Risknowledge
True Risk • Revealed

The Ridge Backtest for Expected Shortfall

Properties and Applications

confidential – restricted distribution

Carlo Acerbi
Larix Risk Consulting and EPFL, Lausanne

Grupo Santander
09/04/2024

Risknowledge
True Risk • Revealed

Clouds over Basel

Basel III: what backtest?

- **2012: BCBS FRTB: <u>replaces</u> $\text{VaR}_{1\%}$ with $\text{ES}_{2.5\%}$**
  - “*Moving from value-at-risk to expected shortfall, a risk measure that better captures "tail risk" …*”
  - Problem: **no valid backtest was known for the ES**
  - **<u>In the absence of an ES backtest, BCBS proposed a mixed test</u>**
    - VaR backtests (at 1% and 2.5%)
    - “P&L Attribution Test” (from 2019: Spearman corr. + Kolmogorov- Smirnov)
  - **Tests are redundant and incomplete at the same time**
    - False positives and false negatives (Type II and I errors)
      - “…*a Russian roulette*” : banks give up IMA

- **2024: Natural question: can ES be directly backtested, at all?**
  - Long-disputed question in the literature

Nov 23, EBA/ECB imposes ES backtesting

- EBA/CP/2023/04 «*Consultation on draft RTS on the assessment methodology under which competent authorities verify an institution's compliance with the internal model approach*»
  - Larix Risk Consulting participated in the consultation. Our response **here**
  - Nov 23, EBA's accepted our recommendation and imposed ES backtesting

- However:
  - EBA don't specify «**which**» ES backtest
  - ES backtest is requested «**in addition to**» the old tests
    - No interest for banks: ISDA pushed back


EBA EUROPEAN BANKING AUTHORITY

EBA/RTS/2023/05

21 November 2023

Final report

Draft regulatory technical standards on the assessment methodology under which competent authorities verify an institution's compliance with the internal model approach as per Article 325az(8) of Regulation (EU) No 575/2013 (Capital Requirements Regulation 2 - CRR2)

Risknowledge
True Risk • Revealed

Jan 24: ISDA-SIFMA response to US Basel III NPR

- **Strong criticism** to several aspects of Basel III NPR

- In particular (pag. 60), they propose .
  - To **remove the PLAT** from the eligibility tests
  - To **replace VaR Backtest with ES Backtest**
    - They mention only **the ridge backtest**

- ISDA-SIFMA observe
  - **Consistency** between the risk measure and the backtest
  - Possibility to directly estimate **capital multipliers**
  - Sensitivity of the backtest to the **magnitude** and not only frequency **of exceedances**

ISDA Safe, Efficient Markets sifma

January 16, 2024

Ann E. Misback
Secretary
Board of Governors of the Federal Reserve System
20th Street and Constitution Avenue N.W.
Washington, D.C. 20551

James P. Sheesley
Assistant Executive Secretary
Attention: Comments/Legal OES (RIN 3064-AF29)
Federal Deposit Insurance Corporation
550 17th Street N.W.
Washington, D.C. 20429

Chief Counsel's Office
Attention: Comment Processing
Office of the Comptroller of the Currency
400 7th Street S.W.
Suite 3E-218
Washington, D.C. 20219

Re: Regulatory capital rule: Amendments applicable to large banking organizations and to banking organizations with significant trading activity

Federal Reserve: Docket No. R-1813, RIN 7100-AG64
FDIC: RIN 3064-AF29
OCC: Docket ID OCC-2023-0008

Dear Sir/Madam,

The International Swaps and Derivatives Association, Inc. ("**ISDA**") and the Securities Industry and Financial Markets Association ("**SIFMA**" and, together with ISDA, the "**Associations**") welcome the opportunity to comment on the proposal referenced above (the "**Proposal**") issued by the Board of Governors of the Federal Reserve System (the "**Federal Reserve**"), the Federal Deposit Insurance Corporation (the "**FDIC**") and the Office of the Comptroller of the Currency (the "**OCC**" and, collectively with the FDIC and the Federal Reserve, the "**Agencies**").

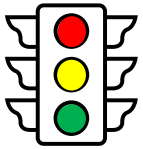
This letter highlights issues arising from the Fundamental Review of the Trading Book ("**FRTB**" or "**market risk**"), credit valuation adjustment ("**CVA**") risk and, with respect to counterparty credit risk ("**CCR**"), aspects of the Proposal relating to securities financing transactions ("**SFTs**") and derivatives. These aspects of the Proposal and the Proposal more generally—in conjunction with the Federal Reserve's stress testing regime and the capital surcharge imposed on U.S. global systemically important banking organizations ("**GSIBs**" and the "**GSIB**

VaR backtest: counting exceptions

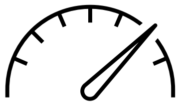
Basel $\mathbf{VaR}_{1\%}$ backtest over $T = 250$ days			
	Number of exceptions	multiplier	Cumulative probability
Green zone	0	1.50	8.106%
	1	1.50	28.575%
	2	1.50	54.317%
	3	1.50	75.812%
	4	1.50	89.219%
Yellow zone	5	1.70	95.882%
	6	1.76	98.630%
	7	1.83	99.597%
	8	1.88	99.894%
	9	1.92	99.975%
Red zone	10 or more	2.00	99.995%

- Basel VaR backtest since 1996

- What **it tells:**
  - **If the model is right or wrong** (and the significance thereof)

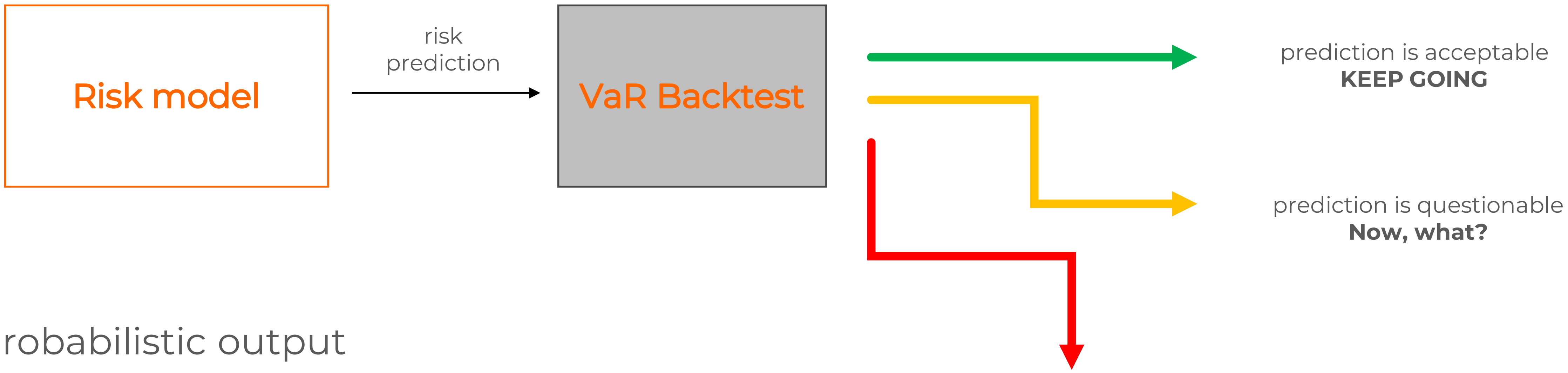


- What **it does not tell:**
  - **The magnitude of the prediction discrepancy,**
    - **hence an estimate of the actual VaR**



* Notice: Basel « multipliers » are just conventional and rosy. Calibrated under Gaussian assumptions for the alternative hypothesis

VaR backtest: model acceptance/rejection



- Probabilistic output
- **No quantification** of prediction gap
- **No control** of the flow: a lottery

Risknowledge
True Risk • Revealed

Backtesting Risk Measures

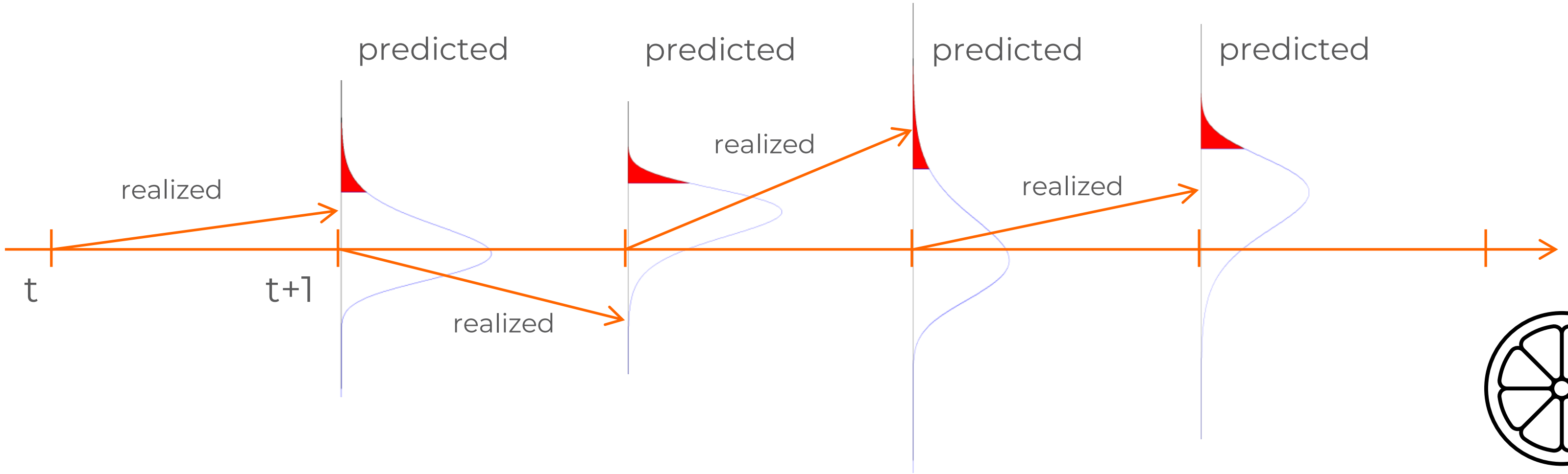
Backtesting: the art of comparing apples to oranges

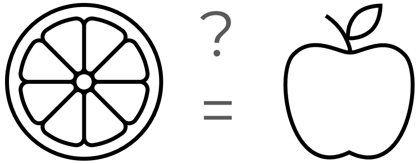
- Backtesting risk measures means validating*
  - **Predictions** $\rho_t = \boldsymbol{\rho}(P_t)$ of the risk measure $\boldsymbol{\rho}(F_t)$

    versus

  - **Realizations** x_t of the P&L r.v. $X_t \sim F_t$

the only quantities you can observe





- ... an **apples** to **oranges** comparison: possible only for certain risk measures

* P_t is the model distribution and F_t is the real-world (unknowable) distribution

A closer look to the VaR backtest

- What makes the quantile (VaR) backtestable?

- $F(X \leq \boldsymbol{q}_\alpha(F)) = \alpha$ (for F cont. in $\boldsymbol{q}_\alpha$)

- The function $Z(q,x) = \alpha - (x \leq q)$
  - <u>Depends only on the observables x and q</u>
  - $\mathbb{E}_F[Z(q,X)]$ is strictly increasing in q (for F strictly increasing)
  - and $\mathbb{E}_F[Z(q,X)] \begin{cases} < 0 & & q < \boldsymbol{q}_\alpha(F) \\ = 0 & iff & q = \boldsymbol{q}_\alpha(F) \\ > 0 & & q > \boldsymbol{q}_\alpha(F) \end{cases}$

- We draw inspiration for a general definition of *backtestability*

Definition of Backtestability (Acerbi and Szekely, 2017)

- $\boldsymbol{y}$ is said to be ***$\mathcal{F}$-backtestable*** if there exists a *backtest function $Z(y, x)$* such that, $\forall F \in \mathcal{F}$
  - $\mathbb{E}_F[Z(y, X)] = 0$ iff $y = \boldsymbol{y}(F)$
  - $y \mapsto \mathbb{E}_F[Z(y, X)]$ strictly increasing

- Intuition*: a backtest must tell if risk is over/under/well- estimated, based only on risk predictions and return realizations*

- Examples:

$\mathbf{y}$	$Z_{\mathbf{y}}(y, x)$	$\mathcal{F}_Z$
$\boldsymbol{\mu}$	$y - x$	maximal
$\mathbf{q}_{1/2}$	$(y > x) - (y < x) + c(x = y)$	$F(x)$ cont. in $\mathbf{q}_{1/2}$ and str. incr.
$\mathbf{q}_{\alpha}$	$(1 - \alpha)(y > x) - \alpha(y < x) + c(x = y)$	$F(x)$ cont. in $\mathbf{q}_{\alpha}$ and str. incr.
$\mathbf{e}_{\alpha}$	$(1 - \alpha)(x - y)_- - \alpha(x - y)_+$	maximal

Risknowledge
True Risk • Revealed

Hypothesis testing: just the usual setup

- Backtestability allows **hypothesis testing** for **model validation**
  - **Test statistic:** *mean backtest realized* over a trailing test window T

$$\bar{z} = \frac{1}{T}\sum_{t=1}^{T} Z(y_t, x_t)$$

  - **Null hypothesis distribution**: *model distribution* $P_{\bar{Z}}$ of the test statistic

$$\bar{Z} = \frac{1}{T}\sum_{t=1}^{T} Z(y_t, X_t) \quad \text{with} \quad X_t \sim P_t,$$

  - **Significance level(s):** e.g. Basel traffic light thresholds $\eta = 95\%, 99{,}99\%$
  - ***p*-value:** $p = \ P_{\bar{Z}}(\bar{z})$
  - **Acceptance/rejection** if $p \gtrless (1 - \eta)$

- Notice the need to store all predictive distributions P_t for computing $P_{\bar{Z}}$
  - Only for $\boldsymbol{VaR}$ this is not necessary because $P_{\bar{Z}}$ is binomial, independent on P

Risknowledge
True Risk • Revealed

Elicitability (Osband 1985, Lambert et al. 2008)

- y is said ***$\mathcal{F}$-elicitable*** if there exists a *scoring function $S(y,x)$* such that, $\forall F \in \mathcal{F}$

$$y(F) = \arg\min_{y} \mathbb{E}_F[S(y,X)]$$

- Examples:

$\mathbf{y}$	$S_{\mathbf{y}}(y,x)$	$\mathcal{F}_S$
$\boldsymbol{\mu}$	$(y-x)^2$	maximal
$\mathbf{q}_{1/2}$	$\lvert y-x \rvert$	maximal
$\mathbf{q}_{\alpha}$	$\alpha(x-y)_+ + (1-\alpha)(x-y)_-$	maximal
$\mathbf{e}_{\alpha}$	$\alpha(x-y)_+^2 + (1-\alpha)(x-y)_-^2$	maximal

- Elicitability necessary to backtestability: $S(y,x) = \int^y Z(t,x)\,dt$
- **Variance (Lambert et al 2008) and ES (Gneiting 2011) are not elicitable,**
  - **hence are not backtestable (Acerbi and Szekely 2017)**

ES is non backtestable!

- (Gneiting 2011) : ES is not elicitable $\Rightarrow$ **ES is not backtestable**

- **Intuition**:
  - For the $\boldsymbol{ES}$ there exists no expression of the type $\mathbb{E}_F[f(\boldsymbol{ES}, X)] = 0$ where f is a function of $\boldsymbol{ES}$ and X only
  - You must include also some other argument, for instance $\boldsymbol{VaR}$:

$$\mathbb{E}_F\left[\boldsymbol{ES}_\alpha - \boldsymbol{VaR}_\alpha - \frac{1}{\alpha}(X + \boldsymbol{VaR}_\alpha)_-\right] = 0$$

- **Consequence:** any attempt to backtest $\boldsymbol{ES}$ necessarily bears some spurious sensitivity to something else (e.g. to $\boldsymbol{VaR}$)

Risknowledge
True Risk • Revealed

L’ES est mort, vive l’ES

However, ES admits a « Ridge backtest »

Model independent !

- However, thanks to Uryasev and Rockafellar's (2001) extremality relationship

$$ES_\alpha = \min_v \left[v + \frac{1}{\alpha}\mathbb{E}_F[(X+v)_-] \right]$$

$$VaR_\alpha = \arg\min_v \left[v + \frac{1}{\alpha}\mathbb{E}_F[(X+v)_-] \right]$$

- ES admits a **"ridge backtest"** (A.Sz. 2017, 2019)

$$Z_{\boldsymbol{ES}_\alpha}(e,v,x) = e - v - \frac{1}{\alpha}(x+v)_-$$

$$\mathbb{E}_F[Z_{\boldsymbol{ES}_\alpha}(e,v,X)] = e - \boldsymbol{ES}_\alpha(F) - B(v)$$

with bias $B(v) = \mathbb{E}_F\left[v + \frac{1}{\alpha}(X+v)_- \right] - \boldsymbol{ES}_\alpha(F) \geq 0$

- … whose sensitivity to VaR is zero at 1st order, and one-sided
- Small, prudential sensitivity to mispredictions $v \neq \boldsymbol{VaR}_\alpha$

Déjà vu?

find ES: replace variance
find VaR: replace mean

- Perfect analogy with variance

$$\boldsymbol{\sigma}^{2} = \min_{m} \mathbb{E}_F[(X-m)^2]$$

$$\boldsymbol{\mu} = \arg\min_{m} \mathbb{E}_F[(X-m)^2]$$

- The variance σ^2 admits a "ridge backtest"

$$Z_{\boldsymbol{\sigma}^2}(v,m,x) = v - (x-m)^2$$

$$\mathbb{E}_F\left[Z_{\boldsymbol{\sigma}^2}(v,m,X)\right] = v - \boldsymbol{\sigma}^2(F) - B(m)$$

with bias $B(m) = \mathbb{E}_F[(X-m)^2] - \boldsymbol{\sigma}^2(F) \geq 0$

- Whose sensitivity to the mean $\boldsymbol{\mu}$ is zero at 1st order, and one-sided
- Small, prudential sensitivity to mispredictions $m \neq \mu$

Risknowledge
True Risk • Revealed

Non backtestable, de facto backtestable

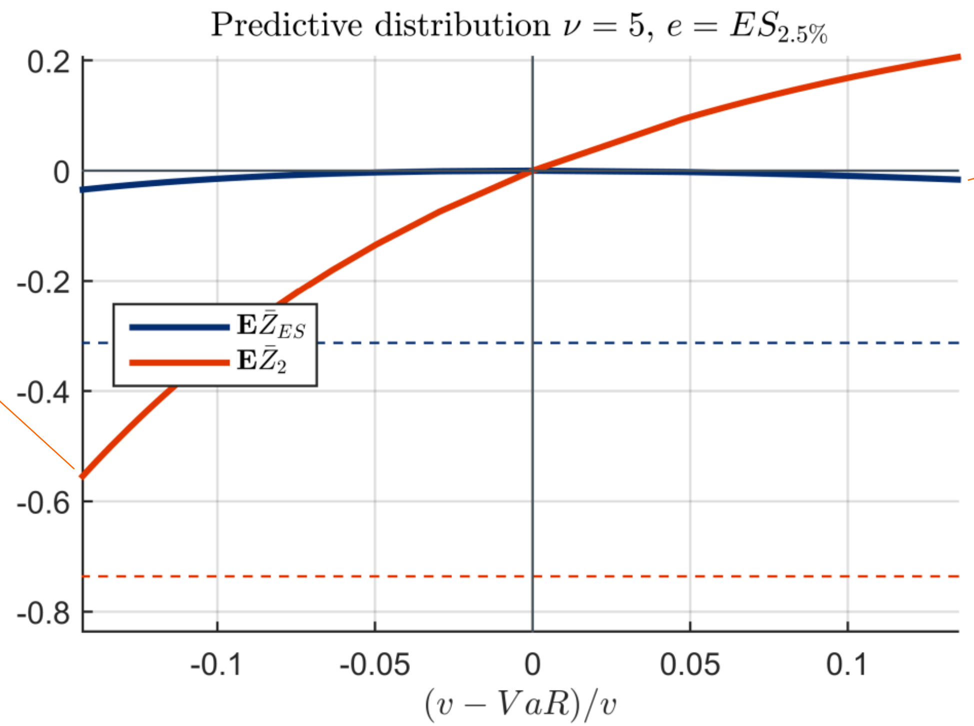
- ES and Variance are **backtestable up to a small and prudential bias**
  - Here's why Variance is commonly used/backtested without much drama
- The ridge backtest for ES is **unique**:

Any other backtest for ES suffers from larger and 2-sided bias

Example: sensitivity to VaR

- Varying F's with identical $ES(F)$ and different $VaR(F)$
- **We assume that ES predictions are correct**: $e = ES_\alpha(X)$
  - Z_2 (A.Sz. 2014) linear strong sensitivity; Z_{ES} (A.Sz. 2017) muted, quadratic sensitivity

Any other backtest
Significant bilateral bias



Ridge backtest
Small and prudential bias

Risknowledge
True Risk • Revealed

Risknowledge
True Risk • Revealed

Mind the Gap

”Realized Expected Shortfall”

The ES backtest is the difference between realized and predicted risk !

- The **realized backtest function** $\bar{z}_{ES}$ can be written as

$$\bar{z}_{ES}(e,v,x) \equiv \sum_{t=1}^{T} Z(e_t,v_t,x_t) = \frac{1}{T}\sum_t e_t - \widehat{ES}_\alpha$$

Where the **“realized ES”** (in perfect analogy with *realized variance*)

$$\widehat{ES}_\alpha = \frac{1}{T}\sum_t \left[v_t + \frac{1}{\alpha}(x_t+v_t)_-\right]$$

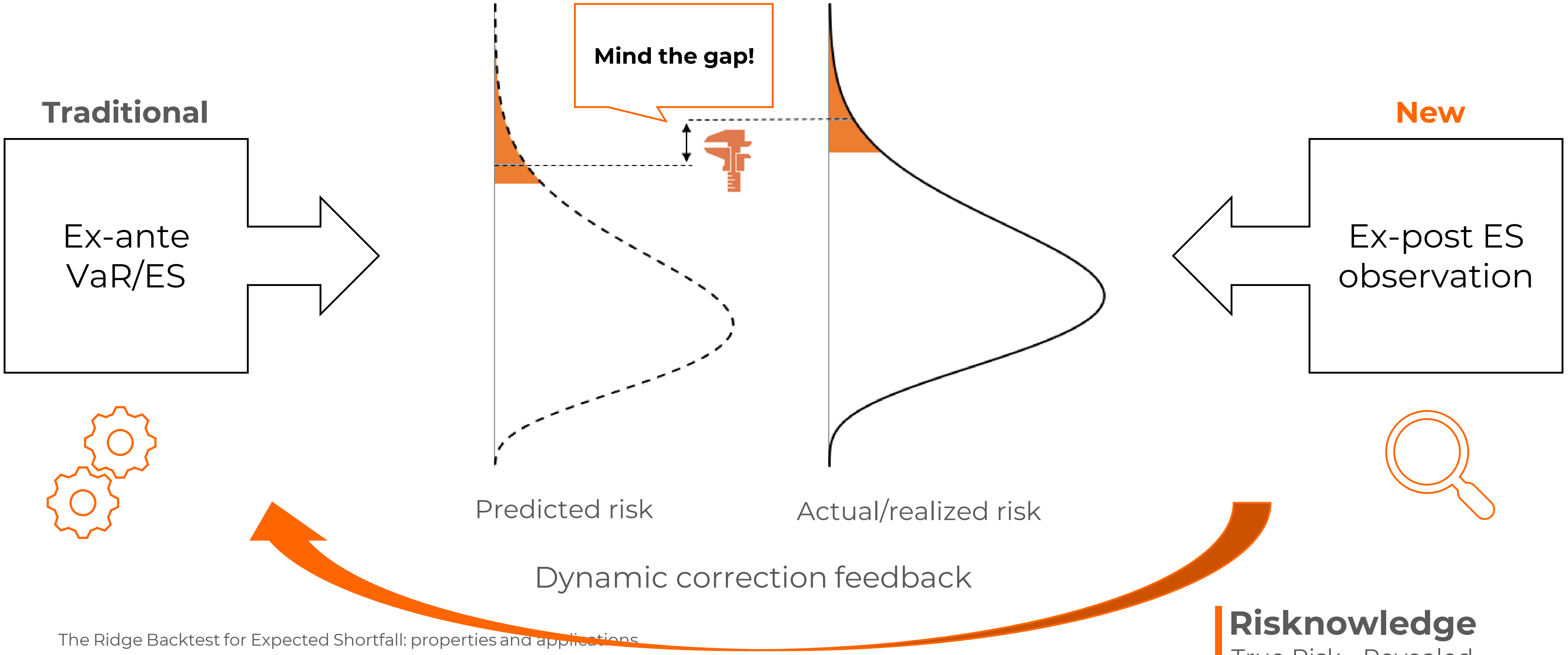
is a biased estimator of the **average true ES** : $\mathbb{E}_F\left[\widehat{ES}_\alpha\right] \geq \frac{1}{\mathrm{T}}\sum_t ES_\alpha(F_t)$

- Important consequences:
  - **ES** can be “observed” ex-post, on average, as opposed to **VaR**
  - The backtest is an apples to apple comparison: it is a measure of discrepancy between average predictions and realizations of **ES**
  - Follows from the *sharpness* of the ridge backtest (A.Sz. 2017)

Risknowledge
True Risk • Revealed

Ex-post Risk Analytics

✓ **Ridge ES backtest**: a comparison between **predicted** and **realized** risk



Absolute and relative ES backtests

- (**Absolute**) **backtest :** denominated in monetary terms: absolute discrepancy between predicted and realized ES

$$Z_{\boldsymbol{ES}_\alpha}(e,v,x) = e - v - \frac{1}{\alpha}(x+v)_-$$

$$\mathbb{E}_F[Z_{\boldsymbol{ES}_\alpha}(e,v,X)] = e - \boldsymbol{ES}_\alpha(F) - B(v) \ \leq e - \boldsymbol{ES}_\alpha(F)$$

- **Relative backtest :** dimensionless, renormalised test: relative discrepancy between predicted and realised ES

$$Z^{Rel}_{\boldsymbol{ES}_\alpha}(e,v,x) \ \equiv \frac{Z_{\boldsymbol{ES}_\alpha}(e,v,x)}{e}$$

Less obvious than it seems: still monotonic wrt e?

$$\mathbb{E}_F[Z^{Rel}_{\boldsymbol{ES}_\alpha}(e,v,X)] = \frac{e - \boldsymbol{ES}_\alpha(F) - B(v)}{e} \leq \frac{e - \boldsymbol{ES}_\alpha(F)}{e}$$

Dynamic multipliers

- From the relative realized backtest we obtain a **realised prediction ratio**

$$\hat{\phi}_{ES} \equiv \frac{1}{T}\sum_{t}\left[\frac{v_t + \frac{1}{\alpha}(x_t + v_t)_-}{e_t}\right]$$

which is a positively biased estimator of the **average prediction ratio** $\boldsymbol{ES/e}$

$$\mathbb{E}_F\left[\hat{\phi}_{ES}\right] = \frac{1}{\mathrm{T}}\sum_{t}\frac{\boldsymbol{ES}_\alpha(F_t) + B(v_t)}{e_t} \geq \frac{1}{\mathrm{T}}\sum_{t}\frac{\boldsymbol{ES}_\alpha(F_t)}{e_t}$$

- $\hat{\phi}_{ES}$: portfolio/model-specific dynamic multiplier

Risknowledge
True Risk • Revealed

Adaptive models: dynamic capital multpliers

- The ES Ridge backtest directly measures portfolio-specific multipliers
  - $\hat{\phi}$ prudentially biased
  - γ bias-free

Risk scaling factors										
		$\mathbf{VaR}_{1\%}$	$\mathbf{ES}_{2.5\%}$							
Model distribution			$\nu = \infty$	$\nu = 10$	$\nu = 5$	$\nu = 3$	$\nu = \infty$	$\nu = 10$	$\nu = 5$	$\nu = 3$
	$\eta = 1 - p$	Basel	$\hat{\phi}_{\mathbf{ES}}$				γ			
Green zone	8.106%	1.00	0.89	0.85	0.81	0.72	0.85	0.80	0.72	0.60
	28.575%	1.00	0.94	0.92	0.89	0.83	0.94	0.91	0.87	0.80
	54.317%	1.00	1.00	1.00	0.99	0.96	1.00	0.99	0.99	0.96
	75.812%	1.00	1.05	1.07	1.09	1.12	1.05	1.07	1.08	1.11
	89.219%	1.00	1.11	1.15	1.21	1.31	1.09	1.13	1.18	1.26
Yellow zone	95.882%	1.13	1.17	1.24	1.34	1.55	1.13	1.19	1.26	1.41
	98.630%	1.17	1.23	1.33	1.49	1.86	1.17	1.24	1.35	1.59
	99.597%	1.22	1.29	1.42	1.66	2.31	1.21	1.29	1.44	1.82
	99.894%	1.25	1.35	1.52	1.86	3.01	1.24	1.34	1.54	2.13
	99.975%	1.28	1.42	1.62	2.14	4.24	1.27	1.39	1.65	2.61
Red zone	99.995%	1.33	1.48	1.72	2.52	6.36	1.30	1.43	1.78	3.34

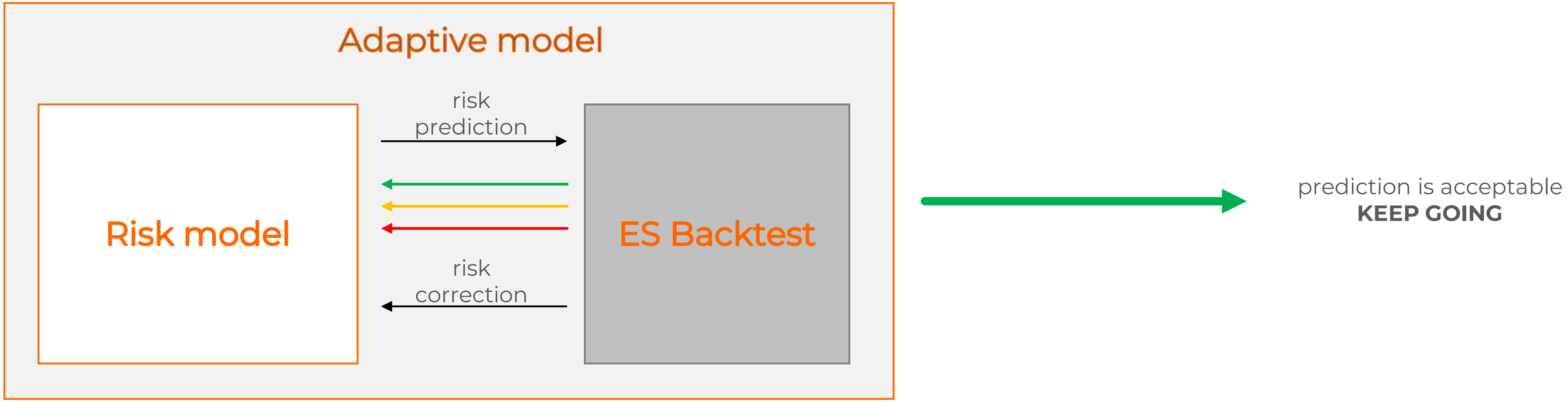
Prudential

Unbiased

Conventional VaR multipliers

Model-specific ES multipliers

ES Ridge backtest: model correction



- **Actionable feedback** for model correction
- **Control**: business continuity, no risk of model fall-back
- Capital **planning** possible
- The backtest is part of the model

Risknowledge
True Risk • Revealed

Actual vs Predicted Risk: Now we can

Traditional Financial Risk Models: limitations

- **Banks[1]: risk models limitations**
  - Current risk models output only **ex-ante predictions** for VaR and ES, not **true risk**
  - **Risk is managed based only on (unvalidated) predictions**
    - Potentially distorted decisions
  - VaR backtest tells you if predictions are right/wrong
  - But **true VaR is fundamentally unknowable**
    - Acerbi and Szekely, “Backtestability and the Ridge Backtest”, Frontiers of Mathematical Finance Dec 2023.
    - See proposition 3.17 and remark 3.18

Frontiers of Mathematical Finance
doi:10.3934/fmf.2023019

BACKTESTABILITY AND THE RIDGE BACKTEST

Carlo Acerbi[*1,2] and Balázs Székely[†3]
[1]Larix Risk Consulting, Genève, Switzerland
[2]École Polytechnique Fédérale de Lausanne, Switzerland
[3]Morgan Stanley, Budapest, Hungary

(Communicated by Wim Schoutens)

Abstract. We propose a formal definition of backtestability for a statistical functional of a distribution: a functional is backtestable if there exists a backtest function depending only on the forecast of the functional and the related random variable, which is strictly monotonic in the former and has zero expected value for an exact forecast. We discuss the relationship with elicitability and identifiability which turn out being necessary conditions for backtestability. The variance and the expected shortfall are not backtestable for this reason. We compare (absolute) model validation in the context of hypothesis tests via backtest functions, versus (relative) model selection between competing forecasting models, via scoring functions. We define a backtest to be sharp when it is strictly monotonic with respect to the real value of the functional and not only to its forecast. This decides whether the expected value of the backtest determines also the prediction discrepancy and not only its significance. We show that the quantile backtest is not sharp and in fact it provides no information whatsoever on its true value. The expectile is also not sharp; we provide bounds for its true value, which are looser for outer confidence levels. We then introduce the notion of ridge backtests, applicable to particular non–backtestable functionals, such as the variance and the expected shortfall, which coincide with the attained minimum of the scoring function of another elicitable auxiliary functional (the mean and the value at risk, respectively). This permits approximated sharp backtests up to a small and one–sided sensitivity to the prediction of the auxiliary variable. The ridge mechanism explains why the variance has always been de–facto backtestable and allows for similar efficient ways to backtest the expected shortfall. We discuss the relevance of this result in the current debate of financial regulation (banking and insurance), where value at risk and expected shortfall are adopted as regulatory risk measures.

1. **Motivation and objectives.** Backtest procedures for testing forecasts of a risk measure are common in the statistical and probabilistic disciplines, notably in the field of financial risk management. Backtesting in general, however, remains to date a collection of disparate practices in the wait for a clear definition. Whether

2020 *Mathematics Subject Classification.* 91G70, 91G60, 91B05, 62G10, 60G25.
Key words and phrases. Backtestability, elicitability, expected shortfall, value at risk, sharp backtests, ridge backtests.
*Corresponding author: Carlo Acerbi.
†This paper is the private opinion of the author and does not necessarily reflect the policy and views of Morgan Stanley.

1

1. Not only banks, but financial institutions more generally, including banks, insurances, funds, hedge funds, clearinghouses, prime brokers, etc.

Risknowledge
True Risk • Revealed

Report: Model Validation : Predicted vs Actual ES

Risknowledge

_hyp
_emp
_10
30/03/2023

	VaR	ES			multipliers			realized ES	
	RAG	average	RAGr	RAGa	phi	gamma	delta	raw	unbiased
Ptf000		5,368,760.70			1.653	1.307	2.120	6,936,292.14	7,019,329.46
Ptf001		5,125,851.88			2.076	1.444	2.649	8,359,460.28	7,400,556.29
Ptf002		4,620,981.40			2.019	1.425	2.603	7,706,009.65	6,583,515.82
Ptf003		2,367,235.57			0.882	0.795	0.835	2,026,095.25	1,881,935.87
Ptf004		2,296,751.76			0.935	0.918	0.897	1,933,999.91	2,108,812.34
Ptf005		31,776.82			0.931	0.894	0.943	27,639.79	28,417.04
Ptf006		269,697.45			1.442	1.249	1.332	395,996.26	336,885.35
Ptf007		15,266.41			0.963	0.954	0.927	13,954.02	14,563.82
Ptf008		682,806.75			1.601	1.314	1.280	1,131,933.60	897,350.10
Ptf009		29,944.54			0.947	0.932	0.964	26,689.65	27,921.47
Ptf010		683,088.02			1.590	1.306	1.279	1,126,757.85	891,911.15
Ptf011		46,039.04			1.059	1.045	1.035	49,366.04	48,097.29
Ptf012		694,526.72			1.657	1.341	1.302	1,189,388.39	931,330.40
Ptf013		24,139.56			1.185	1.131	1.148	25,108.69	27,299.73
Ptf014		1,644,278.61			2.584	1.567	1.583	4,103,716.15	2,575,992.11
Ptf015		557,413.11			10.234	2.824	3.076	1,429,089.05	1,573,948.01
Ptf016		1,249,400.59			2.870	1.624	1.742	3,506,561.10	2,029,262.92
Ptf017		306,108.59			1.522	1.320	1.940	423,150.83	404,149.03

predicted

actual

unbiased multiplier

All what VaR could have told us

Strictly speaking, we should close desk ptf014

Risknowledge

_hyp
_emp
_10
30/03/2023

Portfolio	VaR RAG	average	
Ptf000		5,368,760.70	
Ptf001		5,125,851.88	
Ptf002		4,620,981.40	
Ptf003		2,367,235.57	
Ptf004		2,296,751.76	
Ptf005		31,776.82	
Ptf006		269,697.45	dubious
Ptf007		15,266.41	
Ptf008		682,806.75	dubious
Ptf009		29,944.54	
Ptf010		683,088.02	dubious
Ptf011		46,039.04	
Ptf012		694,526.72	dubious
Ptf013		24,139.56	
Ptf014		N/A	
Ptf015		N/A	
Ptf016		N/A	
Ptf017		306,108.59	

ES ridge backtest: Use cases

- Model validation: FRTB IMA but not only
- Prediction error quantification
- Estimation of actual average ES
  - Actual risk management
    - Risk budgeting
    - Risk Appetite Framework limits
    - Risk-adjusted performance measurement
- Adaptive models
  - ES models corrected via estimated multipliers
    - Daily risk controls and limits

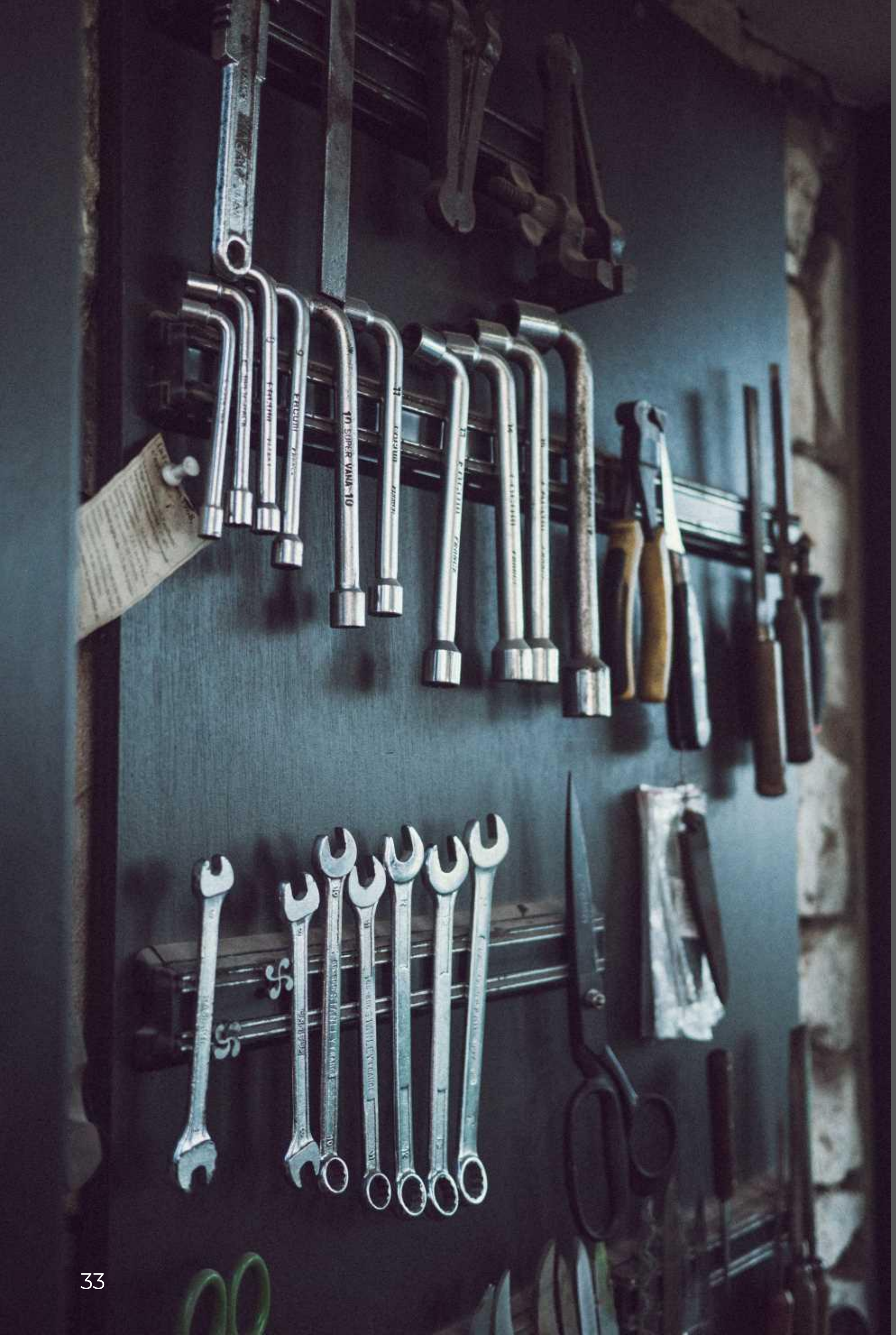
CRR III: Internal models be aligned with daily Risk models

- **Alignment** between Internal Risk Management models and Regulatory Capital models

- **CRR III : Article 325bi**
  - *1. Any internal risk-measurement model used for the purposes of this Chapter shall be conceptually sound, shall be calculated and implemented with integrity, and shall comply with all the following qualitative requirements:*

    a) any internal risk-measurement model used to calculate capital requirements for market risk shall be closely integrated into the daily risk management process of the institution and shall serve as the basis for reporting risk exposures to senior management;

- **If FRTB imposes ES for Pillar I, banks must ensure consistency with Pillar II metrics.**



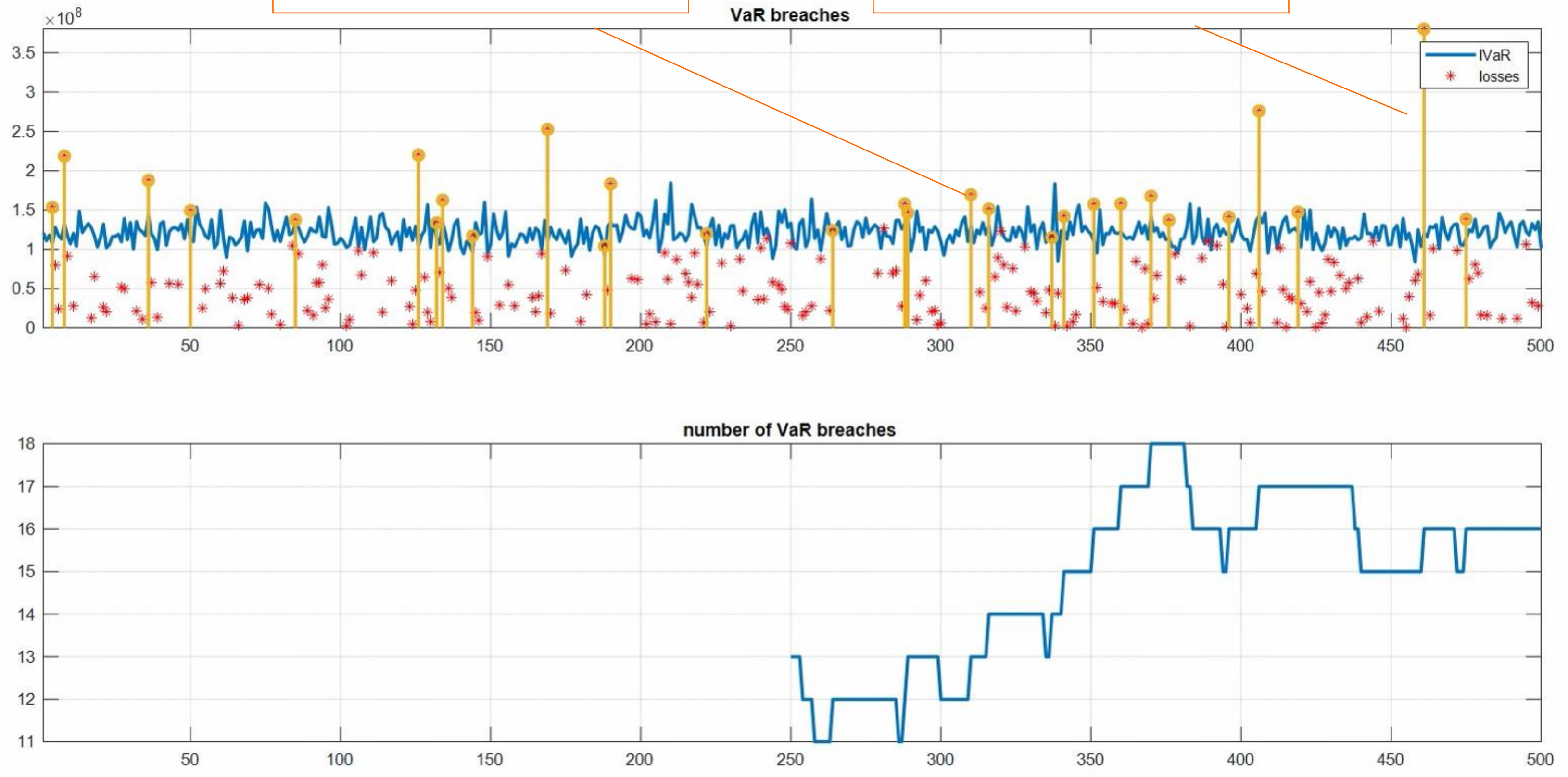
Risknowledge
True Risk • Revealed

A stylized example

VaR breaches (expected 6.25 per year)

Many small breaches. A serious issue for VaR Backtest

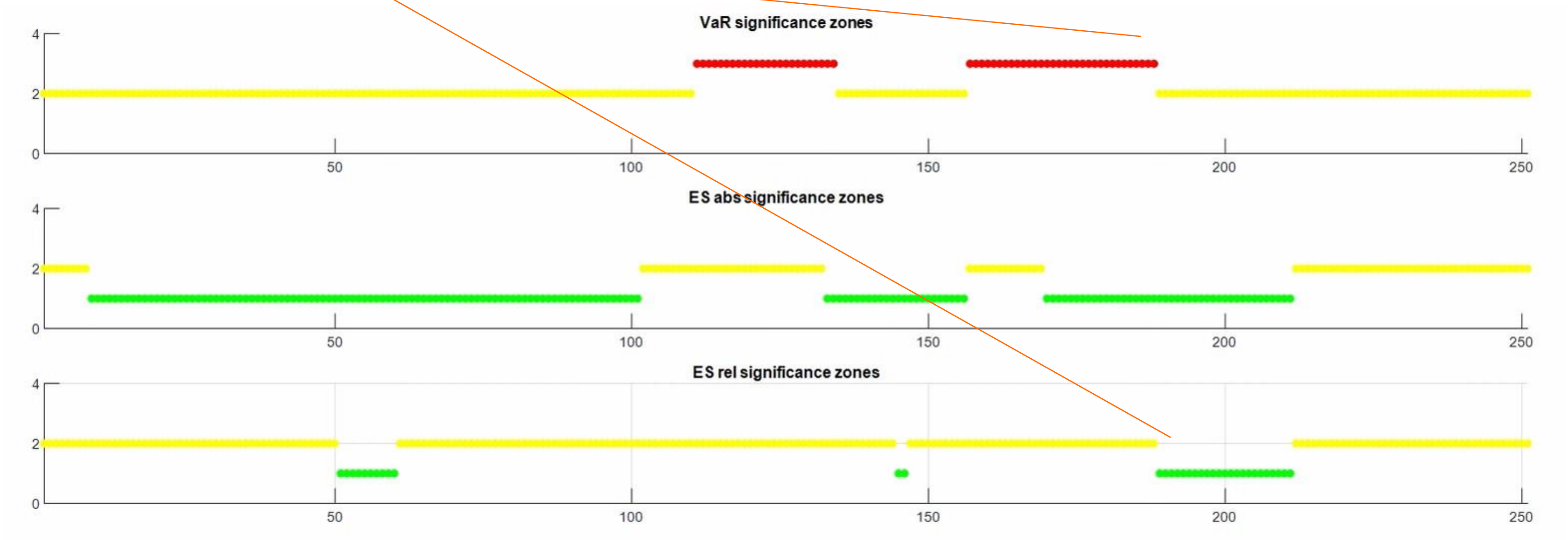
Huge breaches make no difference for VaR



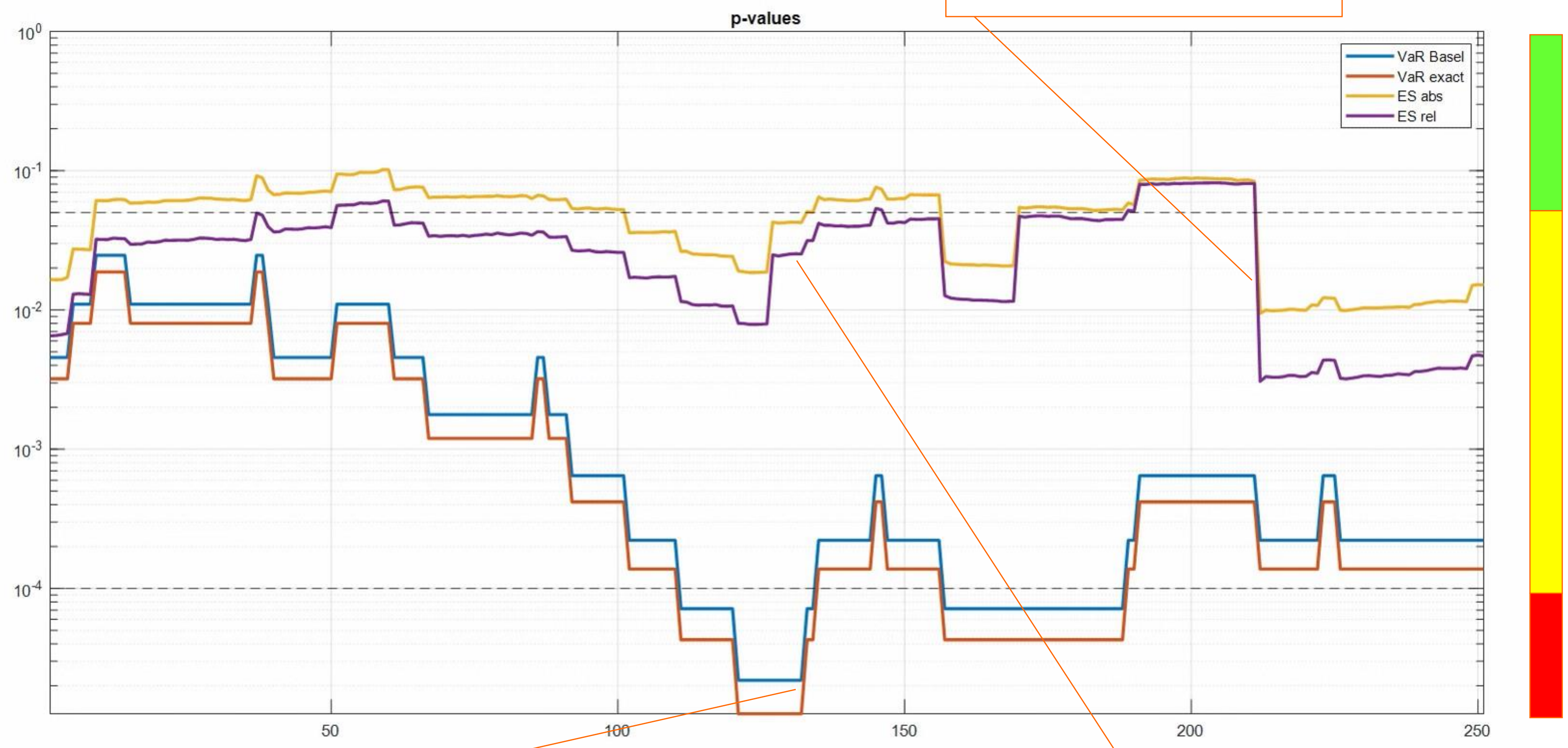
Risknowledge
True Risk • Revealed

Traffic light: VaR and ES (abs and rel)

ES and VAR RAG may be very different: small breaches almost insignificant for ES



p-values



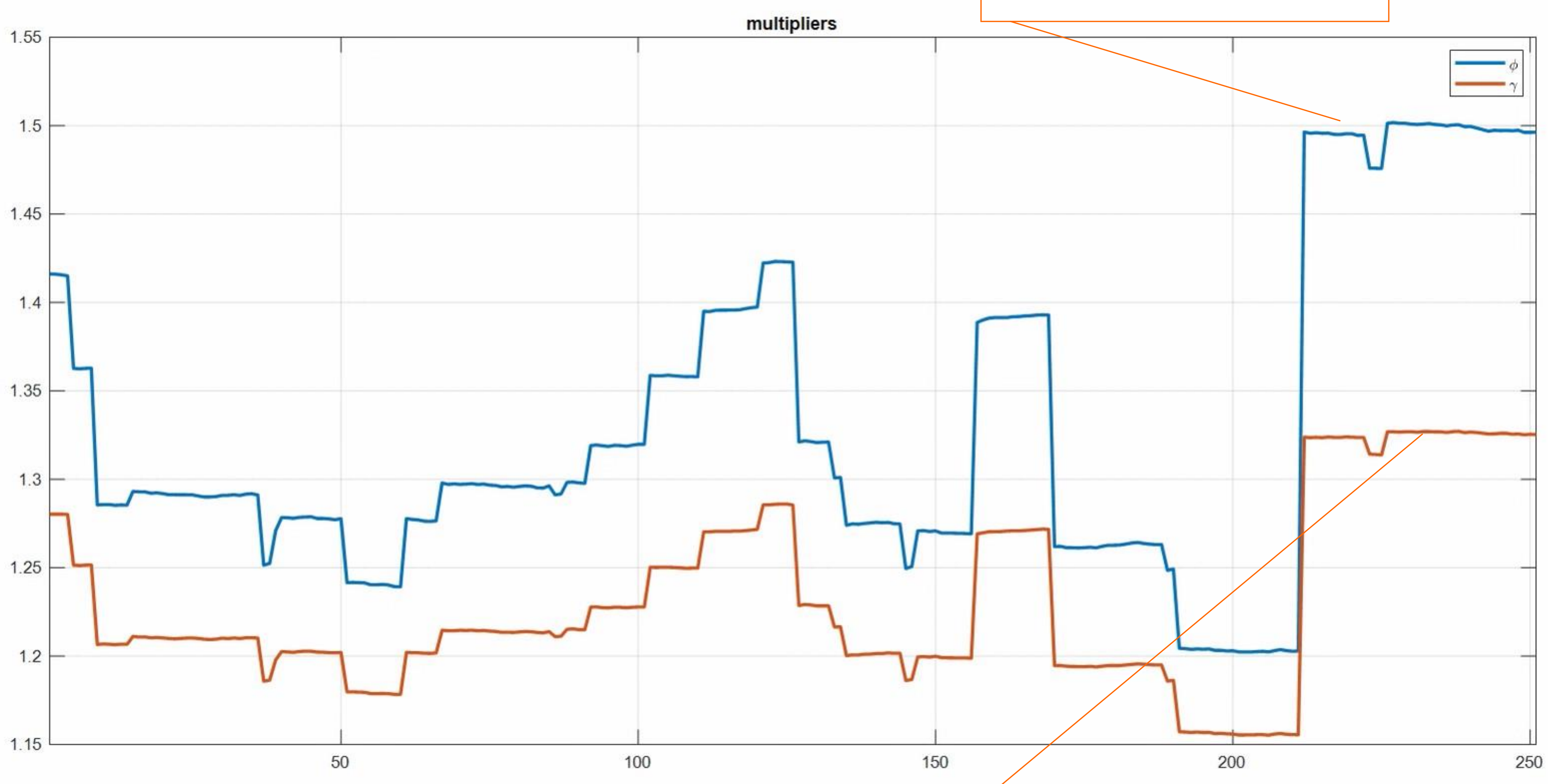
Huge breaches make a big difference for ES

Persistent red zone: pointless « *game over* » message from VaR

ES backtests ensure business continuity, be it in the green, amber or red zones

Dynamic multipliers

Multiplier ϕ
Prudentially biased



Multiplier γ
Bias-neutral

Risknowledge
True Risk • Revealed

Risknowledge
True Risk • Revealed

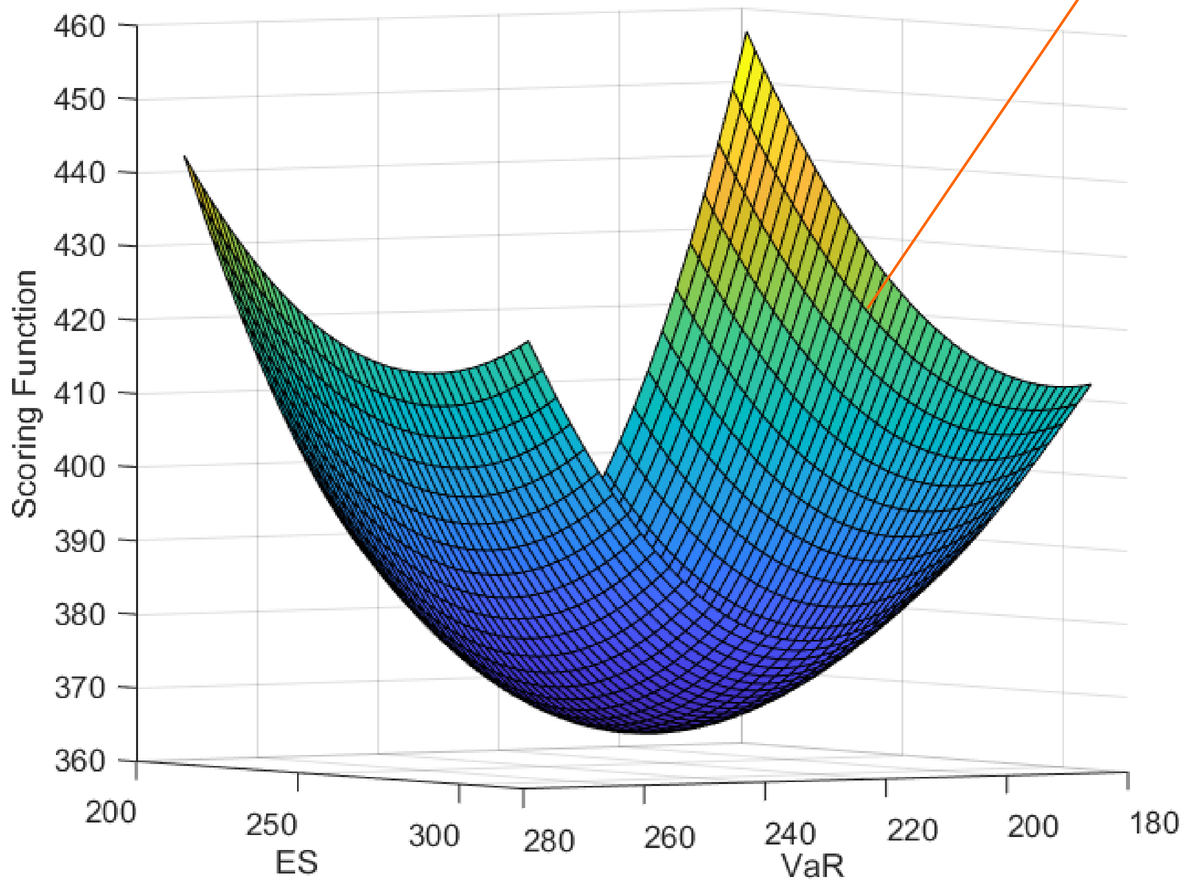
Model selection example: exp-weighted vs eq-weighted

Model Selection for joint {VaR, ES} predictions

- Recent results on **elicitability** permit to define **advanced methods for model selection**, based on **realized scoring functions** of elicitable risk measures
- For competing predictive models, a **lower scoring function defines a better model**
  - **Relative comparison** between alternative candidate models
    - The realized scoring function doesn't provide an absolute assessment of the quality of predictions of a single model. For that you need a backtest.
  - Scoring functions **penalize over- and under-estimation bilaterally**. A "better model" in this selection is not necessarily the more prudent.

- ES is not elicitable, hence doesn't have a scoring function
- However, **the pair {VaR, ES} admits joint scoring functions** (A., Sz., 2014), related to the ridge backtest, which permits model selection for better predictions of the two risk measures.

Joint elicitability



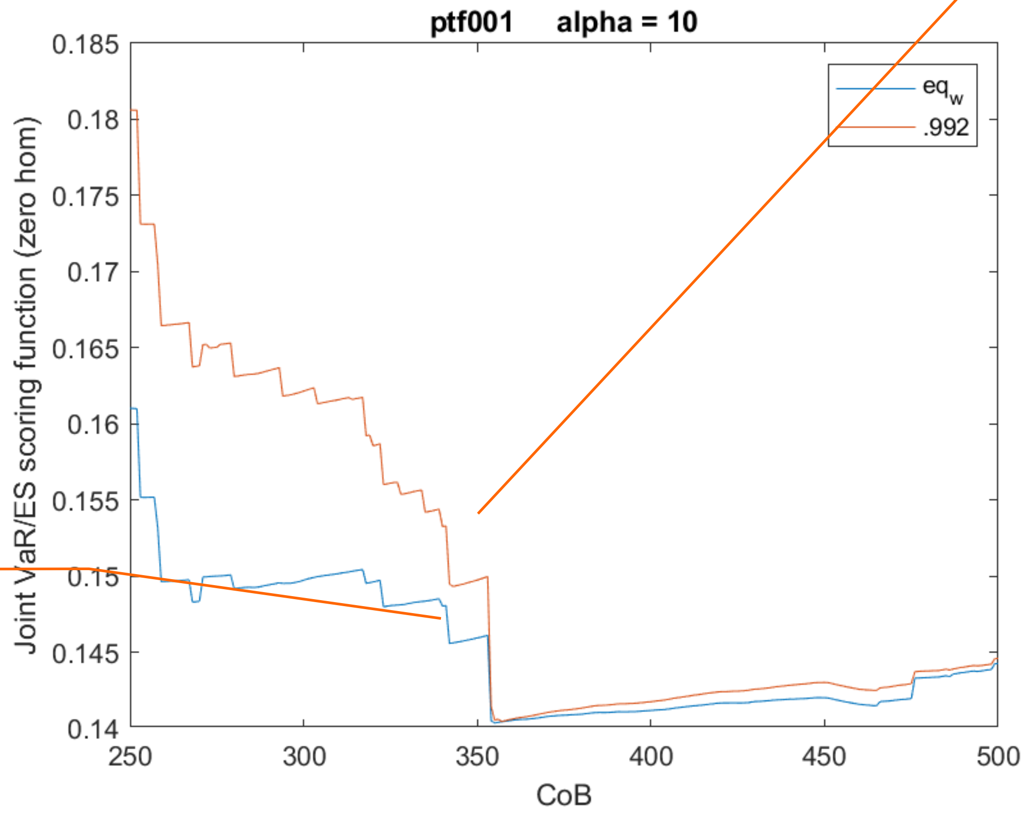
Joint VaR-ES scoring function

$$\{VaR, ES\} = \arg\min_{v,e} \mathbb{E}[S(v,e,X)]$$

Acerbi, Szekely 2014, RISK

Model selection: ewma vs equally weighted. Real case.

The lower the realized scoring function, the better the model



In this specific case, eq_w performs consistently better across the period

Risknowledge
True Risk • Revealed

Risknowledge
True Risk • Revealed

Conclusions

Conclusions

- **The ridge backtest** is the only possible prudential backtest for ES. It is affected by the lowest possible bias, independently on the model.

- Moreover, this test automatically **measures the actual ES** (« realized ES »)
  - And **metrics of prediction discrepancy,** in relative or absolute terms
  - Apple-to-apple comparison between **actual and predicted ES**

- **Banks:**
  - Can directly **backtest ES for ES-based risk models**
  - Can finally reveal and manage **actual risk**
  - Can use backtest results for **correcting model predictions**

Risknowledge
True Risk • Revealed

References

Main References

- Acerbi, C. and Szekely, B. (2014), «Backtesting ES» RISK
- Acerbi, C. and Szekely, B. (2017), «General Properties of Backtestable Statistics», working paper, SSRN (final version published in 2023)
- Acerbi, C. and Szekely, B. (2019), «The minimally biased backtest for ES» RISK
- Acerbi, C. and Szekely, B. (2023), «Backtestability and the Ridge Backtest», Frontiers of Mathematical Finance
- Basel Committee on Banking Supervision (2012): «Fundamental Review of the trading book»
- Basel Committee on Banking Supervision (2019): «Minimum capital requirements for market risk»
- Gneiting, T. (2001), «Making and evaluating point forecasts», JASA
- Osband, K.H. (1985), « Providing incentives for better cost forecasting». PhD Thesis. Univ. Calif., Berkeley

Risknowledge
True Risk • Revealed

Q&A