Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Set-Theoretic Framework for Enterprise Risk Taxonomy Classification with Categorical Semantics

Marc Faria

marc.c.faria@gmail.com

Research Article

Keywords: risk taxonomy, classification, set theory, category theory, HHI, factorization, concentration ratio, risk driver, sub-risk

Posted Date: March 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9212565/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Enterprise risk taxonomies face a persistent classification challenge: deciding which candidate concepts belong in a vertical risk hierarchy and which operate as cross-cutting drivers that overlay many categories. This article presents a reproducible, auditable classification methodology that uses set-theoretic operations over attributed structures, while using categorical concepts as a semantic organizing language. The framework defines three quantitative tests—*Relevance* (Concentration Ratio), *Distinction* (factorization), and *Materiality* (Herfindahl–Hirschman Index)—and a simple decision rule to classify taxonomy candidate concepts as hierarchical sub-risks or cross-cutting drivers. We demonstrate the approach with a worked application to geopolitical risk and validate behavior on representative edge cases (model risk, climate risk). The method is computationally tractable for typical enterprise taxonomies and supports governance-ready documentation of classification decisions.

1. Introduction

Financial institutions and large corporations maintain hierarchical risk taxonomies—typically structured as L1 → L2 → L3 cascades—to organize risk identification, reporting, and capital aggregation. Yet persistent ambiguity arises when emerging phenomena such as climate risk, cyber risk, or geopolitical instability defy neat placement within a single vertical branch, instead propagating effects across multiple categories simultaneously.

Two failure modes recur in practice. First, governance bodies may proliferate spurious top-level (L1) categories for each emergent concern, fragmenting oversight and diluting accountability. Second, committees may force a genuinely cross-cutting driver into a single hierarchical slot, thereby obscuring its systemic reach and inviting double-counting when exposures are later aggregated.

Contemporary practice leans heavily on committee judgment—valuable, but increasingly insufficient as regulators demand replicable, auditable decision trails. The present work responds by advancing a practical algorithmic methodology that: (i) operationalizes a controlled attribute vocabulary; (ii) quantifies impact breadth and concentration; (iii) executes three objective tests yielding a classification recommendation; and (iv) preserves provenance for governance review.

Two design constraints guide the framework. First, implement-ability: enterprise risk teams should be able to deploy the method without recourse to exotic mathematical infrastructure or specialized software. Second, semantic transparency: the categorical language employed clarifies structural intent for auditors and knowledge architects, even as the underlying computations reduce to finite set and order operations.

The remainder of this article proceeds as follows. Section 2 surveys related literature. Section 3 develops the methodology, introducing structural components (Section 3.1), the attribute vocabulary and meet operation (Section 3.2), impact breadth and concentration measures (Section 3.3), the three classification tests (Section 3.4), and categorical semantics (Section 3.5). Section 4 presents a worked

application to geopolitical risk. Section 5 validates the framework against model risk and climate risk and reports sensitivity analyses. Section 6 addresses implementation considerations. Section 7 concludes.

2. Related Work

Risk taxonomy guidance (e.g., Basel, COSO, ISO 31000) provides descriptive frameworks but not objective classification rules. Mathematical and network approaches—graph-theoretic networks used for systemic risk, Bayesian dependency models or copulas for joint risk measurement—address dependence and quantification but not taxonomy membership decisions. Nevertheless, the body of literature offers no principled criterion for taxonomy membership decisions.

Category Theory has emerged as a language for compositional systems and knowledge representation (Spivak, Baez & Pollard), offering intuition about presheaves, profunctors and universal constructions. This article uses categorical terms as semantics to clarify the structural role of drivers and risks, while the implementation remains grounded in set-theoretic operations.

The Herfindahl–Hirschman Index (HHI) originates in industrial organization economics, where it gauges market concentration (Rhoades, 1993). U.S. DOJ/FTC Horizontal Merger Guidelines (2010) codify thresholds—HHI below 0.15 for unconcentrated markets, above 0.25 for highly concentrated ones—that we adapt as intuitive benchmarks rather than legal mandates. Caldara and Iacoviello (2022) demonstrate how geopolitical risk can be quantified through textual analysis of news flows, reinforcing the feasibility of operationalizing emergent risk concepts.

3. Methodology

3.1 Structural Components

· Risk set: denote the taxonomy as a graded poset R with level function H: R → N that assigns levels (L1, L2, L3).

· Driver set: denote cross-cutting candidates as D.

· Attributes: controlled vocabulary V and an attribution function A: R ∪ D → P(V) that assigns keywords to each element.

· Impact assignment: a bipartite weighted edge set E between D (or taxonomy candidate) and L1 risks, with weights recording impact strength; the tuple G = (R, D, E, w) is a tractable representation of the profunctor-like structure.

To ground these definitions concretely, **Table 1** enumerates the L1 risk categories employed throughout this article; **Figure 1** depicts the bipartite graph G for two illustrative drivers.

Table 1: Reference L1 Risk Categories

L1 RISK CATEGORY	DESCRIPTION (generic)
Credit Risk	Risk of loss arising from a borrower's failure to meet contractual obligations
Market Risk	Risk of losses in positions arising from movements in market prices
Operational Risk	Risk of loss resulting from inadequate or failed processes, people, systems, or external events
Strategic Risk	Risk arising from adverse business decisions, improper implementation of decisions, or lack of responsiveness to industry changes
Liquidity Risk	Risk that an entity will be unable to meet short-term financial demands
Reputational Risk	Risk of damage to an organization's standing resulting from negative public perception

Reference definitions for L1 risk categories employed in the classification framework. These

categories form the columns of the impact matrix, and the target set R in the bipartite

structure G = (R, D, E, w).

Figure 1: Bipartite Structure Illustration

A graph showing Driver set D on the left, Risk set R (L1 risks) on the right, with weighted edges representing impact assignments.

Bipartite structure G = (R, D, E, w). Each driver in D connects via weighted edges to multiple L1 risks in R. Edge weights w(c, r) represent the materiality of driver c's impact on L1 risk r.

3.1.1 Notation

The following notation is used throughout the methodology and examples:

· R: the set (poset) of risks in the taxonomy.

· H: level function H: R → N mapping each risk to its hierarchy level (e.g., H(r)=1 for L1 risks).

· D: the set of candidate drivers or cross-cutting concepts under consideration.

· V: controlled vocabulary of attribute terms (domains such as Scope, Nature, Trigger).

· A: attribution map A: R ∪ D → P(V); for x ∈ R ∪ D, A(x) ⊆ V is the attribute set assigned to x.

· E: set of impact edges between candidates (D) and L1 risks; an edge (c, r) ∈ E indicates c impacts L1 risk r.

· w: weight function w: E → R≥0 assigning non-negative impact strengths to edges; for convenience we write w(c,r) for the weight on edge (c,r).

· G: tuple G = (R, D, E, w) describing the bipartite weighted structure.

· c: a candidate concept whose classification is under test (c ∈ D or a proposed new risk).

· impact(c, r): impact specification of c on L1 risk r (encoded via E and w).

· s_i: normalized impact share of candidate c on L1 risk i, defined by s_i = w(c,i) / Σ_j w(c,j), so that Σ_i s_i = 1.

· BREADTH(c): number of L1 risks affected by c, i.e. |{ r ∈ L1 : impact(c, r) ≠ ∅ }|.

· CR(c): concentration ratio, CR(c) = max_i s_i.

· HHI(c): Herfindahl–Hirschman Index, HHI(c) = Σ_i s_i^2.

· R*: a putative parent risk produced by the attribute meet operation (see below).

· R_i: an indexed collection of risks used when computing meets or attribute intersections.

3.2 Attribute Vocabulary and Meet Operation

The vocabulary V is organized into three semantic domains:

Scope (internal, external, cross-cutting);

Nature (financial, operational, strategic, behavioral);

Trigger (market-driven, event-driven, process-driven, decision-driven).

Table 2: Controlled Vocabulary (V)

DOMAIN	KEYWORDS
Scope	internal, external, cross-cutting, enterprise-wide
Nature	financial, operational, strategic, behavioral
Trigger	market-driven, event-driven, process-driven, decision-driven

Controlled vocabulary V used for attribute assignment. Each risk or driver is assigned 2–3 keywords from this vocabulary.

Each risk or taxonomy candidate has an attribute set A(·) ⊆ V. We define an attribute meet operation to detect inheritance relationships.

For a collection of risks {R_i}, the attribute meet returns a candidate parent R* when the intersection of their attribute sets equals the parent's attribute set.

In notation:

meet_A({R_i}) = R*

where,

$A(R^*) = \bigcap_i A(R_i)$ and, for all i: $A(R_i) \supseteq A(R^*)$.

Here, R* denotes a hypothetical or constructed parent risk whose attributes equal the intersection of the children's attributes; R* may or may not be an explicit member of the original risk set.

A non-empty meet indicates the existence of a parent risk whose attributes subsume those of children and supports a sub-risk interpretation.

3.3 Impact Breadth and Concentration

We compute two simple measures to capture the impact of a candidate (c) influence:

· **Breadth**: $BREADTH(c) = |\{r \in L1 : impact(c, r) \neq \emptyset\}|$. A value greater than 1 signals cross-cutting influence.

· **Concentration and materiality**: using L1 impact weights normalized to shares s_i across L1 risks, define the concentration ratio and Herfindahl–Hirschman Index (HHI):

$CR(c) = \max_i s_i$, (concentration ratio)

$HHI(c) = \Sigma_i s_i^2$ (Herfindahl–Hirschman Index)

Economic intuition guides our methodology's threshold selection. A concentration ratio CR ≥ 0.50 implies that a single L1 category absorbs the majority of impact—characteristic of a sub-risk. Moreover, an HHI ≥ 0.25 indicates "moderately concentrated" materiality, (consistent with DOJ/FTC guidance ranges), used here as a heuristic rather than a statutory constraint.

3.4 Three Classification Tests

Each candidate is run through the following tests.

· **Relevance Test**: if CR(c) ≥ 0.50 then the candidate concentrates the majority of its impact on a single L1 and this supports classification as a sub-risk.

· **Distinction Test**: tests factorization — there exists $R^* \in R$ such that $A(c) \supseteq A(R^*)$ (equivalently, $A(R^*) \subseteq A(c)$) and impact patterns align; this indicates the candidate factors through an existing parent risk and is evidence of sub-risk status.

· **Materiality Test**: compute HHI(c) on impact shares; if HHI(c) ≥ 0.25 the candidate exhibits concentrated materiality and this supports sub-risk classification. Lower HHI values favor driver classification.

Table 3 codifies test thresholds and rationale. **Figure 2** presents the decision flowchart integrating these tests.

Table 3: Classification Test Thresholds and Rationale

TEST	METRIC	THRESHOLD	CLASSIFICATION IF MET	RATIONALE
Relevance	CR(c)	≥ 0.50	SUB-RISK	Majority-share principle (anti-trust economics)
Distinction	Factorization	Exists	SUB-RISK	Candidate Factors through existing parent
Materiality	HHI(c)	≥ 0.25	SUB-RISK	Moderately concentrated (DOJ/FTC guidance)

Quantitative thresholds for the three classification tests. A majority (2/3 or 3/3) determines the final classification recommendation.

Figure 2: Classification Decision Flowchart

A decision-tree diagram illustrating the two-stage classification algorithm.

Classification decision flowchart. Candidates first undergo an L1 qualification test (four conditions). Those failing proceed to attribute-meet and impact-breadth checks, then to the three quantitative tests (Relevance, Distinction, Materiality) that determine sub-risk vs. driver classification.

Decision rule: a majority-of-tests rule (2/3 or 3/3) determines the recommendation. Ties default to administrative assignment with documented rationale.

3.5 Semantic Framing

Categorically, the attribution function resembles a presheaf assigning data to objects in a risk category; the impact assignment resembles a profunctor mapping drivers to risk-weighted contexts. The attribute meet corresponds to a limit in the attribute lattice and factorization tests correspond to existence of representable factorings in the slice category over a parent risk.

This semantic framing helps auditors and knowledge architects reason about structure-such as, composition, inheritance, factorization- while the implementation remains set-theoretic.

4. Case Study: Geopolitical Risk

4.1 Motivation

Geopolitical Risk has become prominent since 2022 and has been proposed as a standalone L1 by some practitioners. We apply the methodology to show how a transparent, quantitative test set supports a defensible outcome.

4.2 Attribute Assignment and L1 Qualification

Assigning attributes per the vocabulary, we obtain:

· A(Geopolitical) = {external, event-driven, cross-cutting}.

L1 risk qualification rules require (i) sufficient attribute count, (ii) novelty relative to existing risk categories, (iii) absence of cross-cutting attribute flag, and (iv) evidence of containment (i.e., identifiable L2 risk children). *Geopolitical Risk* fails on novelty/containment and elicits the a cross-cutting attribute; it therefore does not qualify as a new L1 risk classification under the prescriptive L1 filter and proceeds to Driver testing.

4.3 Breadth and Quantitative Tests

Subject-matter experts provided the impact weight assessments summarized in the upper row of **Table 4**.

Table 4: Impact Weight Matrix

CANDIDATE	CREDIT	MARKET	OPERATIONAL	STRATEGIC	LIQUIDITY	REPUTATIONAL	TOTAL
Geopolitical	0.15	0.20	0.15	0.25	0.15	0.10	1.00
Model Risk	0.00	0.14	0.72	0.08	0.06	0.00	1.00
Climate Risk	0.35	0.30	0.20	0.15	0.00	0.00	1.00

Impact weight matrix for the three case-study candidates. Geopolitical Risk distributes relatively evenly across all six L1 risk categories (max weight 0.25), whereas Model Risk concentrates heavily on Operational (0.72). Climate Risk shows moderate concentration on Credit (0.35) but still spans multiple categories.

From the empirical inputs above, we yield:

· BREADTH = 6 (affects all six L1 categories in our exemplar taxonomy)

· CR = 0.25 (no single L1 receives majority share)

· HHI = 0.20 (dispersed materiality)

All three tests return **DRIVER** outcomes: the *Relevance* test fail the majority-share criterion; the *Distinction* test finds no factorization through an existing parent; the *Materiality* test registers dispersed concentration (HHI below threshold). The algorithm therefore classifies *Geopolitical Risk* as a **DRIVER** and assigns its administrative home (for governance) to *Strategic Risk* due to overlapping attributes.

4.4 Practical Implications

Classifying *Geopolitical Risk* as a **DRIVER** carries several operational consequences:

· Captures impacts as Driver → L1 Risk edges with materiality weights, rather than creating a new L1 risk category.

· Avoids duplication or double-counting when aggregating exposures.

· Preserves visibility by recording driver impacts, facilitating scenario construction and stress testing across the full breadth of affected L1 Risks.

5. Additional Validation Cases and Sensitivity

To assess discriminatory power and boundary behavior, we apply the framework to two additional candidates: *Model Risk* (a canonical sub-risk) and *Climate Risk* (a contested cross-cutting phenomenon).

5.1 Sub-risk Example: Model Risk

A canonical sub-risk candidate, Model Risk, receives attributes {internal, process-driven, operational}, and concentrates overwhelmingly on *Operational Risk*:

BREADTH = 1,

CR = 0.72, and

HHI = 0.52.

The three tests unanimously support SUB-RISK under *Operational Risk*. Notably, this result aligns with Basel Committee guidance and industry practice.

5.2 Edge Case: Climate Risk

Climate risk often spans credit, market, operational and strategic channels (TCFD, 2017). With attributes {external, market-driven, cross-cutting, financial}, the framework computes:

BREADTH = 4,

CR = 0.30,

HHI = 0.19

These metrics indicate **DRIVER** across all three tests, with differentiated materiality weights across affected L1s (e.g., credit > market > operational>strategic). *Climate Risk* is therefore a **DRIVER,** consistent with emerging regulatory treatment that eschews a siloed L1 Risk in favor of augmented scenario analysis across existing risk categories.

5.3 Comparative Results

Table 5 consolidates classification outcomes for the three candidates, while **Figure 3** visually contrasts their impact distributions.

Table 5: Case Study Results

CANDIDATE	ATTRIBUTES	BREADTH	CR	HHI	RELEVANCE	MATERIALITY	DISTINCTION	RESULT
Geopolitical	{external, event-driven, cross-cutting}	6	0.25	0.20	DRIVER	DRIVER	DRIVER	DRIVER
Model Risk	{interna, process-driven, operational}	1	0.72	0.52	SUB-RISK	SUB-RISK	SUB-RISK	SUB-RISK
Climate	{external, market-driven, cross-cutting, financial}	4	0.30	0.19	DRIVER	DRIVER	DRIVER	DRIVER

Classification results for three case-study candidates. Geopolitical and Climate are classified as drivers (cross-cutting); Model Risk is classified as a sub-risk under Operational Risk.

Figure 3: Impact Distribution Comparison

Impact distribution profiles for the three case study candidates.

The visual contrast underscores the framework's discriminatory logic: *Model Risk* exhibits a single dominant bar (Operational), characteristic of a sub-risk, whereas *Geopolitical* and *Climate Risks* display dispersed profiles consistent with cross-cutting drivers.

5.4 Sensitivity Analysis

Threshold calibration does require normative judgement and therefore, we conducted a sensitivity analysis, perturbing CR and HHI thresholds. We observed by perturbing thresholds ±10% we maintained classification stability in our examples:

Table 6: Sensitivity Analysis

CANDIDATE	METRIC	BASE VALUE	THRESHOLD	FLIP RANGE	STABLE?
Geopolitical	CR	0.25	0.50	Would need CR $\geq$ 0.50	Yes
Geopolitical	HHI	0.20	0.25	HHI $\in$[0.22, 0.28]	Yes
Model Risk	CR	0.72	0.50	Would need CR < 0.40	Yes
Model Risk	HHI	0.52	0.25	Would need HHI < 0.20	Yes
Climate Risk	CR	0.30	0.50	Would need CR $\geq$ 0.50	Yes
Climate Risk	HHI	0.19	0.25	Would need HHI $\geq$ 0.25	Yes

Sensitivity analysis showing classification stability under ±10% threshold perturbations. All tested candidates-geopolitical, model, and climate-maintain their classification across reasonable threshold variations.

All three candidates retain their classifications across the tested perturbation range. For instance, *Geopolitical Risk* would require CR ≥ 0.50—double its observed value—to flip to sub-risk; Model Risk would require CR < 0.40, well below its observed 0.72. These margins indicate robust discriminatory power under plausible threshold variation.

6. Implementation Considerations

6.1 Data and Governance Requirements

a. Deploying the framework requires the following data requirements:

· **Taxonomy mapping:** a catalog of current L1/L2/L3 elements and attributes.

· **Controlled vocabulary:** maintenance via periodical reviews and versioned lexicon **Table 1**).

· **Impact assessment matrix:** qualitative scores or quantified weights with documented provenance and expert sign-off.

b. Governance integration proceeds naturally:

· Classification operates as a Second-Line activity with a documented appeals pathway to the Risk Committee/Risk Executive(s).

· Annual vocabulary refresh and periodic sensitivity checks.

· Recordkeeping for classification decisions and supporting metrics satisfies audit and regulatory review.

6.2 Computational complexity and tooling

All core operations are tractable at enterprise scale. Complexity:

· **Breadth computation**: O(|D| × |L1|).

· **Attribute meet**: O(|V|^2) worst-case for naive comparisons, typically trivial for small vocabularies.

The method can be implemented in simple spreadsheets, Python scripts or as a lightweight service; no advanced libraries are required. For reproducibility and auditability, store the attribute assignments, impact matrices and computed metrics in versioned artifacts alongside decision memoranda.

6.3 Limitations

There are several caveats that warrant acknowledgement:

· **Vocabulary drift:** Without disciplined maintenance, the vocabulary may lose coherence and/or reproducibility.

· **Threshold normativity**: CR and HHI carry normative judgments; institutions should calibrate them to strategic tolerance (e.g., risk appetite) and regulatory context.

· **Scope Constraint**: This methodology addresses classification, not measurement or capital allocation; it complements measurement frameworks rather than replacing them.

7. Conclusion

We have presented a pragmatic, auditable approach to resolving a central taxonomy problem—distinguishing vertical sub-risks from cross-cutting drivers—by blending set-theoretic computation with categorical semantics for interpretability. The method's three quantitative tests (Relevance, Distinction, Materiality) and a clear decision rule provide a defensible default recommendation that preserves domain judgment through documented governance overrides.

Organizations that adopt this approach gain a repeatable, transparent process that reduces ad-hoc taxonomy proliferation, improves auditability, and better captures the systemic role of drivers such as geopolitics or climate when integrated into risk and scenario analysis workflows.

Declarations

Author Contribution

I, Marc Faria, am solely responsible for all research and contents of this manuscript.

References

1. Basel Committee (2006). International Convergence of Capital Measurement.
2. COSO (2017). Enterprise Risk Management—Integrating with Strategy and Performance.
3. ISO (2018). ISO 31000:2018 Risk Management Guidelines.
4. Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk.
5. Baez, J. C., & Pollard, B. S. (2017). A compositional framework for reaction networks.
6. Spivak, D. I. (2014). Category Theory for the Sciences.
7. Caldara, D., & Iacoviello, M. (2022). Measuring Geopolitical Risk. American Economic Review.
8. U.S. DOJ/FTC (2010). Horizontal Merger Guidelines (HHI thresholds).
9. Rhoades, S. A. (1993). The Herfindahl-Hirschman Index. Federal Reserve Bulletin.

Figures

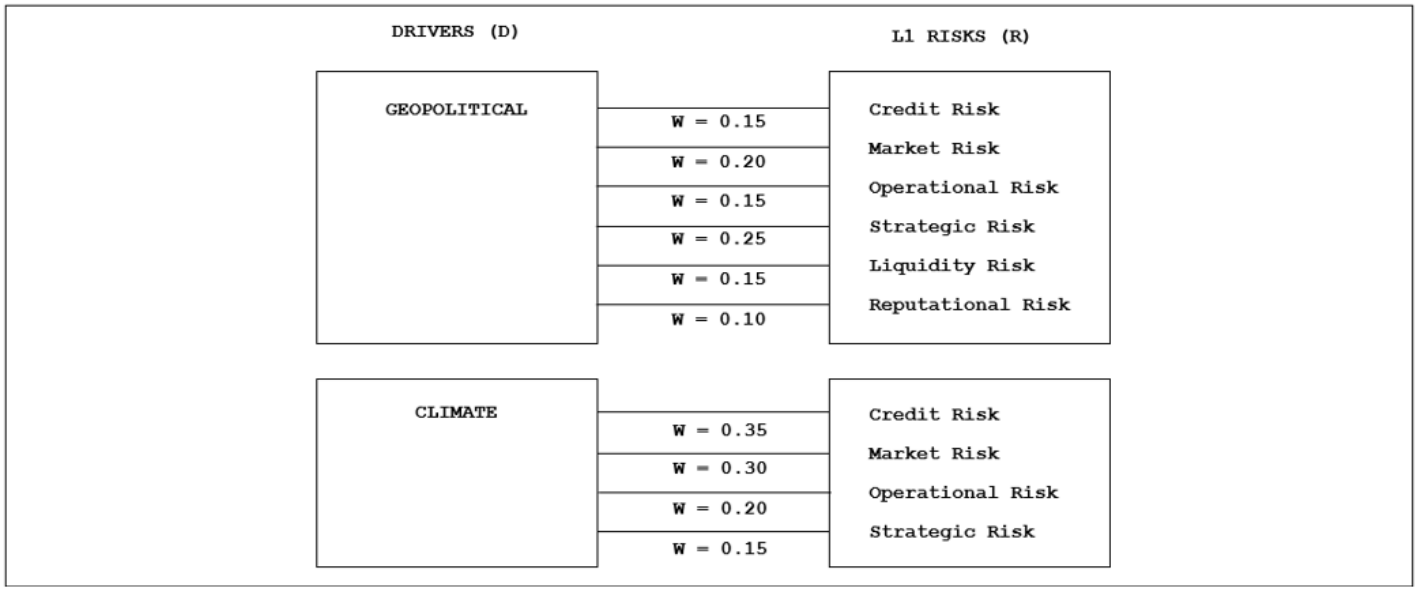


Figure 1

Bipartite Structure Illustration

A graph showing Driver set D on the left, Risk set R (L1 risks) on the right, with weighted edges representing impact assignments.

Bipartite structure G = (R, D, E, w). Each driver in D connects via weighted edges to multiple L1 risks in R. Edge weights w(c, r) represent the materiality of driver c's impact on L1 risk r.

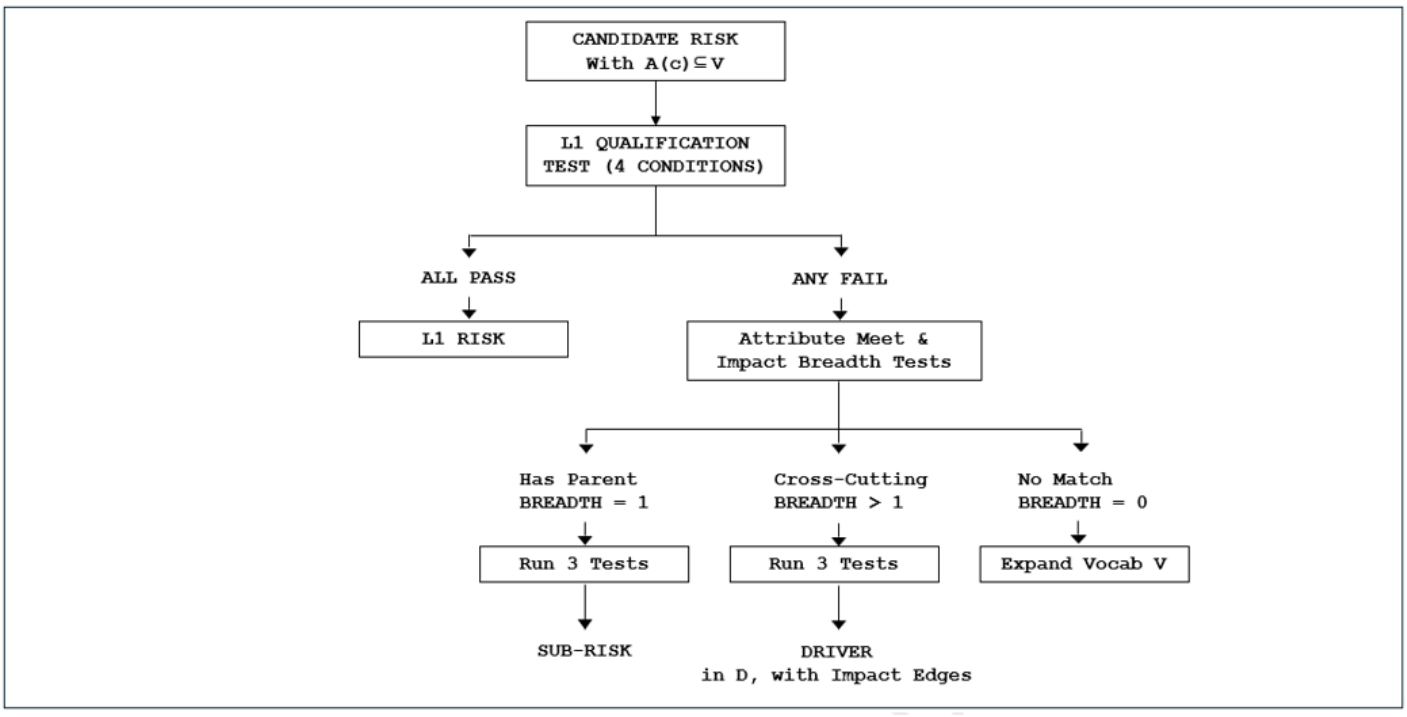


Figure 2

Classification Decision Flowchart

A decision-tree diagram illustrating the two-stage classification algorithm.

Classification decision flowchart. Candidates first undergo an L1 qualification test (four conditions). Those failing proceed to attribute-meet and impact-breadth checks, then to the three quantitative tests (Relevance, Distinction, Materiality) that determine sub-risk vs. driver classification.

CANDIDATES	CREDIT	MARKET	OPERATIONAL	STRATEGIC	LIQUIDITY	REPUTATIONAL
Geopolitical Risk	0.15	0.20	0.15	0.25	0.15	0.10
Model Risk	0.00	0.14	0.72	0.08	0.06	0.00
Climate Risk	0.35	0.3	0.20	0.15	0.00	0.00

Figure 3

Impact Distribution Comparison