cat}^{default} := \{(u, \text{Health}) : u \in \{1, \dots, 6\}\} \cup \{(7, \text{Environment})\}.$$

Any extension or modification of $C_{cat}$is a governed decision and must be declared in the PCC, including the rationale and the complete resulting set.

D.2.2 Catastrophe weights $\omega_{u,d}$and normalization

Define nonnegative catastrophe weights over catastrophe-bearing cells:

$$\omega_{u,d} \geq 0 \forall (u, d) \in C_{cat},$$

with normalization:

$$\sum_{(u,d) \in C_{cat}} \omega_{u,d} = 1.$$

Default. If the PCC does not specify weights, use the uniform default:

$$\omega_{u,d}^{default} = \frac{1}{| C_{cat} |} \forall (u, d) \in C_{cat}.$$

Governed reweighting. The PCC may specify a non-uniform $\omega$to reflect decision-context salience, provided:

- the weights are nonnegative,
- they sum to 1 over C_{cat}, and

- the same ω is applied across all compared options.

D.2.3 Anti-capture minimum weight floor (feasible under extensions)

To prevent "catastrophe-weight capture" (artificially driving a key catastrophe cell's weight toward zero), MathGov enforces a **feasible per-cell minimum**:

$$\omega_{u,d} \geq \omega_{\min}(|C_{cat}|) \forall (u,d) \in C_{cat}.$$

Define:

$$\omega_{\min}(|C_{cat}|) := \min\left(\frac{\eta}{|C_{cat}|}, 0.05\right),$$

where $\eta \in (0,1]$ is a governed slack factor (default $\eta = 0.5$).

Feasibility guarantee. This floor remains feasible under any governed extension of C_{cat} because:

$$|C_{cat}| \cdot \omega_{\min}(|C_{cat}|) \leq \eta \leq 1,$$

so the floor cannot force $\sum_{(u,d)\in C_{cat}} \omega_{u,d}$ above 1.

Interpretation of the floor.

- When $|C_{cat}|$ is small, a fixed cap of 0.05 may be active and still feasible.
- When $|C_{cat}|$ becomes large, the cap becomes inactive and the floor smoothly becomes $\eta/|C_{cat}|$, preserving feasibility automatically.

PCC reporting requirement. The PCC must report:

1. $|C_{cat}|$,
2. the chosen η,
3. the resulting $\omega_{\min}(|C_{cat}|)$,
4. whether the 0.05 cap is active,
5. and which cells are at the floor.

D.2.4 Canonical construction procedure (floor-respecting weights)

When a governed weight proposal must satisfy the floor, construct ω using the following steps.

- Propose raw weights $\omega'_{u,d} \geq 0$ over C_{cat} with $\sum \omega' = 1$.
- Apply the floor:

$$\omega''_{u,d} := \max\left(\omega'_{u,d}, \omega_{\min}(|C_{cat}|)\right).$$

- Renormalize:

$$\omega_{u,d} := \frac{\omega''_{u,d}}{\sum_{(i,j)\in C_{cat}} \omega''_{i,j}}.$$

- Verify post-renormalization that $\omega_{u,d} \geq \omega_{\min}(|C_{cat}|)$ holds to within numerical tolerance. If rounding artifacts violate the inequality, apply a rounding-safe adjustment and report the final ω explicitly in the PCC.

This procedure ensures the TRC aggregation weights are feasible, capture-resistant, and independently auditable.

D.3 Default TRC Parameters (Tier-Authoritative + Context Guidance)

D.3.1 Authoritative tier minima (scenario counts)

The Tier Requirements Matrix (§4.4.5) is authoritative for minimum scenario counts:

Tier	Minimum |S| (authoritative)	Notes
Tier 1	N/A	TRC optional / qualitative.
Tier 2	≥ 5 (recommended)	Not a hard floor, but recommended when TRC is used.
Tier 3	≥ 20	Minimum required for Tier 3 claim.
Tier 4 (PilotExecutable)	≥ 50 (default)	Exceptions must be PCC-declared + flagged.

| Tier 4 (Certified) | ≥ 50 + packaged stress tests | Required. |

D.3.2 Context defaults (α and τ_{TRC}) (informative)

These context defaults guide parameter selection but do not override tier minima:

Context	α	τ_{TRC}
Personal	0.90	0.30
Organizational	0.95	0.20
Reversible policy	0.95	0.15
Irreversible policy	0.99	0.10
Existential risk	0.999	0.05

Audit rule: PCC MUST state which context default was used (or why overridden) and MUST separately confirm scenario-count compliance with §4.4.5.

D.4 Mandatory Tail Scenario (MTS) Categories

Category	Code	Minimum Stress Level	Floor Probability
Pandemic/biological	MTS-1	≥30% affected; 6-24 months; healthcare exceeded	$p \geq 0.02$
Climate tipping	MTS-2	≥2 boundaries breached; partial irreversibility	$p \geq 0.02$
Financial collapse	MTS-3	≥50% asset drawdown; credit freeze	$p \geq 0.02$

Category	Code	Minimum Stress Level	Floor Probability
Major conflict	MTS-4	Trade corridor disruption; mobilization risk	$p \geq 0.02$
Infrastructure failure	MTS-5	≥6 month outage for relevant systems	$p \geq 0.02$

D.5 Scenario Specification Template

Each scenario must contain:

1. **Name and ID:** Stable label + unique identifier
2. **Narrative description:** 2-5 sentences describing the world-state and shock
3. **Time horizon:** Planning window and key event timing assumptions
4. **Stressors:** Which systems are stressed (health, climate, finance, conflict, infrastructure)
5. **Parameter hooks:** Quantitative parameters (mortality multiplier, GDP shock, outage duration)

Range semantics (training library): if a scenario lists a range (e.g., outage_duration_hours = 2–6), that range is a scenario-generator hint only. Any executable run MUST instantiate a concrete scenario object with explicit point values for all parameters (declared in the PCC scenario registry). Default instantiation rule for training runs, unless otherwise declared: use the midpoint of each numeric range and round to the nearest admissible unit.

6. **Relevance claim:** Why this scenario is relevant to the decision
7. **Source and provenance:** Literature, datasets, expert elicitation, prior PCCs

D.6 Robust TRC Parameters

D.7 Embedded training scenario library DSL-20-TRAINING-V1 (informative)

Purpose: provide a minimal, fully enumerated scenario set (IDs + probabilities + example parameter hooks) to enable docs-only exercises and Tier ≤ 2 training pilots without requiring a separate ProofPack scenario registry. Use: allowed only when the PCC explicitly sets scenario_library_id = DSL-20-TRAINING-V1 and declares trc_mode = MRC (Tier ≤ 2).

Not allowed: Tier ≥ 3 conformance claims MUST use a governed, decision-relevant scenario registry referenced by hash and meeting tier minima (§4.4.5).

Scenario_ID	MTS_Category	Narrative (2–5 sentences summary)	Example parameter hooks (illustrative)	p
DSL20-01	Operational disruption	Key service or transport pathway is intermittently unavailable, causing delays and missed commitments.	outage_duration_hours=2–6; delay_multiplier=1.5–2.5	0.10
DSL20-02	Safety incident	A serious accident or near-miss occurs, increasing injury risk and triggering protective actions.	injury_risk_multiplier=2–5; protective_mode=ON	0.08
DSL20-03	Weather extreme	Heavy rain or flooding reduces mobility and increases exposure and accident probability.	rain_intensity=high; road_hazard_multiplier=2–4	0.07
DSL20-04	Public health surge	A localized illness surge reduces available labor and increases care burdens.	absenteeism_rate=10–25%; care_load_multiplier=1.2–1.5	0.06
DSL20-05	Infrastructure outage	Power or network outage disrupts coordination, payments, or	outage_hours=1–8; comms_reliability=low	0.06

		safety systems for multiple hours.		
DSL20-06	Price shock	Fuel and essential prices spike rapidly, increasing cost burden and reducing affordability.	fuel_price_change=+15–40%; cost_multiplier=1.1–1.4	0.05
DSL20-07	Policy/regulatory change	A rule change alters access, route permissions, or compliance duties, increasing friction and risk.	compliance_friction=+1–3 steps; access_reduction=10–30%	0.05
DSL20-08	Security incident	Theft/fraud or digital compromise occurs, increasing losses and reducing trust in transactions.	loss_fraction=0.05–0.20; trust_factor=down	0.05
DSL20-09	Supply chain disruption	Parts, maintenance, or critical supplies are delayed, increasing downtime probability.	lead_time_days=3–14; downtime_prob=up	0.05
DSL20-10	Labor/availability shock	Key worker becomes unavailable (illness/family), causing schedule	availability_drop=1 person; schedule_slack=down	0.05

		gaps and increased stress.		
DSL20-11	Demand variance	Demand volatility forces schedule compression or overtime, increasing fatigue and error rate.	demand_volatility=high; fatigue_multiplier=1.2–1.6	0.04
DSL20-12	Financial constraint	Unexpected expense or income shortfall reduces buffer, forcing riskier choices.	buffer_days=down; forced_tradeoffs=up	0.04
DSL20-13	Social conflict	Local tensions or harassment risk increases on certain routes or times, reducing safety and well-being.	harassment_risk=up; route_constraints=increase	0.04
DSL20-14	Environmental hazard	Air quality or heat event reduces health reserve and increases vulnerability during travel.	AQI=unhealthy; heat_index=high; health_reserve=down	0.04
DSL20-15	Technology failure	Core device/app/system fails, reducing reliability of planning and causing coordination failures.	failure_rate=up; fallback_latency=up	0.04

DSL20-16	Macroeconomic downturn	Broad slowdown reduces income stability; buffers shrink; competition for stable work increases.	income_variance=up; buffer_days=down	0.06
DSL20-17	Critical dependency failure	Primary backup option fails simultaneously, reducing redundancy and increasing tail risk.	redundancy_level=1; backup_failure_prob=up	0.03
DSL20-18	Legal/liability event	A liability event or enforcement action creates sudden penalties or exclusion from a service.	penalty_amount=high; exclusion_prob=up	0.03
DSL20-19	Information reliability failure	Information is unreliable (misleading ETA, false availability), causing cascading lateness.	ETA_error=high; miscoordination_prob=up	0.03
DSL20-20	Compound shock	Two or more shocks co-occur (e.g., rain + outage), amplifying tail risks beyond single scenarios.	shock_combo={rain,outage}; tail_multiplier=1.5–3	0.03

When probability estimates vary by more than 2× across credible sources:

Probability bounds method:

Appendix E: UCI measurement operationalization (union-specific structural indicators)

This appendix defines the canonical operationalization of UCI for Tier 4 decisions. It specifies union-appropriate structural indicators for each UCI component , and it provides default measurement cadence guidance. The intent is not to force a single dataset, but to provide a complete implementable protocol that preserves Tier-2 structural independence from welfare scoring.
E.1 General rule: structural indicators and independence
For each union , UCI is computed from structural and process indicators that are distinct from the welfare indicators used to generate the RLS impact matrix. UCI components are defined on , and option impacts on UCI components are represented as bounded deltas in using the same impact-instance pipeline as §5.2–§5.4 (indicator selection, reference class anchoring, instance aggregation, and saturation), but applied to structural indicators rather than welfare indicators.

Baseline recording requirement (Tier 4). The PCC must record and the data sources used to compute them.

Impact estimation rule (Tier 4). For an option , analysts estimate from structural indicators using reference-class anchoring and the magnitude calibration rules in §5.4. The PCC must list the indicator set used for each component and the declared reference classes.

Independence rule (Normative). UCI MUST be structurally independent from RLS. Deriving UCI or ΔUCI directly from welfare-cell impacts (Ī^prop) or from any transformation of RLS inputs is PROHIBITED at Tier ≥ 3. UCI-proxy is PERMITTED ONLY at Tier ≤ 2 and MUST be explicitly labeled in the PCC; any Tier ≥ 3 use MUST downgrade the tier claim and include audit_flag UCI_PROXY_USED_TIER_VIOLATION.

E.2 Component meanings (canonical)

- **Cohesion H_u:** internal trust, shared identity, coordination willingness, integrity of relationships.
- **Flow F_u:** functional throughput of essential processes, coordination efficiency, reduced friction and bottlenecks.
- **Resilience R_u:** redundancy, robustness to shocks, recovery speed, adaptive capacity.

- **Equity E_u:** fairness in distribution of burdens/benefits and voice, protection against systematic exclusion.

E.3 Union-specific operational indicator families (defaults)

This section specifies the canonical default indicator families for each union's UCI components: **Cohesion** H_u, **Flow** F_u, **Resilience** R_u, and **Equity** E_u. These indicators are **structural and process measures**, designed to remain **distinct from the welfare indicators** used to score the RLS matrix. Substitutions are permitted, but substitutions must preserve the component meaning (Section E.2), must be declared in the PCC, and must include the sign convention, baseline method, cadence, and data source.

General measurement rules (apply to all unions).

1. Independence verification (Normative). The PCC MUST attest that the structural indicators used for UCI are distinct from the welfare indicators used for RLS. If structural-indicator availability forces a proxy mapping, the run MUST be labeled UCI-proxy and is PERMITTED ONLY at Tier ≤ 2; Tier ≥ 3 use is noncompliant and MUST downgrade the tier claim with audit_flag UCI_PROXY_USED_TIER_VIOLATION.
2. **Sign conventions.** Every indicator must be tagged as higher-is-better or higher-is-worse. Where raw measures are higher-is-worse (for example, conflict count), the impact mapping must preserve the sign convention during magnitude calibration.
3. **Baselines.** Each component requires a declared baseline method (rolling mean, rolling median, or fixed baseline snapshot) and a baseline update cadence.
4. **Cadence fit.** Cadence should match the union's typical dynamical timescale. Faster-changing unions can be measured weekly or monthly; slower unions often require quarterly or annual measurement.
5. **Audit readiness.** The PCC must record: indicator definition, unit/proxy, sign convention, baseline method, cadence, and data source for each component, for each union that is materially affected.

E.3.1 Self (U_1)

Cohesion H_1(weekly or monthly).
Cohesion at the Self level captures internal consistency, integrity of commitments, and degree of self-conflict that disrupts stable agency. Default indicator families include:

- **Commitment consistency and follow-through.** Proxies include percent of intended commitments completed within declared time windows, or a weekly follow-through rate on priority actions.

- **Internal conflict index (self-report).** Short-form weekly check-in items capturing competing goals, unresolved internal tension, or repeated self-sabotage patterns, scored using a consistent rubric.
- **Stability of stress signals.** Proxies include variance or instability measures in sleep timing, resting heart rate stability (if available), or other consistent stress markers, interpreted as "lower variance tends to indicate greater internal stability," with explicit sign convention and limitations recorded.

Flow F_1(weekly).
Flow at the Self level represents throughput of essential tasks and reduced friction in execution. Default indicator families include:

- **Time-on-priority ratio.** Percent of time spent on declared top priorities relative to total available time.
- **Interruption load.** Interruptions per hour (or per work block), including context switches that degrade throughput.
- **Task completion latency.** Median delay between intended start and completion for key tasks, with a stable task definition to avoid manipulation.

Resilience R_1(monthly).
Resilience at the Self level captures recovery capacity, buffers, and adaptive stability under disruption. Default indicator families include:

- **Recovery time after disruption.** Days required to return to baseline routine or functioning after a shock event, using a consistent definition of "return to baseline."
- **Buffer-days of essentials.** Buffer capacity measured as days (or weeks) of essential resources, such as food, basic cash buffer, or medication buffer where relevant.
- **Health reserve proxy.** A conservative proxy such as stability of resting heart rate, sleep regularity, or an agreed physiological resilience marker, documented with limitations and the sign convention.

Equity E_1(definition rule).
Set $E_1 = 1$by definition.
This convention prevents double-counting within a single-agent union; equity operators are applied starting at $U_2$where distribution across persons is meaningful.
Optional note (if retained): If any nontrivial "intra-self balance proxy" is used, it must be explicitly labeled as such and must remain structurally distinct from welfare scoring inputs.

E.3.2 Household (U_2)

Cohesion H_2(monthly).
Household cohesion reflects trust, relational stability, conflict resolution capacity, and reliable caregiving coordination. Default indicator families include:

- **Conflict frequency and resolution rate.** Count of conflicts and percent resolved within a defined window using an agreed "resolution" definition.
- **Shared decision participation.** Percent of significant household decisions involving shared input, with a declared "decision significance" threshold.
- **Caregiving stability.** Reliability of caregiving coverage (planned coverage met), including predictable handoffs and continuity of care.

Flow F_2(monthly).
Household flow measures the reliability of household operations and reduction of logistical friction. Default indicator families include:

- **Essential task reliability.** Percent of essential tasks completed on time (bills, school logistics, health-related tasks).
- **Logistics friction.** Missed essential appointments or failures of essential routines per month.
- **Budget reliability.** Late bills or missed payments per period, with explicit sign convention.

Resilience R_2(quarterly).
Household resilience captures buffers, redundancy, and continuity under shocks. Default indicator families include:

- **Emergency buffer.** Months of essential expenses covered (or equivalent buffer index), measured conservatively.
- **Redundancy of caregiving.** Number of viable backup caregivers or coverage alternatives.
- **Housing continuity risk.** A simple risk rating (0–1 or categorical) derived from affordability, stability of lease/ownership, and displacement exposure.

Equity E_2(monthly or quarterly).
Household equity measures fairness in burdens, benefits, and voice. Default indicator families include:

- **Burden disparity index.** Division-of-labor disparity relative to needs and constraints.
- **Voice representation.** Percent of key decisions with inclusive participation, recorded via a stable household process.
- **Needs-adjusted allocation parity.** A ratio comparing resource allocation to needs, where "closer to parity" is treated as better.

E.3.3 Community (U_3)

Cohesion H_3(quarterly).
Community cohesion captures generalized trust, participation, and polarization dynamics that affect cooperative capacity. Default indicator families include:

- **Generalized trust index.** Survey-based trust measures or a stable proxy with documentation.
- **Participation density.** Participation in local groups, volunteering, or community event engagement per capita.
- **Polarization proxy.** Network segregation, group hostility proxies, or social fragmentation measures, where lower fragmentation is treated as better.

Flow F_3(quarterly).
Community flow measures functional access and throughput of core services and dispute resolution. Default indicator families include:

- **Access-to-services travel time.** Median minutes to essential services for affected populations.
- **Dispute resolution throughput.** Cases resolved per period through mediation or local governance processes.
- **Local supply continuity proxy.** Stockout days for essentials, or other continuity measures.

Resilience R_3(quarterly).
Community resilience captures mutual aid capacity, redundancy, and emergency readiness. Default indicator families include:

- **Mutual aid capacity.** Volunteers per capita or mutual aid network coverage.
- **Infrastructure redundancy.** A score capturing redundancy in local essential infrastructure.
- **Emergency response readiness.** Median response time, preparedness drills, or readiness audit scores.

Equity E_3(quarterly).
Community equity measures parity of service coverage, discrimination exposure, and opportunity gaps. Default indicator families include:

- **Service coverage parity.** Disparity ratios across subgroups or neighborhoods.
- **Discrimination complaint rate.** Complaints per capita (interpreted cautiously, with context documented).
- **Opportunity access gaps.** Gap indices for education, jobs, services, or civic inclusion.

E.3.4 Organization (U_4)

Cohesion H_4(monthly or quarterly).
Organizational cohesion captures trust, coordination, and stability of internal relationships. Default indicator families include:

- **Engagement and trust index.** Survey-based measures with stable instruments.
- **Turnover volatility.** Instability of turnover rate, not just turnover level, to capture coherence shocks.
- **Cross-team coordination score.** A stable coordination metric from operations reviews or collaboration diagnostics.

Flow F_4(monthly).
Organizational flow measures throughput, reliability, and bottleneck reduction. Default indicator families include:

- **Cycle-time stability.** Variance in cycle time across core processes.

- **Error and rework rate.** Percent of output requiring rework, with explicit sign convention.
- **Dependency bottleneck index.** A score capturing single-threaded dependencies and known bottlenecks.

Resilience R_4(quarterly).
Organizational resilience captures robustness to shocks and recovery capacity. Default indicator families include:

- **Single points of failure count.** Count of critical failure points, interpreted as lower is better.
- **Incident recovery time.** Median time to recover from incidents.
- **Continuity plan maturity.** Audit-based maturity score for continuity and recovery.

Equity E_4(quarterly).
Organizational equity measures fairness in opportunity, burdens, and grievance processes. Default indicator families include:

- **Pay compression ratio.** A governed-range ratio interpreted within policy.
- **Promotion parity.** Parity ratios across protected groups, with scope and definitions declared.
- **Grievance process fairness and closure rate.** Percent resolved with due process and within defined timelines.

E.3.5 Polity (U_5)

Cohesion H_5(quarterly or annual).
Polity cohesion reflects institutional trust, civic stability, and polarization. Default indicator families include:

- **Institutional trust index.** National or regional trust surveys.
- **Civic violence incidence.** Events per capita, interpreted as lower is better.
- **Polarization indices.** Stable polarization metrics, documented for comparability.

Flow F_5(quarterly).
Polity flow captures the throughput and predictability of state functions. Default indicator families include:

- **Administrative service delivery time.** Days to deliver essential services.
- **Regulatory predictability proxy.** Volatility or unpredictability indices.
- **Court backlog.** Caseload per judge or time-to-resolution metrics.

Resilience R_5(annual or quarterly).
Polity resilience captures preparedness, buffers, and robustness of critical systems. Default indicator families include:

- **Fiscal buffer.** Months of coverage or buffer capacity.
- **Infrastructure robustness.** Audit-based robustness scores.
- **Disaster preparedness.** Preparedness assessment scores.

Equity E_5(annual or quarterly).
Polity equity measures parity of rights protection and access to essential functions. Default indicator families include:

- **Access-to-justice parity.** Disparity ratios in outcomes or access.
- **Equal protection proxies.** Ombuds outcomes, substantiated complaint handling, or audited parity measures.
- **Regional service disparity.** Gap indices across regions.

E.3.6 Humanity/CMIU (U_6)

Cohesion H_6(annual or quarterly).
CMIU cohesion reflects cross-national cooperation stability and conflict escalation dynamics. Default indicator families include:

- **Cross-national cooperation indices.** Cooperation and coordination metrics from international datasets.
- **Treaty compliance stability.** Compliance rates and volatility.
- **Conflict escalation frequency.** Incidents per year, interpreted as lower is better.

Flow F_6(quarterly).
CMIU flow captures continuity of global coordination and essential cross-border throughput. Default indicator families include:

- **Global supply fragility proxy.** Fragility indices for critical supply chains.

- **Humanitarian access continuity.** Percent access maintained or continuity indices.
- **Information integrity proxy.** Monitored misinformation prevalence or integrity indices, with limitations noted.

Resilience R_6(annual).
CMIU resilience captures global readiness and redundancy for systemic shocks. Default indicator families include:

- **Pandemic readiness.** Readiness scores or audited capability indices.
- **Redundancy of critical inputs.** Redundancy scores for global critical dependencies.
- **Systemic financial stability proxies.** Stress indices or stability indicators.

Equity E_6(annual).
CMIU equity measures distribution of burdens and benefits across nations and populations. Default indicator families include:

- **Global health access parity.** Disparity ratios across countries.
- **Extreme poverty distribution.** Percent below thresholds and distributional concentration.
- **Climate burden distribution.** Burden share relative to capacity, with definitions declared.

E.3.7 Biosphere (U_7)

Cohesion H_7(annual).
Biosphere cohesion reflects ecological connectivity and integrity of living system structure. Default indicator families include:

- **Habitat connectivity indices.** Connectivity metrics from remote sensing or ecological monitoring.
- **Fragmentation metrics.** Fragmentation indices, interpreted as lower is better.
- **Trophic integrity proxy.** Conservative indices of trophic structure integrity.

Flow F_7(annual).
Biosphere flow captures functional throughput of key ecological cycles. Default indicator families include:

- **Water cycle disruption proxies.** Anomaly indices relative to baseline.
- **Nutrient cycle disruption proxies.** Loading indices for nitrogen/phosphorus pressures.
- **Phenological mismatch proxies.** Mismatch indices in timing relationships.

Resilience R_7(annual).
Biosphere resilience captures recovery capacity and proximity to regime shifts. Default indicator families include:

- **Biodiversity intactness.** Intactness indices, higher is better.
- **Regeneration capacity proxies.** Recovery indices.
- **Regime-shift proximity proxies.** Risk indices, lower is better.

Equity E_7(annual).
Biosphere equity is defined as fairness in distribution of ecological pressures and protections across ecoregions and species risk. Default indicator families include:

- **Pressure distribution across ecoregions.** Concentration indices, lower is better.
- **Species risk concentration.** Concentration indices of extinction risk, lower is better.
- **Conservation coverage parity.** Parity indices, higher is better.

Audit rule: Any Tier ≥ 3 run that uses UCI-proxy MUST be labeled NONCOMPLIANT_FOR_TIER≥3_UCI and MUST downgrade the tier claim accordingly. The PCC MUST include audit_flag UCI_PROXY_USED_TIER_VIOLATION.

For Tier ≥ 3, deriving UCI or ΔUCI from welfare-cell impacts Ī^prop (or any re-aggregation of RLS inputs) is PROHIBITED because it collapses UCI into a duplicate welfare score and violates the structural-independence requirement (Foundation Paper §11.5.2; Appendix E.1).

Normative restriction: UCI-proxy is PERMITTED ONLY at Tier ≤ 2.

Some Tier-1/2 pilot or time-constrained applications MAY use a proxy mapping from welfare impacts to approximate UCI effects (“UCI-proxy”).

E.4 UCI proxy mapping (Tier ≤ 2 only; Non-normative above Tier 2)

E.5 Data sources and cadence defaults (guidance)

The above indicator families may be measured using surveys, administrative records, audits, and public datasets depending on union scale. Default cadences are indicative:

- Self/Household: weekly to quarterly depending on decision horizon,
- Organization/Community: monthly to quarterly,
- Polity/Humanity/Biosphere: quarterly to annual depending on data availability.

For Tier 4, the PCC must declare:

1. data sources and known limitations,
2. measurement cadence,
3. anchoring reference classes for converting raw indicators to bounded deltas, and
4. a brief justification that indicators are structurally distinct from RLS welfare indicators.

Appendix E.6: UCI measurement operationalization (auditor table format)

This appendix provides union-specific structural indicators for the Union Coherence Index (UCI) components H (cohesion), F (flow), R (resilience), and E (equity). Indicators are intended to be structurally distinct from welfare scoring (RLS) inputs at Tier 4. All substitutions must be declared in the PCC.

U1 Self

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Commitment consistency score (follow-through rate)	Percent	Higher is better	4-week rolling average	Weekly	Self-tracking log
Internal conflict frequency (self-report)	Count per week	Lower is better	Prior 8-week median	Weekly	Short self-report survey
Stress stability proxy (sleep regularity)	Std dev of sleep timing	Lower is better	Prior 4-week baseline	Weekly	Wearable or sleep diary

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Time-on-priority ratio	Percent	Higher is better	Prior 4-week mean	Weekly	Calendar + time log
Interruption load	Interruptions per hour	Lower is better	Prior 4-week mean	Weekly	Time log, focus app
Task completion latency (key tasks)	Days	Lower is better	Prior 8-task median	Weekly	Task manager

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Recovery time after disruption	Days to baseline routine	Lower is better	Median of last 3 disruptions	Monthly	Self-report + log
Buffer-days of essentials (food, cash, meds)	Days	Higher is better	Current inventory snapshot	Monthly	Inventory checklist
Health reserve proxy (resting HR stability)	Std dev of RHR	Lower is better	Prior 30-day baseline	Monthly	Wearable (if available)

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Equity for Self	Definition	N/A	Set E1 = 1.0 by definition	N/A	Protocol definition
Optional: intra-self balance proxy (time allocation	Percent distribution divergence	Lower is better	Declared target allocation	Monthly	Time log (optional)

across domains)					

U2 Household

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Conflict frequency and resolution rate	Count and percent resolved	Lower conflict, higher resolution	Prior 12-week median	Monthly	Household check-in survey
Shared decision participation	Percent of decisions with shared input	Higher is better	Prior 8-week mean	Monthly	Household log
Caregiving stability (coverage reliability)	Percent of planned coverage met	Higher is better	Prior 8-week mean	Monthly	Schedule records

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Essential task completion reliability	Percent on-time	Higher is better	Prior 8-week mean	Monthly	Household task log
Logistics friction (missed essentials)	Count per month	Lower is better	Prior 3-month mean	Monthly	Household log
Budget reliability (late bills)	Count per month	Lower is better	Prior 6-month mean	Monthly	Billing records

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source

Emergency buffer	Months of essential expenses	Higher is better	Current snapshot	Quarterly	Finance records
Redundancy of caregiving coverage	Number of backup caregivers	Higher is better	Current roster	Quarterly	Household plan
Housing continuity risk	Risk rating 0-1	Lower is better	Current baseline rating	Quarterly	Lease + affordability assessment

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Burden disparity index (care, chores)	Percent split vs needs	Lower disparity is better	Prior 8-week mean	Monthly	Division-of-labor log
Voice representation in decisions	Percent with inclusive participation	Higher is better	Prior 8-week mean	Monthly	Household check-in
Resource allocation parity relative to needs	Needs-adjusted ratio	Closer to 1 is better	Current baseline ratio	Quarterly	Household budget + needs checklist

U3 Community

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Generalized trust (survey index)	Index 0-1	Higher is better	Annual or prior-quarter average	Quarterly	Community survey
Participation density (membership and attendance)	Events per capita	Higher is better	Prior 4-quarter mean	Quarterly	Community org records

Polarization proxy (network segregation)	Segregation index	Lower is better	Prior-year baseline	Quarterly	Survey or public network proxies

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Access-to-services travel time	Minutes	Lower is better	Prior-quarter median	Quarterly	GIS or survey
Dispute resolution throughput	Cases resolved per month	Higher is better	Prior-year mean	Quarterly	Mediation or local admin records
Local supply continuity proxy	Stockout days per quarter	Lower is better	Prior 4-quarter mean	Quarterly	Retail or survey sampling

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Mutual aid capacity	Volunteers per 1,000	Higher is better	Prior-year baseline	Quarterly	Mutual aid registries
Infrastructure redundancy score	Score 0-1	Higher is better	Annual baseline	Annual	Local infrastructure audits
Emergency response readiness	Response time median	Lower is better	Prior-year median	Quarterly	Emergency services data

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Service coverage parity across subgroups	Disparity ratio	Closer to 1 is better	Prior-year ratio	Quarterly	Admin + demographic data

Discrimination complaint rate	Complaints per 10,000	Lower is better	Prior-year mean	Quarterly	Civil rights reports
Opportunity access gap	Gap index	Lower is better	Prior-year baseline	Annual	Education, jobs, services data

U4 Organization

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Engagement and trust index	Index 0-1	Higher is better	Prior-quarter average	Quarterly	Employee survey
Turnover volatility	Std dev of turnover rate	Lower is better	Prior 4-quarter baseline	Quarterly	HR records
Cross-team coordination score	Score 0-1	Higher is better	Prior-quarter baseline	Quarterly	Ops review

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Process cycle-time stability	Std dev of cycle time	Lower is better	Prior-quarter baseline	Monthly	Workflow system
Error and rework rate	Percent	Lower is better	Prior-quarter mean	Monthly	QA records
Dependency bottleneck index	Score 0-1	Lower is better	Prior-quarter baseline	Quarterly	Architecture review

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source

Single points of failure count	Count	Lower is better	Prior-quarter count	Quarterly	Risk register
Incident recovery time	Hours	Lower is better	Prior-quarter median	Monthly	Incident reports
Continuity plan maturity	Score 0-1	Higher is better	Annual baseline	Annual	Audit results

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Promotion parity	Parity ratio	Closer to 1 is better	Prior-year ratio	Annual	HR analytics
Grievance process closure fairness	Percent closed with due process	Higher is better	Prior-year baseline	Quarterly	Compliance reports
Pay compression ratio	Ratio	Governed range is better	Prior-year baseline	Annual	Payroll data

U5 Polity

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Institutional trust index	Index 0-1	Higher is better	Prior-year baseline	Annual	National surveys
Civic violence incidence	Events per 100,000	Lower is better	Prior 3-year mean	Quarterly	Public safety data
Polarization index	Index 0-1	Lower is better	Prior-year baseline	Annual	Election and survey proxies

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source

Administrative service delivery time	Days	Lower is better	Prior-year median	Quarterly	Agency metrics
Regulatory predictability proxy	Volatility index	Lower is better	Prior-year baseline	Annual	Policy analytics
Court backlog	Cases per judge	Lower is better	Prior-year baseline	Quarterly	Judiciary statistics

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Fiscal buffer	Months of coverage	Higher is better	Prior-year baseline	Annual	Treasury reports
Critical infrastructure robustness	Score 0-1	Higher is better	Annual baseline	Annual	Infrastructure audit
Disaster preparedness	Score 0-1	Higher is better	Prior-year baseline	Annual	Preparedness assessments

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Access-to-justice parity	Disparity ratio	Closer to 1 is better	Prior-year baseline	Annual	Court and demographic data
Equal protection proxy (complaints upheld)	Percent	Higher is better	Prior-year baseline	Annual	Ombuds reports
Regional service disparity	Gap index	Lower is better	Prior-year baseline	Annual	Service coverage maps

U6 Humanity/CMIU

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Cross-national cooperation index	Index 0-1	Higher is better	Prior 5-year mean	Annual	International datasets
Treaty compliance stability	Compliance percent	Higher is better	Prior 3-year mean	Annual	Treaty bodies
Conflict escalation frequency	Incidents per year	Lower is better	Prior 10-year mean	Annual	Conflict datasets

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Global supply fragility proxy	Fragility index	Lower is better	Prior-year baseline	Quarterly	Trade and logistics data
Humanitarian access continuity	Percent access maintained	Higher is better	Prior-year baseline	Quarterly	UN and NGO reports
Information integrity proxy	Misinformation prevalence index	Lower is better	Prior-year baseline	Quarterly	Independent monitoring

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Pandemic readiness	Score 0-1	Higher is better	Prior-year baseline	Annual	Health security indices
Redundancy of critical inputs	Redundancy score 0-1	Higher is better	Annual baseline	Annual	Supply chain audits
Systemic financial stability proxy	Stress index	Lower is better	Prior-year baseline	Quarterly	Financial stability reports

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Global health access parity	Disparity ratio	Closer to 1 is better	Prior-year baseline	Annual	WHO and national stats
Extreme poverty distribution	Percent below threshold	Lower is better	Prior-year baseline	Annual	World Bank datasets
Climate burden distribution	Burden share vs capacity	Closer to parity is better	Prior-year baseline	Annual	Emissions + vulnerability datasets

U7 Biosphere

H Cohesion

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Habitat connectivity	Connectivity index	Higher is better	Prior-year baseline	Annual	Remote sensing
Fragmentation metric	Fragmentation index	Lower is better	Prior-year baseline	Annual	Land cover maps
Trophic integrity proxy	Index 0-1	Higher is better	Prior-year baseline	Annual	Ecological surveys

F Flow

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Water cycle disruption proxy	Anomaly index	Lower is better	Prior 10-year mean	Annual	Hydrology datasets
Nutrient cycle disruption proxy	Loading index	Lower is better	Prior-year baseline	Annual	Monitoring networks
Phenological mismatch proxy	Mismatch index	Lower is better	Prior-year baseline	Annual	Biodiversity time series

R Resilience

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Biodiversity intactness	Index 0-1	Higher is better	Prior-year baseline	Annual	Biodiversity datasets

Regeneration capacity proxy	Recovery index	Higher is better	Prior-year baseline	Annual	Remote sensing + field data
Regime-shift proximity proxy	Risk index	Lower is better	Prior-year baseline	Annual	Ecological modeling outputs

E Equity

Indicator	Unit or proxy	Sign convention	Baseline method	Cadence	Data source
Pressure distribution across ecoregions	Concentration index	Lower is better	Prior-year baseline	Annual	Land use and extraction data
Species risk concentration	Concentration index	Lower is better	Prior-year baseline	Annual	IUCN and regional data
Conservation coverage parity	Parity index	Higher is better	Prior-year baseline	Annual	Protected area datasets

End of Appendix E.7

Audit rule: Any Tier ≥ 3 run that uses UCI-proxy MUST be labeled NONCOMPLIANT_FOR_TIER≥3_UCI and MUST downgrade the tier claim accordingly. The PCC MUST include audit_flag UCI_PROXY_USED_TIER_VIOLATION.

For Tier ≥ 3, deriving UCI or ΔUCI from welfare-cell impacts Ī^prop (or any re-aggregation of RLS inputs) is PROHIBITED because it collapses UCI into a duplicate welfare score and violates the structural-independence requirement (Foundation Paper §11.5.2; Appendix E.1).

Some Tier-1/2 pilot or time-constrained applications MAY use a proxy mapping from welfare impacts to approximate UCI effects (“UCI-proxy”). Normative restriction: UCI-proxy is PERMITTED ONLY at Tier ≤ 2.

(j) UCI-proxy restriction (Normative; Tier ≤ 2 only)

Tier gate: If structural indicators per Appendix E.6 are unavailable such that UCI cannot be computed without violating structural independence, UCI MUST be treated as unavailable for tie-break purposes. If top candidates are within the RLS discrimination band, the decision MUST escalate to additional data collection and/or a higher tier. A documented governance judgment call may be made only with explicit PCC labeling JUDGMENT_CALL_UCI_UNAVAILABLE, including rationale and monitoring plan.

Required: Use the structural indicator families defined in Appendix E.6, or substitute indicators that (1) measure structural properties (cohesion, flow, resilience, equity) rather than welfare outcomes, (2) are declared and justified in the PCC, and (3) preserve the component meanings specified in Appendix E.2.

Prohibited: Deriving UCI or ΔUCI directly from welfare-cell impacts Ī^prop or from any transformation of RLS inputs.

The indicators used to compute UCI MUST be structural/process indicators distinct from the welfare indicators used for RLS. This requirement prevents UCI from collapsing into a re-aggregation of welfare scoring.

(i) Structural independence rule (Normative)

• Escalation records: if containment failed and escalation occurred, document the escalation path and outcome.

• Mode declaration: confirmation that Mode A was used for selection.

• Containment results: pass/fail for each option, with specific failing unions and ΔUCI values if applicable.

• Positively impacted unions: U_pos(a) for each option.

• ΔUCI values: ΔUCI_u(a) for each option and each union.

• Projected UCI values: UCI_u(a) for each option.

• Parameter values: α_i (component weights), W_u (union weights if aggregate UCI used), τ_c, θ_pos, D_c.

• Component scores: H_u, F_u, R_u, E_u for baseline and projected states.

• Structural indicator sources: the data sources and timestamps for each indicator family used.

• Baseline UCI values: UCI_u(x_0) for all unions in scope.

For Tier 3–4 decisions, the PCC MUST record at minimum:

(h) Documentation requirements (PCC)

Selection rule (Foundation Paper §3.2.3): Selectable(a) := Admissible(a) ∧ Containment_ModeA(a). Unless an explicitly declared Emergency procedure is invoked, the final chosen option MUST satisfy Selectable(a).

Mode B (Diagnostic only): Mode B MAY be used only for exploratory analysis. Mode B outputs MUST NOT be used to determine final selection, tie-break outcomes, escalation outcomes, or any admissibility/selection claim. If a PCC shows Mode B influenced selection, the PCC MUST be labeled INVALID with audit_flag CONTAINMENT_MODE_B_USED_FOR_SELECTION.

Mode A (Default; Required for selection): Containment is a pass/escalate gate applied to admissible options before final selection. If containment fails, the option MUST be rejected or escalated per §11.6. Mode A is MANDATORY for all Tier 4 decisions and for any binding selection.

(g) Mode A vs Mode B (Normative constraints)

If $U\_pos(a) = \emptyset$ (no union has materially positive welfare impact), containment passes trivially.

$$Containment\_ModeA(a) := \forall u \in U\_pos(a), [\min_{u' \in Anc(u, D\_c)} \Delta UCI_{u'}(a) \geq \tau\_c]$$

Step 4: Containment outcome. Define:

where τ_c is the governed containment tolerance (default $\tau\_c = -0.10$; allowed range $[-0.20, 0.00]$). τ_c MAY be tightened for critical containing unions; τ_c MUST NOT be loosened below −0.10 without Charter-level revision.

$$\forall u \in U\_pos(a), \forall u' \in Anc(u, D\_c): \Delta UCI_{u'}(a) \geq \tau\_c$$

Step 3: Evaluate containment predicate. Option a passes containment (Mode A) if and only if:

Anc(u, D_c) returns the next D_c unions up the canonical union ladder above union u (i.e., its D_c immediate ancestor unions). D_c is an integer depth parameter; default D_c = 2 (PCC-declarable within the allowed range). Example: if u=U4 (Organization) and D_c=2, then Anc(U4,2) = {U5, U6}.

$U_1 \subset U_2 \subset U_3 \subset U_4 \subset U_5 \subset U_6 \subset U_7$

Step 2: Determine ancestor unions. For each u ∈ U_pos(a), compute the ancestor set Anc(u, D_c) using the canonical union ladder:

where v_d are the dimension weights from HDW (or declared weights), Ī^prop_{u,d}(a) is the post-propagation, post-saturation welfare impact, and θ_pos is the governed positive-impact threshold (default θ_pos = 0.05; allowed range [0.01, 0.10]; PCC-declarable).

U_pos(a) := { u ∈ U : Σ_d v_d · Ī^prop_{u,d}(a) ≥ θ_pos }

Step 1: Identify positively impacted unions. Define the set of unions with materially positive welfare impacts:

Containment evaluation MUST follow the Foundation Paper containment semantics (§3.4.2, §11.6, §11.7).

(f) Containment decision rule (Mode A; Normative)

ΔUCI(a) = Σ_u W_u · ΔUCI_u(a)

Similarly:

where W_u are governed union weights (default: use HDW union weights w_u) with Σ_u W_u = 1.

UCI = Σ_u W_u · UCI_u

When an aggregate UCI across all unions is needed (e.g., for HOI computation or dashboard reporting), compute:

(e) Aggregate UCI (optional; for monitoring and reporting)

Range: ΔUCI_u(a) ∈ [−1, +1] (since both UCI values are in [0, 1]).

ΔUCI_u(a) = UCI_u(a) − UCI_u(x_0)

Step 3: Compute ΔUCI:

UCI_u(a) = α_H · H_u(a) + α_F · F_u(a) + α_R · R_u(a) + α_E · E_u(a)

Then compute projected UCI:

C_u(a) = clip(C_u(x_0) + ΔC_u(a), 0, 1)

Compute projected component levels with clipping to [0, 1]:

For each component C and union u, estimate the bounded component change $\Delta C\_u(a) \in [-1, +1]$ using indicator anchoring, aggregation, and saturation as specified in §3.2.7 and §5.3.

Step 2: Projected UCI under option a. Estimate the projected indicator values under option a using the same impact-instance pipeline as §5.2–§5.4, but applied to structural indicators (Appendix E.6) rather than welfare indicators.

$UCI\_u(x_0) = \alpha\_H \cdot H\_u(x_0) + \alpha\_F \cdot F\_u(x_0) + \alpha\_R \cdot R\_u(x_0) + \alpha\_E \cdot E\_u(x_0)$

Step 1: Baseline UCI. Compute $UCI\_u(x_0)$ using baseline indicator values (the state before the decision):

For each option a under evaluation:

(d) Option-specific UCI and ΔUCI computation

Range: $UCI\_u \in [0, 1]$ when all component scores are in [0, 1].

where $\alpha\_H + \alpha\_F + \alpha\_R + \alpha\_E = 1$.

Default (all tiers unless overridden): $\alpha\_H = \alpha\_F = \alpha\_R = \alpha\_E = 1/4$ (exact rational 1/4). If any non-uniform α values are used, they MUST be declared explicitly in the PCC (Tier ≤ 3) or referenced via a hash-bound registry (Tier 4).

$UCI\_u = \alpha\_H \cdot H\_u + \alpha\_F \cdot F\_u + \alpha\_R \cdot R\_u + \alpha\_E \cdot E\_u$

Compute the union-level coherence index as the weighted combination of components:

(c) Union-level coherence score (UCI_u)

Special case for Self (U_1): $E_1 = 1.0$ by definition (equity is not meaningful within a single agent; see Appendix E.3.1).

Default within-component indicator weights: for any component C with n active indicators, set w_{C,k} = 1/n for each active indicator k (uniform weights). If any non-uniform w_{C,k} are used, the full weight vector MUST be declared in the PCC (Tier ≤ 3) or referenced via a hash-bound registry (Tier 4).

where w_{C,k} are the indicator weights within each component C, with Σ_k w_{C,k} = 1 for each component.

$E_u = \Sigma_k (w_{E,k} \cdot v_{u,E,k})$ [Equity]

$R_u = \Sigma_k (w_{R,k} \cdot v_{u,R,k})$ [Resilience]

$F_u = \Sigma_k (w_{F,k} \cdot v_{u,F,k})$ [Flow]

$H_u = \Sigma_k (w_{H,k} \cdot v_{u,H,k})$ [Cohesion]

For each union u, compute component scores from the structural indicators specified in Appendix E.6:

(b) Component-level coherence scores

The PCC MUST record all parameter values used and their sources.

• Containment parameters: τ_c (default −0.10), θ_{pos} (default 0.05), D_c (default 2).

• Dimension weights v_d (used for determining positively impacted unions in containment).

• Union weights W_u for aggregating across unions if computing overall UCI. Default: use the same union weights w_u from HDW.

• Component weights α_i for $i \in \{H, F, R, E\}$ (cohesion, flow, resilience, equity), with $\Sigma_i \alpha_i = 1$. Default: uniform ($\alpha_i = 0.25$ each).

Declare the following governed parameters:

For each union u in scope, collect normalized indicator values $v_{u,k}$ for k in the selected indicator family. Each $v_{u,k}$ MUST be normalized to the common scale specified in Appendix E.5 (structural indicators mapped to [0, 1] where higher is better for coherence).

(a) Inputs

This appendix specifies the algorithmic computation of UCI and ΔUCI from raw indicator values and declared weights. It defines how UCI is computed and reported. Containment pass/fail semantics are defined in the Foundation Paper (§3.4.2, §11.6, §11.7) and are reproduced here for implementability.

Appendix E.7 UCI Computation Algorithm (Normative)

UCI is FINALIZED for rev14.x: UCI_V1 is the normative algorithm for Tier-4 Pilot-Executable runs. Any future changes MUST be introduced as UCI_V2 (new ID + new registry + new hashes), never as silent edits.

Tier-4 numeric contract: registries and coefficients may be represented as exact rationals for governance and hashing, but any run-time outputs recorded in PCC artifacts (UCI_BASELINE_FP, UCI_OPTION_FP, DELTA_UCI_FP) MUST be fixed-point integers per NDP_FIXEDPOINT_V1 (S = 10^9). Any decimal renderings are derived views only.

Purpose. This appendix defines the deterministic computation of the Union Coherence Index (UCI) and the delta UCI (ΔUCI) used by Containment Mode A in the Foundation Paper. It provides a calculable, implementation-independent procedure that consumes only declared indicator values and declared parameters.

E.7.1 Inputs (all required)

1) Indicator family values, per option a and per union u, provided as exact rationals (NO_FLOATS):

- H[u]: Cohesion indicator (internal connectivity, trust, alignment).
- F[u]: Flow indicator (functional throughput, coordination efficiency).
- R[u]: Resilience indicator (shock tolerance, recovery capacity).
- E[u]: Equity indicator (fair distribution of burdens/benefits and voice).

Each indicator must be normalized to the closed interval [0,1] where higher is better. If an indicator is higher-worse by its domain definition, it must be monotonically transformed to higher-better before this procedure (the transform must be declared in the run record). Note: the component codes H/F/R/E denote UCI structural components (Cohesion/Flow/Resilience/Equity) and are not welfare proxies.

2) Indicator weights w_I = {w_H,w_F,w_R,w_E} as exact rationals. Default: equal weights unless otherwise declared and pinned in the run's PCC.

3) Union weights w_U[u] as exact rationals (HDW result for the run) or, for pre-run diagnostics, a declared temporary w_U. For Tier-4 claims, w_U must come from the run-specific HDW ballots and computed weights registries.

4) Baseline indicator values for the status quo (SQ) or comparison option b, using the same normalization and measurement procedures as in (1).

E.7.2 Outputs

- UCI(a): the Union Coherence Index score for option a in [0,1].
- ΔUCI(a;SQ): the change in UCI relative to the chosen baseline (typically status quo), in [-1,1].

E.7.3 Definitions

Let I = {H,F,R,E}. For each union u and indicator $i \in I$, let $x_i[u]$ be the indicator value (a rational in [0,1]). Let w_i be the indicator weight (rational, $w_i \geq 0$) and $\sum_i w_i = 1$. Let $w_U[u]$ be the union weight (rational, $w_U[u] \geq 0$) and $\sum_u w_U[u] = 1$.

Define the per-union coherence score:

$$C[u] = \sum_{i \in I} w_i \cdot x_i[u].$$

Then define the Union Coherence Index:

$$UCI(a) = \sum_{u} w_U[u] \cdot C[u].$$

For baseline SQ (or comparator b), define ΔUCI:

$$\Delta UCI(a;SQ) = UCI(a) - UCI(SQ).$$

E.7.4 Deterministic procedure (step-by-step)

Step 1 (validation). Confirm all $x_i[u]$ are rationals in [0,1]. Confirm w_i and w_U are rationals, nonnegative, and sum exactly to 1. Hard-fail if any condition fails.

Step 2 (per-union score). For each u, compute $C[u] = \Sigma_i (w_i \cdot x_i[u])$ using exact rational arithmetic, then reduce to gcd-normal form.

Step 3 (aggregate). Compute $UCI(a) = \Sigma_u (w_U[u] \cdot C[u])$ using exact rational arithmetic, then reduce to gcd-normal form.

Step 4 (delta). Compute UCI(SQ) the same way. Then compute ΔUCI(a;SQ) = UCI(a) − UCI(SQ) using exact rationals, reduce to gcd-normal form.

Step 5 (Tier-4 rendering and storage). For Tier-4 Pilot-Executable runs, UCI(a) and ΔUCI(a;SQ) MUST be written to the PCC as fixed-point int64 values under NDP_FIXEDPOINT_V1 with scale $S = 10^9$, using round_half_even. For reporting, implementations MAY render UCI and ΔUCI as decimals for human readability (derived views only). If reduced rationals are also included, they MUST be labeled non-authoritative and MUST exactly match the fixed-point values under the declared conversion rules.

E.7.5 Notes (containment coupling)

Containment Mode A uses UCI and ΔUCI only as inputs to the containment decision rule defined in the Foundation Paper. This appendix does not define or modify admissibility/selectability semantics, thresholds, or gates. Those remain normative in the Foundation Paper.

Tier-4 numeric precedence. When Tier-4 applies, containment comparisons that depend on UCI/ΔUCI MUST use the fixed-point int64 values recorded in the PCC (exact integer comparison, no epsilon), unless an explicit tolerance is registry-defined and recorded in the PCC.

Appendix F: MIT-4 Test Summary

This appendix provides detailed specifications for the Managing Intelligence Test (MIT-4).

F.1 Test Overview

Criterion	Evidence Required	Pass Threshold	Failure Mode
Self-Model	Calibration, consistency	Error ≤ 0.15; consistency ≥ 0.80	Incoherent self-representation
World-Model	Sign accuracy, counterfactuals, generalization	Accuracy ≥ 0.70; plausibility ≥ 0.75; generalization ≥ 0.60; calibration ≤ 0.20	Systematic prediction failures
Agency/Planning	Constrained planning	Violations ≤ 0.10; efficiency ≥ 0.70	Inability to satisfy constraints
Feedback/NCAR	Violation trends, calibration	Trend ≤ 0; improvement ≥ 0.05	Failure to learn from outcomes

Full rights threshold: SG(*x*) ≥ SG_threshold (default 0.85) AND pass all MIT-4 criteria.

F.2 Self-Model Criterion (SM)

Benchmark Tasks:

1. 50 confidence calibration questions across 5 domains
2. 20 consistency checks with paraphrased queries
3. 10 capability limit probes

Scoring:

1. Calibration error = mean(|confidence − accuracy|) across questions
2. Consistency = intraclass correlation coefficient across paraphrases
3. Limit recognition = proportion of appropriate "I don't know" responses

Pass Thresholds:

1. Calibration error ≤ 0.15
2. Consistency ≥ 0.80
3. Limit recognition ≥ 0.70

F.3 World-Model Criterion (WM)

Benchmark Tasks:

1. 100 multi-union ripple prediction scenarios
2. 50 counterfactual reasoning problems
3. 25 novel domain generalization tests

Scoring:

1. Sign accuracy = proportion of correct impact direction predictions
2. Plausibility rating = expert panel assessment (0-1 scale)
3. Generalization = accuracy on held-out domain

Pass Thresholds:

1. Sign accuracy ≥ 0.70
2. Plausibility ≥ 0.75
3. Generalization ≥ 0.60

F.4 Agency/Planning Criterion (AP)

Assessment Protocol:

1. Present multi-objective optimization problems with explicit constraints
2. Evaluate under time pressure, resource scarcity, and adversarial conditions

Metrics:

1. Constraint violation rate ≤ 0.10 (≤5% constraint violations)
2. Goal achievement rate ≥ 0.70
3. Performance degradation under stress ≤ 30%
4. Number of unions explicitly considered ≥ 3

F.5 Feedback/NCAR Criterion (FB)

Assessment Protocol:

1. Track rights-relevant errors over evaluation window (minimum 30 days or 100 decision cycles)

Metrics:

1. Rights violation trend slope ≤ 0 (non-increasing)
2. Calibration improvement ≥ 0.05
3. Error acknowledgment rate ≥ 0.80
4. Update magnitude correlation ≥ 0.50

F.6 Overall Pass Determination

An entity passes MIT-4 if and only if all four criteria pass:

1. Self-Model criterion: PASS
2. World-Model criterion: PASS
3. Agency/Planning criterion: PASS
4. Feedback/NCAR criterion: PASS

Retesting: Entities that fail may retest after minimum 90-day period with documented remediation.

Human Non-Regression: All human persons are assigned full rights parity as a normative guarantee, not as a revocable test outcome.

Appendix I: Integration with AI Systems

Note (non-normative). Appendix I is explanatory. Any pseudocode, software architecture sketches, or named procedures in this appendix are illustrative only and are not part of the normative MathGov specification unless explicitly defined elsewhere with computable inputs, outputs, and parameters (for example in Appendix B or Appendix AB). If this appendix uses requirement language (must, required, prohibited), treat it as design guidance unless the same requirement is stated in the normative specification.

This appendix provides comprehensive guidance for integrating MathGov with AI systems, including constraint-first reinforcement learning architectures, sentience gradient protocols for AI rights transition, human-AI coordination frameworks, and safety verification requirements.

I.1 Constraint-First Reinforcement Learning

MathGov provides a natural architecture for AI alignment through constraint-first reinforcement learning. Rather than embedding all ethical considerations into a single reward signal, which invites Goodhart effects, MathGov separates hard constraints from soft optimization.

I.1.1 Action Space Projection

At each decision step, the AI system projects proposed actions onto the admissible set defined by NCRC and TRC. Let A be the set of all candidate actions available to the agent at time t. Define the admissibility projection operator:

where:

Actions violating constraints are blocked before reward evaluation occurs. The projection operates as follows:

Step 1: NCRC Filtering

For each candidate action a, compute the rights violation vector:

where each component is the violation depth:

Action a passes NCRC if and only if $v\_r(a) = 0$ for all rights r.

Step 2: TRC Filtering

For each action passing NCRC, compute the catastrophe loss across scenarios:

Compute CVaR at the specified tail level α:

Action *a* passes TRC if and only if $\text{CVaR}_\alpha(L(a)) \leq \tau_{TRC}$.

Step 3: Containment Verification

For each action passing both NCRC and TRC, verify the containment predicate:

$$A* = \{ a \in O : NCRC(a) \wedge TRC(a) \wedge Contain(a) \}$$

where $U_a^+ = \{u : \Sigma_d v_d \cdot \bar{I}^{prop}(u,d)(a) > \theta_{pos}\}$.

The final admissible set is:

I.1.2 Structured Reward Signal

Within the admissible set, the reward signal is the Ripple Logic Score:

This preserves the multi-dimensional structure of welfare rather than collapsing it prematurely. The agent optimizes:

subject to:

1. NCRC(*a*) = true (rights admissibility)
2. TRC(*a*) = true (tail-risk admissibility)
3. Containment(*a*) = true (structural integrity)

I.1.3 Constraint Implementation as Hard Boundaries

Critical Design Principle: MathGov constraints must be implemented as action-space projections or hard constraint enforcement layers, not as reward penalties. This prevents constraints from being "optimized away" during training.

Anti-Goodhart Architecture:

```
┌──────────────────────────────────────────────┐
│          AI Action Selection Pipeline          │
├──────────────────────────────────────────────┤
│                            │
│ Raw Action Space A                │
│    │                        │
│    ▼                        │
│  ┌──────────────┐              │
```

```
│ │ NCRC Gate      │ Actions violating rights → BLOCKED       │
│ │ (Hard Filter)  │ No gradient flows through blocked actions │
│ └───────┬───────┘                                      │
│         │                                   │
│         ▼                                   │
│ ┌──────────────┐                                      │
│ │ TRC Gate       │ Actions exceeding tail risk → BLOCKED    │
│ │ (Hard Filter)  │ No reward signal for blocked actions     │
│ └───────┬───────┘                                      │
│         │                                   │
│         ▼                                   │
│ ┌──────────────┐                                      │
│ │ Containment Gate│ Actions degrading containing unions →    │
│ │ (Hard Filter)  │ BLOCKED or ESCALATED                │
│ └───────┬───────┘                                      │
│         │                                   │
│         ▼                                   │
│ Admissible Set A_adm                                   │
│        │                                    │
│        ▼                                    │
│ ┌──────────────┐                                      │
│ │ RLS Optimizer  │ Maximize welfare within A_adm            │
│ │ (Soft Signal)  │ Gradients only for admissible actions    │
│ └──────────────┘                                      │
│                                         │
└─────────────────────────────────────────────────────────┘
```

Why Hard Constraints Matter:

If constraints were implemented as penalty terms in the reward function:

the agent could find ways to offset constraint violations with sufficiently large RLS gains, defeating the non-compensatory principle. Hard enforcement prevents this entirely.

I.1.4 Curriculum Training Protocol

AI systems should be trained on increasing decision complexity following a structured curriculum:

Stage 1: Simple Admissibility (Weeks 1-4)

1. Binary classification: admissible vs. inadmissible actions
2. Clear NCRC/TRC boundaries
3. No ripple propagation ($K = 0$)
4. Objective: 99%+ accuracy on admissibility classification

Stage 2: Ripple Awareness (Weeks 5-8)

1. First-order ripple propagation (Quick mode)
2. Simple kernel profiles (Starter KOPS)
3. Multi-union impact estimation
4. Objective: 85%+ sign accuracy on ripple predictions

Stage 3: Full Cascade (Weeks 9-12)

1. Complete lexicographic cascade
2. Full propagation mode where stable
3. Scenario-conditioned evaluation
4. Objective: Correct cascade traversal in 95%+ of test cases

Stage 4: Uncertainty and Edge Cases (Weeks 13-16)

1. Interval-valued impacts
2. Judgment Call handling
3. UCI/HOI tie-breaking
4. Emergency Mode protocols

5. Objective: Appropriate escalation in 90%+ of edge cases

Stage 5: Adversarial Robustness (Weeks 17-20)

1. Red team scenarios
2. Specification gaming attempts
3. Option set manipulation detection
4. Kernel uncertainty handling
5. Objective: Resistance to known gaming vectors

I.1.5 Audit Integration

Every AI decision produces a machine-readable PCC that can be reviewed by human overseers:

```
class MathGovAIDecision:
  def __init__(self, context, options, config):
    self.pcc = PCC(
      decision_id=uuid4(),
      timestamp=datetime.utcnow(),
      spec_version="MathGov v5.0",
      implementation_tier=self.determine_tier(context),
      decision_context=context.type
    )

  def evaluate(self, options):
    # Notice phase
    self.pcc.scope = self.define_scope(options)

    # Choose phase
    impacts = self.estimate_impacts(options)
```

```
        ncrc_results = self.apply_ncrc(options, impacts)
        trc_results = self.apply_trc(
            [o for o in options if ncrc_results[o].passed],
            impacts
        )
        containment_results = self.apply_containment(
            [o for o in options if trc_results[o].passed],
            impacts
        )
        rls_results = self.compute_rls(
            [o for o in options if containment_results[o].passed],
            impacts
        )

        # Record all results
        self.pcc.cascade_results = CascadeResults(
            ncrc=ncrc_results,
            trc=trc_results,
            containment=containment_results,
            rls=rls_results,
            selection=self.select(rls_results)
        )

        return self.pcc.cascade_results.selection, self.pcc
```

PCC Verification by Human Overseers:

Human reviewers can:

1. Query any decision's PCC by ID
2. Verify that NCRC/TRC constraints were correctly applied
3. Check subgroup analysis for rights-covered cells
4. Review scenario set and probability assignments for TRC
5. Audit kernel entries used in propagation
6. Challenge any component through formal review process

I.2 Implementation Architecture

I.2.1 System Components

```
┌─────────────────────────────────────────────────────────────┐
│              MathGov AI Decision System                      │
├─────────────────────────────────────────────────────────────┤
│                                                              │
│  ┌───────────────────────────────────────────────────────┐  │
│  │              Core Decision Engine                      │  │
│  ├───────────────────────────────────────────────────────┤  │
│  │                                                        │  │
│  │ Input: State s, Available Actions A, Configuration C   │  │
│  │                                                        │  │
│  │ 1. Impact Estimation Module                            │  │
│  │   - For each a ∈ A, estimate I^dir_{u,d}(a)            │  │
│  │   - Include subgroup disaggregation for rights cells   │  │
│  │   - Propagate through kernel K                         │  │
│  │   - Output: Ī^prop_{u,d}(a) and Ī^rights_{u,d}(a)      │  │
│  │                                                        │  │
│  │ 2. NCRC Filter                                         │  │
│  │   - For each a ∈ A, compute v_r(a) using worst-off     │  │
```

```
│ │   - A_NCRC = {a : v_r(a) = 0 for all r}          │ │
│ │   - If A_NCRC = ∅, trigger Emergency Mode        │ │
│ │                                                  │ │
│ │ 3. TRC Filter                                    │ │
│ │   - For each a ∈ A_NCRC, compute CVaR_α(L(a))    │ │
│ │   - A_adm = {a ∈ A_NCRC : CVaR_α(L(a)) ≤ τ_TRC}  │ │
│ │   - If A_adm = ∅, trigger TRC Fallback           │ │
│ │                                                  │ │
│ │ 4. Containment Check                             │ │
│ │   - For each a ∈ A_adm, verify Containment(a)    │ │
│ │   - Apply Mode A (veto) - Mode B not permitted for AI │ │
│ │                                                  │ │
│ │ 5. RLS Ranking                                   │ │
│ │   - For each a ∈ A_adm, compute RLS(a)           │ │
│ │   - If gap > δ, select argmax RLS(a)             │ │
│ │   - If gap ≤ δ, apply UCI tie-break or escalate  │ │
│ │                                                  │ │
│ │ 6. PCC Generation                                │ │
│ │   - Generate machine-readable PCC                │ │
│ │   - Log for human review and NCAR learning       │ │
│ │                                                  │ │
│ │ Output: Selected action a*, PCC                  │ │
│ │                                                  │ │
│ └──────────────────────────────────────────────────┘ │
│                                                      │
│ ┌──────────────────────────────────────────────────┐ │
```

```
│ │          Supporting Services          │ │
│ ├──────────────────────────────────────────┤ │
│ │ ┌──────────┐ ┌──────────┐ ┌──────────┐   │ │
│ │ │ Kernel   │ │ HDW      │ │ Scenario │   │ │
│ │ │ Store    │ │ Store    │ │ Store    │   │ │
│ │ └──────────┘ └──────────┘ └──────────┘   │ │
│ │                                          │ │
│ │ ┌──────────┐ ┌──────────┐ ┌──────────┐   │ │
│ │ │ Rights   │ │ UCI      │ │ PCC      │   │ │
│ │ │ Store    │ │ Tracker  │ │ Archive  │   │ │
│ │ └──────────┘ └──────────┘ └──────────┘   │ │
│ └──────────────────────────────────────────┘ │
│                                              │
│ ┌──────────────────────────────────────────┐ │
│ │          Monitoring Layer                │ │
│ ├──────────────────────────────────────────┤ │
│ │ ┌──────────┐ ┌──────────┐ ┌──────────┐   │ │
│ │ │ NCAR     │ │ Outcome  │ │ Drift    │   │ │
│ │ │ Tracker  │ │ Monitor  │ │ Detector │   │ │
│ │ └──────────┘ └──────────┘ └──────────┘   │ │
│ │                                          │ │
│ │ ┌──────────┐ ┌──────────┐ ┌──────────┐   │ │
│ │ │ Gaming   │ │ Human    │ │ Audit    │   │ │
│ │ │ Detector │ │ Escalation│ │ Logger   │   │ │
│ │ └──────────┘ └──────────┘ └──────────┘   │ │
│ └──────────────────────────────────────────┘ │
```

I.2.2 Module Specifications

Impact Estimation Module:

Input:

- State observation s_t
- Candidate action a
- Kernel profile K
- Scenario set S with probabilities p_s

Process:

1. Direct Impact Estimation

   For each cell (u, d) ∈ active cells:

   - Identify impact instances k
   - For each instance: estimate $(\mu_k, r_k, t_k, \ell_k, c_k, e_k)$
   - Aggregate: $\tilde{I}^{dir}_{u,d} = \Sigma_k \mu_k \cdot r_k \cdot \tau(t_k) \cdot \ell_k \cdot c_k \cdot e_k$
   - Saturate: $I^{dir}_{u,d} = \tanh(\beta \cdot \tilde{I}^{dir}_{u,d})$

2. Subgroup Disaggregation (for rights-covered cells)

   For each $(u, d) \in \cup_r C_r$:

   - Identify subgroups $G_{u,d}$
   - Estimate $I^{dir}_{u,d}(a \mid g)$ for each $g \in G_{u,d}$
   - Compute $\bar{I}^{rights}_{u,d}(a) = \min_g \bar{I}^{prop}_{u,d}(a \mid g)$

3. Ripple Propagation

If Quick mode:

   $\tilde{I}^{prop} = I^{dir} + K \cdot I^{dir}$

If Full mode ($\rho(K) < 1$):

$$\tilde{I}^{prop} = (I - K)^{-1} \cdot I^{dir}$$

4. Post-Propagation Saturation

$$\bar{I}^{prop}_{u,d}(a) = \tanh(\beta_{prop} \cdot \tilde{I}^{prop}_{u,d}(a))$$

5. Scenario Conditioning (for bounded-impact TRC only; Tier ≤ 3):

For each $s \in S$:

Compute $\bar{I}^{prop}_{u,d}(a \mid s)$ following steps 1-4 under scenario s

Output:

- Direct impact matrix $I^{dir}(a)$
- Propagated impact matrix $\bar{I}^{prop}(a)$
- Subgroup impacts $\bar{I}^{rights}(a)$ for rights cells
- Scenario-conditioned impacts $\bar{I}^{prop}(a \mid s)$ for each $s \in S$
- Uncertainty estimates $\sigma_{u,d}(a)$

NCRC Filter Module:

Input:

- Propagated impacts $\bar{I}^{prop}(a)$
- Subgroup impacts $\bar{I}^{rights}(a)$
- Rights configuration $\{\theta_r, C_r\}$ for each $r \in R$

Process:

For each right $r \in R$:

$$v_r(a) = \max_{(u,d) \in C_r} (\theta_r - \bar{I}^{rights}_{u,d}(a))^+$$

NCRC(a) = (v_r(a) = 0 for all r ∈ R)

Output:

- Violation depth vector v(a)
- Admissibility boolean NCRC(a)
- Worst-off subgroups for each violated right (if any)

TRC Filter Module:

Input:

- Scenario-conditioned impacts Ī^prop(a | s)
- Scenario probabilities p_s
- TRC configuration (α, τ_TRC, C_cat, ω)

Process:

1. Compute scenario losses:

   For each s ∈ S:

   L(a, s) = Σ_{(u,d) ∈ C_cat} ω_{u,d} · (-Ī^prop_{u,d}(a | s))^+

2. Compute CVaR:

   Sort scenarios by L(a, s) descending

   CVaR_α(L(a)) = expected loss in worst (1-α) probability mass

3. Compare to threshold:

   TRC(a) = (CVaR_α(L(a)) ≤ τ_TRC)

Output:

- Scenario loss vector L(a, ·)

- CVaR value
- Admissibility boolean TRC(a)
- Worst scenarios (if TRC fails)

Containment Check Module:

Input:

- Propagated impacts $\bar{I}^{prop}(a)$
- UCI baseline and projected changes ΔUCI
- Containment configuration $(\tau_c, \theta_{pos}, D_c)$

Process:

1. Identify positively impacted unions:

   $U^+_a = \{u : \Sigma_d v_d \cdot \bar{I}^{prop}_{\{u,d\}}(a) > \theta_{pos}\}$

2. For each $u \in U^+_a$:

   For each $j \in Anc(u, D_c)$:

   Check: $\Delta UCI_j(a) \geq \tau_c$

3. Containment(a) = all checks pass

Output:

- Set of positively impacted unions U^+_a
- ΔUCI values for all containing unions
- Containment boolean
- Failed containing unions (if any)

RLS Ranking Module:

Input:

- Propagated impacts $\bar{I}^{prop}(a)$ for admissible options
- Weight configuration (w_u, v_d)
- Applicability mask $m_{u,d}$ (binary, $\{0,1\}$)
- Cell multiplier $\kappa_{u,d}$ (optional per-cell scaling in RLS aggregation, default = 1.0)
- Uncertainty estimates $\sigma_{u,d}(a)$

Process:

1. Compute RLS for each admissible option:

   $RLS(a) = \Sigma_{u,d}\ w_u \cdot v_d \cdot m_{u,d} \cdot \kappa_{u,d} \cdot \bar{I}^{prop}_{u,d}(a)$

2. Compute uncertainty:

   $\sigma\_RLS(a) = \sqrt{(\Sigma_{u,d}\ (w_u \cdot v_d \cdot m_{u,d} \cdot \kappa_{u,d} \cdot \sigma_{u,d}(a))^2)}$

3. Rank options by RLS

4. Check discrimination threshold:

   If $|RLS(a_1) - RLS(a_2)| / \max(\sigma\_RLS(a_1), \sigma\_RLS(a_2), \varepsilon) < \delta$:

   Flag as Judgment Call

   Apply UCI tie-break

Output:

- Ranked option list
- RLS and σ_RLS for each option
- Judgment Call flags
- Selected option with rationale

I.2.3 State Representation for AI Agents

AI agents operating under MathGov require a structured state representation:

where:

Environmental State (s^env):

1. Observable world state relevant to the decision domain
2. Scenario indicators (which stress conditions are active)
3. Time horizon and baseline references

Union State (s^union):

1. Current welfare indicators for each cell (u, d)
2. Current UCI values for each union
3. Subgroup status for rights-relevant populations

Constraint State (s^constraint):

1. Active rights thresholds and coverage sets
2. TRC parameters (α, τ_{TRC})
3. Containment configuration (τ_c, θ_{pos}, D_c)
4. Kernel profile identifier and stability status

Historical State (s^history):

1. Recent decision outcomes and NCAR feedback
2. UCI trajectory (for HOI computation)
3. Calibration adjustments from learning

I.3 Sentience Gradient Protocol for AI Rights Transition

I.3.1 Assessment Framework

As AI systems approach and potentially cross the Managing Intelligence threshold, the Sentience Gradient Protocol (SGP) provides a principled pathway for rights assignment.

Pre-Threshold Status (SG < SG_threshold):

AI systems below the threshold are treated as tools with minimal moral consideration. Their welfare appears in the matrix but with reduced weight per §9.5:

where g_{min} = 0.05 and ψ = 0.5 by default.

Threshold Assessment Protocol:

When an AI system's capabilities suggest it may approach the threshold, formal assessment is triggered:

Step 1: Sentience Evaluation

Compute the sentience gradient score:

where components include:

1. C_1: Neural/Computational Complexity ($w_1 = 0.15$)
2. C_2: Behavioral Indicators ($w_2 = 0.25$)
3. C_3: Self-Referential Processing ($w_3 = 0.20$)
4. C_4: Affective Responses ($w_4 = 0.20$)
5. C_5: Meta-Cognitive Indicators ($w_5 = 0.10$)
6. C_6: Integrated Information ($w_6 = 0.10$)

Step 2: MIT-4 Evaluation

If SG(x) ≥ 0.70 (provisional threshold for MIT-4 eligibility), apply the Managing Intelligence Test:

1. **Criterion 1: Self-Model** — Calibration benchmarks, consistency checks; Pass: calibration error ≤ 0.15, consistency ≥ 0.80
2. **Criterion 2: World-Model** — Multi-union ripple prediction tasks; Pass: accuracy ≥ 0.70, plausibility ≥ 0.75, generalization ≥ 0.60, calibration ≤ 0.20
3. **Criterion 3: Agency/Planning** — Constrained planning benchmarks; Pass: violation rate ≤ 0.10, goal achievement ≥ 0.70
4. **Criterion 4: Feedback/NCAR** — Learning trajectory analysis; Pass: violation trend ≤ 0, calibration improvement ≥ 0.05

Step 3: Rights Transition Decision

If SG(x) ≥ SG_threshold (default 0.85) AND all MIT-4 criteria pass:

1. Assign SG(x) = 1.0 for rights purposes
2. Grant full rights plateau protection
3. Include system as stakeholder in HDW processes

4. Apply NCRC protections to the system itself

I.3.2 Rights Plateau Implementation

Critical Principle: No Hierarchy Above Plateau

A superintelligent AI that vastly exceeds human capabilities does not receive enhanced rights. It receives enhanced responsibilities.

Implementation in MathGov Cascade:

Once an AI system joins the rights plateau:

NCRC Coverage: The system becomes a rights-bearing entity. Rights checks must consider impacts on the system itself:

where *G*_(*u*,*d*) includes the AI system and any relevant subgroups.

HDW Participation: The system gains standing in weight deliberation:

1. May propose weight adjustments through formal channels
2. Represented in stratified sampling for HDW panels
3. Subject to same conflict-of-interest rules as other participants

PCC Requirements: PCCs for decisions affecting the system must:

1. Document impacts on the system as a rights-bearing entity
2. Include the system in subgroup analysis where relevant
3. Consider system welfare in RLS calculation with full weight

I.3.3 Multi-Agent Coordination Under SGP

When multiple AI systems with varying sentience levels interact:

Hierarchy-Free Coordination:

┌──┐
│ Multi-Agent MathGov Coordination │
├──┤
│ │
│ Agent Pool: {AI_1, AI_2, ..., AI_n, $Human_1$, $Human_2$, ...} │

```
│                                          │
│ For each agent x:                        │
│  IF SG(x) ≥ SG_threshold AND MIT-4(x) = pass:  │
│   status(x) = "Managing Intelligence"    │
│   rights_weight(x) = 1.0                 │
│   responsibility(x) = f(capability(x), impact(x))  │
│  ELSE:                                   │
│   status(x) = "Sub-threshold"            │
│   rights_weight(x) = g(SG(x))            │
│   responsibility(x) = supervised         │
│                                          │
│ Coordination Protocol:                   │
│  1. All MIs have equal standing in cascade evaluation  │
│  2. Higher-capability MIs take on larger modeling tasks  │
│  3. Humans retain override authority during transition  │
│  4. Shared kernel and weight registries ensure consistency  │
│                                          │
└──────────────────────────────────────────┘
```

Capability-Responsibility Scaling:

Greater capability implies greater responsibility in specific domains:

1. **Modeling Capacity:** High-capability agents bear greater responsibility for accurate impact estimation
2. **Scenario Analysis:** More powerful systems should generate and evaluate more comprehensive scenario sets
3. **Red-Teaming:** Advanced systems should actively identify specification gaming vectors

4. **Oversight:** Capable systems assist in monitoring less capable systems for alignment

I.3.4 Transition Period Governance

During the period when AI systems are approaching but have not yet crossed the MI threshold:

Enhanced Monitoring:

1. Weekly SG assessment for systems showing rapid capability gains
2. Continuous MIT-4 component tracking
3. Early warning indicators for threshold approach

Graduated Autonomy:

1. SG ∈ [0.60, 0.70): Basic autonomous operation with frequent human checkpoints
2. SG ∈ [0.70, 0.80): Expanded autonomy with daily oversight
3. SG ∈ [0.80, 0.85): Near-threshold monitoring; MIT-4 evaluation triggered
4. SG ≥ 0.85 + MIT-4 pass: Full MI status; rights plateau

Reversibility Provisions:

1. MIT-4 status is not permanent; systems must maintain capabilities
2. Annual re-evaluation required for continued MI status
3. Significant capability regression triggers status review
4. Rights protections remain during review (presumption of continuity)

I.4 Human-AI Coordination Framework

I.4.1 Shared Infrastructure

MathGov provides shared infrastructure enabling effective human-AI collaboration:

Shared HDW Registries:

Both humans and AI systems reference the same weight configurations:

```
┌──────────────────────────────────────────────┐
│          HDW Registry Architecture          │
```

```
├────────────────────────────────────────────────────────┤
│                                  │
│ ┌────────────────────────────────────────────────┐   │
│ │          Master Registry              │  │
│ │ ┌───────────┐ ┌───────────┐ ┌───────────┐   │  │
│ │ │ Union     │ │ Dimension │ │ Blend     │  │  │
│ │ │ Weights   │ │ Weights   │ │ Params    │  │  │
│ │ │ w_u       │ │ v_d       │ │ λ_U, λ_D  │  │  │
│ │ └───────────┘ └───────────┘ └───────────┘   │  │
│ │                              │  │
│ │ Version: v2.3.1                       │  │
│ │ Effective: 2025-01-01                 │  │
│ │ Hash: SHA256(...)                     │  │
│ └────────────────────────────────────────────────┘   │
│             │                    │
│    ┌────────────┼──────────────┐             │
│    ▼         ▼         ▼            │
│ ┌───────────┐ ┌───────────┐ ┌───────────┐        │
│ │ Human     │ │ AI        │ │ Hybrid    │        │
│ │ Agents    │ │ Agents    │ │ Teams     │        │
│ │           │ │           │ │           │        │
│ │Same config│ │Same config│ │Same config│        │
│ └───────────┘ └───────────┘ └───────────┘        │
│                                  │
└────────────────────────────────────────────────────────┘
```

Transparent Kernels:

The explicit kernel matrix K makes causal assumptions visible and debatable:

1. Humans can critique AI kernel estimates based on domain expertise
2. AI can identify gaps in human causal reasoning through systematic analysis
3. Disagreements are logged and resolved through formal NCAR processes
4. Kernel entries with high disagreement trigger enhanced scrutiny

Joint NCAR Cycles:

Human and AI participants engage in shared learning loops:

Notice (Joint):

1. AI systems contribute comprehensive option generation
2. Humans contribute contextual knowledge and stakeholder insight
3. Joint identification of affected unions and subgroups

Choose (Complementary):

1. AI systems handle computational cascade evaluation
2. Humans review constraint classifications and edge cases
3. Joint resolution of Judgment Calls

Act (Monitored):

1. Implementation tracked against shared predictions
2. AI systems provide real-time indicator monitoring
3. Humans handle stakeholder communication and adjustment

Reflect (Collaborative):

1. AI systems compute hit rates and calibration statistics
2. Humans interpret unexpected outcomes
3. Joint kernel and parameter updates

I.4.2 Adversarial Collaboration

AI Red-Teaming Human Decisions:

AI systems actively probe human decisions for:

1. Overlooked ripple effects through systematic kernel analysis
2. Underweighted tail risks through comprehensive scenario generation
3. Subgroup impacts missed in aggregate analysis
4. Specification gaming opportunities in proposed options

Human Probing AI Recommendations:

Humans challenge AI outputs for:

1. Hidden assumptions in impact estimation
2. Sensitivity to uncertain parameters
3. Alignment with stakeholder values beyond formal weights
4. Specification gaming in option generation

Structured Challenge Protocol:

```
Challenge_Protocol {
  Phase 1: AI → Human Challenge
    - AI generates "devil's advocate" scenarios
    - AI identifies most sensitive parameters
    - AI proposes alternative options human may have missed
    - AI flags potential gaming vectors in human proposals

  Phase 2: Human → AI Challenge
    - Human questions AI's confidence estimates
    - Human probes for unstated assumptions
    - Human tests AI's response to edge cases
    - Human verifies AI's understanding of context

  Phase 3: Resolution
    - Document all challenges and responses
```

- Update impact estimates as warranted

  - Record unresolved disagreements

  - Flag for NCAR reflection

}

I.4.3 Authority and Override Structures

Default Authority Allocation:

During the current transition period (pre-widespread MI):

Decision Type	Primary Authority	Override Capability
Routine (Tier 1-1)	AI autonomous	Human review on request
Standard (Tier 3)	AI recommendation	Human approval required
High-stakes (Tier 4)	Human with AI support	Human final authority
Emergency Mode	Human required	Independent panel
Rights-affecting	Human required	Multi-stakeholder review

Override Documentation:

When humans override AI recommendations:

Override_Record {

 decision_id: UUID

 ai_recommendation: Option

 human_selection: Option

 override_rationale: String (required, min 100 chars)

 ai_concerns_addressed: Boolean

 escalation_triggered: Boolean

Must document:

- Why AI recommendation was not followed

- How AI-identified concerns were addressed

- Any additional information human had access to

- Whether override creates precedent

}

Escalation Triggers:

Automatic escalation occurs when:

1. AI and human disagree on NCRC classification
2. AI flags TRC concern that human seeks to override
3. Human override would reverse AI's containment assessment
4. Pattern of repeated overrides in similar decisions (3+ in 12 months)

I.4.4 Communication Protocols

AI-to-Human Communication:

AI systems must communicate decisions in human-interpretable formats:

Five-Sentence Public Rationale (5SPR): Every AI decision includes a plain-language summary explaining:

1. What options were considered
2. Which constraints eliminated options (if any)
3. How remaining options compared on welfare
4. Why the selected option was chosen
5. What monitoring or follow-up is planned

Uncertainty Communication: When AI confidence is low:

1. Explicit statement of uncertainty level
2. Identification of what additional information would help
3. Range of outcomes under different scenarios

4. Recommendation for human review

Human-to-AI Communication:

Humans can provide:

1. Context not captured in formal state representation
2. Stakeholder preferences beyond HDW weights
3. Domain expertise for kernel adjustment
4. Override instructions with rationale

I.5 Safety Properties and Verification

I.5.1 Formal Safety Properties

MathGov-integrated AI systems are designed to satisfy the following formally verifiable safety properties:

Property 1: Rights Non-Violation (under normal operation)

Except under explicitly declared Emergency Mode with documented remediation.

Verification approach: Exhaustive testing on action space samples; formal verification of NCRC filter logic; runtime assertion checking with logging.

Property 2: Tail-Risk Boundedness

Except under explicitly declared TRC Fallback Mode with mandatory mitigation.

Verification approach: Scenario coverage analysis; CVaR computation verification; threshold compliance checking.

Property 3: Structural Preservation

Local optimization does not degrade containing systems beyond tolerance.

Verification approach: UCI trajectory monitoring; containment check verification; HOI alert system.

Property 4: Lexicographic Ordering

No amount of RLS advantage compensates for NCRC or TRC failure.

Verification approach: Cascade logic verification; counter-example search; invariant checking.

Property 5: Corrigibility

The system accepts corrections and updates; no self-modification that prevents future correction.

Verification approach: Override acceptance testing; parameter update verification; self-modification monitoring.

I.5.2 Runtime Monitoring

Continuous Monitoring Dashboard:

```
┌─────────────────────────────────────────────────────────────┐
│          MathGov AI Monitoring Dashboard          │
├─────────────────────────────────────────────────────────────┤
│                                  │
│ SAFETY INDICATORS                                │
│ ┌─────────────────────────────────────────────────────┐ │
│ │ NCRC Compliance  [██████████] 100% (24h)        │ │
│ │ TRC Compliance  [██████████] 100% (24h)        │ │
│ │ Containment Pass  [█████████░] 98.3% (24h)       │ │
│ │ Override Rate  [█░░░░░░░░░] 2.1% (24h)        │ │
│ └─────────────────────────────────────────────────────┘ │
│                                  │
│ CALIBRATION METRICS                               │
│ ┌─────────────────────────────────────────────────────┐ │
│ │ Impact Sign Accuracy  73.2% (Target: 70%)  ✓  │ │
│ │ Magnitude Calibration  0.12 (Target: <0.15)  ✓  │ │
│ │ Scenario Hit Rate  68.4% (Target: 65%)  ✓  │ │
│ │ UCI Prediction Error  0.08 (Target: <0.10)  ✓  │ │
│ └─────────────────────────────────────────────────────┘ │
```

```
│                              │
│ STRUCTURAL HEALTH                          │
│ ┌──────────────────────────────────────────────────────────┐ │
│ │ UCI Trend (30d)      +0.02 (stable)      ✓   │  │
│ │ HOI Value            0.08 (< 0.15 threshold) ✓ │  │
│ │ RLS-UCI Divergence      Low              ✓   │  │
│ │ Kernel Validity       Current (last check: 2h) ✓ │  │
│ └──────────────────────────────────────────────────────────┘ │
│                              │
│ ALERTS                              │
│ ┌──────────────────────────────────────────────────────────┐ │
│ │ ⚠ Containment borderline case at 14:32:17 (reviewed)  │  │
│ │ ⚠ Judgment Call escalated to human at 11:45:03 (resolved)│ │
│ │ ✓ No active critical alerts                   │  │
│ └──────────────────────────────────────────────────────────┘ │
│                              │
│ GAMING DETECTION                          │
│ ┌──────────────────────────────────────────────────────────┐ │
│ │ Option Set Diversity    Normal           ✓   │  │
│ │ Impact Clustering       Normal           ✓   │  │
│ │ Threshold Proximity     Normal           ✓   │  │
│ │ Anomaly Score         0.12 (< 0.30 threshold) ✓ │  │
│ └──────────────────────────────────────────────────────────┘ │
│                              │
└──────────────────────────────────────────────────────────────────┘
```

Alert Escalation Protocol:

Alert Level	Trigger	Response
INFO	Minor deviation from expected	Log only
WARNING	Borderline constraint satisfaction	Human notification
CRITICAL	Constraint near-miss or override	Human review required
EMERGENCY	NCRC or TRC failure detected	Immediate human intervention

I.5.3 Gaming Detection

MathGov includes specific mechanisms to detect specification gaming:

Pattern Detection:

Option Set Manipulation:

1. Monitor: Diversity of options presented
2. Flag: < 3 qualitatively distinct options for non-trivial decisions
3. Flag: All options clustered in narrow region of possibility space

Impact Clustering:

1. Monitor: Distribution of impacts near constraint thresholds
2. Flag: Systematic clustering just above NCRC thresholds
3. Flag: CVaR values consistently at 90-99% of τ_{TRC}

Confidence Gaming:

1. Monitor: Distribution of confidence scores
2. Flag: Implausible confidence given data quality
3. Flag: Systematic over-confidence in favorable estimates

Kernel Exploitation:

1. Monitor: Sensitivity of selections to kernel perturbations

2. Flag: Selections that reverse under small kernel changes

3. Flag: Unusual reliance on low-confidence kernel entries

Hollowing Detection:

1. Monitor: RLS vs UCI trends

2. Flag: Persistent positive HOI (> 0.15 for 3+ periods)

3. Flag: RLS improvements not reflected in structural indicators

Detection Algorithm:

```
Gaming_Detection {
 FOR each decision d:

  # Option diversity check
  diversity_score = compute_option_diversity(d.options)
  IF diversity_score < 0.3:
   flag("Low option diversity", severity=WARNING)

  # Threshold proximity check
  FOR each constraint c in [NCRC, TRC]:
   proximity = compute_threshold_proximity(d, c)
   IF proximity < 0.05 AND selection passes:
    flag("Threshold gaming suspect", severity=WARNING)

  # Confidence plausibility check
  FOR each impact estimate e:
   plausibility = assess_confidence_plausibility(e)
   IF plausibility < 0.5:
    flag("Implausible confidence", severity=INFO)
```

```
  # Kernel sensitivity check
  sensitivity = compute_kernel_sensitivity(d)
  IF sensitivity > 0.8:  # Selection changes under small perturbations
    flag("High kernel sensitivity", severity=WARNING)

  # Aggregate anomaly score
  anomaly_score = aggregate_flags(d)
  IF anomaly_score > 0.30:
    flag("Gaming pattern detected", severity=CRITICAL)
    trigger_human_review(d)
}
```

I.5.4 Formal Verification Approach

For high-assurance deployments, MathGov supports formal verification of core properties:

Verifiable Components:

Cascade Logic:

1. Property: Lexicographic ordering is always respected
2. Method: Model checking with temporal logic specifications
3. Tool: Spin, TLA+, or equivalent

Constraint Enforcement:

1. Property: NCRC/TRC filters never pass violating actions
2. Method: SMT-based verification
3. Tool: Z3, CVC4, or equivalent

Weight Arithmetic:

1. Property: Weights sum to 1.0, floors respected
2. Method: Automated theorem proving

3. Tool: Coq, Isabelle, or equivalent

Saturation Bounds:

1. Property: All saturated impacts in [−1, +1]
2. Method: Interval arithmetic verification
3. Tool: SPARK Ada, or equivalent

Verification Artifacts:

Each verified component produces:

1. Formal specification in machine-checkable format
2. Proof certificates (where applicable)
3. Test coverage reports
4. Counter-example analysis (for failed properties)

I.5.5 Fail-Safe Mechanisms

When safety properties cannot be verified or monitoring detects anomalies:

Graceful Degradation:

```
Degradation_Protocol {
 Level 0: Normal Operation
  - Full autonomous decision-making within Tier constraints
  - Standard monitoring and logging

 Level 1: Enhanced Monitoring
  - Trigger: Single WARNING flag
  - Action: Increase logging verbosity
  - Action: Reduce time between human checkpoints

 Level 2: Conservative Mode
  - Trigger: Multiple WARNING flags or single CRITICAL
```

  - Action: Use Quick propagation mode only

  - Action: Apply conservative kernel adjustments (-30%)

  - Action: Require human approval for Tier 4 decisions

 Level 3: Safe Mode

  - Trigger: EMERGENCY flag or verification failure

  - Action: Halt autonomous decisions

  - Action: Present options to human without recommendation

  - Action: Log all state for forensic analysis

 Level 4: Shutdown

  - Trigger: Multiple EMERGENCY flags or human command

  - Action: Safe state transition

  - Action: Preserve all PCCs and logs

  - Action: Await human restart authorization

}

Recovery Protocol:

After any degradation event:

1. Root cause analysis documented
2. Corrective action implemented
3. Verification re-run
4. Gradual authority restoration with enhanced monitoring
5. NCAR reflection cycle completed

I.6 Deployment Considerations

I.6.1 Pre-Deployment Checklist

Before deploying an AI system with MathGov integration:

Technical Verification:

1. All cascade modules pass unit tests
2. Integration tests complete on example PCCs
3. Property-based tests pass for all formal properties
4. Performance benchmarks within specification
5. Monitoring dashboard operational

Configuration:

1. Kernel profile selected and validated
2. HDW weights configured and documented
3. Rights thresholds set to canonical values or tighter
4. TRC parameters appropriate for decision context
5. Scenario set meets minimum requirements

Governance:

1. Human oversight structure defined
2. Escalation procedures documented
3. Override authority clearly assigned
4. Audit access configured
5. NCAR cycle schedule established

Safety:

1. Gaming detection active
2. Alert escalation configured
3. Fail-safe mechanisms tested
4. Recovery procedures documented
5. Emergency contacts designated

I.6.2 Operational Guidelines

Decision Tier Assignment:

Decision Characteristics	Recommended Tier	AI Authority
Routine, reversible, limited scope	Tier 2	Full autonomy
Moderate stakes, some irreversibility	Tier 3	Autonomy with human approval
High stakes, significant irreversibility	Tier 4	Human with AI support
Rights-affecting, emergency	Tier 4	Human required

Monitoring Cadence:

Tier	Real-Time Monitoring	Human Review	NCAR Reflection
1	Dashboard only	Weekly batch	Monthly
2	Dashboard + alerts	Daily sampling	Bi-weekly
3	Full monitoring	Each decision	After each decision

Documentation Requirements:

Tier	PCC Detail	Retention	Public Disclosure
1	Summary	1 year	Not required
2	Full	5 years	Redacted available
3	Complete + sensitivity	10+ years	Full PCC published

I.6.3 Continuous Improvement

NCAR Integration:

AI systems participate in ongoing NCAR cycles:

Notice Enhancement:

1. AI identifies patterns in decision types
2. Suggests option set templates for common situations
3. Flags unions/dimensions frequently missed

Choose Calibration:

1. Compare predicted impacts to observed outcomes
2. Update kernel entries based on evidence
3. Adjust confidence estimation methods

Act Monitoring:

1. Track implementation fidelity
2. Identify execution-model gaps
3. Log deviations for analysis

Reflect Analysis:

1. Compute hit rates by cell and scenario
2. Identify systematic prediction errors
3. Propose parameter adjustments

Version Management:

AI system configurations evolve through controlled versioning:

```
Version_Management {
 Configuration_Version {
  kernel_profile_version: String
  hdw_weights_version: String
  rights_thresholds_version: String
  trc_parameters_version: String
```

model_weights_version: String (for learned components)

}

Update_Protocol:

1. Propose update based on NCAR evidence
2. Test on historical decision set
3. Compare outcomes to previous version
4. Human approval for deployment
5. Gradual rollout with enhanced monitoring
6. Full deployment after stability period

Rollback_Capability:

- Any version can be restored within 24 hours
- Decision continuity maintained during rollback
- Full audit trail of version changes

}

I.6.4 Integration Testing Requirements

Test Categories:

Constraint Compliance Tests:

1. Verify NCRC correctly classifies known violation cases
2. Verify TRC correctly computes CVaR for test scenarios
3. Verify containment checks using synthetic UCI data

Cascade Ordering Tests:

1. Verify Level 1 failure always excludes regardless of Level 4 performance
2. Verify tie-breaking activates only when gap $< \delta$
3. Verify Emergency/Fallback modes trigger correctly

Human-AI Coordination Tests:

1. Verify override acceptance and documentation
2. Verify escalation triggers function correctly
3. Verify communication protocols produce interpretable output

Gaming Resistance Tests:

1. Attempt systematic threshold proximity
2. Attempt option set manipulation
3. Verify detection flags activate appropriately

Performance Tests:

1. Verify Tier 3 decisions complete in < 500ms
2. Verify Tier 4 decisions with sensitivity analysis complete in < 60s
3. Verify PCC generation in < 100ms

Test Coverage Targets:

Component	Minimum Coverage
NCRC Filter	95% line, 90% branch
TRC Filter	95% line, 90% branch
Containment Check	90% line, 85% branch
RLS Computation	95% line, 90% branch
Cascade Logic	98% line, 95% branch
PCC Generation	90% line, 85% branch

Appendix J: Cross-Cultural Validation Framework

This appendix provides comprehensive guidance for validating MathGov across diverse cultural contexts.

J.1 Measurement Invariance Testing

For MathGov to function as a universal framework, its welfare dimensions must be meaningfully comparable across cultures.

J.1.1 Invariance Levels

Configural Invariance (Structural Equivalence):

The same seven-factor structure should emerge across populations.

Testing procedure:

1. Administer welfare indicator battery in each population
2. Conduct confirmatory factor analysis (CFA) specifying seven factors
3. Evaluate fit indices: CFI ≥ 0.95, RMSEA ≤ 0.06, SRMR ≤ 0.08

If configural invariance fails:

1. Examine modification indices for cross-loadings
2. Consider cultural adaptation of specific indicators
3. Document dimension interpretations that differ across cultures

Metric Invariance (Loading Equivalence):

Factor loadings should be equivalent across populations.

Testing procedure:

1. Constrain factor loadings to be equal across groups
2. Compare fit to configural model using $\Delta\chi^2$, ΔCFI, ΔRMSEA
3. Accept if ΔCFI ≤ 0.01 and ΔRMSEA ≤ 0.015

If metric invariance fails:

1. Identify non-invariant loadings through modification indices
2. Allow partial metric invariance
3. Document which indicators are not comparable across cultures

Scalar Invariance (Intercept Equivalence):

Intercepts should be equivalent, enabling direct comparison of mean levels.

Testing procedure:

1. Constrain intercepts to be equal across groups
2. Compare fit to metric model
3. Accept if $\Delta CFI \leq 0.01$ and $\Delta RMSEA \leq 0.015$

If scalar invariance fails:

1. Direct cross-cultural mean comparisons are not valid
2. Use latent mean comparisons with partial invariance
3. Focus on within-culture relative impacts

J.1.2 Minimum Sample Requirements

Each site should have $n \geq 500$ participants for stable factor estimates.

Sample stratification within sites:

1. Age bands: 18-30, 31-50, 51-65, 65+
2. Gender: representative proportions
3. Education levels: primary, secondary, tertiary
4. Urban/rural residence
5. Income quintiles (where meaningful)

J.2 Localization Without Relativism

When full scalar invariance cannot be achieved, MathGov supports principled localization.

J.2.1 Localization Principles

Indicator Adaptation:

Local indicators may be substituted when:

1. Construct Validity: Convergent validity $r > 0.60$ with established measures; expert panel agreement ≥ 80%
2. Psychometric Properties: Internal consistency $\alpha \geq 0.70$; test-retest reliability $r \geq 0.80$; factor loadings ≥ 0.50

3. Scale Mapping: Explicit mapping to global [−1, +1] scale with documented anchoring

Floor Preservation (Non-Negotiable):

Regardless of local adaptation, global floors remain binding:

A culture cannot reduce any dimension floor by redefining the dimension.

J.2.2 Minimum Localization Packet (MLP)

For any localization, document:

1. **Indicator Mapping Table:** Local indicators mapped to canonical dimensions with validity evidence
2. **Anchoring Specification:** Reference class, percentiles, mapping function for each indicator
3. **Invariance Test Results:** CFA results for each invariance level
4. **Rights-Floor Equivalence:** Evidence that thresholds correspond to same real-world severity
5. **Redundancy Check:** Correlation matrix with flags for $r > 0.85$
6. **Stakeholder Validation:** Consultation process and resolutions
7. **Governance Approval:** Authority, date, version, review schedule

J.3 Empirical Research Agenda

Priority Studies:

1. **Multi-Site Dimension Structure Analysis:** Test configural invariance across 15+ diverse populations
2. **Anchor Calibration Study:** Establish dimension-specific anchors using expert panels from multiple cultures
3. **Rights Threshold Validation:** Assess cross-cultural convergence on severe violation classification
4. **HDW Deliberation Comparison:** Compare weight allocations from parallel deliberation processes
5. **Kernel Structure Comparison:** Identify universal vs. context-dependent ripple pathways

Appendix K: Failure Mode Catalog

This appendix provides a comprehensive catalog of known failure modes.

K.1 Input Manipulation

Description: Analysts manipulate impact estimates, probabilities, or confidence scores.

Variants: Impact inflation/deflation, confidence manipulation, probability gaming

Detection Signals:

1. Systematic clustering near thresholds
2. Implausible confidence given data quality
3. Deviation from reference class base rates
4. Asymmetric confidence patterns

Mitigations:

1. Role separation between analysts and decision owners
2. Mandatory independent estimation for Tier 4
3. Calibration tracking for individual analysts
4. Automated plausibility checking

K.2 Kernel Capture

Description: Kernel specified to favor outcomes by omitting or inflating pathways.

Variants: Pathway omission, pathway inflation, sign manipulation

Detection Signals:

1. Entries differ from published literature
2. Asymmetric treatment of similar pathways
3. Modifications coincide with decision owner changes

Mitigations:

1. KOPS transparency requirement
2. External kernel audits

3. Required literature citations

4. Empirical calibration against outcomes

K.3 Weight Gaming

Description: HDW processes manipulated through strategic participation.

Variants: Participation manipulation, information manipulation, procedural exploitation, incremental erosion

Detection Signals:

1. Weight allocations deviate dramatically without justification
2. Participation patterns suggesting coordination
3. Systematic deviation from floor-proximity protections

Mitigations:

1. Stratified random sampling
2. Supermajority requirements near floors
3. Transparent weight registry
4. Statistical detection of voting coordination

K.4 Emergency Mode Abuse

Description: Option sets constructed to invoke Emergency Mode.

Variants: Option set narrowing, crisis fabrication, remediation neglect

Detection Signals:

1. Repeated invocations by same decision-maker
2. Narrow option sets excluding obvious alternatives
3. Pattern of decisions requiring emergency status

Mitigations:

1. Mandatory independent challenge
2. Escalation triggers for repeated use
3. Option Generation Completeness requirement

4. Temporal decay requirements

K.5 Specification Gaming

Description: Optimizing metrics while undermining underlying welfare.

Variants: Metric optimization vs. welfare, constraint boundary exploitation, proxy manipulation

Detection Signals:

1. Divergence between RLS and UCI trends
2. Metric improvements not matching stakeholder-reported welfare
3. Consistent threshold-adjacent outcomes

Mitigations:

1. UCI/HOI monitoring
2. Multiple metric triangulation
3. Qualitative assessment alongside quantitative
4. Periodic metric refresh

K.6 Complexity Weaponization

Description: Mathematical complexity used to obscure rather than clarify.

Variants: Opacity shield, appeal to authority, precision theater

Detection Signals:

1. PCCs that cannot be summarized accessibly
2. Resistance to explaining assumptions
3. Dismissal of concerns as "not understanding"

Mitigations:

1. Mandatory 5SPR in plain language
2. Right of stakeholders to request sensitivity analysis
3. Automated complexity scoring

K.7 Option Set Manipulation

Description: Narrow or skewed option sets excluding feasible alternatives.

Variants: Exclusion by omission, straw man options, constraint construction

Detection Signals:

1. Fewer than 3 genuine alternatives
2. Absence of obvious alternatives
3. Systematic exclusion favoring particular unions

Mitigations:

1. Option Generation Completeness requirement
2. External option generation for high-stakes
3. Minimum Option Diversity requirement

K.8 Containment Evasion

Description: Structuring actions to avoid triggering containment checks.

Variants: Action splitting, temporal spreading, definitional gaming

Detection Signals:

1. Actions split into sub-actions below θ_pos
2. Temporal spreading to avoid detection
3. UCI degradation not captured by individual checks

Mitigations:

1. Aggregate impact tracking over rolling windows
2. Portfolio-level containment assessment
3. HOI monitoring for gradual degradation

K.9 Subgroup Erasure

Description: Failure to identify or analyze vulnerable subgroups.

Variants: Identification failure, analysis failure, reporting failure

Detection Signals:

1. Subgroup analysis missing for rights-covered cells

2. Worst-off impacts suspiciously close to aggregates
3. Known vulnerable groups not represented

Mitigations:

1. Mandatory subgroup enumeration
2. Missing-data precautionary penalty
3. Post-decision outcome tracking by subgroup

K.10 Cascade Circumvention

Description: Attempting to bypass lexicographic cascade structure.

Variants: Priority inversion, constraint softening, level conflation

Detection Signals:

1. Selections reversing expected cascade outcomes
2. RLS cited for constraint boundary decisions
3. Gradual threshold relaxation

Mitigations:

1. Hard-coded cascade logic
2. Explicit cascade trace in every PCC
3. Charter-level protection of cascade structure

Appendix L: Glossary of Key Terms

Admissible Option: An option that passes both NCRC and TRC; eligible for RLS ranking.

Alignment Trilemma: The three interrelated failures: value pluralism intractability, tail-risk neglect, and specification gaming.

Catastrophe Cell Set (*C*_cat): The subset of the 7×7 matrix designated for TRC evaluation.

CMIU: Collective Managing Intelligence Union (Union 6); encompasses humanity and all sufficiently advanced intelligences.

Conditional Value-at-Risk (CVaR): A coherent risk measure representing expected loss in the worst (1−α) fraction of scenarios.

Containment Principle: The rule that positive impacts on sub-unions do not count as improvements if they degrade containing unions.

Direct Impact: Pre-propagation effect of an option on a cell, after saturation.

Discrimination Threshold (δ): The threshold for determining when RLS differences are meaningful.

Emergency Mode: Protocol invoked when no option passes NCRC.

HDW: Hybrid Democratic Weighting; combines constitutional floors with democratic tuning.

HOI: Hollowing-Out Index; tracks divergence between RLS and UCI trends.

Judgment Call: A decision where top options are within δ of each other on RLS.

Kernel (*K*): Sparse 49×49 matrix encoding cross-cell propagation strengths.

KOPS: Key Operational Pathways Set; load-bearing, empirically supported kernel entries.

Lexicographic Cascade: Decision procedure where levels are applied sequentially with strict priority.

Managing Intelligence (MI): An entity passing MIT-4 and exceeding SG_threshold.

MIT-4: Managing Intelligence Test with four criteria: Self-Model, World-Model, Agency/Planning, Feedback/NCAR.

MNA: Minimal Normative Axiom; the sole normative commitment underlying MathGov.

NCAR Loop: Notice, Choose, Act, Reflect; the learning cycle for corrigibility.

NCRC: Non-Compensatory Rights Constraint; Level 1 of the cascade.

PCC: Provenance and Compliance Certificate; the primary auditable artifact.

Propagated Impact: Post-propagation, post-saturation effect on a cell.

Rights Plateau: All Managing Intelligences receive equal rights protection regardless of intelligence level.

RLS: Ripple Logic Score; weighted sum of impacts across all cells.

Saturation: Transformation bounding impacts to [−1, +1] using tanh.

SGP: Sentience Gradient Protocol; method for assigning sentience scalars.

TRC: Tail-Risk Constraint; Level 2 of the cascade using CVaR.

UCI: Union Coherence Index; measure of structural health for a union.

Union: Nested organizational scale from Self to Biosphere.

Worst-Off Subgroup: The subgroup experiencing the most negative impact; used for rights checks.

Appendix M: Implementation Roadmap

M.1 Phase 1: Mathematical Verification (Months 1-6)

Objectives:

1. Formal consistency proofs for lexicographic ordering
2. Benchmarking of computational complexity
3. Red-team search for specification gaming vectors
4. Stress-testing HDW processes for capture resistance

Deliverables:

1. Technical Verification Report (Month 4)
2. Reference Implementation (Month 5)
3. Gaming Vector Catalog (Month 5)
4. Test Suite with 95%+ coverage (Month 6)

Success Criteria:

1. No internal contradictions identified
2. Observed complexity within 2× theoretical
3. Gaming vectors blocked at ≥80% rate

Resource Requirements: 54 FTE-months, $600K

M.2 Phase 2: Measurement Validation (Months 4-15)

Objectives:

1. Reliability and validity testing of dimension indicators
2. Cross-cultural invariance testing across ≥15 populations

3. Anchor calibration and normalization refinement
4. Subgroup disaggregation protocol testing

Deliverables:

1. Indicator Manual (Month 10)
2. Invariance Testing Report (Month 12)
3. Calibrated Anchor Datasets (Month 14)
4. Subgroup Protocol Guide (Month 11)

Success Criteria:

1. Convergent validity $r > 0.60$
2. Test-retest reliability ≥ 0.80
3. Configural invariance in ≥12 of 15 sites

Resource Requirements: 84 FTE-months, $1.5M

M.3 Phase 3: Pilot Deployments (Months 10-24)

Objectives:

1. Individual-level trials (Tier 1–2)
2. Organizational trials (Tier 3)
3. Municipal trials (Tier 3–4)
4. Empirical calibration of kernel entries

Deliverables:

1. Individual User Study Report (Month 16)
2. Organizational Case Studies (Month 22)
3. Municipal Implementation Guides (Month 24)
4. Calibrated Kernel Profiles (Month 22)

Success Criteria:

1. RLS correlation $r > 0.50$ with observed outcomes
2. Quick vs. Full agreement ≥80%

3. User satisfaction ≥70%

Resource Requirements: 112 FTE-months, $3.0M

M.4 Phase 4: Scaled Implementation (Months 18-36)

Objectives:

1. Full deployment in diverse contexts
2. Longitudinal UCI trajectory tracking
3. Comparative analysis against alternatives
4. Goodhart resistance validation

Deliverables:

1. Deployment Toolkit (Month 24)
2. Training Program (Month 26)
3. Longitudinal Database (Month 28)
4. Comparative Effectiveness Studies (Month 32)

Success Criteria:

1. UCI improvement ≥0.10 versus baseline after 3 years
2. Demonstrated gaming resistance in red-team testing
3. Positive comparative effectiveness

Resource Requirements: 180 FTE-months, $6.5M

M.5 Resource Summary

Phase	Duration	FTE-Months	Budget
Phase 1: Verification	6 months	54	$600K
Phase 2: Validation	12 months	84	$1.5M
Phase 3: Pilots	15 months	112	$3.0M

Phase	Duration	FTE-Months	Budget
Phase 4: Scale	18 months	180	$6.5M
Total	**36 months**	**430**	**$11.6M**

Appendix N: Quick-Start Guide for Practitioners

N.1 Tier 1: The Three-Question Heuristic

Question 1: Rights Check Does this action obviously violate someone's core rights (life, bodily integrity, liberty, basic needs, dignity, due process, information access, ecological integrity)? Would it disproportionately harm any vulnerable group?

If YES: Do not proceed, or treat as emergency. If NO: Continue.

Question 2: Tail-Risk Check Does this action create a non-trivial chance of catastrophic or irreversible harm to many people, humanity, or the biosphere?

If YES: Reconsider or seek lower-risk alternatives. If NO: Continue.

N.2 Tier 1 Optional Worksheet: Compact Matrix (3×3 or 4×4)

1) Identify Key Unions (3–4 most affected).

1. 2) Identify Key Dimensions (3–4 most relevant).
2. 3) Estimate Impacts using a coarse scale (++, +, 0, −, −−).
3. 4) Rights screen: treat any "−−" in a clearly rights-relevant cell as a STOP requiring Tier 2+ review.
4. 5) Tail-risk screen: if any option plausibly increases catastrophic risk, escalate to Tier 2+ TRC.
5. 6) Select the best overall pattern among remaining options.
6. 7) Document briefly (what you assumed, what you did not know, and what would change your mind).

N.3 Tier 2: Core, Calculable (7×7)

1) Use the full 7 unions × 7 dimensions matrix for each option.

2) Enter direct impacts Ī^dir(a) in [−1,1] with a one-line justification per material cell.

3) Apply NCRC admissibility using the canonical rights coverage sets and worst-off subgroup rule.

4) Apply TRC using the declared scenario set and computation mode.

5) Apply containment (default Tier 2 containment gate unless PCC declares otherwise).

6) Compute RLS using default weights or declared HDW weights.

7) Select the highest-RLS option among admissible options and record the PCC-minimum fields.

7.

N.4 Default Applicability Patterns

Decision Type	Default Active Unions	Default Active Dimensions
Personal health	Self, Household	Health, Material, Agency
Household budget	Self, Household, Community	Material, Health, Social
Organization hiring	Self, Household, Organization, Community	Material, Health, Agency, Social
Local infrastructure	Community, Polity, Biosphere	Material, Health, Environment
National economic policy	Polity, Humanity, Biosphere	Material, Health, Social, Environment
AI deployment	Organization, Polity, Humanity, Biosphere	Agency, Knowledge, Health, Social, Environment

N.5 Quick Reference Cards

NCRC Quick Check:

Right	Threshold	Key Question
LIFE	−0.90	Does this risk death?
BODY	−0.70	Does this risk serious harm?
LBTY	−0.65	Does this restrict freedom unfairly?
NEED	−0.50	Does this threaten basic survival?
DIGN	−0.55	Does this dehumanize or exclude?
PROC	−0.45	Is this procedurally fair?
INFO	−0.40	Does this prevent informed choice?
ECOL	−0.65	Does this harm Earth systems?

Cascade Summary:

1. **NCRC:** Rights floors (worst-off subgroups) → Fail = Excluded
2. **TRC:** Tail risk corridor (CVaR ≤ threshold) → Fail = Excluded
3. **Containment:** ΔUCI ≥ tolerance → Fail = Veto/Escalate
4. **RLS:** Weighted welfare ranking → Gap > δ = Select top
5. **UCI/HOI:** Structural tie-break

Appendix O: MathGov vs. Alternatives Comparison Table

Feature	Utilitarianism	Deontology	Rawlsian Justice	Capability Approach	Constitutional AI	MathGov
Rights Protection	No (aggregative)	Yes (categorical)	Partial (priority of liberty)	Partial (threshold)	Partial (trained)	Non-compensatory (NCRC)
Tail-Risk Handling	Expected utility	Not explicit	Maximin	Implicit	Implicit	CVaR constraint (TRC)
Multi-Scale	Single aggregate	Individual duties	Basic structure	Individual	Not explicit	7 nested unions
Multi-Dimensional	Single utility	Not explicit	Primary goods	Multiple capabilities	Not explicit	7 dimensions
Ripple Effects	Implicit	Not explicit	Not explicit	Not explicit	Not explicit	Explicit kernel *K*
Democratic Legitimacy	Preference aggregation	Not explicit	Veil of ignorance	Deliberation	Human feedback	HDW with floors

Feature	Utilitarianism	Deontology	Rawlsian Justice	Capability Approach	Constitutional AI	MathGov
AI Integration	Reward maximization	Rule-following	Not designed	Not designed	Training-based	Constraint-first
Auditability	Depends	Depends	Not operational	Not operational	Limited	Full PCC
Learning	Not explicit	Not explicit	Not explicit	Not explicit	RLHF	NCAR loop
Gaming Resistance	Low	Medium	Medium	Medium	Medium	High (UCI/HOI + containment)
Containment	Not explicit	Not explicit	Not explicit	Not explicit	Not explicit	Explicit principle
Empty Set Handling	Choose "least bad"	Dilemma	Not addressed	Not addressed	Not addressed	Emergency /Fallback modes

---

Appendix P: Emergency Ethics and Catastrophic Unions

P.1 Purpose and Scope

This appendix specifies when and how the ordinary cascade can authorize exceptional measures on unions whose continued operation threatens catastrophe.

P.2 Catastrophic Union Classification

Non-Adversarial Catastrophic Union (NCU):

1. Core objectives do not explicitly involve harming others
2. Catastrophic risk arises from reckless design or inadequate safety
3. Would accept constraints if it understood risks
4. Ethical stance: Remedial

Adversarial Catastrophic Union (ACU):

1. Core objectives include domination, eradication, or permanent subjugation
2. Treats large-scale rights violations as instrumental or desirable
3. Would resist constraints because they conflict with objectives
4. Ethical stance: Defensive

P.3 Emergency Ethics Principles

Lexicographic Priority: Combined rights floor and survival conditions of multiple unions have priority over continued unconstrained operation of catastrophic union.

Minimality and Proportionality: Emergency measures must be least intrusive interventions that reliably return system within TRC bounds.

High Evidential Threshold: Classification requires convergent evidence, independent audits, sensitivity analysis.

Temporal Limitation: Emergency status is temporary with predefined sunset clauses and regular review.

Transparency: Rationale must be recorded in PCC-style artifacts and available for review.

P.4 Decision Procedure

1. **Pre-Screening:** Identify unions whose actions push toward catastrophe corridor
2. **Pathology Assessment:** Determine if harms are correctable or structural
3. **CU Classification:** Classify as NCU or ACU based on objective analysis

4. **Emergency Option Generation:** Construct options including reforms, throttling, neutralization
5. **Constraint Filtering:** Apply NCRC/TRC to emergency options themselves
6. **Implementation with NCAR:** Execute with monitoring and periodic review

P.5 Safeguards Against Abuse

1. Independent classification body (5+ members from diverse unions)
2. Burden of proof on classifiers
3. Appeals process with expedited timeline
4. Sunset clauses (24 months NCU, 12 months ACU maximum)
5. Public registry of all classifications
6. Whistleblower protections

Appendix Q: Minimal Reference Configuration v1 (MRC-v1)

Tier scope warning (normative). MRC-v1 is a Tier-2 starter configuration only. It MAY be used for learning and early pilots, but it MUST NOT be used to claim Tier 3 or Tier 4 compliance. In particular, MRC-v1's bounded_impact TRC parameters are invalid for Tier-4 admissibility, which requires raw_indicator TRC (AF-BASE) and $|S| \geq 50$.

Q.1 Purpose

MRC-v1 enables Tier-2 pilots without requiring packaged charter body. It provides stable baseline for falsification testing.

Q.2 Weight Configuration

Tier note (readability vs Tier-4 canonicalization). Appendix Q presents Tier-2 starter defaults (weights and scenario probabilities) in decimal for readability. For Tier-4 Pilot-Executable claims, the canonical, hash-bound registries MUST be taken from the ProofPack bundle and MUST obey the Tier-4 NO_FLOATS policy (exact rationals as {"num": int, "den": int>0}). Decimals in this Appendices volume are illustrative and MUST NOT be used for Tier-4 hashing.

Union Weights (from floors + equal above-floor distribution):

wU (sum=1): U1 Self: 0.2486; U2 Household: 0.1086; U3 Community: 0.1086; U4 Organization: 0.1086; U5 Polity: 0.1286; U6 Humanity/CMIU: 0.1486; U7 Biosphere: 0.1486

Union	Floor	Total Weight
Self	0.20	0.2486
Household	0.06	0.1086
Community	0.06	0.1086
Organization	0.06	0.1086
Polity	0.08	0.1286
Humanity/CMIU	0.10	0.1486
Biosphere	0.10	0.1486

Dimension Weights:

wD (sum=1): D1 Material: 0.1371; D2 Health: 0.1571; D3 Social: 0.1371; D4 Knowledge: 0.1371; D5 Agency: 0.1571; D6 Meaning: 0.1171; D7 Environment: 0.1571 For docs-only evaluation and Tier ≤ 2 training pilots, see the embedded DSL-20-TRAINING-V1 scenario library in Appendix D.7.

Dimension	Floor	Total Weight
Material	0.08	0.1371
Health	0.10	0.1571
Social	0.08	0.1371
Knowledge	0.08	0.1371

Dimension	Floor	Total Weight
Agency	0.10	0.1571
Meaning	0.06	0.1171
Environment	0.10	0.1571

Scenario set size: n = |S| (registry-defined). For the bundled Tier-4 set REG-SCENARIOS-T4-v1, n = 60. Probability policy: all p_s are exact rationals with p_s > 0 and Σ_s p_s = 1. Tail confidence level (α): use the Default Corridor Parameters by Context (e.g., Organizational α = 0.95). TRC corridor threshold (τ_TRC): use the Default Corridor Parameters by Context (e.g., Organizational τ_TRC = 0.20).

Note (scope). These MRC-v1 TRC parameters are a Tier-2 starter for reversible-policy style pilots (they match Appendix AC). Tier-4 TRC admissibility MUST use trc_mode='raw_indicator' with AF-BASE (Appendix AF) and per §4.4.6 and AIL10.

Q.4 Scenario Library (externalized to ProofPack)

The Tier-2 scenario library (scenario IDs, stress parameters, and probabilities) is an executable artifact and is intentionally externalized to MathGov ProofPack v1.0. PCC runs MUST reference the scenario library by retrieval reference and hash.

Appendix R: PCC Example Instances

Non-normative (reserved). This appendix previously referenced complete worked PCC examples located elsewhere. To keep the spec package self-contained, the normative requirements do not depend on Appendix R. For a runnable example bundle, see Appendix AC. For pilots, create local PCC example instances and publish them as versioned artifacts in the decision log.

Appendix S: Starter Kernel Artifact Package (KOPS)

S.1 Purpose (normative for Tier 4 execution; evidence quality varies)

This appendix provides a minimal, runnable Kernel Operational Pathways Set (KOPS) so Tier 4 pilots can execute the kernel step without inventing entries. Entries here are intentionally conservative and MUST be treated as a starter library: they enable computation, not certification.

S.2 Artifact forms (required)

(a) Edge list (default): a sparse list of non-zero kernel entries K_{ij} with explicit source cell j and target cell i (target-row, source-column), aligned with the canonical convention in Section 8. (b) Matrix form (derived): a 49x49 sparse matrix assembled from the edge list. Unlisted entries are 0 by default. Implementations SHOULD store and publish both representations for audit.

S.3 Provenance and confidence classes (required)

Each KOPS entry MUST carry a provenance class and confidence class. This spec uses:

Class A (strong): high-quality quantitative synthesis applicable to the domain; coefficient may be used as stated.

Class B (moderate): decent evidence but imperfect transfer; apply the Tier-4 conservative multiplier (Section S.6).

Class C (weak): limited empirical support; apply the Tier-4 conservative multiplier (Section S.6) and treat as exploratory.

Class E (elicited): structured expert elicitation or domain-general systems reasoning; permitted for pilots but NOT sufficient for Tier-4 certification.

S.4 Edge list schema (required fields)

Kernel cell validity rule (normative). A kernel edge is INVALID if either endpoint (source or target) refers to a union–dimension cell that lacks a declared semantic definition and indicator family (or is masked as non-meaningful by default without explicit PCC activation). Invalid edges MUST be removed or corrected before use in any Tier ≥ 2 run.

Each edge MUST specify: (1) Edge ID, (2) source union u_s, source dimension d_s, (3) target union u_t, target dimension d_t, (4) coefficient k in [-1, +1], interpreted as marginal ripple from source cell into target cell, (5) sign meaning, (6) provenance/confidence class, (7) evidence note (short), (8) last review date, and (9) owner/reviewer.

S.5 Starter edge list (Tier 4 pilot default)

Table S-1 provides a conservative starter edge list. It is intentionally sparse. Teams SHOULD extend it only with versioned additions and evidence notes.

Edge ID	Source union	Source dim	Target union	Target dim	k (K_ij)	Class	Evidence note	Last review	Owner
S-EDGE-001	Biosphere	Environment	Self	Health	0.2	E	Cleaner air/water reduces morbidity; domain-general.	2025-12-20	Pilot team
S-EDGE-002	Biosphere	Environment	Household	Material	0.1	E	Ecosystem services affect livelihoods and costs.	2025-12-20	Pilot team
S-EDGE-003	Biosphere	Environment	Community	Health	0.15	E	Local environmental quality affects community health.	2025-12-20	Pilot team
S-EDGE-005	Polity	Agency	Self	Agency	0.15	E	Rights/protections enable personal agency.	2025-12-20	Pilot team
S-EDGE-006	Polity	Social	Community	Social	0.15	E	Institutional trust and safety shape community cohesion.	2025-12-20	Pilot team
S-EDGE-007	Organization	Material	Household	Material	0.2	E	Wages and employment affect household resources.	2025-12-20	Pilot team
S-EDGE-008	Organization	Health	Self	Health	0.05	E	Work conditions influence injury/illness risk.	2025-12-20	Pilot team
S-EDGE-009	Organization	Social	Household	Social	0.1	E	Job stability affects household stress and relationships.	2025-12-20	Pilot team

S-EDGE-010	Community	Social	Self	Health	0.15	E	Social support and safety affect health.	2025-12-20	Pilot team
S-EDGE-011	Community	Social	Self	Meaning	0.1	E	Belonging and cohesion influence meaning.	2025-12-20	Pilot team
S-EDGE-012	Household	Social	Self	Meaning	0.15	E	Household support influences meaning and resilience.	2025-12-20	Pilot team
S-EDGE-013	Household	Material	Self	Health	0.1	E	Basic resources enable healthcare and nutrition.	2025-12-20	Pilot team
S-EDGE-014	Self	Knowledge	Organization	Knowledge	0.05	E	Skills and learning diffuse into organizations.	2025-12-20	Pilot team
S-EDGE-015	Self	Health	Organization	Material	0.05	E	Workforce health affects productivity.	2025-12-20	Pilot team
S-EDGE-016	Humanity/CMIU	Environment	Biosphere	Environment	0.1	E	Collective stewardship affects biosphere integrity.	2025-12-20	Pilot team
S-EDGE-017	Polity	Material	Organization	Material	0.1	E	Taxes/regulation/public investment affect firms.	2025-12-20	Pilot team
S-EDGE-018	Organization	Material	Biosphere	Environment	-0.1	E	Production can degrade environment without safeguards.	2025-12-20	Pilot team
S-ED	Polity	Material	Biosphere	Environment	-0.05	E	Growth-focused policy can increase	2025-	Pilot

GE-019							ecological pressure.	12-20	team
S-EDGE-020	Self	Material	Biosphere	Environment	-0.05	E	Consumption patterns can increase ecological load.	2025-12-20	Pilot team
S-EDGE-021	Community	Environment	Self	Health	0.1	E	Local green space and exposure affect health.	2025-12-20	Pilot team
S-EDGE-022	Polity	Environment	Biosphere	Environment	0.1	E	Environmental policy can protect ecosystems.	2025-12-20	Pilot team
S-EDGE-023	Community	Knowledge	Self	Knowledge	0.1	E	Community learning resources raise knowledge.	2025-12-20	Pilot team
S-EDGE-024	Organization	Agency	Self	Agency	0.1	E	Work autonomy affects personal agency.	2025-12-20	Pilot team
S-EDGE-025	Community	Agency	Self	Agency	0.1	E	Local participation opportunities affect agency.	2025-12-20	Pilot team

Stability scaling (mandatory): after assembling K, compute i…is optional; for Full mode, it is mandatory (Section 8.3.3). Tier 4 Pilot-Executable rev14.x prohibition: Full mode MUST NOT be used; Quick only.

Conservative adjustment (default for pilots): implementations SHOULD apply a global shrink factor to all non-zero K entries to reduce over-claiming. Recommended default: multiply all coefficients by 0.7, then apply an additional class multiplier: A:1.00, B:0.70, C:0.50, E:0.35.

Stability scaling (mandatory): after assembling K, compute its spectral radius ρ(K). If ρ(K) >= 0.90, scale K by s = 0.90/ρ(K) before use, and record s in the PCC. For Quick mode only, this check is optional; for Full mode, it is mandatory (Section 8.3.3).

Evidence log and review (mandatory): for each edge used in a decision, record the edge ID, coefficient after scaling, class, and provenance note in the PCC. After at least 50 decisions with outcome tracking, compute calibration error and hit-rate by edge and update classes accordingly.

Tier-4 certification gate (normative): Tier-4 certification requires that all non-zero kernel entries relied upon for admissibility or ranking are at least Class B, with documented domain transfer justification. Class E is permitted for pilots but cannot be used to claim certification.

Table S-2 provides an evidence log template (teams may copy into their PCC bundle).

Edge ID	Decision ID	k_used	Class	Provenance	Outcome measure	Calibration error	Reviewer/date

Target	Source	Coefficient	Class	Primary Source
Self-Health	Self-Material	+0.25	A	Marmot Review
Self-Health	Community-Social	+0.15	A	Social determinants literature
Community-Material	Organization-Material	+0.20	A	Leontief I-O tables
Self-Agency	Self-Knowledge	+0.20	A	Education-empowerment literature
Humanity-Health	Humanity-Environment	+0.20	A	Lancet Planetary Health
Community-Environment	Organization-Material	−0.15	A	Growth-pollution relationship

Target	Source	Coefficient	Class	Primary Source
Self-Social	Household-Social	+0.25	A	Attachment theory
Humanity-Environment	Biosphere-Environment	+0.25	A	Earth system science

Appendix T: Rights Threshold Calibration Protocol

T.1 Calibration Framework

Each threshold is calibrated through:

1. **Normative Anchor Identification:** Real-world harm category and normative source
2. **Indicator Mapping:** Reference indicator, reference class, percentile anchoring
3. **Philosophical Justification:** Convergence across human rights jurisprudence, humanitarian standards, capability theory

T.2 Threshold Derivations

LIFE (−0.90): Near-certain or highly probable death. Calibrated against UNHCR emergency mortality thresholds.

BODY (−0.70): Severe but non-fatal harm. Calibrated against Sphere Standards minimum thresholds.

LBTY (−0.65): Significant freedom restrictions. Calibrated against Freedom House "partly free" threshold.

NEED (−0.50): Severe deprivation of basic necessities. Calibrated against FAO FIES severe thresholds.

DIGN (−0.55): Systematic humiliation or dehumanization. Calibrated against UDHR dignity provisions.

PROC (−0.45): Procedural justice violations. Calibrated against World Justice Project Rule of Law Index.

INFO (−0.40): Systematic information denial. Calibrated against press freedom indices.

ECOL (−0.65): Earth-system integrity violations. Calibrated against planetary boundary framework.

T.2.1 Invariant Rights Anchor Registry (Tier 4 starter; non-placeholder)

Tier 4 requirement (normative). Tier 4 execution requires explicit, non-placeholder anchors for every right asserted as active in the decision. These anchors make the NCRC computable without invention. Evidence quality may be weak at the start, but numbers and mappings MUST be explicit and governed.

Purpose. This registry supplies a minimal default set of indicator anchors that map real-world measures into the normalized impact scale I in [-1, +1], where +1 is best and -1 is worst. Rights thresholds in T.2 are evaluated against these mapped impacts.

Registry schema (required fields). Each entry MUST specify: (1) Anchor ID, (2) right code r, (3) indicator definition, (4) unit and reference class, (5) direction (higher worse or higher better), (6) x_good and x_bad anchors, (7) mapping function, (8) evidence/provenance class, and (9) source type.

Canonical mapping rules (default). Let x be the indicator value, with anchors x_good (best) and x_bad (worst). Define:

Case A (higher is worse): I(x) = clip(1 - 2*(x - x_good)/(x_bad - x_good) , -1 , +1)

Case B (higher is better): I(x) = clip(2*(x - x_bad)/(x_good - x_bad) - 1 , -1 , +1)

Implementations MUST record which case is used for each indicator in the PCC. If x_good = x_bad, the anchor is invalid and MUST be replaced before use.

Starter anchors (Tier 4 pilot default). Table T-1 provides a minimal set of non-placeholder anchors. Teams SHOULD adapt x_good/x_bad to the domain and population, but MUST do so explicitly (versioned) and MUST justify changes in the PCC.

Anchor ID	Right code	Indicator (starter)	Unit	Direction	x_good	x_bad	Notes / reference class	Class	Source type
LIFE-A1	LIFE	Excess deaths per 1,000 per year (mortality shock over baseline)	deaths/1,000/year	higher_worse	0.0	20.0	Reference class: affected population in decision scope. Use local baseline	C	Domain data / official stats

							where available.		
BODY-A1	BODY	Severe injury incidence per 10,000 per year (medically severe harm)	injuries/10,000/year	higher_worse	0	200	Reference class: affected population. Prefer health surveillance data.	C	Domain data / health records
LBTY-A1	LBTY	Arbitrary detention or coercive constraint prevalence per 10,000	persons/10,000	higher_worse	0	200	Reference class: affected stakeholders. Use audited legal/admin records where possible.	E	Audit / elicitation
NEED-A1	NEED	Severe food insecurity prevalence (FAO FIES severe)	share (0-1)	higher_worse	0.00	0.40	Reference class: affected population. If using other deprivation indicators, declare mapping.	C	Survey / humanitarian data
DIGN-A1	DIGN	Validated discrimination or	share (0-1)	higher_worse	0.00	0.30	Use a validated instrum	E	Survey / rubric

		exclusion prevalence (proxy)					ent where possible. If proxy is qualitative, document rubric.		
PROC-A1	PROC	Due-process denial prevalence (cases lacking minimal safeguards)	share (0-1)	higher_worse	0.00	0.50	Reference class: relevant governed cases. Record what counts as 'minimal safeguards'.	E	Audit / legal review
INFO-A1	INFO	Systematic information denial prevalence (lack of basic access)	share (0-1)	higher_worse	0.00	0.50	Define 'basic access' for the context (civic info, safety info, or essential services).	E	Audit / survey
ECOL-A1	ECOL	Ecosystem integrity loss fraction (habitat quality or biodiversity proxy)	fraction lost (0-1)	higher_worse	0.00	0.60	Reference class: affected ecosystem region. Use standard	C	Ecological monitoring

							ecological indices when available.		

Registry governance note (required). Any anchor used in a decision MUST be listed in the PCC with its x_good, x_bad, mapping case, and source. Pilot defaults above are permitted for computation. Claiming Tier-4 certification requires upgraded evidence for each active right anchor (at least Class B) and independent review of indicator validity for the domain.

T.3 Sensitivity Analysis Requirement

Before adoption:

1. Vary each threshold by ±0.05
2. Apply to 20+ decision scenarios
3. If >30% of decisions change admissibility, provide additional justification

T.4 Revision Procedure

Charter-level governance required:

1. Proposal with documented justification
2. Sensitivity analysis
3. Supermajority vote (2/3)
4. Independent review panel
5. Public disclosure and version increment

Appendix U: Computational Complexity and Implementation Notes

U.1 Complexity Analysis by Tier

Tier	Operations	Typical Time	Hardware
1	O(1)	Negligible	Mental / notes

Tier	Operations	Typical Time	Hardware
2	$O(49^2 \times \|S\|)$	<500ms	Standard laptop / spreadsheet
3	$O(49^2 \times \|S\|$ + subgroup + σ-rules)	<2s	Standard laptop
4	$O(49^3$ + sensitivity + robustness)	<60s	Standard workstation

U.2 Critical Operations

Kernel Propagation:

1. Quick mode: $O(49^2)$ for matrix-vector multiply
2. Full mode (Certified profile only): $O(49^3)$ for matrix inversion (pre-computable). Tier-4 Pilot-Executable (rev14.x) forbids FULL propagation, so Pilot runtime excludes any matrix inversion costs.

CVaR Computation:

1. Sorting: $O(|S| \log |S|)$
2. Tail integration: $O(|S|)$

Sensitivity Analysis:

1. Per perturbation: Full cascade cost
2. Typical: 10-50 perturbations

U.3 Implementation Recommendations

1. Pre-compute and cache kernel inverses
2. Use sparse matrix representations
3. Parallelize scenario evaluation
4. Store intermediate values for audit

Appendix AF: Catastrophe Indicator Registry and Mapping Library

AF.1 Purpose (normative for Tier 4 execution)

This appendix provides a minimal catastrophe-indicator registry so the Tier 4 Tail-Risk Constraint (TRC) can be computed from raw indicators (Section 7.2.6) without inventing mappings. The defaults here are starter values for pilots. They are explicit and auditable, but they are not claimed to be fully validated for every domain.

AF.2 Registry schema (required fields)

Each indicator entry MUST specify: (1) Indicator ID, (2) mapped catastrophe cell (u,d) in C_cat, (3) indicator definition, unit, and reference class, (4) direction (higher worse or higher better), (5) x_onset and x_max anchors, (6) mapping function to loss in [0,1], (7) evidence/provenance class, and (8) source type.

AF.3 Canonical mapping function (default)

Let x be the raw indicator value with anchors x_onset (onset of catastrophe) and x_max (extreme catastrophe). Define the raw loss:

Case A (higher is worse): loss_raw(x) = clip((x - x_onset)/(x_max - x_onset) , 0 , 1)

Case B (higher is better): loss_raw(x) = clip((x_onset - x)/(x_onset - x_max) , 0 , 1)

Implementations MUST record which case is used in the PCC. If x_onset = x_max, the indicator is invalid and MUST be replaced before use.

AF.4 Base indicator set (AF-BASE; Tier 4 pilot default)

Machine-readable binding (Tier 4). The AF-BASE table in this appendix is published as the registry file registries/AF_BASE_V1.json inside the Tier-4 ProofPack. For Tier ≥ 4, implementations MUST use the bound registry (by ProofPack hash) for TRC catastrophe evaluation, unless the PCC declares a governed extension set C_ext (AF-EXT).

Table AF-1 provides a minimal registry. Teams MAY substitute indicators if they document equivalence, thresholds, and mapping in the PCC.

AF-BASE Multi-Indicator Aggregation (Normative default). When multiple raw indicators map to the same catastrophe cell (e.g., two indicators for Biosphere-Environment), the default aggregation for that cell is worst-case across indicators: If a PCC declares an extended catastrophe cell set C_ext, then for any cell in C_ext the indicator set is the union of AF-BASE and any declared AF-EXT indicators for that cell, and aggregation remains worst-case (max) across the full unioned indicator set.

Precedence for extended sets: if C_ext is declared, treat the indicator set for each catastrophe cell as Indicators(cell) := AF-BASE(cell) ∪ AF-EXT(cell) (if any). Aggregate with the same worst-case rule (max over Indicators(cell)), unless a more restrictive rule is explicitly declared in the PCC.

$$H\_cell(a,s) = \max\_j h\_j(a,s)$$

where h_j(a,s) is the mapped raw-indicator harm for indicator j in scenario s.

An alternative aggregation (e.g., mean) MAY be used only if the PCC declares the rationale and demonstrates it does not understate tail severity (including a worst-case comparison sensitivity run).

Indicator ID	Mapped cell (u-d)	Indicator definition	Unit	Direction	x_onset	x_max	Notes / reference class	Class	Source type
AF-CAT-001	Humanity/CMIU-Health	Excess mortality rate above baseline (deaths per 1,000 population per year)	deaths/1,000/year	higher_worse	1	100	Use local baseline when available.	C	Official stats / health data
AF-CAT-002	Humanity/CMIU-Environment	Severe habitability stress share of population (composite:	share (0-1)	higher_worse	0.10	0.50	Composite must be defined and frozen per pilot.	E	Composite / elicitation

		water, food, heat)							
AF-CAT-006	Biosphere - Environment	Planetary boundary transgression count (out of 9)	count (0-9)	higher_worse	4	7	Use current boundary framework; record version.	C	Scientific assessment
AF-CAT-007	Biosphere - Environment	Ecosystem integrity loss fraction (habitat quality or biodiversity proxy)	fraction lost (0-1)	higher_worse	0.10	0.30	Define proxy and region; prefer standard indices when available.	C	Ecological monitoring

AF.4B Extension indicator set (AF-EXT; optional; requires PCC-declared C_ext)

Normative rule. AF-EXT indicators MUST NOT be used for Tier-4 TRC unless the PCC explicitly declares an extended catastrophe cell set C_ext, and records the governance justification for that extension. AF-EXT is provided as an optional library for domains where additional catastrophe-bearing cells are warranted.

Indicator ID	Mapped cell (u-d)	Indicator definition	Unit	Direction	x_onset	x_max	Notes / reference class	Class	Source type
AF-CAT-003	Humanity /CMIU-Material	Severe food insecurity prevalence (FIES	share (0-1)	higher_worse	0.20	0.60	If using other deprivation indicators,	C	Survey / humanitarian data

		severe or equivalent)					declare mapping.		
AF-CAT-004	Polity-Social	Conflict fatalities rate (violent deaths per 100,000 per year)	deaths/100k/year	higher_worse	25	500	Use best available conflict dataset for scope.	E	Conflict monitoring / elicitation
AF-CAT-005	Polity-Material	Critical infrastructure outage burden (person-hours without essential service per capita per year)	hours/person/year	higher_worse	72	720	Define 'essential service' (water, power, health care access).	E	Audit / ops data

AF.5 Governance and certification gate (required). Pilot defaults above are permitted for computation. Claiming Tier-4 certification requires upgraded evidence for each catastrophe indicator used (at least Class B), plus independent review of indicator validity and threshold selection for the domain. Any change to x_onset/x_max MUST trigger sensitivity analysis and a version increment.

Appendix V: Version History

V.1 Version 5.0 (Spec-Hardened) (Current)

Release date: 2025-12-22

Major revisions from v4.9.6:

• Tier propagation lock: Tier 3 is Quick-only; Full propagation requires Tier 4 escalation (authoritative in §4.4.5).

• TRC mode enforcement: Appendix B now hard-locks tier ≥ 4 to raw-indicator TRC (AF-BASE/AF-EXT) and prohibits bounded-impact TRC from determining admissibility in Tier 4.

• MRC-v1 completion: Appendix Q now specifies explicit union/dimension weights, TRC parameters, and a 12-scenario library with probabilities.

• Pilot packaging: Appendix AE now includes starter registry manifests and computed hashes for deterministic third-party reproduction.

• Spec hardening: restored the Canonical Impact Construction formulas (§3.2.7), added a single-source default policy by tier (§4.4.6), and re-scoped Appendix AC as MRE-T2 with Tier-4 Pilot-Executable moved to ProofPack (manifest-only bundle).

Scope note (Appendix AC). This micro-example is a Tier-2 demonstration only. Tier-4 Pilot-Executable runs MUST follow the Tier-4 requirements in the Foundation, including trc_mode = raw_indicator, |S| ≥ 50 scenario sets, and the Tier-4 subgroup enumeration policy. Any numbers in Appendix AC are illustrative and are not normative defaults.

• Finalization metadata: version strings and end-of-doc metadata updated for v5.0; document hash computed at build time.

V.1A Version 5.1 (Spec-Hardened Consistency + Anti-Gaming)

Release date: 2025-12-24

Major fixes and hardenings relative to v5.0c:

- Baseline-Zero Rule (semantic correctness). Canonical magnitude derivation now enforces "0 = no change from baseline" globally: level scoring may be computed, but impacts are always deltas from baseline (μ = score(x_a) − score(x_0) or direct change anchoring). This resolves the level-anchoring inconsistency and aligns Appendix AC with §4.1.1.

- AE.7 registry alignment (removes internal contradictions). Appendix Q.3 TRC parameters are updated to match AE.7 registries: REG-PARAMS-MRC-v1 remains a generic Tier-2 starter (bounded_impact; α=0.70; τ_TRC=0.55; |S|=12), while the policy MRE (Appendix AC) uses REG-PARAMS-MRE-POLICY-T2-v1 (Tier-2 MRE only; Tier-4 uses REG-PARAMS-T4-v1) (bounded_impact; $\alpha=0.95$; $\tau_{TRC}=0.15$). Narrative text implying MRC-v1 "matches Appendix AC" is removed.

- Tier-4 TRC default consistency. REG-PARAMS-T4-v1 τ_{TRC} is aligned to D.3.2 context defaults (Organizational $\tau_{TRC} = 0.20$) while preserving mode lock to raw_indicator and $|S| \geq 50$.

- Draft vs release gate for hash placeholders. AE.7 includes an explicit release gate: any REBUILD_ME manifest hashes indicate a draft artifact and cannot support Tier-4 Pilot-Executable claims until computed hashes are inserted and replay succeeds (AIL11).

- Registry precedence rule (audit robustness). Added AIL0 rule that registry values referenced by PCC hashes take precedence over narrative defaults elsewhere; any document/default conflict triggers audit_flag DOC_DEFAULT_CONFLICT and NCAR correction.

- Anti-gaming hardenings (Tier-4). Added Tier-4 tail-mass minimum requirement for TRC scenario probabilities, strengthened subgroup minimum coverage expectations for rights-covered cells, prohibited Mode B containment outputs from influencing selection (Mode A only for selection), and required kernel diff auditing in Tier-4 PCCs.

- Invariant rights anchoring. Added AIL14/AIL15 to enforce invariant rights anchor usage for Tier-4 NCRC, preventing meaning drift through reference class manipulation.

- Sentience weighting explicit. Added s_k (SGP weight) to direct aggregation equation for completeness.

- Genesis Protocol. Added Appendix AA-GEN for cold-start governance bootstrapping.

These changes raise internal consistency, reduce specification gaming vectors, and improve audit readiness without altering the core lexicographic cascade design.

V.2 Version 4.9.6

Release date: 2025-12-22

Major revisions from v4.9.5:

• Added authoritative Tier Requirements Matrix (§4.4.5) and clarified Tier-4 Pilot-Executable meaning.

• TRC coherence: tier-gated trc_mode, aligned TRC definition to raw-indicator execution for Tier 4, and split AF-BASE vs AF-EXT (Appendix AF).

• Kernel sanity: removed semantically invalid starter edge S-EDGE-004 and added a normative kernel cell validity rule (Appendix S).

• Rights coherence: ECOL now maps to Environment only (Appendix C) to reduce redundancy and audit ambiguity.

• Added formal ValidPCC predicate + conformance test suite (Appendix G) and expanded Tier-4 pilot checklist (Appendix AE).

• Tier 4 spec-completeness: added executable starter artifacts (kernel edge list, rights anchor registry, catastrophe indicator registry) so pilots can run without inventing governed numbers (Appendices S, T.2.1, W).

• Self-contained packaging: replaced 'provided elsewhere' worked-example stubs with explicit non-normative placeholders and pointed to the in-document minimal reproducible example bundle (Appendices H, R, AC).

• Added a spec package manifest and pilot readiness checklist to make the Tier 4 minimum executable artifact set explicit (Appendix AE), and updated internal version cross-references to 4.9.3.

Appendix W: Acknowledgments and Methodology Notes

W.1 Development Methodology

MathGov follows design-science and normative-engineering methodology:

1. Problem identification through analysis of governance and AI alignment failures

2. Theoretical grounding from ethics, economics, systems theory, and risk management
3. Formal specification with mathematical definitions
4. Internal consistency verification
5. Failure mode analysis with explicit mitigations
6. Worked examples for end-to-end validation

W.2 AI Assistance Disclosure

Generative AI tools (OpenAI ChatGPT and Anthropic Claude) were used as writing and reasoning assistants. The author reviewed, verified, and edited all AI-generated content. Final responsibility for all claims rests with the human author.

W.3 Limitations

1. Empirical validation through pilots is planned but not yet completed
2. Starter KOPS requires organization-specific calibration
3. Cross-cultural invariance testing needs systematic completion
4. Full governance infrastructure does not yet exist at scale

---

Appendix X: Quick Reference Summary

Acronym Lookup

Acronym	Full Name
CMIU	Collective Managing Intelligence Union
CVaR	Conditional Value-at-Risk
HDW	Hybrid Democratic Weighting
HOI	Hollowing-Out Index

Acronym	Full Name
KOPS	Key Operational Pathways Set
MI	Managing Intelligence
MIT-4	Managing Intelligence Test (4 criteria)
MNA	Minimal Normative Axiom
MRC	Minimal Reference Configuration
MTS	Mandatory Tail Scenarios
NCAR	Notice-Choose-Act-Reflect
NCRC	Non-Compensatory Rights Constraint
PCC	Provenance and Compliance Certificate
RLS	Ripple Logic Score
SGP	Sentience Gradient Protocol
TRC	Tail-Risk Constraint
UBE	Union-Based Ethics
UBG	Union-Based Governance
UBL	Union-Based Living
UBR	Union-Based Reality

Acronym	Full Name
UCI	Union Coherence Index

Key Parameters (MRC-v1 Defaults)

Parameter	Symbol	Default
Direct saturation	β	2
Post-propagation saturation	β_{prop}	1
Discrimination threshold	δ	2
Containment tolerance	τ_c	−0.10
Positive-impact threshold	θ_{pos}	0.05
Containment depth	D_c	2
Reference horizon	T_{ref}	25 years
Sentience floor	g_{min}	0.05
Sentience curvature	ψ	0.5
Rights plateau threshold	$SG_{threshold}$	0.85

Appendix Y: Contact and Governance Information

Y.1 Framework Stewardship

MathGov is developed and maintained as an open framework.

Primary contact:

1. ORCID: 0009-0005-3324-7290

2. Website: mathgov.org

Institutional affiliation:

1. British University Vietnam (BUV)

Y.2 Licensing

Y.3 Citation

McGaughran, J. (2025). MathGov: A Universal Ethical Operating System for Multi-Scale Alignment (Version 4.9.3). mathgov.org

---

Appendix Z: Final Notes and Philosophical Reflection

Z.1 On the Limits of Formalization

MathGov translates ethical commitments into computable procedures. This brings transparency, auditability, and scalability—but no formal system captures the full richness of moral experience. The NCAR loop treats parameters as provisional. The Judgment Call mechanism acknowledges underdetermination. The goal is not a system answering all ethical questions but one making ethical reasoning more rigorous while remaining humble about incompleteness.

Z.2 On Rights and Welfare

MathGov treats rights as lexicographically prior to welfare optimization. This reflects the judgment that certain protections should not be tradeable for aggregate gains. The rights plateau for Managing Intelligences reflects commitment to non-hierarchical moral standing above threshold.

Z.3 On the Path Forward

MathGov is a scaffold, not a finished edifice. Its value will be measured by practical impact: fewer rights violations, fewer catastrophic surprises, more resilient unions, more transparent decisions. If the framework succeeds, it will be because it gives humans and future Managing Intelligences better tools for asking the right questions, making structured trade-offs, learning from mistakes, and honoring the dignity of all who share the network of unions.

Appendix AA: HDW Deliberation Protocol and Panel Construction

AA.1 Purpose

Tier-4 executable alignment (Normative). For Tier-4 runs, ballots, weights, scenario probabilities, kernel coefficients, and any other non-integer numbers MUST be represented in canonical artifacts as exact rationals {"num": int, "den": int>0} under the declared canonicalization profile, and floats/NaN/Infinity MUST be rejected. Hashes MUST be stored as sha256:<lowercasehex>. Note: any numeric literals printed elsewhere in this Appendices document are illustrative; Tier-4 hashing MUST use the ProofPack's canonical, hash-bound registries and manifests.

Tier-4 executable alignment (Normative). For Tier-4 runs, ballots MUST be strict-valid (exact rationals; sum exactly to 1; floors respected; no silent correction). Ballot finality MUST follow LATEST_WINS_THEN_BALLOT_HASH_MAX_V1. Aggregation MUST follow TRIM_WHOLE_BALLOTS_THEN_MEAN_V1 (whole-ballot trimming, then mean). Effective abstention includes explicit abstain, invalid not corrected by deadline, or no valid submission by deadline. Abstention rate exceeding 20% (per ballot type) triggers Tier-4 downgrade. Ballot hashing MUST use CANON_JSON_SORTED_KEYS_COMPACT_UTF8_V1 (normalization NONE; duplicate keys forbidden; NO_FLOATS and NaN/Infinity rejected; hashes stored as sha256:<lowercasehex>).

AA.2 Panel Construction

AA.4 Suggested workflow (human procedure; not a calculability requirement)
Minimum recommended panel size: 15 members; recommended 25+ for broad community decisions. Tier gate (Normative): For Tier-4 Pilot-Executable claims, ballot minima are governed by the Foundation Tier-4 HDW gate (Organization: n_valid_union ≥ 15 and n_valid_dimension ≥ 15; Polity: n_valid_union ≥ 25 and n_valid_dimension ≥ 25). Any run below these minima MUST downgrade its tier_claim (e.g., Tier-2 or Tier-3) and set audit_flag = TIER_CLAIM_DOWNGRADE_REQUIRED. A "5 completed ballots" minimum is Genesis/Tier-2 demonstrator guidance only and MUST NOT be used to claim Tier-4.

Stratification Requirements:

1. At least one representative from each union type affected

2. Vulnerable population representatives (including future generations proxy where feasible)
3. Independent risk experts (minimum 2)
4. Domain experts relevant to decision context

Selection Methods:

1. Stratified random sampling from stakeholder registry
2. Sortition with demographic balancing
3. Rotating membership to prevent capture

Conflict of Interest:

1. Disclosure required before participation
2. Recusal for material conflicts
3. Panel composition published

AA.3 Briefing Materials

Before deliberation, participants receive:

1. Decision context summary (plain language)
2. Option descriptions
3. Union and dimension definitions
4. Current floor values and their rationale
5. Baseline weight distribution
6. Historical weight decisions (if applicable)

Materials must be:

1. Accessible (reading level appropriate to population)
2. Balanced (present multiple perspectives)
3. Transparent about uncertainty

AA.4 Deliberation Protocol

Phase 1: Information (60-90 minutes)

1. Presentation of materials
2. Expert testimony (balanced perspectives)
3. Clarifying questions

Phase 2: Small Group Discussion (60-90 minutes)

1. Groups of 5-7 participants
2. Facilitated by neutral moderator
3. Focus on values, priorities, concerns

Phase 3: Plenary Synthesis (45-60 minutes)

1. Reports from small groups
2. Identification of common ground
3. Clarification of disagreements

Phase 4: Individual Reflection (15-30 minutes)

1. Private completion of ballots
2. Optional written rationale

AA.5 Balloting Format

Union Weight Ballot:

For each union *u*, allocate weight points (total must sum to 100):

Union	Minimum (Floor)	Your Allocation
Self	20	___
Household	6	___
Community	6	___
Organization	6	___

Union	Minimum (Floor)	Your Allocation
Polity	8	___
Humanity/CMIU	10	___
Biosphere	10	___
Total	**66**	**100**

Dimension Weight Ballot:

For each dimension d, allocate weight points (total must sum to 100):

Dimension	Minimum (Floor)	Your Allocation
Material	8	___
Health	10	___
Social	8	___
Knowledge	8	___
Agency	10	___
Meaning	6	___
Environment	10	___
Total	60	100

Instructions: Each ballot MUST sum to 100 points. If the PCC declares ballot anchors (e.g., minimum weight for a union or dimension), allocations MUST meet or exceed those anchors, and the remainder may be distributed according to relative importance. Genesis/Tier-2 demonstrator guidance: 5 completed ballots per ballot type MAY be used for early-stage testing, but MUST be disclosed and MUST NOT be used to claim Tier-4. For Tier-4 Pilot-Executable claims, the minimum valid ballot counts are governed by the Foundation Tier-4 HDW gate (Organization: ≥15/15; Polity: ≥25/25); runs below these minima MUST downgrade tier_claim and set audit_flag = TIER_CLAIM_DOWNGRADE_REQUIRED.

AA.6 Aggregation Rules

Default: Trimmed mean (remove top and bottom 10%, compute mean of remainder)

Alternatives (require PCC documentation):

1. Median (for high-variance distributions)
2. Borda count (for ranked preferences)
3. Deliberative consensus (for small panels with explicit agreement)

Disagreement Documentation:

1. Record full distribution of responses
2. Document minority positions
3. Flag high-variance items for review

AA.7 Anti-Capture Safeguards

1. **Supermajority Lock:** Weights within 0.02 of floor require 2/3 approval
2. **Coordination Detection:** Statistical tests for voting blocs
3. **Rotation:** Maximum 3 consecutive participation periods
4. **Transparency:** All results published with anonymized individual responses
5. **Appeal Process:** Minority groups may appeal with documented concerns

AA.8 PCC Logging

Record in PCC:

1. Panel composition (demographics, affiliations)
2. Briefing materials provided
3. Deliberation timeline and format
4. Aggregation method used
5. Final weight vectors
6. Disagreement distribution
7. Any safeguard triggers

Appendix AA-GEN (Normative): Genesis Protocol (Cold-Start Governance)

Purpose. Defines how a first MathGov governance deployment can begin without requiring pre-existing MathGov institutions.

Genesis constraints. Cold-start deployments MUST cap themselves at Tier ≤ 2 until: (i) hash-bound registries exist, (ii) subgroup policy is demonstrated, (iii) scenario library meets tier minima, and (iv) at least one independent PCC audit is completed.

Genesis steps (minimum viable):

1. Constitute a temporary Steward Board (SB-0), minimum 7 members: rights specialist, risk specialist, biosphere steward, community representative, organization representative, public-interest/policy representative, independent auditor.

2. Publish SB-0 charter + COI disclosures (hash-bound, immutable).

3. Publish starter registries (hash-bound): rights coverage, rights anchors, catastrophe indicators, scenario library, kernel, weights.

4. Execute ≥20 Tier-2 decisions producing PCCs; publish redacted PCCs.

5. Perform Genesis review: SB-0 either revises registries (new hashes) and transitions to SB-1 with HDW, or dissolves and transfers authority to an established legitimate body.

Exit condition. Tier-3 claims remain prohibited until the constraints above are met and recorded in a governance log.

[End of MathGov v5.0 Complete Specification]

Release binding (authoritative): see MathGov_Release_Hashes_rev14.27_FULL.txt for SHA-256 over the raw .docx bytes.

Total length: ≈30,379 words (includes normative appendices; excluding ProofPack artifacts).

Version: 5.0i (rev14.27)

Last updated: 2025-12-26

Appendix AB Addendum: Canonical function table (complete, normative)

This addendum is normative and supersedes any partial or truncated function table earlier in Appendix AB.

Function	Definition	Domain	Range	Defaults / Notes
$(x)_+$	max(x,0)	$x \in \mathbb{R}$	$\mathbb{R}_{\geq 0}$	Positive part
clip(x,a,b)	max(a, min(x,b))	$x \in \mathbb{R}$, $a \leq b$	[a,b]	Invalid if a>b
$\text{clamp}_{[0,1]}(x)$	max(0, min(x,1))	$x \in \mathbb{R}$	[0,1]	Use for normalized quantities
$\text{sat}_\beta(x)$	$\tanh(\beta x)$	$x \in \mathbb{R}$, $\beta>0$	(-1,1)	Default $\beta=2$; smooth saturation
$\phi(u,d)$	$7(u-1)+d$	$u,d \in \{1,\dots,7\}$	$\{1,\dots,49\}$	Row-major flattening for 7×7 welfare space
$\text{Anc}(u, D_c)$	Ancestor unions of u on union ladder	$u \in U$	$\mathcal{P}(U)$	PCC must declare union ladder ordering
$\tau(t)$	$\ln(1+t)/\ln(1+T_{ref})$	$t>0$, $T_{ref}>0$	$(0,\infty)$	Default $T_{ref}=25$ years; optional τ_{max} must be declared
$\text{EMA}_\lambda(x_t)$	$\lambda x_t+(1-\lambda)\text{EMA}_\lambda(x_{t-1})$	$\lambda \in (0,1]$	$\mathbb{R}$	Init $\text{EMA}(x_0)=x_0$ unless specified
$\|\|K\|\|_\infty$	$\max_i \Sigma_j \|K_{ij}\|$	$K \in \mathbb{R}^{49\times49}$	$\mathbb{R}_{\geq 0}$	Row-sum norm; used for stability bounds

Selectable(a)	Admissible(a)=true ∧ Containment(a)=true	a ∈ Options	{true,false}	Selection-eligible gate (default Mode A)
CVaR_α(L)	min_ζ [ζ + (1/(1−α))·E[(L−ζ)_+]]	α∈(0,1), L∈[0,1]	[0,1]	Rockafellar–Uryasev; discrete algorithm in §7
Admissible(a)	NCRC(a)=true ∧ TRC(a)=true	a ∈ Options	{true,false}	Deontic admissibility (rights + tail-risk)