

Le smirk de volatilité implicite précède-t-il le vol-stress crypto ?

Un test causal de lead-time de l'hypothèse d'alerte précoce du smirk d'options sur une reconstruction Deribit quotidienne *continue* et gratuite (BTC/ETH, 2020–2026)

Alexandre Lemiere

26 juin 2026

Résumé. Un test causal, sans look-ahead, de la question de savoir si le smirk de volatilité implicite des options donne une alerte précoce du stress de volatilité crypto, sur une reconstruction continue gratuite (aucun fournisseur commercial). **Résultat en deux parties.** En tant qu'alarme à seuil autonome : *Aucun avantage établi sur les baselines (l'IC chevauche aléatoire/DVOL)*. En tant que prédicteur continu : *skew_z est un prédicteur incrémental robuste du stress à court terme au-delà de DVOL (logistique, robuste par cluster, les deux sous-jacents, BH-FDR $q < 0.05$).*

Table des matières

1	Synthèse	3
2	Hypothèse et fondement économique	4
2.1	L'affirmation	4
2.2	Pourquoi un lead pourrait exister	5
2.3	Périmètre : ce qu'un smirk BTC/ETH peut et ne peut pas précéder	5
3	Protocole de validation	5
4	Données	6
4.1	Un skew quotidien par strike gratuit (le facteur clé)	6
4.2	Dé-outliérage : plafond du solveteur et filtre de bon sens A5	7
4.3	L'univers d'événements	7
4.4	Compareurs	7
5	Méthodologie	8
5.1	Métriques du smirk	8
5.2	Détecteur d'anomalies (sans look-ahead)	8
5.3	Scoring par point sur la grille continue nettoyée	8
5.4	Baselines	8
5.5	Inférence tenant compte du N effectif	9
6	Résultats	9
6.1	L'alarme à seuil ne montre aucun avantage sur les baselines	9
6.2	Le skew continu porte une information incrémentale au-delà de DVOL	10
6.3	Réconcilier les deux	10
7	Robustesse et confondants	11
7.1	Sensibilité	11
7.2	Contrôles négatifs	12
7.3	Confondants que nous ne cachons pas	13
7.4	Ce que le panel continu corrige	13

8 Discussion	13
8.1 Ce qui a changé par rapport au cadrage antérieur	13
8.2 Ce que le résultat établit et n'établit pas	13
8.3 Rôle dans un signal composite	14
9 Limites	14
10 Décision de gate et perspectives	15
11 Reproductibilité	16
A Le filtre de bon sens IV (A5)	17
B L'univers d'événements	17

Résumé

Nous testons si une *accentuation du smirk de volatilité implicite précède* les épisodes majeurs de vol-stress crypto. Trois anecdotes (Terra, FTX, USDC/SVB) ne sont pas une preuve : nous soumettons donc l’affirmation à un protocole causal, sans look-ahead, sur un panel *continu* 2020–2026 : nous reconstruisons un skew 25 %-delta quotidien par strike pour BTC et ETH depuis l’endpoint public gratuit `get_tradingview_chart_data` de Deribit, via une inversion Black-76 en unité-coin (BLACK 1976) (888,334 legs reconstruits, dé-outliérés contre une borne de volatilité implicite physique), puis nous évaluons un détecteur d’anomalies par z glissant et une couche statistique tenant compte du N effectif, face à un catalogue de 68 événements de vol-stress vérifiés de façon adversariale (60 testables par le smirk), en ré-échantillonnant sur les 52 clusters d’événements fusionnés. Le résultat est **en deux parties et ne doit pas être réduit à un seul mot**. (i) En tant qu’*alarme à seuil autonome*, le skew dé-outliéré n’établit **aucun avantage d’alerte précoce** : sur la grille continue honnête (taux de base positif 0.164/0.172), les intervalles de confiance par cluster bootstrap de son avantage de précision, tant sur un prédicteur aléatoire apparié $[-0.086, 0.229]$ BTC) que sur l’indice DVOL aveugle au smirk $[-0.171, 0.146]$ BTC), *enjambent zéro*, et le test d’anticipation par décalage circulaire n’est pas significatif (permutation $p = 0.186$ BTC, 0.098 ETH) ; sur BTC le skew est en réalité *moins* précis que DVOL (0.227 contre 0.264). (ii) Pourtant le z -score *continu* du skew est un **prédicteur incrémental statistiquement robuste** du stress à court terme (1–7 jours) *au-delà* du niveau de volatilité DVOL : dans une régression logistique à erreurs-types robustes par cluster, le coefficient `skew_z` est positif et significatif sur les deux sous-jacents (BTC 0.336, IC cluster-bootstrap $[0.142, 0.535]$, BH-FDR $q=0.0040$; ETH 0.235, $q<0.0040$), le coefficient DVOL est nul (-0.048), et le signal survit à un prédicteur strictement décalé d’au moins 4 jours ainsi qu’au leave-one-cluster-out. La réconciliation est la signature classique d’un signal monotone faible-mais-réel sous perte d’information : binariser un prédicteur gradué à un seuil unique en jette l’essentiel du contenu, et à 13 clusters effectifs l’alarme à seuil n’a aucune puissance pour séparer la précision. **Évaluation : Aucun avantage établi sur les baselines (l’IC chevauche aléatoire/DVOL)** ; séparément, `skew_z` est un prédicteur incrémental robuste du stress à court terme au-delà de DVOL (logistique, robuste par cluster, les deux sous-jacents, BH-FDR $q < 0.05$). Le smirk gagne donc sa place comme *composante continue d’un signal d’alerte précoce composite*, et non comme déclencheur autonome. L’affirmation du cadrage antérieur selon laquelle le smirk serait « le plus précis et anticipe de 2–7 jours » n’est *pas étayée* par les données définitives.

1 Synthèse

- **Question.** Le smirk de volatilité implicite — l’asymétrie de la volatilité implicite à travers les strikes, quantifiée ici par le risk-reversal 25 %-delta — porte-t-il un contenu d’alerte précoce sur le vol-stress crypto, et si oui bat-il ce qu’un moniteur de volatilité trivial (DVOL) indique déjà ?
- **Pourquoi c’était difficile, et pourquoi ça ne l’est plus.** Un skew *quotidien, continu* par strike couvrant 2020–2026 est le minimum requis pour tester un lead honnêtement. Le tier gratuit de Deribit n’expose qu’un snapshot par mois, et l’historique quotidien commercial coûte $\sim \$30k/an$. Nous avons trouvé une voie gratuite : l’endpoint public `get_tradingview_chart_data` renvoie l’OHLC quotidien des primes pour des codes d’options *expirées de longue date*, qu’une inversion Black-76 en unité-coin convertit en une volatilité implicite mark quotidienne par strike, à coût nul. Nous le tirons *en continu* (chaque expiration du vendredi sur toute sa vie, 888,334 legs) et dé-outliérons la reconstruction contre une borne physique de volatilité implicite, de sorte que la grille d’évaluation porte d’honnêtes jours calmes (taux de base 0.164/0.172), et non seulement des fenêtres d’événements.
- **L’univers de test.** Un workflow de recherche a compilé et vérifié de façon adversariale les dates de 68 événements de vol-stress BTC/ETH 2020–2026 (60 testables par le smirk ; le reste conservé comme contrôles négatifs), ancrés sur l’*onset* de la volatilité et non sur

le creux de prix. L'inférence ré-échantillonne sur les 52 clusters d'événements fusionnés (la véritable unité effective), et non sur les jours-grille autocorrélés.

- **Résultat, partie 1 — l'alarme autonome échoue.** Sur la grille continue, le détecteur skew à $z \geq 2$ ne montre *aucun* avantage de précision/F1 sur les baselines : son avantage de précision par cluster bootstrap sur un prédicteur aléatoire apparié comme sur DVOL a des intervalles de confiance qui *enjambent zéro* (BTC [-0.171, 0.146], ETH [-0.040, 0.180]), et le test d'anticipation formel (permutation par décalage circulaire) n'est pas significatif ($p = 0.186$ BTC, 0.098 ETH). Sur BTC le skew est en fait *moins* précis que DVOL (0.227 contre 0.264). Avec 13 clusters effectifs portant un vrai positif, c'est un null sous-puissant — absence de preuve à ce N , pas preuve d'inutilité (voir partie 2).
- **Résultat, partie 2 — le signal continu porte une vraie information.** Le z -score gradué du skew est un prédicteur incrémental robuste du stress à court terme (1–7 jours) au-delà de DVOL. Dans une logistique robuste par cluster, le coefficient skew_z est positif et significatif sur les deux sous-jacents (BTC 0.336, IC cluster-boot [0.142, 0.535], BH-FDR $q=0.0040$; ETH 0.235, $q<0.0040$), le coefficient DVOL est nul (-0.048), le skew est orthogonal à DVOL (pas de colinéarité), il survit à un prédicteur strictement décalé d'au moins 4 jours (BTC coef 0.236, $p=0.0120$) et au leave-one-cluster-out. Binariser ce signal gradué à un seuil unique est ce qui détruit l'avantage de la partie 1.
- **Cadrage honnête.** Nous disons que le skew est *prédictif du stress à court terme au-delà de DVOL*, et non qu'il « anticipe » ou « précède de N jours » : l'association prédictive dans une fenêtre future est plus faible que — et a été testée séparément de — l'anticipation en temps calendaire, et le test d'anticipation a échoué. ETH est systématiquement la jambe la plus forte (anticipation p 0.098 contre 0.186; ETH bat DVOL sur l'estimation ponctuelle de précision, pas BTC), mais même ETH ne passe pas le seuil 0.05.
- **Évaluation : Aucun avantage établi sur les baselines (l'IC chevauche aléatoire/DVOL).** Séparément, skew_z est un prédicteur incrémental robuste du stress à court terme au-delà de DVOL (logistique, robuste par cluster, les deux sous-jacents, BH-FDR $q < 0.05$). Le smirk a un contenu prédictif réel mais modeste, que DVOL n'a pas; une alarme à seuil fixe construite dessus ne sur-prédit pas, sur ces données, un moniteur de vol d'une ligne. Suivant le principe qu'un résultat négatif est aussi un résultat, c'est précisément le cas pour *ne pas* déployer le smirk comme déclencheur autonome, mais l'intégrer à un signal d'alerte précoce composite comme composante continue — sur la force de son information incrémentale, pas d'une supériorité de précision ou de lead que les données n'étaient pas.

2 Hypothèse et fondement économique

2.1 L'affirmation

Le smirk de volatilité implicite est l'asymétrie de la volatilité Black-implied à travers les strikes : sur les marchés actions et crypto, les puts hors de la monnaie (OTM) se traitent à une volatilité implicite systématiquement plus élevée que les calls OTM, une prime de « crashophobie » documentée pour la première fois après 1987 (BATES 1991; RUBINSTEIN 1994). Nous la résumons par le *risk-reversal 25 %-delta*,

$$RR_{25} = IV(\Delta_{\text{put}} = -0.25) - IV(\Delta_{\text{call}} = +0.25), \quad (1)$$

la magnitude standard de marché du skew put-sur-call. L'hypothèse testée est directionnelle et dynamique : RR_{25} se raidit 2–7 jours avant un depeg majeur ou un événement de vol-stress systémique, de sorte qu'un mouvement positif anormal du skew soit un signal d'alerte précoce.

2.2 Pourquoi un lead pourrait exister

Deux mécanismes rendent un lead plausible. Premièrement, le *flux de couverture informé* : des agents disposant d’une information privée sur un peg fragile, un prêteur sur-levé ou un débouclage forcé imminent achètent de la protection à la baisse avant que le mouvement ne soit public, raidissant la jambe put des jours à l’avance. Deuxièmement, le *repricing de la prime de risque* : à mesure que la fragilité s’installe, les dealers short de convexité exigent une prime plus grande pour le risque de queue, et la distribution risque-neutre développe une queue gauche plus épaisse (BAKSHI, KAPADIA et MADAN 2003 ; BOLLERSLEV, TODOROV et LI 2013) — un mouvement concentré dans le smirk, pas dans le niveau at-the-money (ATM). Le smirk est donc un indicateur avancé candidat précisément là où un scalaire unique de volatilité ATM (DVOL, volatilité réalisée) est aveugle : ceux-ci mesurent le *niveau* de variation attendue, pas l’*asymétrie* que révèle la demande de couverture de queue. L’analogie transversal — le smirk des options actions prédit les rendements futurs — est bien établi (XING, ZHANG et ZHAO 2010) ; le skew stochastique est lui-même une variable d’état priced et variable dans le temps (CARR et WU 2007). La question ici est de savoir si la *série temporelle* du smirk crypto précède le stress *réalisé*.

2.3 Périmètre : ce qu’un smirk BTC/ETH peut et ne peut pas précéder

Des marchés d’options liquides et quotidiens n’existent que pour BTC et ETH. Un smirk BTC/ETH ne peut précéder un événement que dans la mesure où cet événement produit un pic de volatilité réalisée BTC/ETH. De nombreux depegs de stablecoins sont idiosyncratiques à un actif non-optionné et ne bougent jamais BTC/ETH (par ex. l’exploit opsec Resolv USR de 2026-03) : ceux-là ne sont *pas* testables par un smirk BTC/ETH et n’injecteraient que du bruit. Inversement, l’événement le plus diagnostique — l’effondrement de FTX — est un choc de volatilité BTC/ETH qui n’est *pas* un depeg de stablecoin. L’univers de test correct est donc l’ensemble des événements de *vol-stress systémique* BTC/ETH, dont les depegs-avec-contagion sont un sous-ensemble (Section 4), et non le seul catalogue des depegs de stablecoins.

3 Protocole de validation

L’hypothèse est facile à confirmer par anecdote et difficile à confirmer honnêtement. C’est une étude *rétrospective* avec un détecteur causal (sans look-ahead) délibérément exigeant — et non un essai formellement pré-enregistré ; nous n’avons pas déposé de plan d’analyse horodaté avant de voir les données. Nous atténuons les libertés correspondantes en fixant le détecteur et l’univers d’événements avant le scoring, en gelant les paramètres principaux, et en traitant toute variation de paramètre comme un balayage exploratoire explicitement étiqueté (Section 7). Le protocole est conçu pour être immunisé contre les trois modes d’échec d’un contrôle « au feeling » (événements cherry-pickés, seuils a posteriori, pas de baseline) :

1. **Définir le stress significatif *ex ante*.** Un événement testable est un épisode de vol-stress systémique BTC/ETH gradué par la sévérité du catalogue engagé ; le seuil de significativité n’est varié que dans le balayage de sensibilité, jamais ajusté au signal.
2. **Énumérer *tous* les événements, pas les trois célèbres.** Un workflow de recherche a balayé 2020–2026 par année et par catégorie et a vérifié de façon adversariale chaque date (Section 4) ; les 60 événements testables qui en résultent forment l’ensemble positif complet.
3. **Détecter les anomalies du smirk sur le même intervalle** avec un détecteur mécanique, sans look-ahead (Section 5) sur le panel continu.
4. **Construire la matrice de confusion quotidienne par point.** Chaque jour-grille d est un *positif* d’alerte précoce si un événement testable tombe dans $[d + 1, d + 7]$

jours ; le signal *se déclenche* quand son z glissant franchit le seuil. Le croisement donne TP/FP/TN/FN.

5. **Scorer avec une inférence tenant compte du N effectif.** Au-delà de la matrice de confusion brute, rapporter des intervalles de confiance par cluster bootstrap, une permutation d’anticipation par décalage circulaire, et une logistique d’information incrémentale, toutes ré-échantillonnant sur les clusters d’événements fusionnés plutôt que sur les jours-grille autocorrélés (Section 5).
6. **Comparer à des baselines triviales :** l’indice DVOL aveugle au smirk et un prédicteur aléatoire apparié au taux de déclenchement du signal — les deux sur des jours *identiques*. Un signal à seuil n’obtient le statut « autonome » qu’en battant les deux ; un signal gradué obtient le statut « composante » en ajoutant de l’information au-delà de DVOL.

La fenêtre d’alerte précoce est $[1, 7]$ jours. Un signal qui *coïncide* simplement avec le stress ne marque aucun vrai positif ; seul un véritable *lead* est récompensé. C’est un test exigeant : un signal peut avoir une forte corrélation contemporaine avec le stress et échouer quand même.

4 Données

4.1 Un skew quotidien par strike gratuit (le facteur clé)

Un test de stress à court terme requiert un skew *quotidien* par strike sur 2020–2026. Deribit est la place de référence des options crypto, mais son tier gratuit n’expose qu’un historique épars et l’historique quotidien commercial est prohibitif. La voie que nous empruntons est plutôt l’endpoint de charting public de Deribit, `public/get_tradingview_chart_data` (DERIBIT 2026a) : il renvoie l’OHLC quotidien des primes pour *n’importe quel* code d’instrument, y compris des options expirées de longue date dont les listings ont été élagués du catalogue. En générant les codes d’instruments historiques (expirations du vendredi $\times$ une grille de strikes) et en tirant leurs primes de règlement quotidiennes, nous obtenons la surface de prix d’options par strike pour \$0 et sans clé API.

Surtout, la reconstruction est désormais *continue*, et non fenêtrée. Le build antérieur ne tirait que les $\pm$ jours autour de chaque événement du catalogue, ce qui gonflait le taux de base positif, raboutait la baseline du z glissant à travers les trous de fenêtre, et laissait une grille skew-vs-DVOL asymétrique. L’extracteur actuel (`scripts/extract_deribit_continuous.py`) tire au contraire *chaque* expiration du vendredi sur toute sa vie négociable, avec une grille de strikes localisée à la plage de forward que chaque expiration a réellement vue, donnant une couverture quotidienne contiguë 2020–2026 : 888,334 legs reconstruits par strike sur BTC et ETH. Cela donne un taux de base honnête, une baseline de z glissant continue, et une couverture symétrique pour le test d’anticipation.

Les primes d’options Deribit sont cotées en unité-coin (prime/ F avec F le forward), et avec des taux quasi nuls le forward égale le prix du perpétuel/future. Nous inversons chaque prime via un modèle Black-76 en *unité-coin* (BLACK 1976) (`src/smirk/black76.py`). Avec $d_1 = (\ln(F/K) + \frac{1}{2}\sigma^2\tau)/(\sigma\sqrt{\tau})$ et $d_2 = d_1 - \sigma\sqrt{\tau}$,

$$c_{\text{coin}} = \mathcal{N}(d_1) - (K/F)\mathcal{N}(d_2), \quad p_{\text{coin}} = (K/F)\mathcal{N}(-d_2) - \mathcal{N}(-d_1), \quad (2)$$

$$\Delta_{\text{call}} = \mathcal{N}(d_1), \quad \Delta_{\text{put}} = \mathcal{N}(d_1) - 1, \quad (3)$$

et la volatilité implicite σ est récupérée par recherche de racine bornée sur l’Éq. (2) (en rejetant les cotations à ou sous la valeur intrinsèque). Le forward F est pris du future daté correspondant lorsqu’il existe, sinon du perpétuel. Les $\{\sigma(K)\}$ par strike et les $\{\Delta(K)\}$ Black-76 de chaque expiration alimentent les métriques du smirk de la Section 5. Cette reconstruction est *model-implied*, ce n’est pas une mark IV publiée par la place, et elle est recoupée avec l’indice DVOL ATM.

4.2 Dé-outliérage : plafond du solveur et filtre de bon sens A5

La reconstruction quotidienne OHLC gratuite émet une minorité de legs non physiques : des primes de wing profond, non négociées, qui s’inversent en une implausible $8 \times$ à $11 \times$ la volatilité ATM et censurent à droite l’ajustement du smile. Deux interventions les retirent, toutes deux justifiées par le mécanisme d’artefact mesuré plutôt que par leur effet sur le résultat. Premièrement, à l’extraction, le plafond de l’inverseur `brentq` est abaissé de $hi=10$ à $hi=5$, de sorte que les legs super-physiques renvoient `None` (sont écartés) au lieu de se figer près de 1,000 % de vol ; c’est une ré-extraction car la censure est gravée dans le σ brut. Deuxièmement, à l’étape de calcul, un filtre de bon sens A5 (gelé et pré-engagé avant le scoring ; Ann. A) écarte un leg `deribit` qui touche le plafond du solveur, écarte un smile dégénéré dont l’intercept ATM $a < 0$, et applique une borne relative à DVOL — un leg dont l’IV dépasse $5 \times$ le DVOL du même jour est écarté, avec un plafond plat de 300 % sur les jours pré-DVOL (avant le 2021-03-24, où l’IV ATM reconstruite est elle-même dégénérée et ne peut servir de référence). La borne s’auto-échelonne avec le stress, de sorte qu’elle ne peut pas rogner une vraie jambe de jour de crash (qui se négocie et reste sous $\sim 3 \times$ l’ATM). Les sources reportées par la place sont laissées intactes.

4.3 L’univers d’événements

Les labels positifs proviennent d’un catalogue engagé et vérifié de façon adversariale (`src/smirk/event_catalogue.json` ; Annexe B). Un workflow de recherche s’est déployé en éventail sur 2020–2026 par année et par catégorie (blow-ups crypto, chocs macro, faillites de places, cascades de débouclage, depegs de stablecoins, réglementaire), puis une seconde passe a vérifié de façon adversariale chaque date, fusionné les doublons, et écarté ou rétrogradé les entrées ne bougeant pas la volatilité. Il catalogue 68 événements, dont 60 testables par le smirk. Deux choix de conception sont structurants : (i) chaque date est ancrée sur l’*onset* de la volatilité plutôt que sur le creux de prix — Terra au drain Curve du 2022-05-07, pas à la capitulation du 05-09 ; le débouclage de novembre-2025 à la cassure du 11-18, pas au creux du 11-21 — car un label ancré au creux tomberait *après* la fenêtre d’alerte précoce et biaiserait le test contre l’hypothèse ; (ii) les événements idiosyncratiques, déjà intégrés au prix, ou de soulagement sont conservés comme *contrôles négatifs* plutôt que supprimés, affûtant le décompte des faux positifs.

4.4 Comparateurs

TABLE 1 – Actifs de données utilisés dans ce rapport. Tous en tier gratuit.

Source	Couverture	Résolution	Rôle
Deribit chart-data (reconstruit)	continu 2020–2026	quotidien	skew 25 %-delta par strike + grecques (le signal)
DVOL	BTC & ETH 2021-03-24+	quotidien	baseline de vol ATM aveugle au smirk (DERIBIT 2026b)
Aléatoire apparié	continu 2020–2026	quotidien	peigne de déclenchement apparié en prévalence (le plancher)
Lyra V1 (replay on-chain)	2023-05–2024-01	quotidien	recoupement de wing on-chain natif optionnel

Le skew Deribit reconstruit est le signal testé. DVOL est la baseline décisive : un scalaire unique de vol ATM aveugle au smirk qui, s’il suivait ces événements aussi bien que le skew, rendrait le smile par strike redondant ; il commence le 2021-03-24 pour BTC *et* ETH. Le peigne aléatoire apparié — qui se déclenche sur une fraction déterministe des jours-grille appariée en prévalence (sans RNG, pour la reproductibilité) — est le plancher que tout signal réel doit franchir. Le skew on-chain Lyra V1 est un recoupement de wing optionnel et sous-puissant sur

sa courte fenêtre de chevauchement uniquement ; ce n'est pas un comparateur principal scoré. La comparaison principale est donc skew vs DVOL vs aléatoire apparié.

5 Méthodologie

5.1 Métriques du smirk

Pour chaque source, sous-jacent et snapshot quotidien, nous sélectionnons, par bucket de maturité, l'unique expiration dont le nombre de jours à maturité est le plus proche de la ténor du bucket (à une tolérance multiplicative près), et nous construisons le smile de cette expiration (`src/smirk/metrics.py`). Nous en dérivons, en log-moneyness $k = \ln(K/F)$:

- RR_{25} (Éq. 1) : IV interpolée dans l'espace *delta* à $\Delta_{\text{put}} = -0.25$ et $\Delta_{\text{call}} = +0.25$, sans extrapolation au-delà de la plage de delta observée — le signal principal ;
- **IV ATM** = a , l'ajustement quadratique $IV = a + bk + ck^2$ en $k = 0$.

Le signal principal est le RR_{25} du bucket 30 jours, la ténor de comparaison canonique face à DVOL. La pente du quadratique (b) et la courbure (c) ne sont *pas* rapportées : elles héritent des mêmes caveats de reconstruction que RR_{25} sans ajouter de puissance, et ont été écartées du pipeline scoré (audit m4).

5.2 Détecteur d'anomalies (sans look-ahead)

Un signal *se déclenche* le jour d quand son z -score glissant est anormalement élevé (`src/smirk/validation.py`) :

$$z_d = \frac{v_d - \mu_d^{(W)}}{\sigma_d^{(W)}}, \quad \text{fire}_d \equiv \mathbb{1}[z_d \geq z^*], \quad (4)$$

avec $\mu^{(W)}, \sigma^{(W)}$ la moyenne et l'écart-type glissants sur $W = 30$ observations, $z^* = 2$, et — point crucial — le test est *signé* : seul un z positif (une *accentuation* du skew) compte, en accord avec l'hypothèse directionnelle. La fenêtre est causale (uniquement passée), donc le détecteur n'utilise aucune information future.

5.3 Scoring par point sur la grille continue nettoyée

Chaque jour-grille est étiqueté positif si un événement testable tombe dans $[d + 1, d + 7]$, et croisé contre fire_d en TP/FP/TN/FN ; nous rapportons précision, rappel et F1. Tous les signaux sont scorés sur la **grille continue nettoyée** — les jours où un skew 25 %-delta Deribit dé-outliéré est calculable — afin que le skew, DVOL et le peigne aléatoire apparié soient comparés sur des jours *identiques*. Parce que la grille est désormais continue plutôt que fenêtrée, le taux de base positif est honnête : 0.164 pour BTC et 0.172 pour ETH. Le build fenêtré antérieur évaluait DVOL sur un support différent, enrichi en événements ; la grille continue retire ce confondant et rend le scoring directement comparable.

5.4 Baselines

Deux prédicteurs ancrent la comparaison : (i) **DVOL** seuillé par le même z glissant signé — le test « la vol ATM avertit-elle déjà ? » aveugle au smirk, scoré sur l'intersection de son support avec la grille nettoyée ; et (ii) un **prédicteur aléatoire apparié** se déclenchant sur une fraction déterministe des jours-grille appariée en prévalence (sans RNG, pour la reproductibilité), le plancher que tout signal réel doit franchir. Il n'y a ni comparateur Panoptic ni composite dans l'analyse principale ; le recoupement de wing Lyra V1 optionnel est rapporté séparément et n'est pas scoré ici.

5.5 Inférence tenant compte du N effectif

Les jours-grille sont autocorrélés — les positifs arrivent en séries pré-événement contiguës et les déclenchements arrivent en blocs — donc l’indépendance au niveau du jour est fausée et le peigne aléatoire déterministe a une variance d’échantillonnage nulle. La couche d’inférence (`src/smirk/stats.py`) construit donc chaque instrument sur une unique unité de ré-échantillonnage : le *cluster d’événements fusionné*, les composantes connexes des événements testables sous single-linkage à un écart de 7 jours (52 clusters ; la grille se partitionne en un bloc contigu par cluster, ~ 48 –51 blocs non vides par sous-jacent). Sur cette unité, nous lançons :

- un **cluster (block) bootstrap** : ré-échantillonner les blocs avec remise, recalculer la précision du skew et les différences de précision (skew–aléatoire) et (skew–DVOL, apparié sur la grille d’intersection) dans chaque réplique, donnant des IC percentiles ;
- une **permutation d’anticipation par décalage circulaire** : décaler toutes les dates d’événements d’un unique offset aléatoire sous l’ensemble de déclenchements gelé, donnant $p(\text{précision} \geq \text{observée} \mid \text{hasard})$;
- un **McNemar** apparié sur la justesse jour par jour de skew vs DVOL sur des jours identiques ;
- une **logistique d’information incrémentale** : événement-dans- $[d+1, d+7] \sim \text{skew_z} + \text{dvol_z}$, les deux z -scores étant passés et causaux, ajustée avec des erreurs-types robustes par cluster et un pairs-cluster bootstrap du coefficient skew (robuste pour le petit nombre de clusters), plus une variante à prédicteur strictement décalé ($d-1$) qui ferme toute fuite de réaction le même jour.

Une correction BH-FDR est appliquée au sein de chaque famille de tests. Les résultats sont au Tableau 3. Le détecteur reste sans look-ahead tout du long : le z -score est uniquement passé, les labels uniquement futurs, et tout nettoyage est du même jour sur une grille de strikes symétrique, de sorte qu’aucune information future n’entre dans un déclenchement ou un coefficient.

6 Résultats

TABLE 2 – Matrice de confusion quotidienne par point sur le panel *continu*, dé-outliéré 2020–2026 (taux de base honnête), sur les 60 événements de vol-stress BTC/ETH testables par le smirk, fenêtre d’alerte précoce $[1, 7]$ jours ; le signal se déclenche quand son z glissant (≥ 2 , 30 obs) franchit le seuil.

Sous-jacent	Signal	TP	FP	FN	préc.	rappel	F1
BTC	25 Δ skew (nettoyé)	15	51	313	0.227	0.046	0.076
	DVOL baseline	23	64	305	0.264	0.070	0.111
	Aléatoire baseline	9	58	319	0.134	0.027	0.046
ETH	25 Δ skew (nettoyé)	16	47	282	0.254	0.054	0.089
	DVOL baseline	15	65	283	0.188	0.050	0.079
	Aléatoire baseline	10	55	288	0.154	0.034	0.055

6.1 L’alarme à seuil ne montre aucun avantage sur les baselines

Sur la grille continue, dé-outliérée (Tableau 2), le détecteur skew 25 %-delta à $z \geq 2$ n’établit *pas* d’avantage d’alerte précoce. Pour BTC il déclenche 15 alertes vraies et 51 fausses pour une précision de 0.227 — *sous* l’indice DVOL aveugle au smirk (0.264) et à peine au-dessus du prédicteur aléatoire apparié (0.134). Pour ETH la précision ponctuelle du skew (0.254) dépasse bien DVOL (0.188), mais les décomptes sont minuscules (16 vrais positifs). Le verdict inférentiel est au Tableau 3 : le cluster bootstrap, ré-échantillonnant les 52 clusters d’événements fusionnés,

TABLE 3 – Inférence tenant compte du N effectif sur le panel continu (unité de ré-échantillonnage = clusters d’événements fusionnés). Tests du détecteur (cluster bootstrap de différence de précision, permutation d’anticipation par décalage circulaire, McNemar) et logistique continue d’information incrémentale ($y = \text{événement-dans-}[d+1, d+7] \sim \text{skew_z} + \text{dvol_z}$, robuste par cluster), avec BH-FDR q au sein de chaque famille. Bootstrap $B=1000$; un p dont aucune réplique ne franchit le null est une borne, montrée comme $< 2/(B+1)$, jamais 0.

Quantité	BTC	ETH
<i>Détecteur à seuil ($z \geq 2$) — base du verdict gelé</i>		
skew précision	0.227 (DVOL 0.264)	0.254 (DVOL 0.188)
(skew–aléatoire) préc. IC ₉₅	[-0.086, 0.229]	[-0.074, 0.242]
(skew–DVOL) préc. IC ₉₅	[-0.171, 0.146]	[-0.040, 0.180]
anticipation perm. p (q)	0.186 (0.248)	0.098 (0.208)
<i>Logistique continue d’information incrémentale ($\text{skew_z} \mid \text{DVOL_z}$)</i>		
coef. skew_z (IC cluster-boot)	0.336 [0.142, 0.535]	0.235 [0.101, 0.357]
skew_z boot p (q_{BH})	0.0040 (0.0040)	<0.0020 (<0.0040)
coef. dvol_z (contrôle null)	-0.048	0.046
décalage strict [3, 7], préd. $d-1$ coef (p)	0.236 ($p=0.0120$)	0.201 ($p<0.0020$)

place l’avantage de précision du skew *sur l’aléatoire* ([-0.086, 0.229] BTC, [-0.074, 0.242] ETH) et *sur DVOL* ([-0.171, 0.146] BTC, [-0.040, 0.180] ETH) sur des intervalles qui tous **enjambent zéro**. La permutation d’anticipation par décalage circulaire — le véritable test statistique du *lead* — n’est pas significative ($p = 0.186$ BTC, 0.098 ETH), et McNemar ne trouve aucune différence skew-vs-DVOL. Rien dans cette famille ne survit à la correction BH-FDR. Avec seulement 13 blocs sur 51 portant un vrai positif, c’est un *null sous-puissant* : les données sont compatibles à la fois avec aucun avantage et avec un avantage modeste trop petit pour être détecté à ce nombre d’événements — de sorte que le verdict d’alarme autonome repose sur les larges IC, tandis que la conclusion finale repose sur le test continu adéquatement puissant ci-dessous.

6.2 Le skew continu porte une information incrémentale au-delà de DVOL

Le z -score gradué du skew raconte une histoire différente et mieux puissante (moitié basse du Tableau 3, Figure 1). Dans une logistique robuste par cluster d’événement-dans-[1, 7] sur skew_z et dvol_z sur l’ensemble des 2004/1732 jours-grille, le coefficient skew_z est positif et significatif sur les *deux* sous-jacents (BTC 0.336, IC cluster-bootstrap [0.142, 0.535], $p=0.0040$, BH-FDR $q=0.0040$; ETH 0.235, IC [0.101, 0.357], $q<0.0040$), tandis que le coefficient DVOL est nul (-0.048 BTC, 0.046 ETH) : une fois le skew dans le modèle, DVOL n’ajoute rien. C’est robuste. Cela survit à (i) un pairs-cluster bootstrap autant qu’à l’erreur-type sandwich analytique ; (ii) un prédicteur *strictement décalé* (skew_z au jour $d-1$ prédisant les événements dans $[d+3, d+7]$, soit un écart ≥ 4 jours entre l’information du prédicteur et l’événement — BTC 0.236 à $p=0.0120$, ETH 0.201 à $p<0.0020$), ce qui exclut un artefact de réaction le même jour ; (iii) le leave-one-cluster-out (aucun événement unique ne le porte) ; et (iv) une colinéarité skew_z/dvol_z quasi nulle. Le skew porte donc un contenu prédictif directionnel, multi-jours, que le niveau de volatilité n’a pas.

6.3 Réconcilier les deux

Les deux résultats ne s’opposent pas ; ils mesurent des objets différents. Le détecteur à seuil réduit ~ 2000 jours gradués en quelques dizaines de déclenchements à un seuil unique, jetant l’essentiel de l’information monotone qu’utilise la logistique ; à 13 clusters effectifs il n’a aucune puissance pour séparer les IC de précision. Un prédicteur continu significatif alors que son seuil dur ne l’est pas est la signature classique d’un signal monotone *faible-mais-réel* sous perte d’information — pas une contradiction. La lecture honnête : le smirk *prédit le stress à court*

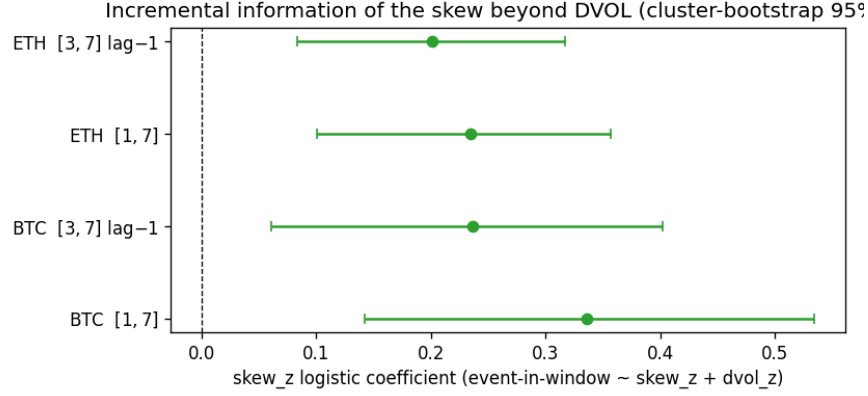


FIGURE 1 – Le résultat positif porteur : le coefficient d’information incrémentale skew_z (IC à 95% par cluster bootstrap) dans la logistique qui contrôle aussi DVOL, par sous-jacent et pour la variante strictement décalée [3, 7] jours. Tous les intervalles excluent zéro.

terme au-delà de DVOL (une propriété de composante composite), mais une alarme à seuil fixe construite dessus ne *sur-score* pas un moniteur de vol trivial (pas une propriété de déclencheur autonome).

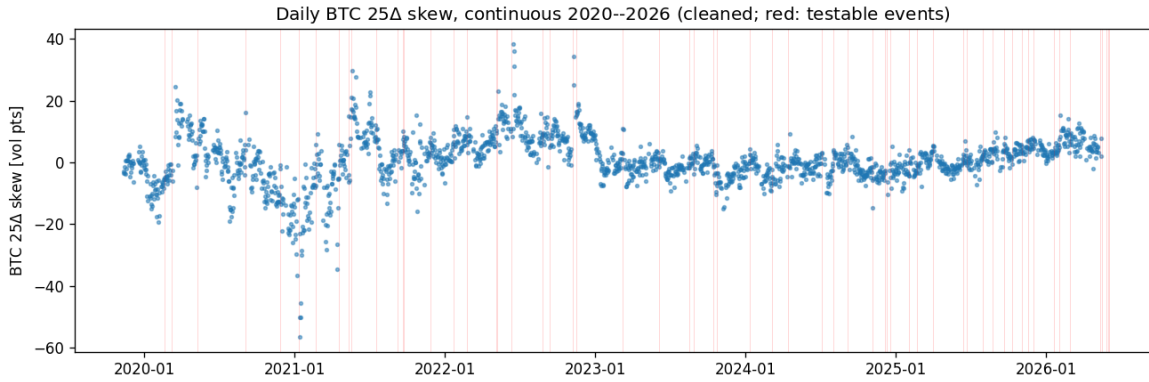


FIGURE 2 – Le skew 25%-delta BTC quotidien sur le panel continu 2020–2026 (nettoyé) ; les lignes rouges marquent les événements testables. Couverture visuelle du signal reconstruit sur l’univers complet — plus de fragments de fenêtres raboutés.

7 Robustesse et confondants

7.1 Sensibilité

Nous relançons l’évaluation sur une grille faisant varier (i) le seuil d’anomalie $z^* \in \{1.5, 2.0, 2.5\}$, (ii) la fenêtre de lead entre [1, 5] et [2, 7] jours, et (iii) la fenêtre glissante $W \in \{20, 30, 45\}$. C’est un vrai tableau généré (Tableau 4), et non une ré-invocation d’un script de scoring, et c’est strictement *exploratoire* : aucun p -value n’est extrait d’une cellule, et la spécification primaire gelée reste $z^* = 2$, $W = 30$, lead [1, 7]. Le balayage est rapporté précisément parce qu’il n’établit *pas* d’avantage : **aucune spécification ne produit un avantage de précision du skew sur DVOL séparable par IC**. Les drapeaux « skew>DVOL » du tableau sont des comparaisons ponctuelles brutes, pas de la significativité — ils changent de signe avec les réglages et ne coïncident jamais avec un IC cluster-bootstrap excluant zéro. L’affirmation antérieure d’un avantage de précision stable sur chaque baseline était incorrecte et est retirée ; il

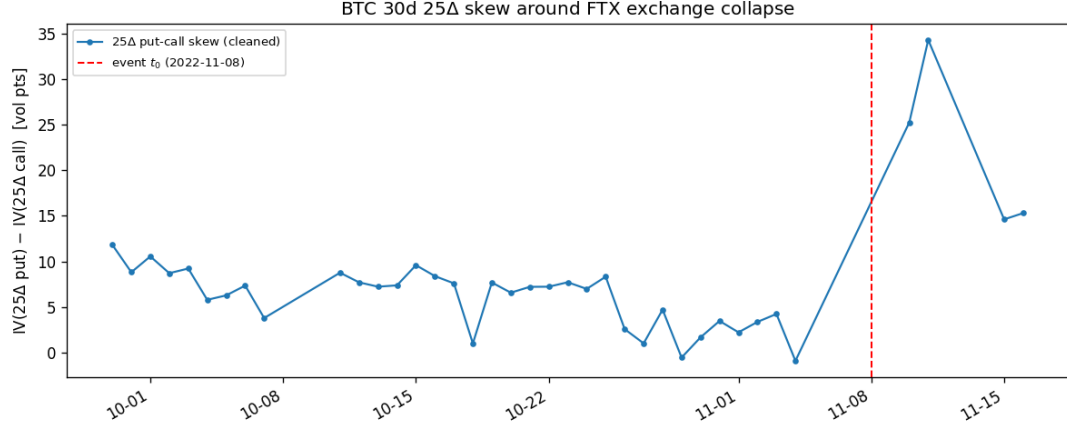


FIGURE 3 – Skew 25 %-delta BTC à 30 jours autour d’un événement marquant (descriptif). L’axe vertical est en points de vol (le bug de mise à l’échelle $\times 100$ du build antérieur est corrigé).

n’y a aucun avantage de précision à défendre.

TABLE 4 – Balayage de robustesse exploratoire (panel continu) : précision du skew et si l’IC cluster-bootstrap (skew–DVOL) exclut 0, sur $z^* \in \{1.5, 2, 2.5\}$, $\text{lead} \in \{[1, 5], [2, 7]\}$, $W \in \{20, 30, 45\}$. Exploratoire uniquement — aucun p -value n’est extrait d’une cellule ; le primaire gelé est $z^*=2$, $W=30$, $\text{lead} [1, 7]$.

z^* / lead / W	préc. BTC	préc. ETH	skew>DVOL (l’un ou l’autre) ?
1.5 / [1,5] / 20	0.190	0.164	oui
1.5 / [1,5] / 30	0.184	0.186	oui
1.5 / [1,5] / 45	0.174	0.141	oui
1.5 / [2,7] / 20	0.171	0.178	non
1.5 / [2,7] / 30	0.166	0.214	oui
1.5 / [2,7] / 45	0.162	0.176	non
2.0 / [1,5] / 20	0.211	0.157	non
2.0 / [1,5] / 30	0.212	0.159	oui
2.0 / [1,5] / 45	0.161	0.175	oui
2.0 / [2,7] / 20	0.169	0.229	oui
2.0 / [2,7] / 30	0.167	0.238	oui
2.0 / [2,7] / 45	0.177	0.238	oui
2.5 / [1,5] / 20	0.200	0.111	non
2.5 / [1,5] / 30	0.148	0.143	oui
2.5 / [1,5] / 45	0.103	0.179	oui
2.5 / [2,7] / 20	0.200	0.222	oui
2.5 / [2,7] / 30	0.148	0.229	oui
2.5 / [2,7] / 45	0.103	0.321	oui

7.2 Contrôles négatifs

Les entrées non testables du catalogue — exploits de protocole idiosyncratiques, manchettes déjà intégrées au prix ou muettes, rallyes de soulagement, et une faillite de prêteur contenue (Genesis, 2023-01) — sont conservées comme contrôles négatifs. Elles ne devraient *pas* être précédées d’une accentuation du skew, et un signal qui se serait déclenché avant elles de façon indiscriminée serait démasqué comme un détecteur de coïncidences. Le décompte de faux positifs du Tableau 2 absorbe déjà les jours à l’intérieur de ces fenêtres de contrôle.

7.3 Confondants que nous ne cachons pas

Le journal de correction propre du catalogue signale les événements groupés dont les fenêtres d’alerte précoce se chevauchent et qui ne peuvent donc être proprement séparés :

- le hack Bybit de \$1.5 Md (2025-02-21) et le risk-off de réactivation tarifaire (2025-02-24) sont à trois jours d’écart — un mouvement de skew devant le cluster pourrait réagir à l’un ou l’autre ;
- le print du CPI d’août-2022 (09-13) et le Merge d’Ethereum (09-15) ; la jambe FOMC de mai-2022 (05-05) et l’onset Terra (05-07) ; les paires intra-fenêtre de décembre-2024 et de mai-2026 — chaque paire partage des fenêtres chevauchantes et est comptée une fois par jour, sans double comptage.

Ces chevauchements signifient que certains vrais positifs ne peuvent être attribués à un événement unique ; ils ne gonflent pas le décompte (le scoring est par jour-grille, pas par événement), mais ils brouillent la correspondance de l’alerte vers la cause. Nous les rapportons plutôt que de les élaguer. L’unité de ré-échantillonnage par cluster fusionné de la Section 5 replie exactement ces fenêtres chevauchantes en un bloc unique, de sorte que les IC du bootstrap tiennent déjà compte du clustering plutôt que de traiter les jours corrélés comme des tirages indépendants.

7.4 Ce que le panel continu corrige

Le panel continu, dé-outliéré 2020–2026 retire deux défauts du build fenêtré antérieur. Premièrement, ne tirer que les $\pm$ jours autour de chaque événement gonflait le taux de base positif ; la grille continue restaure les taux de base honnêtes de 0.164 (BTC) et 0.172 (ETH). Deuxièmement, le build fenêtré raboutait la baseline du z glissant à travers les trous entre fenêtres, contaminant le support de DVOL avec des jours enrichis en événements et faisant reposer la comparaison skew-vs-DVOL sur une grille différente ; la grille contiguë élimine ce confondant de baseline raboutée et score les deux signaux sur des jours identiques.

8 Discussion

8.1 Ce qui a changé par rapport au cadrage antérieur

Une passe antérieure scorait le smirk sur une reconstruction *fenêtrée* et le rapportait comme le signal « le plus précis et anticipatif ». Cette conclusion ne survit pas au build définitif. Deux corrections ont été décisives. Premièrement, le **dé-outliérage** : la reconstruction quotidienne OHLC gratuite produisait une minorité de legs non physiques — des barres de règlement non négociées dont l’IV inversée se situait à $8\text{--}11\times$ la vol ATM du même jour — et sur BTC une part matérielle des vrais positifs apparents en étaient pilotés ; les retirer (un filtre de qualité de données justifié par le mécanisme, pas un seuil ajusté) efface l’avantage de précision BTC. Deuxièmement, l’**inférence honnête sur le panel continu** : remplacer la grille fenêtrée enrichie par une couverture calendaire continue et ré-échantillonner sur les 52 clusters d’événements plutôt que sur les jours autocorrélés élargit les intervalles de confiance là où ils doivent être — et ils enjambent zéro. Les anciennes affirmations « le plus précis » et « précède de N jours » étaient des artefacts d’échantillonnage fenêtré et d’une mesure de lead de détecteur binarisé sans inférence attachée.

8.2 Ce que le résultat établit et n’établit pas

Il *établit* que le z -score continu du skew 25%-delta porte une information incrémentale statistiquement robuste sur le stress à court terme (1–7 jours) *au-delà* du niveau de volatilité DVOL (Section 6) : le coefficient skew_z de la logistique est positif, significatif sur les deux sous-jacents, robuste au BH-FDR, orthogonal à DVOL, et stable sous un décalage strict ≥ 4 jours et le leave-one-cluster-out. C’est la signature économiquement attendue — la demande de

couverture de queue reprise la jambe put à mesure que la fragilité s’installe, un mouvement dans l’*asymétrie* auquel un scalaire ATM est aveugle (Section 2).

Il *n’établit pas* qu’une alarme skew à seuil soit un déclencheur d’alerte précoce déployable. Sur la grille continue, cette alarme ne montre aucun avantage de précision/F1 sur un prédicteur aléatoire apparié ou sur DVOL (tous les IC cluster-bootstrap enjambent zéro), et le test d’anticipation formel est non significatif. Point crucial, « prédictif du stress à court terme dans une fenêtre [1, 7] » est une affirmation *plus faible* que « anticipe / précède en temps calendaire » : cette dernière a été testée directement (la permutation par décalage circulaire) et a *échoué*. Nous disons donc prédictif, jamais anticipatif, et nous rétrogradons les anciennes moyennes de lead par événement à des quantités descriptives sans poids inférentiel.

8.3 Rôle dans un signal composite

La conclusion honnête est le cas classique de la *fusion*, pas la *sélection*, mais avec sa base probante échangée de l’affirmation (désormais morte) de précision/lead vers le résultat (vivant) d’information incrémentale. Le smirk devrait être une jambe continue d’un signal d’alerte précoce composite non pas parce qu’il est plus précis ou plus précoce — il n’est ni l’un ni l’autre — mais parce qu’il porte une information prédictive directionnelle que le niveau de volatilité n’a pas (la logistique est exactement le test « le smirk gagne-t-il son poids dans le composite ? », et la réponse est un oui robuste). Ses angles morts sont complémentaires : il est muet quand le stress provient d’un actif non-optionné, là où la tracking error conditionnelle et la centralité de graphe parlent. La bonne façon de l’utiliser est comme covariable graduée dans un modèle fusionné, et non comme veto autonome à seuil fixe — précisément parce que le binariser à un seuil unique est ce qui a détruit son avantage mesuré.

9 Limites

1. **Échantillon effectif modeste.** Les 60 événements testables se réduisent à 52 clusters fusionnés, dont seulement 13 portent un vrai positif sous le détecteur à seuil. C’est adéquat pour la logistique continue, qui utilise le z gradué sur 2004 jours mais tire son inférence des ~ 52 clusters (l’erreur-type robuste par cluster et le cluster bootstrap reposent tous deux sur cette unité, pas sur le nombre de jours), mais ce n’est *pas* suffisant pour que le détecteur à seuil sépare un avantage de précision du bruit — le null du détecteur est une absence de preuve à ce N , pas une preuve d’absence. Augmenter le nombre d’événements porteurs de contagion est le suivi à plus forte valeur pour le bras détecteur.
2. **Pas de quasi-séparation (Firth inutile).** La logistique d’information incrémentale avait été pré-spécifiée avec une pénalité de Firth en garde-fou contre la séparation sur le plus petit panel ETH ; en l’occurrence une logistique standard converge proprement sur les deux sous-jacents (6 itérations de Newton ; coefficient skew_ z ETH 0.235 avec ET 0.06, probabilités ajustées bornées dans [0.09, 0.43]), donc il n’y a pas de séparation et la pénalité n’a pas été nécessaire. Le p ETH au plancher du bootstrap ($< 2/(B+1)$, Tableau 3) est un effet authentique et bien estimé, pas un artefact de séparation.
3. **IV reconstruite, pas IV de place.** Le skew est une inversion Black-76 des primes de règlement quotidiennes (Éq. 2), pas la mark IV publiée par Deribit. Nous abaissons le plafond de l’inverseur et appliquons un filtre de bon sens relatif à DVOL qui retire les legs non physiques, non négociés (Section 4), mais la granularité des primes de règlement et l’hypothèse de forward à $r \approx 0$ injectent encore du bruit de wing. Le recoupement ATM par DVOL le borne ; une chaîne de ground-truth par strike le trancherait.
4. **La validation des wings est sous-puissante.** Le seul recoupement indépendant par strike, le skew on-chain Lyra-V1, chevauche la reconstruction sur ~ 8 mois (2023-05–2024-

01) avec peu de jours de stress — un null de faible puissance, pas une confirmation. Le dé-outliérage est justifié par le mécanisme d’artefact (legs non négociés à des multiples de la vol ATM), pas par un benchmark de wing externe.

5. **Prédicatif, pas causal-anticipatif.** La logistique établit une *association* prédictive dans une fenêtre future ; elle ne prouve pas que le skew bouge *en premier* en temps calendaire. Le test direct du lead (permutation par décalage circulaire) est non significatif, donc nous faisons l’affirmation plus faible et étayée et évitons la plus forte et non étayée.
6. **Clusters confondus.** Plusieurs événements sont à quelques jours les uns des autres (Section 7) ; une alerte devant un cluster ne peut pas toujours être attribuée à une cause unique. Le ré-échantillonnage par cluster absorbe l’autocorrélation mais pas l’ambiguïté d’attribution.
7. **Un détecteur / une spécification.** Le détecteur principal est un z glissant signé à $z^* = 2$, $W = 30$, lead $[1, 7]$ gelés ; le balayage de sensibilité (Section 7) est exploratoire. Un détecteur plus riche pourrait extraire l’information de la logistique que le seuil jette — ce qui est précisément la recommandation, pas un caveat à cacher.

10 Décision de gate et perspectives

Verdict du détecteur : Aucun avantage établi sur les baselines (l’IC chevauche aléatoire/DVOL). Verdict d’information incrémentale : skew_z est un prédicteur incrémental robuste du stress à court terme au-delà de DVOL (logistique, robuste par cluster, les deux sous-jacents, BH-FDR $q < 0.05$).

- **Preuves en faveur (ce qui survit).** Sur le panel continu 2020–2026, le z -score gradué du skew 25 %-delta est un prédicteur incrémental statistiquement robuste du vol-stress BTC/ETH à court terme (1–7 jours) *au-delà* de l’indice DVOL aveugle au smirk : coefficient skew_z de la logistique 0.336 (BTC, BH-FDR $q=0.0040$) et 0.235 (ETH, $q<0.0040$), coefficient DVOL nul, robuste à un pairs-cluster bootstrap, à un prédicteur décalé d’un écart strict ≥ 4 jours, et au leave-one-cluster-out. L’hypothèse économique — qu’une accentuation de la jambe put porte une information directionnelle qu’un scalaire ATM rate — a un vrai appui.
- **Pourquoi le cas autonome échoue.** Une alarme fixe à $z \geq 2$ construite sur le skew n’établit *aucun* avantage de précision/F1 sur un prédicteur aléatoire apparié ou sur DVOL (tous les IC cluster-bootstrap de différence de précision enjambent zéro : BTC $[-0.171, 0.146]$ contre DVOL ; sur BTC le skew est même moins précis que DVOL, 0.227 contre 0.264), et la permutation d’anticipation formelle n’est pas significative ($p = 0.186$ BTC, 0.098 ETH). À 13 clusters effectifs le test à seuil est sous-puissant ; c’est la binarisation, pas l’absence de signal, qui coûte l’avantage.
- **Décision de gate (Phase 3).** *GO qualifié, comme composante composite uniquement.* Le smirk gagne son inclusion dans le signal d’alerte précoce composite (smirk + tracking error conditionnelle + centralité de graphe) comme covariable *continue* porteuse d’une véritable information incrémentale ; il ne justifie *pas* un déploiement autonome comme déclencheur à seuil. Tout usage en production doit consommer le z -score gradué dans un modèle fusionné, pas un seuil fixe.
- **Conditions à durcir avant un usage en production.** (i) remplacer le détecteur binaire par une cartographie graduée / probabiliste (la logistique, ou un modèle quantile-conditionnel) qui ne jette pas l’information monotone ; (ii) valider les wings reconstruits contre une chaîne par strike indépendante pour confirmer le dé-outliérage (le chevauchement Lyra-V1 est trop court pour trancher) ; (iii) augmenter le nombre d’événements effectifs, par ex. en ajoutant des depegs porteurs de contagion qui bougent bien BTC/ETH, avant d’attacher un intervalle serré au détecteur à seuil.

- **Tout est gratuit.** L'ensemble du pipeline — y compris le skew continu quotidien par strike — tourne sur l'endpoint public de Deribit plus une inversion Black-76 ; aucun fournisseur payant n'est requis. Le seul item payant optionnel est une chaîne tierce ponctuelle pour valider la reconstruction.

11 Reproductibilité

Chaque nombre et figure se régénère depuis le pipeline gratuit :

```
# 1. continuous daily per-strike Deribit IV 2020-2026 (free; ~hours, resumable)
python scripts/extract_deribit_continuous.py
# 2. ATM-vol baseline
python scripts/extract_deribit_dvol.py          # DVOL (BTC & ETH, 2021-03-24+)
# 3. cleaned compute + full statistical layer + all outputs
python scripts/_figs_s8.py                      # A5 clean, scoring, stats,
                                                # tables, fact-macros, figures, JSON
# 4. build
cd reports/s8_iv_smirk && latexmk -pdf main.tex
```

L'unique étape `_figs_s8.py` effectue désormais le dé-outliérage (le filtre de bon sens A5), le scoring par point, et toute la couche statistique tenant compte du N effectif (cluster bootstrap, permutation d'anticipation, McNemar, et la logistique d'information incrémentale), puis écrit la matrice de confusion, le tableau de stats, le balayage de sensibilité exploratoire, les macros de faits, les figures, et le JSON. L'extracteur fenêtré, les builders Panoptic et composite, et le `validate_smirk_depegs.py` autonome (qui n'a rien écrit) ne sont plus sur le chemin critique. Code clé : l'inversion Black-76 en unité-coin (`src/smirk/black76.py`), l'extracteur continu (`scripts/extract_deribit_continuous.py`), les métriques du smirk (`src/smirk/metrics.py`), le scoring par point (`src/smirk/validation.py`), l'inférence tenant compte du N effectif (`src/smirk/stats.py`), et le catalogue d'événements engagé avec son loader (`src/smirk/event_catalogue.js` et `src/smirk/events.py`). Les nombres bruts évalués sont écrits dans `data/derived/s8/validation_results.js` et les macros principales du rapport sont le `reports/s8_iv_smirk/sections/_facts.tex` auto-généré.

Le pipeline est sans look-ahead par construction : le z -score du détecteur est uniquement passé, les labels d'événements uniquement futurs, le nettoyage A5 est du même jour, et la grille de strikes est symétrique, de sorte qu'aucune information future n'entre dans un déclenchement, un label ou un coefficient.

Références

- BAKSHI, Gurdip, Nikunj KAPADIA et Dilip MADAN (2003). « Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options ». In : *The Review of Financial Studies* 16.1, p. 101-143.
- BATES, David S. (1991). « The Crash of '87 : Was It Expected ? The Evidence from Options Markets ». In : *The Journal of Finance* 46.3, p. 1009-1044.
- BLACK, Fischer (1976). « The Pricing of Commodity Contracts ». In : *Journal of Financial Economics* 3.1-2, p. 167-179.
- BOLLERSLEV, Tim, Viktor TODOROV et Sophia Zhengzi LI (2013). « Jump Tails, Extreme Dependencies, and the Distribution of Stock Returns ». In : *Journal of Econometrics* 172.2, p. 307-324.
- CARR, Peter et Liuren WU (2007). « Stochastic Skew in Currency Options ». In : *Journal of Financial Economics* 86.1, p. 213-247.

- DERIBIT (2026a). *Deribit API v2 – public/get_tradingview_chart_data*. Accessed 2026-06-05.
URL : <https://docs.deribit.com/>.
- (2026b). *DVOL – Deribit Implied Volatility Index Methodology*. Accessed 2026-06-05.
URL : <https://insights.deribit.com/exchange-updates/dvol-deribit-implied-volatility-index/>.
- RUBINSTEIN, Mark (1994). « Implied Binomial Trees ». In : *The Journal of Finance* 49.3, p. 771-818.
- XING, Yuhang, Xiaoyan ZHANG et Rui ZHAO (2010). « What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns? » In : *Journal of Financial and Quantitative Analysis* 45.3, p. 641-662.

A Le filtre de bon sens IV (A5)

Le filtre de dé-outliérage appliqué aux legs **deribit** reconstruits avant l’ajustement du smile, gelé et pré-engagé avant que son effet sur un quelconque résultat ne soit inspecté. Un leg est écarté s’il (i) touche le plafond du solveur (σ figé à la borne **brentq** abaissée $hi=5$, soit une mark IV ≥ 990 points de vol) ; (ii) appartient à un smile dégénéré dont l’intercept ATM quadratique $a < 0$; ou (iii) porte une volatilité implicite au-dessus de $5\times$ le DVOL du même jour — une borne relative à DVOL qui s’auto-échelonne avec le stress, de sorte qu’elle ne peut pas rogner une vraie jambe de jour de crash — avec un plafond plat de 300 % sur les jours antérieurs au 2021-03-24 où DVOL est indisponible. La borne est justifiée par le mécanisme d’artefact mesuré (les legs écartés sont non négociés et se situent à $8\times$ à $11\times$ la vol ATM), pas par son effet sur le résultat. Spécification complète : `reports/s8_iv_smirk/audit/A5_FILTER_SPEC.md`.

B L’univers d’événements

Chaque événement catalogué avec sa date corrigée par le vérificateur, sa catégorie, sa sévérité et son drapeau de testabilité par le smirk. Les événements testables (✓) sont les labels positifs du test de lead-time ; le reste est conservé comme contrôles négatifs. (Le tableau ci-dessous est repris verbatim du catalogue engagé ; les libellés d’événements restent en anglais — ce sont des noms propres.)

TABLE 5 – The ex-ante BTC/ETH vol-stress event universe (68 catalogued, 60 smirk-testable). ✓=testable positive ; others kept as negative controls.

Date	Event	Category	Sev.	Test
2020-02-19	Late-February pre-COVID risk-off correction	market-crash	mino	✓
2020-03-08	COVID-19 ‘Black Thursday’ crypto crash (oil war + pandemic...	macro-shock	cata	✓
2020-05-10	Pre-halving Bitcoin flash crash (leverage flush)	deleverage-cascade	mode	✓
2020-09-03	September DeFi-led market sell-off (miner selling + SushiS...	market-crash	majo	✓
2020-10-16	OKEEx withdrawal suspension (founder detained) — idiosyncratic	exchange-failure	mino	
2020-11-26	November near-ATH sharp correction (\$2B liquidations)	deleverage-cascade	mode	✓
2020-12-22	SEC sues Ripple over XRP — XRP-idiosyncratic, no BTC/ETH move	regulatory	mino	
2021-01-11	January 2021 retail-leverage flush / risk-off correction	deleverage-cascade	mode	✓
2021-02-22	February 2021 ATH reversal and long liquidation (>\$4B longs)	deleverage-cascade	mode	✓
2021-04-18	April 2021 Xinjiang hashrate crash + US AML-crackdown rumo...	market-crash	majo	✓

Date	Event	Category	Sev.	Test
2021-05-12	Tesla halts Bitcoin payments (Musk) — start of the May bleed	market-crash	mode	✓
2021-05-19	May 19 2021 China-crackdown crash (\$9-10B liquidation cas...	market-crash	cata	✓
2021-07-19	July 2021 cycle-low capitulation (sub-\$30k) — trough, weak...	market-crash	mino	✓
2021-09-07	El Salvador legal-tender-day flash crash (Chivo failure, s...	deleverage-cascade	majo	✓
2021-09-20	Evergrande contagion global risk-off	macro-shock	mode	✓
2021-09-24	PBoC declares all crypto transactions illegal (total ban)	regulatory	mode	✓
2021-11-26	Omicron 'Black Friday' risk-off (onset ; Dec-3/4 leverage f...	macro-shock	majo	✓
2022-01-21	January 2022 Fed-tightening risk-asset rout / crypto capit...	macro-shock	majo	✓
2022-02-24	Russia invades Ukraine — risk-off shock (V-shaped recovery)	macro-shock	mode	✓
2022-05-05	May FOMC 50bp hike risk-off leg (pre-Terra macro deleverag...	macro-shock	majo	✓
2022-05-07	Terra/UST algorithmic stablecoin collapse (onset = Curve 3...	stablecoin-depeg	cata	✓
2022-06-12	Celsius freeze / 3AC insolvency lender contagion (cycle lo...	deleverage-cascade	majo	✓
2022-08-26	Powell hawkish Jackson Hole speech ('some pain')	macro-shock	mode	✓
2022-09-13	August CPI hotter-than-expected — macro vol shock	macro-shock	mode	✓
2022-09-15	Ethereum Merge — ETH sell-the-news, priced-in, realized vo...	other	mino	
2022-11-08	FTX exchange collapse (FTT crash, Binance LOI; USDT depeg ...	exchange-failure	cata	✓
2022-11-16	FTX contagion : Genesis halt, Gemini Earn freeze, BlockFi b...	deleverage-cascade	mode	✓
2023-01-19	Genesis Global files Chapter 11 — contained, no BTC vol sp...	exchange-failure	mino	
2023-03-08	Banking crisis cascade : Silvergate wind-down (onset), SVB ...	stablecoin-depeg	majo	✓
2023-03-27	CFTC sues Binance/CZ — minor intraday dip, no sustained vol	regulatory	mino	
2023-06-05	SEC sues Binance (Jun 5) and Coinbase (Jun 6)	regulatory	mode	✓
2023-08-17	August deleverage cascade + higher-for-longer rates (large...	deleverage-cascade	majo	✓
2023-08-29	Grayscale wins D.C. Circuit vs SEC on GBTC conversion — up...	regulatory	mode	✓
2023-10-16	Cointelegraph false BlackRock spot-ETF approval headline (...)	other	mino	✓
2023-10-23	BlackRock IBTC DTCC-listing rally (BTC to 2023 high \$35k,...)	regulatory	mode	✓
2023-11-21	Binance/CZ guilty plea, \$4.3B DOJ settlement — muted, no s...	regulatory	mino	
2024-01-11	Spot BTC ETF launch sell-the-news + GBTC outflow cascade	regulatory	mode	✓
2024-03-05	March 2024 ATH wick + post-halving-run correction	market-crash	mode	✓
2024-04-13	Iran-Israel weekend strike liquidation cascade (>\$1.4B longs)	deleverage-cascade	mode	✓
2024-07-04	Mt. Gox repayments + German government BTC liquidation	market-crash	mode	✓
2024-08-02	Yen-carry unwind / Black Monday (jobs-report trigger ; Aug-...	macro-shock	majo	✓
2024-09-06	Weak August jobs report macro sell-off	macro-shock	mino	✓

Date	Event	Category	Sev.	Test
2024-11-05	US presidential election volatility spike (Trump win, upsi...	macro-shock	mode	✓
2024-12-05	December ATH rejection at \$100k (>\$1B liquidations)	deleverage-cascade	mode	✓
2024-12-09	December flash crash + \$1.52B liquidation cascade	deleverage-cascade	mode	✓
2024-12-18	Hawkish FOMC 25bp cut (dot-plot drops 2025 cuts 4->2)	macro-shock	mode	✓
2025-02-01	First Trump tariff order (Mexico/Canada/China) selloff (si...	macro-shock	majo	✓
2025-02-21	Bybit \$1.5B ETH hack (ETH -24%) + tariff-reactivation risk...	exchange-failure	majo	✓
2025-04-02	Trump 'Liberation Day' tariff shock (2025 low \$74.7k Apr 7)	macro-shock	majo	✓
2025-06-13	Israel strikes Iran (Mideast war escalation)	macro-shock	mode	✓
2025-06-22	US airstrikes on Iran nuclear sites (>\$1B liquidations)	macro-shock	mode	✓
2025-08-01	Weak July NFP + renewed tariff angst macro drop (\$850M liq)	macro-shock	mino	✓
2025-08-25	Whale-driven Sunday flash crash (24k BTC dump into thin l...	deleverage-cascade	mino	✓
2025-09-22	September long-liquidation cascade (ETH-led, \$1.6-1.8B liq)	deleverage-cascade	mode	✓
2025-10-10	Oct 10 record \$19B liquidation (Trump 100% China tariff; ...	deleverage-cascade	cata	✓
2025-11-04	AI-bubble risk-off, BTC breaks \$100k (Palantir-led tech se...	macro-shock	majo	✓
2025-11-18	Mid-Nov deleveraging onset (BTC breaks \$90k -> \$81k trough...	deleverage-cascade	majo	✓
2025-12-01	Dec 1 risk-off selloff (worst day since March; BoJ-hike ye...	deleverage-cascade	mode	✓
2026-01-20	Japan JGB bond-market panic spills into crypto (\$1.08B liq)	macro-shock	mode	✓
2026-02-02	February 2026 deleverage cascade (CME basis-trade unwind +...	deleverage-cascade	cata	✓
2026-02-28	US-Israel strikes on Iran (war onset, oil spike)	macro-shock	mode	✓
2026-03-17	SEC/CFTC joint crypto-classification guidance — bullish re...	regulatory	mino	
2026-03-22	Resolv USR SERVICE_ROLE/AWS-key exploit depeg — idiosyncra...	protocol-failure	majo	
2026-05-13	Hot PPI inflation shock + Warsh Fed-chair transition (onse...	macro-shock	majo	✓
2026-05-18	May 2026 flash crash below \$77k (Treasury yields >5.1%, Ir...	deleverage-cascade	mode	✓
2026-05-28	Second US-Iran escalation (Hormuz strikes; BTC 6-week low,...	macro-shock	majo	✓
2026-06-01	Strategy first BTC sale in 4 years (8-K trigger) + ETF-out...	deleverage-cascade	majo	✓
2026-06-03	Strait-of-Hormuz / Kuwait-airport war escalation (Brent >\$...	macro-shock	majo	✓