

Turning a Telescope on Quant: Part B – The Canals of Mars

AI, False Patterns and the Risk of Seeing Too Much

In [Part A](#), astronomy gave us the hopeful version of low signal-to-noise: stack enough weak observations and hidden structure can become visible. Part B is the warning label. When evidence is faint, humans and models can also invent structure where none exists. Along the way, we'll peer at Martian civilizations through Victorian telescopes, count how many canals a computer can find in pure noise, see what AI does to that count, and ask who is most likely to build them.

Seeing Canals

Low signal-to-noise environments are dangerous because they hide true patterns and encourage false ones. When evidence is faint or ambiguous, the mind fills the gaps. A flexible model can do the same, only faster and with greater confidence.

Astronomy learned this lesson the hard way. In 1877, Giovanni Schiaparelli, director of the Brera Observatory in Milan, used a close approach of Mars to map its surface and recorded a network of faint straight lines he called *canali*, or channels (Figure 1). In English the word became “canals,” which implies that someone dug them. Over the following decade Schiaparelli mapped more of them, and reported that some had split into parallel pairs.

The man who did the digging was Percival Lowell, a wealthy Bostonian who in 1894 built an observatory in Flagstaff, Arizona, with his own money, largely to study the canals. In three popular books he turned faint lines into the irrigation network of a dying civilization, carrying water from the polar caps. The telescope offered ambiguity; Lowell supplied the narrative, and paid for it.¹

Incorrectly identifying canals was not a failure of intelligence. It was perception failing in a low signal-to-noise regime. The eye saw features at the threshold of detectability, the brain

¹ Lowell's Mars also gave us some classic science fiction: H. G. Wells's *The War of the Worlds* (1898), Edgar Rice Burroughs's Barsoom novels, and some less stellar movies... *John Carter*, anyone?

imposed structure, and culture amplified the result. Once the story existed, observers knew what they were looking for. Bias entered before measurement, not after it.

It took a controlled experiment to show where the lines came from. In 1903, Edward Maunder and Joseph Evans at the Royal Observatory, Greenwich, asked classes of schoolboys to copy discs marked with scattered smudges and dots, but no lines. Seated far enough back, the boys drew canals: their eyes joined the dots, just as Lowell's had. Mariner 4's first close-up photographs, in 1965, showed craters, not canals.

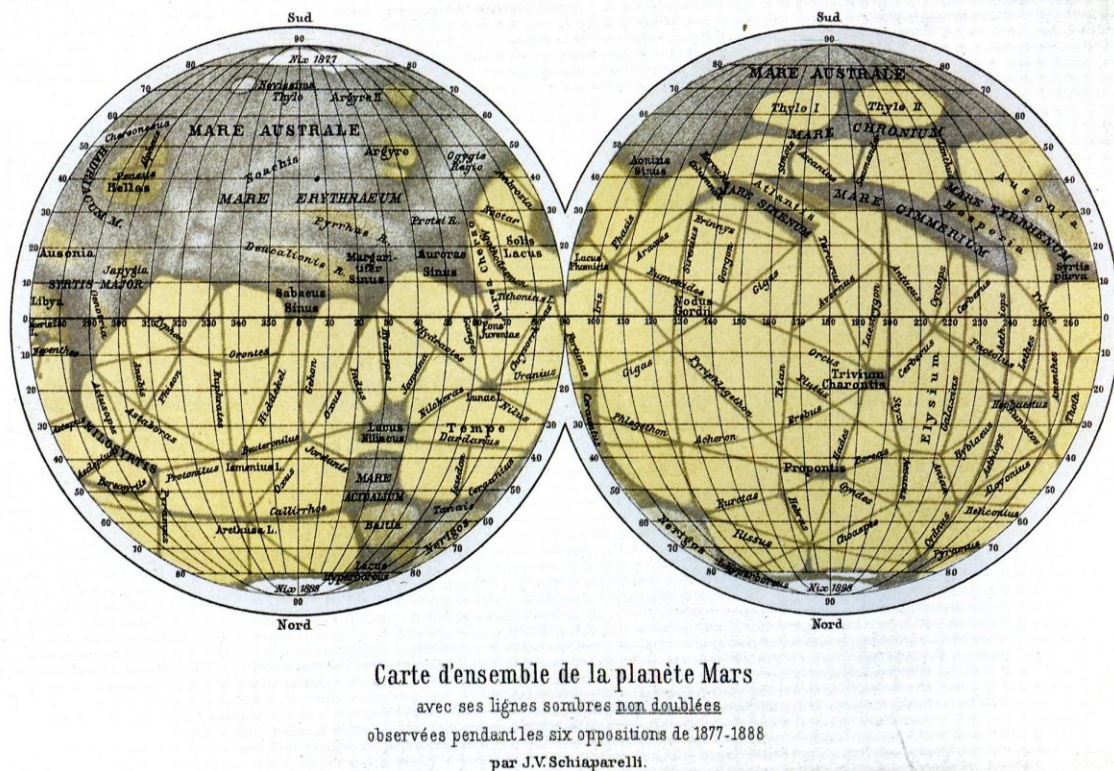


Figure 1. Schiaparelli's map of Mars, from observations over the six oppositions of 1877–1888: the dark regions he took for seas, joined by the straight *canali*. South is at the top. Source: Wikimedia Commons, public domain.

Nor is this a Victorian failing. In 2014, the BICEP2 team announced that swirls in the cosmic microwave background were the imprint of gravitational waves from the Big Bang; within a year, a joint analysis with the Planck satellite showed the signal was largely dust in our own galaxy. In December 2015, both large experiments at CERN reported a small excess of photon pairs at an energy of 750 GeV. The experiments were careful not to claim a discovery, but theorists wrote around 500 papers explaining the bump before more data made it vanish in the summer of 2016. A paper explaining a new particle is worth writing; a paper saying “probably noise” is not.

The Canals of Finance

Quantitative finance is built on faint evidence. At the portfolio level, a strategy with a true Sharpe ratio of 0.5 needs about 16 years of data before its record clears the conventional hurdle of a t-statistic of two, and 36 years to clear the stricter t-statistic of three that Harvey, Liu and Zhu (2016) proposed for new factors. Very few strategies, except perhaps trend or carry, can point to persistent behavior of over 200 years (e.g. [Lempérière et al., 2014](#)). In astronomical terms, image exposure is too short: few markets have that much clean history, regimes change within it, and every discovered pattern risks being traded away or crowded. This is fertile ground for canal-building. A backtest can become a telescope pointed at noise; a parameter search, an observatory for wishful thinking.

Humans have always been the main source of this risk. We begin with theories, hunches and institutional biases, then use statistics to test whether any of them survive contact with data. Good research practice was designed to discipline the human tendency to see what it wants to see: out-of-sample testing, pre-specified hypotheses and independent validation.

How Many Canals Can You Find?

To see how easily noise becomes conviction, we simulated trading strategies whose true Sharpe ratio is exactly zero, pure coin flips, and gave each 20 years of daily returns. Then we played researcher: search N of them, keep the one with the best Sharpe ratio over the first ten years, and see how it does over the next ten. Figure 2 shows the result.

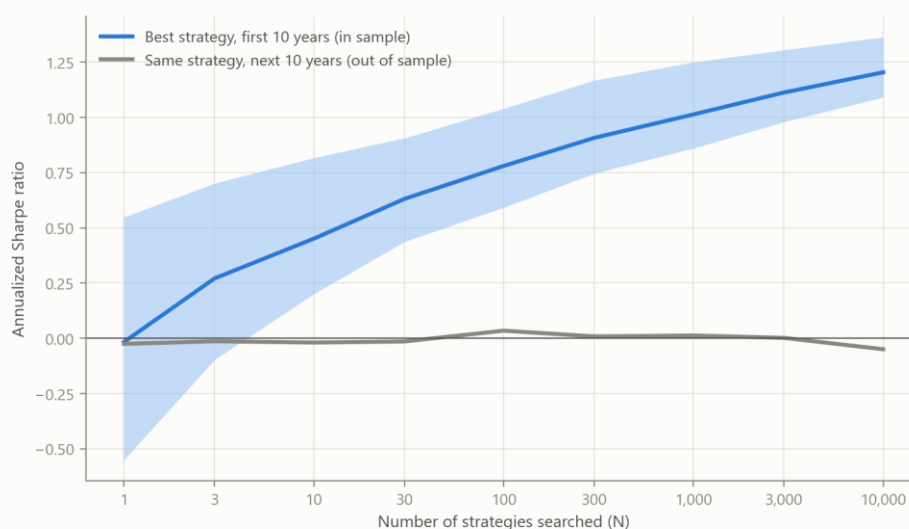


Figure 2. Simulated strategies with a true Sharpe ratio of zero. The line shows the median, across 400 searches, each over N fresh strategies, of the best in-sample Sharpe ratio over ten years; the band spans the 5th to 95th percentiles. The gray line is the same strategy's Sharpe ratio over the following ten years. Source: GQ, simulation.

The best of 100 has a median Sharpe ratio of 0.8 in sample, a t-statistic of about 2.5 over ten years, and nine times out of ten it clears the conventional hurdle of two. The best of 1,000 has a Sharpe ratio of 1.0, a t-statistic of about 3.2, and usually clears even the stricter hurdle of three. Over the following ten years, every one of them earns what it was always going to earn: nothing. In practice the trials in a parameter grid are correlated, so its effective N is smaller than its count; but a research program that has been trying ideas for a decade has a larger N than it remembers. And correlation shrinks the count, not the danger: every trial runs on the same history, so the noise they share is what the search selects for. It is Part A's stacking run in reverse, accumulating the shared error rather than the signal.

AI Canals

AI adds two ways of fooling yourself. The first issue is over-parametrization. Quant models used to be made of relatively few parameters, usually stacked linearly. Each component could be tested and understood on its own. AI has made it possible to combine thousands of parameters in a non-linear fashion, paving the way to financial hallucinations. The second is multiplicity: try many simple models and keep the best. This is the risk that AI multiplies. A hypothesis that used to take an analyst a week to code and test now takes minutes, so the number of trials explodes while the length of history does not. In our simulation, going from 100 to 10,000 trials lifts the best ten-year Sharpe ratio found in pure noise from 0.8 to 1.2. And the same model that ran the search will write a fluent, plausible story for whichever strategy wins. It supplies both the canals and the civilization.

None of this makes AI the enemy. The same tools can run the controls, placebo tests and out-of-sample checks that people skip because they are tedious. But that discipline has to be designed into the process; it will not arrive on its own, and the introduction of AI demands a full redesign of the quant process.

Who Builds Canals?

Surviving a low signal-to-noise world takes three things at once: statistical rigor, the incentive to be right, and the ability to implement. Figure 3 maps what tends to go wrong when one or two of them are missing.

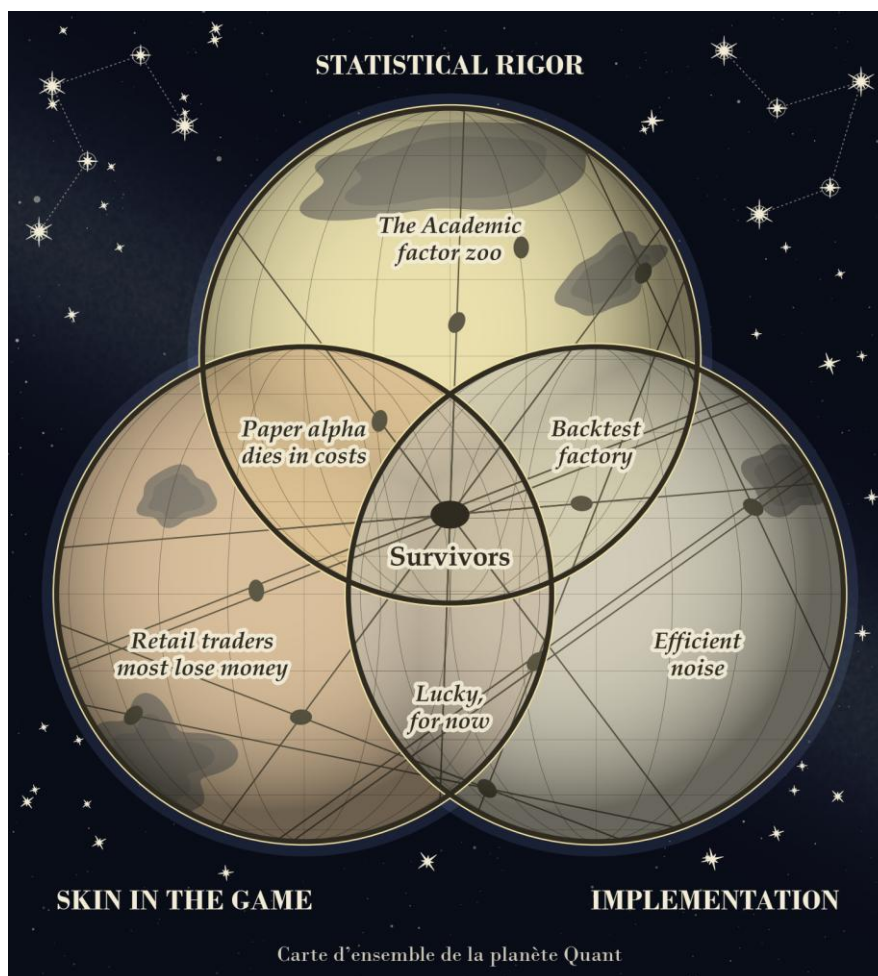


Figure 3. Three things a systematic strategy needs to survive, and what happens with only one or two of them, drawn in the manner of Schiaparelli's map. Source: GQ.

Academic finance has the statistics, but its incentives reward finding patterns rather than filtering them out. Very few journals publish a failed test, just as nobody runs a “No Life on Mars” headline. Harvey, Liu and Zhu (2016) counted 316 published factors that claimed to predict stock returns. When Hou, Xue and Zhang (2020) tried to replicate 452 anomalies, 65% failed even the conventional hurdle. McLean and Pontiff (2016) found that the returns to 97 published predictors fell by 26% out of sample and by 58% after publication. The further you sit from running money, the less it costs you to be wrong.

Retail traders have the right incentive, since it is their own money, but rarely the tools. Of the Brazilian day traders who kept at it for more than 300 days, 97% lost money, and only 1.1% earned more than the minimum wage (Chague, De-Losso and Giovannetti, 2019). Writing the strategy in code does not fix this. Quantopian, a platform for crowd-sourced systematic strategies, studied 888 of its users' algorithms and found that backtest Sharpe ratios

explained less than 2.5% of the variation in their out-of-sample Sharpe ratios (Wiecki et al., 2016).

Rigor and implementation without the incentive make a backtest factory: a platform that can build and run strategies but is paid on the assets they gather rather than on whether they work.

The third circle is the one most often left out. Complex models can find patterns that do not exist, and they can also find something real but too fragile to trade. McLean and Pontiff's post-publication decay was largest for predictors concentrated in illiquid stocks with high idiosyncratic risk, where trading is most expensive. Astronomers have a word for what stands between a real object and a sharp image: seeing, the air that blurs the light on its way to the telescope. In trading, it is cost. The costs that kill a strategy, such as turnover, market impact and capacity, are rarely in the backtest that sold it.

Don't Fund the Canals!

Pattern recognition is where science and investing both begin, and the Martian canals show what it does without control. Stacking weak signals until something real emerges is good magic; mistaking a coherent story for a coherent signal is not. Low signal-to-noise data demand humility: validation and a bias toward robustness.

Lowell had the telescope, the money and the conviction. What he lacked was a control: someone to seat the schoolboys at the back of the room. When we reviewed the first decade of GQ trading, we ran that control on ourselves: the configuration our first strategy launched with in 2017, frozen and run forward over years it had never seen, beside the live record. AI will make that discipline more important. It will help researchers search more widely and test hypotheses more quickly, and it will draw elegant canals wherever ambiguity allows them. The future of quant research belongs to the process that knows which patterns are pure make-believe, and which deserve funding.

References

- ATLAS Collaboration (2016), "Search for Resonances in Diphoton Events at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Journal of High Energy Physics* 2016(9).
- BICEP2 Collaboration (2014), "Detection of B-Mode Polarization at Degree Angular Scales by BICEP2", *Physical Review Letters* 112(24).
- BICEP2/Keck Array and Planck Collaborations (2015), "Joint Analysis of BICEP2/Keck Array and Planck Data", *Physical Review Letters* 114(10).
- Chague, F., De-Losso, R. and Giovannetti, B. (2019), "Day Trading for a Living?", working paper, University of São Paulo.
- CMS Collaboration (2016), "Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s} = 8$ and 13 TeV", *Physical Review Letters* 117(5).
- CMS Collaboration (2017), "Search for High-Mass Diphoton Resonances in Proton-Proton Collisions at 13 TeV and Combination with 8 TeV Search", *Physics Letters B* 767.

- Harvey, C. R., Liu, Y. and Zhu, H. (2016), "...and the Cross-Section of Expected Returns", *Review of Financial Studies* 29(1).
- Hou, K., Xue, C. and Zhang, L. (2020), "Replicating Anomalies", *Review of Financial Studies* 33(5).
- Lempérière, Y., Deremble, C., Seager, P., Potters, M. and Bouchaud, J.-P. (2014), "Two Centuries of Trend Following", *Journal of Investment Strategies* 3(3).
- Maunder, E. W. and Evans, J. E. (1903), "Experiments as to the Actuality of the 'Canals' Observed on Mars", *Monthly Notices of the Royal Astronomical Society* 63(8).
- McLean, R. D. and Pontiff, J. (2016), "Does Academic Research Destroy Stock Return Predictability?", *Journal of Finance* 71(1).
- Wiecki, T., Campbell, A., Lent, J. and Stauth, J. (2016), "All That Glitters Is Not Gold: Comparing Backtest and Out-of-Sample Performance on a Large Cohort of Trading Algorithms", *Journal of Investing* 25(3).

Important Disclosures

This material is for general informational and educational purposes only and is not investment, legal, tax, accounting, or other professional advice; an offer or solicitation to buy or sell any security, commodity interest, derivative, advisory service, or private fund interest; or a recommendation regarding any instrument, strategy, or product. It does not consider any reader's objectives, financial circumstances, risk tolerance, restrictions, tax status, or other needs. Investment decisions should be made only after consulting appropriate advisers and reviewing applicable offering, advisory, account, and risk materials.

Trading commodity interests, including futures, options, swaps, and other derivatives, is speculative and involves a high degree of risk. These instruments may be volatile, leveraged, or illiquid and may cause substantial losses, including losses exceeding the amount invested where applicable. Using derivatives to increase notional exposure or combine return streams can magnify losses, liquidity demands, financing costs, and adverse-correlation effects. Past performance is not necessarily indicative of future results.

Sharpe ratio is a risk-adjusted measure calculated by subtracting a risk-free rate from an investment's return and dividing by return volatility over the applicable period. A higher ratio generally indicates more return per unit of measured volatility. Results depend on the period, return frequency, risk-free rate, annualization, fees and expenses, and return distribution; comparisons may not be meaningful unless methodologies are consistent. Sharpe ratio does not capture all risks, including liquidity, leverage, concentration, tail, counterparty, and operational risk.

References to trend following, return stacking, thematic or factor exposures, signals, portfolio construction, Sharpe ratios, and related concepts describe general observations and research considerations only and, unless expressly stated, do not describe the actual or expected performance, risk, holdings, trading, or investment process of any Gresham product, fund, account, or strategy.

Any signals, factors, strategies, backtests, modeled or hypothetical results, academic studies, numerical examples, or other non-actual results are illustrative only and, unless expressly stated, do not represent actual Gresham trading. Such information has inherent limitations, may use hindsight or assumptions that prove inaccurate, and may omit material economic, market, and implementation effects, including fees, expenses, taxes, financing, transaction costs, slippage, liquidity, position limits, changing correlations, market impact, and operational constraints. No investor, account, fund, or strategy is represented as likely to achieve similar results.

Gresham, its affiliates, clients, funds, accounts, or personnel may hold or trade, directly or indirectly and long or short, instruments, sectors, themes, issuers, or strategies discussed herein. Positions and exposures may change without notice, and this material does not describe all such activity.

Information from academic research, market commentary, index providers, data vendors, public websites, and other third parties is believed reliable but has not been independently verified; no representation is made as to its accuracy, completeness, timeliness, or reliability. References and hyperlinks are for convenience and context, do not imply endorsement, adoption, or sponsorship, and Gresham is not responsible for third-party website content, availability, or privacy practices.

Forward-looking statements, identified by terms such as "may," "will," "should," "expect," "anticipate," "project," "estimate," "intend," "continue," or "believe," involve risks and uncertainties; actual events, results, or performance may differ materially. Nothing herein is a guarantee, promise, assurance, or representation as to the future.