

7543001-fit-is-not-enough-an-empirical-study-of-essvi-calibrations-through-the-local-vol

SUMMARY: The article examines the construction and calibration of an equity volatility surface in the extended SSVI (eSSVI) parameterization, focusing on how skew can vary across maturities. The study compares eSSVI, SSVI, and raw SVI methods, using four weeks of daily and intraday option chains on four equity-index ETFs and four single names. The research evaluates the fit, static arbitrage, parameter stability, and the smoothness of the Dupire local volatility induced by each surface. The study also prices six-month vanilla options and knock-out barrier options under the induced local volatility. The results show that fit alone is not sufficient to judge volatility-surface calibration, as the way a surface is constructed can significantly affect local-volatility dynamics and path-dependent prices.

METHOD: The study uses four weeks of daily and intraday option chains on four equity-index ETFs and four single names. The eSSVI, SSVI, and raw SVI methods are calibrated sequentially under penalties, jointly with γ fixed at 1/2 or estimated. The study scores each surface on in-sample and out-of-sample fit, static arbitrage, parameter stability, and the smoothness of the Dupire local volatility it induces. The induced local volatility is scored on roughness and containment, and the price of knock-out barrier options is scored on shared Monte Carlo paths.

KEY FINDINGS:

- The eSSVI calibrations disagree on a six-month option by up to 3.5% of its value at the money on the index chains and by up to 7.5% half a standard deviation out of the money.
- The SSVI surface induces the smoothest local volatility.
- The guaranteed eSSVI constructions are clean at the fitted pillars, although interpolation can still cross between them.
- A knock-out barrier option amplifies the disagreement between eSSVI calibrations to 16 to 68% of the price where the surface is inadmissible between pillars.

IMPLICATION FOR TRADING: The study highlights that volatility-surface calibration cannot be judged by fit alone, as the way a surface is constructed can significantly affect local-volatility dynamics and path-dependent prices. Practitioners should consider the implications of different calibration methods and ensure that the induced local volatility is smooth and admissible across all maturities. This can have practical implications for pricing exotic options and structuring financial products.

Fit Is Not Enough: An Empirical Study of eSSVI Calibrations Through the Local Volatility They Induce

Siddharth Singh*

September 2026

Abstract

We study the construction and calibration of an equity volatility surface in the extended SSVI (eSSVI) parameterization. The eSSVI provides a flexible, arbitrage-aware framework in which skew can vary across maturities. We calibrate it sequentially under penalties, by a greedy forward pass and jointly, with the power-law exponent γ fixed at $\frac{1}{2}$ or estimated. We benchmark these structured approaches against SSVI and raw SVI. We use four weeks of daily and intraday option chains on four equity-index ETFs and four single names, with two thinner single names as an extension. We score each surface on in-sample and out-of-sample fit, static arbitrage, parameter stability and the smoothness of the Dupire local volatility it induces, and then price six-month vanillas in closed form and knock-out barrier options under that local volatility.

Raw SVI fits best in sample but degrades most out of sample and is unstable on single-name chains; SSVI induces the smoothest local volatility; the guaranteed eSSVI constructions are clean at the fitted pillars, although interpolation can still cross between them. In price, the four eSSVI calibrations disagree on a six-month option by up to 3.5% of its value at the money on the index chains and by up to 7.5% half a standard deviation out of the money, against about 2% on single names. A knock-out adds little on a well-behaved surface, but where the surface is inadmissible between pillars it amplifies the disagreement to 16 to 68% of the price. These results show that volatility-surface calibration cannot be judged by fit alone; the way an admissible surface is constructed can materially affect the local-volatility dynamics and path-dependent prices derived from it.

Keywords: implied volatility surface; eSSVI; SSVI; no-arbitrage calibration; Dupire local volatility; barrier options; out-of-sample validation; equity options.

1 Introduction

The implied volatility surface is the market’s coordinate system. Quoting, hedging, structuring and every smile-consistent model downstream start from it, from Dupire local volatility to stochastic-local-volatility leverage functions to variance-swap replication. The market provides something coarser, a discrete and noisy set of bid and ask quotes on a strike–expiry grid that

*Manager, Quantitative Advisory Services, EY; siddharth_qs@outlook.com. The views expressed are those of the author alone and do not represent those of EY or any of its member firms. No client data or proprietary firm resources were used.

is occasionally arbitrageable. Turning those quotes into a surface means choosing a *smoothing parameterization*, and usually that choice rests on the in-sample fit error.

In-sample fit is not a sufficient criterion. Most often the surface is not used where it was fitted but downstream, in the local-volatility model that prices the exotic book. So structure imposed across expiries matters most at the level of the local volatility, which differentiates the surface once in time t and twice in strike k before it reaches a price. Two surfaces fitted to the same quotes can induce local volatilities that differ severalfold in roughness, and a barrier option, whose value depends on σ_{loc} along the path, turns that difference into a price. Fit alone is thus a weak guide to what a parameterization does downstream.

We therefore evaluate smile parameterizations where their consequences are realized. We fit six calibrations to identical, cleaned, de-Americanized chains from a retail-accessible feed, and follow each one through the same pipeline:

parameterization $\rightarrow$ Dupire local volatility $\rightarrow$ barrier price.

We score the fitted surface on in-sample fit, on a leave-one-expiry-out test that asks each calibration to predict a smile it never saw, and on admissibility, which we measure on a dense grid rather than assume from the construction. We score the induced local volatility on roughness and containment. And we price the barriers on shared Monte Carlo paths so that only the surface differs, then read how far the calibrations disagree relative to each price, beside a vanilla at the same strike priced in closed form from each surface.

The six arms. We evaluate three parameterizations across six distinct calibration methods, set out in Table 1. Three of them are eSSVI slices (Hendriks and Martini, 2019) that carry the Hendriks–Martini no-arbitrage conditions *by construction*. Two come from the global recursion of Mingone (2022), with the power-law backbone exponent pinned at the theoretically distinguished value $\gamma = \frac{1}{2}$ and fitted freely (G-eSSVI). The third, we introduce greedy forward pass calibration, which keeps each slice inside its predecessor’s admissible region at per-expiry speed with a single free parameter per slice (S-eSSVI, Section 3.4).

Three more stand against them. The Hendriks–Martini calibration as practised fits slices in time order under penalties against calendar crossing and butterfly violation. SSVI (Gatheral and Jacquier, 2014) is the fully global comparison point, calibrated inside the exponent domain $(0, \frac{1}{2}]$ on which its no-arbitrage conditions are sufficient, so its guarantee holds by construction. And raw SVI (Gatheral, 2004) spends five parameters per expiry with no cross-expiry structure at all. SVI is not a candidate for production use. It serves as the *unstructured benchmark*.

We call the three eSSVI arms whose no-arbitrage conditions hold by construction the *guaranteed* arms. SSVI is also calibrated within its sufficient-condition domain and therefore has a construction-level no-arbitrage guarantee. That guarantee covers the *fitted pillar slices*. Whether it survives the interpolation between expiries is a separate question, which we answer by measuring the between-pillar surface on a dense strike–maturity grid (Sections 4 and 6.1.4). Throughout the paper, “guaranteed” refers specifically to the fitted maturity pillars.

Barrier pricing is a downstream sensitivity test, with monitoring, correction and barrier conventions held fixed across calibrations. Alongside the comparison, the paper releases an open implementation of the Mingone (2022) global eSSVI construction and introduces S-eSSVI, a greedy one-parameter-per-slice variant that carries the same pillar guarantee at per-expiry cost (Section 3.4).

The construction literature is mature: Homescu (2011) surveys methodologies, Fengler (2009) smooths with arbitrage-free splines, and Gatheral and Jacquier (2014) introduce SSVI together

	S-eSSVI $\gamma=\frac{1}{2}$	G-eSSVI $\gamma=\frac{1}{2}$	G-eSSVI γ free	eSSVI (seq.)	SSVI	SVI (bench.)
Free parameters	1/expiry + 1 global	$2N+1$ global	$2N+2$ global	2/expiry	3 global	5/expiry
Calibration	greedy	joint	joint	sequential	joint	per expiry
Butterfly at pillars	by constr.	by constr.	by constr.	penalized	by constr.	none
Calendar at pillars	by constr.	by constr.	by constr.	penalized	by constr.	none
Skew term structure	power law	power law	power law	free, twists	one global ρ	free/expiry
Interpolant in family	yes	yes	yes	yes	yes	no
Closed-form derivatives	yes	yes	yes	yes	yes	yes
ATM level	quoted	joint fit	joint fit	quoted	quoted	by-product

Table 1: The six arms. “By constr.” denotes conditions that hold for every admissible parameter vector; “penalized” denotes conditions encouraged by a calibration penalty. N is the number of fitted expiries.

with the butterfly bounds enforced below. The eSSVI line runs from [Hendriks and Martini \(2019\)](#) through the slice-level no-arbitrage domain of [Corbetta et al. \(2019\)](#) and [Martini and Mingone \(2022\)](#) to the global parametrization of [Mingone \(2022\)](#), which this paper implements. [Cont and da Fonseca \(2002\)](#) study the dynamics of fitted surfaces.

2 Setting and no-arbitrage preliminaries

2.1 Notation

For an underlying with forward curve F_t and an option quote with strike K and expiry t , write the Black–Scholes implied volatility $\sigma_{\text{BS}}(K, t)$, the *forward log-moneyness*

$$k = \ln \frac{K}{F_t}, \quad (1)$$

and the *total implied variance*

$$w(k, t) = \sigma_{\text{BS}}^2(F_t e^k, t) t. \quad (2)$$

All three parameterizations live naturally in (k, w) coordinates, so we compare them there. Measuring k against the *forward* fixes the drift under which the local volatility is induced (Section 2.3).

2.2 Static arbitrage in (k, w) coordinates

Following [Roper \(2010\)](#) and [Gatheral and Jacquier \(2014\)](#), a surface $w(k, t)$ that is (i) free of *butterfly* arbitrage in every slice and (ii) free of *calendar* arbitrage across slices, and satisfies mild limit conditions, defines arbitrage-free call prices.

Definition 1 (Butterfly condition). *For a smile $w(\cdot)$ at fixed expiry, define*

$$g(k) = \left(1 - \frac{k w'(k)}{2 w(k)}\right)^2 - \frac{w'(k)^2}{4} \left(\frac{1}{w(k)} + \frac{1}{4}\right) + \frac{w''(k)}{2}. \quad (3)$$

The slice is free of butterfly arbitrage iff $g(k) \geq 0$ for all k (and call prices tend to zero as $k \rightarrow \infty$). The risk-neutral density is recovered as

$$p(k) = \frac{g(k)}{\sqrt{2\pi w(k)}} \exp\left(-\frac{1}{2} d_-(k)^2\right), \quad d_-(k) = -\frac{k}{\sqrt{w(k)}} - \frac{\sqrt{w(k)}}{2}, \quad (4)$$

so $g < 0$ is exactly a negative density (Breedon and Litzenberger, 1978; Gatheral and Jacquier, 2014).

Definition 2 (Calendar condition). *The surface is free of calendar arbitrage iff $w(k, t)$ is non-decreasing in t at every fixed k .*

Two classical constraints recur below. First, Lee (2004): an arbitrage-free smile has asymptotically linear total variance with wing slopes at most 2,

$$\limsup_{|k| \rightarrow \infty} \frac{w(k)}{|k|} \leq 2. \quad (5)$$

Second, Dupire’s local variance in Gatheral’s form: if the surface is smooth,

$$\sigma_{\text{loc}}^2(k, t) = \frac{\partial_t w(k, t)}{g(k, t)}, \quad (6)$$

so butterfly violations ($g \leq 0$) and calendar violations ($\partial_t w < 0$) are exactly the failure modes that make a fitted surface unusable for local-volatility pricing. Equation (6) ties the pipeline together: the admissibility scans look for the points where its numerator or denominator turns negative, and the roughness measure reads what it produces everywhere else.

2.3 Local volatility and the drift it implies

Dupire’s local volatility (Dupire, 1994; Derman and Kani, 1994) is the unique one-factor Markovian diffusion that reprices an entire arbitrage-free vanilla surface. Written in the forward log-moneyness coordinate used above, (6) defines σ_{loc} for the dynamics

$$\frac{dS_t}{S_t} = \frac{d \ln F(t)}{dt} dt + \sigma_{\text{loc}}(S_t, t) dW_t. \quad (7)$$

The drift here comes from the *chain’s own* forward curve, not from the risk-free rate. For a dividend payer, or a name with borrow, the two differ materially.

Two consequences drive the design.

First, (6) is a *second-order* functional of the surface: it divides $\partial_t w$ by g , and g itself contains w'' . A small curvature error, or a kink in the term structure at a pillar expiry, is invisible in a fitted-vol comparison and can dominate the local volatility.

Second, the model carries its own acceptance tests. Because (7) is *constructed* to reproduce the calibrating surface, any repricing error is calibration or numerical error rather than model error, which makes round-trip repricing an unambiguous criterion (Section 6.2.1). The spot must likewise satisfy $\mathbb{E}[S_T] = F_T$ under (7), which fails when σ_{loc} is clamped or badly extrapolated in the wings and probability mass leaks away.

3 Three smile parameterizations

3.1 Raw SVI

The raw SVI parameterization (Gatheral, 2004) of one expiry’s total variance is

$$w_{\text{SVI}}(k) = a + b \left[\rho(k - m) + \sqrt{(k - m)^2 + \sigma^2} \right], \quad (8)$$

with level a , wing slope $b \geq 0$, skew $\rho \in (-1, 1)$, shift m , and ATM curvature $\sigma > 0$; admissibility additionally requires the minimum total variance $a + b\sigma\sqrt{1 - \rho^2} \geq 0$. The wings are asymptotically linear,

$$w_{\text{SVI}}(k) \sim b(1 \pm \rho)|k| \quad (k \rightarrow \pm\infty), \quad (9)$$

so Lee's bound (5) is respected whenever $b(1 + |\rho|) \leq 2$. SVI's decisive practical advantage is that (8) is C^∞ with *closed-form* derivatives. Both $g(k)$ in (3) and the Dupire numerator and denominator in (6) are analytic, which makes the induced local volatility smooth by construction *within* a slice. Butterfly arbitrage, though, is *not* guaranteed: the Vogt parameters in [Gatheral and Jacquier \(2014\)](#) satisfy every constraint above and still give $g < 0$ on an interval, so $g(k)$ has to be checked on every fitted slice.

Five parameters per expiry is the largest budget among the three parameterizations, and on a slice with a dozen quoted strikes it absorbs nearly half the degrees of freedom in the data. The parameters are also weakly identified when the smile is close to symmetric, with (a, b, σ) trading off against one another. That is the documented instability [Zeliade Systems \(2009\)](#) discuss, and one of the motivations for SSVI.

3.2 SSVI

SSVI ([Gatheral and Jacquier, 2014](#)) parameterizes the whole surface from the ATM total-variance term structure $\theta_t = w(0, t)$ and a curvature function φ :

$$w_{\text{SSVI}}(k, t) = \frac{\theta_t}{2} \left[1 + \rho \varphi(\theta_t) k + \sqrt{(\varphi(\theta_t)k + \rho)^2 + 1 - \rho^2} \right], \quad (10)$$

with the power-law curvature used throughout this paper,

$$\varphi(\theta) = \frac{\eta}{\theta^\gamma (1 + \theta)^{1-\gamma}}, \quad \gamma \in (0, 1], \quad \eta > 0. \quad (11)$$

The two shape parameters divide the work. Since $\theta\varphi(\theta) \rightarrow \eta$ as θ grows, η is the level the curvature scale approaches at the long end, and it multiplies the smile of every slice; γ then sets how fast the at-the-money skew decays with maturity, like $t^{-\gamma}$ at the short end. Three global parameters (ρ, η, γ) shape every expiry; the term structure enters only through θ_t , taken from the quoted ATM variances. Its appeal is that Theorems 4.1 and 4.2 of [Gatheral and Jacquier \(2014\)](#) give the two kinds of admissibility separately.

Calendar spread. The surface is free of calendar-spread arbitrage when θ_t is non-decreasing in t and

$$0 \leq \partial_\theta(\theta\varphi(\theta)) \leq \frac{1}{\rho^2} \left(1 + \sqrt{1 - \rho^2} \right) \varphi(\theta), \quad \theta > 0. \quad (12)$$

For the power law (11) one has $\theta\varphi(\theta) = \eta(\theta/(1 + \theta))^{1-\gamma}$, so the two sides of the right-hand inequality stand in the ratio $\partial_\theta(\theta\varphi)/\varphi = (1 - \gamma)/(1 + \theta) < 1$, while the bracketed factor is at least 1 for every ρ , with equality only at $|\rho| = 1$. It follows that calendar admissibility holds for *every* $\gamma \in (0, 1]$, and only the monotonicity of θ_t can fail.

Butterfly. A sufficient condition for freedom from butterfly arbitrage is, at each pillar,

$$\theta\varphi(\theta)(1 + |\rho|) < 4, \quad \theta\varphi(\theta)^2(1 + |\rho|) \leq 4. \quad (13)$$

For the power law both are implied by

$$\gamma \in (0, \tfrac{1}{2}] \quad \text{and} \quad \eta(1 + |\rho|) \leq 2, \quad (14)$$

because $\theta\varphi(\theta) < \eta$ for every θ , and $\sup_{\theta>0} \theta\varphi(\theta)^2 \leq \eta^2$ precisely on $\gamma \leq \frac{1}{2}$: that supremum equals η^2 at $\gamma = \frac{1}{2}$ and grows without bound above it.

Calendar admissibility asks $\gamma \leq 1$; it is the butterfly *sufficient* condition (14) that asks $\gamma \leq \frac{1}{2}$. We impose (14) in calibration, through the parameter transform of Section 4, and the validator re-checks (12) and (13) on the fitted parameters afterwards.

Three properties of the family follow. First, (12) and (13) constrain φ as a *function* of θ , not a finite set of slices. So once (14) holds and θ_t is non-decreasing, it covers every t , interpolated expiries included. Second, one global ρ fixes the skew sign at every tenor, and, up to φ , the shape along with it. Third, SSVI does not escape the time seam that every per-expiry smoother carries, but it does localize it. Differentiating (10) gives

$$\partial_t w(k, t) = \dot{\theta}_t \left[\frac{w}{\theta_t} + k \frac{\varphi'(\theta_t)}{\varphi(\theta_t)} \partial_k w \right], \quad \frac{\varphi'}{\varphi} = - \left[\frac{\gamma}{\theta} + \frac{1 - \gamma}{1 + \theta} \right] \quad (15)$$

for the power law (11). We interpolate θ_t linearly between ATM pillars, so $\dot{\theta}_t$ is piecewise constant and $\partial_t w$ still steps at each one. But the bracketed shape factor stays continuous in both k and t . SSVI's step therefore rescales the whole smile uniformly, where a per-expiry smoother's step changes its shape.

3.3 eSSVI

eSSVI (Hendriks and Martini, 2019) keeps SSVI's slice shape and frees its parameters expiry by expiry. Writing $\psi_t = \theta_t \varphi_t$, so that the ATM skew is exactly $\rho_t \psi_t$, each slice is

$$w_{\text{eSSVI}}(k, t) = \frac{\theta_t}{2} \left[1 + \rho_t \frac{\psi_t k}{\theta_t} + \sqrt{\left(\frac{\psi_t k}{\theta_t} + \rho_t \right)^2 + 1 - \rho_t^2} \right], \quad (16)$$

with $w(0, t) = \theta_t$ and $\partial_k w(0, t) = \rho_t \psi_t$ by construction. Setting $\rho_t \equiv \rho$ and $\psi_t = \theta_t \varphi(\theta_t)$ recovers (10) exactly, so SSVI is the constrained special case.

Three properties distinguish it from “SVI with fewer parameters”. First, taking θ_t from the quoted ATM variance rather than fitting it leaves **two** free parameters per slice. Second, interpolation stays inside the family. The surface between pillars comes from interpolating (θ, ψ, ρ) , so every intermediate expiry is itself an eSSVI slice, whereas interpolating two SVI slices in total variance yields a smile that is not SVI. Third, admissibility splits cleanly by kind and is stated on the slice parameters themselves. With $\varphi_t = \psi_t / \theta_t$, the butterfly bounds (13) carry over one slice at a time as

$$\psi_t(1 + |\rho_t|) < 4, \quad \frac{\psi_t^2}{\theta_t}(1 + |\rho_t|) \leq 4, \quad (17)$$

while calendar admissibility couples adjacent slices. Two eSSVI slices at $t_1 < t_2$ can only avoid crossing if their asymptotic wing slopes of total variance, $\psi_t(1 \pm \rho_t)$, do not decrease in maturity.

For $\theta_2 > \theta_1 > 0$, [Pasquazzi \(2023\)](#) shows that this wing condition, together with a second condition on the curvature ratio φ_2/φ_1 , is also sufficient ([Appendix G](#)). Taking the simpler branch of the second condition, $\varphi_2 \leq \varphi_1$, gives the pair

$$\psi_1 p_2 \leq \psi_2 \leq \psi_1 \frac{\theta_2}{\theta_1}, \quad p_2 := \max\left(\frac{1+\rho_1}{1+\rho_2}, \frac{1-\rho_1}{1-\rho_2}\right), \quad (18)$$

whose lower bound is the wing condition written as a single ratio and whose upper bound is $\varphi_2 \leq \varphi_1$ expressed through ψ . Slices satisfying (18) never cross. [Hendriks and Martini \(2019\)](#) originally stated the wing condition in a squared form that loses the sign of the comparison and admits crossing pairs; [Appendix G](#) records the original proposition and the correction. Both conditions constrain the slice parameters $(\theta_t, \psi_t, \rho_t)$ directly and say nothing about how ψ_t varies with θ_t . Once a power-law backbone sets $\psi_t = \theta_t \varphi(\theta_t)$, the exponent γ does enter (18) through ψ_1 and ψ_2 , exactly as it enters the SSVI conditions of [Section 3.2](#).

3.4 The global parametrization, and the power-law backbone

[Mingone \(2022\)](#) reparameterize the N slices so that the Hendriks–Martini conditions cannot be violated by any admissible parameter vector, and calibrate all maturities in one joint least-squares problem.

For successive slices $(\rho_1, \theta_1, \psi_1)$ and $(\rho_2, \theta_2, \psi_2)$ the conditions are

$$\theta_2 > \theta_1 > 0, \quad \psi_1 p_2 < \psi_2 \leq \min\left(\frac{\psi_1}{\theta_1} \theta_2, \frac{4}{1+|\rho_2|}, \sqrt{\mathfrak{f}(\theta_2, |\rho_2|)}\right), \quad p_2 := \max\left(\frac{1+\rho_1}{1+\rho_2}, \frac{1-\rho_1}{1-\rho_2}\right), \quad (19)$$

with $\mathfrak{f}$ the butterfly bound (we use the Gatheral–Jacquier form $\mathfrak{f}_{GJ} = 4\theta/(1+|\rho|)$; the sharper Martini–Mingone bound requires an inner minimization per evaluation).

Mingone’s construction takes free parameters $\rho_1, \dots, \rho_N \in (-1, 1)$, $\theta_1 > 0$ and $a_2, \dots, a_N > 0$, and builds the ladder slice by slice:

$$\theta_i = \theta_{i-1} p_i + a_i, \quad A_{\psi_i} = \psi_{i-1} p_i, \quad C_{\psi_i} = \min\left(\frac{\psi_{i-1}}{\theta_{i-1}} \theta_i, S_i\right), \quad S_i = \min(f_i, S_{i+1}/p_{i+1}), \quad (20)$$

where f_i is slice i ’s own butterfly cap. The interval (A_{ψ_i}, C_{ψ_i}) , the *tube*, holds the skew scales slice i may take: its floor is the wing-slope condition of (19), and its ceiling the smaller of the calendar bound $\psi_{i-1}\theta_i/\theta_{i-1}$ and S_i , the butterfly cap carried back from every later slice. That backward recursion is the look-ahead: a ψ_i inside its tube can never strand a later slice, so any choice strictly inside every tube yields N slices free of butterfly and calendar-spread arbitrage, which is Mingone’s [Proposition 3.1](#).

The backbone. Where [Mingone \(2022\)](#) places each ψ_i anywhere in its tube, with a free coefficient $c_i \in (0, 1)$ per slice, we tie it to SSVI’s power-law curvature,

$$\hat{\psi}(\theta) = \eta \left(\frac{\theta}{1+\theta} \right)^{1-\gamma}, \quad (21)$$

projected into the tube. Two global parameters replace one coefficient per slice, and the skew scale follows one smooth, increasing curve in θ instead of a set of admissible points, leaving ρ_i and θ_i as the only per-slice freedom. η and γ keep their SSVI meaning ([Section 3.2](#)), so the eSSVI and SSVI fits compare parameter by parameter, and the pin $\gamma = \frac{1}{2}$ becomes a hypothesis the free arm can test. The cost falls on slices whose target lands outside the tube: ψ_i is projected onto the nearest edge and leaves the backbone there.

The greedy variant. With the backbone supplying ψ , the joint solve is not the only route to the guarantee. The conditions (19) tie each slice only to the one before it, and calendar ordering is transitive: $w_1 \leq w_2$ and $w_2 \leq w_3$ give $w_1 \leq w_3$. A slice inside its own tube is therefore butterfly-free and lies above every earlier slice. A forward pass confines ρ_i to the interval on which slice i 's tube is non-empty, and that interval always contains ρ_{i-1} (Appendix G).

S-eSSVI walks the ladder this way. It reads θ_i off the quotes, fixes η by a preliminary pass and $\gamma = \frac{1}{2}$, sets ψ_i by projecting the backbone target into the tube, and fits ρ_i as the one free parameter per slice (Section 4). It keeps the same guarantee at per-expiry cost, and gives up three things relative to the joint solve: each ρ_i is optimal given its predecessor rather than jointly, η is fixed before the ρ_i rather than fitted with them, and θ_i is quoted rather than fitted.

4 Calibration, local volatility, and the barrier price

4.1 Fitting the smiles

Every fit except the SVI benchmark is a Levenberg–Marquardt least squares in an *unconstrained transform space*, so every iterate is a *domain-valid* parameter set ($\theta > 0$, $|\rho| < 1$, and so on). Only two kinds of arm make every iterate an *arbitrage-admissible* surface as well: the guaranteed arms, whose transforms map into the Mingone tube, and SSVI, whose transforms map into its sufficient set. The sequential eSSVI's penalties shape the objective away from violations but guarantee nothing, at any iterate or at the optimum. Arm by arm:

- **SVI** (per expiry): the quasi-explicit calibration of [Zeliade Systems \(2009\)](#), not Levenberg–Marquardt. At fixed (m, σ) the total-variance residuals are *linear* in $(a, d, c) = (a, \rho b \sigma, b \sigma)$, which solve exactly as a tiny constrained least squares whose box ($a \geq 0$, $|d| \leq c$, $c + |d| \leq 4\sigma/t$) enforces non-negative variance and Lee's wing bound by construction; only $(m, \ln \sigma)$ are searched numerically, by a deterministic grid seed refined with Nelder–Mead.
- **eSSVI** (per expiry, in time order): residuals in implied volatility; θ_t fixed to the quoted ATM total variance, leaving (ψ_t, ρ_t) free, with a butterfly penalty (weight 10) on the slice-wise bounds and a calendar penalty (weight 1000) against crossing the previously fitted slice.
- **SSVI** (global): residuals in total variance over all quotes; ρ , γ and η through bounded periodic (sine) maps onto $(-0.999, 0.999)$, $(0, \frac{1}{2}]$ and $(0, 2/(1 + |\rho|)]$ respectively, so that every iterate satisfies (14); a fit that reaches the exponent bound is finished with γ pinned there and (ρ, η) re-solved. The θ -pillars are fixed to quoted ATM total variances (not fitted), isolating the three shape parameters.
- **G-eSSVI, both arms** (global): residuals in implied volatility on the shared pillars; the Mingone parameters in unconstrained transforms (ρ_i through $\tanh$, θ_1 and the a_i in log space anchored at the observed ATM term structure, η in log space, and — in the γ -free arm — γ through a logistic map onto $(0, 1)$), so every Levenberg–Marquardt iterate maps to an arbitrage-free slice set by construction. The standard start takes $\rho_i = 0$, the a_i from the observed ATM variance increments, and η at mid-tube.
- **S-eSSVI** (greedy): a penalty-free per-slice pre-pass sets η by closed-form regression of the free ψ_i on the backbone; the guaranteed pass then fits one bounded ρ_i per slice (Levenberg–

Marquardt on the slice’s vol residuals, ρ mapped by tanh into the tube-preserving interval), with ψ_i projected onto the backbone inside the tube.

The eSSVI family fits implied-volatility residuals, while SVI and SSVI fit total-variance residuals, which SVI has no choice about because its quasi-explicit inner solve is linear only in total variance. Since $dw = 2\sigma t d\sigma$, a total-variance objective weights each expiry’s vol errors by $(2\sigma t)^2$, so the long expiries dominate it; SSVI, a joint fit, carries that weighting across the whole surface, and part of its short-end fit gap is the objective rather than its three parameters. We report implied-volatility RMSE for every arm regardless, so the comparison metric stays uniform even where the objectives are not. That metric is defined once for the whole paper. Let expiry j carry n_j repaired pillar quotes (K_i, σ_i) , let $\hat{\vartheta}$ be an arm’s parameter vector fitted on every pillar, and let $\hat{\vartheta}_{-j}$ be the same arm refitted with expiry j removed. The in-sample and leave-one-expiry-out errors of expiry j are

$$\begin{aligned} \text{RMSE}_j &= 10^4 \sqrt{\frac{1}{n_j} \sum_{i \in j} (\sigma_{\text{fit}}(K_i, t_j; \hat{\vartheta}) - \sigma_i)^2}, \\ \text{RMSE}_j^{\text{LOO}} &= 10^4 \sqrt{\frac{1}{n_j} \sum_{i \in j} (\sigma_{\text{fit}}(K_i, t_j; \hat{\vartheta}_{-j}) - \sigma_i)^2}, \end{aligned} \tag{22}$$

in basis points of implied volatility, unweighted, at the expiry’s own strikes. The two differ only in which fit is scored: the leave-one-out error of Section 6.1.2 is evaluated at the held-out expiry alone, on a surface that never saw its quotes, using the arm’s own interpolation between the neighbouring pillars to reach t_j .

4.2 From fitted surface to local volatility

We build a surface in total implied variance from the fitted slices, and then we differentiate it. Collecting (3) and (6), the quantity evaluated at each grid point is

$$\sigma_{\text{loc}}^2(k, t) = \frac{\partial_t w(k, t)}{\left(1 - \frac{k \partial_k w}{2w}\right)^2 - \frac{(\partial_k w)^2}{4} \left(\frac{1}{w} + \frac{1}{4}\right) + \frac{\partial_k^2 w}{2}}, \tag{23}$$

so the whole construction runs on total implied variance, its first two strike derivatives and its time derivative, and on nothing else. The denominator is the butterfly function g of (3). In this construction a negative density and an unusable local volatility are the same event.

The strike derivatives come in closed form from each arm’s own slice formula, so no numerical k -differentiation enters the comparison at all. The time derivative comes from the interpolation rule, and every arm uses the same one: linear in t in its own natural coordinates, with proportional accrual below the first pillar. The benchmark interpolates total variance itself at fixed log-moneyness, the eSSVI arms interpolate (θ, ψ, ρ) and re-evaluate the slice formula, and SSVI interpolates θ_t with its global shape parameters held fixed.

4.3 Pricing the knock-out barrier

The pipeline's last stage is a continuously monitored single-barrier knock-out. Write H for the barrier level and $D(0, T)$ for the discount factor. The two contracts we report are

$$V_{\text{DO}}^{\text{put}} = D(0, T) \mathbb{E} \left[(K - S_T)^+ \mathbf{1} \left\{ \min_{0 \leq t \leq T} S_t > H \right\} \right], \quad (24)$$

$$V_{\text{UO}}^{\text{call}} = D(0, T) \mathbb{E} \left[(S_T - K)^+ \mathbf{1} \left\{ \max_{0 \leq t \leq T} S_t < H \right\} \right], \quad (25)$$

and they read opposite wings of the same smile. Strike and barrier are set in volatility units from the parity forward F of the maturity,

$$K = F \exp(\pm c \sigma_{\text{ATM}} \sqrt{T}), \quad H = F \exp(\pm \sigma_{\text{ATM}} \sqrt{T}), \quad (26)$$

the upper signs for the call and the lower for the put, with σ_{ATM} the market at-the-money vol of the repaired surface at T , identical for every arm. We price two designs at six months: at the money, $c = 0$, and out of the money, $c = \frac{1}{2}$, where the strike sits half a standard deviation into the wing and the contract reads the smile between strike and barrier. We price the knock-outs rather than the knock-ins because in-out parity makes the latter redundant, a knock-in and a knock-out summing to the vanilla.

The expectation is taken under (7), which we discretize in logs on a grid $0 = t_0 < \dots < t_N = T$ with the local volatility frozen at the left node:

$$\ln S_{n+1} = \ln S_n + \ln \frac{F(t_{n+1})}{F(t_n)} - \frac{1}{2} \sigma_n^2 \Delta t_n + \sigma_n \sqrt{\Delta t_n} Z_{n+1}, \quad \sigma_n = \sigma_{\text{loc}}(S_n, t_n), \quad (27)$$

with the Z_{n+1} independent standard normals.

A knock-out monitored only at the nodes misses the excursions between them and prices too high, with a bias of order $\sqrt{\Delta t}$ (Broadie et al., 1997) comparable to the cross-arm gaps we are trying to resolve. So we do not carry an indicator. Conditional on its two endpoints and with σ_n frozen, the log-path across a step is a Brownian bridge, and the bridge's probability of never touching H is known in closed form. Each path therefore carries a survival weight

$$w_T = \prod_{n=0}^{N-1} p_n, \quad p_n = 1 - \exp \left(- \frac{2 \ln(S_n/H) \ln(S_{n+1}/H)}{\sigma_n^2 \Delta t_n} \right), \quad (28)$$

with the logarithms inverted to $\ln(H/S_n)$ for an up barrier, and with w_T set to zero the moment a node itself lands on the far side. The estimator for (24) is $D(0, T) M^{-1} \sum_m w_T^{(m)} (K - S_T^{(m)})^+$ over M paths, and the knock-out probability recorded with it is $1 - \mathbb{E}[w_T]$. The bridge weight removes the monitoring bias and, being continuous in the path where an indicator is not, lowers the estimator's variance as well.

5 Data and pipeline

5.1 Universe, collection, and the size of the panel

The panel is CBOE option chains¹ for eight liquid US names, the four most liquid index ETFs (SPY, QQQ, IWM and DIA) and four liquid single names (META, AMD, TSLA and GOOGL).

¹CBOE's public feed is delayed. That matters for trading on the quotes and not for comparing calibrations of them, since every capture is analyzed after the fact and all six arms fit the identical chain. We treat each capture as a timestamped snapshot throughout.

Two thinner single names, JPM and XOM, are run through the identical pipeline as an extension (Section 6.4) and enter no pooled statistic. We collect the chains every 15 minutes during regular trading hours over the nineteen sessions from 2026-08-14 to 2026-09-11, four calendar weeks.

Index ETF chains are where smile parameterizations are most heavily deployed. They carry hundreds to thousands of live quotes per capture, dense weekly and daily expiries at the short end, and listed maturities past a year. The single names extend the panel beyond that setting. Their volatility is higher, their carry contains real dividends and borrow, and their early-exercise premia are large enough to make the de-Americanization step below quantitatively material. They test whether conclusions calibrated on index surfaces carry over to single-name chains. Expiries from one day to 1.5 years enter the fits, so the panel spans the ultra-short smiles and the LEAP wings in one surface.

Two slices are used, and they play different roles.

- The **daily slice** takes one capture per name-session (the capture nearest 15:30 ET) and carries the headline cross-sectional results.
- The **intraday slice** retains every quarter-hour capture and serves as a stability instrument, repeating the daily measurements on roughly an order of magnitude more captures through the identical pipeline. More observations do not imply proportionally more information: the feed refreshes unevenly, and consecutive captures can share the same quote timestamp. Where a measurement requires a genuine refresh, we select captures explicitly by timestamp and report the surviving count.

Nineteen sessions give eighteen day-over-day transitions per name, and the intraday slice supplies up to 27 snapshots per name per session for the stability instrument.

5.2 From quotes to fitting inputs

Each chain passes through identical preparation for all six arms, with every parameter fixed before the models are run:

1. **Discounting:** zero curve bootstrapped from US Treasury par yields for the date and rolled to the chain’s as-of date.
2. **De-Americanization:** every option in the panel is American-style, so put–call parity — which the forwards below rely on — holds only as an inequality in the raw bids. Before any implied vol is solved, each quote’s bid is lowered by its early-exercise premium, estimated per expiry in four steps: the parity forward $\hat{F}$ is implied from the current bids; the carry $\hat{q} = r - \ln(\hat{F}/S)/t$ is read off it, so that dividends, borrow and basis enter through the quotes rather than through an assumed yield; each quote’s current bid is inverted to a vol with the Black formula on $\hat{F}$, and the quote is priced at that vol under $(r, \hat{q})$ both American, on a 201-step CRR tree, and European, in closed form; the difference, floored at zero, is subtracted from the raw bid. The corrected bids feed a second round, in which the forward, the carry and the premium are re-estimated; the corrected pairs then satisfy parity as an identity, and the premium barely moves. The premium inherits the tree’s discretization error, a few tenths of a basis point of spot at 201 steps and falling as $1/N$, which is well inside the put premia it measures and below the resolution of every statistic reported. Per-expiry outcomes (forward shifts, implied carries, premium sizes) are released in `dea.csv`.

3. **Quote filters:** quoted and uncrossed; relative spread $\leq 25\%$; $K/F \in [0.6, 1.6]$; expiries between 1 day and 1.5 years; OTM side only; at least 5 surviving quotes per expiry; implied vol inside $[0.5\%, 500\%]$.
4. **Forwards:** put–call parity on the de-Americanized mids, using the five most ATM strike pairs.
5. **Implied vols:** safeguarded Newton with arbitrage-bound rejection; quotes violating static bounds are dropped and counted.
6. **Pre-repair:** a minimal-change static-arbitrage repair (per-expiry convex-hull projection of call prices plus a forward total-variance monotonicity sweep), applied *once*, so all six arms fit the same arbitrage-clean pillar set.

5.3 Sample attrition

The full cleaning cascade is reported in Appendix B. Table 11 shows quote-level attrition, while Table 12 summarizes the inputs received by each fit. Roughly a quarter of the quotes reach the fits, and half the loss is the out-of-the-money selection we impose by design.

6 Results

Tables report medians, with interquartile ranges or 90th percentiles where indicated; every distribution here is skewed, so a mean would follow tails that say nothing about the typical capture. The daily panel contains 152 captures ($8 \text{ names} \times 19 \text{ sessions}$), all of which survive cleaning. The intraday panel contains 3,877 captures, of which 3,776 survive. A median index capture carries 30 expiry slices and about 2,350 cleaned quotes; a median single-name capture carries 20 slices and about 1,050 quotes.

6.1 The fitted surface and its cost

6.1.1 Fit quality

Table 2 shows fit quality across the six arms. Raw SVI fits tightest at 23.2 bp per slice, followed by sequential eSSVI at 60.2 bp. The guaranteed construction adds 7.2 bp with a free exponent, and another 8.3 bp once $\gamma = \frac{1}{2}$ is pinned. The greedy variant adds 4.3 bp more, reaching 80.1 bp, because its slice-local ρ path forces the backbone target to be projected into the tube (Section 3.4) far more often than the joint fit does: a median 22.6% of slices per capture, against 4.8% and 5.3% for the two joint arms. SSVI is loosest at 102.2 bp, with three shape parameters for the full surface and its exponent held inside the sufficient-condition domain. Every guaranteed variant remains within roughly 20 bp of sequential eSSVI. Figures 1 and 2 show the fits on one capture for every name. Inside the quoted span the arms mostly agree to within the bid–ask. The visible departures are the joint G-eSSVI arms sitting below the quotes at the money on AMD, META and TSLA, the re-leveling of θ taken up in Section 6.1.2, and the steepest one-month put wings, which rise above every arm on SPY and QQQ.

Across the sample, the structured methods fit single names 19–36 bp tighter than the index ETFs, while raw SVI barely notices the difference (23.5 against 22.7 bp). Single-name captures carry fewer expiries, about 20 against about 30, which gives the global backbone fewer slices to reconcile.

Table 2: **Fit quality across the six arms.** Per-slice root-mean-square residual against the shared repaired pillars, in implied-volatility basis points: medians and 90th percentiles. Bucket n is the number of expiry slices, identical across arms because every arm fits the same pillars.

Stratum	Slices	S-eSSVI $\gamma=\frac{1}{2}$		G-eSSVI $\gamma=\frac{1}{2}$		G-eSSVI γ free		eSSVI (seq.)		SSVI		SVI (bench.)	
		Med.	p90	Med.	p90	Med.	p90	Med.	p90	Med.	p90	Med.	p90
<i>By tenor bucket (daily)</i>													
≤1W	399	138.0	455	107.8	286	119.0	282	90.1	297	208.8	566	27.5	125
1W–1M	1 022	151.4	365	125.4	243	125.3	241	129.6	258	174.2	433	37.7	140
1–6M	1 201	84.3	188	77.7	162	69.7	141	68.8	151	104.4	228	24.9	66
>6M	982	24.3	59	43.2	85	24.2	56	17.9	51	31.8	67	8.7	31
<i>By ticker (daily)</i>													
SPY	577	126.4	252	108.6	214	113.4	207	104.1	218	157.9	298	42.9	79
QQQ	561	125.0	210	104.3	175	107.9	175	105.9	185	150.3	263	27.9	58
IWM	569	77.4	141	68.7	129	68.7	129	67.4	134	98.1	192	17.8	33
DIA	346	57.4	115	43.4	80	41.1	78	35.5	66	68.2	167	8.0	40
META	383	35.9	138	42.3	114	33.3	110	26.8	113	58.2	201	16.6	50
AMD	398	87.7	298	67.0	167	52.9	172	48.9	185	120.9	391	17.5	55
TSLA	383	152.7	636	173.9	344	128.6	336	128.2	369	163.8	669	80.7	238
GOOGL	387	56.0	181	48.6	126	45.8	120	41.5	112	69.8	225	24.0	62
All daily slices	3 604	80.1	244	75.8	187	67.5	182	60.2	191	102.2	300	23.2	74
All intraday slices	88 613	76.8	232	74.8	184	65.5	179	58.1	186	100.9	290	23.2	73
<i>Solver outcome (% of fitted slices; daily / intraday)</i>													
Converged		99.2	99.0	69.7	69.9	75.5	74.2	97.3	96.9	98.3	98.3	100.0	100.0
Passes own conditions		100.0	100.0	100.0	100.0	100.0	100.0	65.4	67.6	100.0	100.0	100.0	100.0

“Passes own conditions” applies each model’s own parameter-level sufficient no-arbitrage conditions to its own fitted parameters. It is a certificate the model issues about itself rather than a measured violation count (Table 5 measures the surfaces), and the conditions differ by model, so the column is not a like-for-like comparison across arms.

“Converged” in the table means that the solver reached its own tolerance criterion before the iteration limit. Convergence is not treated as a quality criterion for the joint fits. Alternative starting points reach parameter vectors that are distant in the transformed parameter space, a median $\|x\|_\infty$ distance of 100–139, at RMSE differences of less than 0.01 bp.

Stability through the session. The intraday refits add two findings about parameters. Raw SVI’s instability is a single-name phenomenon: on the index chains no arm’s worst within-session move exceeds 15 panel IQRs and the benchmark’s is 7.6, while on the single names GOOGL’s SVI b jumps 6 800 IQRs between consecutive snapshots against 8 to 15 for every structured arm. The one structured arm the refits flag is the greedy one, whose snapshot-to-snapshot ATM skew moves carry a p99 of 0.30 against 0.07 to 0.15 elsewhere, the time-series counterpart of its slice-local ρ path.

6.1.2 Out-of-sample interpolation

Table 3 reports the exhaustive leave-one-expiry-out test on the daily slice. Within each capture we hold out one interior expiry at a time: its quotes are removed, all six arms are refitted on the remaining pillars, the held-out smile is evaluated at its original strikes against the same repaired pillar vols the in-sample fit uses, and the step repeats for the next expiry. The first and last pillars are never held out, since that would test extrapolation.

The joint arms show a systematic ATM bias. The greedy arm, penalty fit, and SSVI, which anchor θ to the quoted ATM variance, remain within 3 bp of the held-out ATM level; the

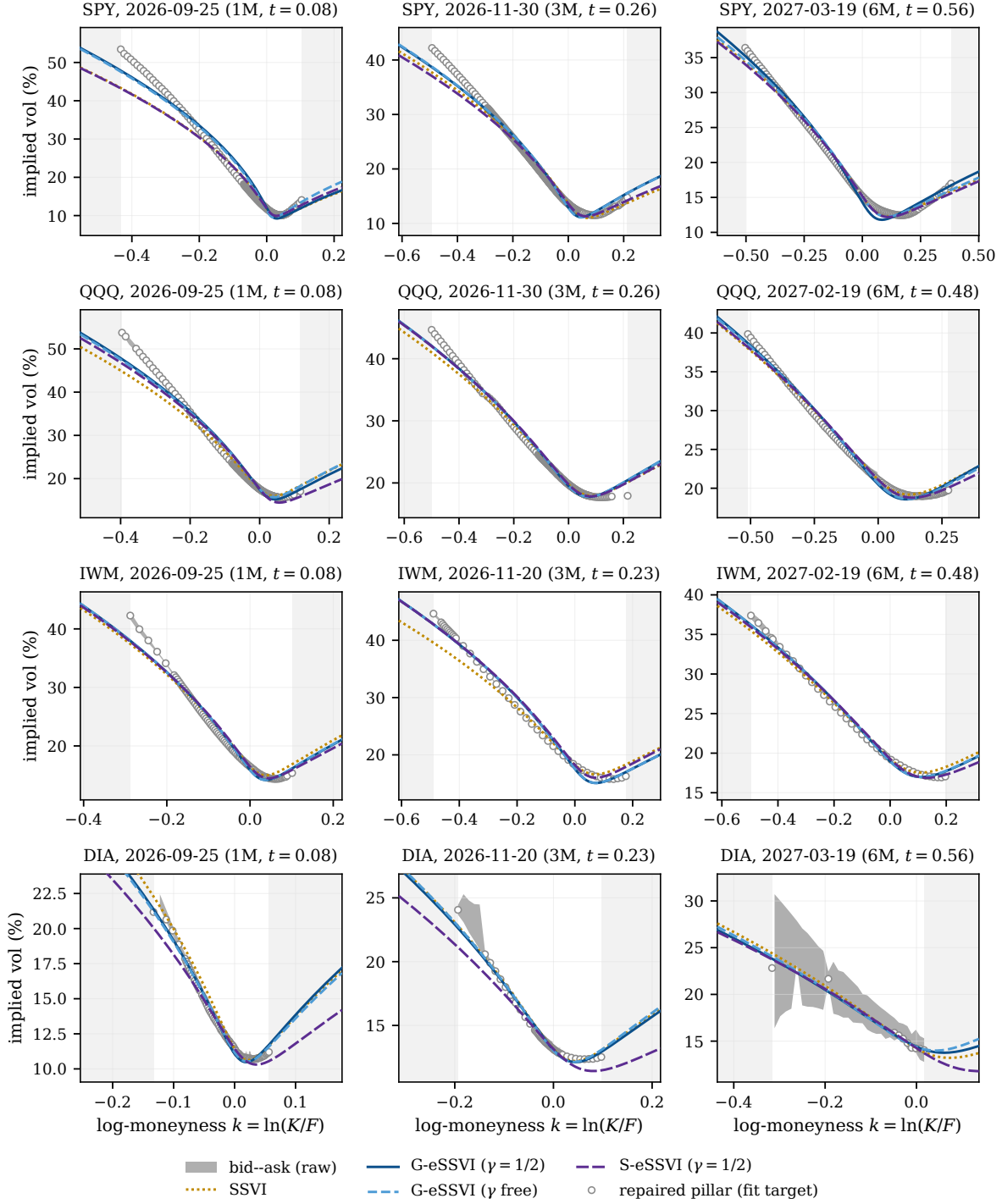


Figure 1: **Fitted smiles against the market: index ETFs (daily slice, 2026-08-27).** Each name at the pillar expiries nearest one, three and six months. Dots are the repaired pillar vols every arm is fitted to; the grey band is the raw bid-ask in implied vol; the curves are SSVI, the two joint G-eSSVI arms and S-eSSVI at their fitted parameters. Outside the quoted span, shaded, the curves are extrapolation.

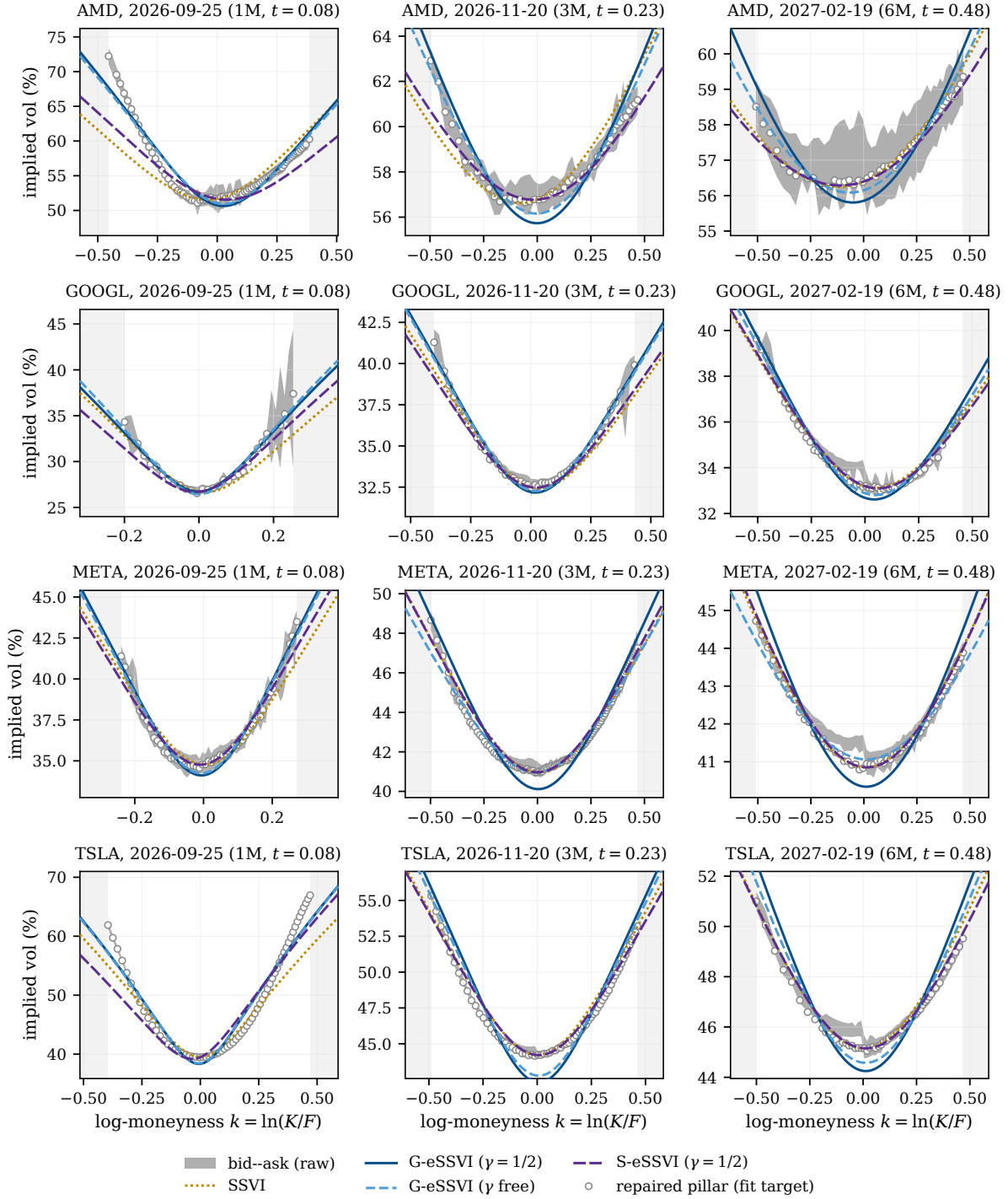


Figure 2: Fitted smiles against the market: single names (daily slice, 2026-08-27). As Figure 1.

Table 3: **Out of sample: the leave-one-expiry-out test (daily slice).** Each interior pillar is held out in turn, every arm refitted on the rest and evaluated at the held-out strikes. Both errors are measured against the same repaired pillar vols, in implied-volatility basis points, so their ratio measures what interpolation costs.

Model	n	In-sample		LOO RMSE (bp)		LOO/IS	ATM bias	LOO median by tenor (bp)		
		med. bp	med.	[IQR]	med.			$\leq 1M$	1–6M	$> 6M$
S-eSSVI ($\gamma = \frac{1}{2}$)	3 300	82.0	108.0	[52.1, 187.6]	1.13	2.6	183.7	109.8		37.5
G-eSSVI ($\gamma = \frac{1}{2}$)	3 300	77.4	103.8	[63.1, 167.1]	1.14	-52.2	163.9	106.3		57.4
G-eSSVI (γ free)	3 300	69.1	99.4	[52.2, 160.0]	1.19	-39.8	162.0	99.7		38.6
eSSVI (seq.)	3 300	63.9	101.2	[47.1, 170.1]	1.26	2.6	173.1	103.6		31.5
SSVI	3 300	103.6	124.0	[57.2, 213.0]	1.05	2.6	210.0	125.0		41.7
SVI (benchmark)	3 300	24.4	52.7	[24.6, 97.4]	1.90	4.3	88.1	53.8		20.2

One row per (capture, arm, held-out interior expiry). *In-sample* is the full fit’s RMSE at the same held-out expiries, so the LOO/IS ratio compares like with like (medians of per-row ratios). *ATM bias* is the signed error, model minus market, at the quote nearest the forward; negative means the interpolated smile sits below the market at the money.

benchmark is within 5 bp. The two jointly calibrated arms instead lie 39.8 and 52.2 bp below the held-out level. This is systematic: the joint arms are below the market on 78.7% and 82.4% of held-out expiries, compared with about 45% for the other arms. The effect persists across tenor buckets, with both joint arms above 68% in every bucket. Their jointly fitted θ term structure smooths through locally elevated ATM levels in the quotes, and that smoothing carries into the interpolated surface. The effect is confined to dense chains: on the two thin chains of Section 6.4 the joint arms sit within about 6 bp of the held-out level.

The benchmark shows the largest out-of-sample degradation. Its median LOO error is 1.90 times its in-sample error, while the structured arms remain much closer to their in-sample performance, with LOO-to-in-sample ratios between 1.05 and 1.26.

The effect of fixing γ is most visible at the long end. The pinned joint arm has a median LOO error of 57.4 bp there, compared with 38.6 bp for the free- γ arm.

6.1.3 The fitted exponent

Table 4 reports the freely fitted G-eSSVI exponent name by name, alongside an SSVI refit with the sufficient-condition bound removed, a diagnostic of the exponent the data favour. The median free exponent sits above $\frac{1}{2}$ for every name but DIA, so the pin imposes a slower decay of the at-the-money skew than the data prefer; Appendix D extends the analysis to sixteen further names and finds two liquid names below $\frac{1}{2}$.

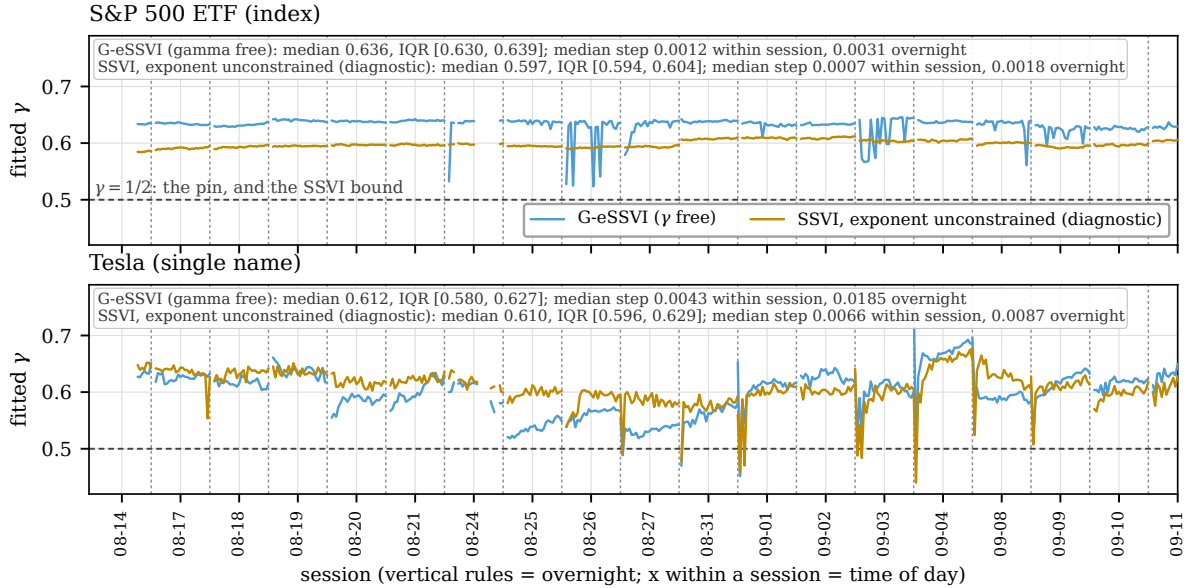
DIA illustrates the identification issue. It has the lowest free exponent in the panel and the highest unconstrained SSVI exponent. The two should not be read separately, because γ and η are strongly correlated, from -0.83 to -0.99 across captures in both arms, and DIA sits at the extreme of that trade-off with the largest G-eSSVI η (1.52) and the smallest SSVI η (0.19). The cause is the limited variation in its ATM variance: a median ratio of 309 between its largest and smallest pillar, against 970 to 1 880 on the other index names, over 18 pillars rather than 30.

The exponent is stable through time. Every name’s median day-over-day move is at most 0.04, against a spread of 0.19 between the name medians (Figures 3 and 7); for TSLA the overnight steps are about four times the within-session ones.

SSVI frequently reaches its sufficient-condition bound. The constrained optimum is at $\gamma = \frac{1}{2}$

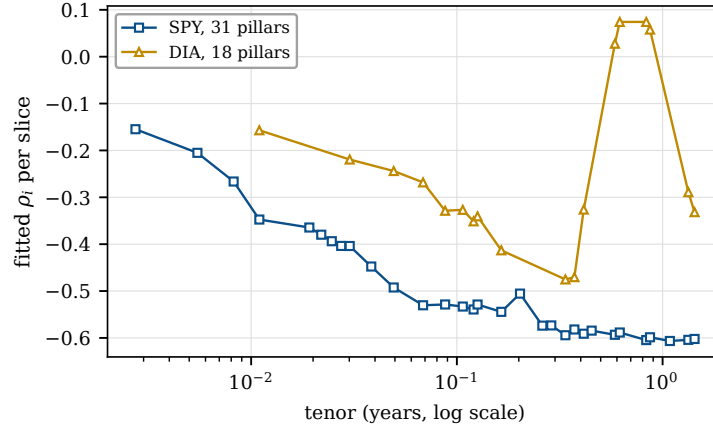
Table 4: **The fitted exponent, name by name.** Median [IQR] of the fitted power-law exponent per capture on the intraday slice, and the share of captures above $\frac{1}{2}$. The second block is the unconstrained SSVI refit, a diagnostic: the SSVI arm itself is calibrated inside $(0, \frac{1}{2}]$ and sits on the bound on 87% of these captures.

Ticker	G-eSSVI (γ free)			SSVI, exponent unconstrained		
	Med.	IQR	$> \frac{1}{2}$ (%)	Med.	IQR	$> \frac{1}{2}$ (%)
SPY	0.636	[0.630, 0.639]	100	0.597	[0.594, 0.604]	100
QQQ	0.527	[0.508, 0.546]	81	0.597	[0.590, 0.615]	100
IWM	0.508	[0.499, 0.521]	72	0.604	[0.595, 0.611]	100
DIA	0.450	[0.422, 0.474]	1	0.851	[0.795, 0.874]	100
META	0.573	[0.549, 0.597]	100	0.485	[0.451, 0.513]	36
AMD	0.562	[0.541, 0.592]	97	0.539	[0.494, 0.579]	72
TSLA	0.612	[0.580, 0.627]	99	0.610	[0.596, 0.629]	99
GOOGL	0.570	[0.558, 0.592]	100	0.573	[0.543, 0.598]	94
All	0.557	[0.512, 0.605]	81	0.596	[0.557, 0.616]	88



Intraday captures over 19 sessions, 940 captures across the two names. The constrained SSVI arm (not drawn) sits on its bound $\gamma = 1/2$ on 99% of these captures. Source: results/gamma_ext.csv (study --panel gamma); identical ingest to the intraday panel, so the G-eSSVI series is the intraday panel's own fit.

Figure 3: **The fitted exponent through four weeks (intraday slice).** Every intraday capture for SPY and TSLA: the G-eSSVI backbone exponent fitted freely, and the unconstrained SSVI refit. Vertical rules are overnights; the dashed line is the pin and the SSVI bound. Each panel reports the median step between consecutive snapshots within a session and across the overnight.



Session 2026-08-17 1530, G-eSSVI with γ free.

Figure 4: **Why DIA crosses between pillars (daily slice).** The fitted skew parameter ρ_i slice by slice on one session, for DIA and SPY, using G-eSSVI with γ free.

on 87% of captures. The unconstrained refit lowers the total-variance objective by construction, yet the implied-vol slice RMSE comes out 0.3 bp higher at the median on the intraday slice.

6.1.4 Static arbitrage

Table 5 counts fitted name-days with at least one butterfly violation or calendar crossing on the full (k, t) grid, as counts and rates only. We do not summarize severity with $\min g$: degenerate benchmark slices dominate it, and it gives no meaningful scale for how far a fit extends beyond the admissible region.

At the fitted pillars the three guaranteed arms are clean on every capture, as their construction requires, whereas the sequential penalty fit leaves residual crossings on 87 of 152 daily captures, up to 4×10^{-5} in total variance.

Interpolation introduces calendar crossings between pillars. Every crossing the scans find for the guaranteed arms lies *between* fitted pillars, where the surface interpolates (θ, ψ, ρ) linearly in t . Such crossings appear on 12 to 14% of name-days for the three guaranteed arms and on 23% for the penalty fit (Table 5, Panel A).

The crossings are concentrated in DIA, and the fitted parameters show why. Panel B of Table 5 breaks the counts down by name: DIA drives the panel rate, and no single name crosses under any guaranteed arm. Two admissible slices can still cross between pillars, because the floor in (19) rises whenever ρ moves ($p > 1$ for $\rho_{\text{mid}} \neq \rho_1$), so the interpolated parameters can leave the tube while both endpoints lie inside it. A crossing therefore needs a small ψ increment and a large ρ step across the same adjacent pair. Crossing name-days show both: their smallest ψ increment is 0.4–1.6% against 2.2–2.7% on clean ones, and their largest ρ step 0.33–0.46 against 0.09–0.14.

DIA supplies the ρ step. Its largest adjacent $|\Delta\rho|$ is 0.38 at the median capture, against 0.07–0.14 for every other name. Figure 4 shows a representative session: DIA’s ρ moves by 0.35 between neighbouring long-dated expiries and turns positive at the long end, where SPY’s largest step is 0.08. DIA’s thinner book pins the long-end skew less tightly, so we read that movement cautiously, and thinness alone does not produce it (Section 6.4). The greedy arm

Table 5: **Static arbitrage of the fitted surfaces.** Share of fitted name-days with at least one butterfly violation or calendar crossing, measured on the full (k, t) grid. Panel B gives, for each name, the share of its captures carrying at least one calendar crossing.

Panel A: by arm, share of name-days with at least one violation

Model	Days	Butterfly		Calendar	
		n	%	n	%
<i>Daily slice</i>					
S-eSSVI ($\gamma=\frac{1}{2}$)	152	0	0.0	21	13.8
G-eSSVI ($\gamma=\frac{1}{2}$)	152	0	0.0	18	11.8
G-eSSVI (γ free)	152	0	0.0	21	13.8
eSSVI (seq.)	152	0	0.0	35	23.0
SSVI	152	0	0.0	0	0.0
SVI (benchmark)	152	24	15.8	83	54.6
<i>Intraday slice</i>					
S-eSSVI ($\gamma=\frac{1}{2}$)	3 776	0	0.0	525	13.9
G-eSSVI ($\gamma=\frac{1}{2}$)	3 776	0	0.0	499	13.2
G-eSSVI (γ free)	3 776	0	0.0	550	14.6
eSSVI (seq.)	3 776	0	0.0	896	23.7
SSVI	3 776	0	0.0	0	0.0
SVI (benchmark)	3 776	602	15.9	2 095	55.5

Panel B: calendar crossings by name, share of captures

Name	Caps.	S-eSSVI $\gamma=\frac{1}{2}$	G-eSSVI $\gamma=\frac{1}{2}$	G-eSSVI γ free	eSSVI (seq.)	SSVI	SVI (bench.)
<i>Daily slice</i>							
SPY	19	5%	11%	11%	11%	0%	74%
QQQ	19	0%	0%	0%	11%	0%	68%
IWM	19	5%	5%	5%	53%	0%	100%
DIA	19	100%	79%	95%	100%	0%	100%
META	19	0%	0%	0%	5%	0%	32%
AMD	19	0%	0%	0%	5%	0%	5%
TSLA	19	0%	0%	0%	0%	0%	0%
GOOGL	19	0%	0%	0%	0%	0%	58%
All	152	14%	12%	14%	23%	0%	55%
<i>Intraday slice</i>							
SPY	466	6%	12%	11%	8%	0%	85%
QQQ	467	0%	0%	0%	12%	0%	63%
IWM	468	5%	9%	9%	58%	0%	96%
DIA	480	98%	83%	94%	100%	0%	100%
META	470	0%	0%	0%	1%	0%	29%
AMD	479	0%	0%	0%	7%	0%	14%
TSLA	474	0%	0%	0%	0%	0%	1%
GOOGL	472	0%	0%	0%	3%	0%	57%
All	3 776	14%	13%	15%	24%	0%	55%

Table 6: **The induced Dupire local volatility, name by name.** Panel A: curvature roughness of σ_{loc} per ticker. Panel B: clamp activity against the $[0.01, 3.0]$ bounds, the local-volatility range, and the finite-difference round trip back to each arm’s own smile. Definitions in the footnote.

<i>Panel A: roughness $\times 10^4$, median by name (daily slice)</i>										
Model	SPY	QQQ	IWM	DIA	META	AMD	TSLA	GOOGL	All [IQR]	
S-eSSVI ($\gamma=\frac{1}{2}$)	4.70	3.14	3.08	4.13	0.98	0.59	1.39	1.13	2.10	[1.09, 3.24]
G-eSSVI ($\gamma=\frac{1}{2}$)	6.05	3.73	3.50	3.73	1.19	1.07	2.56	1.48	2.96	[1.31, 3.73]
G-eSSVI (γ free)	6.04	3.71	3.50	3.83	1.17	1.03	2.39	1.57	3.05	[1.39, 3.79]
eSSVI (seq.)	4.65	3.03	3.04	6.12	1.27	0.92	1.69	1.48	2.46	[1.45, 4.25]
SSVI	4.43	2.93	2.85	2.65	0.82	0.61	1.29	1.00	1.75	[0.93, 2.93]
SVI (benchmark)	10.00	6.06	5.70	15.36	3.80	3.52	3.04	3.63	5.57	[3.57, 10.43]
<i>Panel B: containment and round trip</i>										
Model	Daily slice				Intraday slice					
	Clamped	$\sigma_{\min}$	$\sigma_{\max}$	Round trip bp	Rough. All [IQR]	Clamped	$\sigma_{\max}$	Round trip bp		
S-eSSVI ($\gamma=\frac{1}{2}$)	0.00%	0.153	1.618	43.5	1.74 [1.07, 3.38]	0.00%	1.674	45.1		
G-eSSVI ($\gamma=\frac{1}{2}$)	0.00%	0.145	1.708	47.0	2.97 [1.30, 3.83]	0.00%	1.763	48.7		
G-eSSVI (γ free)	0.00%	0.143	1.796	48.5	3.07 [1.35, 3.90]	0.00%	1.841	47.9		
eSSVI (seq.)	0.00%	0.078	1.560	49.4	2.60 [1.45, 4.28]	0.00%	1.643	50.7		
SSVI	0.00%	0.163	1.251	45.3	1.53 [0.91, 2.91]	0.00%	1.300	46.7		
SVI (benchmark)	0.90%	0.010	2.432	45.6	5.74 [3.63, 10.39]	1.08%	2.614	47.2		

Roughness is $10^4 \text{mean}_{(k,t)} |\sigma_{\text{loc}}(k + \Delta k, t) - 2\sigma_{\text{loc}}(k, t) + \sigma_{\text{loc}}(k - \Delta k, t)|$ over the shared grid, at $\Delta k = 0.01$; a value of 3 means neighbouring-strike deviations of about 3 vol basis points from the local line. *Clamped* is the median share of grid cells pinned at the $[0.01, 3.0]$ bounds. $\sigma_{\min}$ and $\sigma_{\max}$ are medians of the per-day extremes, so day-level tail events do not appear in them. *Round trip bp* is Crank–Nicolson local-vol repricing of the calibrating vanillas, re-inverted to implied vol, against the arm’s own fitted smile.

crosses on every DIA capture, the joint arms on 79% and 95%, because its slice-by-slice ρ path takes larger steps. The pillar-level guarantee is the same for all three; on the evidence of one name, the joint calibration interpolates more smoothly.

SSVI’s construction-level guarantee extends across the interpolated surface. With its exponent and η restricted to (14), SSVI shows no butterfly violation and no calendar crossing on any capture of either slice. It is the only arm whose full-grid scan is clean, and the reason is structural: SSVI constrains φ as a function of θ , so its sufficient conditions cover the continuous term structure rather than a finite set of fitted slices.

The benchmark is the only arm with butterfly violations, and its induced local volatility reaches the upper Dupire clamp on 35% of captures, against at most 2% for any other arm.

6.2 The induced local volatility

6.2.1 Roughness and the round trip

Table 6 shows that the fit ranking does not carry directly into local-volatility roughness. SVI, the tightest fit, is the roughest surface, 1.8 to 3.2 times the structured arms, while SSVI, the arm with the strongest structural restriction, produces the smoothest local volatility, first in the pooled ranking and on seven of the eight names. The other structured arms sit close together with overlapping interquartile ranges, and the two joint eSSVI arms are indistinguishable on every name.

Table 7: **Martingale check, per arm.** Absolute relative error $|\mathbb{E}[S_T]/F_T - 1|$ in basis points, and $|z|$, pooled across the 1M, 3M and 6M horizons; both are defined in Section 6.2.2.

Model	Daily				Intraday			
	n	Med. bp	p99 bp	max $ z $	n	Med. bp	p99 bp	max $ z $
S-eSSVI ($\gamma=\frac{1}{2}$)	456	1.62	9.3	2.67	11 328	1.58	9.6	3.36
G-eSSVI ($\gamma=\frac{1}{2}$)	456	1.43	9.4	2.39	11 328	1.39	9.7	3.11
G-eSSVI (γ free)	456	1.46	9.1	2.41	11 328	1.40	9.6	2.68
eSSVI (seq.)	456	1.50	9.0	2.53	11 328	1.46	9.2	2.71
SSVI	456	1.27	8.2	1.64	11 328	1.22	8.2	1.64
SVI (benchmark)	456	1.20	8.7	1.85	11 328	1.20	9.5	4.21
All	2 736	1.39	9.1	2.67	67 968	1.37	9.4	4.21

The name effect is as large as the arm effect. On the index names the common linear-in- t interpolation puts a seam in $\partial_t w$ at each of roughly thirty pillars, and that seam term, identical across arms, dominates the roughness; the single names, with fewer and sparser pillars, are smoother despite their higher volatility. Within the index names the arms that take θ from the quoted ATM variance are generally smoother, except on DIA, where the joint arms are smoother than the other eSSVI arms. Both patterns are consistent with the role of $\partial_t w$ in the Dupire numerator: increments of the ATM term structure feed straight into the time derivative of total variance, so roughness reflects how θ is calibrated and interpolated at least as much as the skew shape. Clamp activity is likewise concentrated by name (Figure 6).

The round trip does not distinguish the six arms. Medians from the fitted surfaces run from 43.5 to 49.4 bp, with overlapping ranges and no systematic ordering, so the test serves as a consistency check on the shared Dupire implementation rather than as a measure of surface quality.

6.2.2 The martingale gate

The gate simulates each arm’s local volatility with the library’s own Monte Carlo, separate from the barrier engine, and asks whether the spot is a martingale under it: at each of the 1M, 3M and 6M horizons the simulated mean $\hat{\mathbb{E}}[S_T]$ must recover the parity forward F_T the surface was built in. The simulation is log-Euler with antithetic pairs, 32 768 paths and 104 steps per year, one seed and identical draws across arms. We report the absolute relative error $|\hat{\mathbb{E}}[S_T]/F_T - 1|$ and its z -score, $z = (\hat{\mathbb{E}}[S_T] - F_T)/\text{SE}$ with SE the Monte Carlo standard error of the simulated mean, so that $|z| > 3$ flags a cell whose miss exceeds what sampling noise explains at the 0.27% level.

All six local volatilities pass, and they pass identically. Across 2 736 daily and 67 968 intraday arm-horizon cells the median $|\mathbb{E}[S_T]/F_T - 1|$ is 1.39 and 1.37 bp respectively. Eleven cells exceed $|z| = 3$, all of them intraday and none further than 20 bp from the forward, which is fewer than the 0.27% normality alone would produce. Table 7 gives the per-arm detail, pooled across horizons.

6.3 The barrier price

Table 8 reports the pipeline’s final stage: how much the four eSSVI calibrations disagree on a six-month price, relative to that price, for the two designs of Section 4.3, on the index chains

Table 8: **Calibration disagreement on six-month prices.** $(\max - \min)/\text{mean}$ over the four eSSVI arms (S-eSSVI, the two G-eSSVI arms, sequential eSSVI), in percent of the price; median and 90th percentile over name-days. Barrier $H = F \exp(\pm \sigma \sqrt{T})$ in both panels, σ the market at-the-money vol; knock-outs priced by every arm on one shared path set, the vanilla at the same strike in closed form at each arm’s fitted vol. *Premium* is the median mean price, in percent of spot. *Index* is SPY, QQQ and IWM; DIA is in Table 9.

	Index			Single names		
	Premium (% spot)	Median (%)	p90 (%)	Premium (% spot)	Median (%)	p90 (%)
<i>Panel A: at the money, $K = F$</i>						
Vanilla (call = put)	5.44	3.5	6.4	12.20	1.4	3.2
Up-and-out call	1.99	3.1	5.1	2.36	2.1	3.7
Down-and-out put	0.56	2.0	9.2	2.15	2.1	3.1
<i>Panel B: out of the money, $K = F \exp(\pm \frac{1}{2} \sigma \sqrt{T})$</i>						
Vanilla call	2.32	5.9	14.0	6.64	1.8	3.2
Up-and-out call	0.29	7.5	21.7	0.32	2.2	6.9
Vanilla put	3.19	3.6	5.4	5.69	1.7	2.6
Down-and-out put	0.06	5.4	21.1	0.25	2.3	10.5

Table 9: **The same contracts on DIA.** As Table 8. *bp* is the spread in basis points of spot; *Quoted θ* is the median spread between the two eSSVI arms that take θ from the quoted ATM variance (S-eSSVI and sequential eSSVI), which share the at-the-money level and differ only in the wings and between pillars.

	Premium (% spot)	Median (%)	p90 (%)	bp	Quoted θ (%)
<i>Panel A: at the money, $K = F$</i>					
Vanilla (call = put)	4.01	1.6	3.1	6.4	0.0
Up-and-out call	1.00	36.1	62.9	36.1	22.6
Down-and-out put	0.52	15.8	34.2	8.7	10.5
<i>Panel B: out of the money, $K = F \exp(\pm \frac{1}{2} \sigma \sqrt{T})$</i>					
Vanilla call	1.90	13.5	24.2	23.4	5.7
Up-and-out call	0.13	67.6	109.7	7.5	39.9
Vanilla put	2.25	4.1	7.0	9.2	1.5
Down-and-out put	0.06	16.9	34.3	1.0	14.4

and the single names. Table 9 gives DIA, and Table 15 in Appendix F the out-of-the-money contracts in basis points of spot, name by name and one choice at a time.

At the money, the calibrations agree to a few percent. Struck at the forward, the vanilla disagrees by 3.5% of its price on the index chains and 1.4% on the single names at the median, and by 6.4 and 3.2% at the 90th percentile over name-days. All of it is the joint arms’ re-levelling of θ , the same re-levelling behind their held-out bias (Section 6.1.2): the two arms that take θ from the quotes price the at-the-money vanilla identically. The knock-outs disagree by 2 to 3% at the median and 3 to 5% at the 90th percentile, except the index down-and-out put, whose 0.56% premium stretches its tail to 9%.

Out of the money, the index chains disagree more. Half a standard deviation into the wing, the index vanilla call disagrees by 5.9% and the knock-out beyond it by 7.5% at the median, and by 14 and 22% at the 90th percentile; on the put side the medians are 3.6 and 5.4%. The single names barely move, 1.7 to 2.3% at the median and at most 11% at the 90th percentile. The index knock-outs are cheap, 0.06 to 0.3% of spot, so in money these percentages are 0.3 to 2 basis points.

DIA is where the path matters. Its vanillas disagree like the other index chains, by 1.6 to 13.5%, but its knock-outs by 16 to 68% at the median and 34 to 110% at the 90th percentile (Table 9). Most of that survives between the two arms that share the quoted at-the-money level, which differ on DIA’s knock-outs by 10 to 40% against at most 6% on its vanillas; on every other chain the same pair differs by under 2% on every contract. That is the between-pillar crossing of Section 6.1.4 priced: where the interpolated surface is inadmissible, the local volatility differentiates it and the knock-out reads the result, while a vanilla, which sees only the terminal density, does not.

Three choices, one at a time. In basis points of spot (Table 15), on the out-of-the-money vanilla, enforcement and calibration mode cost the same, 12 bp, because both comparisons set a joint arm against a quoted- θ one; the two arms that agree on θ , greedy and penalty, differ by 0.6.

6.4 Two thin chains

Most of the guaranteed arms’ between-pillar crossings fall on DIA, which is also the thinnest index chain in the panel. To separate thinness from whatever else DIA has, we ran the full daily pipeline and the leave-one-out test on two thinner single names, JPM and XOM, the liquid names with the fewest surviving pillars in the cross-section of Appendix D. Table 10 places them beside DIA and SPY.

Thinness does not reproduce DIA. With fifteen pillars and a quarter of DIA’s quotes, the three guaranteed arms are clean between pillars on every capture of both names, project less often, and give local volatility as smooth as the panel’s single names. What they lack is DIA’s combination: their smallest ψ increment is four to six times DIA’s and their ρ steps are SPY-sized. A crossing needs a small increment and a large step across the same pair (Section 6.1.4), and DIA alone has both. Pillar count, quote count and the compression of the ATM term structure, which is stronger on JPM and XOM than on DIA, do not predict it. Figure 5 shows their smiles on the same capture as Figure 1: the arms coincide at six months and part only on the one-month slice, inside a wide bid–ask band.

Thinness breaks the penalty fit and degrades interpolation. The sequential arm crosses between pillars on 74% and 63% of captures, against 11% on SPY, with adjacent ρ steps of 0.45 to 0.64, as large as DIA’s. Out of sample every arm degrades: leave-one-out error is 1.6 to 2.0 times in-sample for the joint arms, 1.9 to 2.2 for the penalty fit and 2.4 to 3.2 for the benchmark, against 1.0 to 1.2 on SPY. In money the cost is small: the choice of arm moves the six-month knock-outs by under 1 bp of spot and the vanillas by under 4 (Table 10). Thinness costs admissibility and interpolation, not price.

7 Limitations

The results are subject to four limitations.

Table 10: **Two thin chains beside two from the panel (daily slice).** JPM and XOM are run through the identical pipeline as an extension and enter no pooled statistic. Panel A describes the chain and the pinned joint arm’s fit; Panel B the consequences for admissibility, local volatility and the six-month down-and-out put. Medians over the nineteen sessions.

Panel A: the chain, and the pinned joint arm

Name	Pillars	Quotes	$\min \Delta\psi/\psi$ (%)	$\max \Delta\rho $ joint	$\max \Delta\rho $ seq.	Fit bp	LOO/IS	ATM bias bp
<i>Extension</i>								
JPM	15	305	3.9	0.15	0.64	23.7	1.81	1.3
XOM	15	237	3.1	0.11	0.45	23.5	1.59	-6.0
<i>Panel</i>								
DIA	18	932	0.7	0.35	0.46	43.4	1.21	-8.6
SPY	30	4123	1.1	0.13	0.10	108.6	1.07	-86.0

Panel B: consequences

Name	Crossings (%)			Rough. $\times 10^4$		6M down-and-out put (bp)		
	G-eSSVI			joint	seq.	Enforce gap	Greedy gap	Menu disp.
	S-eSSVI	$\gamma=\frac{1}{2}$	seq.					
<i>Extension</i>								
JPM	0	0	74	1.16	1.42	0.1	0.1	0.3
XOM	0	0	63	0.98	1.49	0.4	0.1	0.4
<i>Panel</i>								
DIA	100	79	100	3.73	6.12	0.5	0.4	1.0
SPY	5	11	11	6.05	4.65	0.4	0.5	0.5

$\min \Delta\psi/\psi$ is the smallest relative increment of the fitted skew scale between adjacent pillars and $\max |\Delta\rho|$ the largest adjacent step in ρ , each per capture, for the pinned joint arm (joint) and the penalty fit (seq.). *Fit*, *LOO/IS* and *ATM bias* are as in Tables 2 and 3. *Crossings* is the share of name-days with a calendar crossing on the grid, as in Table 5. The three gaps are the enforcement gap, the greedy gap and the spread over the four eSSVI arms of Table 15, in bp of spot.

The comparison is within one family. All six arms are parametric members of the SVI family, so the paper ranks calibrations of that family and says nothing about how it compares with non-parametric smoothers, such as the arbitrage-free splines of Fengler (2009), which trade a closed-form slice for flexibility and induce a local volatility of different smoothness. Nor is there an unsmoothed baseline, so the value of smoothing itself is not priced.

One interpolation rule. Every between-pillar result, from the calendar crossings to the local volatility and the knock-out prices, rests on linear interpolation in t (Section 4.2). Other rules, such as total variance at fixed log-moneyness or a smooth spline in t , are not tested.

The sample. Four weeks, one regime, eight names, four correlated index ETFs and four large-cap single names, with two thin large-cap chains as an extension. Stressed markets and genuinely illiquid or small-cap chains lie outside it.

The backbone. Tying ψ to a power law is our restriction of Mingone (2022), which reduces the ψ ladder to two parameters. A free coefficient per slice has strictly more freedom, so the fit gaps we report bound the construction’s cost from above.

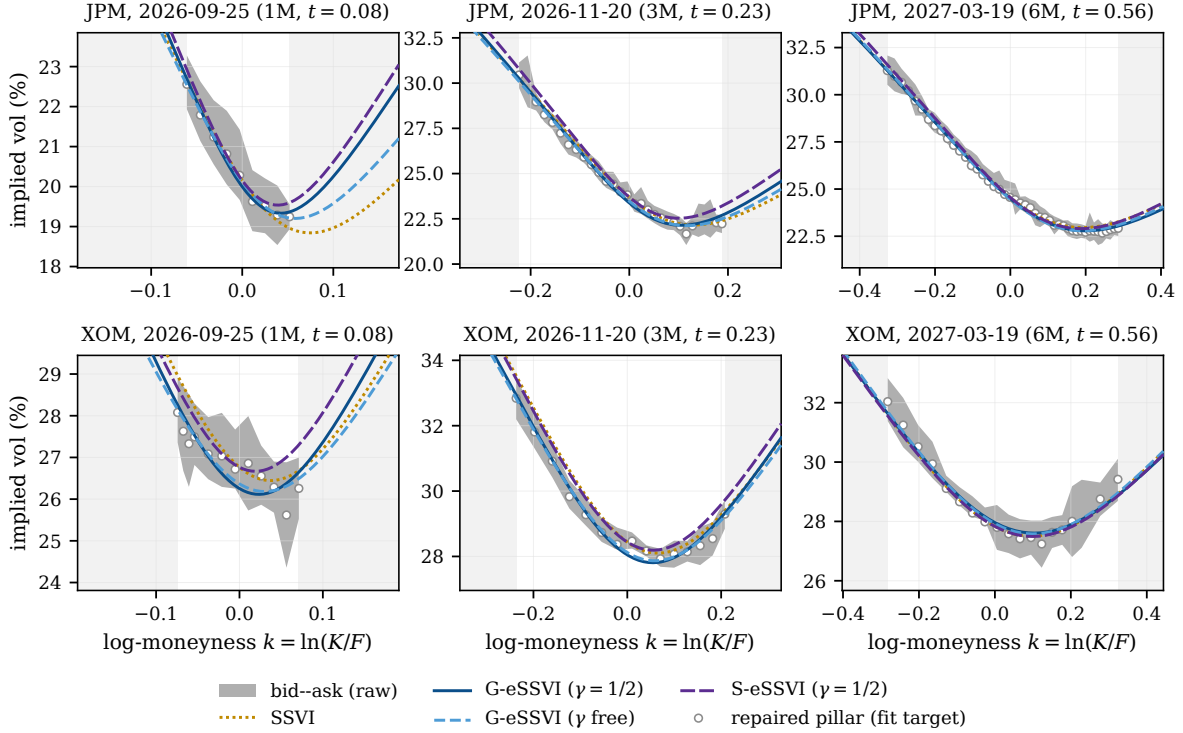


Figure 5: **Fitted smiles against the market: the two thin chains (daily slice, 2026-08-27).** As Figure 1.

8 Conclusion

The arbitrage guarantee of [Mingone \(2022\)](#) does not need a joint solve: a greedy forward pass with one fitted parameter per slice is free of butterfly and calendar arbitrage at every fitted pillar, as the joint calibrations are and the penalty calibration is not. The joint solve carries a cost of its own. Its fitted θ term structure smooths through the quoted at-the-money levels, so on dense chains it interpolates held-out expiries 40 and 52 bp below the market, where every arm that takes θ from the quotes stays within 3 bp. Fit does not reveal this, nor does it identify the better surface: raw SVI fits tightest in sample and degrades most out of sample, and the guaranteed arms still cross between pillars on 12 to 14% of name-days. Held-out expiries and arbitrage scans on a finer grid are what separate the surfaces, and SSVI, with its sufficient condition imposed, is the only one clean on the full grid.

The local volatility reverses the fit ranking. Raw SVI, the tightest fit, induces the roughest local volatility and the only one that hits the Dupire clamp; SSVI, the most restricted, induces the smoothest. Among the eSSVI arms roughness depends more on the name than on the calibration, because the shared linear-in- t interpolation puts a seam in $\partial_t w$ at every pillar and the index chains carry the most pillars. Neither the round trip nor the martingale gate separates the arms: both check the shared Dupire implementation rather than the surface.

Different fits of the same quotes give different prices, and a badly behaved surface amplifies the disagreement. Across the four eSSVI calibrations a six-month index option differs by up to 3.5% of its value at the money and 7.5% out of the money, against about 2% on single names.

On the vanillas most of it is the joint solve’s at-the-money bias again; the guarantee itself, at a common at-the-money level, barely moves the price. A knock-out adds little on a well-behaved surface, but on DIA, whose interpolated surface crosses between pillars, the same knock-outs differ by 16 to 68% of their price. Only the fit error misses DIA, which fits more tightly than any other index chain. The rest of the pipeline flags it stage by stage: its fitted exponent is the panel’s lowest, a sign of weak identification; the full-grid scan finds calendar crossings on 79 to 100% of its captures, where a small ψ increment meets a large ρ step between adjacent pillars; and the knock-out price shows what those crossings cost.

A Reproducibility

The repository

<https://github.com/siddharthqs/smile-smoothing-shootout>

holds the Rust driver, the Python scripts behind every table and figure, the daily option-chain captures and the daily result CSVs; its README gives the commands and settings. Every result except the wall clocks reproduces bit for bit with RustyQLib at commit ee2103d. The intraday captures and results are available from the author on request.

B The panel in detail

This is the full accounting behind Section 5.3: the cleaning cascade rule by rule (Table 11) and the fitting input it produced (Table 12). Two readings follow.

First, the largest single loss, the in-the-money side of every strike, is a *design choice*. The tenor panel of Table 12 shows where the fitted surface is actually anchored. Sub-week slices carry a median 36 pillar quotes on a span of barely $[-0.08, +0.04]$ in log-moneyness, while the 1–6M slices reach $k = -0.49$, where the moneyness window binds on the put wing, but only $k \approx +0.26$ on the call wing, where the spread cap removes the quotes.

C Supporting figures

Figure 6 reports what the Dupire clamp absorbed, arm by arm, and Figure 7 the day-to-day path of the fitted exponent across the whole panel.

D The fitted exponent over a wider cross-section

Table 13 extends Table 4 to sixteen further single names, six liquid large-capitalisation names and ten thinner ones, to ask where γ sits across a wider cross-section. It shows three things.

On liquid chains the exponent varies materially by name, from 0.41 on JPM to 0.64 on SPY, so the pin $\gamma = \frac{1}{2}$ restricts some surfaces much more than others.

The variation does not split by index against single name. Eight of the ten liquid single names fit above $\frac{1}{2}$, but JPM and XOM sit below it, near DIA, and the unconstrained SSVI refit places them lower still. The study panel’s four single names are all large-capitalisation growth names, so their shared preference for a skew decaying faster than $T^{-1/2}$ is not a property of single names in general; four weeks of data cannot say what drives it.

Table 11: **Quote attrition: where the raw chain goes.** Median quote counts per capture through the shared cleaning cascade, applied once upstream of all six arms. Rules in application order: unquoted, crossed or zero-bid (unq.); relative spread above 25% (spread); expiry outside 1 day–1.5 years (expiry); K/F outside $[0.6, 1.6]$ (mny.); the in-the-money side of each strike (ITM); and the residual rules (other: vol bounds, failed inversions, expiries left with fewer than five quotes). Failures are captures on which no expiry survives.

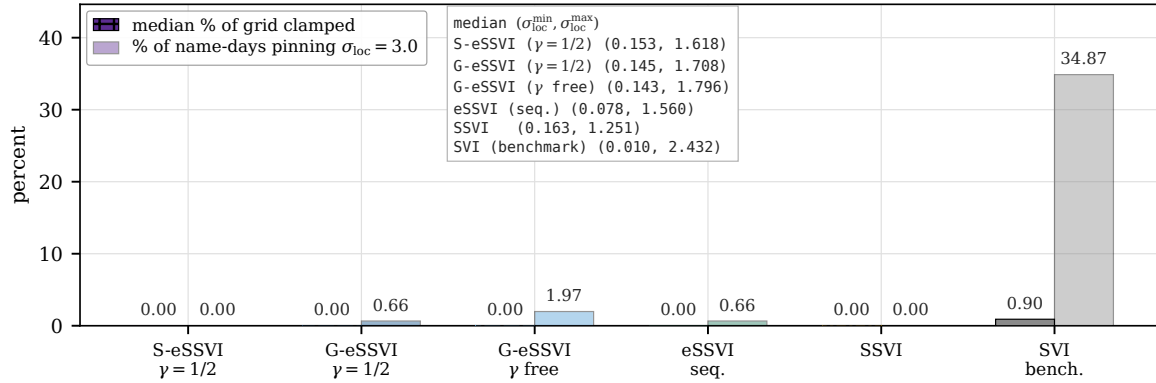
Ticker	Capt.	Fail	Raw	Median quotes dropped, by rule						Used	Kept %
				unq.	spread	expiry	mny.	ITM	other		
<i>Daily (15:30) slice</i>											
SPY	19	0	13 316	769	699	1 146	1 399	5 195	0	4 123	31.1
QQQ	19	0	11 588	790	1 051	858	712	4 783	0	3 418	29.6
IWM	19	0	5 442	547	581	186	261	2 361	0	1 506	27.7
DIA	19	0	4 870	428	982	305	130	2 032	0	932	19.2
META	19	0	6 836	704	848	730	1 506	1 858	0	1 173	17.2
AMD	19	0	6 100	630	741	626	1 663	1 522	0	1 056	16.9
TSLA	19	0	5 112	342	448	514	1 558	1 279	0	1 105	21.2
GOOGL	19	0	4 642	472	631	475	977	1 289	0	871	18.4
All daily	152	0	5 749	566	712	523	1 124	2 011	0	1 163	20.8
All intraday	3 877	101	5 744	561	761	523	1 049	1 999	0	1 131	20.4

Table 12: **What the fits received.** Top: the fitting input and the static-arbitrage repair per name (daily slice, medians per capture): expiry slices, pillar quotes per capture and per expiry, butterfly and calendar violations in the *raw* quotes, repair adjustments applied, and the largest single vol change the repair made (absolute implied vol; 0.01 is one volatility point). Bottom: the same pillar quotes pooled by tenor bucket.

<i>Per name (daily)</i>							
Ticker	Expiries	Quotes	Quotes/expiry	Raw bfly	Raw cal.	Repairs	Max Δ vol
SPY	30	4 123	141 [103, 189]	747	109	1 955	0.148
QQQ	30	3 418	120 [97, 143]	840	91	1 521	0.232
IWM	30	1 506	44 [38, 74]	213	38	433	0.086
DIA	18	932	52 [28, 73]	264	24	474	0.089
META	20	1 173	59 [42, 70]	258	13	312	0.077
AMD	21	1 056	49 [46, 55]	267	13	356	0.129
TSLA	20	1 105	59 [40, 69]	100	24	166	0.221
GOOGL	20	871	40 [23, 67]	191	11	257	0.101
<i>Per tenor bucket (daily, all names)</i>							
Bucket	Slices	Quotes/expiry		Median $[k_{lo}, k_{hi}]$			
$\leq 1W$	399	36 [23, 57]		[-0.08, 0.04]			
1W–1M	1 022	62 [40, 91]		[-0.22, 0.09]			
1–6M	1 201	74 [53, 111]		[-0.49, 0.26]			
$>6M$	982	58 [40, 109]		[-0.50, 0.45]			

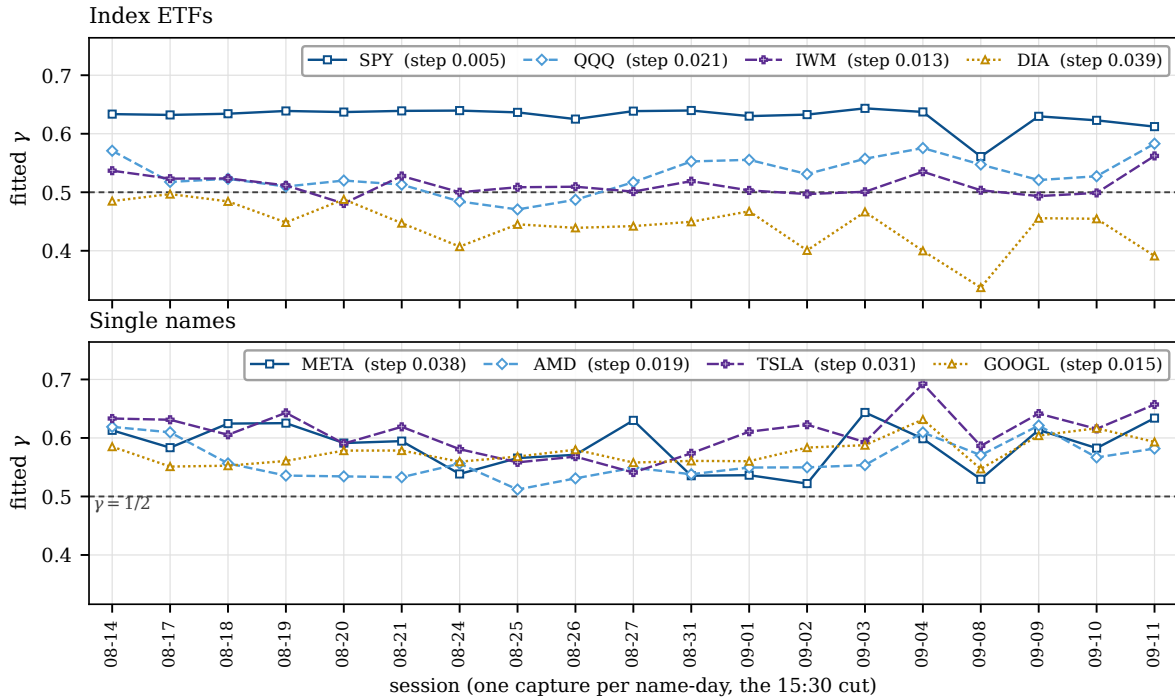
What the Dupire clamp absorbed

Daily slice, 152 name-days per arm. Clamp fractions are medians over name-days; the pinning bar counts name-days whose local volatility reaches the upper clamp anywhere on the grid.



Counts and medians only. $\min_g$, the worst butterfly density on the grid, is not plotted and must not be quoted as a headline: degenerate benchmark slices give it neither a usable scale nor a meaningful mean.

Figure 6: **What the Dupire clamp absorbed (daily slice).** Median share of grid cells the $[0.01, 3.0]$ clamp touched, and the share of name-days whose local volatility reached the upper bound anywhere on the grid. The inset gives each arm's median local-volatility extremes.



Daily slice, 19 sessions, G-eSSVI with the exponent fitted freely. "step" is the median absolute day-over-day move in γ .

Figure 7: **The fitted exponent session by session, all eight names (daily slice).** One capture per name-day, the G-eSSVI backbone exponent fitted freely; index ETFs above, single names below, on a shared axis. The dashed line is the pin. Each legend entry gives that name's median absolute day-over-day move in γ .

Table 13: **The fitted exponent over a wider cross-section.** As Table 4, extended to sixteen further single names fitted for the exponent only, with identical ingest. n is the number of intraday captures that survive the quote filters and fit; names with fewer than 100 (CRI, OLN, THO, WEN) are omitted. JPM and XOM also run through the full pipeline in Section 6.4.

Ticker	n	G-eSSVI (γ free)			SSVI, exponent unconstrained		
		Med.	IQR	$> \frac{1}{2}$ (%)	Med.	IQR	$> \frac{1}{2}$ (%)
<i>Index ETFs (study panel)</i>							
SPY	466	0.636	[0.630, 0.639]	100	0.597	[0.594, 0.604]	100
QQQ	467	0.527	[0.508, 0.546]	81	0.597	[0.590, 0.615]	100
IWM	468	0.508	[0.499, 0.521]	72	0.604	[0.595, 0.611]	100
DIA	480	0.450	[0.422, 0.474]	1	0.851	[0.795, 0.874]	100
<i>Single names (study panel)</i>							
META	470	0.573	[0.549, 0.597]	100	0.485	[0.451, 0.513]	36
AMD	479	0.562	[0.541, 0.592]	97	0.539	[0.494, 0.579]	72
TSLA	474	0.612	[0.580, 0.627]	99	0.610	[0.596, 0.629]	99
GOOGL	472	0.570	[0.558, 0.592]	100	0.573	[0.543, 0.598]	94
<i>Liquid single names</i>							
AAPL	468	0.542	[0.529, 0.560]	99	0.560	[0.545, 0.574]	100
MSFT	468	0.539	[0.521, 0.553]	97	0.557	[0.531, 0.580]	91
AMZN	469	0.577	[0.565, 0.592]	100	0.621	[0.600, 0.641]	100
NVDA	469	0.632	[0.590, 0.671]	99	0.687	[0.631, 0.737]	100
JPM	485	0.411	[0.391, 0.441]	7	0.367	[0.336, 0.402]	0
XOM	478	0.445	[0.409, 0.474]	10	0.361	[0.315, 0.396]	3
<i>Thinner single names</i>							
CLX	478	0.145	[0.000, 0.331]	13	0.000	[0.000, 0.172]	5
DPZ	485	0.546	[0.295, 0.909]	55	0.271	[0.139, 0.618]	35
ETSY	485	0.365	[0.149, 0.546]	31	0.022	[0.000, 0.204]	8
FIVE	481	0.637	[0.213, 0.831]	67	0.082	[0.000, 0.277]	16
KMX	485	0.433	[0.149, 0.884]	48	0.087	[0.021, 0.233]	6
WHR	485	0.174	[0.016, 0.409]	18	0.000	[0.000, 0.232]	10

On the thinner names the exponent is not identified. Those chains carry a median of five fitted pillars against twenty-one on the study names, their interquartile ranges span most of the unit interval, and the fits reach the edge of the admissible range on a fifth of captures for G-eSSVI and nearly half for the unconstrained SSVI refit. A power-law backbone needs a term structure to identify it, and five pillars are not one.

E Calibration cost

Table 14 and Figure 8 report single-threaded wall clocks on one laptop, which we read as an ordering: the multiples would survive a different machine, the milliseconds would not. Every iterative fit runs through one shared Levenberg–Marquardt implementation with forward-difference Jacobians, from the standard starting point of Section 4. Analytic Jacobians could cut the joint arms’ cost severalfold without changing the structure: N one- or two-parameter problems against one of $2N+1$ parameters.

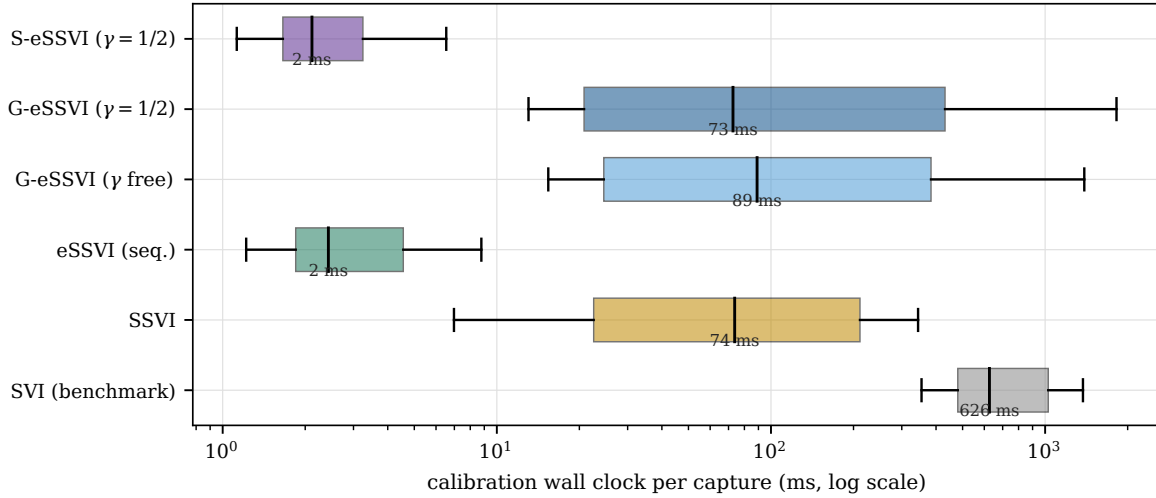
The greedy pass is the cheapest arm, slightly faster than the penalty fit it replaces. The joint solve costs about two orders of magnitude more on an index chain and about five times

Table 14: **What each calibration costs.** Wall-clock milliseconds per capture, median [IQR], for one independent fit from the standard starting point, single-threaded within a fit. Split by family because the joint problem scales with the expiry count.

Model	Index ETFs				Single names			
	Dim	ms	IQR	Iter.	Dim	ms	IQR	Iter.
<i>Daily slice</i>								
S-eSSVI ($\gamma=\frac{1}{2}$)	–	3	[2, 5]	142	–	2	[2, 2]	83
G-eSSVI ($\gamma=\frac{1}{2}$)	60	399	[202, 830]	36	41	23	[18, 27]	7
G-eSSVI (γ free)	61	398	[213, 592]	32	42	27	[21, 38]	8
eSSVI (seq.)	–	5	[4, 6]	–	–	2	[2, 2]	–
SSVI	–	84	[54, 198]	–	–	88	[15, 233]	–
SVI (benchmark)	–	1 058	[771, 1 286]	–	–	557	[501, 650]	–
<i>Intraday slice</i>								
S-eSSVI ($\gamma=\frac{1}{2}$)	–	3	[2, 6]	136	–	2	[1, 2]	80
G-eSSVI ($\gamma=\frac{1}{2}$)	59	424	[245, 881]	40	41	21	[17, 27]	7
G-eSSVI (γ free)	60	369	[207, 573]	33	42	25	[19, 34]	8
eSSVI (seq.)	–	5	[4, 6]	–	–	2	[2, 2]	–
SSVI	–	74	[40, 163]	–	–	74	[12, 224]	–
SVI (benchmark)	–	1 030	[683, 1 233]	–	–	500	[420, 605]	–

Dimension of the global problem: $2N+1$ parameters for N expiries with γ fixed, one more with γ free. Iterations are Levenberg–Marquardt steps.

Calibration cost per capture, in milliseconds



Intraday slice, 3776 captures per arm, both families pooled -- the joint problem scales with the expiry count, so the global arms' boxes straddle their two regimes: median 424 ms on the ~ 30 -expiry index chains (dim 59) against 21 ms on the single names (dim 41). Where the pin is misspecified the optimizer grinds: on SPY the pinned arm runs a median 38 iterations vs the free arm's 25.

Figure 8: **Calibration cost per capture.** Wall clock on the intraday slice (boxes: IQR; whiskers: 5th–95th percentile; log axis). The global arms' boxes straddle the index and single-name regimes, because the joint problem grows with the expiry count.

SSVI, a gap that narrows on the single names because the joint problem grows with the expiry count; the guarantee needs the backbone and the forward pass, not the joint solve's cost. Raw SVI is the slowest arm, since each of its per-slice fits runs a gridded two-dimensional search. The pinned joint arm needs more iterations than the free one on the index chains, despite one parameter fewer.

F Barrier prices by name and by choice

Table 15 gives the out-of-the-money design of Table 8 in basis points of spot, name by name, with the three one-variable gaps that Section 6.3 discusses and the Monte Carlo standard error of one price.

G The eSSVI construction in detail

This appendix states the calendar condition in its original and corrected forms, shows why Mingone's tube is never empty, and proves that the greedy pass is feasible.

The calendar condition. Take two slices of the form (16) at $t_1 < t_2$, with $\varphi_i = \psi_i/\theta_i$, and write

$$\Theta := \frac{\theta_2}{\theta_1}, \quad \Phi := \frac{\varphi_2}{\varphi_1}, \quad \Psi := \frac{\psi_2}{\psi_1} = \Theta \Phi. \quad (29)$$

The slices do not cross when $w_1(k) \leq w_2(k)$ for every $k \in \mathbb{R}$. Since $w_i(k) \sim \frac{1}{2}\psi_i(1 \pm \rho_i)|k|$ as $k \rightarrow \pm\infty$, the wings of slice 2 lie above those of slice 1 exactly when $\Psi(1 \pm \rho_2) \geq 1 \pm \rho_1$, that is $\Psi \geq p_2$ with p_2 as in (18); $p_2 \geq 1$, with equality only when $\rho_1 = \rho_2$.

The original proposition and its correction. Hendriks and Martini (2019) state that the slices do not cross if and only if

$$N': \quad \Theta \geq 1 \quad \text{and} \quad (\Psi\rho_2 - \rho_1)^2 \leq (\Psi - 1)^2, \quad (30)$$

$$S: \quad \Phi \leq 1 \quad \text{or} \quad (\Psi\rho_2 - \rho_1)^2 \leq (\Theta - 1)(\Theta\Phi^2 - 1). \quad (31)$$

Pasquazzi (2023) shows this wrong in two places. The squared inequality in N' reads $|\Psi\rho_2 - \rho_1| \leq |\Psi - 1|$, which for $\Psi < 1$ holds exactly when both wing slopes of slice 2 lie *below* those of slice 1. Such a pair has $\Phi = \Psi/\Theta < 1$, so it passes S by its first branch, yet it sits above slice 1 at the money and below it in the wings: it crosses. And N' admits $\Theta = 1$, where the characterization differs. The corrected statement splits by Θ . For $\Theta = 1$ the slices do not cross if and only if $\rho_1 = \rho_2 = 0$ and $\Phi \geq 1$, or $\Phi = \rho_1/\rho_2$ with $\rho_1^2 \geq \rho_2^2$, a tangency at the money. For $\Theta > 1$ they do not cross if and only if S holds together with

$$N: \quad 1 - \Psi \leq \Psi\rho_2 - \rho_1 \leq \Psi - 1, \quad \text{equivalently} \quad \Psi \geq p_2, \quad (32)$$

the wing condition unsquared, which keeps the sign of the comparison and forces $\Psi \geq 1$.

Table 15: **The out-of-the-money design of Table 8, by name and by choice.** Struck half an at-the-money standard deviation into the wing with the barrier a full one from the forward, $K = F \exp(\pm \frac{1}{2} \sigma \sqrt{T})$, $H = F \exp(\pm \sigma \sqrt{T})$, priced by every arm on one shared path set; the vanilla at the same strike is priced in closed form at each arm’s fitted vol. *Premium* is the mean price over the four eSSVI arms. *Spread* is max – min over the four eSSVI arms (S-eSSVI, the two G-eSSVI arms, sequential eSSVI); the three gaps are $|a - b|$ for one choice at a time. Everything in basis points of spot, medians over the nineteen sessions; MC σ is the median Monte Carlo standard error of one price. 400 000 antithetic paths per price, seed 20260824.

Name	Premium	Spread	Pin $\gamma=\frac{1}{2}$ vs free	Enforce joint vs pen.	Greedy vs joint	MC σ
<i>Panel A: down-and-out put</i>						
SPY	3	0.5	0.4	0.4	0.5	0.0
QQQ	6	0.3	0.1	0.3	0.2	0.1
IWM	6	0.2	0.0	0.2	0.1	0.1
DIA	6	1.0	0.1	0.5	0.4	0.1
META	24	0.4	0.3	0.1	0.1	0.2
AMD	35	0.9	0.4	0.8	0.5	0.2
TSLA	25	2.6	2.4	2.6	2.0	0.2
GOOGL	18	0.4	0.1	0.2	0.2	0.1
All	13	0.5	0.2	0.3	0.3	0.1
<i>Panel B: up-and-out call</i>						
SPY	29	3.3	2.5	2.9	3.3	0.1
QQQ	30	1.8	0.5	1.5	1.1	0.2
IWM	28	1.8	0.3	1.5	1.2	0.1
DIA	13	7.5	1.1	5.0	2.1	0.1
META	32	0.4	0.2	0.2	0.1	0.2
AMD	42	0.9	0.3	0.8	0.4	0.3
TSLA	33	2.2	1.9	2.0	1.5	0.3
GOOGL	27	0.6	0.3	0.5	0.2	0.2
All	30	1.4	0.5	1.4	0.9	0.2
<i>Panel C: vanilla call at the same strike, closed form at the fitted vol</i>						
SPY	153	20.5	12.3	19.9	20.5	–
QQQ	259	13.1	4.8	12.5	12.4	–
IWM	232	9.5	1.3	8.9	8.8	–
DIA	190	23.4	3.5	12.9	6.1	–
META	627	8.5	7.3	6.9	7.4	–
AMD	902	7.3	3.3	7.2	7.1	–
TSLA	696	20.4	13.3	19.7	20.4	–
GOOGL	492	10.9	5.6	10.1	10.9	–
All	377	12.5	5.7	11.6	11.5	–

What the construction uses. The tube of (19) takes $\Theta > 1$, the corrected N as its floor $\psi_1 p_2 \leq \psi_2$, and the first branch of S , $\Phi \leq 1$, as its calendar ceiling $\psi_2 \leq \psi_1 \Theta$: a sufficient subset of the sharp domain, the one Corbetta et al. (2019) and Mingone (2022) also use. The error does not touch it, because on $\Psi \geq p_2 \geq 1$ the original N' and the corrected N coincide. The dropped branch of S is one the backbone never needs: under (11), φ decreases in θ for every $\gamma \in (0, 1]$, so $\Phi \leq 1$ holds at every backbone target, and the ceiling binds only after an earlier projection has moved ψ off the backbone. The tangency cases of $\Theta = 1$ never arise, since θ strictly increases along the ladder.

Why the tube is never empty. In (20), $a_i > 0$ gives $\theta_i/\theta_{i-1} > p_i$, so the calendar ceiling $\psi_{i-1}\theta_i/\theta_{i-1}$ clears the floor $\psi_{i-1}p_i$ whatever ρ_i is. The backward cap does the same on the butterfly side: $\psi_{i-1} < S_{i-1} \leq S_i/p_i$, so $\psi_{i-1}p_i < S_i$. This is Mingone’s Proposition 3.1, which our implementation checks over 500 randomized parameter draws as a unit test.

Feasibility of the greedy pass. Without the backward cap, the greedy pass restricts ρ_i to the values at which slice i ’s tube is non-empty: $p_i < \theta_i/\theta_{i-1}$ on the calendar side and $\psi_{i-1}p_i < f_i(\rho_i)$ on the butterfly side, the second binding where dense short-dated listings make $f_i \sim \sqrt{4\theta_i}$ small. Both hold strictly at $\rho_i = \rho_{i-1}$, where $p_i = 1$, θ has strictly increased, and ψ_{i-1} sits below a butterfly cap that only grows with θ ; both tighten as ρ_i moves away from ρ_{i-1} . The feasible set is therefore an interval containing ρ_{i-1} , whose edges the pass finds by bisection outward. When two pillars carry near-equal ATM variance the interval collapses to a point, and the previous slice’s ρ is held.

References

- Douglas T. Breeden and Robert H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978.
- Mark Broadie, Paul Glasserman, and Steven Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7(4):325–349, 1997.
- Rama Cont and José da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60, 2002.
- Jacopo Corbetta, Pierre Cohort, Ismail Laachir, and Claude Martini. Robust calibration and arbitrage-free interpolation of SSVI slices. arXiv:1804.04924, 2019.
- Emanuel Derman and Iraj Kani. Riding on a smile. *Risk*, 7(2):32–39, 1994.
- Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.
- Matthias R. Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.
- Jim Gatheral. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. Presentation, Global Derivatives & Risk Management, Madrid, 2004.

- Jim Gatheral and Antoine Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.
- Sebastien Hendriks and Claude Martini. The extended SSVI volatility surface. *Journal of Computational Finance*, 22(5):25–39, 2019.
- Cristian Homescu. Implied volatility surface: Construction methodologies and characteristics. arXiv:1107.1834, 2011.
- Roger W. Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.
- Claude Martini and Arianna Mingone. Explicit no-arbitrage domain for sub-SVIs via death and resurrection of a Vogt smile. arXiv:2109.02418, 2022.
- Arianna Mingone. No arbitrage global parametrization for the eSSVI volatility surface. *Quantitative Finance*, 22(12):2205–2217, 2022. doi: 10.1080/14697688.2022.2117076.
- Leo Pasquazzi. A note about characterization of calendar spread arbitrage in eSSVI surfaces. *Theoretical Economics Letters*, 13(5):1341–1358, 2023. See also the companion calibration study, arXiv:2304.02106.
- Michael Roper. Arbitrage free implied volatility surfaces. Working paper, University of Sydney, School of Mathematics and Statistics, 2010.
- Zeliade Systems. Quasi-explicit calibration of Gatheral’s SVI model. Technical report, Zeliade White Paper, 2009.