blackholeindex.com
ivan.savich@blackholeindex.com
ORCID 0009-0003-9542-7022 · SSRN 6555158

EVIDENCE NOTE · EUROPEAN COMMISSION · TARGETED CONSULTATION 22 AUGUST 2026

Evidence Note: Measuring Structural Provider Dependency

Specification, validation and limits of the counterfactual sensitivity analysis cited in the Black Hole Index response to the European Commission's targeted consultation on safeguarding the EU's data sovereignty.

Submitted by: Ivan Savich, Creator of the Black Hole Index (independent researcher, responding in a personal capacity)
Accompanies: Response to the targeted consultation "Safeguarding the EU's Data Sovereignty", open until 8 September 2026
Scope: Structural dependency on digital providers and practical exit capacity. This note does **not** address jurisdictional sovereignty, foreign-government access, data location or applicable law.
Declaration: I have no commercial relationship with any system assessed. All data used are publicly available. This note is non-confidential and may be published in full.
Permanent URL: blackholeindex.com/BHI-Evidence-Note-EU-Data-Sovereignty-22-August-2026.pdf
Instrument: Black Hole Index V3.1, Q3-2026 release — 184 systems across 18 sectors, eleven parameters, published rubric

1. Purpose

1.1. This note accompanies my response to the Commission's targeted consultation on safeguarding the EU's data sovereignty. Its sole purpose is to make every quantitative statement in that response independently reproducible, and to state the limits of the instrument without requiring the reader to ask.

1.2. I respond in a personal capacity as an independent researcher. I am not an EU operator and have no operational experience of cross-border data flows. This note is confined to measurement method.

1.3. Nothing here measures jurisdictional sovereignty, foreign-government access, data location, applicable law or control of encryption keys. The instrument contains no parameter for any of them.

2. The instrument, in full

2.1. The Black Hole Index expresses structural dependency as a dimensionless ratio of capture to exit capacity, computed from eleven separately scored parameters, each on a 0–10 rubric with published anchors at 0, 3, 5, 7 and 10.

2.2. **Capture:** Data Capture (d), Memory Capture (m), Action Capture (a), Process Capture (p), Network (n), Closeness (c). **Escape:** Portability (x), Substitutability (s), Human Fallback (h). **Extended:** Org Depth (o), Momentum (t).

2.3. Every parameter is normalised by dividing by 10. Writing CES_4 and CES_3 for the constant-elasticity aggregators at σ = 2, which reduce to the square of the mean of the square roots, the model is:

> $core = CES_4(d, m, a, p) \cdot synergy = 1 + 0.3 \cdot (\sqrt{(d \cdot m)} + \sqrt{(a \cdot p)}) / 2$
> $capture = c \cdot (1 + 0.5n) \cdot core \cdot synergy \cdot (1 + 0.5o) \cdot (1 + 0.3t)$
> $feedback = max(0.3, 1 - 0.35 \cdot core) \cdot escape = 0.1 + CES_3(x, s, h) \cdot feedback$
> $B = capture / escape$

2.4. Zone thresholds are 0.4, 0.7, 1.3 and 2.5. The threshold at B = 1 is definitional: capture equals exit capacity. Whether it corresponds to an observable behavioural transition is an untested empirical question.

3. Validation of the reimplementation

3.1. Before any counterfactual was run, the model above was implemented from scratch against the published rubric and the raw parameter values served by the public API. It reproduces the published score for **184 of 184 systems** to a tolerance of 0.005, with 0 mismatches.

3.2. This step matters more than it appears. Without it, a counterfactual computed from a divergent implementation would produce plausible figures that no reader could reconcile with the published dataset.

4. The counterfactual, specified before computation

4.1. In accordance with the project's measurement protocol, the intervention mapping was written down before any figure was produced.

> ***Mapping.*** *Universal high-quality portability → set the portability parameter x to its maximum for every system; leave all ten remaining parameters unchanged; recompute.*
> ***Second mapping, added on review.*** *A portability regime may also enlarge the field of alternatives, so substitutability is raised alongside portability; and as an absolute ceiling, every exit parameter is raised at once.*
> ***Prediction.*** *If capture and exit capacity are genuinely separable within the model, a material share of systems remains in the highest dependency band at x = 10.*
> ***Falsification condition.*** *If almost nothing remains above B = 1.3 at x = 10, portability is nearly sufficient for exit within the model, the proposition is unsupported by this instrument, and it is withdrawn from the submission.*

4.2. The single-parameter test alone would be open to an obvious objection: portability and substitutability are not independent, and a measure that improves the first may improve the second. That objection is met below by reporting the more demanding scenarios rather than only the most favourable one.

4.3. Maximum portability is the framework's own definition — full one-click export of all state to a competing system, including history and configuration. This is **deliberately stronger than any realistic regulatory portability measure.** The test asks not what a particular law would achieve, but whether portability alone suffices under conditions extremely favourable to it. The intermediate scenario at x = 7 is reported for sensitivity only and is not claimed to correspond to any specific policy requirement.

5. Result

SCENARIO	MEDIAN B	B ≥ 1.3	B ≥ 2.5
Baseline (as scored)	3.71	167	135
Counterfactual, x = 7	3.13	165	119
Counterfactual, x = 10	2.78	162	106

5a. The full ladder, including the scenarios least favourable to the argument

EXIT PARAMETERS SET TO MAXIMUM	MEDIAN B	B ≥ 2.5	SHARE OF THE 135 SURVIVING
None — baseline	3.71	135	—
Portability (x)	2.78	106	79%
Portability and substitutability (x, s)	2.15	72	53%
Every exit channel (x, s, h)	1.76	30	22%

5a.1. Two things follow. First, the denominator is not inert: setting every exit channel to its maximum removes 105 of the 135 systems from the highest band and halves the median index. An objection that the instrument is dominated by its capture side, and that exit capacity cannot move a score, is not supported.

5a.2. Second, 30 systems remain in the highest band even at that ceiling. It would be tempting to read this as showing that the binding constraint for those cases is the capture side rather than the provision of exit routes. That reading is **not supported**, for the reason set out in §5b, and it is withdrawn.

5b. Sensitivity to the model's feedback assumption

5b.1. The aggregation contains a term that reduces the effectiveness of exit routes as capture deepens: *feedback = max(0.3, 1 − 0.35 · core)*. It encodes a substantive claim — that deep capture degrades the routes out — and it is not a neutral scaling factor. Any result about what survives at a high level of exit capacity is therefore partly a statement about that coefficient. Recomputing with it neutralised isolates how much:

EXIT PARAMETERS AT MAXIMUM	AS PUBLISHED	FEEDBACK NEUTRALISED
None — baseline	3.71 / 135	3.02 / 115
Portability (x)	2.78 / 106	2.14 / 72
Portability and substitutability (x, s)	2.15 / 72	1.65 / 25
Every exit channel (x, s, h)	1.76 / 30	1.35 / 1

Each cell gives the median index and the count of systems at B ≥ 2.5. Neutralising the coefficient is not a claim that the model would be better without it; it is a way of showing how much of each figure the coefficient carries.

5b.2. Three conclusions follow, and the third is unfavourable to the argument this note supports.

5b.3. **The direction is robust.** The ladder is monotone under both specifications — 3.71, 2.78, 2.15, 1.76 as published, and 3.02, 2.14, 1.65, 1.35 with the coefficient neutralised. Each additional exit channel reduces measured dependency, and portability alone reduces it least.

5b.4. **The central figure is partly, but not mainly, carried by the coefficient.** Of the 106 systems remaining under maximum portability, 34 do so only because of it; 72 remain under either specification. The claim that maximising portability alone leaves a large share of systems in the highest band therefore does not depend on this assumption.

5b.5. **The ceiling residual is almost entirely an artefact of it.** Of the 30 systems remaining when every exit channel is maximised, 29 remain only because of the coefficient; with it neutralised, one does. No conclusion about capture being a binding constraint can be drawn from that figure, and none is drawn. This is reported because a reader who implements the published formula would find it, and would be right to ask why it had not been mentioned.

5.1. Returning to the single-parameter test: under maximum portability, 106 of 184 systems remain in the highest dependency band against 135 at baseline, and the median index falls by roughly 25 per cent.

5.2. Among the systems that remain, low substitutability and deeper process and organisational integration are particularly prominent characteristics relative to the population, while data and memory capture also remain elevated. This describes the surviving group; it is not an attribution of contribution to individual parameters, which would require a separate one-variable-at-a-time analysis that has not been performed.

6. Distribution across sectors, with denominators

6.1. The result is not confined to a single sector. Denominators are given because proportions alone would be misleading: a sector holding two systems can show complete persistence without that carrying meaningful weight.

SECTOR	SYSTEMS	B ≥ 2.5 AT BASELINE	STILL B ≥ 2.5 AT X = 10
Saas	24	20	16
Banks	15	12	12
Tech	12	12	12
Infra	11	11	9
Fintech	13	9	6
Social	15	8	6
Cybersecurity	5	5	5
Ecommerce	12	6	5
Gaming	11	7	5
Pharma	11	9	5
Telecom	8	7	5
Ai	10	5	4
Education Research	8	6	4
Automotive	6	6	3
Crypto Assets	10	4	3
Industrial Software	3	3	3
Healthcare It	2	2	2
Crypto Infra	8	3	1
All sectors	**184**	**135**	**106**

6.2. The defensible classification is *cross-sector within the modelled dataset*. That is a model result. It is not observational evidence that the underlying market phenomenon is systemic.

7. What was excluded, and why

7.1. An earlier version of this analysis proposed to report the observed association between the portability parameter and the index across the dataset. That association is strong, but it is **mechanically guaranteed**: portability is a component of the escape aggregate, which is the denominator of B. It is a part-whole correlation and carries no information about the world. It was removed from the evidence base before the response was drafted.

7.2. A second earlier proposal was a sector-specific worked example asserting that enterprise cloud systems combine high formal portability with high dependency. Tested against the dataset, no such case exists at the top of the scale. The claim was withdrawn rather than reformulated.

7.3. Both are recorded here because a reader is entitled to know what an author discarded, not only what survived.

8. Limitations

8.1. **Single evaluator.** All 184 scores were produced by one person. Inter-rater reliability has not been established. This is the principal limitation of the work, and until it is addressed no dependency index — including this one — should carry decisive weight in a policy decision.

8.2. **Contestable design assumption.** The result depends on exit channels being imperfect substitutes for one another. A reviewer who rejects that assumption would obtain a different answer.

8.3. **Thin validation cohort.** The published validation rests on 14 systems with observed retention data.

8.4. **Ranks are robust, bands are not.** Under simulated ±1 scoring perturbation, rank ordering is stable (Spearman ρ 0.962–0.964) while 9.3–9.6 per cent of systems change band per draw. This is a sensitivity simulation, not inter-rater reliability.

9. How to reproduce every figure above

9.1. Download the pinned dataset for this note: **blackholeindex.com/bhi-scores-q3-2026.csv**. It carries one row per system with all eleven raw parameters, the published index and the band. Its SHA-256 is `576c93ca3755aee652220ca507c5c0ca93db7db12441f31d8ad58630ce853265`. Implement the model set out in §2, confirm agreement with the published index column, then set x to 10 for every row, leave the other ten parameters untouched, recompute, and count rows by band. Repeat for the further scenarios in §5a and §5b. No private data, credential or permission is involved at any step.

9.2. **Use the pinned file rather than the live endpoint.** The index is re-scored quarterly, and blackholeindex.com/api/bhi always serves the current release. A reader who follows these steps against the live endpoint after the next quarterly release will obtain different figures and would be entitled to conclude that this note is wrong. The pinned file is the Q3-2026 release as it stood when this note was written, and later releases do not overwrite it. As at the date of this note the two agree exactly, parameter for parameter.

Ivan Savich

CREATOR, BLACK HOLE INDEX · INDEPENDENT RESEARCHER

ivan.savich@blackholeindex.com · blackholeindex.com · ORCID 0009-0003-9542-7022

References

Black Hole Index (2026). Methodology, scoring rubric and aggregation function. blackholeindex.com/methodology

Black Hole Index (2026). Robustness note: scoring-noise Monte Carlo and aggregation-form ablation, Q3-2026 dataset. blackholeindex.com/robustness

Black Hole Index (2026). Validation and sensitivity. blackholeindex.com/validation

Savich, I. (2026). Response to the CMA's Consultation on a Proposed Steering Conduct Requirement for Apple's Mobile Platform. 21 July 2026. Published by the Competition and Markets Authority, 14 August 2026.

European Commission (2026). Commission reaches preliminary position that Amazon's and Microsoft's market leading cloud services should be designated under the DMA. 25 June 2026.

European Commission (2026). Commission imposes interim measures on Meta to preserve free access to WhatsApp for rival AI assistants. IP/26/1276, 9 June 2026.

blackholeindex.com · CC-BY 4.0