ADVANCES IN QUANTITATIVE META-STRATEGIES

Marcos López de Prado

Lawrence Berkeley National Laboratory
Computational Research Division

Key Points

- Quantitative Meta-Strategies (QMS) are quantitative strategies designed to manage investment strategies.
- As a field, QMS is the mathematical study of the decisions made by the supervisor of a team of investment managers, regardless of whether their investment style is systematic or discretionary.
- Some advantages that QMS offer are:
  – Algorithmized investment processes can be tested and improved before being applied to a business.
  – They provide objective and consistent oversight, and help prevent repeated mistakes.
  – They are scalable and speed up quality improvement by limiting managerial frictions and biases.

SECTION I
The Art & Science of Investing

Investing is Not an Art

- <u>Popular view</u>: Investing is an Art because…
  - It does not follow fixed rules (…unlike Art!)
  - Like a sport or game, excellence cannot be taught, but it can <u>only</u> be learned through practice.
- <u>Reality</u>: Sports, Games, Arts increasingly rely on Math.

Any game can be studied mathematically. For decades it was thought that the sheer number of combinations involved in Chess would mean that computers would never beat top players. In 1996, IBM's Deep Blue settled that question.

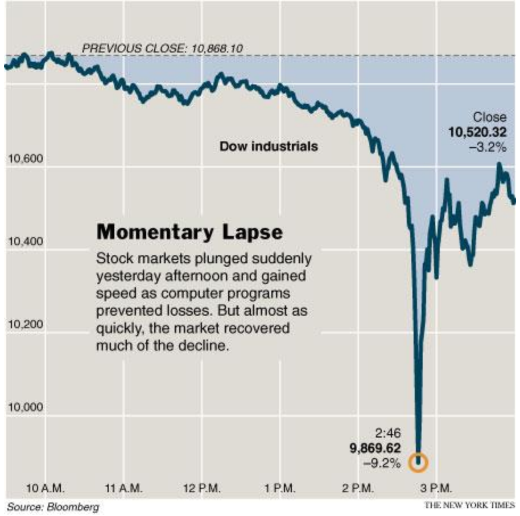
Today, [some of the most successful hedge funds are math-oriented].

Investing is Not an Academic Endeavor

- Academic view: Like physical objects, Markets follow fundamental principles that can be studied.
- Reality: Investing differs from Physics in several aspects:
  - There are no "laboratories": We cannot reproduce experiments under controlled conditions.
  - Low signal-to-noise leads to the proliferation of false positives
  - Effects are not immutable: Competition arbitrages them away.

We will never know if Mr. Sarao caused the Flash Crash. Unlike in physics, we cannot repeat the events of that day in absence of Mr. Sarao's spoofing.

In the words of Prof. Campbell Harvey (President-Elect of the American Finance Association) *"most claimed research findings in financial economics are likely false"*.



Investing is an Industrial Science

- In contrast with Academic Finance:
  - Financial firms can conduct research in terms analogous to Scientific laboratories. E.g., deploy an execution algorithm and experiment with alternative configurations (market interaction).
  - Financial firms can control for the increased probability of false positives that results from multiple testing, because they can take into account the results from all trials.
  - Financial firms do not necessarily report their discoveries, thus discovered effects are more likely to persist.
- <u>Conclusion #1</u>: *Empirical* Finance discoveries are more likely to occur in the Industry than in Academia.
- QMS are investment processes geared towards taking advantage of those industrial discoveries.

SECTION II
Strategy Selection / PM Hiring

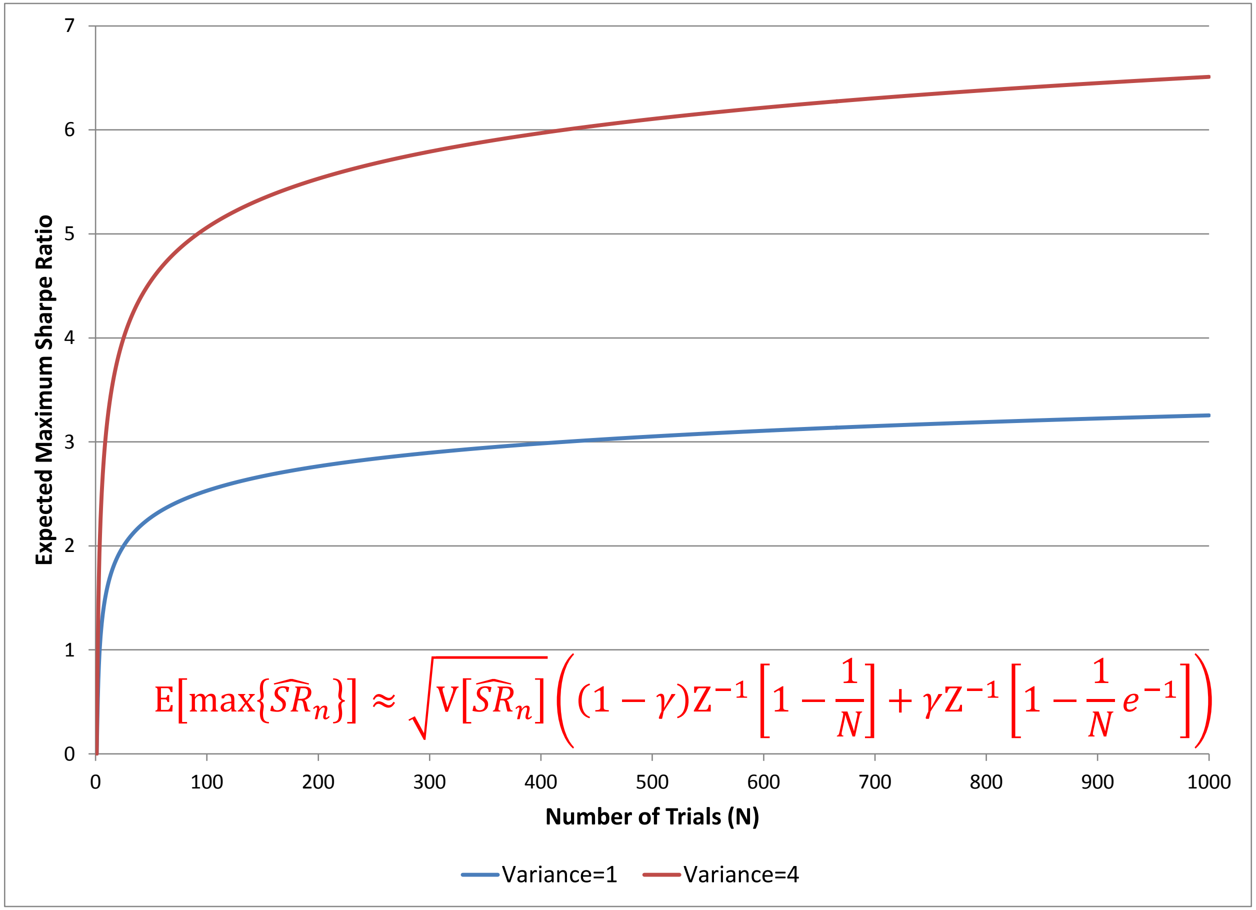
Strategy Selection Under Multiple Testing

- <u>PROPOSITION #1</u>: *Suppose N independent trials following a Normal distribution, with mean $E\left[\widehat{SR}_n\right] = 0$ and variance $V\left[\widehat{SR}_n\right]$. Then, the expected maximum Sharpe Ratio is*

$$E\left[max\{\widehat{SR}_n\}\right] \approx \sqrt{V\left[\widehat{SR}_n\right]}\left((1-\gamma)Z^{-1}\left[1-\frac{1}{N}\right]+\gamma Z^{-1}\left[1-\frac{1}{N}e^{-1}\right]\right)$$

where $\boldsymbol{\gamma}$ is the Euler-Mascheroni constant (approx. 0.5772), $\boldsymbol{Z}$ is the CDF of the Standard Normal and $\boldsymbol{e}$ is Euler's number.

Backtest Overfitting

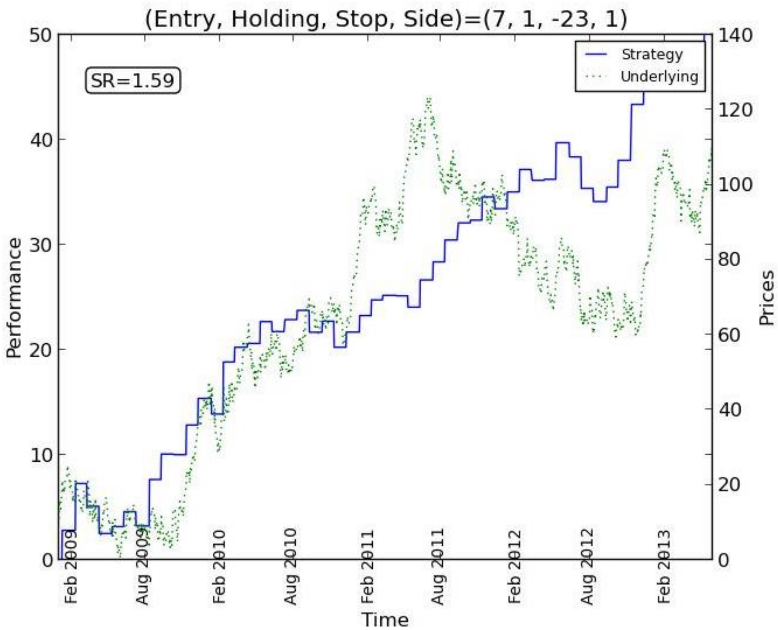


[Expected Maximum Sharpe Ratio](#) as the number of independent trials N grows, for $E\left[\widehat{SR}_n\right] = 0$ and $V\left[\widehat{SR}_n\right] \in \{1,4\}$.

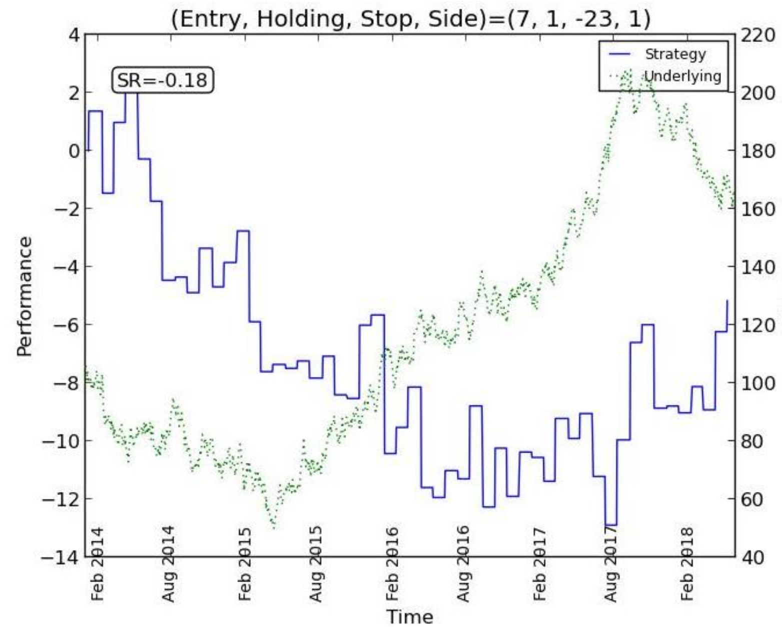
Searching for empirical findings regardless of their theoretical basis is likely to magnify the problem, as $\mathrm{V}\left[\widehat{SR}_n\right]$ will increase when unrestrained by theory.

This is a consequence of pure random behavior. We will observe better candidates *even if there is no investment skill associated with this strategy class* ($E\left[\widehat{SR}_n\right] = 0$).

The Backtest Overfitting Simulation Tool



An “optimized” investment strategy (in blue) making steady profit while the underlying trading instrument (in green) gyrates in price. It is trivial to make Financial “discoveries” if enough variations are tried.



The same investment strategy performs poorly on a different sample of the same trading instrument.

http://datagrid.lbl.gov/backtest/index.php

The Deflated Sharpe Ratio (1/2)

- The Deflated Sharpe Ratio (DSR) corrects the inflationary effect of multiple trials, non-normal returns and shorter sample lengths:

$$\widehat{DSR} \equiv \widehat{PSR}\left(\widehat{SR}_0\right) = Z\left[\frac{\left(\widehat{SR} - \widehat{SR}_0\right)\sqrt{T-1}}{\sqrt{1 - \hat{\gamma}_3\widehat{SR} + \frac{\hat{\gamma}_4 - 1}{4}\widehat{SR}^2}}\right]$$

where

$$\widehat{SR}_0 = \sqrt{V\left[\widehat{SR}_n\right]}\left((1-\gamma)Z^{-1}\left[1 - \frac{1}{N}\right] + \gamma Z^{-1}\left[1 - \frac{1}{N}e^{-1}\right]\right)$$

DSR is a Probabilistic Sharpe Ratio where the rejection threshold is adjusted to reflect the multiplicity of trials.

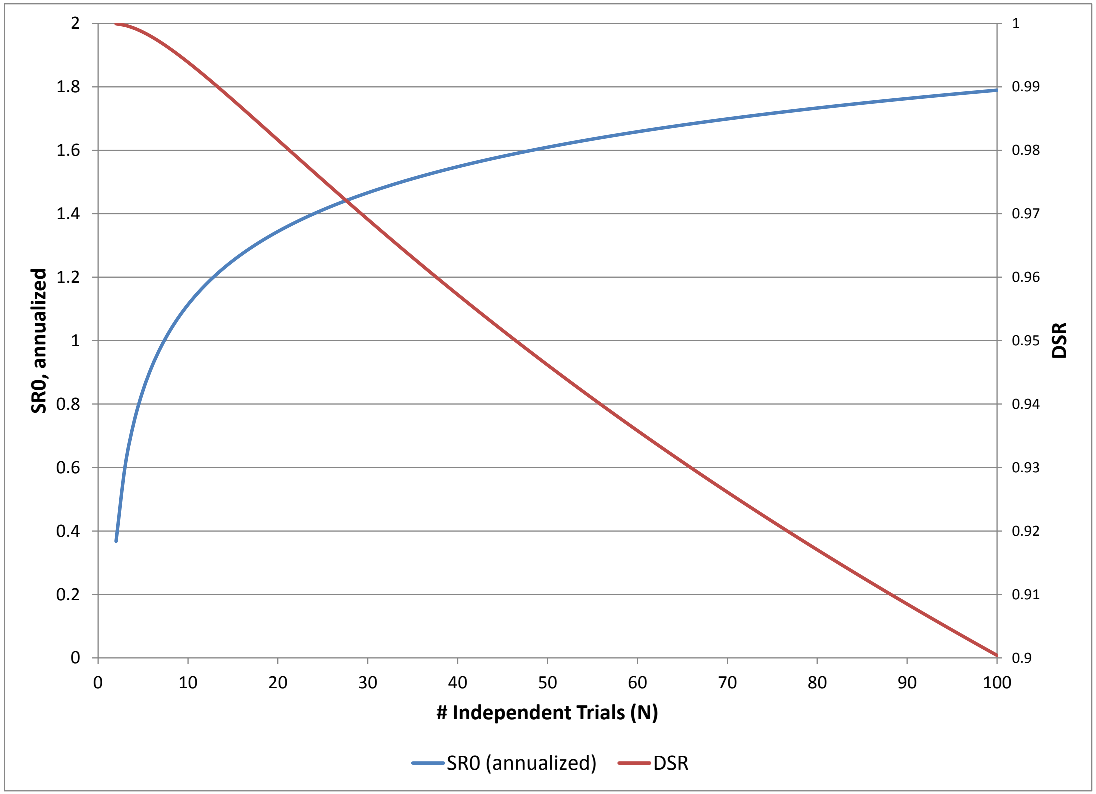
The Deflated Sharpe Ratio (2/2)

- The standard Sharpe Ratio (SR) is computed as a function of two estimates:
  - Mean of returns
  - Standard deviation of returns.
- DSR deflates SR by taking into consideration five additional variables (it packs more information):
  - The non-Normality of the returns $(\hat{\gamma}_3, \hat{\gamma}_4)$
  - The length of the returns series (T)
  - The variance of the SRs tested$\left(V\left[\widehat{SR}_n\right]\right)$
  - The number of independent trials involved in the selection of the investment strategy (N)

Numerical Example (1/2)

- An analyst uncovers a daily strategy with annualized *SR=2.5,* after running *N=100* independent trials, where $V\left[\widehat{SR}_n\right] = \frac{1}{2}$, *T=1250,* $\hat{\gamma}_3 = -3$ and $\hat{\gamma}_4 = 10$.
- QUESTION: Is this a legitimate discovery, at a 95% conf.?
- ANSWER: **No.** There is only a 90% probability that the true Sharpe ratio is above zero.
  - $\widehat{SR}_0 = \sqrt{\frac{1}{2 \cdot 250}}\left((1-\gamma)Z^{-1}\left[1-\frac{1}{100}\right] + \gamma Z^{-1}\left[1-\frac{1}{100}e^{-1}\right]\right) \approx 0.1132$
  - $\widehat{DSR} \approx Z\left[\frac{\left(\frac{2.5}{\sqrt{250}}-0.1132\right)\sqrt{1249}}{\sqrt{1-(-3)\frac{2.5}{\sqrt{250}}+\frac{10-1}{4}\left(\frac{2.5}{\sqrt{250}}\right)^2}}\right] = 0.9004.$

Numerical Example (2/2)



Should the strategist have made his discovery after running only *N=46,* then $\widehat{DSR} \approx 0.9505$.

Non-Normality also played a role in discarding this investment offer: For $\hat{\gamma}_3 = 0, \hat{\gamma}_4 = 3$, then $\widehat{DSR} = 0.9505$ after *N=88* independent trials.

It is critical for investors to account for both sources of performance inflation <u>jointly</u>, as DSR does.

SECTION III
Portfolio Oversight

Finance and the Theory of Evolution

- Standard structural break tests (see Maddala and Kim [1999]) attempt to identify a *“break” or permanent shift* from one regime to another within a time series.
- In contrast, the methodology we present here signals the emergence of a new regime as it happens, *while it co-exists* with the old regime (thus the mixture).
- This is a critical advantage, in terms of providing an *early signal* that a new investment style is emerging in a fund or portfolio.
- Conclusion #2: Evolutionary divergence attempts to signal the emergence of a new investment style before it is so prevalent that a “break” can be detected.

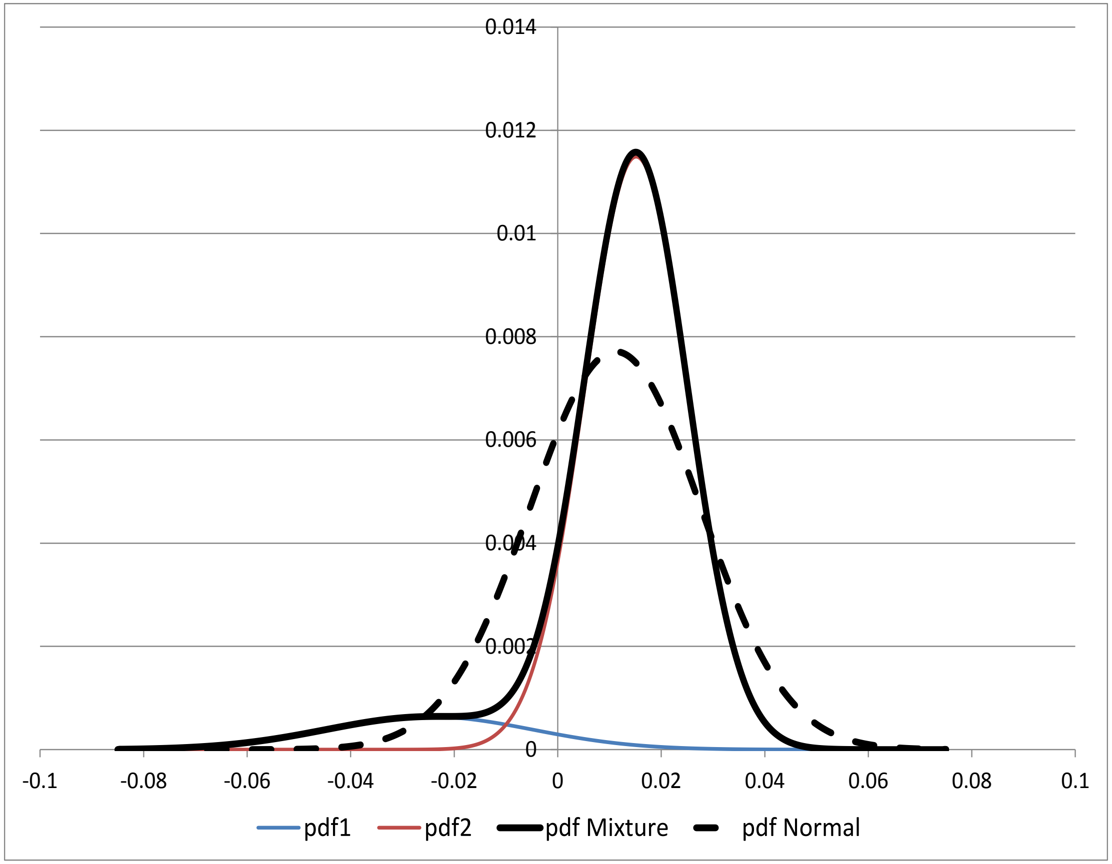
Action plan

1. Apply the EF3M algo for matching the track record's moments (we have already seen this step).
2. Simulate path scenarios consistent with the matched moments.
3. Derive a distribution of scenarios based on that match.
4. Probability of Divergence (PD): Evaluate what percentile of the distribution corresponds with the PM's recent performance.

<u>Conclusion #3</u>: PD assesses the evolutionary divergence by taking into account the entire *distribution* of the possible mixture parameters, based on the reliable moments.

An Example

- Suppose a mixture of 2 Gaussians with true parameters $(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{p}) = (-0.025, 0.015, 0.02, 0.01, 0.1)$.



True raw moments

Moments	Origin	Mean
1	1.10E-02	0
2	3.95E-04	2.74E-04
3	2.53E-06	-7.85E-06
4	4.31E-07	5.63E-07
5	-7.48E-09	-3.28E-08

Fitted Parameters

Parameter	Average	StDev
Mu1	-0.0245	0.0027
Mu2	0.0150	0.0001
Sigma1	0.0201	0.0009
Sigma2	0.0100	0.0002
Prob1	0.1026	0.0144

Estimation Errors

Errors	Average	StDev
1	0.00E+00	0.00E+00
2	0.00E+00	0.00E+00
3	0.00E+00	0.00E+00
4	-3.80E-11	2.54E-11
5	-5.93E-12	3.01E-11

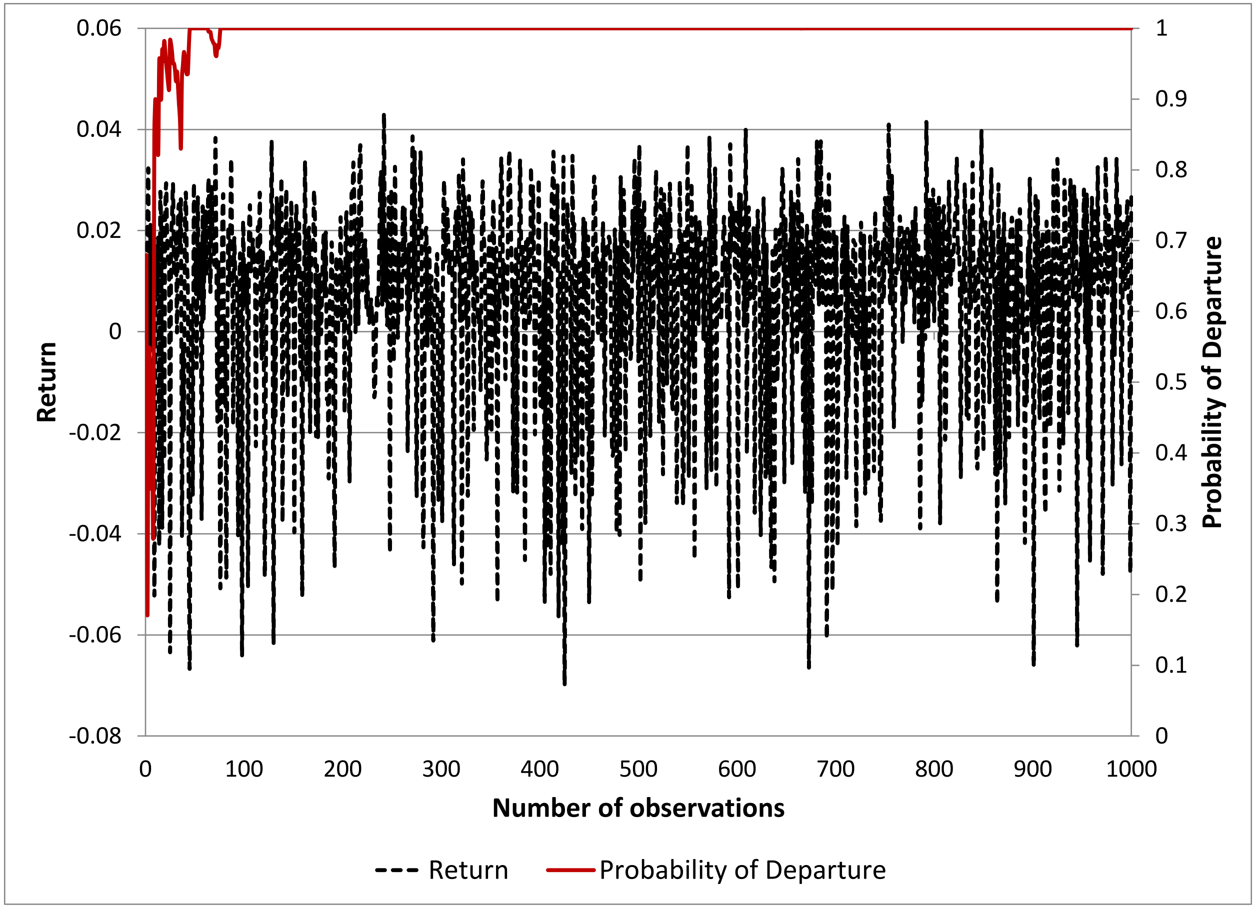
Example 1 (1/2)

- Reference: Suppose that a PM has a track record consistent with the following Mixture of 2-Gaussians

$$(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{p}) = (-0.025, 0.015, 0.02, 0.01, 0.1)$$

- Divergence: What would happen if draws from the first Gaussian become more likely? For example, if *p=0.2* instead of *p=0.1*, the mixture's distribution would become more negatively skewed and fat-tailed.

Example 1 (2/2)



That situation is distinct from the approved track-record, and PD slowly but surely converges to 1.

Example 2 (1/2)

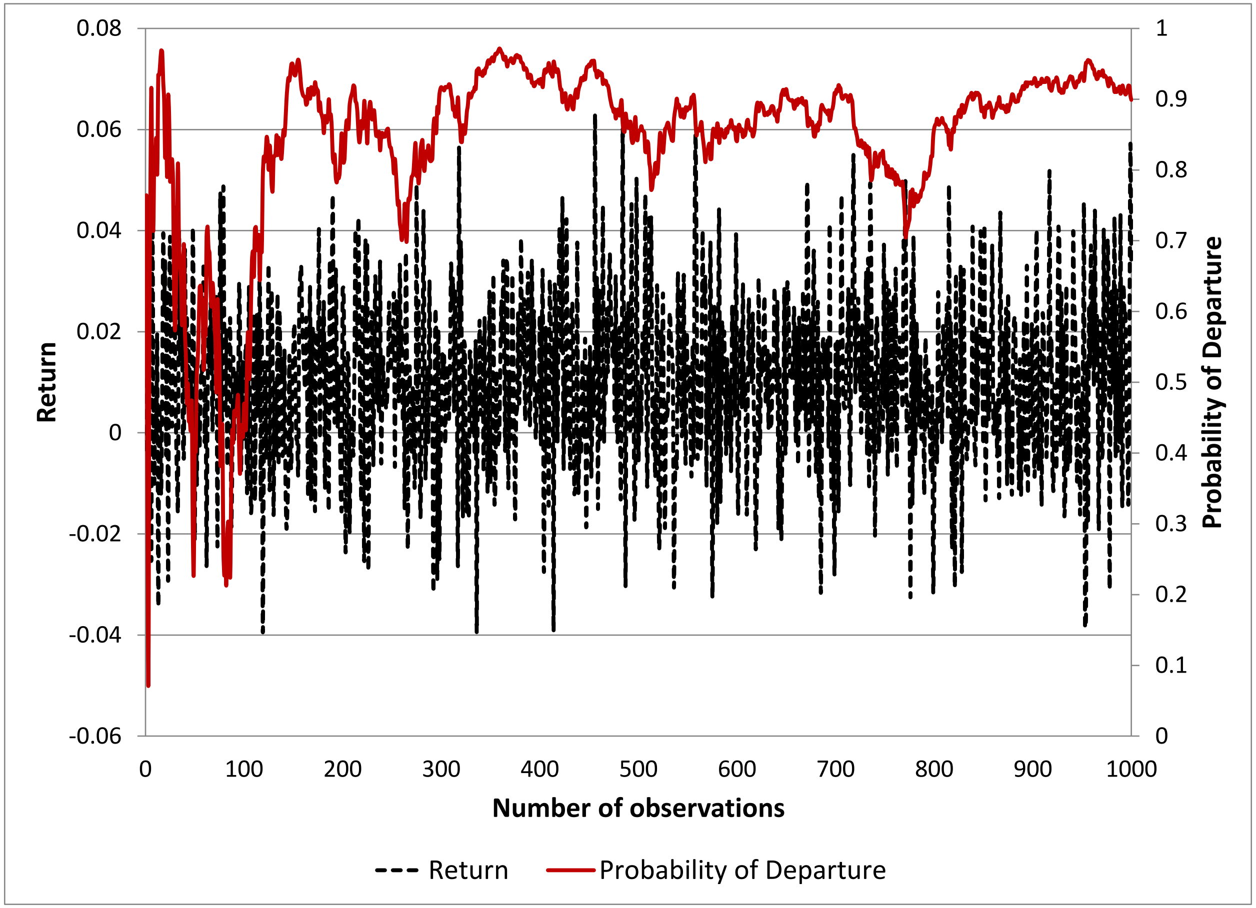
- Reference: Suppose that a PM has a track record consistent with the following Mixture of 2-Gaussians

$$(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{p}) = (-0.025, 0.015, 0.02, 0.01, 0.1)$$

- Divergence: What would happen if, after capital is allocated, returns are IID Normal, matching the mixture's mean and variance, i.e.

$$N(\tilde{\mu}, \tilde{\sigma}^2) = N(1.10E-02, 2.74E-04)$$

Example 2 (2/2)



PD approaches 1, although the model cannot completely discard the possibility that these returns in fact were drawn from the reference mixture.

Example 3 (1/2)

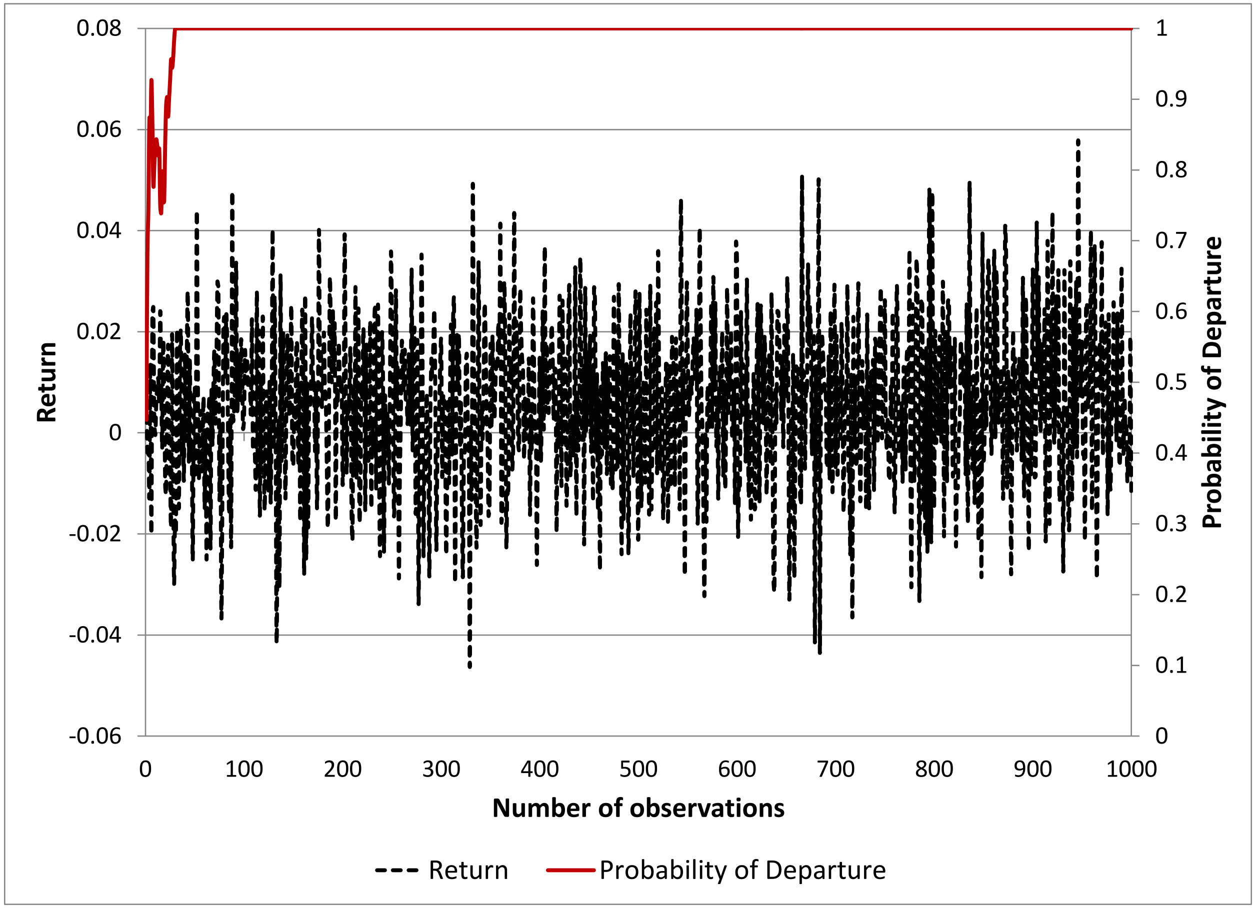
- Reference: Suppose that a PM has a track record consistent with the following Mixture of 2-Gaussians

$$(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{p}) = (-0.025, 0.015, 0.02, 0.01, 0.1)$$

- Divergence: What would happen if, after capital is allocated, returns are IID Normal with a mean half the mixture's and the same variance as the mixture…

$$N(\tilde{\mu}, \tilde{\sigma}^2) = N(5.5E-03, 2.74E-04)$$

Example 3 (2/2)



PD quickly converges to 1, as the model recognizes that those Normally distributed draws do not resemble the mixture's simulated paths.

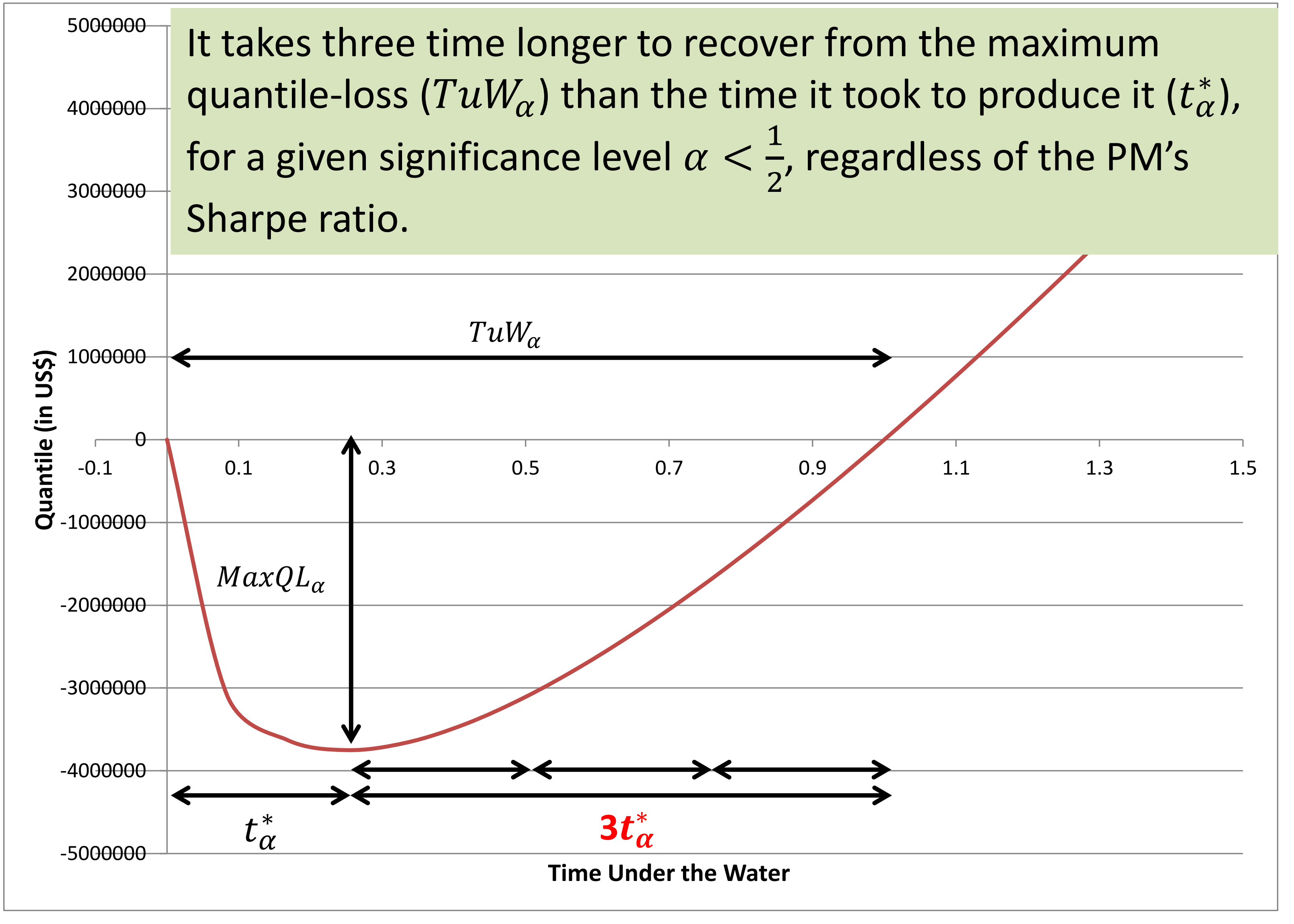
SECTION IV
Decommissioning of Strategies / PMs

The Triple Penance Rule (1/2)

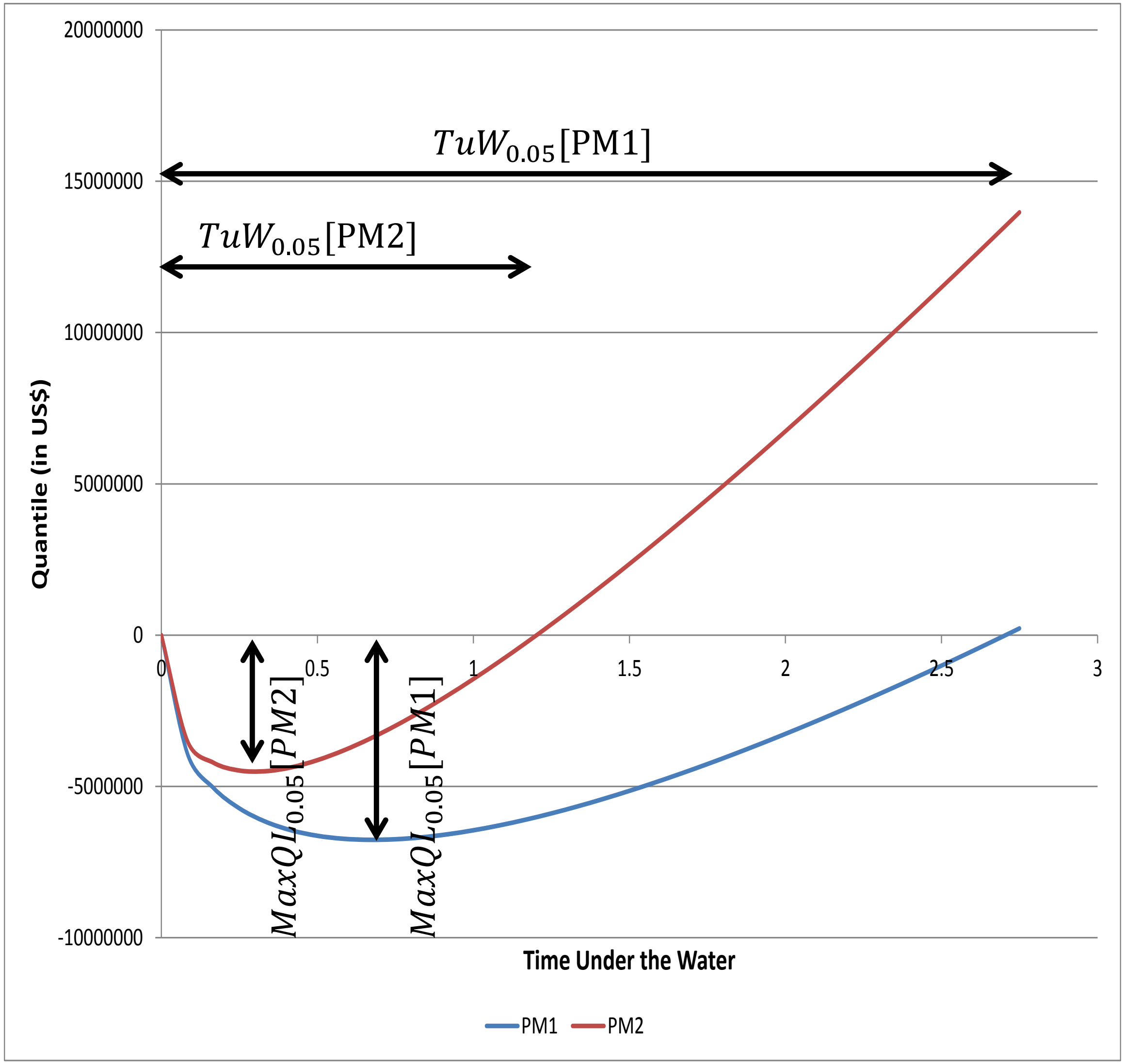
- <u>THEOREM #1</u>: *Under IID Normal outcomes, a strategy's maximum quantile-loss $MaxQL_\alpha$ for a significance level α occurs after t_α^* observations. Then, the strategy is expected to remain under water for an additional $3t_\alpha^*$ after the maximum quantile-loss has occurred, with a confidence $(1-\alpha)$.*
- If we define $Penance := \frac{TuW_\alpha}{t_\alpha^*} - 1$, then the "triple penance rule" tells us that, assuming independent $\Delta\pi_\tau$ identically distributed as Normal (which is the standard portfolio theory assumption), $\boldsymbol{Penance = 3}$, **regardless of the Sharpe ratio of the strategy**.

The Triple Penance Rule (2/2)

It takes three time longer to recover from the maximum quantile-loss (TuW_α) than the time it took to produce it (t_α^*), for a given significance level $\alpha < \frac{1}{2}$, regardless of the PM's Sharpe ratio.



Example 1

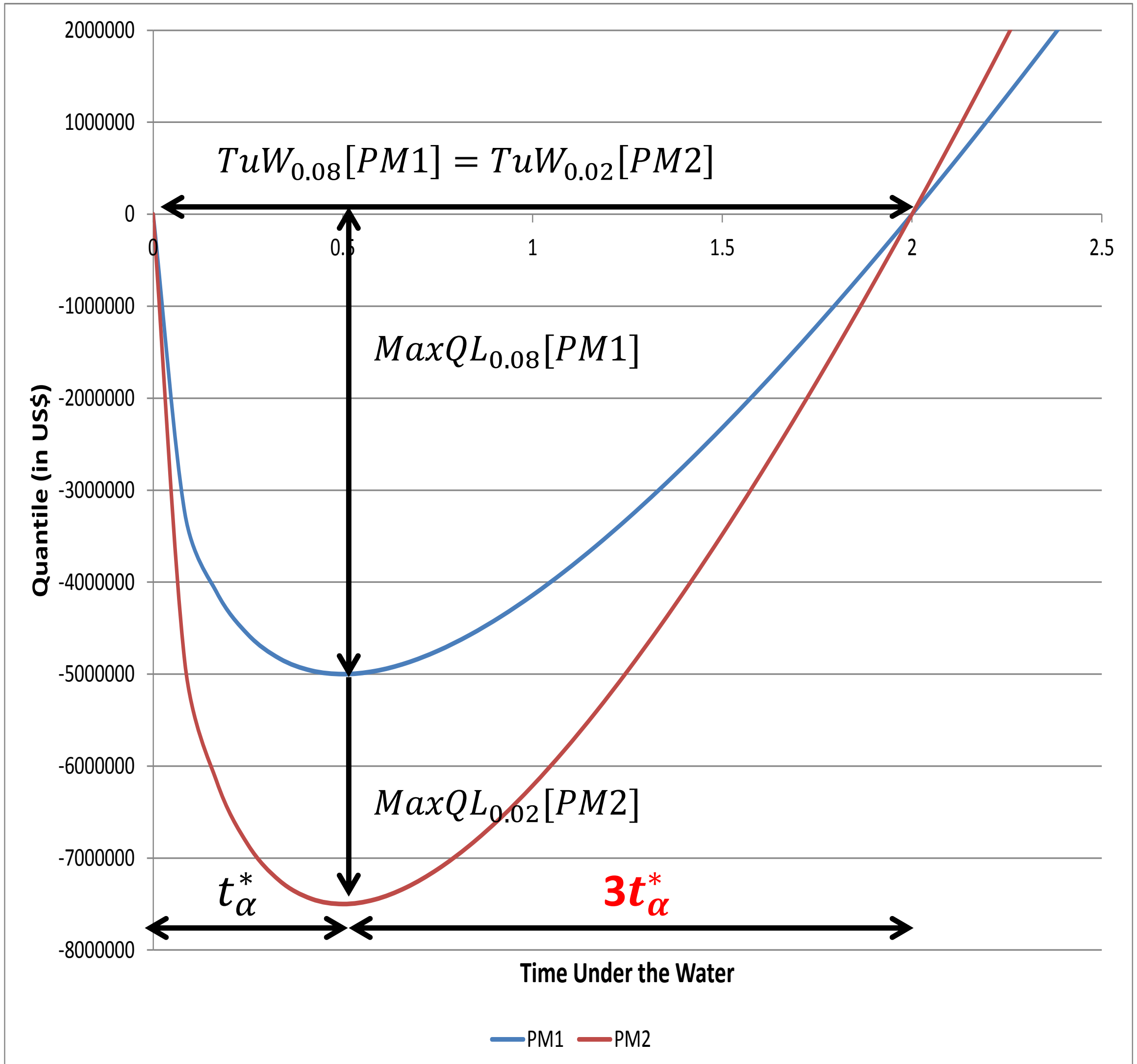


PM1 has an annual mean and standard deviation of US$10m (SR=1), and PM2 has an annual mean of US$15m and an annual standard deviation of US$10m (SR=1.5).

For a 95% confidence level, PM1 reaches a maximum drawdown at US$6,763,859 after 0.676 years, and remains up to 2.706 years under water.

PM2 reaches a maximum drawdown at US$4,509,239 after 0.3 years, and remains 1.202 years under water.

Example 2



PM1 has an annual mean and standard deviation of US$10m (SR=1), and PM2 has an annual mean of US$15m and an annual standard deviation of US$10m (SR=1.5).

For a ~92% confidence level, PM1 reaches a maximum drawdown at US$5,000,000 after 0.5 years, and remains up to 2 years under water.

For a ~98% confidence level, PM2 reaches a maximum drawdown at US$7,500,000 after 0.5 years, and remains up to 2 years under water.

A Better Way to Stop-Out

- Given a realized performance $\tilde{\pi}_t < 0$ and assuming IID Normal returns with mean $\mu > 0$, the Implied Time Under Water (ITuW) is

$$ITuW_{\tilde{\pi}_t} = \frac{{\tilde{\pi}_t}^2}{\mu^2 t} - 2\frac{\tilde{\pi}_t}{\mu} + t$$

- The above equation translates a realized loss into time under water.
- It makes sense stopping-out strategies based on their expected recovery time, rather than waiting for a fixed loss threshold to be hit.

Implications of the Triple Penance Rule

1. It makes possible the translation of drawdowns in terms of time under water.
2. It sets expectations regarding how long it may take to earn performance fee (for a certain confidence level).
   – The remaining time under water may be so long that withdrawals are expected. This has implications for the firm's cash management.
3. It shows that the penance period is independent of the Sharpe ratio (in the IID Normal case).
   – E.g., if a PM makes a fresh new bottom after being one year under water, it may take him 3 years to recover, under the confidence level associated with that loss. This holds true whether that PM has a Sharpe of 1 or a Sharpe of 10.

SECTION V
Conclusions

Pros & Cons of Classic Approaches vs. QMS

Practical Application	Classic approach	Quantitative Meta-Strategy
Hiring **(Example 1)**	Interview candidates with SR (or any other performance statistic) and track record length above a given threshold. Pros: Trivial to implement. Cons: Unknown (possibly high) probability of hiring unskilled PMs.	Design an interview process that recognizes the variables that affect the probability of making the wrong hire: • False positive rate. • False negative rate. • Skill-to-unskilled odds ratio. • Number of independent trials. • Sampling mechanism. Pros: It is objective and can be improved over time, based on measurable outcomes. Cons: More laborious.
Oversight **(Example 2)**	Allocate capital as if PMs were asset classes. Pros: Trivial to implement. Cons: Correlations are unstable, meaningless. Risks are likely to be concentrated.	Recognize that PMs styles evolve over time, as they adapt to a changing environment. Pros: It provides an early signal while the style is still emerging. Allocations can be revised before it is too late. Cons: Allocation revisions may be needed on an irregular calendar frequency.
Stop-Out **(Example 3)**	Stop-out a PM once a certain loss limit has been exceeded. Pros: Trivial to implement. Cons: It allows preventable problems to grow until it is too late.	For any drawdown, large or small, determine the expected time underwater and monitor every recovery. Even if a loss is small, a failure to recover within the expected timeframe indicates a latent problem. Pros: Proactive. Address problems before they force a stop-out. Cons: PMs may feel under tighter scrutiny.

THANKS FOR YOUR ATTENTION!

SECTION VI
The stuff nobody reads

Bibliography

- Bailey, D., J. Borwein, M. López de Prado and J. Zhu (2015): “The Probability of Backtest Overfitting.” **Journal of Computational Finance**, forthcoming. Available at: http://ssrn.com/abstract=2326253
- Bailey, D., J. Borwein, M. López de Prado and J. Zhu (2014): “Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-Of-Sample Performance.” **Notices of the American Mathematical Society**, 61(5), May. Available at: http://ssrn.com/abstract=2308659
- Bailey, D. and M. López de Prado (2012): “The Sharpe Ratio Efficient Frontier.” **Journal of Risk**, 15(2), Winter. Available at http://ssrn.com/abstract=1821643
- Bailey, D. and M. López de Prado (2015): “Stop-Outs Under Serial Correlation and 'The Triple Penance Rule’”, **Journal of Risk**, forthcoming. Available at http://ssrn.com/abstract=2201302
- López de Prado, M. (2015): “Quantitative Meta-Strategies”, **Practical Applications (IIJ)**, 2(3), Spring. Available at http://ssrn.com/abstract=2547325
- López de Prado, M. and M. Foreman (2013): A Mixture of Gaussians Approach to Mathematical Portfolio Oversight: The EF3M Algorithm”, **Quantitative Finance**, 14(5): 913-930. Available at http://ssrn.com/abstract=1931734

Bio

Marcos López de Prado is Senior Managing Director at *Guggenheim Partners*. He is also a Research Fellow at *Lawrence Berkeley National Laboratory*'s Computational Research Division (U.S. Department of Energy's Office of Science), where he conducts unclassified research in the mathematics of large-scale financial problems and supercomputing.

Before that, Marcos was Head of Quantitative Trading & Research at Hess Energy Trading Company (the trading arm of *Hess Corporation*, a Fortune 100 company) and Head of Global Quantitative Research at *Tudor Investment Corporation*. In addition to his 17 years of trading and investment management experience at some of the largest corporations, he has received several academic appointments, including Postdoctoral Research Fellow of *RCC-Harvard University* and Visiting Scholar at *Cornell University*. Marcos earned a Ph.D. in Financial Economics (2003), a second Ph.D. in Mathematical Finance (2011) from *Complutense University*, is a recipient of the National Award for Excellence in Academic Performance by the Government of Spain (National Valedictorian, 1998) among other awards, and was admitted into *American Mensa* with a perfect test score.

Marcos serves on the Editorial Board of the *Journal of Portfolio Management* (IIJ) and the *Journal of Investment Strategies* (Risk). He has collaborated with ~30 leading academics, resulting in some of the most read papers in Finance (SSRN), four international patent applications on High Frequency Trading, three textbooks, numerous publications in the top Mathematical Finance journals, etc. Marcos has an Erdös #2 and an Einstein #4 according to the *American Mathematical Society*.

Disclaimer

- The views expressed in this document are the authors' and do not necessarily reflect those of the organizations he is affiliated with.
- No investment decision or particular course of action is recommended by this presentation.
-

Notice:

The research contained in this presentation is the result of a continuing collaboration with

David H. Bailey, Berkeley Lab
Jon M. Borwein, FRSC, AAAS
Matthew Foreman, UC-Irvine

The full paper is available at:
http://ssrn.com/abstract=2547325

For additional details, please visit:
http://ssrn.com/author=434076
www.QuantResearch.info