



COMPUTATIONAL FINANCE & RISK MANAGEMENT

UNIVERSITY *of* WASHINGTON

Department of Applied Mathematics

Walk Forward Analysis An Introduction

CFRM 522

Introduction to Trading Systems (020)

- Pardo: Ch 11
- Jaekle & Tomasini: Ch 6
- Aronson: Ch 6

Walk Forward Analysis (Pardo)

Walk-Forward Analysis provides answers to four essential questions that are necessary to begin and to continue trading a strategy with intelligence and confidence:

1. Is the trading strategy robust? Will it make money in real-time trading?
2. What rate of return will the trading strategy produce in real-time trading?
3. How will the typical changes in market behavior such as trend, volatility and liquidity affect trading performance?
4. What are the best parameters to use in real-time trading to produce maximum profit with minimum risk?

Is the Trading Strategy Robust? (Pardo)

- A robust trading strategy is, by definition, one that is durable under changing market conditions, and will produce trading profits in real time consistent with those produced in its historical simulation.
- The primary goal of the walk-forward test is to determine whether the performance of a trading model under optimization is the result of a robust and repeatable process or the illusory result of overfitting.
- The primary purpose of a Walk-Forward Analysis is to determine whether a trading strategy's performance is real or not (skill vs luck).
- A trading strategy that does not stand up to Walk-Forward Analysis is most likely a trading strategy that is not going to produce profits in real-time trading either.
- A Walk-Forward Analysis that includes a large number of walk-forwards will greatly diminish the likelihood that the resulting performance is a product of random chance.

- Robusto:



- Walk Forward Efficiency (WFE): the ratio of an annualized walk-forward profit and loss to an annualized optimization profit and loss.
- The Walk-Forward Analysis also provides a more accurate insight into maximum drawdown, providing a more in-depth view of drawdown than that provided by one in-sample optimization.
- The Walk-Forward Analysis test also verifies the validity of the optimization process itself. An unsound strategy will simply fail a Walk-Forward Analysis.

More about WFA (Pardo, Jaekle & Tomasini, Aronson)

- Jaekle & Tomasini: Explanations for out-of-sample deterioration (also in previous train/test discussions on the Luxor strategy):
 1. The trading logic generates less than 100 trades, therefore the results are not statistically significant
 2. The system was published at the beginning of 2006 and the trading logic has been adopted by too many traders which destroyed its performance
 3. The market dynamics within the training data range are different from the one in the test data range
 4. The system has been adapted too much to market noise within the training period, i.e. curve over-fitting
- 1 & 2, however, are refuted by the authors (and Aronson)
- Re 2: Bollinger bands, for example are still widely used
- 3 is a possible explanation
- 4 is most likely, and the motivation for out-of-sample testing, including Walk Forward Analysis

- Aronson discounts points 3 & 2:
 - It is simply implausible to suggest that a market's dynamics change as frequently as rules fail out of sample.
 - Given that there are an almost infinite number of rules that could be formulated, it would be unlikely that a large enough number of participants would adopt any particular rule. Even in the case with futures trading funds that employ objective trend-following methods, reduced rule performance seems to be due to changes in market volatility that are not related to the usage of technical systems.

- Aronson instead posits:
 - A more complete account of out-of-sample performance deterioration is based on the data-mining bias.
 - It names two villains:
 - (1) randomness, which is a relatively large component of observed performance (related to J & T's point 4, and
 - (2) the logic of data mining, in which a best-performing rule is selected after the back-tested performances of all tested rules are available for the data miner's examination (loosely related to J & T's point 4).

- J & T, on the topic of sufficient data:
 - 20 years of daily data consists of about 5000 bars.
 - A year of 15 minute cash session bars also consists of about 5000 bars.
 - Does this mean that a year's worth of 15 minute data is equivalent to 20 years of daily data?
 - No, it doesn't. The probability of recording a crash in six months' of data would not be any better if its sampling interval were three minutes instead of 60 minutes or daily bars.
 - Be careful with short data samples regardless of timeframe or trading frequency. If your optimisation window is short, it is more likely you will miss important data outside of the window.
 - You will be generalising on limited historical data, and this will lead to a weaker potential for predicting the future.

More about WFA (Pardo, Jaekle & Tomasini, Aronson)

- Pardo also weighs in:
- Trading strategies typically have their worst drawdowns when trends or market conditions change.
- The rolling windows of the Walk-Forward Analysis provide a unique perspective showing exactly what has happened historically when the strategy encounters changing market conditions.
- The robust trading strategy is more capable of weathering changing conditions without catastrophic losses.
- The impact of such changes, however, can be masked by standard optimization. They are both revealed by the Walk-Forward Analysis and accommodated by its built-in adaptation to changing conditions through periodic reoptimization.

Parameter Set Optimization (Pardo)

- Two major points:
 - The parameter set provided by the WFA for real-time trading is more likely to be better suited to the current market environment.
 - The parameter set thereby provided comes with an expiration date.
- The length of the optimization window and of the walk-forward window *both become parameters of or part of the very structure of the trading strategy*.
- Our trading strategy will perform best if it can be tailored to current market conditions.
- *However*, the other axiom of trading strategy development demands that we use sufficient price data to produce robust results.
- So, how do we know we have sufficient data?

Choosing Sufficient Amount of Data (Pardo)

- Relevant data should be as large as possible as dictated by the restrictions imposed by:
 1. Statistical requirements
 2. Current market conditions
 3. The nature of the trading strategy
- A WFA should be as long as possible, typically at least 10 to 20 years, whenever possible. A WFA of this length can produce 10, 20, or more *walkforwards*. The combined performance of these multiple walkforwards will often be sufficient to yield statistical rigor.
- Consequently, a long Walk-Forward Analysis is able to use shorter length in-sample windows to achieve both optimal trading performance from its use of the most relevant data and statistical rigor at the same time (thus also satisfying point 2, and to some extent point 3).

Balancing Act (Pardo)

- Between relevant data and statistical rigor
- Markets change their behaviors from time to time
- Market conditions do shift and change and they shift and change enough to make it advantageous for some trading strategies to periodically adjust to these shifting conditions
- Examples:
 - Weather conditions, particularly relevant for agricultural commodities
 - Business cycle
 - Supply and demand (temporary shortages, etc)
 - Strong tendency to trend (eg currency markets); this effect may dominate tradable price swings
 - Mean reversion (S&P and Nasdaq futures)
 - Volatility effects (may be related to business cycle and bull/bear market)
 - Liquidity fluctuations
- Strong points of WFA:
 - Able to identify the best optimization and trading windows for a trading strategy
 - Achieves maximum trading performance by adapting to current conditions while maintaining statistical robustness

The Walk Forward

- A two-step process:
 1. Traditional optimization on a historical sample of market data
 2. Performance of the optimized parameter set is evaluated on an additional, adjacent sample of price history that was not included in the optimization price sample
- Point 2 essentially means simulation of real-time trading
- WFA is one of the very best methods available to evaluate the robustness of a trading strategy



[Keep On Truckin' \(Eddie Kendricks\)](#)

Conclusions after Running WFA

- Pardo mentions there are three possible conclusions following the running of a WFA test:
 1. If the strategy loses money in the majority of the out-of-sample tests, then one does not have a tradable strategy
 2. If the strategy performs with some degree of moderate or limited profit in a majority of these tests, this may indicate a poor-to-overfit trading strategy
 3. Performance of out-of-sample tests in proportion to its in-sample profit in the majority of tests indicates a likely robust trading strategy

- A walk forward requires the following components:
 1. Scan ranges for the variables to be optimized (paramset)
 2. An objective or search function (default is `max Net.Trading.PL`)
 3. Size of the optimization window
 4. Size of the walk-forward, trading or out-of-sample window
- Length of the optimization window:
 - In general, the faster-paced a trading strategy, the more likely it is to benefit from shorter optimization (and walk-forward) windows.
 - one to two years in length
 - Conversely, the slower that strategy, the longer the window needed. three to six years in length
 - These values are not strict, however and may be highly variable because of a number of (other) factors

- Length of the walk-forward (test) window:
 - Typically a function of the size of optimization window.
 - A walk-forward window should be somewhere in the area of 25 to 35 percent the size of the optimization window
- The size of these windows is best determined empirically
 - Author describes in further detail later in the chapter; however,
 - Author also says this is a topic far too complex and in-depth for the book.
- In practice:
 - a strategy optimized properly on two years of price history will likely remain usable for between three and six months of real-time trading
 - a model built on five years of price history will likely remain usable for between one and two years of real time trading

Example in Pardo

- S&P 500 futures strategy
- 36-month optimization window
- 6-month test window
- Roll 36-month optimization window forward six months at a time
- Historical data: 1990.01.01 – 2007.12.31 (See Table 11.2)
- Average annualized profit for 90 years of training: $\$745,000/90 = \8278 (very simple calculation)
- Average annualized profit for 15 total years of test (walk forward) periods: $\$168,113/15 = \$11,208$
- Author set a 75% target of test/train profit
- End up at 135%
- Conclusion: the strategy is robust

TABLE 11.2 Walk-Forward Analysis is Performance Summary

Summary Report for Session

RSI_CT_WF_sp SP-9967.TXT 1/1/1990 to 12/31/2007. System is RSI_CT()

Performance Summary: All Trades

Total net profit	\$154,350.00	Open position P/L	(\$19,625.00)
Gross profit	\$756,050.00	Gross loss	(\$601,700.00)
Total number of trades	218	Percent profitable	51.38%
Number winning trades	112	Number losing trades	106
Largest winning trade	\$90,150.00	Largest losing trade	(\$57,875.00)
Average winning trade	\$6,750.45	Average losing trade	(\$5,676.42)
Ratio average win/average loss	1.19	Average trade (win & loss)	\$708.03
Maximum consecutive winners	8	Maximum consecutive losers	8
Average number bars in winners	15	Average number bars in losers	18
Maximum intraday drawdown	(\$112,875.00)	Maximum number contracts held	1
Profit factor	1.26	Yearly return on account	8.05%
Account size required	\$112,875.00		

- Risk – Drawdown (**Important**)
 - a walk-forward maximum drawdown is an important measure of risk and must be closely watched
 - a trading strategy that produces a real-time drawdown that exceeds optimization (train) drawdown by a wide margin—all things being equal—may be showing signs of trouble

[END]