

When a Proxy Becomes the Target It Stops Being a Proxy

— Four Things Break the Link, and the Fence Is the Tail of the Error — [Paper 186]

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

August 2026

— *From thinking to syncing.* —

ABSTRACT

A measure that becomes a target stops being a good measure — and this is usually explained by saying that people game the system. This paper writes down three ways it happens with nobody there at all. **NO NEW MATHEMATICAL THEOREM AND NO NEW STATISTICAL METHOD IS CLAIMED.** First, selecting on a noisy proxy picks up the noise. With a noise of unit size, taking the top thousandth by proxy gave a mean target of **2.378** against the **3.359** obtained by selecting on the target itself (Section 2). Second, this is the fence. With a light-tailed error, optimising the proxy harder always improves the target. With a heavy-tailed one there is a best amount of optimisation, and past it things get worse — under a Cauchy error the peak is **0.528** at the top fifth, and by the top ten-thousandth it is **0.005**, no better than choosing at random (Section 3). Third, the correlation itself stops working in the tail. Even at a correlation of **0.99** the agreement at the top fell steadily from **0.901** to **0.765** (Section 4). Fourth, and this is the core. A proxy predicts the target and does not control it. At a correlation of **0.995**, observing the proxy at **2.5** gives a mean target of **2.477**, while writing the proxy to **2.5** gives **−0.0001** — the gap is the whole of the value (Section 5).

Scope of this paper: **NO NEW MATHEMATICAL THEOREM AND NO NEW STATISTICAL METHOD IS CLAIMED** — selection bias, the absence of tail dependence for a gaussian pair, and the distinction between observing and intervening are all standard. Nothing is claimed about any institution, organisation or policy — what is treated is only the mathematics between a proxy quantity and a target quantity. On Paper 183: that paper showed one table carrying two correct answers and wrote that an arrow decides. The fourth root here uses that machinery unchanged — not a new pattern but a report that the same pattern appears at the proxy. On a word: throughout this corpus 指標 has meant a character in representation theory (Papers 70, 152, 181); the proxy here has nothing to do with it. What is added is separating the four roots by what breaks the link, measuring that the fence lies in the tail of the error rather than its size, following the absence of tail dependence across five decades, and setting observation and intervention side by side in numbers.

1. Notation and position

Call the quantity one actually wants to improve the **target** and the quantity one measures the **proxy**. Throughout, the target is standard normal and the proxy is the target plus an error. Everything sampled was sampled for this paper, with sample sizes stated.

Position: Paper 164 wrote that “probability” has four addresses and Paper 183 showed one table carrying two correct answers. This paper stands beside them and measures the gap between measuring and improving.

2. Selecting on a noisy proxy picks up the noise

Verification 2.1: Target standard normal, proxy the target plus an error, two million points.

error size	top	mean target, chosen by proxy	chosen by target
0.0	0.1%	3.359	3.359
0.5	0.1%	2.996	3.359
1.0	10%	1.241	1.755
1.0	1%	1.890	2.667
1.0	0.1%	2.378	3.359
2.0	0.1%	1.503	3.359

(QED)

Observation 2.2

The harder the selection, the larger the share of noise among those selected. There is nobody here — no one gaming anything, no institution, no intent. It happens because the proxy has an error.

3. The fence is the tail of the error, not its size

Verification 3.1: The same set-up with a normal error (light tail) and a Cauchy error (heavy tail), pushing the selection harder.

top	normal error	Cauchy error
50%	0.566	0.418
20%	0.991	0.528
10%	1.244	0.336
1%	1.885	0.024
0.1%	2.383	−0.007
0.01%	2.752	0.005

(QED)

Observation 3.2

With a normal error, optimising harder always improves the target: the column is monotone. With a Cauchy error the peak is at the top fifth, and beyond it the target falls, reaching 0.005 at the top ten-thousandth — no better than choosing at random.

Observation 3.3

What separates them is not the size of the error. In Section 2 the noise was raised to 2.0 and the column stayed monotone. What separates them is the tail — with a heavy tail the very top of the proxy is occupied not by cases with a high target but by cases whose error happened to be enormous.

Observation 3.4

So both “optimising is good” and “optimising is dangerous” are half-sentences. The full one is that how far one may optimise is settled by the tail of the error. The fence stands in the same place, for the same reason, as the one where Paper 157 broke least squares with a Cauchy error.

4. The correlation itself stops working in the tail

Verification 4.1: Draw eight million gaussian pairs with correlation ρ and measure how often Y is in the top q given that X is.

ρ	top 10%	top 1%	top 0.1%	top 0.01%
0.50	0.325	0.128	0.057	0.015
0.90	0.689	0.542	0.442	0.381
0.99	0.901	0.850	0.810	0.765

Every row rests on at least 800 samples. (QED)

Observation 4.2

Even at 0.99 the agreement keeps falling towards the tip. A gaussian pair has no tail dependence — for any ρ below one the agreement tends to 0 in the limit. This computation does not reach that limit and does not claim to. What it shows is that five decades in, the agreement is still falling.

Observation 4.3

A correlation is measured in the middle. Optimising carries you to the tip. Where it was measured and where it is used are different places — this is not even about error, it is about regime.

5. The core: a proxy predicts and does not control

Observation 5.1

Suppose the proxy **arises from** the target, the arrow running from target to proxy with $P = T + 0.1\epsilon$.
Observationally the proxy predicts the target well, at a correlation of **0.995**.

Verification 5.2: Four million points.

value of P	samples	observed $E[T \mid P]$	intervened $E[T \mid \text{do}(P)]$
1.0	38645	0.989	−0.0001
2.0	8912	1.977	−0.0001
2.5	2789	2.477	−0.0001

(QED)

Observation 5.3

Seeing a high proxy means the target is high. Writing a high proxy moves the target not at all — the arrow runs from the target to the proxy. The gap is 2.477 against −0.0001, which is the whole of the value.

Observation 5.4

This pattern is not new to this corpus. Paper 183 showed one table carrying two correct answers and wrote that what decides is an arrow rather than an entry in the table. All that is added here is that the same pattern appears at the proxy.

6. The census: what breaks the link

Observation 6.1

root	what breaks it	how it shows	verified here
regressional	the error itself	selection picks up noise	Sections 2, 3
extremal	leaving the measured regime	correlation fails in the tail	Section 4
causal	the direction of an arrow	predicts but does not control	Section 5
adversarial	somebody is there	the proxy is moved directly	not measured

Observation 6.2

Three of the four happen with nobody there. The explanation that people game the system fits only the fourth. The first three need no intent and no agent: error, regime and an arrow are enough.

Observation 6.3

And the four call for four different responses. The regressional root calls for reducing the error or deciding where to stop optimising, the fence being the tail. The extremal root calls for measuring again in the regime one will use. The causal root calls for measuring by intervention — no amount of observation will do. The adversarial root calls for reasoning about the other party. One name over four things makes one remedy seem enough.

7. Conclusion

1. Selection picks up the error (Ver. 2.1): **2.378** against **3.359** at the top thousandth.
2. The fence is the tail (Ver. 3.1): monotone for a normal error, peaking at the top fifth for a Cauchy one.
3. Correlation fails in the tail (Ver. 4.1): even $\rho = 0.99$ falls from 0.901 to 0.765.
4. Predicting is not controlling (Ver. 5.2): **2.477** against **−0.0001**.
5. Three of them need nobody (Obs. 6.2).

Closing: When a measure is said to have broken the moment it became a target, the next question is what broke the link. The error, the regime, an arrow, or an opponent. And in the first three there was never anybody trying to game anything.

References

1. Goodhart, C. A. E. (1975). *Problems of Monetary Management: The U.K. Experience*. Papers in Monetary Economics, Reserve Bank of Australia.
2. Campbell, D. T. (1979). *Assessing the Impact of Planned Social Change*. Evaluation and Program Planning, 2(1).
3. Sibuya, M. (1960). *Bivariate Extreme Statistics*. Annals of the Institute of Statistical Mathematics, 11(3).
4. Capen, E. C., Clapp, R. V., Campbell, W. M. (1971). *Competitive Bidding in High-Risk Situations*. Journal of Petroleum Technology, 23(6).

5. Pearl, J. (1995). *Causal Diagrams for Empirical Research*. Biometrika, 82(4).
6. Manheim, D., Garrabrant, S. (2018). *Categorizing Variants of Goodhart's Law*. arXiv:1803.04585.
7. Yamagishi, Y. (2026). *$\hbar$ Is Not on the Mathematical Side of the Inequality* [Paper 129]. Zenodo. 10.5281/zenodo.22079681
8. Yamagishi, Y. (2026). *The KL Divergence Is Not a Distance, but Its Second Derivative Is a Metric* [Paper 137]. Zenodo. 10.5281/zenodo.22081369
9. Yamagishi, Y. (2026). *Measuring Overlap, or Measuring Displacement* [Paper 139]. Zenodo. 10.5281/zenodo.22081801
10. Yamagishi, Y. (2026). *A Census of the Twos* [Paper 157]. Zenodo. 10.5281/zenodo.22091931
11. Yamagishi, Y. (2026). *Probability Has Four Addresses* [Paper 164]. Zenodo. 10.5281/zenodo.22092815
12. Yamagishi, Y. (2026). *The Fence around Least Squares Is Linearity, Not Normality* [Paper 171]. Zenodo. 10.5281/zenodo.22094287
13. Yamagishi, Y. (2026). *“The Curse of Dimensionality” Is Not One Curse* [Paper 182]. Zenodo. 10.5281/zenodo.22104170
14. Yamagishi, Y. (2026). *One Table, Two Correct Answers* [Paper 183]. Zenodo. 10.5281/zenodo.22104230
15. Yamagishi, Y. (2026). *Incompatibility Begins with the Third Option* [Paper 184]. Zenodo. 10.5281/zenodo.22105459
16. Yamagishi, Y. (2026). *What Classifies a Defect Is the Target Space* [Paper 185]. Zenodo. 10.5281/zenodo.22105496

— *From thinking to syncing.* —

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)