The Second Moment of the Index Effect: Comovement, Liquidity, and the Statistical Identity of STOXX Europe 600 Constituents

A quantitative study of quarterly index reviews, 2014–2025 — companion to "The Index Effect Before the Announcement"

v0.9 — 5 July 2026 · all numbers from audited artifacts; see the provenance manifest

Contents

Abstract

Does joining an index change what a stock is, or only how it trades? We answer for Europe's most-tracked benchmark, the STOXX Europe 600, using 289 usable scheduled additions and 285 usable surviving demotions (of 406 scheduled deletions) across 46 quarterly review cycles, own-stock-free index returns, and the Chen–Singal–Whitelaw correction battery as the default design rather than a robustness appendix. Scheduled additions show a real rise in daily index beta — median +0.095, review-cycle bootstrap interval [+0.03, +0.16], above all 2,000 cycle-conditional placebo draws — which survives a matched-peer control (ruling out market-wide beta drift) and then dies under every synchronicity correction: a median of +0.014 with Dimson lead–lag betas, +0.023 at weekly frequency. The anatomy is mechanical: the correction lifts the pre-inclusion beta

from 0.95 to 1.03 and leaves the post-inclusion beta untouched, and additions gain roughly 26% in genuine share turnover at the mean — inclusion speeds up price discovery rather than rewiring comovement. The one exception is 2018–2021, where the shift's magnitude is preserved across the entire battery on the mean (0.12–0.17; medians 0.08–0.13), marginal at sixteen review cycles — and it coincides with the flat segment of the passive-AUM path, while the effect disappears during the fastest passive growth on record; the alignment runs backwards from a habitat account at both ends. Surviving demotions show no detectable comovement decline on any battery leg in any era and no loss of trading activity: the demotion-side "liquidity loss" is price mechanics. STOXX Europe 600 membership changes how a stock trades, not what it is.

---

1. Introduction

The first volume of this study measured what a STOXX Europe 600 membership change does to a stock's price: a repricing of roughly ±4% that happens, almost entirely, before the change is announced. This volume asks a different question — what membership does to the stock itself. A price effect is a one-time transfer; a change in a stock's *statistical identity* — its beta to the benchmark, its share of variance explained by the index, its turnover — reprices every hedge, every margin model, and every pairs book that touches the name for as long as the change persists. For a trading desk, the second moment is not a curiosity: the post-inclusion beta *is* the hedge ratio.

The claim that indexing rewires comovement has a distinguished pedigree. Barberis, Shleifer and Wurgler (2005) documented that stocks added to the S&P 500 begin to comove more with the index and less with their old habitat, and read this as evidence for category- and habitat-based trading: if investors trade "the index" as a thing, membership itself changes a stock's return dynamics, with no change in fundamentals. This finding became a load-bearing wall for behavioral views of comovement. It also acquired an equally distinguished critique. Chen, Singal and Whitelaw (2016) showed that most of the measured beta shift in the U.S. evidence is mechanical rather than behavioral: newly added stocks trade non-synchronously with the index before inclusion (biasing pre-betas down), enter the very index they are regressed on, and drift in size and industry composition around the event. Corrected with lead–lag (Dimson) estimators, lower-frequency returns, and controls for that composition drift, the habitat effect largely evaporates in their data.

Europe has never had this autopsy. The comovement literature is overwhelmingly S&P-centric; the European evidence that exists predates the correction battery, and none of it exploits the institutional feature that makes the STOXX Europe 600 the cleanest possible laboratory: a rule-based, committee-free review whose scheduled deletions are *demotions* — stocks that fall to the mid-cap segment and keep trading — rather than deaths. A two-sided design is therefore available in Europe in a way it never quite is in the S&P 500, where deletions are contaminated by M&A and distress: if index membership changes what a stock is, surviving demotions should un-change it, on the same tape, under the same rules.

This paper runs that design at the same evidentiary standard as its companion volume, and with the post-2016 literature's discipline built in rather than appended. The Chen–Singal–Whitelaw correction battery — Dimson lead–lag betas, weekly-frequency re-estimation, matched-peer difference-in-differences, and a characteristics-drift check — is the *default* here: no raw comovement shift is reported as a finding unless it survives the battery. Two legs of our chain go beyond what the U.S. critique itself had. First, a matched-peer DiD on the beta shift rules out the one alternative it can rule out — market-wide beta drift from volatility regimes — so the raw effect we then dissect is not an artifact of *when* reviews happen. Second, we attach a mechanism: a liquidity decomposition that separates genuine trading activity (share turnover) from price-level mechanics, telling us *why* daily betas move when they do.

The answer the data give is one sentence long: **STOXX Europe 600 membership changes how a stock trades, not what it is.** Scheduled additions show a real, placebo-corroborated rise in daily beta — a median shift of +0.095 with a review-cycle bootstrap interval that excludes zero, surviving the peer control (+0.102, cluster-t 4.46). But the shift dies under every synchronicity correction: mean shifts of +0.024 under Dimson (cluster-t 0.61) and +0.055 at weekly frequency (t 1.09). The anatomy is exactly the one Chen–Singal–Whitelaw predicted: the correction operates entirely on the pre-event window — the Dimson estimator lifts the pre-inclusion beta from 0.95 to 1.03 and leaves the post-inclusion beta untouched — so the raw rise is a pre-window nonsynchronicity artifact, not a post-event transformation. The mechanism side agrees: additions gain roughly 26% in genuine share turnover at the mean — inclusion speeds up price discovery, which raises measured daily comovement without touching the lower-frequency covariance structure that a habitat story requires.

The era dimension sharpens rather than softens this reading. In the middle of our sample — 2018 through 2021 — the addition-side shift is the one cell the battery does *not* kill: the point estimate sits between 0.08 and 0.13 on the median (0.12–0.17 on the mean) on every leg, raw, Dimson, and weekly alike, a pattern we report with directional language because sixteen review cycles cannot carry more. And the passive-ownership alignment runs exactly backwards from the habitat prediction: that battery-robust window coincides with the *flat* segment of the passive-AUM path (+€0.23bn per year), while the era of fastest passive growth (+€5.3bn per year) is the era in which the corrected effect disappears. Whatever briefly moved European comovement in 2018–21, the alignment points away from index demand.

The deletion side completes the argument with a null that talks. Surviving demotions — 285 usable events of 406 scheduled deletions, a funnel we disclose to the last stock — show no detectable comovement decline on any battery leg in any era, and no loss of trading activity: their EUR-volume decline decomposes entirely into the price leg , with share turnover statistically flat. Demoted stocks keep trading; nothing mechanical reverses; and the asymmetry with additions is precisely what a trading-synchronicity account predicts and a habitat account does not.

Five things in this volume are, to our knowledge, new. It is the first two-sided, CSW-corrected comovement study of Europe's flagship benchmark, with surviving demotions as the mirror sample. It adds two legs the U.S. critique lacked: the matched-peer DiD control and the turnover mechanism. It reframes the index liquidity effect: additions gain real activity, while the demotion-side "liquidity loss" is price mechanics — a distinction with direct consequences for how execution desks should read post-demotion depth. It delivers an era-resolved attenuation test whose passive-AUM alignment cuts against the demand story on both ends. And it maintains the companion volume's inference discipline throughout: 46 review cycles are the unit of inference, every significance statement is cluster-robust or block-bootstrapped at the cycle level, and every diagnostic that falls short of that standard is labeled as such where it appears.

The paper proceeds as follows. Section 2 describes the data layer and the sample funnels, including the disclosures that make the deletion design honest. Section 3 specifies the comovement regressions, the correction battery, and the inference doctrine. Section 4 dissects the addition-side result. Section 5 resolves it across eras and confronts the passive-AUM path. Section 6 reads the deletion null. Section 7 decomposes the liquidity effect. Section 8 closes the habitat question and reports the cross-sectional non-result. Sections 9–11 present the robustness matrix, the limitations inventory, and the practitioner reading.

2. Data and sample construction

2.1 Upstream inheritance

This project consumes, read-only and at a frozen commit, the data layer built and audited for the companion volume: a survivorship-free daily price spine of 34.4 million stock-day rows across three vendor eras (a LSEG-sourced era through 2017-10-11, covering roughly 1,400 constituent-adjacent keys; a broad legacy era through 2023-06-13; and native STOXX data thereafter); an event panel of 1,219 constituent changes across 721 stocks and 49 quarterly review cycles, of which 421 scheduled additions and 406 scheduled deletions form this paper's headline cohorts; the benchmark return series; and the matched peer baskets and passive-AUM panel described there. Where this paper quotes a property of the inherited layer, the number comes from the companion volume's own provenance manifest and is not re-derived. The seams between vendor eras — different FX conventions, corporate-action treatment, and coverage — are inherited as disclosed limitations, and every era-stratified result in this paper carries them (Section 10).

2.2 The daily membership calendar

The comovement design needs something the upstream layer deliberately does not provide: a daily answer to "which stocks are in the SXXP today, at what weight?" We reconstruct it by replaying the complete review history against 122 quarterly composition snapshots. The replay validates exactly: after applying a signed, five-defect override table (two silent identifier remaps and three changes absent from the source change log), all 122 snapshots reconcile with zero unexplained divergences, daily membership counts stay within the index's 600–603 band, and reconstructed weights sum to one to machine precision. The three log-omitted changes deserve one honest sentence each in any use of this calendar: they are absent from both the STOXX change log and the parent event panel (a 2020 deletion of Quilter, a 2020 addition of Allegro, and a 2020 addition of THG), and their dates are re-derived from selection-list brackets and asserted at build time. One of them matters downstream: THG's 2021 scheduled deletion is a headline sample member whose membership spell opens on the reconstructed addition. We flag it rather than hide it; a median null over 285 events does not move on one spell, but a reader auditing our membership spells against the public log will find exactly this discrepancy, and should find it disclosed.

2.3 Own-stock-free index returns and the outside-Europe leg

An added stock enters the very index it is regressed on, and part of any measured beta rise is that arithmetic. Every comovement regression in this paper therefore uses the index *excluding the stock itself*:

$$R_t^{\mathrm{ex}\,i} \;=\; \frac{R_t^{\mathrm{idx}} - w_{i,t}\,R_{i,t}}{1 - w_{i,t}},$$

with daily weights from the membership calendar and free-float market capitalization; for nonmembers the formula degrades to the plain index return. The strip is exact in simple-return space and is validated event-by-event (679,485 event-day rows across all 1,219 events; the median within-event correlation between the stripped and unstripped index return is 1.000 to three decimals, as it should be with individual weights all below 5%). The bivariate habitat test needs a second regressor: the return on investable Europe *outside* the index, which we build from the selection-list universe (the same 1,700–3,000 name pool the index selects from) minus current members, free-float weighted, and correlated 0.907 with the index itself — a collinearity we handle by design rather than denial (Section 3.1).

During construction we found, and worked around, a genuine upstream defect (its full inventory is in Section 10): in the pre-2017-10 vendor era, sixteen stocks carry EUR prices and market capitalizations inflated by unit errors of ×6 to ×431. Log returns are scale-invariant and are unaffected; the outside-index weights are built from selection-list capitalizations rather than the defective panel column; the one affected stock that was an index *member* during the defect window has its weight repaired against a selection-list anchor; and seventeen EUR-volume events on seven of these keys are excluded from the liquidity section's EUR-denominated cells. The comovement and liquidity sections therefore run on samples that differ by construction, and we never present them as one (Section 7.3).

2.4 Windows

All dates run on the STOXX Europe trading calendar, and all windows are anchored on the two event dates of the review timeline documented in the companion volume: the announcement (T_ann) and the effective date (T_eff).

Window	Definition	Role
PRE	[T_ann - 273, T_ann - 21] trading days	pre-event beta and liquidity baseline
EVENT (excluded)	[T_ann - 20, T_eff + 20]	announcement drift, MOC pressure, reversal — never estimated over
POST	[T_eff + 21, T_eff + 273]	post-event beta and liquidity
POST-SHORT (C4 only)	[T_eff+1, +21], [+22, +63], [+64, +273]	temporary-vs-permanent liquidity split

The excluded EVENT window is the entire subject of the companion volume; leaving it inside an estimation window would contaminate betas with exactly the price pressure that paper measures. The PRE window deliberately ends where the companion volume's market-model estimation window ends (T_ann - 21), and a robustness leg re-runs everything with PRE ending five days before the selection-list date, since paper 1 documents pre-announcement repricing inside [T_sl, T_ann] . Estimation requires 150 clean daily observations per window (30 weekly observations for the weekly legs).

One stock can appear in several events — added, later demoted, sometimes re-added. Windows of consecutive same-stock events are truncated at each other's EVENT boundaries, and where two windows of the same stock still conflict, both events are voided: 352 events are clipped by a neighbour, 100 events (50 same-stock pairs) are voided outright, and the usable count for the univariate design lands at 669 of the panel's 1,219 events. The headline samples are the scheduled cohorts within that set: **289 usable scheduled additions and 285 usable surviving demotions (365 survivors / 406 scheduled; Section 2.5), each spanning 46 review cycles**.

2.5 The deletion funnel

Deletion results in this design are conditional on survival, by construction: a post-event beta requires 150 post-event trading days, which a stock delisted in a merger does not have. The scheduled-deletion cohort is where this bites least — scheduled deletions are demotions to the mid-cap segment, not deaths — but "least" is not "not", and the funnel from 406 scheduled deletions to 285 usable events is the single most attackable object in the paper, so we print it in full. Of 406 scheduled deletions, 41 are non-survivors: 17 structurally (one death at the effective date, eleven delistings inside the post-window, five stocks alive but trading below the observation floor), and 24 because they are recent (2025H2 and later) and the data simply end before their post-windows do — right-censored, not selection-excluded. That leaves 365 surviving demotions. Of these, 80 fail estimation requirements — 48 on the pre-window floor (overlap truncation), 37 on the post-window floor (34 of these because the stock was *re-added* to the index inside the post-window, which truncates the demoted spell); five events fail both floors, so the union is 80. The result is 285 usable surviving demotions.

Two disclosures make this funnel honest rather than merely visible. First, survival is not random: the five alive-but-below-floor exclusions have a median pre-event EUR volume of €1.08M against €6.79M for survivors — the deletion sample under-represents exactly the least liquid demotions, and every deletion claim in this paper is a claim about *surviving demotions*, a phrase we use verbatim throughout. Second, the re-add truncation is a conditioning channel — dropping demotions that later return could bias the sample toward "unsuccessful" demotions — which we close empirically in Section 9: readmitting the truncated events at a shorter observation floor (augmenting the cell to 307 events) leaves the deletion null unchanged.

3. Methodology

3.1 Comovement regressions

For each event *i* we estimate, separately on the PRE and POST windows, the univariate market model on daily log returns against the own-stock-free index return,

$$R_{i,t} = \alpha_i + \beta_i \, R^{\text{ex}\,i}_{\text{SXXP},t} + \varepsilon_{i,t},$$

and record the event-level shift $\Delta\beta_i = \beta_i^{\text{POST}} - \beta_i^{\text{PRE}}$ (and the analogous ΔR_i^2, aggregated in Fisher-z space). The benchmark is the net-return index (SXXR) matched to net-return stock series, with the price-return pair (SXXP) as a robustness leg; the gross-return index is excluded on the tax grounds documented in the companion volume. The bivariate habitat specification adds the outside-index Europe return,

$$R_{i,t} = \alpha_i + \beta_i^{\text{in}} R^{\text{ex}\,i}_{\text{SXXP},t} + \beta_i^{\text{out}} R_t^{\text{out}} + \varepsilon_{i,t},$$

under which the habitat prediction is $\Delta\beta^{\text{in}} > 0$ *and* $\Delta\beta^{\text{out}} < 0$ for additions. The two regressors are correlated at 0.907 , which is not a defect but a feature of the design inherited from Barberis, Shleifer and Wurgler: with regressors this collinear, coefficient *levels* are unstable and we never interpret them — only the pre-to-post *shifts*, whose estimation error is common across windows, following the presentation convention of that literature.

3.2 The correction battery

The paper's central discipline (and, we would argue, its licence to speak after 2016) is that no raw comovement shift is a finding until it survives all four legs of the correction battery — three aimed at the artifact channels Chen, Singal and Whitelaw identified, plus the matched-peer control this paper adds:

Dimson lead–lag betas. Small and newly added stocks trade less synchronously with the index before inclusion; same-day OLS betas are biased down in the PRE window. We re-estimate with one lead and lag of the index return (±2 as sensitivity), summing the coefficients.

Frequency escalation. A synchronicity artifact lives at daily frequency and dies at weekly; a genuine change in comovement does not. We re-estimate the full design on Wednesday-to- Wednesday weekly returns (with an alternative Wednesday-close compounding as a robustness leg).

Matched-peer difference-in-differences. Each event's $\Delta\beta_i$ is differenced against the same-window shift of its matched non-event peer basket (same country and size band, inherited from the companion volume). This kills one specific alternative — market-wide beta drift from volatility regimes that happen to straddle review dates — and *only* that alternative: a DiD that survives says the shift is event-specific, not that it is non-mechanical. We hold that line wherever the DiD appears.

Characteristics drift. Size composition changes around events; we regress $\Delta\beta$ on the change in log market capitalization and read how much of the raw effect the control absorbs.

3.3 Inference

Events arrive in batches: all events of a review cycle share one market realization, one volatility regime, one crisis if there is one. Everything inferential in this paper is therefore clustered at the review-cycle level (46 cycles for the pooled headline samples), in two deliberately different forms, and it is worth being explicit because the two answer different questions. The **cluster-robust t** tests whether the cross-event *mean* shift is zero. The **review-cycle block bootstrap** resamples whole cycles and delivers a confidence interval for the *median* shift — the estimator robust to the fat event-level tails that betas estimated on 150–250 observations produce. We report both, always labeled, and never splice them into a single "estimate [CI]" pair: where the two reads disagree (they do in a handful of disclosed cells, most consequentially in Section 5, and we say so at each), the disagreement is information about estimand and tail shape, not a menu to choose from. Sign tests appear as directional corroboration only — under cycle clustering a pooled sign test is anti-conservative, and no significance statement in this paper rests on one. The same honesty applies to the peer placebo of Section 9: it redraws pseudo-events within *fixed* cycles, so its randomization band is roughly 0.6× as wide as the cluster-robust interval, and we quote it exclusively as rank evidence, never as a p-value with cluster-robust standing.

Multiple testing is handled by architecture rather than adjustment: the raw-to-corrected sequence is one nested family whose headline claims come only from the corrected end (P2-6); the era stratification is pre-specified, inherited verbatim from the companion volume's periodization (Early 2014–17, Middle 2018–21, Late 2022–); and the single pre-registered cross-sectional specification is the only one carrying inferential weight in Section 8, with everything else labeled exploratory in the artifact itself.

3.4 Liquidity measures

Volume data cover 717 of the 721 event stocks at daily frequency. EUR volume is raw share volume times the *raw* EUR close — never the adjusted price, which would smear splits across the volume series. The primary liquidity object is **turnover** (volume over shares outstanding), which is price-free; EUR volume is decomposed into turnover, shares-outstanding, and price legs precisely because the composite conflates activity with valuation; and the Amihud illiquidity ratio is the price-impact companion. All window statistics are medians over window days (the effective-date volume spike is the companion volume's subject, not a liquidity level), compared pre-to-post as log ratios, with the three post-event horizons of Section 2.4 separating the temporary from the settled level — a split that exists only for the EUR-volume measure, a point Section 7.2 keeps explicit. Days with no trading are absent rows, not zeros, and the code counts rows rather than trusting NA-handling; spread- and quote-based measures are out of scope for want of quote data, which is the one leg of the classic S&P 500 liquidity design (Hegde–McDermott) we cannot mirror and say so.

3.5 Cohorts, eras, and scrubs

Additions and deletions are never pooled (they are different economic events with different selection); mid-cycle changes (fast entries, M&A deletions, spin-offs) never enter headline results and appear only in the descriptive appendix; and every headline cell is reported both on the full sample and with the companion volume's contamination scrub (corporate-action and spin-off windows), which nowhere changes a conclusion. Era cuts are the companion volume's, to the day, hard-coded and byte-checked; era cells rest on 14–16 cycles each, and the paper's language discipline scales with that power.

4. The inclusion beta rise, dissected

4.1 The raw fact, and that it is a fact

Scheduled additions comove more with the index after inclusion, at daily frequency, by any uncorrected read. The median event-level beta shift is +0.095, with a review-cycle bootstrap interval of [+0.032, +0.163] that excludes zero; the mean shift is +0.100 with a cluster-robust t of 3.57 across 46 cycles. The bivariate specification says the same thing louder ($\Delta\beta$_in median +0.152, mean +0.168, cluster-t 3.82). And the shift is not an artifact of how we build peer comparisons or estimate betas: in a placebo exercise that re-estimates the identical statistic on 2,000 pseudo-event draws from the matched peer pool within fixed review cycles, the observed median sits above every one of the 2,000 placebo medians (rank q = 1.000, against a placebo band of roughly ±0.04). That band is a cycle-conditional randomization — it does not resample cycles, and we never quote it as cluster-robust significance — but as rank evidence it settles one question completely: the raw addition-side shift is real, not noise mining (Figure 5).

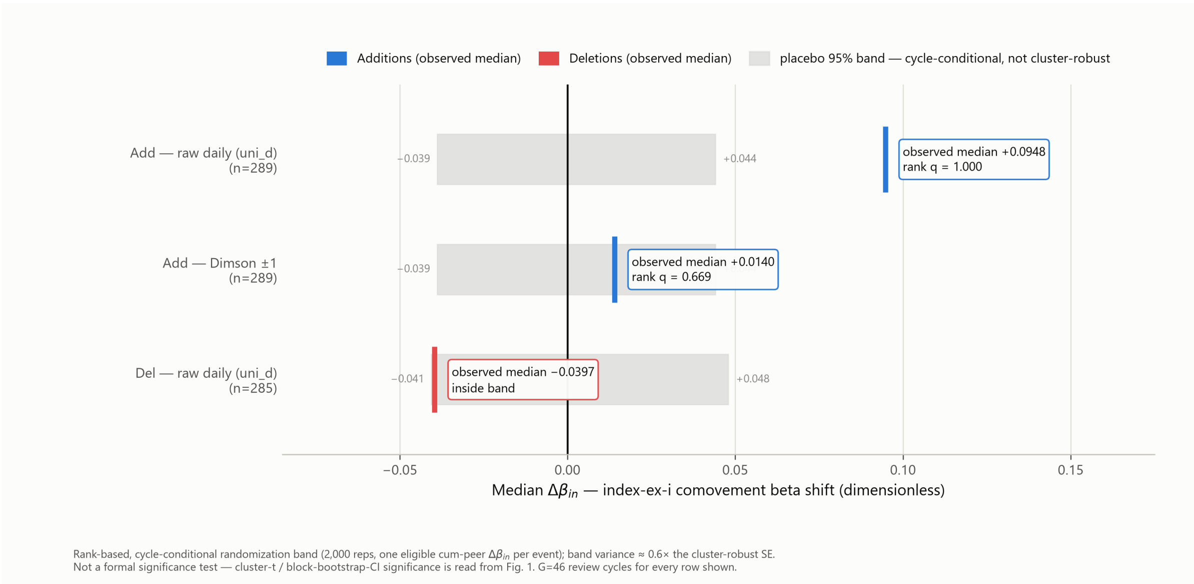

Figure 5. Observed median $\Delta\beta$_in against the peer pseudo-event placebo band (2,000 draws, peers redrawn within fixed review cycles). The raw addition median sits above every placebo draw (rank q = 1.000); the Dimson-corrected median sits mid-band; the deletion median sits inside the band. The band is a cycle-conditional randomization — variance roughly 0.6× the cluster-robust interval — and is quoted as rank evidence only, never as significance; cluster-t and bootstrap-CI significance is read from Figure 1.

4.2 Not market drift

The first alternative any referee should raise is that betas drifted for everyone — review dates straddle volatility regimes, and a beta estimated into a crisis differs from one estimated out of it. The matched-peer difference-in-differences closes exactly this channel: differencing each event's shift against its country- and size-matched non-event basket over the identical calendar windows leaves the effect intact — DiD median +0.102, bootstrap CI [+0.033, +0.149], mean +0.102 with a cluster-t of 4.46, on peer baskets whose own average shift is statistically zero. We state the scope of this result as precisely as the design permits: the DiD rules out market-wide beta drift, and nothing more. It does not certify the shift as non-mechanical — that is the battery's job, and the battery reaches the opposite verdict.

4.3 The kill, and its anatomy

Under the Dimson lead–lag correction the pooled shift collapses to +0.024 (cluster-t 0.61; median +0.014, CI [-0.073, +0.102]); at ±2 leads and lags it is +0.052 (t 1.13) with a median of +0.001; at weekly frequency it is +0.055 (t 1.09; median +0.023, CI [-0.062, +0.107]). These are same-sample comparisons — the five legs run on materially identical events (raw 289, peer-DiD 289, Dimson ±1 288, Dimson ±2 285, weekly 289) across the same 46 cycles — so the kill is not a composition change. The corrected estimate is placebo-sized: against the same 2,000-draw placebo band, the Dimson median sits at rank q = 0.669, indistinguishable from a pseudo-event.

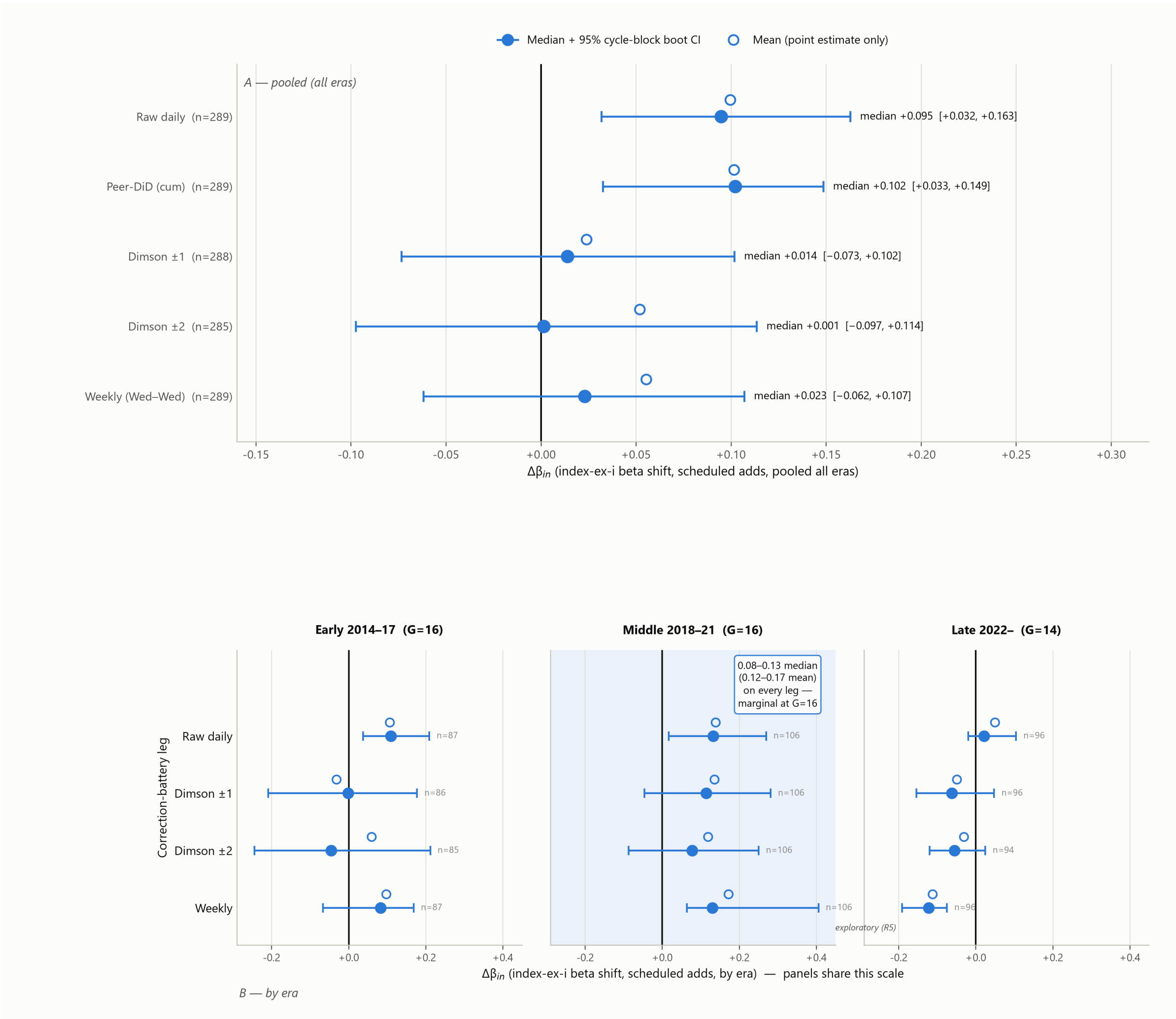

Figure 1. The correction battery for scheduled additions. Panel A (pooled, 46 cycles): median $\Delta\beta$_in with review-cycle block-bootstrap CI (filled dot and whisker) and the mean as an open marker; the raw daily shift and the peer-DiD are the only legs whose CI excludes zero. Panel B (by era, shared scale): the same four legs per era; in the Middle era the magnitude is preserved on every leg (medians 0.08–0.13, means 0.12–0.17), marginal at G = 16; the Late-era weekly cell is exploratory (R5). Per-leg n as printed; the Dimson kill is a same-sample comparison.

Figure 2 shows *where* the correction bites, which is the anatomy that separates a mechanical story from a behavioral one. On the matched 288-event sample, the Dimson estimator lifts the **pre-event** beta from 0.950 to 1.028 and moves the post-event beta from 1.049 to only 1.052 . The raw +0.10 was never a post-inclusion transformation; it was a pre-inclusion *measurement bias*. Before inclusion these stocks trade non-synchronously enough with the index — the same-day OLS beta misses covariance that arrives at a one-day lag — that the pre-beta is understated by almost exactly the size of the “effect”. Correct the measurement, and the stock’s comovement is flat through the event. This is the Chen–Singal–Whitelaw signature, reproduced on a two-sided European sample.

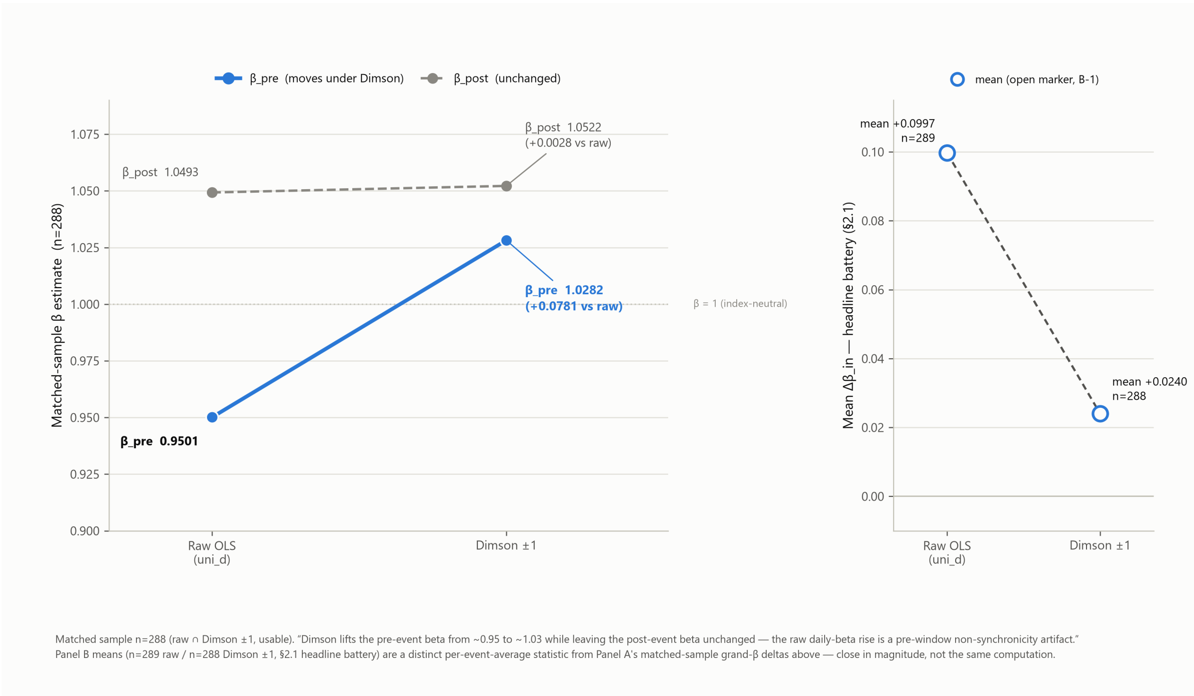

Figure 2. Anatomy of the Dimson kill on the matched 288-event sample. Left: the lead–lag correction lifts the pre-event beta from 0.950 to 1.028 while the post-event beta is essentially unchanged (1.049 $\to$ 1.052) — the raw daily-beta rise is a pre-window non-synchronicity artifact. Right: the corresponding headline mean $\Delta\beta$_in falls from +0.100 to +0.024 (open markers denote means; the two panels are different statistics, as the footnote in the figure states).

The characteristics-drift channel, the battery's fourth item, absorbs essentially nothing: regressing the raw shift on the change in log market capitalization leaves the adjusted shift, if anything, slightly larger, with an insignificant drift coefficient (-0.20, cluster-t -0.65, n = 267 across 45 cycles). Composition change is not the story either.

4.4 The mechanism: faster trading, not new identity

Why would daily-frequency comovement rise mechanically on inclusion? Because inclusion changes how fast a stock's price incorporates index-level information. The liquidity section establishes (Section 7, on the volume-covered subsample of 278 additions across 45 cycles) that additions gain roughly +24% in median genuine share turnover (+26% at the mean; bootstrap CI of the median log-shift [+0.140, +0.273], cluster-t 3.99) . A stock that trades more stops lagging the index by a day — which raises the same-day daily beta and does precisely nothing to weekly covariance. That is the pattern in the data: turnover up, daily beta up, Dimson-corrected and weekly betas flat. Inclusion accelerates price discovery; it does not re-wire what the stock is.

4.5 Reading

The pooled addition-side result is therefore a replication, with two upgrades, of the modern U.S. verdict: the measured comovement effect of index inclusion is dominated by trading synchronicity, consistent with Chen, Singal and Whitelaw (2016) — with a matched-peer control those authors lacked (ruling out the drift alternative), and a turnover mechanism connecting the correction's anatomy to an observable change in how the stock trades. Every conclusion survives the contamination scrub unchanged (raw +0.101, DiD +0.103, weekly +0.058 on 283 events, Dimson +0.028 on 282) and the price-return benchmark (+0.098 vs +0.100).

5. When the effect was real: the middle era

5.1 Three eras, three verdicts

Stratifying the battery by the companion volume's periodization turns one pooled kill into three distinct stories (Figure 1B). In the **Early era (2014–17)** the raw shift is +0.107 (t 1.95) and the Dimson correction annihilates it (-0.032, t -0.37) : these events ride the sample's largest pre-window lead–lag mass (median +0.146, falling to +0.059 in the middle era and +0.013 late), which is the non-synchronicity signature at its purest. We characterize this era with deliberate care: the diagnostic that would blame *data quality* in its coarsest form — full-day frozen closes — comes back empty (stale-run share identically zero in every era; near-zero-return-day share zero at the median with no era gradient), so frozen-price artifacts are excluded; but thin-trading partial adjustment and economic cross-exchange non-synchronicity are observationally equivalent at daily frequency, and we do not claim to separate them. The timing is at least suggestive: the early/middle boundary coincides with MiFID II (January 2018), after which European closing-auction consolidation reduced cross-venue non-synchronicity — corroboration for why the lead–lag mass falls, not a causal claim.

In the **Late era (2022–)** every corrected leg is flat to negative, and the weekly leg flips sign outright (-0.112, cluster-t -1.47, median CI [-0.19, -0.07]) — a cell we flag in Section 9's spirit as exploratory (it clears the bootstrap read but not the cluster-t, one cell inside a large family) and leave for future work as a possible "passive-era beta compression".

5.2 The middle era: the one cell the battery does not kill

In 2018–21 the addition-side shift survives everything we throw at it — as a magnitude. On the same 106 events across 16 cycles for all four legs, the median shift is +0.133 raw, +0.114 at Dimson ±1, +0.078 at Dimson ±2, and +0.130 weekly (means +0.139, +0.135, +0.119, +0.172). The Dimson correction, which erases the pooled effect, removes essentially nothing here *on the mean* — and, we note under our own estimand discipline, does shrink the Dimson-±2 median visibly (0.133 $\rightarrow$ 0.078), so the magnitude-preservation statement is a mean statement and we confine it to one. Significance is where sixteen cycles show their limits, and we report the full set rather than the best member: the weekly leg's median bootstrap CI excludes zero ([+0.06, +0.41]) while its cluster-t does not reach conventional levels (p.086); the Dimson-±1 leg's own bootstrap CI includes zero ([-0.05, +0.28], p.090); Dimson-±2 is p $\approx$.20. The two sanctioned reads disagree on the weekly leg — the fat-tailed mean sits a third above the median — and that disagreement is part of the result. Our language throughout is therefore the memo's: *a shift whose magnitude is preserved across the full correction battery, concentrated in 2018–2021, marginal at conventional levels given 16 cycles* — Dimson-robust, directional, and pre-specified (the middle era is one of three inherited cuts, not a searched breakpoint). The cell survives the Wednesday-close compounding variant (weekly mean +0.176, median CI [+0.070, +0.395]) and the contamination scrub (raw mean +0.140 on 105 events).

5.3 Attenuation is a hump, not a slope

Put the three eras together on the corrected estimand and the "attenuation" question inherited from the U.S. literature gets a more interesting answer than a trend: null $\rightarrow$ real $\rightarrow$ gone. The Dimson-±1 addition-side path of mean shifts runs -0.032 / +0.135 / -0.049 across eras, and the event-level trend regression agrees quietly: the linear term is nothing, while the concave quadratic term on the corrected estimand has a clustered p of.105 — shape-consistent, directional, and not significant, which is precisely how we phrase it. (The artifact also carries heteroskedasticity-robust columns that look stronger; under this paper's doctrine only the clustered read speaks.)

5.4 The AUM alignment runs backwards from habitat

If the middle-era shift were passive demand at work — the habitat story's modern form — it picked a strange decade. Figure 3 places the era hump against the passive-AUM path from the companion volume's ETF panel: over the full sample, identified passive AUM tracking the index grows 7.7-fold (first review cycle €4.4bn, last €33.8bn); but the **middle era is the flat segment** — cycle-mean €11.0bn, ramp +€0.23bn per year, a path that includes the 2018 drawdown, 2019 outflows and the COVID crash — while the **late era is the steepest ramp** (+€5.3bn per year, mean €19.2bn), exactly when the corrected effect disappears. An exploratory regression of the corrected shift on log passive AUM is null (t 0.29). The alignment cuts against a demand/habitat reading at both ends — the effect thrives without AUM growth and dies during the fastest AUM growth on record — and is comfortable with the synchronicity frame, in which what matters is market microstructure regime, not ownership stock. We present this as an alignment, not an identification; sixteen-cycle eras do not carry causal claims.

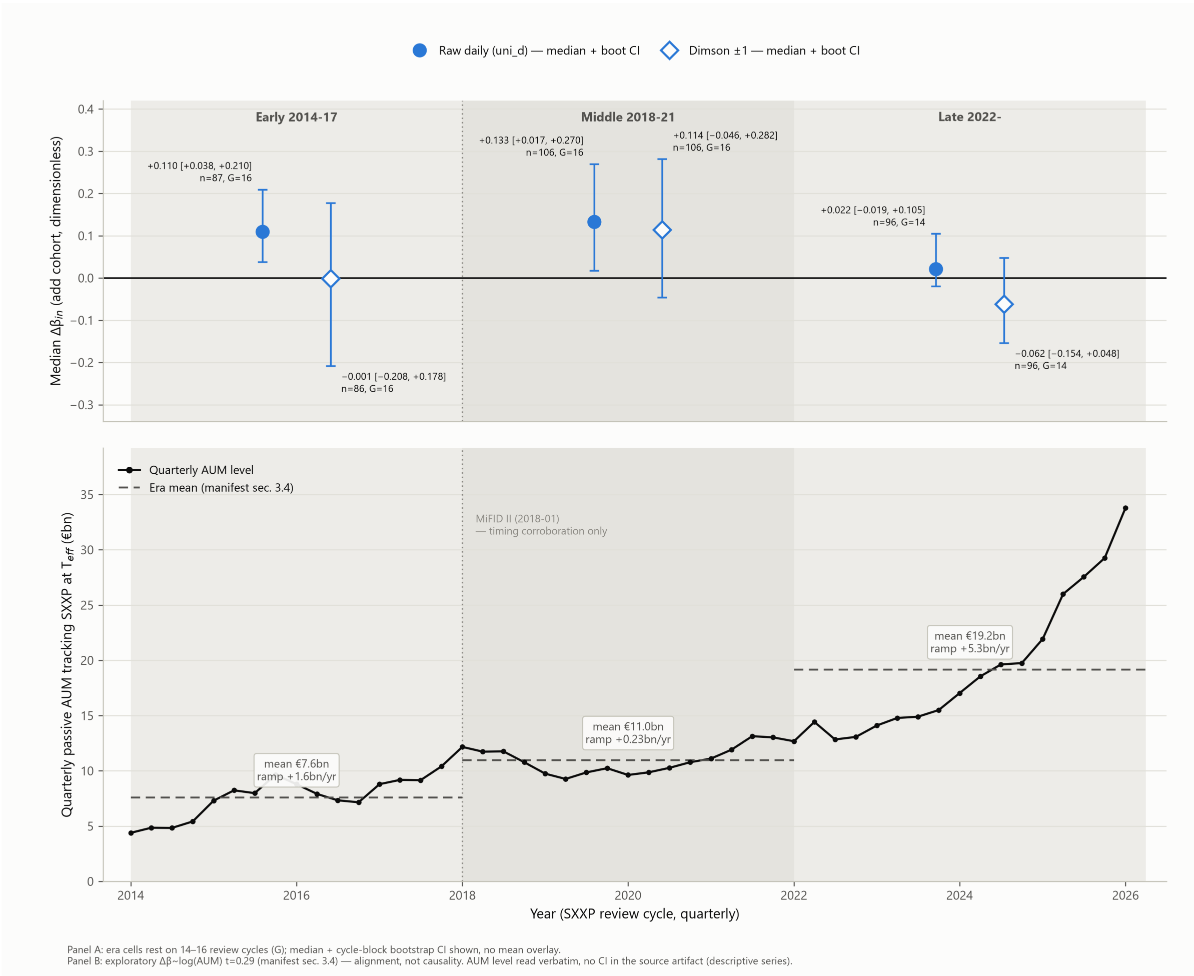

Figure 3. Top: median $\Delta\beta$_in by era for the raw daily and Dimson-corrected estimands (bootstrap CIs; 14–16 cycles per era) — the corrected effect is null, real, gone. Bottom: quarterly passive AUM tracking the index (€bn) with era means and ramps. The battery-robust Middle-era shift coincides with the flat AUM segment (+€0.23bn/yr); the effect disappears during the steepest ramp (+€5.3bn/yr). The MiFID II marker is timing corroboration only. Alignment, not causality.

6. Demotions: the null that talks

6.1 No detectable decline, on any leg

If inclusion rewires a stock, demotion should un-wire it. It does not. Across 285 usable surviving demotions (365 survivors / 406 scheduled), every univariate battery leg is flat: raw daily -0.034 (cluster-t -0.85; median -0.040, CI [-0.114, +0.025]), Dimson ±1 -0.025 (t -0.41), Dimson ±2 +0.019 (t 0.29, n = 283), weekly -0.062 (t -0.92), peer-DiD -0.034 (t -1.05) — every bootstrap interval straddling zero, in every era (the largest single cell, the late-era weekly at -0.228, has a cluster-p of.049 but a bootstrap CI of [-0.29, +0.05] and a null sign test — it clears one of three reads and we treat it as the exploratory cell it is). One directional wrinkle is disclosed rather than buried: in the *bivariate* specification the raw daily $\Delta\beta$_in is -0.124 with a cluster-t of -2.20 (scrubbed -2.56), but its bootstrap CI includes zero ([-0.178, +0.012]) and the effect is gone at weekly frequency (+0.029, t 0.23) — a daily-frequency shadow of the addition-side mechanics, failing the same battery the addition effect fails. The claim we make is exactly this: **no detectable comovement decline for**

surviving demotions — not “deletions have no effects” (they have a large price-mechanical one; Section 7).

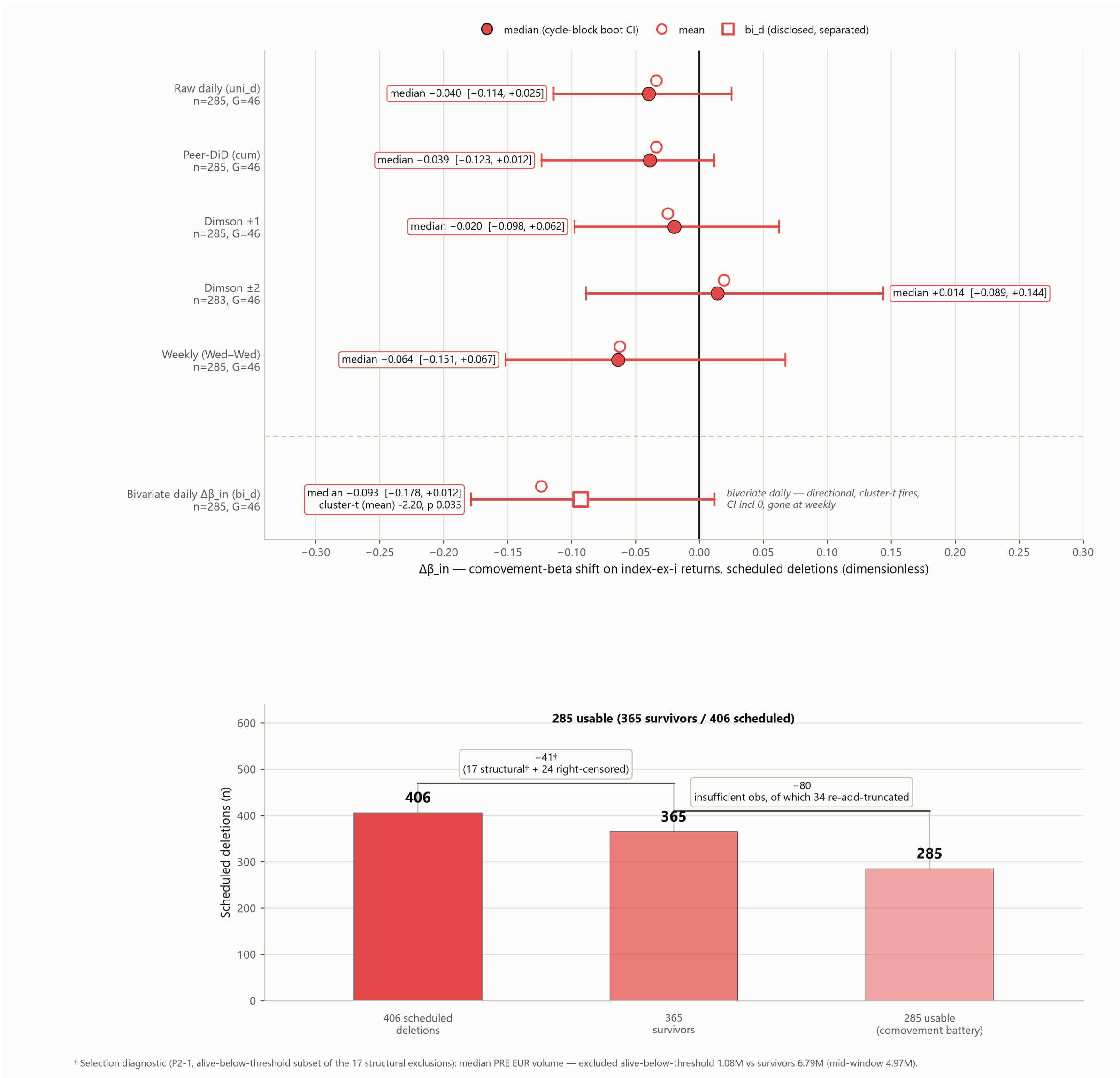

Figure 6. Top: surviving-demotion $\Delta\beta$_in across the battery — every leg’s bootstrap CI straddles zero; the bivariate daily cell (open square, visually separated) is the disclosed directional cell whose cluster-t fires while its CI includes zero and which is gone at weekly frequency. Bottom: the sample funnel — 285 usable (365 survivors / 406 scheduled) — with the structural, right-censored, and insufficient-observation drops labeled and the liquidity-selection diagnostic in the footnote.

6.2 Who is missing, and whether it matters

The funnel of Section 2.5 (Figure 6B) is this null’s fine print, and two of its channels can be tested rather than merely disclosed. The re-add conditioning — 34 of the 80 insufficient-observation drops truncate because the stock re-entered the index — is closed empirically: readmitting truncated demotions at a 60-day observation floor augments the cell to 307 events and moves nothing (median -0.036, cluster-t -0.60, CI [-0.100, +0.027]) ; the new-only companion cell is too small to sign anything (n = 22, CI [-0.225, +0.294]) and we do not. The liquidity-selection channel cannot be closed, only quantified: the excluded below-floor demotions are roughly six times less liquid than the survivors (median pre-event EUR volume €1.08M vs €6.79M) , so this null is a statement about demotions liquid enough to keep estimating — the population a desk actually trades.

6.3 The asymmetry is the point

Additions gain daily synchronicity and genuine turnover; surviving demotions lose neither. That asymmetry is exactly what the mechanical account predicts — an added stock starts trading faster and stops lagging the index; a demoted stock keeps trading (Section 7 shows its activity is statistically unchanged), so there is nothing mechanical to reverse — and it is awkward for a habitat account, under which exit from the category should mirror entry. We note the correspondence qualitatively and only qualitatively: the comovement and liquidity deletion samples differ by construction (285 usable events under the strict 150-observation floor vs 270–278 volume-covered events including two alive-below-threshold non-survivors), and no table in this paper pretends otherwise.

7. Liquidity: real activity vs price mechanics

7.1 Turnover first

The liquidity results reframe what "the index liquidity effect" is (Figure 4). On the price-free primary measure — share turnover — **additions gain and demotions lose nothing**: the addition-side median log-shift is +0.215 (+24% in levels; mean +26%; bootstrap CI [+0.140, +0.273], cluster-t 3.99, n = 278 of 421 scheduled across 45 cycles) , while the demotion-side shift is -0.027 at the median with a cluster-t of -0.43 and a CI of [-0.108, +0.017] — a genuine null, robust to removing the one upstream-defective volume series in the cell (DE101T; the mean moves -0.018 $\to$ -0.028).

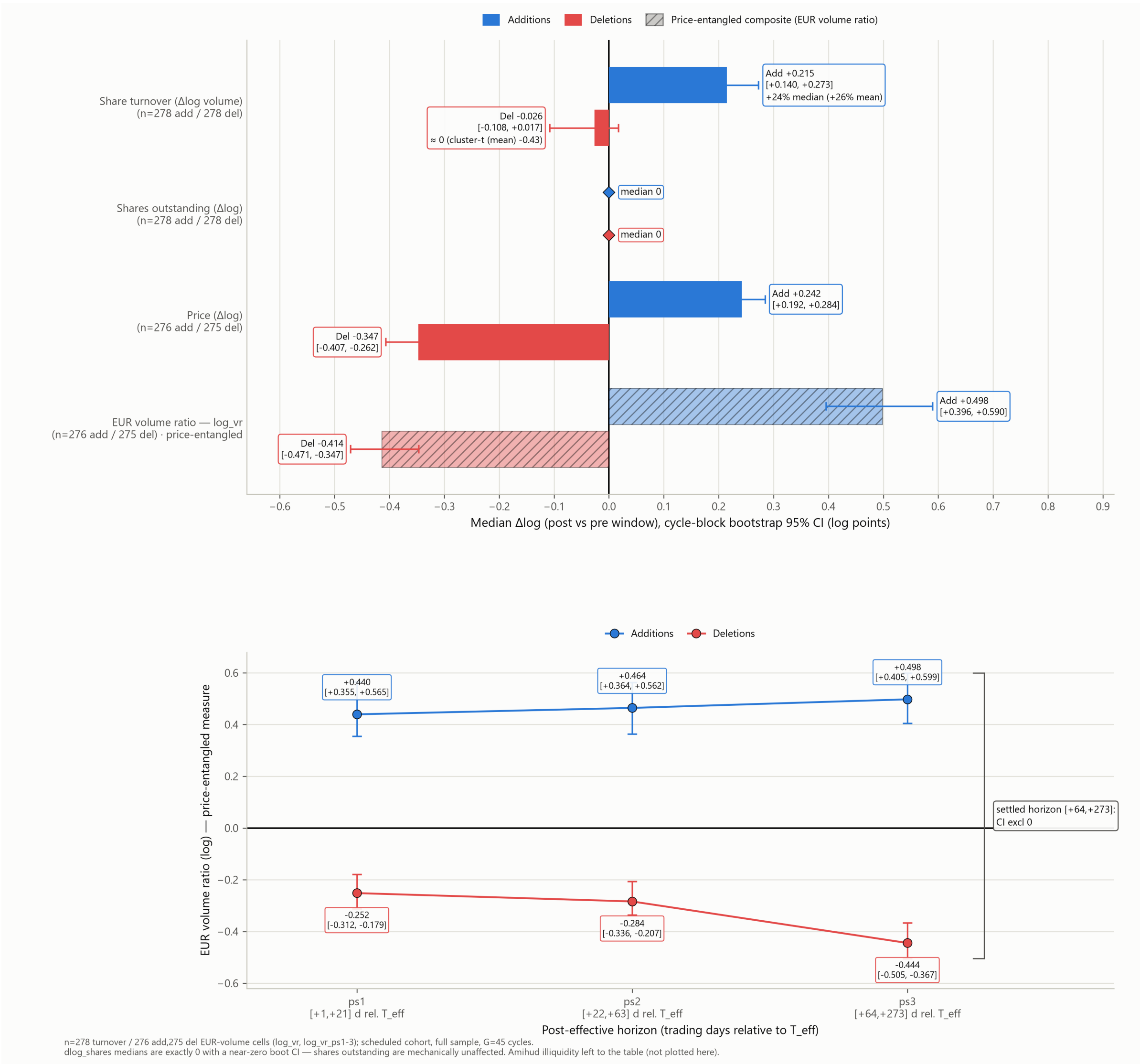

Figure 4. Top: median log-shifts with bootstrap CIs for the price-free turnover, the shares and price legs, and the price-entangled EUR-volume composite (hatched), additions vs demotions. Additions gain genuine turnover (+24% median); the demotion decline loads entirely on the price leg. Bottom: the EUR-volume horizon profile — the shift persists to the settled horizon on both sides; horizon language attaches only to this price-entangled measure.

The composite EUR-volume measure that a first pass would headline tells a different story only because it conflates activity with valuation. Demotion-side EUR volume falls hard (median -0.414, cluster-t -9.95) — but decomposing the log ratio into turnover, shares-outstanding and price legs loads the entire decline onto the **price** leg (-0.347 median, cluster-t -10.4; shares ≈ 0). Demoted stocks are worth less (the companion volume's subject, plus the selection into demotion itself — demotion is triggered by low market cap, so this leg is partly the sorting variable, and we make no causal claim on it); they do not *trade* less. Symmetrically, the addition-side EUR gain (+0.498 median) is roughly half genuine turnover and the rest price and float legs. The Amihud price-impact companion moves with activity, in the right directions on both sides (additions -0.441 median log-illiquidity; demotions +0.499), with the demotion-side worsening again inheriting the price leg's mechanics.

What we do *not* claim is the full classic scoreboard. Hegde and McDermott's S&P 500 design found liquidity improvements for additions and deterioration for deletions across their measure set; on our price-free primary measure the deletion side simply does not deteriorate, and the honest European summary is one line: **additions gain ~26% genuine turnover at the mean; the demotion-side "liquidity loss" is price mechanics, not activity loss.**

7.2 Temporary or permanent

The horizon split — available for the EUR-volume measure only, since the decomposition legs have no horizon variants — shows the addition-side shift arriving immediately and persisting: the median EUR-volume ratio is +0.440 in the first post-month, +0.464 in months two–three, and +0.498 at the settled horizon ([+64, +273] days), whose bootstrap CI excludes zero ([+0.405, +0.599]). The demotion-side EUR decline likewise persists to the settled horizon (-0.445 median, CI [-0.505, -0.367]). We phrase persistence exactly this way — *the shift persists to the settled horizon* — and not as strengthening or decay across horizons, which the design does not test (the horizon CIs overlap); and we attach no horizon statement at all to the price-free turnover result.

7.3 Fine print

The EUR-denominated cells exclude seventeen events on seven stocks whose pre-2017 EUR prices carry the upstream unit defect of Section 2.3; the turnover and comovement analyses retain them (log returns and share ratios are unit-invariant), so the liquidity section's EUR sample (275–276 events per side) is deliberately not the comovement sample (285–289) nor quite the turnover sample (278). Volume *coverage* — 80% of scheduled additions and 87% of scheduled deletions clear the raw availability floor — is a data-availability statement, not an analysis n, and the two are never interchanged. Spread- and quote-based measures do not exist in this data layer, which bounds what we can say about liquidity *cost*; turnover, EUR volume and Amihud bound what happened to liquidity *activity* and *impact*.

8. The habitat postscript and the cross-section

8.1 The habitat test is dead on arrival

The bivariate design was built to give habitat its best shot: if additions migrate into an "index habitat", β_in should rise *and* β_out — the loading on non-index Europe — should fall. The second half never happens. The addition-side $\Delta\beta$_out is -0.072 at daily frequency (cluster-t -1.23; median -0.088, CI [-0.168, +0.003]) and -0.017 at weekly (t -0.16) — directional at best, never clearing either sanctioned read, and gone exactly where a habitat effect would have to live (lower frequencies). The paper accordingly claims no habitat migration; the cells live in Appendix B as the directional rows they are.

8.2 The cross-section is an absence of evidence, and says so

Who gets the (raw) shift? The pre-registered cross-sectional regression — the peer-netted DiD shift on log size, index weight, arbitrage risk and cycle passive-flow change, with country and supersector fixed effects — comes back null on all four characteristics (largest |cluster-t| 1.39). We report it as an *absence of evidence* and mean it: the realized sample is the weights-available subsample only — 103 additions spanning 2015–2019 across 19 cycles — and the specification carries 28 parameters against 19 clusters, so the cluster covariance is rank-deficient and HC3 is the primary standard error (a disclosed, conservative-for-a-null deviation from the paper's clustering doctrine). No individual fixed- effect loading is interpretable (the two sector cells with cluster-p <.05 are rank-deficiency artifacts, insignificant under HC3); a reduced-FE variant that restores degrees of freedom flips the signs of the headline loadings outright (index weight +398 $\rightarrow$ -347; arbitrage risk -51 $\rightarrow$ +31) while R^2 collapses from 0.27 to 0.05 — the full-FE loadings are noise, and we say so instead of reading tea leaves. The battery's characteristics-drift item is closed by the powered drift check of Section 4.3 (n = 267, 45 cycles, a single size-drift control — $K \ll G$), not by this regression; the two nulls are different objects and the paper keeps them apart.

9. Robustness

Every load-bearing cell in the paper was re-estimated under an eight-item variant matrix, and none of them moves. The price-return benchmark reproduces the net-return headline (addition raw +0.098 vs +0.100). Ending the pre-window five days before the selection-list date — sterilizing the front-running window paper 1 documents — changes the addition raw shift by 0.003. The Wednesday-close weekly compounding variant preserves the pooled weekly null, the middle-era addition cell, and the deletion nulls (one middle-era deletion cell flips its consistency flag on a near-zero estimate, 0.026 vs 0.001 — a labeled power artifact, not a sign change). Substituting exact-calendar peer windows into the DiD leaves it at +0.102; the small cycle-window subcells (n = 22–30) are labeled power artifacts where their flags fire. The peer pseudo-event placebo delivers the two rank statements already used in Section 4 — the raw addition median above all 2,000 cycle-conditional draws, the Dimson-corrected median mid-band — under a frozen label that its quantiles are not cluster-robust significance (a randomization band roughly 0.6× as wide as the cluster-robust interval), which is also why the deletion-side placebo quantile is never quoted against Section 6's null. The re-add augmentation (Section 6.2) retires the one conditioning channel the funnel could not disclose away. And the era and scrub reference blocks reproduce every headline cell within rounding . The full matrix, with the frozen interpretation labels, is Appendix C.

10. Limitations and data honesty

Five disclosures bound what this paper can claim, and we prefer stating them to being caught by them. **Membership reconstruction.** The daily membership calendar is a snapshot-anchored replay plus a signed five-defect override table; three entry/exit changes are absent from both the vendor change log and the parent event panel and are injected at selection-list-bracketed dates. One headline deletion (THG, 2021) opens its member spell on such a reconstructed addition; no other headline event does. **The pre-2017 unit defect.** Sixteen stocks carry EUR price/market-cap unit errors (×6–431) in the earliest vendor era. Log returns are immune; outside-index weights are built from selection-list capitalizations; one member-era weight is repaired against a selection-list anchor; seventeen EUR-volume events on seven stocks are excluded from EUR-denominated liquidity cells — so the comovement, turnover, and EUR-liquidity samples differ by construction and are never presented as one. One further volume series (DE101T, 2017) mixes exchange lines upstream; the affected cell's null is robust to its removal. **The C6 weights void.** The pre-registered cross-section runs only where index weights exist — 103 additions, 2015–2019, 19 cycles — and its nulls are statements about that subsample's power, nothing more. **Era seams.** The daily spine stitches three vendor eras (seams at 2017-10 and 2023-06) with different FX, corporate-action and coverage conventions; era is a first-class stratum throughout, the early era carries the non-synchronicity / data-quality qualification of Section 5.1 (the frozen-price form is excluded; the partial-adjustment form is not separable from economic non-synchronicity), and late-era anomalies are exploratory. Era cells rest on 14–16 review cycles, and every era statement in this paper is phrased at the power those numbers buy. **Scope.** No quote data exists in this layer, so liquidity *cost* (spreads) is out of scope; deletion results describe surviving demotions, with the survivorship funnel and its liquidity selection quantified in Section 2.5. Upstream defects identified here (the log omissions, the unit defect, the line-mixed volume series, the weights void) are documented for the parent data layer's maintainers and none of them, by the audits and exclusions above, carries a headline result.

11. What a practitioner should take from this

For a hedging or structured-products desk, the operational sentence is: **index membership moves the daily beta you measure, without moving the stock's weekly identity.** A post-inclusion daily-beta rise of +0.10 is real at the frequency where execution, margin, and short-horizon hedge ratios live — it reflects the stock genuinely trading faster and more synchronously with the benchmark — but it is a measurement-horizon phenomenon: at weekly frequency, and in lead–lag-corrected daily terms, the stock is the same animal it was before the review. A desk that re-estimates hedge ratios on daily windows after an index event is measuring the synchronicity change; a desk that hedges at longer horizons should not chase it. Symmetrically, for surviving demotions, exit from the index neither un-indexes the stock's risk (no detectable comovement decline) nor dries up its trading (no turnover loss): post-demotion depth is repriced, not withdrawn — the EUR-volume decline is a price-level effect.

For the passive-investing debate, Europe's most-tracked benchmark offers no evidence that indexing rewires fundamental comovement. The one era in which a battery-robust shift exists (2018–21, directional at 16 cycles) coincides with *flat* passive AUM, and the effect disappears precisely during the fastest passive growth on record — an alignment that the demand/habitat story predicts backwards on both ends. What indexing demonstrably changes is market microstructure: how fast a stock trades and how quickly its price absorbs index-level information. Combined with the companion volume — a price effect that has migrated upstream of the announcement, and a rebalancing cost of 2.8–5.1 basis points per year — the two papers describe an index whose mechanics are efficient, whose constituents' identities are stable, and whose remaining anomalies live in narrow, disclosed, and mostly pre-announcement windows.

Future work is disciplined by what we flagged and did not claim: the late-era weekly sign-flips ("passive-era beta compression") are two exploratory cells awaiting more cycles; quote-based liquidity awaits quote data; and the cross-sectional "who gets the shift" question awaits index weights beyond 2019.

Appendix A. Mid-cycle cohorts (descriptive)

Mid-cycle changes — fast entries, M&A deletions, spin-offs, forced deletions — never enter headline results (they differ in selection, timing and information environment, and their windows are anchored differently). Their descriptive comovement rows are reproduced from the results artifacts with the same column conventions and no inference language.

Appendix B. Bivariate and $\triangle R^2$ companions

Full bivariate tables ($\Delta\beta$_in, $\Delta\beta$_out by side and era, daily and weekly) and the Fisher-z ΔR^2 companion for the univariate design, from `comovement_results.csv` / `csw_results.csv` per the manifest's keying. The habitat rows of Section 8.1 appear here with their directional labels.

Appendix C. Robustness matrix and era/scrub references

The eight-item matrix of Section 9 verbatim from the manifest (`02_paper_numbers.md` §7), including the frozen placebo interpretation labels and the consistency flags with their power-artifact annotations, plus the era and scrubbed reference blocks.

Appendix D. Terminology and replication map

Artifact keys $\rightarrow$ paper terms (window names, spec names `uni_d`/`dimson1`/`uni_w`/`bi_d`, cohort labels), the window table of Section 2.4 with trading-day arithmetic, era boundaries, and the artifact inventory (`data/processed/*.csv`) with the manifest generator (`docs/whitepaper/tools/dump_numbers.py`) that reproduces every printed scalar.