Rebelative

MINI WHITE PAPER · H2 2026

METHODOLOGY · RESEARCH STANDARDS

Causation Before Backtests: Why Strong Alt-Data Results Deserve Scrutiny

A strong backtest can reflect a dataset's history as much as a signal's power. Association is not a mechanism, and in alternative data the difference is expensive.

The dataset has a history, and the backtest inherits it

Commercial alt-data panels are living datasets. As massive data panels grow and churn, histories get backfilled, stale names could fall away, and extrapolation weights are recalibrated, often as part of routine maintenance to keep a panel stable, continuous, and usable. Evaluation has to account for these technical updates. A backtest run on today's version of history could inherit **look-ahead and survivorship artifacts**, and could look strongest precisely where the artifacts were largest. Then add the research community's multiple-testing problem: hundreds of signal variants evaluated, the best one reported, a challenge every quant team (ours included) must manage. A spurious association can survive every conventional check while encoding no economic content at all.

Rising markets obscure the picture further. Beta and momentum carry weak names, flattering signals that would fail in a flat or stressed tape. That is one more reason regime breakdowns belong in every evaluation.

Mechanism first, backtest second

Recent work on causal factor investing (López de Prado and others) reaches the same conclusion from theory: without a causal hypothesis, factor research produces associations that fail out-of-sample. Our publication guidelines therefore require, for every signal:

- **A documented economic mechanism.** Why should invoice delinquency precede distress? Which cash-flow channel connects the panel to the outcome?
- **True point-in-time reconstruction,** including vendor-version control: signals are tested on the data as it existed then, not as it was later restated.
- **Explicit treatment of revenue-recognition timing,** i.e. how transaction data maps to reported revenue across companies, industries and geographies.
- **Out-of-sample evaluation by sector and regime,** with robustness (not peak performance) as the primary KPI.
- **Multi-vendor corroboration:** no single panel determines a signal, and sources that only "work" in-sample automatically lose weight.
- **Published limitations and disclaimers:** where the analysis applies, where it does not, and why backtests are not predictions.

Why credit raises the stakes

Equity mispricing costs relative performance; credit mispricing costs principal. Distress signals are consumed at the tail of the distribution, exactly where over-fit associations break first. A credit signal must therefore explain itself (mechanism, lineage, out-of-sample behavior) before it earns a place in the workflow. That is the standard we build to, and the standard buyers should demand of any provider, ourselves first.

TALK TO US **hello@rebelative.com** · rebelative.com · Rebelative

FOLLOW US New research, company news and product updates: https://www.linkedin.com/company/rebelative