Building a Supplier-Side Sovereignty Index for the EuroStack Solution Finder

Stefane Fermigier (Abilian) · sf@abilian.com

Draft working paper v0.2, 2026-11-09. This paper specifies the index. The criterion basis, the rulebook, every coding of it and the analyses below are generated from repository data and recompute when a row is contested. An earlier version of the index is deployed in a public directory; Section 12 states where it departs from this one.

Abstract

The European Commission has awarded a public contract on a sovereignty score: it rated the bids for one of its own cloud tenders, worth up to EUR 180M over six years and awarded in April 2026, on eight sovereignty objectives. A Regulation before the Council and the Parliament would make a four-level assurance scale the condition of selling to any Union public body, and national and private frameworks multiply. A score of this kind has to settle three things: what it measures, which evidence it admits, and how it combines that evidence. The frameworks published by others state dimensions and weights without deriving either, count an answer the same whichever way it was obtained, and report a level without saying how much of the supplier they have established. This paper settles the three for the index of the EuroStack Solution Finder, before it scores any service.

This paper specifies a supplier-side index built from the question a buyer needs answered: when it obtains a digital service from a given supplier, how much of the decisions that matter about that service remain its own, or can be taken back? Six decisions define the dimensions: who can read or use the buyer's data, who can alter the service, who can stop it, who sets the terms, who decides what the service becomes, and who can maintain or replace it. A second family records what the service contributes to European capacity, beginning with the contribution share: the European value-added content of a euro the buyer pays, read through people, capital, inputs and returns. Autonomy on each dimension is a degree: retained, recoverable or ceded. A criterion is evidence about a degree; it enters the index when it is informative, checkable and hard to fake.

Four blind coders, AI agents on two models, classified the 94 distinct criteria of eighteen frameworks under a written rulebook. They agree substantially on the dimension and the channel and less on strength and durability. The rulebook has not been tested on criteria held out of its revision. Fifteen criteria are strong evidence under every coding. The channel a criterion is read through bounds what a reading can show: a pass on a supplier's declaration can shift the odds between two degrees by little, a pass on a registry extract or a test by much more. For a reader who scores each reading on its own, a supplier that holds an attribute gains from the more informative channel and one that does not gains from the less informative one, so the channel a supplier chooses is itself evidence. Readings combine by their likelihoods into a degree on each dimension, with its confidence and the coverage behind it.

Consequences for the scoring follow. The fifteen strong criteria are the core of what the index collects, while criteria that are evidence on no dimension, about three tenths of the weight in the Commission's assessment, are collected and displayed and never scored. Each question requests evidence through the strongest channel that reaches it, since a declaration without other evidence barely moves the estimate. No question aimed at who can alter the service is strong evidence under every coding, and on who decides what it becomes only the veto rights in the ownership chain are. Those questions have to be written, and the rulebook has yet to grade existing criteria on the other dimensions they bear on. Suppliers need a route to correct the record, since a registry that wrongly fails a supplier misreports it by name. The result is a profile, one degree per dimension shown as a radar; a dimension shows a degree, or a bound such as at least recoverable, only once the evidence makes it more than 80 percent likely. Composite figures

for kinds of buyer state their weights and appear only where evidence beyond the supplier's word establishes a degree or a bound on every dimension they weight. The index builds in no gate: a buyer who wants a minimum on one dimension sets it. A service whose evidence there is missing or inconclusive is then reported as not established. The classification predicts how often services at each degree meet each criterion, which a review of twenty services is designed to test. The sharpest limit is node scope: a good score for one supplier is consistent with a captured stack.

Keywords: sovereignty, strategic autonomy, procurement, composite indicators, aggregation, verification.
JEL: D82, H57, L15, L86.

1. Evidence before scores

A sovereignty score has entered public procurement, in one tender so far. The Commission's Cloud Sovereignty Framework scores a service on a Sovereignty Effectiveness Assurance Level across eight objectives and takes the lowest as the verdict. The Commission has used it in a cloud tender of its own, worth up to EUR 180M over six years and awarded in April 2026 (European Commission 2025, 2026a). Annex II of the proposed Cloud and AI Development Act sets four Union assurance levels that would decide what any Union public body may lawfully buy (European Commission 2026b). The Netherlands has a sovereignty assessment framework for cloud services (DICTU 2026). Germany has criteria enabling cloud computing autonomy (Bundesamt für Sicherheit in der Informationstechnik 2026). France has a certification scheme with a chapter on jurisdictional exposure (ANSSI 2022), and a dozen vendors and investors publish scores of their own. A register we maintain for this work lists seventeen such frameworks beside six of our own; the criteria of thirteen of them, and of five of ours, are mapped one by one.

A score of this kind has to settle three things: what it measures, which evidence it admits, and how it combines the evidence. The frameworks published by others settle the first by listing dimensions and the third by a weighted sum or a minimum, and leave the second implicit. A question answered by a signed statement counts as much as one answered by a registry extract. A service assessed on complete evidence and a service assessed on two public facts and a filled-in form arrive at the reader looking identical, since none of the outside frameworks reports how much of the supplier it has established. Applied to the Commission's assessment, the test of Section 3 finds about two thirds of its weight on questions that are not strong evidence of the buyer's autonomy. *Ask Me Something Hard* gives the policy reading of that result and the questions to ask in their place (Fermigier 2026a).

This paper specifies the index the EuroStack Solution Finder will score with. It states the question the index answers, asks of each criterion which decision about the service it is evidence on and how strongly, combines the evidence by a stated rule, and publishes its coverage beside every result. The specification decides four things about the scoring that follows it.

1. **What to ask.** Fifteen criteria are strong evidence under every coding (Section 4). They are the core of what the index requests from a supplier and of what its probes look for. Criteria that are evidence on no dimension, about three tenths of the Commission's weight, are collected and displayed for the procurement officer who needs them, and never scored (Section 3).
2. **How to ask.** At the channel rates three AI respondents gave, a declared pass shifts the odds between two degrees by at most about 1.3, a registry or probe pass by up to 7 (Section 6). Each question therefore requests evidence through the strongest channel that reaches it. A supplier's own statements on one dimension count once, however many there are (Section 7).

3. **What is missing.** No question aimed at who can alter the service is strong evidence under every coding; on who decides what it becomes, only the veto rights in the ownership chain are (Section 4). The index has to write those questions and test them, and to grade existing criteria on the other dimensions they bear on (Section 12). The build and signing chain, controls against remote features and rights over the roadmap are where the codings point.
4. **What has to work.** At the elicited rates, a supplier that holds an attribute prefers being checked to declaring until a registry wrongly fails more than 70 percent of holders (Section 6), so the sorting tolerates a poor registry. A wrong failure still misreports a named supplier, so a route to correct the record belongs to the scoring (Section 9), and so does a published error rate for each channel.

What is imported and what is new. The screening result is in *Screening for Sovereignty* (Fermigier 2026h): a requirement reveals a hidden property only where holding the property genuinely costs differently and verification reaches the difference. Here it becomes the test of a criterion's strength. The result that a count of well-rated dimensions identifies the wrong quantity is in *Counting Dependencies* (Fermigier 2026c). The decomposition of exit cost across managed cloud services is in the cloud services matrix paper (Fermigier 2026l). The result that a supplier's markup over a locked-in buyer equals that buyer's cost of leaving is Proposition 1 of *Exit, Pursued by a Bill* (Fermigier 2026j). The two-stage composite, a minimum followed by a weighted mean, comes from *Scoring the Seller* (Fermigier 2026g) and is reworked and proved again here for degrees. Six things are new here, of unequal weight. The object is new: autonomy as who holds each of six decisions over one service, read as a degree. It applies the property-rights view of control, under which owning an asset means holding the residual rights to decide about it (Grossman and Hart 1986; Hart and Moore 1990), to a service the buyer does not own. So is the rulebook that grades each criterion as evidence on a degree, with its coding of the field. The analytical core of the design is the bound a channel sets on what one reading shows, with the finding that the channel a supplier chooses is evidence too. Propositions 1 and 2 are known results put to work here. The second is also partial: the equilibrium in which readers interpret the choice of channel is not worked out. The rule that combines readings into degrees, with coverage and a group for what is not established, is assembled from diagnostic testing and the construction of composite indicators. The contribution share takes the value-added measure of trade statistics down to one supplier. The crosswalk maps eighteen frameworks onto one basis. Propositions 3 and 4 establish properties of the specification and claim no novelty as theory. Each piece is stated so that someone outside the project can contest it row by row; the review of twenty services (Section 12) tests the rulebook and the channel rates.

Related work. Six literatures bear on the design, and none has been applied to a sovereignty index. The property-rights theory of the firm defines control as holding the residual rights to decide what no contract settles (Grossman and Hart 1986; Hart and Moore 1990); the six dimensions of Section 2 list those decisions for a service. The construction of composite indicators sets out the choices of normalisation, weighting and aggregation, and when a compensatory rule misleads (OECD and Joint Research Centre 2008); the profile of Section 7 takes the non-compensatory side and leaves the weights to the buyer. The economics of certification and disclosure studies when quality claims separate sellers, when certifiers inform, and when sellers reveal what they know (Akerlof 1970; Grossman 1981; Milgrom 1981; Lizzeri 1999; Dranove and Jin 2010). Section 6 adds a case in which the seller chooses how its claim is checked, and holders and non-holders choose opposite channels. It is close to ratings shopping, where an issuer picks which rating to disclose (Skreta and Veldkamp 2009), and to tiered certification, where a seller picks which certifier to face (Farhi, Lerner and Tirole 2013). The choice exists only where the supplier controls the route: registries and probes run without it, so the choice is between declaring and offering something checkable, a document tied to a record it does not control or a test the buyer can run. Measures used

for reward bend toward what is measured (Holmström and Milgrom 1991); the durability classes of Section 5 sort criteria by whether bending them delivers the attribute. Combining imperfect tests is routine in diagnostic medicine, where dependence between tests distorts their estimated accuracy (Vacek 1985) and a reference standard that shares inputs with the test inflates it (Whiting et al. 2004); Sections 7 and 12 apply both warnings. Value added in trade measures where value is produced along a chain (Johnson and Noguera 2012; Koopman, Wang and Wei 2014); Section 8 takes the measure down to one supplier and counts profit by who receives it.

2. The object being measured

The question. When a buyer obtains a digital service from a given supplier, how much of the decisions that matter about that service remain the buyer's own, or can be taken back by the buyer? This is strategic autonomy at the scale of one service and one buyer: the capacity to make one's own choices about the service and to act on them. Protection against a foreign state is one part of it.

Six decisions. Each dimension is a decision about the service that someone other than the buyer could take in the buyer's place. Each is defined by the decision, before any criterion is looked at.

Dimension	The decision	Who may hold it
Access	Who can read, copy, use, retain or hand over the buyer's data, including metadata and keys	the supplier, its parent, a state that can compel either, an attacker
Integrity	Who can alter the service, the software it runs or the data without the buyer's consent	whoever controls the build and update chain, a state that can compel it, an attacker
Continuity	Who can stop the service, degrade it, or withdraw a component it depends on	the supplier, a supplier of the supplier, a licensor, a state that can order either
Terms	Who sets the price and the conditions, who can change them and how far given the cost of leaving, and whose courts decide a dispute	the supplier; the owner of the stack it runs
Direction	Who decides what the service becomes: roadmap, features, deprecations, the licence it is offered under	the supplier, its owners, the owner of a licensed stack
Capability	Who can maintain, modify, operate or replace the service if the supplier will not or cannot, including when it fails	the supplier alone, or also the buyer or a third party it chooses

Access, integrity and continuity are the dimensions on which a foreign parent's position weighs most. Terms, direction and capability belong to the supplier relationship itself, whoever owns the supplier. On those three, open source, open standards, interoperability and a working exit path keep decisions with the buyer. They are also the conditions under which a European supply base can compete on its merits. The other party can be anyone the buyer does not control, an attacker included: an intruder who can read the data holds that decision as surely as a state that can compel the supplier. Security criteria are therefore evidence on access, integrity and continuity, judged by the same test as any other criterion.

The six dimensions cover the exposures the screening model distinguishes: legal access, covert access by compelled changes to code, denial of service, and the lock-in rent. They also cover the six dimensions of the Franco-German joint paper (France and Germany 2026) apart from economic value creation, and the Commission's eight objectives apart from

environmental sustainability (Fermigier 2026i, Section 8, maps the joint paper's dimensions onto measurable objects).

Contribution, a second family. Autonomy over one service is not the whole of sovereignty. A buyer's choice of supplier also decides where what it pays ends up. Of each euro, some becomes value added in Europe (payroll, research, intellectual property, taxes) and some leaves as licence fees, royalties and margins to the owner of a stack abroad. A European operator reselling a licensed foreign stack and a European operator running its own can score alike on autonomy and very differently here. The supply also builds capability in Europe, or does not: development, skills, contributions to the open source others rely on. This family is reported next to the six dimensions and never folded into them; Section 8 defines the measure.

A degree on each dimension. Autonomy on a dimension is a degree, defined the same way on each:

- **Retained:** the buyer takes the decision, or can take it back at a cost it can bear in the ordinary course.
- **Recoverable:** another party takes the decision, but the buyer can take it back with time and effort it could plan for: a tested migration path, a maintained fork, escrowed source with the skills to use it.
- **Ceded:** another party takes the decision and the buyer cannot take it back on any timescale it could plan for.

The reference buyer. Whether a decision can be taken back depends on who is taking it: a ministry with an engineering team can maintain a fork that a small firm cannot. The published degrees are read for a reference buyer able to run and administer the service and to commission work from third parties, with no team of its own able to maintain the software. A buyer with more capacity may read recoverable where the reference buyer reads ceded, mostly on capability and terms; the evidence behind each degree is published so that such a buyer can read it again for itself.

The profile follows from the definition. The definition treats sovereignty as a set of options and the cost of moving between them, as *Path Sovereignty* states it and *Cuts and Options* formalises it (Fermigier 2026k, 2026f). An option set is destroyed by its binding constraint, not by the average of its constraints, so a strong position on one decision does not repair a ceded one on another. The arithmetic of Section 7 follows: each dimension is shown on its own, with no average built into the result.

The limits below each appear on the face of every published result, because each is a way a reader could take the number for more than it is.

Node scope is the sharpest of them. The index scores one supplier at a time, and sovereignty is a property of stacks. The number of controller-disjoint ways to operate a stack is at most the minimum over its layers (Fermigier 2026f, which proves this and the certification result that follows). The inequality becomes strict as soon as one controller appears in the option sets of two layers. For any m, stacks exist with m independent suppliers at every layer and exactly one way to operate the whole thing. An index that inspects one layer at a time therefore certifies a quantity other than the one the buyer needs. A good supplier score is consistent with a captured stack. Scoring more suppliers does not engineer the limit away: no sum, minimum or product of per-node scores recovers the stack-level quantity, because the shared controller is invisible in every individual score. The correct response is to say so on the result and to point the reader at a path-level measure.

The index estimates autonomy and identifies no outcome. No result in this paper shows that a higher degree causes a better procurement outcome. The index estimates how much of each decision a buyer would keep, which is a different claim and a weaker one.

A result describes a point in time. Enforcement and audit stay endogenous and costly. A supplier satisfying a criterion on the day of assessment has told the buyer about that day.

Control changes. An observation of where a service runs is a guess about where it runs six months later. Section 6 therefore dates every reading and repeats those that can be repeated. Section 5 classifies criteria by what happens to them once suppliers begin optimising against the published rule.

One further scope statement belongs here because it is routinely assumed away. The index is silent on product quality and on fitness for the buyer's need. The buyer's own procurement has to ask about those. Within the directory that deploys it, the index contributes, by the directory's methodology specification, roughly fifteen percent of the default ranking signal, alongside need-matching, maturity, community health and recency. Positioning it as *the* criterion is an explicit anti-goal.

3. Criteria as evidence

The specification the project inherited had five pillars of five criteria each, which is the shape such frameworks usually take; the deployed version scores eleven criteria in the same five pillars (Section 12). Twenty-five is a comfortable number: it looks thorough, it fills a form, every pillar gets equal visual weight. The design here starts from the criterion.

A criterion earns its place by three properties. It is *informative*: learning that a service meets it should change the estimate of the degree on some dimension. It is *checkable*: the buyer, or someone acting for it, can verify it in a public record the supplier cannot edit, by a test the buyer runs, by observation, or in a document someone independent can check. And it is *hard to fake*: once the rule is published, a supplier cannot produce the evidence without the autonomy being real.

The test of strength. Take a service at a lower degree on the dimension, where someone else holds the decision. Could it meet the criterion genuinely and stay at that lower degree? If it could not, because meeting the criterion moves the decision back toward the buyer, the criterion is strong evidence. If it could, at a modest or cosmetic cost, the evidence is weak. If services at every degree meet it equally, it is no evidence. The size of the supplier's expense does not decide; whether the decision has moved does. This is the screening result applied to a degree in place of a type.

Criterion	Could a service at a lower degree meet it and stay there?	Evidence
Data stored in an EU region	Yes: the parent can still be compelled to hand data over	weak, on access
Supplier incorporated in the EU	Yes: an EU subsidiary of a foreign parent	weak, on access
Signed statement that no foreign law applies	Yes: it can sign	weak, and unchecked
Customer holds the keys, checked by a revocation test	No: a supplier that can be made to hand over readable data fails the test	strong, on access
Migration off the service completed by a real customer	No: a service the buyer cannot leave has no such migration	strong, on terms
No parent outside the EU, in company registers	No: the register shows the controller	strong, on access and continuity

The rulebook and its coding. The test is written down as a rulebook that someone else can apply. It has a column for the dimension, one for the strength, one for the strongest channel that reaches the evidence, and one for durability: whether the evidence would keep its strength once suppliers optimise against a published rule. Every criterion the deployed frameworks ask, reduced to 94 framework-neutral atoms (102 rows once eight compound criteria are coded as their parts), was coded blind four times. The coders are AI agents,

two on each of two models, each working from the rulebook and the criterion list alone. On the 81 criteria that bear on the supplier, agreement between coders on different models, measured by kappa, is 0.85 on the dimension, 0.70 on the strength, 0.82 on the channel, 0.62 on durability, and 0.70 on whether a criterion is strong evidence. Two coders on the same model agree more closely (0.95 on the dimension, 0.80 on the strength), so the four codings are nearer two independent readings than four. Language models agree with each other more closely than crowd workers do on annotation tasks (Gilardi, Alizadeh and Kubli 2023), and agreement between them is weaker evidence of a rule's clarity than agreement between people. A criterion's strength is coded on its main dimension; the other dimensions it bears on are listed without one. The rulebook reached that form through earlier rounds, revised from the disagreements between codings and never by discussing a criterion. Agreement is measured on the same criteria the rulebook was revised against, with none held out; some of the rulebook's examples describe criteria of the basis. Human coders, and criteria held out of revision, are the stronger test.

Classified by that test, most of the field's weight is weak or no evidence. The Commission's assessment asks 48 questions under eight weighted objectives. Spreading each objective's weight over its questions, as its calculator does, and reading each question through the criteria it feeds:

Coding	Strong evidence	Weak evidence	Evidence on no dimension
First	36.9%	29.7%	33.4%
Second	36.9%	31.7%	31.4%
Third	39.1%	33.1%	27.9%
Fourth	40.7%	30.0%	29.3%

Under every coding, between 59 and 63 percent of the weight rests on questions that are not strong evidence, and between 28 and 33 percent on questions that are evidence on no dimension at all. Strong evidence in the table includes the contribution family; counting evidence of autonomy only, the share that is not strong evidence is 65 to 68 percent. Under the weights printed in the framework document, which differ from the calculator's on three objectives, the share that is not strong evidence is 61 to 64 percent. An earlier classification against the narrower test of separating two supplier types put the non-separating share between 52 and 60 percent across five codings (Fermigier 2026a); the wider question asked here gives a similar share for a different reason. *Ask Me Something Hard* draws the policy consequences (Fermigier 2026a). For the index, the classification decides which criteria are scored, as the next paragraphs set out.

The classification is a prediction. Section 7 defines strength by pass rates: a criterion is weak evidence when services at every degree meet it at nearly the same rate. Those rates are not yet measured (Section 12), so the codings predict them. Criteria coded strong should show pass rates further apart across degrees than criteria coded weak, on services whose degree is fixed independently. The distance is the ratio of the rates, so that a rare criterion met only at the top degree counts as strong. A criterion coded strong whose rates do not differ, or one coded weak that separates degrees cleanly, is a row the rulebook has wrong.

Evidence on nothing has costs, none of them neutral. A criterion that is evidence on no dimension still has to be collected, stored, refreshed and disputed, so the audit budget a framework imposes grows faster than its criterion count. Inside a weighted sum it dilutes the criteria that do bear on the buyer's position, so a supplier can offset a ceded decision with a certification every supplier holds. The reader of a single number cannot tell which of the two they are reading.

The resolution keeps both functions and separates them. Evidence on a dimension enters the degree on that dimension, each criterion for what it is worth: strong evidence moves the estimate a long way, weak evidence a little (Section 7). Criteria that are evidence

on no dimension are collected, displayed and searchable, and never scored. A procurement officer who needs to see a subprocessor list, a retention policy or a security certificate finds it, structured and comparable.

4. What the evidence covers

Fifteen criteria are strong evidence under all four codings.

Family	Criteria
Access	the ownership and control chain; exposure to extraterritorial compelled-access law; customer-held keys; the identity root of trust
Continuity	continuity and operational autonomy if the supplier is cut off; exposure to export controls; single points of dependency outside the reference jurisdiction
Terms	data portability; exit cost with a tested migration path
Capability	interface openness; maintenance without the originating vendor; documentation and source sufficient for a takeover; open-source availability and licence
Contribution	where intellectual property is created and held; AI models developed and trained in Europe

No criterion aimed at integrity is strong evidence under every coding; one aimed at direction is. That one is part of the ownership chain: veto, board-appointment or blocking rights held from outside the reference jurisdiction, which every coder places on direction, and which the table above counts under access with the rest of the chain. Every coder also classes it fragile, since it is read through a document. The build and signing chain is strong evidence for three coders of four, controls against remote features for two; locks against acquisition or enclosure for two, stakeholders' say over the roadmap for one. The frameworks do ask who can alter the service and who decides what it becomes. They ask in forms a supplier can meet without moving the decision: a statement, an advisory board, a document. Some criteria strong on another dimension also bear on these two. Every coder lists the identity root of trust as bearing on integrity, as strong evidence. Three of four list open-source availability as bearing on integrity and direction, as strong evidence; the fourth places it principally on direction. Three of four list exposure to compelled-access law as bearing on integrity or direction. The rulebook does not grade them there. The codings therefore show only that no question aimed at integrity, and none aimed at direction apart from the veto rights, is strong evidence under every coding. How strong the field's evidence on those two dimensions is overall remains to be graded (Section 12).

Terms is the economic dimension the field leaves implicit. A supplier's markup over a locked-in buyer equals that buyer's cost of leaving, at the margin under a posted price and buyer by buyer under individual pricing (Fermigier 2026j, Proposition 1). A high exit cost is therefore a high rent capacity. The magnitude of realised lock-in rents remains an open empirical question; the capacity does not. External figures point the same way. In a survey Asterès conducted for Cigref, 54 chief information officers in France, Belgium and the Netherlands report the cost of their cloud and software services rising 8.7 percent a year over three years; how much of that growth is rent is not identified (Asterès 2026). The repricing of VMware after its acquisition is the single-case version. Apart from those we maintain, no supplier-side framework in the register has a first-class economic dimension; the Commission folds rent and lock-in into its technology and operational objectives behind a strategic weight.

Contribution is the family the frameworks barely ask about. All four coders place five criteria principally under it: the economic footprint in the reference jurisdiction, where intellectual property is created, which jurisdiction holds it, AI models developed and trained in Europe, and contributions to the open source the service relies on. All four also read

skills and the location of software development as bearing on it. Five of the thirteen outside frameworks mapped criterion by criterion, and four of our five, ask at least one criterion whose principal dimension is contribution (Section 11). None asks how much of what the buyer pays stays in Europe. That quantity comes from accounts, country-by-country reports and the licence terms between an operator and the owner of the stack it runs, and a questionnaire reaches it only through the licence terms. At the level of whole blocks of the stack, *An Open Internet Stack* computes the EU share of value added with sources and confidence flags (Fermigier 2026b). Section 8 brings the quantity down to one supplier and one euro.

5. Scoring a criterion

Every criterion scores in [0, 1] with a distinguished value for absent evidence, higher meaning more autonomy, against five anchors: fully meets, substantially meets, partially meets, minimal, does not meet. Criteria with evidence too thin for five-way discrimination admit a reduced subset. The scale is one; the granularity is per criterion.

Absent evidence and inapplicability are different, and conflating them corrupts coverage. A criterion marked ∅ means the index looked and found nothing, which lowers coverage and is visible to the reader. A criterion marked N/A means the question is meaningless for this category, as hosting location is for a desktop application. It leaves the dimension's evidence set entirely, changing the denominator of coverage and not its numerator. An index that scored ∅ as zero would read every fact left unexamined as a failure; one that scored N/A as ∅ would report ignorance where there is none.

The exit criterion shows why anchors have to be argued and not assigned. The obvious anchoring for exit cost runs on data export: full export at the top, no export at the bottom. The measurement says that anchoring is wrong. Across thirty-five managed cloud services, policy acting on data export alone reaches on average a fifth of the exit tax scored there, and under thirty percent on the deeply coupled services where the problem lives (Fermigier 2026l). The cost sits in re-pointing an application off proprietary interfaces, not in moving the bytes. So the anchors score that instead, and treat a complete export as the precondition it is:

Score	Condition
1.00	A second supplier serves the same interfaces, with a migration completed by a real customer
0.75	Interfaces are standard or widely implemented; the migration path is documented and untested
0.50	Complete export in a documented format; the application has to be rewritten against another interface
0.25	Export only, partial or requiring the supplier's assistance
0.00	No documented exit

A supplier can meet the export obligations of Chapter VI of the Data Act (European Union 2023), which also phases out switching charges, without lowering its customers' exit cost by much. An index anchored on export would have certified compliance with that obligation and called it portability.

Criteria behave differently once the index is consulted at award, and the rulebook records which. This is the Goodhart problem, handled by classification:

- **Durable.** The evidence is the attribute, or moving the evidence means changing the attribute: the ownership chain in a register, a licence in a repository, a test only genuine autonomy passes. These criteria keep their strength.
- **Self-defeating.** Cheap to satisfy. A supplier that satisfies one has thereby given the buyer what the criterion is for: interfaces opened, prices published with capped egress,

an export that works. These are the criteria an index should want most, the one class where gaming is a success condition.

- **Fragile.** Cheap to satisfy, with nothing in it for the buyer: a plan written for the audit, a region chosen for the probe, a signed statement. They detect washing today because no supplier is yet optimising against them. Treat them as a one-shot audit of the current population and expect their strength to decay.

Under the four codings, between three eighths and a little over half of the strong criteria are durable, and nearly all the fragile ones are read through a document. Durability is also the property the coders agree on least.

The classification predicts; successive editions show what happened. A pass rate that climbs from one edition to the next has two causes with opposite remedies. Under saturation, services at all degrees come to hold the attribute, the pass rates converge across degrees, and the criterion stops being evidence on the dimension. That is the policy working; the criterion should leave the profile. Under gaming, suppliers learn to pass without holding the attribute and the channel's false-pass rate rises. The criterion should then stay, read through a stronger channel. The two can be told apart without knowing any service's degree. Under saturation, passes rise on all channels, registries and tests included. Under gaming, they rise only on declarations and documents, and readings of one supplier that conflict across channels multiply. A saturated criterion leaves the profile when, on the strongest channel that reads it and in two consecutive editions, the ratios of its pass and failure rates at the highest and lowest degrees stay within a margin of one that the rulebook fixes. The rule is announced one edition ahead. It stays in the disclosure layer, where a reversal would show.

6. Channels of evidence

Claims arrive by two routes. Self-declaration follows the deployed submission form: a blocking eligibility gate answered before scoring opens, per-line evidence on each claim, a signed attestation where no third-party certificate exists, and sections conditional on the solution's category. Black-box observation watches the supplier's public estate without its cooperation, which makes it cheap, repeatable and difficult to refuse.

The black-box route does more than raise confidence on observable criteria. The difference between what a supplier declares and what a probe observes is the washing gap, measured. A vendor that declares the service hosted on infrastructure it operates, while the service's endpoints resolve to a hyperscaler's network, has produced a datum about itself that no questionnaire would have elicited. Eight probe families are in use. They cover hosting and network ownership, mail exchange, and the naming layer with the content delivery and certificate issuers. Four more reach embedded third-party dependencies on the vendor's front end, the company registry, repository signals and certification registers. The last watches telemetry endpoints for software running on the buyer's machines. Two rules govern every probe. Only actual embeds count, since a link is not a dependency. Every finding also records the signal that fired, so a supplier can dispute the observation itself.

A probe reads what it reaches, which is not always the service. Across 761 domains of European-established suppliers in the directory's catalogue, probed on 8 November 2026, the mail exchange or name service of 67 percent sits with an operator under a non-EU parent, Google above all, then Cloudflare, Amazon and Microsoft. Thirteen percent sit entirely with European operators or on servers the supplier runs; 20 percent could not be classified. A probe of a supplier's corporate mail therefore fails most European suppliers on the strength of their office tools. It reads a proxy for single points of dependency, the supplier's corporate domain, and is weak evidence on continuity, scored at the weak rates although the criterion it stands in for is coded strong (Section 7). A failure on it says almost nothing about the service.

A claim's channel caps its confidence, on an ordering that is the operational form of audit accuracy. A registry hit, a certificate confirmed in the issuer's register or a test the buyer runs sits at the top. An independent probe observation, or a document someone independent can check, sits below it. Below that comes a signed attestation, or a declaration with a document link that resolves. A bare self-declaration is at the bottom. A criterion established only by the supplier's word about itself does not reach the threshold at which the screening result says a requirement bites, which is why evidence resting on it counts as weak however real the underlying property.

The ordering becomes a bound once the pass rate is split. The pass rate of Section 7, $p_c(d)$, combines a property of the criterion with a property of the channel:

$$p_c(d) = \beta\, h_c(d) + \gamma\big(1 - h_c(d)\big),$$

where $h_c(d)$ is the share of services at degree d that hold the attribute criterion c reads, and the channel passes a holder with probability β and a non-holder with probability γ.

Proposition 1 (channel cap). Let the channel read only the attribute, so that β and γ do not depend on the degree, with $0 < \gamma < \beta < 1$. For any holding rates, the likelihood ratio a reading gives between two degrees d_1 and d_2 lies in $[\gamma/\beta,\ \beta/\gamma]$ on a pass and in $[(1-\beta)/(1-\gamma),\ (1-\gamma)/(1-\beta)]$ on a failure. The bounds are reached when the attribute separates the two degrees perfectly.

Proof. $p_c(d) = \gamma + (\beta - \gamma)\, h_c(d)$ is increasing in $h_c(d)$ and lies in $[\gamma, \beta]$, so the ratio of two such values lies between γ/β and β/γ. The failure case applies the same argument to $1 - p_c(d)$. ∎

The bound is the attenuation of a test's likelihood ratio under misclassification (Bross 1954). A brute-force check on 20,000 random channels and holding rates confirms it. The channel rates themselves are unmeasured. Three AI respondents, in the roles of a buyer, an engineer and an analyst, each in a fresh session given only a role description and the questions, gave these plausible values for ownership and continuity attributes:

Channel	Passes a holder (β)	Passes a non-holder (γ)	Largest pass ratio	Largest failure ratio
Declaration by the supplier	0.90	0.70	1.29	3.0
Document a third party can check	0.80	0.40	2.0	3.0
Registry extract or direct probe	0.70	0.10	7.0	3.0

At these values a declared pass shifts the odds between two degrees by at most 1.29 and a registry or probe pass by up to 7. At the point values a failure on any channel shifts them by at most 3, because the respondents put the strong channels' false failures higher: a registry is only as good as the record-keeping behind it. That a registry failure weighs no more than a document failure holds in only 42 percent of the draws below, so it is not relied on. Drawing each value 20,000 times from the ranges the respondents gave, and keeping the 93 percent of draws in which every channel passes holders more often than non-holders, the ordering of the pass ratios holds in 85 percent of them. A declared pass stays below a ratio of 1.5 in 80 percent.

The choice of channel is itself evidence. Suppose a reader with prior q that a supplier holds an attribute scores one reading by its posterior and ignores how the channel was chosen.

Proposition 2 (choice of channel). For a channel with $0 < \gamma < \beta < 1$, write $p = q\beta + (1-q)\gamma$ for its pass rate and $I = (\beta - \gamma)^2/\big(p(1-p)\big)$. A holder's expected posterior is $q + q(1-q)^2 I$ and a non-holder's is $q - q^2(1-q)\, I$. A holder therefore prefers the channel with the larger I, and a non-holder the one with the smaller.

Proof. The posterior μ after a reading has expectation q, so $q\,E_H[\mu] + (1-q)\,E_N[\mu] = q$. Since $\Pr(y)\,\mu(y) = q\,\Pr(y \mid H)$ for each reading y, the second moment of μ is $q\,E_H[\mu]$ and its variance $q\,(E_H[\mu] - q)$. For one binary reading the variance is also $p(1-p)\,(\mu_1 - \mu_0)^2$, with $\mu_1 - \mu_0 = q(1-q)(\beta - \gamma)/\big(p(1-p)\big)$, which gives $q^2(1-q)^2 I$. The two expressions follow. ∎

I is the squared correlation between holding the attribute and passing, divided by $q(1-q)$. A brute-force check on 20,000 random pairs of channels and priors confirms both expressions. At the respondents' point values and $q = 1/2$, I is 0.25 for a declaration, 0.67 for a document and 1.5 for a registry or probe, so holders sort toward the registry, non-holders toward declaring, and the result prints the channel a supplier chose. With the other point values held, a holder prefers declaring to the registry once the registry fails holders more than 70 percent of the time, and prefers a document to the registry from 53 percent. Over the draws of the previous paragraph, the comparison of the registry with the declaration holds in 97 percent at $q = 1/2$, 99 percent at $q = 0.2$ and 92 percent at $q = 0.8$. This is the logic of voluntary disclosure (Grossman 1981; Milgrom 1981) with the seller choosing the test where it would otherwise choose the message. A reader who also interprets the choice changes the incentives the proposition rests on; the equilibrium of that game is not worked out here. The proposition also takes the supplier's payoff to be linear in the posterior, as the composite is; a supplier facing a buyer's threshold may rank channels differently near it. Since the sorting fails when the strong channel fails honest suppliers too often, a supplier is shown what the registry says and may correct the record with evidence.

Readings drawn from one source, such as a declaration and a document the same supplier drafted, count as one reading, since they do not err independently. A criterion read through a document can be strong evidence under the rulebook while a single passing reading of it shifts the odds by about two at most and a failing one by about three. Strength classifies a criterion; the cap bounds one reading of it.

Declarations are adversarial inputs and are treated as such. They are evidenced, discounted or perturbed, never taken on trust. A robustness pass that perturbed only the observed fields while trusting the declarations would manufacture confidence it has not earned, so the pass specified for this index perturbs the declared flags first, because those are the outcome-determining inputs.

Readings are dated and repeated. The facts the criteria read change by events: a change of control, a migration, a new licence. Between events they stay true, so a reading of an unchanged fact loses nothing with age, and a reading of a changed one is wrong in full. A reading's worth therefore depends on how soon a change is seen. A reading the index can repeat, a probe, a registry extract or a test, is repeated on a published schedule, and the newest reading replaces the older ones. A reading only the supplier can renew, a declaration or a document it drafted, stands for a year from its date and then expires unless renewed. A failure recorded on the supplier's own word does not expire; it stands until a newer reading of the same criterion replaces it, so a supplier cannot clear an unfavourable answer by letting it lapse. A reading counts in full while it stands. An expired reading reverts to ∅ and lowers coverage, which is the index reporting that it no longer knows something it once knew. Every reading is published with its date. The deployed specification discounted readings by age, to four fifths after three months and to half after six; a discount treats an unchanged fact as less true and a changed one as still partly true. Its rates were never measured.

7. From evidence to degrees

Five requirements fix the rule, most of them by elimination.

Requirement	Why it binds
Monotone in evidence	An adverse finding must never raise a degree

Requirement	Why it binds
No compensation across dimensions	A strong position on one decision does not repair a ceded one on another
Unknown is not zero	Missing evidence is reported as missing, scored neither as a failure nor as a pass, except a question the supplier was asked and left blank
Coverage-complete	Depth of evidence has to be visible
Retains gradient	Results have to discriminate above the floor

Within a dimension, evidence accumulates by its likelihood. A criterion read on a service gives a reading s_c on the criterion's anchors (Section 5); pass or fail is the case of two anchors. Write $\ell_c(s \mid d)$ for the share of services at degree d on the dimension that read s on criterion c as a buyer checks it, so that $\ell_c(1 \mid d) = p_c(d)$, the pass rate, for a criterion read as pass or fail. For a service with current readings (Section 6) on the criteria in its evidence set K (Section 5), the estimate over degrees is

$$\Pr(d \mid s) \propto \pi(d) \prod_{c \in K} \ell_c(s_c \mid d),$$

with π the prior over degrees, uniform until the review of Section 12 measures how degrees are spread. Every likelihood is floored at 0.01, so that no single reading rules a degree out and the posterior is always defined. This is the standard combination of conditionally independent signals. A criterion met at nearly the same rate at every degree moves the estimate by almost nothing, which is the formal content of weak evidence. A criterion met only at higher degrees moves it a long way. An unevidenced criterion is absent from the product, so unknown is not zero. A failed criterion counts against the service only by as much as failing is rarer at higher degrees. A criterion coded strong on two dimensions enters the product of each.

Until the review, the rates are interim values. Scoring starts before the review of Section 12 measures anything, so the rule runs on placeholders, published with every result and replaced criterion by criterion as measurements arrive. The holding rates $h_c(d)$ at ceded, recoverable and retained are 0.1, 0.5 and 0.9 for a criterion coded strong by at least three of the four codings, and 0.4, 0.5 and 0.6 for any other criterion that is evidence on a dimension. The channel rates are those of the table in Section 6: a registry extract, a test or a probe takes the row for registries and probes, a document a third party can check the row for documents, and a declaration or a signed attestation the row for declarations. A graded criterion is read as a pass at substantially meets or above and as a failure below, until rates per anchor are measured; its anchor is still displayed. A criterion enters only its main dimension, the one most codings give it, until the rulebook grades it on the others. A reading through a proxy takes the weak rates whatever its criterion's code. The probe of a supplier's corporate mail and name service is one: it reads the supplier's office tools and not the service (Section 6), and at the weak rates one failure on it sets no bound. A supplier's merged declarations (below) take the holding rates of the weakest criterion among them. At these values every criterion satisfies the ordering assumption below, and the floor never binds, so Propositions 3 and 4 apply from the first reading; the suite checks both. They also fix how much evidence a result needs. One pass of a strong criterion through a registry, a test or a probe makes recoverable or better 87 percent likely, enough for the bound at least recoverable. Through a checkable document it takes two; a supplier's declarations, counted once, never suffice. Establishing retained itself takes four independent strong passes through the strongest channel. A weak criterion needs three such passes for a bound. One failure of a strong criterion lands exactly on τ and sets no bound; two independent failures through a registry, a test, a probe or a checkable document set at most recoverable. On integrity and direction, where few criteria reach that bar (Section 4), a service will rarely show more than a bound until the questions of Section 1 are written. The middle degree is the hardest to establish: the interim values

place recoverable halfway between its neighbours, so no single criterion separates it from both. A service at recoverable will mostly show one of the two bounds. Criteria that separate recoverable from ceded, such as a tested migration path, or from retained, such as keys the buyer alone holds, are what the review has to measure.

One fact counts once. Conditional independence fails for readings that depend on one underlying fact, and counting both would overstate the confidence. Copies are merged by the rulebook's groups. Readings that one event would change together, such as two probes of the same host, are treated the same way, and so are readings that share a source. On one dimension, all the declarations a supplier makes and the documents it drafts itself count as one reading, at the declaration's channel, which passes only if every one of them passes. A document counts at the document channel only when what it states is checked against a record the supplier does not control, such as a register, a countersigned contract or an auditor's report. A question the self-declaration form asks and the supplier leaves blank makes that merged reading a failure. The supplier knows the answer and chose silence, which a reader takes as unfavourable, as in the economics of disclosure (Grossman 1981; Milgrom 1981); without the rule a supplier would leave its one unfavourable answer blank while the others pass. Everywhere else unknown stays unknown: a probe that gets no answer, or a criterion nobody asked about, is ∅. Three statements are therefore worth one: the suite checks this on a ceded service and an attribute held at recoverable and retained and not at ceded. Three passing declarations would meet a minimum of recoverable if each counted, and leave the service not established under the rule. At the declaration's elicited rates the probability of recoverable or better is 0.72 under the rule and 0.81 if each counted, against $\tau = 0.8$ the example illustrates the rule and claims no margin. Exposure to compelled-access law through a parent is a reading of the ownership chain. The criterion on compelled-access exposure is therefore read only on the supplier's own exposure; its parent route counts once, inside the chain. Dependence that remains between distinct facts still overstates the confidence somewhat; the review of Section 12 measures it from joint pass rates.

What is published. Write the degrees ceded, recoverable and retained as 0, ½ and 1. The most probable degree, ties going to the lower, is published when its probability, the confidence, exceeds $\tau = 0.8$. Short of that, the dimension shows the bound the evidence establishes above τ: at least recoverable when $\Pr(d \geq \text{recoverable}) > \tau$, otherwise at most recoverable when $\Pr(d \leq \text{recoverable}) > \tau$. Every threshold in the index is strict, so a probability of exactly τ establishes nothing; at the interim values below, one failing reading of a strong criterion lands there on every channel. With neither, and on a dimension with no reading, it displays as not established, with the most probable degree and its confidence beside it. The radar never shows missing or thin evidence as a degree. Its bounds rest on the probability the buyer's filter uses. Composites use the expected value of the degree, $E[v(d) \mid s]$. A buyer's minimum m on a dimension uses $\Pr(d \geq m \mid s)$.

Assumption (ordering). For each criterion and any two degrees $d' > d$, the ratio $\ell_c(s \mid d')/\ell_c(s \mid d)$ does not fall as s rises. For pass or fail it means that the pass rate does not fall as the degree rises; a criterion whose rate falls is read with its sense reversed.

Proposition 3 (monotone in evidence). Under the ordering assumption, lowering one reading, or adding a reading at a criterion's lowest anchor, never raises the published degree, the expected value of the degree, or $\Pr(d \geq m \mid s)$ for any m.

Proof. Write the ordering assumption in cross-product form, $\ell_c(s' \mid d')\,\ell_c(s \mid d) \leq \ell_c(s' \mid d)\,\ell_c(s \mid d')$ for $s' < s$ and $d' > d$, which also holds where a likelihood is zero. Summing it over s with $s' = s_{\min}$ gives $\ell_c(s_{\min} \mid d') \leq \ell_c(s_{\min} \mid d)$. Let g be the unnormalised posterior weights before the revision and g' after. Lowering a reading from s to s' replaces a factor $\ell_c(s \mid d)$ by $\ell_c(s' \mid d)$ adding a reading at the lowest anchor multiplies by $\ell_c(s_{\min} \mid d)$. In both cases the two inequalities above give $g'(d')\,g(d) \leq g'(d)\,g(d')$ for every $d' > d$. The new posterior is therefore lower in the likelihood-ratio order, and so in first-order stochastic dominance (Milgrom 1981): the expected value and every $\Pr(d \geq m)$ fall weakly. For the degree, let d^* be

the lowest maximiser of g, so $g(d^*) > 0$. For $d > d^*$, $g'(d)\,g(d^*) \le g'(d^*)\,g(d) \le g'(d^*)\,g(d^*)$, hence $g'(d) \le g'(d^*)$: no degree above d^* beats it after the revision, and the lowest maximiser does not rise. ∎

The same argument read upward shows that a better reading never lowers any of the three. The bounds the radar shows inherit the proposition, since each is a probability $\Pr(d \ge m \mid s)$ or its complement: an adverse reading never adds at least recoverable and never removes at most recoverable. A brute-force check on 20,000 random profiles, with ordered likelihoods over two to five anchors, confirms the proposition. With the ordering assumption dropped, an adverse revision raises the published degree in 16 percent of random draws, the expected value in half of them and the probability of meeting a minimum in two thirds; that is the failure region. A criterion with a pass rate not ordered by degree has to be split or reversed before it enters. The likelihoods are a measurement, made on services whose degree is established independently. Section 6 splits each into the share of services at a degree that hold the attribute, which such a review measures, and the rates of the channel, which are the same for all criteria read through it. Both remain to be measured (Section 12).

Across dimensions, no combination is built in. The natural candidates each fail a requirement.

Averaging smears the chokepoint. A profile at the ceiling on every dimension but one, and at the floor on that one, outranks a uniformly mediocre profile, although the first has a fatal single point of failure and the second has none. That a count of well-rated dimensions identifies something other than a weakest-link quantity is the result of *Counting Dependencies* (Fermigier 2026c), which applies to any mean.

The geometric mean collapsed in practice. It was used in an earlier prototype over a handful of display axes and produced a near-binary outcome under sparse coverage. The geometric mean of a single observed axis is that axis. Any axis at the floor sends the whole product to zero, which made very different estates indistinguishable. Worse, under the floor convention the prototype used, discovering an additional non-fatal dependency could *raise* the score, which is backwards for an exposure measure.

The bare minimum is too punitive across the non-zero range. The same prototype showed that (95, 70, 55) reads no better than (55, 55, 55), so a service with two strong dimensions and one adequate one is indistinguishable from one adequate everywhere. The requirement the minimum satisfies is real; the ranking it produces above the floor is empty.

So the result is the profile. Each dimension has its own degree. The reader sees all six. A buyer who needs a chokepoint reading can take the minimum over the dimensions it cares about, on the O-ring logic that a production chain is as strong as its weakest component (Kremer 1993), and require a minimum degree on them. That requirement is the buyer's, set for its own use. The filter returns three groups. A service fails when, on some named dimension, the evidence puts the degree below the minimum with probability above τ. It meets when, on every named dimension, the evidence puts it at or above the minimum with probability above τ. Otherwise it is not established, which includes every service with no evidence on a named dimension. Silence on the weakest dimension does not pass. An unprobed service does not read as a failed one, nor does a dimension resting on one weak reading pass on the strength of the prior. The Commission's framework builds the gate in, taking the lowest objective as the verdict; this index leaves it to the buyer, because which dimension is fatal depends on what the buyer runs on the service.

Composite figures are offered beside the profile, never in place of it. A composite for a kind of buyer (by size, by sector, or by what it runs on the service) is the weighted mean $V = \sum_j w_j\, E_j[v(d)]$ over the six dimensions, with its weights stated. The figure is computed only for a service with evidence beyond its supplier's own word on every dimension the kind of buyer weights: at least one reading through a document someone independent can check, an observation, a registry or a test. Each of those dimensions must also show a degree or a bound above τ. Any other service shows "not established" in its place, as the filter does.

Entering a dimension at its prior mean would give a service that has ceded the decision ½ for saying nothing. Admitting a declaration alone would give it slightly more for one sentence, since a passing reading lifts the expected degree above the prior's middle. The rule admits neither. A kind of buyer may also name dimensions and minima; the composite then ranks the services the filter says meet them. Services without a composite are listed apart, unranked.

Proposition 4 (filter and composite). Let each minimum the buyer sets be recoverable or retained, and let τ exceed the prior probabilities $\pi(d \geq m)$ and $\pi(d < m)$ of each. Under the ordering assumption, an adverse revision on any dimension, of either kind in Proposition 3, never moves a service up the order fails, not established, meets, and never raises a composite it changes. A named dimension without evidence never meets and never fails, and a weighted dimension without a degree or a bound above τ leaves the service without a composite.

Proof. By Proposition 3, an adverse revision on dimension j lowers $\Pr_j(d \geq m_j)$ and $E_j[v]$ weakly and leaves the other dimensions unchanged, so meeting becomes harder, failing easier, and V, a sum with fixed weights, does not rise. A reading that establishes a degree or a bound on a weighted dimension that had none does not revise a composite: it can only complete the evidence that the figure requires. With no evidence on a named dimension, $\Pr_j(d \geq m_j) = \pi(d \geq m_j) < \tau$ and $\Pr_j(d < m_j) = \pi(d < m_j) < \tau$. ∎

Under the uniform prior both probabilities are at most two thirds for those minima, so $\tau = 0.8$ meets the condition; a minimum of ceded constrains nothing and is not offered. Proposition 4 reworks the two-stage rule of *Scoring the Seller* (Fermigier 2026g, Proposition 1) for degrees, with one change. There the weighted mean runs over the evidenced objectives only, so a first adverse finding on an unevidenced objective could raise the mean when it landed above the others. Requiring evidence on every weighted dimension, with weights that do not move, removes that case. A brute-force check on 5,000 random buyers and profiles over six dimensions, with graded criteria, confirms the proposition, the evidence conditions included. Two services with the same figure and different evidence depth stay distinguishable through their coverage. The weights for each kind of buyer are placeholders until elicited from buyers of that kind.

The result depends on the rulebook, which every result names. Which criteria are evidence on which dimension is fixed with a version of the rulebook and changes only by issuing a new one. A supplier facing a published rule could otherwise lobby for its criteria to be reassigned at no cost in evidence. Every result names the rulebook version it was computed under.

8. Value captured per euro

The quantity. Follow one euro of the service's revenue to where it goes: wages, infrastructure, licences, other inputs, depreciation of the supplier's own capital, tax and profit. Part k has a cost share s_k and a European fraction f_k. The contribution share is

$$C = \sum_k s_k f_k, \qquad \sum_k s_k = 1, \quad 0 \leq f_k \leq 1.$$

For a part the supplier buys in, f_k is the European value-added content of the purchase: a rack rented from a European host running on imported hardware has a fraction below one. For profit, f is the share that accrues to European owners; tax counts where it is paid.

The accounting basis. The measure is an income reading. It departs from value added in trade on one point. Trade statistics count value added where it is produced (Johnson and Noguera 2012; Koopman, Wang and Wei 2014), so the profit a foreign group books in a member state is European value added however much of it leaves as dividends. On that basis an index would reward where a group books its European profit, a choice transfer pricing makes cheap. Counting profit by who receives it ties the measure to where the returns go,

which is the question the contribution family puts. Annex A gives the archetypes on both bases: the domestic-production basis raises the two US archetypes by up to their profit share and leaves the European ones unchanged.

The published result shows it through four channels: people (wages), capital (infrastructure and depreciation), inputs (licences and other purchases) and returns (tax and profit).

Two bounds and one consequence follow from the definition.

- A change of owner that moves only where tax and profit go moves C by at most $s_{\text{tax}} + s_{\text{profit}}$. Ownership can also move the inputs channel, through royalties and licence fees paid inside a group; in the archetypes of Annex A it is larger than the profit outflow for US software sold through a sales office, and smaller for the hyperscaler region.
- Moving only the infrastructure moves C by s_{infra} times the change in the infrastructure's fraction. Where a software firm hosts is a small term in the firm's contribution.
- As a consequence, a token presence earns a small share. A sales office contributes the wages of its staff and what it buys locally, so no separate test against token presence is needed: a criterion enters the contribution family by measuring one part of the split.

Five illustrative cost structures order as expected (Annex A). On round numbers chosen to be plausible, a European software firm on European hosting keeps 0.93 of each euro in Europe and the same firm on a US hyperscaler 0.88. A European operator of a licensed US stack keeps 0.64, a US hyperscaler's European region 0.43, and US software sold through a European sales office 0.19. These illustrate the measure; none is a measurement. Annex A reports how far apart they sit: perturbed by up to half, the pairs that change place are neighbours.

The two families point different ways on one supplier. Moving a European software firm to a US hyperscaler costs it about five points of contribution, between two and eight in 90 percent of the draws, and can move the service's access and continuity by a degree. A hyperscaler's European region contributes through European staff and capital spending, while the profit and the licence paid to the parent leave; the autonomy profile of that service is a separate question. The licensed stack is the case the ownership chain misses. Delos Cloud, a German subsidiary of SAP SE that runs Microsoft technology under licence (Delos Cloud, Microsoft and Arvato Systems 2024), passes a reading of its ownership chain in company registers, since every link is European. The licence shows twice: on direction, continuity and terms, because the licensor decides what the service becomes and can withdraw or reprice it, and here, as the largest outflow of each euro.

What reaches the quantity. The parts that move the share are wages, licences and profit, and the probes of Section 6 reach none of them. Groups above EUR 750 million of consolidated revenue publish employees, revenue, profit before tax and tax per member state for financial years beginning on or after 22 June 2024 (European Union 2021). Groups reporting under IFRS split non-current assets between the country of domicile and all foreign countries (IASB 2006), and groups reporting under US GAAP make the same split under its segment standard (FASB 2009). Those figures are registry-grade evidence on headcount, from which wages are estimated, on capital, and on returns as booked after transfer pricing. Member states may let a group defer commercially sensitive items for up to five years. The asset split names individually material countries and otherwise lumps all foreign countries together, so for a group headquartered outside Europe it may not isolate Europe. A small European firm is nearly settled by its register entry, which gives domicile, headcount and ownership. The hard case is the middle. The licence fees an operator pays a technology owner are unpublished, so they come as a declaration, at the weakest channel, until the supplier attaches evidence, and a supplier with a small licence outflow has reason to attach it. Group figures stand in for service figures; a published share says so.

9. The published result

The result is a profile, not a scalar. A service is reported as a radar over the six dimensions, each with its degree, the confidence in that degree and the evidence behind it, with the contribution share and its channels next to the radar. Composite figures for kinds of buyer appear with their weights, below the profile. A buyer's minimum, where one is set, sorts services into those that meet it, those that fail it and those not established.

Publishing a profile instead of a number is a cost, paid on purpose. A single scalar is what a procurement system wants, and supplying it would mean choosing which decisions may offset which. The profile refuses the choice and asks the reader to read six degrees where they wanted one.

Coverage is reported at two grains, because the two do different jobs. Dimension coverage is the share of dimensions with any evidence behind them, which tells a reader how much of the profile rests on something. Criterion coverage is the share of the criteria the rulebook scores for the service, those marked N/A excluded and the disclosure layer left out, that have evidence behind them. A supplier moves it by answering a question: one answer raises criterion coverage at once, and raises dimension coverage only where it lights up a dimension that was previously empty.

Confidence stays separate from the degree. It is never silently multiplied into it, which would fold two different quantities into one number and make the result unreadable. A dimension whose most probable degree has a confidence of 0.8 or less shows the bound the evidence establishes above that level, or else displays as not established, with the most probable degree beside it; an expired reading reverts to ∅ and lowers coverage.

Two further requirements follow from scoring named vendors on a politically charged subject. Neither is negotiable. Every degree traces down to the criteria and every criterion to the sources, with timestamp and channel; any dimension expands down to the observation that produced it. And any provider can challenge any claim: a challenged criterion enters a disputed state, displays as disputed until review resolves it, and the probe finding names the signal that fired so the dispute can address a specific observation and not a verdict. An index that scores companies and offers them no route of reply has no way to correct its own errors.

10. What the index leaves out

Evidence on no dimension. Thirty-two rows of the basis fall outside the six dimensions and contribution under every coding, because the rulebook assigns them by class: ten visibility criteria, four on environmental performance, five on programme participation, labels and language, and thirteen on the buyer's own posture. A further seven to fourteen rows, depending on the coding, sit on a dimension and are no evidence on it. Several of these criteria are legally mandatory; some matter to a buyer for other reasons. Failing to bear on the buyer's autonomy is a different fact about them.

Visibility is evidence about evidence. Access logs, architecture and dependency disclosures, a software bill of materials, notice of a change of control, audit rights: these are not decisions about the service. They are what makes the other dimensions checkable. They enter the index through the channel and confidence of the evidence they make reachable, not as scores of their own.

Environmental names two different objects. In the Commission's framework it means sustainability reporting: efficiency, hardware reuse, carbon disclosure, which sit outside the index. In the IRN, a client-side index published by the Digital Resilience Initiative (Digital Resilience Initiative 2026), it means energy dependency and continuity under hazard. The second is a continuity question. No framework in the field asks a supplier

whether it can power its own infrastructure without a dependency outside the jurisdiction. That gap in the field is the one this debate should close.

The buyer's own posture. Thirteen criteria score the buyer and not the supplier: the criticality of the asset, the alternatives available to this buyer, the concentration across its portfolio. They come from client-side indices, which are doing what a client-side index should do. A supplier-side profile feeds a buyer's own picture without standing in for it, and the two compose dimension by dimension.

Effects on other nodes. A score attached to one supplier cannot register what its adoption does to others, such as the network effects that make a market tip toward one standard; *The Contribution Criterion* shows why no per-node index can (Fermigier 2026m). The contribution family reports what a supplier does with the money it is paid. The result states that it leaves out what choosing the supplier does to the market.

11. Crosswalk to the deployed frameworks

Each framework's criteria map through the basis, which records the rows of every framework that ask each atom, onto the dimension the codings assign each atom; client-side frameworks enter through their supplier-side criteria only. SEAL's weight is spread over its questions as the calculator does, and each question's weight over the dimensions of the criteria it feeds.

Dimension	Share of SEAL's weight, mean of four codings
Access	21%
Capability	16%
Continuity	11%
Integrity	8%
Contribution	7%
Terms	6%
Direction	6%
Outside the dimensions	25%

The codings move no share by more than nine points. These shares follow where the criteria fall. The last row counts criteria outside the six dimensions and contribution; the 28 to 33 percent of Section 3 also counts criteria that sit on a dimension without being evidence on it, and is larger for that reason. Objective by objective, the strategic objective falls on direction, continuity, contribution and access, the legal objective on access, contribution and continuity, and data and AI on access. The operational objective falls on capability and access, supply chain on integrity and continuity, technology on capability, terms and continuity, and security and compliance on access, integrity and capability. Environmental sustainability falls on none. The Franco-German joint paper states dimensions without criteria, so the crosswalk with it stays at their level. Implementing and enforcing, and protecting data, map onto access and continuity; design, deployment and use onto capability and direction; substitutability onto terms; infrastructure resilience onto continuity; economic value creation onto contribution.

Eighteen frameworks are mapped criterion by criterion, thirteen from outside and five of our own. Access is the largest dimension in all but one, a set of repository badges. The Union assurance levels go furthest, with thirteen of their twenty-one criteria on access and none on terms or direction. Among the outside frameworks, the assurance levels and the repository badges ask nothing on terms or direction, and Red Hat's readiness assessment nothing on direction. Integrity is reached through supply-chain security. Five outside frameworks ask at least one criterion on contribution, about capital, intellectual property or development; none asks what share of the price stays in Europe.

The frameworks converge on access, and less fully on continuity and capability. They are thin on integrity, terms and direction. That pattern is the empirical basis for proposing a shared register of criteria and evidence in place of a shared score. Such a register identifies a supplier and a service by public register identifiers. It records each reading with the result, channel, source and date, and publishes each channel with two error rates. It fixes the dimensions, because a minimum moves when dimensions are split or merged. Each framework keeps its weights, its composites and any minimum.

12. Limits of the deployed version

The deployed version predates this rulebook. It scores eleven criteria grouped in five pillars (jurisdictional, operational, technological, economic, data) with fixed weights, and gates eligibility on the lowest pillar. It departs from this paper on four counts. Three of its criteria (governing law, data location, governance locks) are not strong evidence under every coding here. Two of its criteria both read the ownership chain, so one fact is counted twice. It builds the gate in. And it has no contribution family. Its pillars differ from the six dimensions, and mapping one onto the other is the work of its next version.

It runs one route of the two. Self-attestation opens later, so the current display rests on black-box observation alone. Five of its eleven criteria are reachable without supplier cooperation, in all but one of its pillars. The ceilings are therefore 0.45 criterion coverage and 0.80 pillar coverage.

The catalogue falls well short of those ceilings. Scored across 1,717 catalogue solutions on 15 September 2026, mean pillar coverage is 0.32 and mean criterion coverage 0.19, because an inherited entry has a country and sometimes a licence, with no hosting observation, no repository metadata and no declaration behind it. An index built the usual way would have reported a level for all 1,717 and been silent about the 0.19. A supplier raises coverage by answering, the coverage travels on the face of the result, and the incentive is legible without anyone paying for it.

Seven questions are open. The holding rates the combination rule needs are unmeasured. A review of about twenty services, whose degree is fixed by legal and technical analysis independent of the criteria, measures them, and measures the channels on real readings, which then replace the AI respondents' values. Twenty services spread over three degrees show which criteria separate degrees at all; they do not estimate each rate with any precision. The review's reference standard reads some of the facts the criteria read: ownership and control, licences, the build chain. Rates measured against a standard that includes the test overstate how well it separates, the incorporation bias of studies of diagnostic accuracy (Whiting et al. 2004). The review therefore fixes each degree before its assessors see the criterion readings, and fixes access from contracts and architecture before the registry extract is read; for the ownership chain, which is both test and standard, the bias cannot be removed and is reported. The same review records joint pass rates, which measure how far readings depart from conditional independence. Once enough services are scored, latent-class models offer a second route: they estimate the error rates of several tests read together without a reference standard, given groups that differ in prevalence (Hui and Walter 1980). The rulebook's reliability is measured with AI coders on two models; human coders are the next test. The rulebook grades a criterion on its main dimension only; the next round grades it on every dimension it bears on, which the combination rule needs and which settles how strong the field's evidence on integrity and direction is. The weights of the composite figures for each kind of buyer are placeholders until a pilot with buyers elicits them, in the same exercise that tests whether a buyer can explain a profile and whether it changes a shortlist. No criterion aimed at integrity is strong evidence under every coding, and none aimed at direction beyond a reading of the ownership chain; the criteria that would be have yet to be written and tested. The contribution share has a definition and named data sources; no

supplier has been measured. The schedule on which readings are repeated is a convention, as is the year a declaration stands; how fast hosting and control change in fact is unmeasured.

13. Conclusion

An index that scores suppliers for a buyer has three questions in front of it: what it measures, which evidence it admits, and how it combines that evidence. The frameworks published by others answer the first by listing dimensions and the third by a weighted sum or a minimum, and leave the second implicit.

The index should measure the buyer's autonomy over the service: on six decisions, who holds each, and whether the buyer can take it back. Beside that, it should report what the service contributes to European capacity, beginning with where the money goes: the European value-added content of each euro the buyer pays.

The evidence to admit is evidence that is informative, checkable and hard to fake. Fifteen criteria pass that test under every coding. They are the core of what the index collects. About three tenths of the weight in the Commission's assessment rests on criteria that bear on no decision about the service; the index displays them and does not score them. Each reading counts for what its channel can show, a supplier's own statements once, and each question requests evidence through the strongest channel that reaches it. No question aimed at who can alter the service is strong evidence under every coding. The index has to write it.

The combination is a profile, one degree per dimension with its evidence and coverage, and composites for kinds of buyer that state their weights. The gate belongs to the buyer. An index that will not say how much of the supplier it has established is asking to be trusted on a point where it has no evidence to offer.

Disclosure

The author designs the index this paper specifies and the directory that deploys it, in an engagement funded by SPRIND, the German Federal Agency for Breakthrough Innovation. The author is a co-founder and board member of the EuroStack Foundation, co-president of CNLL, which represents French open-source companies, and a co-founder and management board member of APELL. The author's company, Abilian, is a European open-source software vendor, of the kind the contribution share and the open-source criteria score well.

References

- Akerlof, G. A. (1970). The market for "lemons": quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84(3), 488–500.
- ANSSI (2022). *Prestataires de services d'informatique en nuage (SecNumCloud), référentiel d'exigences*, version 3.2, 8 March 2022.
- Asterès (2026). *From technological dependency to economic capture: the cost of cloud and software services inflation to Europe.* Study commissioned by Cigref, May 2026 (source for the survey of 54 chief information officers).
- Bross, I. (1954). Misclassification in 2×2 tables. *Biometrics* 10(4), 478–486.
- Bundesamt für Sicherheit in der Informationstechnik (2026). *Criteria enabling Cloud Computing Autonomy (C3A).* BSI, 27 April 2026.
- Delos Cloud, Microsoft & Arvato Systems (2024). *Erste souveräne Cloud-Plattform für die deutsche Verwaltung auf der Zielgeraden: Delos Cloud, Microsoft und Arvato Systems schließen finale Verträge.* Press release, Walldorf, Munich and Gütersloh, 23 September 2024. https://assets.dm.ux.sap.com/contenthub/delos-cloud/pdf/pressemitteilung-delos-cloud-microsoft-und-arvato-systems-23september2024.pdf

- DICTU (2026). *Soevereiniteit clouddiensten: Toetsingsinstrument*, v1.0.1. Ministerie van Economische Zaken, January 2026.
- Digital Resilience Initiative (2026). *Référentiel d'évaluation IRN (Indice de Résilience Numérique)*, version 1.1. Association for Digital Resilience Initiative (aDRI). Licence CC BY-NC-ND. https://digitalresilience.org/
- Dranove, D. & Jin, G. Z. (2010). Quality disclosure and certification: theory and practice. *Journal of Economic Literature* 48(4), 935–963.
- European Commission (2025). *Cloud Sovereignty Framework*, version 1.2.1, October 2025.
- European Commission (2026a). *Commission advances cloud sovereignty through strategic procurement.* News, 17 April 2026 (EUR 180M over six years, four providers, Cloud III Dynamic Purchasing System).
- European Commission (2026b). *Proposal for a Regulation on harmonised rules for cloud and artificial intelligence development (Cloud and AI Development Act).* COM(2026) 502, 3 June 2026, procedure 2026/0138(COD).
- European Union (2021). Directive (EU) 2021/2101 of the European Parliament and of the Council of 24 November 2021 amending Directive 2013/34/EU as regards disclosure of income tax information by certain undertakings and branches. *Official Journal of the European Union* L 429, 1 December 2021, 1–14. https://eur-lex.europa.eu/eli/dir/2021/2101/oj
- European Union (2023). Regulation (EU) 2023/2854 on harmonised rules on fair access to and use of data (Data Act). *Official Journal*, 22 December 2023.
- Farhi, E., Lerner, J. & Tirole, J. (2013). Fear of rejection? Tiered certification and transparency. *RAND Journal of Economics* 44(4), 610–631.
- Fermigier, S. (2026a). *Ask Me Something Hard: Why Three Fifths of Europe's Sovereignty Score Cannot Tell Suppliers Apart: Eighteen instruments measured criterion by criterion, and the eleven questions that separate a genuine supplier from a controlled one.* Draft position paper, 18 September 2026. Abilian Econ Lab. https://econ.lab.abilian.com/papers/ask-me-something-hard/
- Fermigier, S. (2026b). *An Open Internet Stack: What to Build, in What Order, and Why.* Draft working paper v0.8, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/tbb-criterion/
- Fermigier, S. (2026c). *Counting Dependencies: Aggregation Rules and What Official Dependency Indices Identify.* Draft working paper v0.1, 2026-07-18. Abilian Econ Lab. https://econ.lab.abilian.com/papers/counting-dependencies/
- Fermigier, S. (2026f). *Cuts and Options: A Theory of Strategic Autonomy in Production Networks.* Draft working paper v0.12, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/theory/
- Fermigier, S. (2026g). *Scoring the Seller: An Economic Foundation for a Supplier-Side Sovereignty Index.* Draft working paper v0.7, 2026-09-16. Abilian Econ Lab. Unpublished.
- Fermigier, S. (2026h). *Screening for Sovereignty: Optimal Procurement when Compliance is Cheap to Fake and Costly to Verify.* Draft working paper v0.20, 2026-10-21. Abilian Econ Lab. https://econ.lab.abilian.com/papers/screening-for-sovereignty/
- Fermigier, S. (2026i). *The Weakest Link and the Final Say: An Economic Commentary on the Franco-German Joint Paper on Digital Sovereignty.* Draft working paper v0.3, 2026-09-30. Abilian Econ Lab. https://econ.lab.abilian.com/papers/franco-german-commentary/
- Fermigier, S. (2026j). *Exit, Pursued by a Bill: Lock-in Rents and Portability in European Cloud and Software.* Draft working paper v0.14, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/rent-channel/

- Fermigier, S. (2026k). *Path Sovereignty: Dependence, Control, and Options in the European Digital Stack.* Draft working paper v0.19, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/power-network/
- Fermigier, S. (2026l). *A Two-Dimensional Taxonomy of Cloud Service Integrations: Modeling Architectural Coupling and Portability Friction in Managed Infrastructures.* Draft working paper v1.2, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/cloud-services-matrix/
- Fermigier, S. (2026m). *The Contribution Criterion: Scoring Policies and Organisations by Their Effect on System-Level Digital Sovereignty.* Draft working paper v0.7, 2026-09-23. Abilian Econ Lab. https://econ.lab.abilian.com/papers/contribution-criterion/
- FASB (2009). *Accounting Standards Codification Topic 280, Segment Reporting.* Financial Accounting Standards Board.
- France and Germany (2026). *Franco-German Joint Paper on Digital Sovereignty.* Joint paper of the French Direction générale des entreprises (DGE) and the German Federal Ministry for Digital Transformation and Government Modernisation (BMDS), 17 June 2026. https://www.entreprises.gouv.fr/files/files/presse/2026/20260617-franco-german-joint-paper-digital-on-sovereignty.pdf
- Gilardi, F., Alizadeh, M. & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences* 120(30), e2305016120.
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24(3), 461–483.
- Grossman, S. J. & Hart, O. D. (1986). The costs and benefits of ownership: a theory of vertical and lateral integration. *Journal of Political Economy* 94(4), 691–719.
- Hart, O. D. & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy* 98(6), 1119–1158.
- Holmström, B. & Milgrom, P. (1991). Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7 (special issue), 24–52.
- Hui, S. L. & Walter, S. D. (1980). Estimating the error rates of diagnostic tests. *Biometrics* 36(1), 167–171.
- IASB (2006). *IFRS 8 Operating Segments.* International Accounting Standards Board.
- Johnson, R. C. & Noguera, G. (2012). Accounting for intermediates: production sharing and trade in value added. *Journal of International Economics* 86(2), 224–236.
- Koopman, R., Wang, Z. & Wei, S.-J. (2014). Tracing value-added and double counting in gross exports. *American Economic Review* 104(2), 459–494.
- Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics* 108(3), 551–575.
- Lizzeri, A. (1999). Information revelation and certification intermediaries. *RAND Journal of Economics* 30(2), 214–231.
- Milgrom, P. R. (1981). Good news and bad news: representation theorems and applications. *Bell Journal of Economics* 12(2), 380–391.
- OECD & Joint Research Centre (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide.* OECD Publishing.
- Skreta, V. & Veldkamp, L. (2009). Ratings shopping and asset complexity: a theory of ratings inflation. *Journal of Monetary Economics* 56(5), 678–695.
- Vacek, P. M. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics* 41(4), 959–968.
- Whiting, P., Rutjes, A. W. S., Reitsma, J. B., Glas, A. S., Bossuyt, P. M. M. & Kleijnen, J. (2004). Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Annals of Internal Medicine* 140(3), 189–202.

The basis, the rulebook, its codings and every analysis in this paper are generated from the project's data by a verification suite. Contesting a row recomputes every quantity that depends on it.

Annex A. Contribution archetypes

Each archetype gives each part of a euro of revenue a cost share and a European fraction. The values are round numbers chosen to be plausible and are not measurements. For a bought-in part, a fraction of one where the purchase is European overstates the value-added content. The share counts profit by who receives it; the last column counts it where it is booked, as value added in trade does, on the assumption that a group books its European profit in Europe, which bounds that basis from above.

Supplier	Share	People	Capital	Inputs	Returns	Largest outflow	Domestic basis
European software firm, European hosting	0.93	0.50	0.11	0.10	0.22	other inputs, 0.04	0.93
European software firm, US hyperscaler	0.88	0.50	0.06	0.10	0.22	infrastructure, 0.05	0.88
European operator of a licensed US stack	0.64	0.30	0.20	0.04	0.10	licences, 0.30	0.64
US hyperscaler, European region	0.43	0.15	0.23	0.03	0.03	profit, 0.35	0.79
US software sold through a European sales office	0.19	0.12	0.04	0.01	0.02	licences, 0.40	0.51

Depreciation is counted as European in every archetype, which overstates it where the servers are imported: at a fraction of 0.3, the hyperscaler region falls from 0.43 to about 0.33 with the ordering unchanged. On the domestic basis the hyperscaler's region rises above the licensed stack. A group that routes its European sales through a member state where it books the profit looks more than two and a half times as European. That is the dependence on booking the income reading of Section 8 avoids.

Robustness. Every cost share and every fraction is multiplied by a factor drawn uniformly between 0.5 and 1.5, each archetype independently, over 20,000 draws; fractions are capped at one and cost shares renormalised. Seven of the ten pairs keep their order in at least 95 percent of draws. The full ordering holds in 54 percent. The three pairs that fail are neighbours: the two software firms (64 percent), the software firm on a hyperscaler above the licensed stack (93.5 percent), and the licensed stack above the hyperscaler region (94.7 percent). Treated as one firm whose hosting is the only change, the first pair keeps its order in all draws by construction, with a loss of 0.02 to 0.08 of each euro in 90 percent of them; treated as two independent firms, the hosting term is smaller than the noise in the other parts.