Median-Anchored Adjustment of Joint VaR–ES Forecasts

Dingli Wang Tae-Hwy Lee

Department of Economics, University of California, Riverside

Motivation

- For a bank or risk manager, rebuilding a risk model is costly (new specification, validation, approval) $\Rightarrow$ the model is kept and its tail forecasts are adjusted
- Example — Basel backtesting: too many VaR exceptions $\Rightarrow$ a larger capital multiplier on the model's VaR; the model itself is not changed

Question

Using only the model's reported median, VaR and ES, how should the tail be adjusted — with **one parameter**, keeping the model's dynamics and the VaR–ES ordering?

Recalibrating by regression can break the VaR–ES order

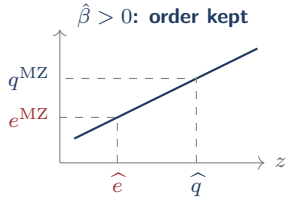
Quantile Mincer–Zarnowitz (MZ) regression (Gaglianone et al., 2011, *JBES*): regress the return on the VaR forecast at level τ_0; the fitted line is the new VaR,

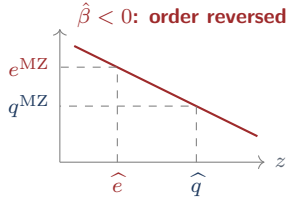
$$q_t^{\mathrm{MZ}} = \hat{\alpha} + \hat{\beta}\,\widehat{q}_t.$$

ES must move with VaR, so it gets the same line, $e_t^{\mathrm{MZ}} = \hat{\alpha} + \hat{\beta}\,\widehat{e}_t$:

$$e_t^{\mathrm{MZ}} - q_t^{\mathrm{MZ}} = \hat{\beta}\,(\underbrace{\widehat{e}_t - \widehat{q}_t}_{\leq 0})$$

- $\hat{\beta} > 0$: order kept
- $\hat{\beta} < 0$: ES above VaR (invalid)





z: baseline forecast; line: $\hat{\alpha} + \hat{\beta} z$

This paper

Median-anchored adjustment

Keep the fitted median; multiply the median-to-VaR and median-to-ES distances by one common $\kappa > 0$.

- **Always a valid pair:** ES stays below VaR for every κ
- **Simple:** one parameter, estimated by sorting; needs only the reported median, VaR and ES
- **Theory:** consistency and asymptotic normality; coverage-error bound and lower FZ0 risk

The rule: rescale tail distances from the median

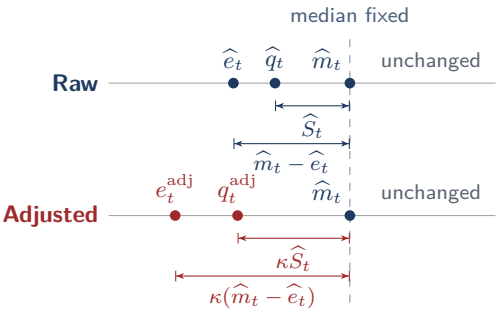
Baseline at level $\tau_0 < 0.5$: median $\widehat{m}_t$, VaR $\widehat{q}_t$, ES $\widehat{e}_t$, with $\widehat{e}_t \le \widehat{q}_t < \widehat{m}_t$.

$$\mathcal{A}_{\kappa,t}(z) = \begin{cases} z, & z \ge \widehat{m}_t, \\ \widehat{m}_t - \kappa(\widehat{m}_t - z), & z < \widehat{m}_t. \end{cases}$$

Proposition 1

$$q_t^{\mathrm{adj}}(\kappa) = \widehat{m}_t - \kappa(\widehat{m}_t - \widehat{q}_t)$$
$$e_t^{\mathrm{adj}}(\kappa) = \widehat{m}_t - \kappa(\widehat{m}_t - \widehat{e}_t)$$

$\Rightarrow \; e_t^{\mathrm{adj}} \le q_t^{\mathrm{adj}} < \widehat{m}_t$ for *every* $\kappa > 0$.



$\widehat{S}_t := \widehat{m}_t - \widehat{q}_t$; illustration with $\kappa = 1.5$

$\kappa > 1$ widens, $\kappa < 1$ narrows the lower tail; Raw is $\kappa = 1$.

MZ form: $q_t^{\mathrm{adj}} = (1-\kappa)\widehat{m}_t + \kappa\widehat{q}_t$, slope $\kappa > 0$ by construction.

What does κ estimate?

- The baseline is misspecified, so we have the **pseudo-true** value m_t^0, q_t^0, e_t^0 and the spread $S_t^0 = m_t^0 - q_t^0$

Lemma 1: the ideal multiplier on date t

$$\kappa_t^\star := \arg\min_{\kappa \in \mathcal{K}} \mathbb{E}[\rho_{\tau_0}\{Y_{t+1} - (m_t^0 - \kappa S_t^0)\} \mid X_t]$$

- The optimal κ makes the adjusted VaR exact: $m_t^0 - \kappa_t^\star S_t^0 = q_{\tau_0,t}^\star$
- $\kappa_t^\star = \dfrac{m_t^0 - q_{\tau_0,t}^\star}{m_t^0 - q_t^0}$

Inside the date-t target

$$\kappa_t^\star = \underbrace{\frac{m_t^0 - q_{0.5,t}^\star}{S_t^0}}_{a_t\text{: median error}} + \underbrace{\frac{q_{0.5,t}^\star - q_{\tau_0,t}^\star}{S_t^0}}_{b_t\text{: true median-to-VaR distance}}$$

both pieces are measured in units of the model's own spread $S_t^0 = m_t^0 - q_t^0$

- Model right in median and tail: $a_t = 0$, $b_t = 1 \Rightarrow \kappa_t^\star = 1$, no adjustment
- Correct median ($a_t = 0$): $\kappa_t^\star = b_t =$ true tail distance / model tail distance
  - $b_t > 1$: the true tail reaches farther $\Rightarrow$ model tail too thin $\Rightarrow$ stretch ($\kappa > 1$)
  - $b_t < 1$: the true tail reaches less far $\Rightarrow$ model tail too thick $\Rightarrow$ shrink ($\kappa < 1$)

One κ is enough: right volatility, wrong tail shape

Truth $Y_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim G$. The baseline has the same μ_t, σ_t but innovation law $\widehat{G}$ (e.g. Gaussian GARCH, fat-tailed truth).

$$q^{\star}_{\tau,t} = \mu_t + \sigma_t G^{-1}(\tau); \qquad \text{baseline: } m^0_t = \mu_t + \sigma_t \widehat{G}^{-1}(0.5), \quad q^0_t = \mu_t + \sigma_t \widehat{G}^{-1}(\tau_0).$$

Plug into a_t and b_t:

$$a_t = \frac{\sigma_t\{\widehat{G}^{-1}(0.5) - G^{-1}(0.5)\}}{\sigma_t\{\widehat{G}^{-1}(0.5) - \widehat{G}^{-1}(\tau_0)\}}, \qquad b_t = \frac{\sigma_t\{G^{-1}(0.5) - G^{-1}(\tau_0)\}}{\sigma_t\{\widehat{G}^{-1}(0.5) - \widehat{G}^{-1}(\tau_0)\}}.$$

- σ_t cancels: $\kappa^{\star}_t = a_t + b_t$ is a constant
- This is the case the paper targets: the overall distribution is estimated well, the tail less so. Then one κ is enough

▸ Details

Estimation

Population target

$$\kappa_0 = \arg\min_{\kappa \in \mathcal{K}} \; \mathbb{E}\, \rho_{\tau_0}\{Y_{t+1} - (m_t^0 - \kappa S_t^0)\}, \qquad \rho_\tau(u) = u\{\tau - \mathbf{1}(u < 0)\}$$

the constant κ with the smallest expected check loss; $\kappa_0 = \kappa^\star$ when one κ is enough

Rolling-window analogue at origin T, window $\mathcal{W}_{T,n} = \{T-n, \ldots, T-1\}$

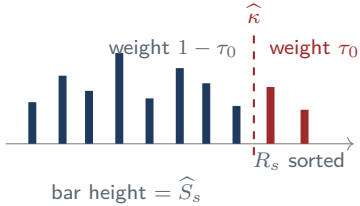
$$\widehat{\kappa}_{T,n} = \arg\min_{\kappa \in \mathcal{K}} \sum_{s \in \mathcal{W}_{T,n}} \rho_{\tau_0}(Y_{s+1} - [\widehat{m}_s - \kappa \widehat{S}_s]), \qquad \widehat{S}_s = \widehat{m}_s - \widehat{q}_s > 0$$

$\widehat{m}_s, \widehat{q}_s$: the current fit $\widehat{\theta}_{T,n}$ evaluated at past states; data up to T only

From an optimization to a sort

With $R_s := \dfrac{\widehat{m}_s - Y_{s+1}}{\widehat{S}_s}$ (depth below the median, in spread units), the same minimizer solves

$$\widehat{\kappa}_{T,n} = \arg\min_{\kappa \in \mathcal{K}} \sum_{s \in \mathcal{W}_{T,n}} \widehat{S}_s \, \rho_{\tau_0}(\kappa - R_s)$$



Lemma 2

$\widehat{\kappa}_{T,n}$ is the weighted $(1 - \tau_0)$-quantile of $\{R_s\}$ with weights $\{\widehat{S}_s\}$.

- An optimization becomes a sort: order the R_s, cumulate the weights to $1 - \tau_0$, read off $\widehat{\kappa}$ — exact, no iterations
- Exceedance $\iff$ $R_s > \kappa$: $\widehat{\kappa}$ sets the spread-weighted exceedance rate to τ_0

Assumption 1 (two-step plug-in regularity)

(i) The data process is stationary and α-mixing, with moments sufficient for a uniform law of large numbers for the fixed-θ_0 criterion.

(ii) $\widehat{\theta}_n \to_P \theta_0$. The fitted median and target-quantile paths are locally Lipschitz in θ, with a stationary integrable envelope, so replacing θ_0 by $\widehat{\theta}_n$ changes the criterion uniformly by $o_P(1)$.

(iii) The population criterion has a unique interior minimizer $\kappa_0 \in \mathrm{int}(\mathcal{K})$, and the spread-weighted distribution of the ratios has a positive density at κ_0.

High-level conditions; the supplement gives sufficient regularity for the GARCH and quantile-regression paths.

Consistency and asymptotic normality

Proposition 2 (Consistency)

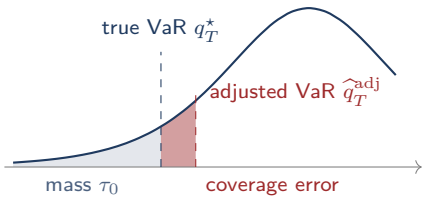
Under Assumption 1, $\widehat{\kappa}_n \to_P \kappa_0$ as $n \to \infty$.

Asymptotic distribution (under high-level two-step conditions)

$$\sqrt{n}\,(\widehat{\kappa}_n - \kappa_0) \;\Rightarrow\; N\big(0,\; J^{-2}\Omega\big)$$

J: limiting derivative of the second-stage score; Ω: long-run variance of the complete two-step influence function, including the first-stage (baseline) contribution.

Coverage



- Coverage error = probability between the adjusted and the true VaR $\approx$ density $\times$ distance
- True VaR today (Lemma 1): $q_T^\star = m_T^0 - \kappa_T^\star S_T^0$

$$|\widehat{q}_T^{\,\mathrm{adj}} - q_T^\star| \;\le\; \Delta_{T,n} = \underbrace{r_{T,n}}_{\text{today's fit error}} + \underbrace{S_T^0\,|\widehat{\kappa} - \kappa_{T,n}^0|}_{\text{sampling error}} + \underbrace{S_T^0\,|\kappa_{T,n}^0 - \kappa_T^\star|}_{\text{one } \kappa \text{ vs. today's } \kappa_T^\star}$$

$r_{T,n} = |\widehat{m}_T - m_T^0| + \kappa_U |\widehat{S}_T - S_T^0|$; $\kappa_{T,n}^0$: the window version of κ_0

Theorem 1: coverage error

Theorem 1

If X_T summarizes $\mathcal{I}_T$, Lemma 1 applies at T, and the density of $Y_{T+1} \mid \mathcal{I}_T$ is at most $\bar{f}$ within δ_0 of $q_T^\star$, then with $\Delta_{T,n}$ from the previous slide

$$\left| \Pr\left\{ Y_{T+1} \le \hat{q}_T^{\mathrm{adj}} \,\middle|\, \mathcal{I}_T \right\} - \tau_0 \right| \;\le\; \bar{f}\, \Delta_{T,n}\, \mathbf{1}\{\Delta_{T,n} \le \delta_0\} + \mathbf{1}\{\Delta_{T,n} > \delta_0\}.$$

- Reading: when the distance is small, coverage error $\le$ density bound $\times$ distance (the second term only says the bound is trivial when the distance is large)
- Today's fit error $\to 0$; sampling error $= O_P(n^{-1/2})$
- The gap $|\kappa_{T,n}^0 - \kappa_T^\star|$ is 0 when one κ is enough; otherwise it is a bias that a longer window does not remove

▸ Full statement

Theorem 2: FZ0 risk

Theorem 2

If the truth is the baseline with its lower tail stretched by one constant $\kappa^\star \neq 1$, the same for VaR and ES,

$$q^\star_{\tau_0,t} = m^0_t - \kappa^\star(m^0_t - q^0_t), \qquad e^\star_{\tau_0,t} = m^0_t - \kappa^\star(m^0_t - e^0_t),$$

then (under Assumption 1 and regularity)

$\widehat{\kappa} \to_P \kappa^\star$ and $\Pr\{\text{FZ0 risk of adjusted pair} < \text{FZ0 risk of raw pair}\} \to 1$

- **Why:** $\widehat{\kappa} \to \kappa^\star$ takes the adjusted pair to the true pair; the raw pair ($\kappa = 1$) stays a fixed distance away; FZ0 is strictly consistent, so the closer pair has the lower expected score

FZ0 score ($e < 0$): $s_{\tau_0}(q, e; y) = -\frac{1}{\tau_0 e}\mathbf{1}\{y \le q\}(q - y) + \frac{q}{e} + \log(-e) - 1$. The fitted median need not be the true median; $\kappa^\star$ is $\kappa^\dagger$ in the paper.

▸ Full statement

Monte Carlo: four designs

one κ is enough

one κ is not enough

DGP 1	DGP 2	DGP 3	DGP 4
Theorem 2 holds *exactly*	right volatility, wrong tail shape	state-dependent median error	state-dependent tail shape too
baseline given	baseline estimated	baseline estimated	baseline estimated
tests estimating κ	our target case	one κ is not enough	one κ is not enough

$B = 500$ replications per cell.

DGP 1: Theorem 2's condition holds exactly

Baseline *given*, not estimated: isolates the estimation of κ and its use on a future sample (t_5 innovations, dynamic volatility, $\tau_0 = 0.05$)

	RMSE of $\widehat{\kappa}$			Future sample, $n = 1000$		
$\kappa^\star$	$n = 250$	$n = 1000$	$n = 2000$	Hit rate	FZ0: Raw $\to$ Anch.	Impr. share
0.75	0.080	0.041	0.028	0.050	$-0.388 \to -0.478$	0.992
1.25	0.135	0.067	0.048	0.050	$0.111 \to 0.034$	0.988
1.75	0.197	0.090	0.066	0.050	$0.986 \to 0.368$	1.000

- $\widehat{\kappa}$ recovers $\kappa^\star$ (bias at most 0.008, RMSE falling in n); coverage at 5%; FZ0 falls in 94–100% of replications

▸ Full tables

DGP 2: estimated baselines, wrong tail shape

Truth: stochastic volatility with a fixed skew-t innovation. GARCH is *estimated*: useful volatility dynamics, wrong tail shape. $\tau_0 = 0.05$, $n = 1600$, in-window.

Baseline	κ_0	Mean $\widehat{\kappa}$	Hit: Raw $\to$ Anch.	FZ0: Raw $\to$ Anch.
skew-t QR	1.000	1.000	$0.050 \to 0.050$	$-0.021 \to -0.021$
Student-t GARCH	1.251	1.251	$0.081 \to 0.050$	$0.065 \to -0.017$
Gaussian GARCH	1.125	1.126	$0.064 \to 0.050$	$0.034 \to -0.007$

- $\widehat{\kappa}$ recovers κ_0; GARCH hit rates move to the nominal 0.050 and FZ0 falls
- Skew-t QR already has the right tail shape: $\widehat{\kappa}$ stays at 1.000

▸ Full tables

DGPs 3–4: when one κ is not enough

Both break the location–scale case. DGP 3: the median error moves with the state (an omitted state-dependent mean). DGP 4: the tail shape moves too. Estimated baselines, future sample of 1,000.

DGP	Baseline	Hit: Raw $\to$ Anch.	FZ0 gain
3	Gaussian GARCH	$0.0616 \to 0.0502$	0.0258
3	Student-t GARCH	$0.0729 \to 0.0499$	0.0471
3	skew-t QR	$0.0511 \to 0.0512$	-0.0001
4	Gaussian GARCH	$0.0543 \to 0.0507$	0.0067
4	Student-t GARCH	$0.0648 \to 0.0506$	0.0195
4	skew-t QR	$0.0503 \to 0.0504$	-0.0003

FZ0 gain: mean Raw $-$ Anchored loss; $\tau_0 = 0.05$.

Coverage by state quintile (DGP 3, Gaussian GARCH)

Quintile	Pooled $\widehat{\kappa}$	Raw
1 (low)	0.084	0.105
3	0.046	0.057
5 (high)	0.027	0.033
All	0.050	0.062

- The *average* correction survives for GARCH
- Coverage stays monotone in the state: one κ shifts the profile but cannot flatten it

Application: daily S&P 500

- Jan 1995 – Dec 2025; **6,800** rolling one-step origins
- 1,000-day window for both the baseline and κ; refit at every origin
- $\tau_0 \in \{1\%, 5\%\}$; baselines: Gaussian GARCH, Student-t GARCH, skew-t QR density
- Evaluation: FZ0 loss with DM tests; Kupiec and Christoffersen coverage tests

Median check

Share of returns below the fitted median: 0.504 / 0.509 / 0.502
(skew-t QR / Student-t / Gaussian); binomial $p = 0.528$ / 0.152 / 0.771

▸ Stability check

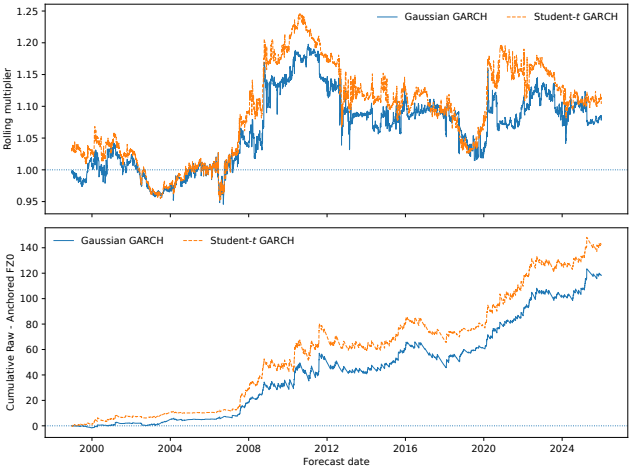
Results: FZ0 loss falls for the GARCH baselines

		FZ0 loss			Hit rate		CC p	
τ_0	Baseline	Raw	Anch.	DM p	Raw	Anch.	Raw	Anch.
1%	skew-t QR	**2.015**	2.036	0.615	0.013	0.015	<0.001	<0.001
	Student-t GARCH	1.259	**1.237**	0.026	0.015	0.013	<0.001	0.022
	Gaussian GARCH	1.389	**1.245**	<0.001	0.022	0.012	<0.001	0.048
5%	skew-t QR	1.198	**1.194**	0.038	0.055	0.054	<0.001	<0.001
	Student-t GARCH	0.849	**0.828**	0.003	0.065	0.054	<0.001	0.392
	Gaussian GARCH	0.858	**0.840**	0.002	0.060	0.053	0.001	0.559

- Largest gain: Gaussian GARCH at 1%, -10.4%; hit rate 2.2% $\rightarrow$ 1.2%
- 5% GARCH: coverage tests no longer reject; the DQ check still does
- Skew-t QR: little or no gain

▸ Full table ▸ CC & DQ

The gain accumulates over time ($\tau_0 = 0.05$)



- **Top:** rolling $\widehat{\kappa}$; mean 1.065 (Gaussian), 1.092 (Student-t)
- **Bottom:** cumulative FZ0 gain of Anchored over Raw
- Gain in all four subperiods, largest in the GFC and COVID/post-COVID periods

▸ Subperiods

Robustness

- **Longer HAC bandwidths, fixed-b:** all four GARCH gains stay significant
- **Family-wise (Romano–Wolf):** Gaussian GARCH 1% survives every block length; the 5% gains survive only the shortest block
- **NASDAQ:** same 5% GARCH gains (both $p < 0.001$); **FTSE 100:** weaker
- **ES check**, Student-t GARCH 5%: p from 0.025 to 0.700
- **Window** of 500 or 2,000 days: 5% results change little
- **Alternatives:** MZ with $\beta \geq 0$ never better; anchoring at 0 instead of the median indistinguishable

▸ DM & RW ▸ Markets ▸ Window ▸ ES check ▸ Comparison

Conclusion

- **Method:** keep the median, rescale both tail distances by one κ; an ordered pair and a closed-form $\widehat{\kappa}$
- **Theory:** a coverage bound and an FZ0-risk gain at the forecast origin
- **Evidence:** helps when scale dynamics are good but the tail is too thin (GARCH)
- **Use it when** the median is well centered and one κ serves VaR and ES — and check both

Backup slides

Related literature

- **VaR-side adjustment:** quantile MZ (Mincer–Zarnowitz 1969; Giacomini–Komunjer 2005; Gaglianone et al. 2011; Fosten et al. 2024); no ordered joint VaR–ES update by itself.
- **Model-risk buffers** from backtesting outcomes (Boucher et al. 2014).
- **Richer transformations or combinations of VaR–ES pairs:** Taylor (2020); Storti–Wang (2023); joint VaR–ES regression (Dimitriadis–Bayer 2019).
- **More information:** PIT diagnostics need the full predictive distribution (Diebold et al. 1998; Berkowitz 2001); FHS needs standardized residuals and conditional scales (Barone-Adesi et al. 1999; McNeil–Frey 2000).
- **New tail equations** instead of adjusting a reported pair: Taylor (2019); Gatta et al. (2024); Gerlach et al. (2025); D'Innocenzo et al. (2026); Liu–Luger (2026).
- **Closest:** Zhang, Gong & Xue (2023): two positive multipliers on the reported VaR and ES levels, estimated by rolling FZ0 minimization under an ordering constraint; the two coordinates move independently.

Here: one multiplier on distances from the fitted median $\Rightarrow$ ordering without an additional constraint, a closed-form estimator, and direct coverage and FZ0-risk results at the forecast origin.

Theory in full: Assumption 1 and Theorem 1

Assumption 1 (Two-step plug-in regularity; high level)

(i) Stationary α-mixing data; a ULLN for the fixed-θ_0 criterion. (ii) $\widehat{\theta}_n \to_P \theta_0$; fitted median and target-quantile paths locally Lipschitz in θ with an integrable envelope. (iii) A unique interior κ_0, at which the spread-weighted distribution of the ratios has positive density.

Theorem 1 (Coverage error at the forecast origin)

Let $F^\star(\cdot \mid \mathcal{I}_T) = F^\star(\cdot \mid X_T)$. If $\kappa_{T,\mathrm{int}} \in \mathcal{K}$ (conditions of Lemma 1), $S_T^0 = O_P(1)$, and the conditional density is bounded by $\bar{f}$ on $[q^\star_{\tau_0,T} - \delta_0,\, q^\star_{\tau_0,T} + \delta_0]$, then

$$\left| \Pr\left\{ Y_{T+1} \le \widehat{m}_{T|T,n} - \widehat{\kappa}_{T,n} \widehat{S}_{T|T,n} \,\middle|\, \mathcal{I}_T \right\} - \tau_0 \right| \le \bar{f}\, \Delta_{T,n} \mathbf{1}\{\Delta_{T,n} \le \delta_0\} + \mathbf{1}\{\Delta_{T,n} > \delta_0\},$$

$$\Delta_{T,n} := r_{T,n} + S_T^0\big(|\widehat{\kappa}_{T,n} - \kappa^0_{T,n}| + |\kappa^0_{T,n} - \kappa^\star_T|\big), \quad r_{T,n} := |\widehat{m}_{T|T,n} - m^0_T| + \kappa_U |\widehat{S}_{T|T,n} - S^0_T|.$$

Proposition 2: $\widehat{\kappa}_n \to_P \kappa_0$. Under the supplement's high-level two-step conditions, $\sqrt{n}(\widehat{\kappa}_n - \kappa_0) \Rightarrow N(0, J^{-2}\Omega)$, Ω the long-run variance of the complete two-step influence function.

Back

Theory in full: FZ0 risk

$\mathcal{R}_T(q, e)$: conditional FZ0 risk; true pair $(q_T^\star, e_T^\star)$; raw $(\widehat{q}_T, \widehat{e}_T)$; adjusted $(\widehat{q}_T^{\mathrm{adj}}, \widehat{e}_T^{\mathrm{adj}})$ at $\widehat{\kappa}_{T,n}$.

Proposition 3 (Local FZ0 risk comparison)

Both pairs lie, w.p.a. 1, in a common convex neighborhood within the score domain with Hessian eigenvalues in $[\lambda_{\min}, \lambda_{\max}]$. Let $D_T := \frac{\lambda_{\min}}{4} \|(\widehat{q}_T - q_T^\star,\ \widehat{e}_T - e_T^\star)\|^2$ and $\mathcal{E}_T^{\mathrm{adj}} := (\widehat{q}_T^{\mathrm{adj}} - q_T^\star)^2 + (\widehat{e}_T^{\mathrm{adj}} - e_T^\star)^2$. If $\mathcal{E}_T^{\mathrm{adj}} = o_P(D_T)$, then $\mathcal{R}_T(\widehat{q}_T, \widehat{e}_T) - \mathcal{R}_T(\widehat{q}_T^{\mathrm{adj}}, \widehat{e}_T^{\mathrm{adj}}) \geq D_T + o_P(D_T)$.

Theorem 2 (Same multiplier adjusts VaR and ES)

Assumption 1 and the score-domain and moment conditions of Proposition 3. Suppose $\kappa^\dagger \in \operatorname{int} \mathcal{K}$, $|\kappa^\dagger - 1| \geq c > 0$, and a.s. $(q_t^\star, e_t^\star) = \big(m_t^0 - \kappa^\dagger(m_t^0 - q_t^0),\ m_t^0 - \kappa^\dagger(m_t^0 - e_t^0)\big)$, with $m_t^0 - q_t^0$ and $m_t^0 - e_t^0$ bounded and bounded away from zero. If the current fitted median, VaR and ES are consistent for their pseudo-true values, then $\kappa_0 = \kappa^\dagger$, $\widehat{\kappa}_{T,n} \to_P \kappa^\dagger$, $\mathcal{E}_T^{\mathrm{adj}} = o_P(1)$, D_T is bounded away from zero in probability, and $\Pr\{\mathcal{R}_T(\widehat{q}_T^{\mathrm{adj}}, \widehat{e}_T^{\mathrm{adj}}) < \mathcal{R}_T(\widehat{q}_T, \widehat{e}_T)\} \to 1$.

Back

One multiplier: details

- In general $\kappa_t^\star = \Pi_{\mathcal{K}}(a_t + b_t)$; interior case $\kappa_t^\star = \kappa_{t,\mathrm{int}} = a_t + b_t$.
- Window target: $\kappa_{T,n}^0 = \arg\min_{\kappa \in \mathcal{K}} \mathbb{E}\, \frac{1}{n} \sum_{s \in \mathcal{W}_{T,n}} \rho_{\tau_0}\{Y_{s+1} - (m_s^0 - \kappa S_s^0)\}$ ($= \kappa_0$ under stationarity). If every $\kappa_s^\star$ in the window lies within δ of a common $\kappa^\star$, then $|\kappa_{T,n}^0 - \kappa^\star| \le \delta$ (supplement).
- Location–scale: truth $Y_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim G$; the baseline has the same (μ_t, σ_t) and innovation law $\widehat{G}$. Then a_t is state-free and

$$b_t = \frac{G^{-1}(\bar{\tau}) - G^{-1}(\tau_0)}{\widehat{G}^{-1}(\bar{\tau}) - \widehat{G}^{-1}(\tau_0)}.$$

- ES error left once the VaR is exactly corrected: $B_{e,t} = (m_t^0 - e_{\tau_0,t}^\star) - \kappa_{t,\mathrm{int}}(m_t^0 - e_t^0)$; zero only under additional compatibility.
- Population FOC of the estimator: $\mathbb{E}\big[S_t^0\{\tau_0 - \mathbf{1}(Y_{t+1} < m_t^0 - \kappa_0 S_t^0)\}\big] = 0$, a spread-weighted hit moment.
- Application: $\mathcal{K} = [0.25, 4]$; the bounds never bind in the 122,556 adjusted equity forecast rows.

Back

Monte Carlo: DGP 1 and DGP 2

DGP 1: one $\kappa^\dagger$ maps the baseline to the true VaR *and* ES (Theorem 2). t_5 innovations, dynamic volatility, $\tau_0 = 0.05$, $n = 1000$, independent future sample of 1,000; $B = 500$.

$\kappa^\dagger$	Bias of $\widehat{\kappa}$	RMSE of $\widehat{\kappa}$	Future hit, Anch.	FZ0 loss: Raw $\to$ Anch.	Improvement share
0.75	0.002	0.041	0.050	$-0.388 \to -0.478$	0.992
1.25	0.005	0.067	0.050	$0.111 \to 0.034$	0.988
1.75	-0.002	0.090	0.050	$0.986 \to 0.368$	1.000

DGP 2: $Y_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}$ with state-driven volatility and a fixed skew-t innovation: a time-invariant tail-shape error for GARCH. In-window, $\tau_0 = 0.05$, $n = 1600$.

Baseline	κ_0	Mean $\widehat{\kappa}$	Hit: Raw $\to$ Anch.	FZ0 loss: Raw $\to$ Anch.
skew-t QR	1.000	1.000	$0.050 \to 0.050$	$-0.021 \to -0.021$
Student-t GARCH	1.251	1.251	$0.081 \to 0.050$	$0.065 \to -0.017$
Gaussian GARCH	1.125	1.126	$0.064 \to 0.050$	$0.034 \to -0.007$

DGP 1 over $n = 250$–2000: |bias| ≤ 0.008, RMSE falls with n, improvement share 0.94–1.00.

◂ Back

Monte Carlo: misspecification designs at both tails

DGP	Baseline	Hit: Raw → Anch.	FZ0 loss: Raw − Anch. (MC s.e.)	DM reject
Panel A: 5% tail				
3	Gaussian GARCH	$0.0616 \to 0.0502$	0.0258 (0.0010)	0.376
3	Student-t GARCH	$0.0729 \to 0.0499$	0.0471 (0.0017)	0.400
3	skew-t QR	$0.0511 \to 0.0512$	-0.0001 (0.0000)	0.164
4	Gaussian GARCH	$0.0543 \to 0.0507$	0.0067 (0.0006)	0.284
4	Student-t GARCH	$0.0648 \to 0.0506$	0.0195 (0.0010)	0.206
4	skew-t QR	$0.0503 \to 0.0504$	-0.0003 (0.0002)	0.148
Panel B: 1% tail				
3	Gaussian GARCH	$0.0253 \to 0.0104$	0.2402 (0.0062)	0.530
3	Student-t GARCH	$0.0209 \to 0.0105$	0.1055 (0.0041)	0.270
3	skew-t QR	$0.0114 \to 0.0120$	-0.0012 (0.0019)	0.150
4	Gaussian GARCH	$0.0212 \to 0.0104$	0.1800 (0.0057)	0.380
4	Student-t GARCH	$0.0158 \to 0.0102$	0.0343 (0.0028)	0.144
4	skew-t QR	$0.0115 \to 0.0115$	-0.0015 (0.0018)	0.130

DGP 3: median misspecification; DGP 4: median and innovation misspecification. $n_{\text{est}} = 1600$, future length 1,000, $B = 500$. "DM reject": share of replication-level paired DM tests rejecting at 5%.

S&P 500: full results (Table 6)

		FZ0 loss		Anch. vs. Raw	Hit rate		Kupiec p		CC p	
τ_0	Baseline	Raw	Anch.	DM p	Raw	Anch.	Raw	Anch.	Raw	Anch.
1%	skew-t QR	2.015	2.036	0.615	0.013	0.015	0.011	<0.001	<0.001	<0.001
	Student-t GARCH	1.259	**1.237**	0.026	0.015	0.013	<0.001	0.020	<0.001	0.022
	Gaussian GARCH	1.389	**1.245**	<0.001	0.022	0.012	<0.001	0.060	<0.001	0.048
5%	skew-t QR	1.198	1.194	0.038	0.055	0.054	0.055	0.124	<0.001	<0.001
	Student-t GARCH	0.849	**0.828**	0.003	0.065	0.054	<0.001	0.187	<0.001	0.392
	Gaussian GARCH	0.858	**0.840**	0.002	0.060	0.053	<0.001	0.295	0.001	0.559

Skew-t QR: FZ0 and DM on $N_{\text{score}} = 6{,}797$ (three positive-ES exclusions); hit rates on all 6,800 origins.

Back

S&P 500: conditional coverage and dynamic-quantile checks

Baseline	Tail, method	Hit	Indep. p	CC p	DQ p
Gaussian GARCH	5%, Raw	0.060	0.766	0.001	$<$0.001
	5%, Anchored	0.053	0.802	0.559	0.008
	1%, Raw	0.022	0.390	$<$0.001	$<$0.001
	1%, Anchored	0.012	0.111	0.048	$<$0.001
Student-t GARCH	5%, Raw	0.065	0.782	$<$0.001	$<$0.001
	5%, Anchored	0.054	0.720	0.392	0.004
	1%, Raw	0.015	0.324	$<$0.001	$<$0.001
	1%, Anchored	0.013	0.141	0.022	$<$0.001
skew-t QR	5%, Raw	0.055	$<$0.001	$<$0.001	$<$0.001
	5%, Anchored	0.054	$<$0.001	$<$0.001	$<$0.001
	1%, Raw	0.013	$<$0.001	$<$0.001	$<$0.001
	1%, Anchored	0.015	$<$0.001	$<$0.001	$<$0.001

DQ regression: constant, four lagged centered hits, current VaR; χ^2_6 reference. Coverage tests are specification checks, not the ranking criterion; the forecasts embed estimation error (Escanciano–Olmo 2010).

Back

Multiplicity and bandwidth: all six S&P comparisons

Tail	Baseline	Mean Raw − Anch.	RW 19	RW 100	RW 250	RW 500
5%	Gaussian GARCH	0.0172	0.0281	0.0767	0.0939	0.0837
5%	Student-t GARCH	0.0207	0.0344	0.0862	0.1053	0.0989
5%	skew-t QR	0.0040	0.1362	0.2517	0.2817	0.2823
1%	Gaussian GARCH	0.1447	0.0010	0.0080	0.0113	0.0088
1%	Student-t GARCH	0.0220	0.1202	0.2195	0.2434	0.2350
1%	skew-t QR	-0.0208	0.6672	0.6425	0.6377	0.6751

Tail	Baseline	$M = 10$	$M = 100$	$M = 250$	Fixed-b reject (5%)
5%	skew-t QR	0.038	0.212	0.244	no
5%	Student-t GARCH	0.003	0.001	$<$0.001	yes
5%	Gaussian GARCH	0.002	0.001	0.001	yes
1%	skew-t QR	0.615	0.656	0.651	no
1%	Student-t GARCH	0.026	0.016	0.012	yes
1%	Gaussian GARCH	$<$0.001	$<$0.001	$<$0.001	yes

Top: shared-index Romano–Wolf stepdown p-values with circular blocks on the common 6,797-date calendar ($19 = \lceil 6{,}797^{1/3} \rceil$). Bottom: two-sided DM p with Bartlett bandwidth M; fixed-b critical values at $M = 250$.

Back

S&P 500 subperiods: Raw − Anchored FZ0 loss

Baseline	τ_0	Pre-GFC	GFC	Post-GFC	COVID/post-COVID
Gaussian GARCH	0.01	0.0348 (0.190)	0.3181 (0.002)	0.1453 (0.005)	0.2474 (0.008)
Gaussian GARCH	0.05	0.0057 (0.016)	0.0383 (0.046)	0.0118 (0.254)	0.0378 (0.010)
Student-t GARCH	0.01	0.0091 (0.441)	0.0399 (0.127)	0.0110 (0.463)	0.0549 (0.094)
Student-t GARCH	0.05	0.0098 (0.004)	0.0523 (0.062)	0.0127 (0.306)	0.0420 (0.029)
skew-t QR	0.01	-0.0314 (0.223)	-0.5824 (0.048)	0.0078 (0.470)	0.1268 (0.004)
skew-t QR	0.05	0.0030 (<0.001)	0.0307 (<0.001)	-0.0042 (<0.001)	0.0117 (0.010)

Mean Raw-minus-Anchored FZ0 differential with the unadjusted DM p in parentheses; positive favors Anchored. Pre-GFC 1998-12-18 to 2007-07-31; GFC 2007-08-01 to 2009-06-30; post-GFC/pre-COVID 2009-07-01 to 2020-02-19; COVID/post-COVID 2020-02-20 to 2025-12-31. Skew-t QR reverses at 1% in the GFC and slightly at 5% post-GFC, consistent with its stability-check rejections.

◂ Back

Other markets: NASDAQ Composite and FTSE 100

		NASDAQ (6,799 origins)					FTSE 100 (6,827 origins)				
		FZ0		Hit		DM	FZ0		Hit		DM
Tail	Baseline	Raw	Anch.	Raw	Anch.	p	Raw	Anch.	Raw	Anch.	p
1%	skew-t QR	3.475	3.495	0.011	0.016	0.500	2.094	1.988	0.016	0.015	0.200
	Student-t GARCH	1.476	1.436	0.016	0.013	0.003	1.196	1.168	0.015	0.011	0.074
	Gaussian GARCH	1.585	1.438	0.021	0.013	$<$0.001	1.281	1.176	0.020	0.011	0.001
5%	skew-t QR	1.673	1.672	0.057	0.057	0.607	1.173	1.173	0.052	0.051	0.816
	Student-t GARCH	1.099	1.073	0.069	0.054	$<$0.001	0.796	0.788	0.060	0.051	0.126
	Gaussian GARCH	1.102	1.078	0.066	0.055	$<$0.001	0.804	0.797	0.056	0.050	0.050

Same rolling design as the S&P 500 application (1,000-day window, refit at every origin); unadjusted pairwise DM p, Anchored vs. Raw. NASDAQ skew-t QR 1%: Raw coverage is already near nominal and Anchored raises the hit rate — the spread-weighted moment is not ordinary coverage.

Back

Sensitivity to the estimation window (S&P 500, GARCH)

			Rolling $\widehat{\kappa}$			Hit rate		Mean FZ0		
Baseline	τ_0	Window	mean	min	max	Raw	Anch.	Raw	Anch.	DM p
Gaussian GARCH	0.01	500	1.171	0.913	1.687	0.023	0.014	1.3961	1.2116	$<$0.001
		1000	1.157	0.946	1.459	0.023	0.013	1.3723	1.2083	$<$0.001
		2000	1.156	0.999	1.332	0.022	0.013	1.3600	1.2250	$<$0.001
Gaussian GARCH	0.05	500	1.076	0.893	1.264	0.062	0.053	0.8235	0.7966	$<$0.001
		1000	1.075	0.946	1.198	0.060	0.051	0.8183	0.7980	0.002
		2000	1.065	0.972	1.154	0.059	0.050	0.8136	0.7950	0.002
Student-t GARCH	0.01	500	1.063	0.837	1.382	0.017	0.014	1.2246	1.1974	0.083
		1000	1.048	0.910	1.254	0.016	0.014	1.2298	1.2008	0.009
		2000	1.040	0.947	1.141	0.015	0.013	1.2351	1.2182	0.116
Student-t GARCH	0.05	500	1.102	0.904	1.352	0.067	0.054	0.8138	0.7813	$<$0.001
		1000	1.104	0.951	1.246	0.064	0.052	0.8102	0.7868	0.004
		2000	1.091	0.993	1.192	0.063	0.051	0.8081	0.7882	0.006

Baseline and multiplier windows change together; common evaluation span from the 2,000-day burn-in onward ($N_{\text{eval}} = 5{,}800$). At 5%, the mean multipliers differ by at most 0.013, adjusted hit rates stay within 0.050–0.054, and every DM comparison remains significant.

Back

ES side: exceedance-residual check (S&P 500)

		$\tau_0 = 0.01$			$\tau_0 = 0.05$		
Baseline	Method	n_{viol}	Mean z	p_{boot}	n_{viol}	Mean z	p_{boot}
skew-t QR	Raw	88	-0.306	0.024	373	-0.120	0.014
	Anchored	97	-0.297	0.014	366	-0.119	0.015
Student-t GARCH	Raw	105	-0.023	0.398	441	-0.039	0.025
	Anchored	88	-0.026	0.353	364	-0.007	**0.700**
Gaussian GARCH	Raw	151	-0.143	$<$0.001	409	-0.140	$<$0.001
	Anchored	84	-0.130	$<$0.001	359	-0.112	$<$0.001

On violation days, $z_t = (y_{t+1} - e_t)/|e_t|$, centered at zero under adequate ES forecasts (McNeil–Frey 2000). Two-sided centered iid bootstrap with 9,999 resamples. Supporting evidence only: violation residuals can be serially dependent. Formal ES backtests: Du–Escanciano (2017), Kratz et al. (2018), Bayer–Dimitriadis (2022).

◂ Back

Alternative adjustments (Table 7, Panel A)

		Anchored	Equal-weight	FZ0-targeted	MZ ($\beta_{\mathrm{MZ}} \geq 0$)	Proportional
$\tau_0 = 0.05$	skew-t QR	1.194	1.196 (0.011)	1.252 (0.315)	1.202 (0.510)	1.194 (0.643)
	Student-t GARCH	0.828	0.830 (0.159)	0.828 (0.602)	0.839 (0.012)	0.829 (0.234)
	Gaussian GARCH	0.840	0.841 (0.428)	0.834 (0.056)	0.847 (0.147)	0.841 (0.095)
$\tau_0 = 0.01$	skew-t QR	2.036	2.000 (0.109)	2.112 (0.280)	1.803 (0.107)	2.015 (0.063)
	Student-t GARCH	1.237	1.234 (0.559)	1.237 (0.946)	1.230 (0.657)	1.239 (0.332)
	Gaussian GARCH	1.245	1.243 (0.705)	1.241 (0.649)	1.242 (0.848)	1.245 (0.358)

FZ0 loss; DM p against Anchored in parentheses. All alternatives use the same current-fit retrospective values.

Back

Affine MZ: score domain and losses (S&P 500)

Share of origins at which the affine MZ extension leaves the FZ0 domain

	$\tau_0 = 0.01$			$\tau_0 = 0.05$		
Baseline	$\Pr(e > q)$	$\Pr(e \geq 0)$	Joint invalid	$\Pr(e > q)$	$\Pr(e \geq 0)$	Joint invalid
skew-t QR	0.469	0.190	0.469	0.128	0.038	0.128
Student-t GARCH	0.000	0.000	0.000	0.000	0.000	0.000
Gaussian GARCH	0.000	0.000	0.000	0.000	0.000	0.000

FZ0 loss: Raw, affine MZ, Anchored

	$\tau_0 = 0.01$					$\tau_0 = 0.05$				
Baseline	Raw	+MZ	Anch.	MZ valid	DM p	Raw	+MZ	Anch.	MZ valid	DM p
skew-t QR	2.015	1.842	2.036	0.531	0.142	1.198	1.209	1.194	0.872	0.684
Student-t GARCH	1.259	1.230	1.237	1.000	0.655	0.849	0.840	0.828	1.000	0.012
Gaussian GARCH	1.389	1.242	1.245	1.000	0.829	0.858	0.847	0.840	1.000	0.144

$e^{\mathrm{MZ}} > q^{\mathrm{MZ}}$ is equivalent to a negative fitted MZ slope. The +MZ mean is on its own score-valid subset; DM p: Anchored vs. MZ on common valid dates. When the slope is positive and the pair stays in the domain, MZ can be competitive.

Back

How does it compare? (alternatives)

Alternative	Difference from Anchored	Result vs. Anchored (S&P 500)
Quantile MZ, $\beta \geq 0$	free intercept and slope	never better; worse for Student-t GARCH (5%)
Proportional scaling	anchor at 0, not the median	statistically indistinguishable in every cell
FHS	needs residuals and conditional scales	5%: $\approx$ Anchored (Student-t), lower loss (Gaussian)

- The gain comes from one common multiplier on the tail spread
- The median anchor adds translation equivariance and a closed form; FHS can do better, but only with residuals

Back

Information: historical vintages and FHS (numbers)

	Less: archived vintages only origins 1,001–6,800			**More:** residuals and scales (FHS) common dates within each cell		
Cell	Raw	Anchored	Hist. vintage	Raw	Anchored	FHS
Student-t GARCH, 5%	0.810	0.787	0.800 ($<$0.001)	0.8493	0.8283	0.8286
Gaussian GARCH, 5%	0.818	0.798	0.810 (0.010)	0.8578	0.8404	0.8326
Student-t GARCH, 1%	1.230	1.201	1.220 (0.150)	1.2593	1.2374	1.2386
Gaussian GARCH, 1%	1.372	1.208	1.217 (0.436)	1.3892	1.2447	1.2364

FZ0 loss; parentheses: DM p of the historical-vintage variant against Anchored. The two panels use different evaluation samples.

Back

Specification checks: median anchor and multiplier stability

Nyblom-type check on partial sums of the check-loss score residuals at the full-span $\widehat{\kappa}_{\text{in}}$, with Bartlett HAC $\widehat{\sigma}^2$: $\text{Ny}_n = \frac{1}{n^2\widehat{\sigma}^2} \sum_{s=1}^{n} \left(\sum_{t=1}^{s} \psi_t(\widehat{\kappa}_{\text{in}}) \right)^2$. Rejection is evidence against a constant κ; non-rejection is *not* evidence of constancy.

S&P 500

	Median anchor		Nyblom, $\tau_0 = 0.01$		Nyblom, $\tau_0 = 0.05$	
Baseline	Share below	p	Stat	Sim. p	Stat	Sim. p
skew-t QR	0.504	0.528	0.327	0.078	0.537	0.020
Student-t GARCH	0.509	0.152	0.212	0.194	0.149	0.312
Gaussian GARCH	0.502	0.771	0.157	0.296	0.145	0.316

Finite-sample rejection rates

Scenario	$n = 300$	600	1200
Constant $\kappa = 1$ (size)	0.028	0.038	0.026
Drift $0.8 \to 1.2$	0.076	0.104	0.194
Break $1.0 \to 1.5$	0.042	0.020	0.026
Break $1.0 \to 2.0$	0.008	0.062	0.538

Size is conservative; power is moderate against smooth drift and limited against moderate breaks. The check is read together with hit rates, FZ0, median and MZ checks.

Back

Why the skew-t QR 1% cell fails: tail extrapolation

Baseline	Method	N_{hit}	N_{score}	Mean FZ0	Hit rate	Mean $\widehat{\kappa}$
Direct QR01	Raw	6800	6800	1.7017	0.0138	–
Direct QR01	Anchored	6800	6800	1.7018	0.0138	1.000
Existing skew-t QR	Raw	6800	6797	2.0149	0.0132	–
Existing skew-t QR	Anchored	6800	6797	2.0356	0.0146	0.969

Common-date loss differential (left minus right)	N_{score}	$\bar{d}$
Direct QR01 Raw − Direct QR01 Anchored	6800	-0.0001
Direct QR01 Raw − Existing skew-t QR Raw	6797	-0.3582
Direct QR01 Anchored − Existing skew-t QR Anchored	6797	-0.3789

Direct rolling QR at $\tau = 0.01$ and at the median (1,000-day window); ES from an empirical tail-gap plug-in, so this is not a fully specified density forecast. Direct QR01 has much lower FZ0 than the skew-t 1% path; adjusting it is essentially neutral (partly mechanical when the median and tail QRs share the sample and regressors).