

From Trading Rules to Reliable Backtests

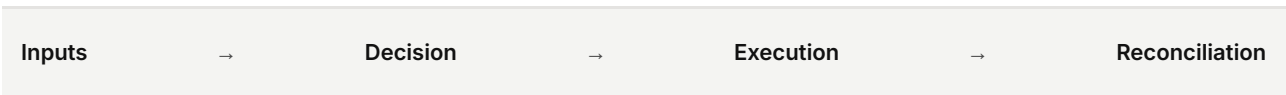
Demonstration report

Abstract

Reliability in a backtest is agreement between a stated rule, an executable timeline and reconciled accounting, not merely an attractive return. Evaluation separates data eligibility, decision timing, execution feasibility and cash/position reconciliation. This methodological worked example examines those dependencies rather than measured strategy performance.

The evaluation contract

A trading rule becomes testable when its observable inputs and permitted information set are explicit. Each observation needs an event timestamp and an availability timestamp: a price stamped at the close is not necessarily usable before that close. Eligibility also records missing values, trading calendars and corporate-action treatment. A signal can depend only on information available at the decision time.



The position rule maps eligible observations and current holdings into a target quantity. It must distinguish a target position from an incremental order; otherwise a repeated signal can accidentally compound exposure. The contract specifies when a decision is made, when an order may be submitted and which fill convention is admissible. A next-bar fill is a convention to be evaluated, not a claim of executable liquidity.

Execution converts orders into quantities, prices and costs. The cost model identifies the fee base and applies each charge once. A cash ledger then records every debit and credit while a position ledger records shares acquired or released. Invariant checks require that ending holdings equal starting holdings plus signed fills, and that cash movement equals signed notionals less explicit costs. Valuation is separate from realized cash: an open position needs a marked value, not an assumed sale.

What the contract establishes

These checks assign failures to their source. A timeline violation is not repaired by correct arithmetic, and a balanced ledger does not prove that a modeled fill was feasible. The contract makes assumptions inspectable before economic interpretation begins.

From Trading Rules to Reliable Backtests

A round trip with explicit costs

Worked hypothetical example. Start with \$20,000 cash and no shares. Buy 100 shares at \$100, then sell all 100 at \$102. The fee rate is 0.05% of notional on each side. Slippage, financing and taxes are ignored solely for this worked case; the fill prices are stipulated.

Transaction	Shares held	Notional	Fee	Cash balance
Opening balance	0	—	—	\$20,000.00
Buy 100 at \$100	100	\$10,000.00	\$5.00	\$9,995.00
Sell 100 at \$102	0	\$10,200.00	\$5.10	\$20,189.90

Reconciliation

Buy notional = $100 \times \$100 = \$10,000$

Buy fee = $\$10,000 \times 0.0005 = \$5.00$

Cash after buy = $\$20,000 - \$10,000 - \$5 = \$9,995$

Sell notional = $100 \times \$102 = \$10,200$

Sell fee = $\$10,200 \times 0.0005 = \$5.10$

Final cash = $\$9,995 + \$10,200 - \$5.10 = \$20,189.90$

Account measure	Reconciled value
Gross profit	\$200.00
Total costs	\$10.10
Net profit	\$189.90
Net account return	0.9495%

Net profit = $\$200 - \$10.10 = \$189.90 = \$20,189.90 - \$20,000$. Net account return = $\$189.90 / \$20,000 \times 100 = 0.9495\%$. Ending shares reconcile to $0 + 100 - 100 = 0$; no residual asset value is needed.

The denominator is part of the claim

The account return uses initial cash, not the \$10,000 invested notional. Dividing by invested notional answers a different question about deployed capital and would overstate this account-level percentage. A price series alone cannot establish tradable performance: it does not specify information availability, order execution, capital committed or costs.

From Trading Rules to Reliable Backtests

Diagnostic checks and failure attribution

Test	Deliberately introduced error	Expected check
Signal timing	Same-close execution after a close-only signal	Timing invalid
Cash accounting	Duplicate fee deduction	Cash fails \$189.90 net reconciliation
Price basis	Mix adjusted and unadjusted prices	Corporate-action consistency fails
Data eligibility	Forward-fill a missing bar into a tradable price	Availability flag fails
Fill feasibility	Treat a market order as guaranteed last price	Execution assumption explicitly unsupported

A next-bar execution convention separates a close-only signal from its earliest modeled action. The implementation must still compare availability timestamps with the decision and submission times; a bar label alone does not establish what was known. Missing observations remain ineligible unless a separately justified rule permits their use. Forward-filling a value for analysis must not silently manufacture an executable quote.

Corporate actions require one consistent basis for prices, quantities and cash adjustments. Mixing an adjusted historical price with unadjusted share accounting can create artificial profit or loss. Fees should be attributed at the transaction level so a mismatch can be traced to a duplicate charge, a missing fill or an incorrect sign rather than hidden in a total return.

Costs, sensitivity and limits

Sensitivity analysis varies fee and execution assumptions separately from the signal rule. If a small change in costs reverses the economic interpretation, that dependence is a result of the assumptions, not proof of a robust strategy. This contract does not model liquidity, slippage, borrow availability, financing or market impact. Those omissions limit what can be inferred about capacity, short positions and achievable fills.

Conclusion

A backtest is a chain of auditable decisions. Eligibility checks, explicit timing, feasible execution assumptions and reconciled accounting support traceability across that chain. Passing arithmetic alone does not validate strategy economics; it establishes only that the stated transaction assumptions have been accounted for consistently.