

Backtest Overfitting

Professor : Daniel Totouom

Presenter:

XiaoYi(Roly).Yang (N12903840)

Yixing(Florence).Chen (N19064360)

Zhongyuan(Lucas).Bi (N12939762)

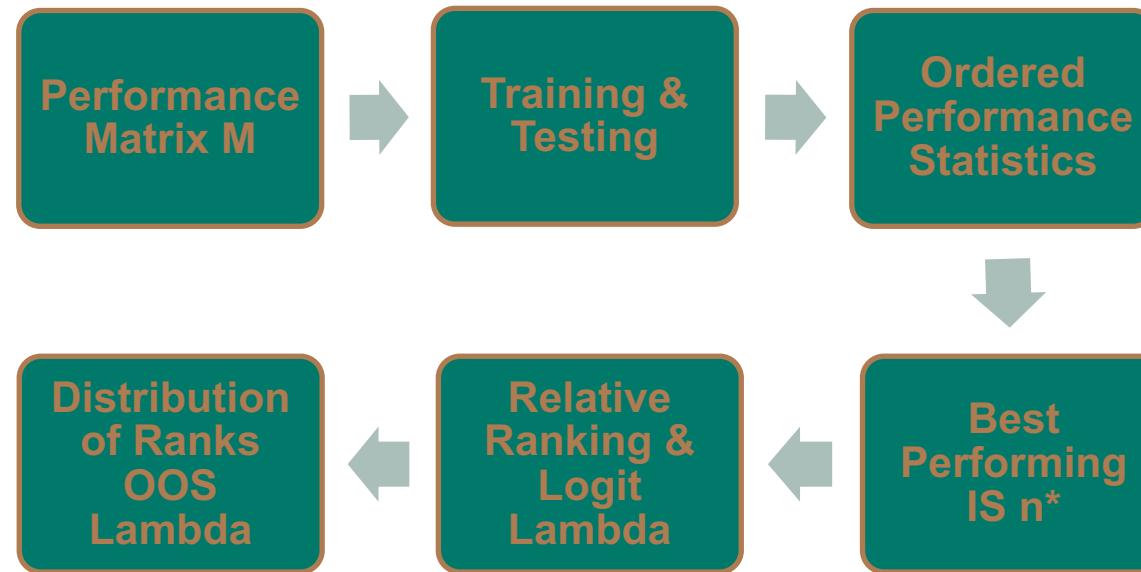
Overview/Objective

- **Cross Validation -- Hold Out**
 - May use “hold out” as part of the IS
 - Information (know how financial variables performed over OOS interval) will be used in strategy design
 - Clearly inadequate for small samples
 - IS too short to fit
 - Can’t conclude OOS sufficient confidence
 - Always present overfitting - >1 strategy

“Adapts recent advances in experimental mathematics, machine learning and decision theory to the very particular problem of assessing the representativeness of a backtest”

Model-Free / Non-parametric

CSCV PROCEDURE



Strategies

- **CSCV**
- **Monte Carlo**
- **Extreme Value Theory**
 - **From the actual distribution of the maximum SR**
- Median Overfitting
 - Overestimate --Conservative
- Low Overfitting
 - Probability is lower than EVT --accurate

Out [15] :		SRCASE	T	N	CSCV	MC	EVT	CSCV-EVT
0	0	500	500	1.000	1.000	1.000	-0.000	
1	0	500	100	1.000	1.000	1.000	-0.000	
16	1	1000	500	0.909	0.885	0.869	0.039	
17	1	1000	100	0.687	0.721	0.715	-0.029	
18	1	1000	50	0.647	0.617	0.628	0.019	
19	1	1000	10	0.214	0.389	0.375	-0.161	
20	1	2500	500	0.500	0.463	0.477	0.023	
21	1	2500	100	0.353	0.271	0.295	0.058	
22	1	2500	50	0.254	0.228	0.220	0.034	
23	1	2500	10	0.119	0.092	0.091	0.028	
24	2	500	500	0.635	0.614	0.599	0.036	
25	2	500	100	0.417	0.442	0.402	0.014	
41	3	1000	100	0.020	0.002	0.002	0.018	
42	3	1000	50	0.000	0.000	0.001	-0.001	
43	3	1000	10	0.000	0.000	0.000	-0.000	
44	3	2500	500	0.000	0.000	0.000	-0.000	
45	3	2500	100	0.000	0.000	0.000	-0.000	
46	3	2500	50	0.000	0.000	0.000	-0.000	
47	3	2500	10	0.000	0.000	0.000	-0.000	

Combinatorially Symmetric Cross Validation

CSCV FEATURES

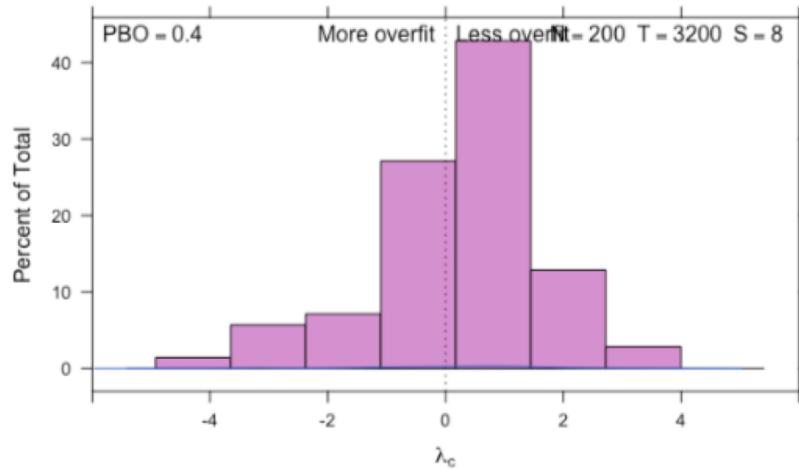
- Training and Testing sets are of equal size--provide comparable accuracy (IS & OOS)
- Symmetric -training sets are reused as testing sets
 - Decline in performance can only result from overfitting
- Respect the time-dependence and other seasonalities present in the delta
- Derives a non random distribution of logits
- Dispersion conveys information -- robustness of the strategy selection procedure
- Model-Free
- Choose double (observations)

Statistics of Overfitting -

EVALUATION STATISTICS

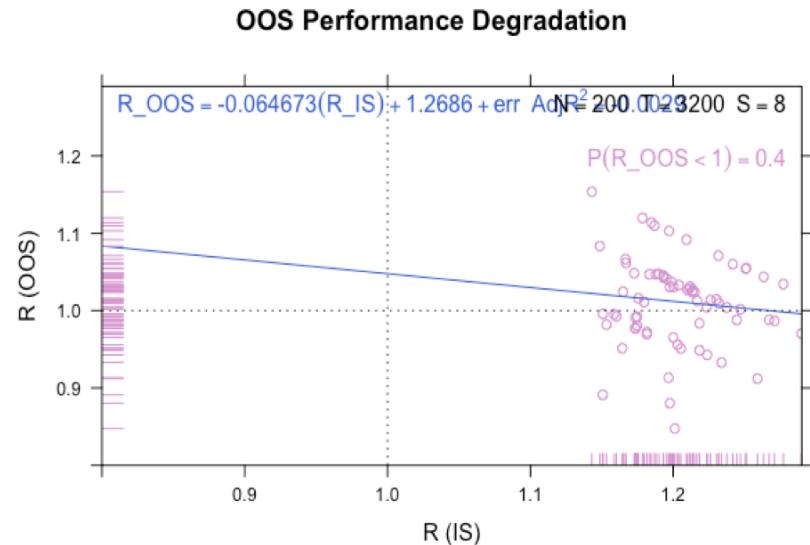
- **PBO**
 - The probability that the model configuration selected as optimal IS will **underperform** the median of the N model configurations OOS.
- **Stochastic Dominance**
 - Whether the procedure used to select a strategy IS is preferable to randomly choosing one model configuration among the N alternatives.
- **Performance Degradation**
 - determines to what extent greater performance IS leads to lower performance OOS

CSCV -- Generate random matrix of return

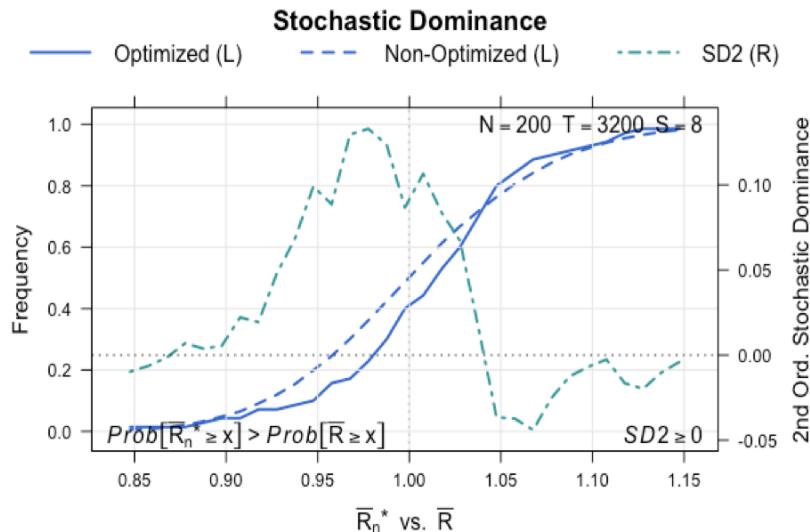


- **Performance Degradation:**
 - **Linear relationship** between the ranking of IS strategies and the ranking of OOS strategies.
 - **Positive** : strategy ranking suits for both IS and OOS pool.

- **Rank Logits**
 - The **higher** the expectations of Logits, the **lower** PBO.



CSCV -- Generate random matrix of return



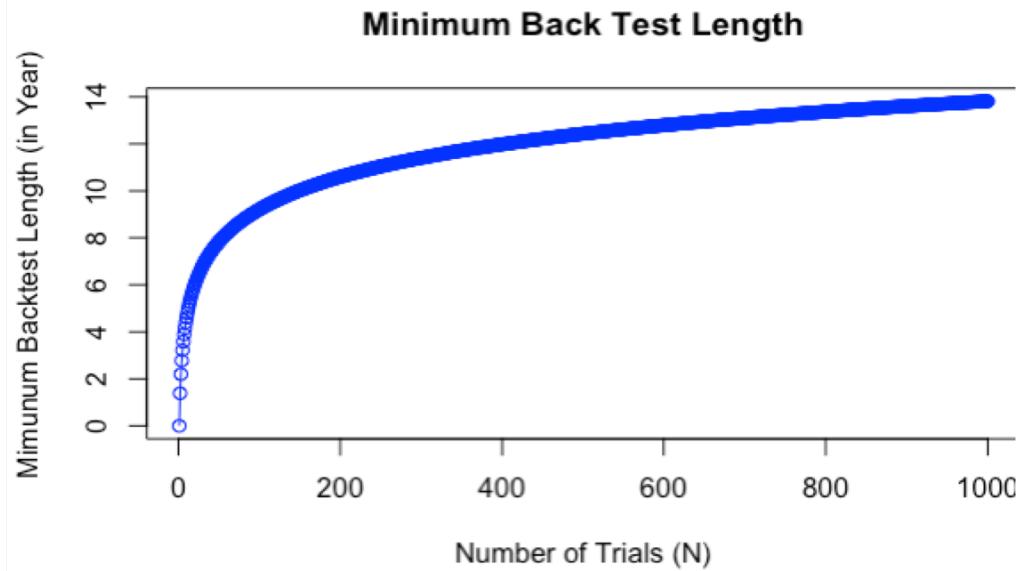
- **Stochastic Dominance**
 - The strategy is not persistant

Minimum Back Test Length

$$\text{MinBTL} = \left(\frac{(1 - \gamma)N^{-1}(1 - 1/N) + \gamma N^{-1}(1 - \frac{1}{N}e^{-1})}{E(\max)} \right)^2 < \frac{2 \ln(N)}{E(\max)^2}$$

$\gamma \approx 0.5772$
Euler-Mascheroni constant

$E(\max)$
Expected Max IS SR



Test Strategies – Industry Volatility Arbitrage Strategy

To construct an arbitrage portfolio:

- Long a strip of option and short a dynamic position of $1/S$ unit of the underlying instrument adjusted daily (short term).
- Short a strip of option and hold a dynamic position of $1/S$ adjusted monthly (long term).

Minimum Backtest Length needed:

4.39 years

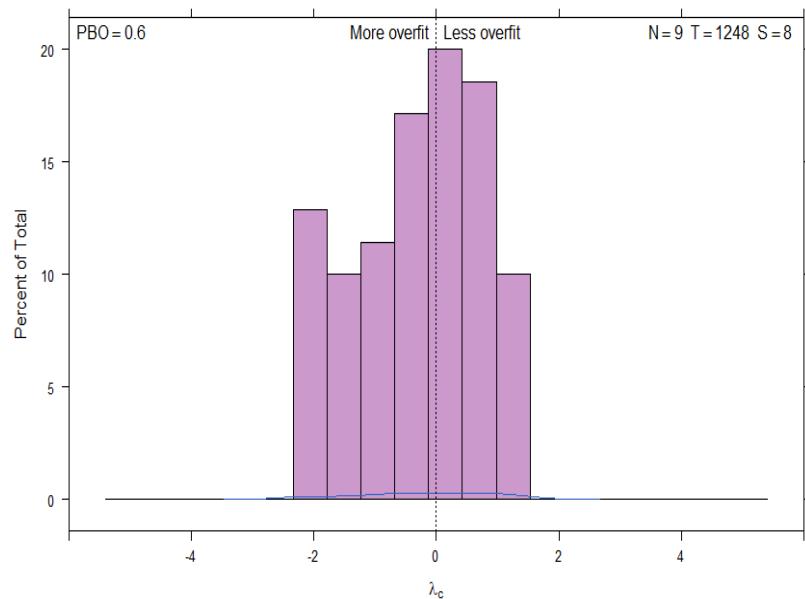
Stocking Selecting (By Industry)

5 Year Stock Price:

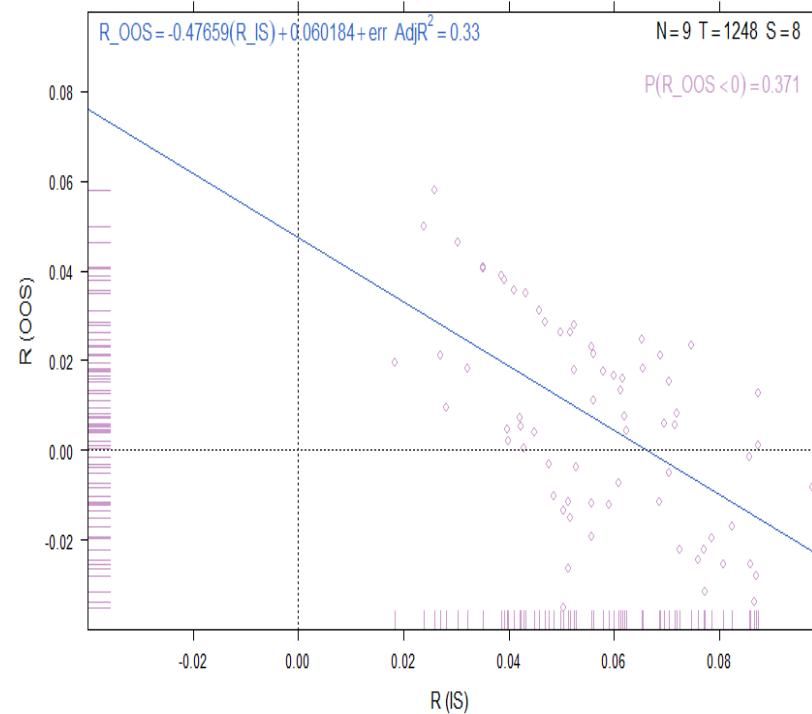
- Financial: MS
- Healthcare: UNH
- Service: AMZN
- Utilities: AEP
- Industrial Goods: EMR
- Basic Materials: XOM
- Conglomerates: HRG
- Consumer Goods: AAPL
- Technology: GOOG

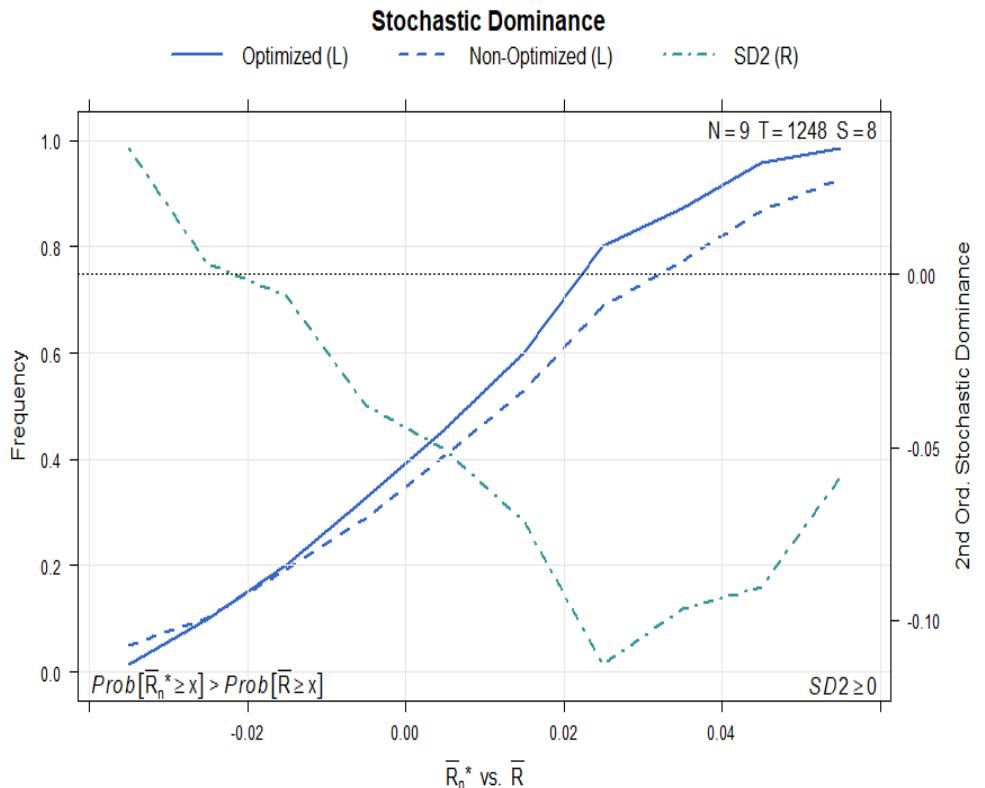
Decreasing Volatility w.r.t Frequency

PBO=0.086
Positive and High Expectation of Logits



OOS Performance Degradation

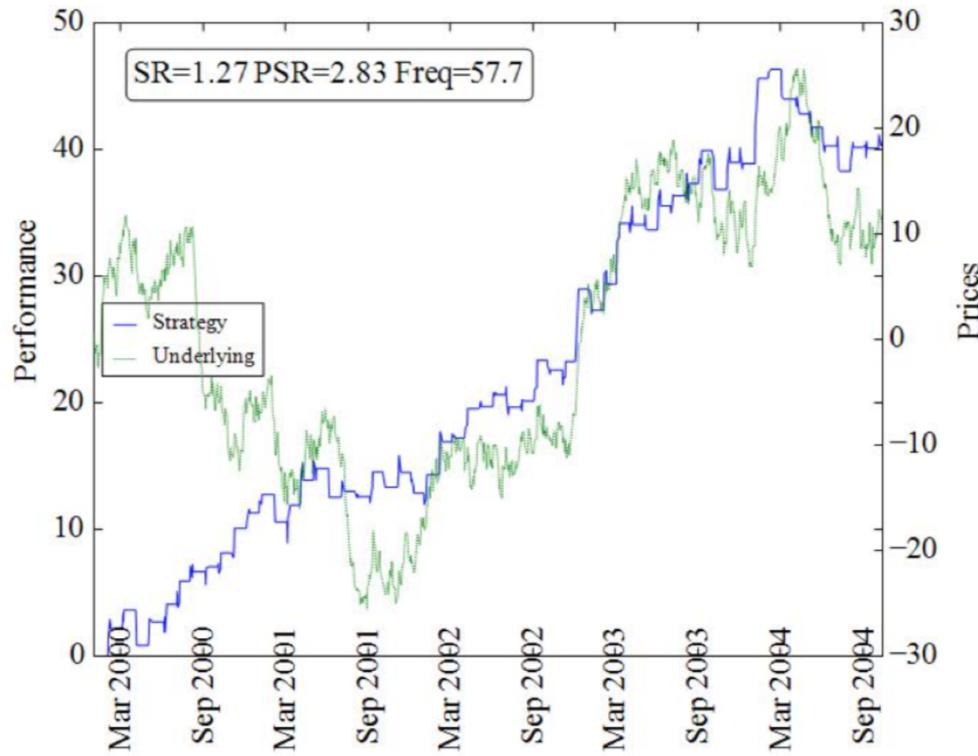




OOS pool does not outperform optimized IS strategy.

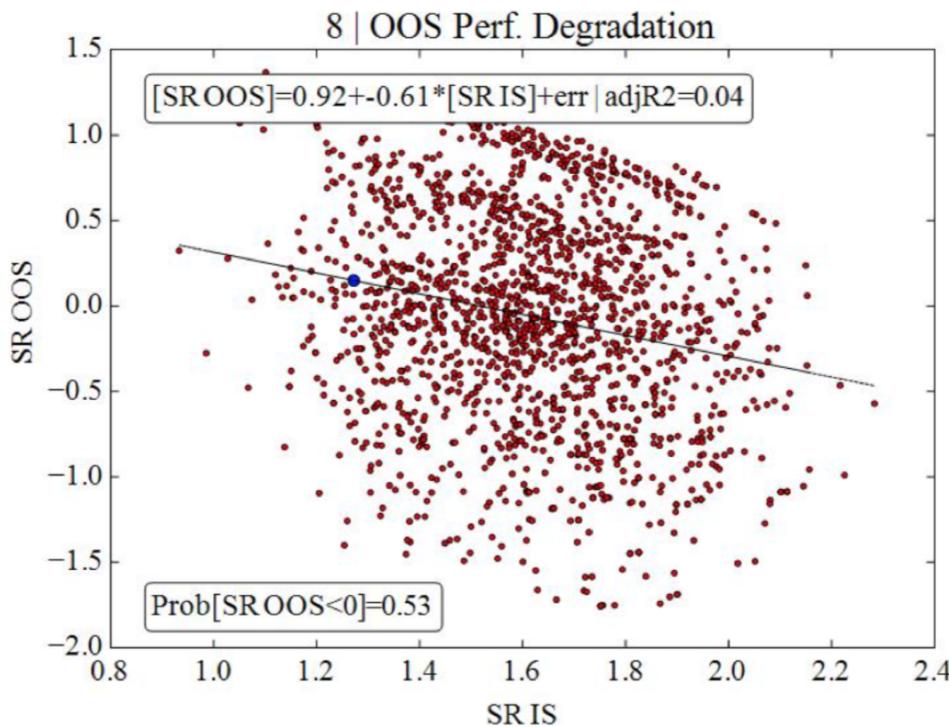
Volatility Arbitrage Strategy can be easily overfitted.

Test Strategies - Seasonal Strategy



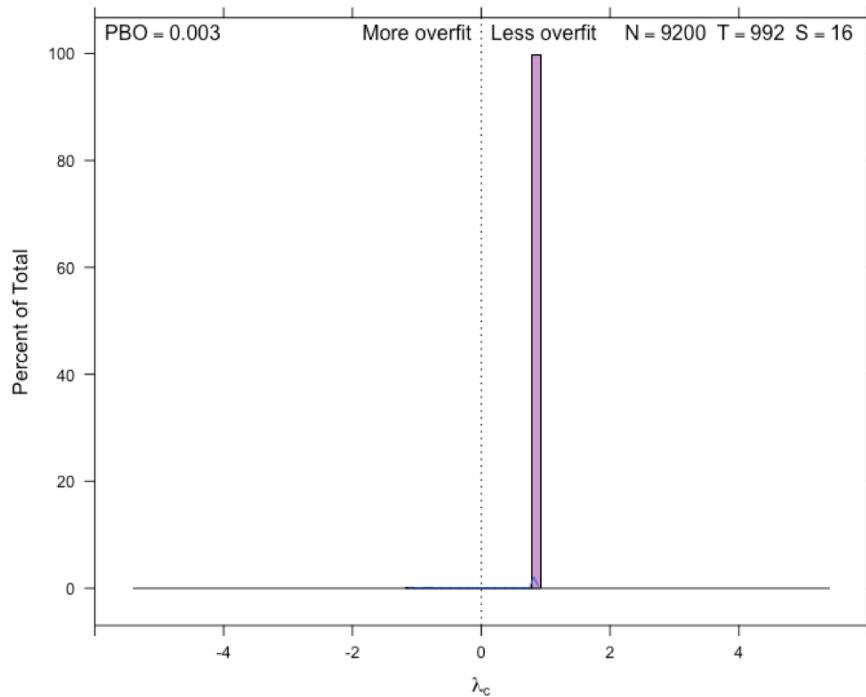
- **PSR(Probability that a Sharpe Ratio is inflated) = 2.83**
 - Max PSR is very close to Normal
 - less than 1% probability that the true Sharpe ratio is below 0

Test Strategies - Seasonal Strategy



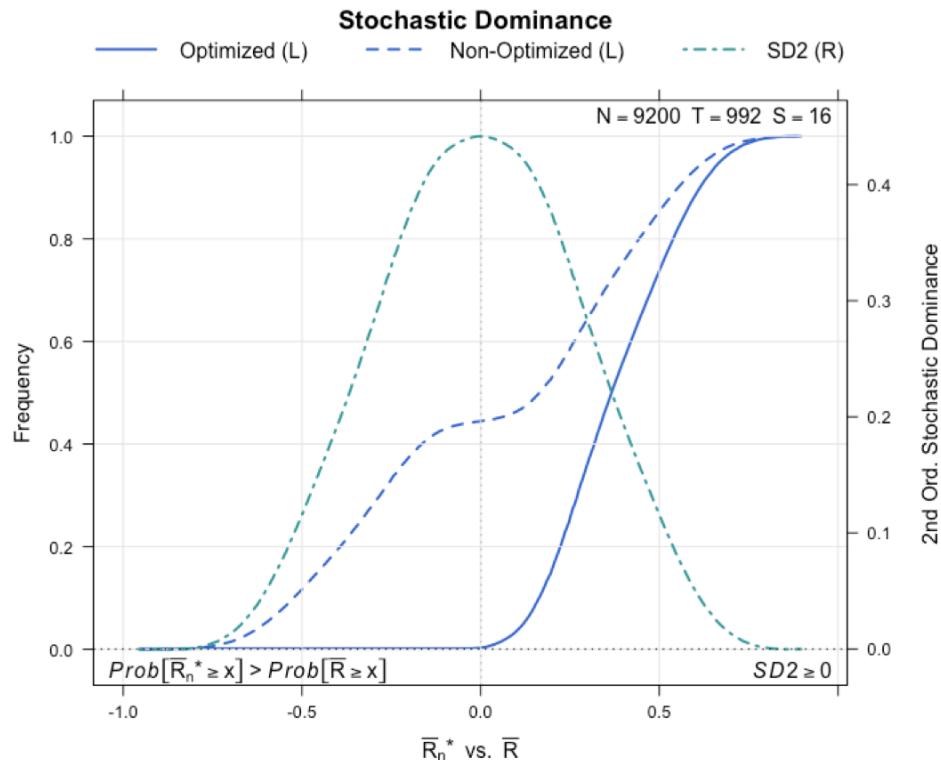
- Use CSCV procedure to estimated PBO
 - **53% of the SR OOS are negative, despite all SR IS being positive and ranging between 1 and 2.2.**

Test Strategies - Seasonal Strategy



- Despite the elevated SR IS, the PBO is **less than 3%**.

Test Strategies - Seasonal Strategy



Distribution of OOS SR **dominate** the overall distribution of optimized IS SR

Combinatorially Symmetric Cross Validation

CSCV Limitations/ Challenge

- TEST as many as profits and losses series (N)
- TEST as many as reasonable and feasible configurations
- Does nothing to evaluate the correctness of a back test
- Skillful strategies may still exist
 - Have high but similar SR
- Cannot apply to guide the search for an optimal strategy
- Difficult to fit in models that use continuous variables

Conclusion