

Quantifying the Alignment Gap Between Battery Arbitrage and Energy System Resilience

All site-resolvable GB grid-scale batteries in the Balancing Mechanism: 87 sites, 6,234 MW, observed 2018 to 2026 from public Elexon and system-operator data.

Energy prices already secure about four-fifths of the modelled battery’s high-load alignment. Response during scarcity and readiness before scarcity emerge as distinct dimensions of battery behaviour.

**UTILISATION** A profit-optimised 50 MW / 2 h reference battery delivers **83%** of its achievable energy in the highest-residual-load periods. Reaching the modelled maximum costs **£11/MW/day** in foregone market value. **UTILISATION** For the observed fleet, registered power, declared capability, notified schedules and acceptance-adjusted dispatch differ at every step: **54%** declared, 85% of that scheduled, 73% remaining after netting accepted bids and offers. **SCARCITY** Since April 2024 the fleet responds more strongly during scarcity but enters events with **less stored energy**: median charge at onset down **19 points**, or 6 under the alternative reconstruction.

1 · Population and measurement basis

**Utilisation: the system working hardest.** Top decile of residual load, 24.8 GW. A utilisation measure, not shortage. Sections 2 and 3, 60 days.

**Scarcity: the operator short of slack.** Loss-of-load probability at or above 10<sup>-4</sup>, 2,075 half-hours, with de-rated margin below 1 GW (n = 39) and Capacity Market Notices (n = 13) as rarer instruments. Section 4, 2018 to 2026.

No number from one basis is quoted against the other.

Scope

Sections 3 and 4 measure declared availability, notified schedules and acceptance-adjusted dispatch. Reliability outcomes, such as avoided unserved energy, are not estimated.

Constructing the population

No public register labels a Balancing Mechanism unit a battery. The population here is units declaring symmetric import and export capability, corroborated against public registers. 22 sites publish no energy capacity and are excluded from every state-of-charge figure.

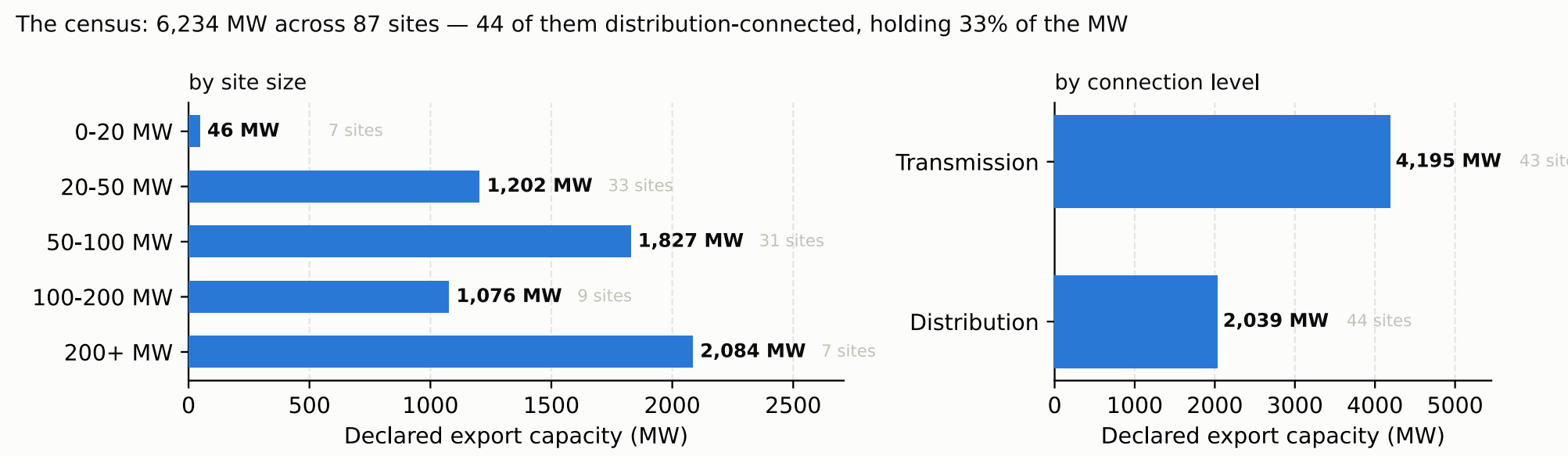


Figure 1. The census by site size and connection level. 44 distribution-connected sites hold 33% of declared capacity, as do the 7 largest sites.

Method

A day-ahead schedule on cleared prices, re-optimised intraday with perfect foresight. Efficiency 0.94 each way, 10% charge floor, 1.5 cycles per day, £2/MWh execution cost. Stored energy holds no terminal value, so modelled end-of-day depletion is an upper bound and the schedule is never rewarded for pre-event charge. Fleet behaviour comes from half-hourly physical notifications, with charge integrated rather than metered; Balancing Mechanism acceptances are netted as a measured sensitivity on both bases.

How other markets price tightness

Some European markets transmit system tightness into a price batteries schedule against, or pay directly for the charging behaviour that GB leaves unrewarded.

**Belgium computes a continuous scarcity price** from an operating-reserve demand curve, published D+1 since October 2019 and still run as a simulation. Its settled imbalance adder, alpha, is an incentivising component driven by system imbalance, not a scarcity adder.

**France pays for the charging behaviour directly.** TURPE 7, in force since August 2025, carries an **optional** and **locational** injection-withdrawal component: from 1 August 2026, for storage at HTA, HTB 1 and HTB 2, in CRE-designated constrained zones. Where it applies, batteries receive up to €69/MWh to charge in midday solar hours and are penalised for discharging outside them. Unlike Belgium's, this settles.

**GB applies neither instrument.** Scarcity reaches the battery only through cash-out. The REMA Summer Update of 10 July 2025 rejected zonal pricing for reformed national pricing, routing locational signals through network and connection-charge reform covering generation, **storage** and demand: the instrument TURPE 7 already applies to batteries.

Each instrument acts on a different gate. Scarcity signals affect scheduling, availability mechanisms declared capability, readiness products pre-event energy, and duration mechanisms the investment that sets the ceiling in Figure 3. Their welfare and investment effects are not estimated here.

Limitations

- **The utilisation basis is one 60-day window**, because no single public source spans the period: ENTSO-E stopped publishing GB day-ahead prices. A winter re-run puts the gap at 15%–22% against this window's 20%.
- **Charge is inferred, not metered.** Re-anchoring shifts levels by about ten points, which is why Figure 6 reports both schemes.
- **Top-decile load is utilisation, not shortage.** Published loss-of-load probability sums to 1.0 h per year against the 3-hour standard, an expectation rather than realised events. The reserve-scarcity component is episodic rather than absent: zero in 99.5% of this window, but £2,223/MWh in winter 2019–20, and still settling only through cash-out.
- **Timing is not identified**, and the dispatch-to-load correlation is mostly diurnal shape (+0.25 falling to +0.11). No causal claim is made about ancillary contracts or the break date.
- **Unchanged by** load threshold, cost model, day resampling, initial charge, and a period-resolved declared bound.

2 · The modelled incentive

**UTILISATION** · top-decile residual load · 2026-06-26 to 2026-08-24

Most peak delivery arrives without being paid for

The day-ahead objective carries no system-value term. The schedule it produces still delivers **83%** (74%–92%) of the top-decile energy this asset reaches when paid to prioritise it. Against the system-optimal schedule, an upper bound no payment attains, it leaves **20%** (8–29%) undelivered.

Scarcity settles outside the scheduling price

GB prices it as the value of lost load times loss-of-load probability, paid in cash-out, so it never enters the day-ahead objective. It was exactly zero in **99.5%** of settlement periods.

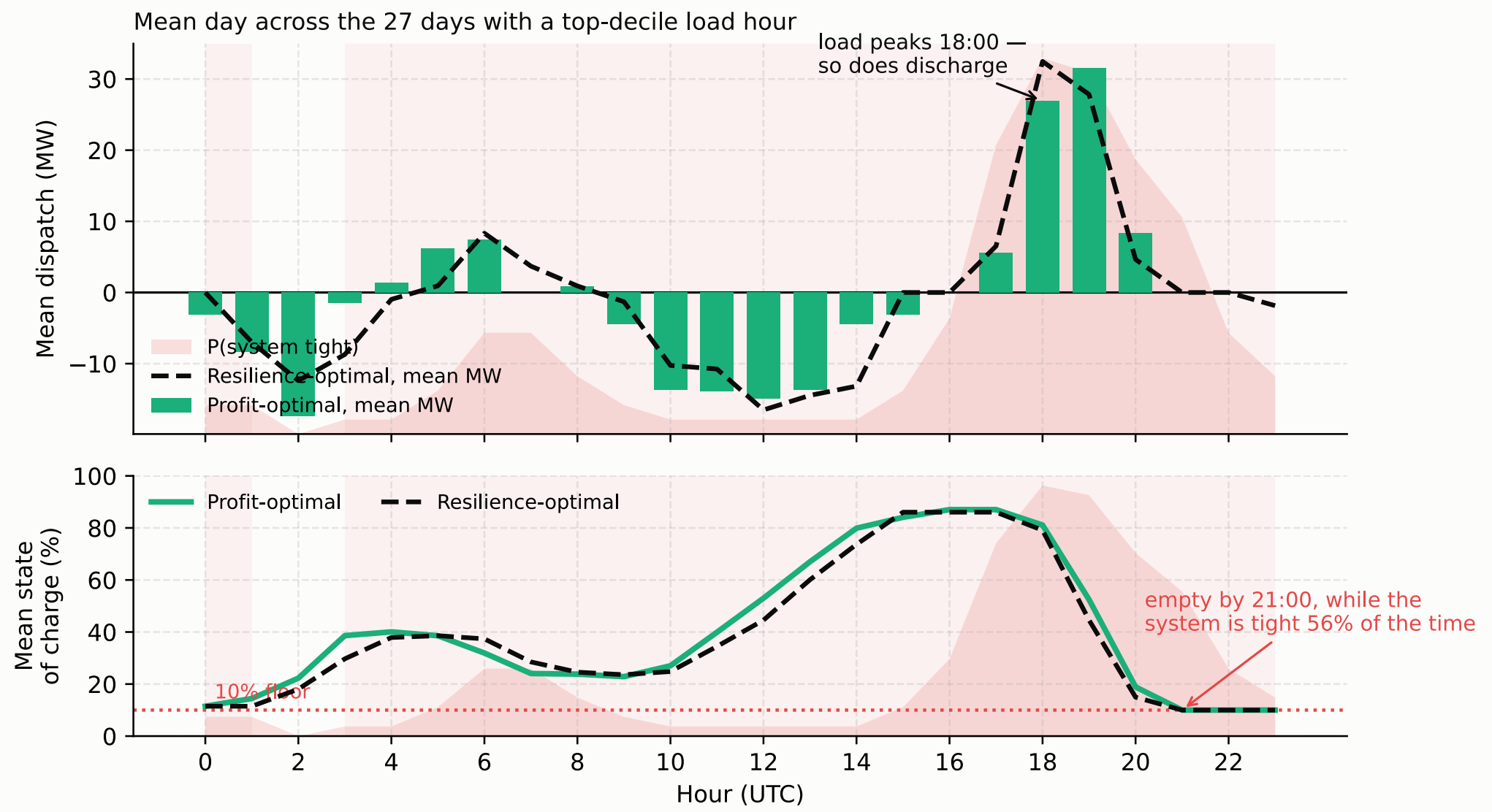


Figure 2. Discharge peaks at 19:00 with the load. Both stores reach the 10% floor by 21:00 while the system is still tight on 56% of days, so the evening shortfall is a duration limit rather than an incentive one.

The cost of alignment and the duration constraint

Buying every top-decile hour costs **£11/MW/day** (£6.8 to £16.4), which is 11% of bench-mark revenue. Buying the entire system-value schedule, a broader objective that also credits surplus absorption, costs **£39/MW/day** (£27–51/MW/day). Within this conditional grid, no payment lifts the two-hour asset past **43%** of what six hours of storage reaches, a level the longer asset attains at **89%** before any payment.

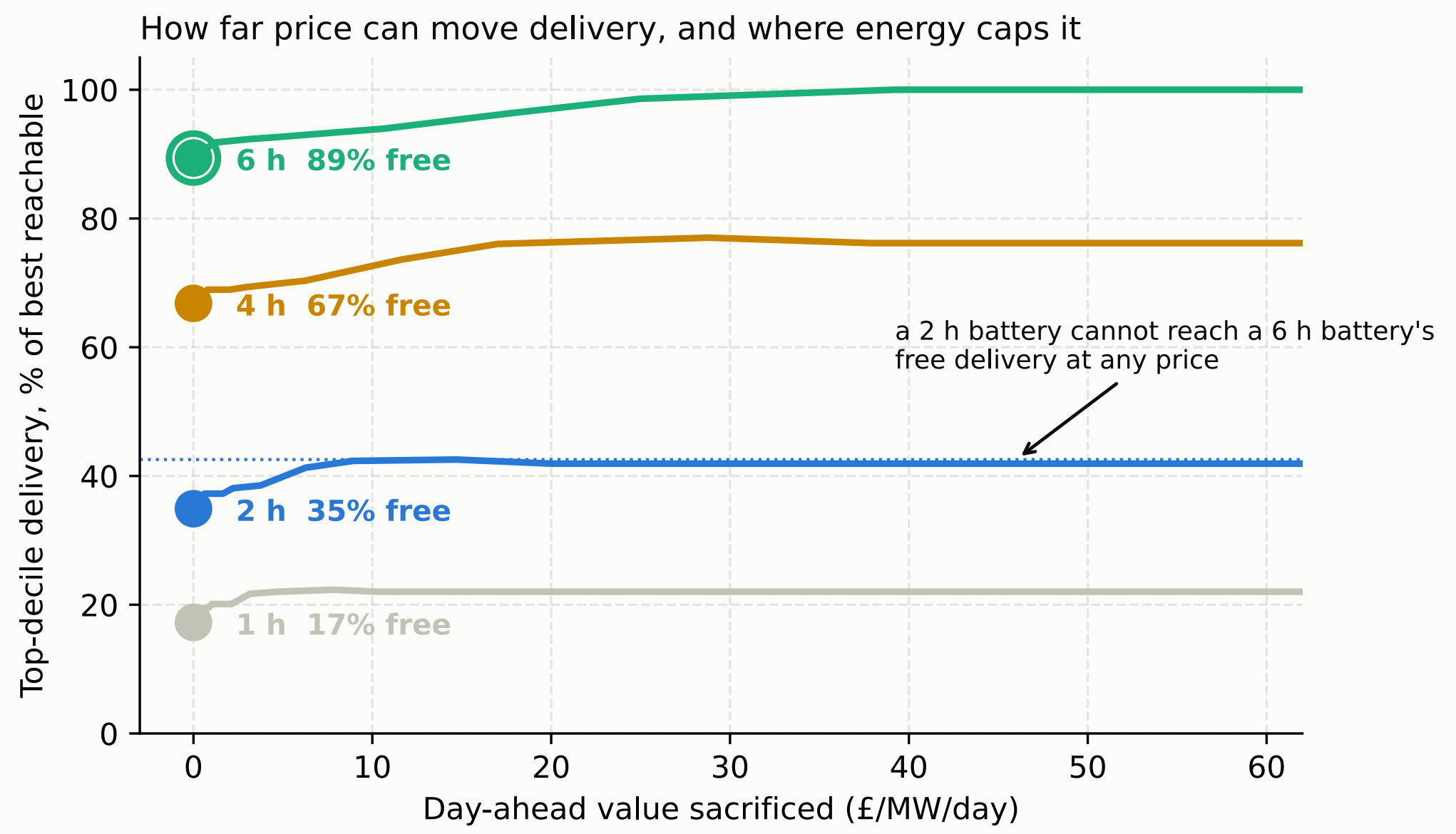


Figure 3. Each curve sweeps the weight on system value for one duration; the dot marks its profit schedule. Each day restarts at 50% SOC; inventory resets and extra energy capacity are unpriced. These are conditional daily comparisons, not continuous strategies or procurement bounds.

3 · The observed fleet on the same basis

**UTILISATION** · top-decile residual load · 2026-06-26 to 2026-08-24

The shortfall accumulates at three gates on this basis

**256,332 MWh**, the most registered power could deliver  
↓ **54% declared** available to the operator  
↓ **85%** of that **scheduled** in notifications  
↓ **73%** of that remaining after **netting accepted bids and offers**  
**86,056 MWh** in the acceptance-adjusted profile

The largest loss occurs at the first gate, and it does not concern dispatch: declared availability supports 54% of registered power. Given what was declared, the fleet scheduled **85%** of it, close to the modelled optimum's 80%. Netting acceptances then removes **27%**, concentrated in the evening peak, with the operator moving batteries **down**.

An availability requirement could affect the first gate but would not by itself address the third. Why declared availability falls so far below registered power is not identified in public data: outage, derating, reserved response headroom and commercial choice are not separable here.

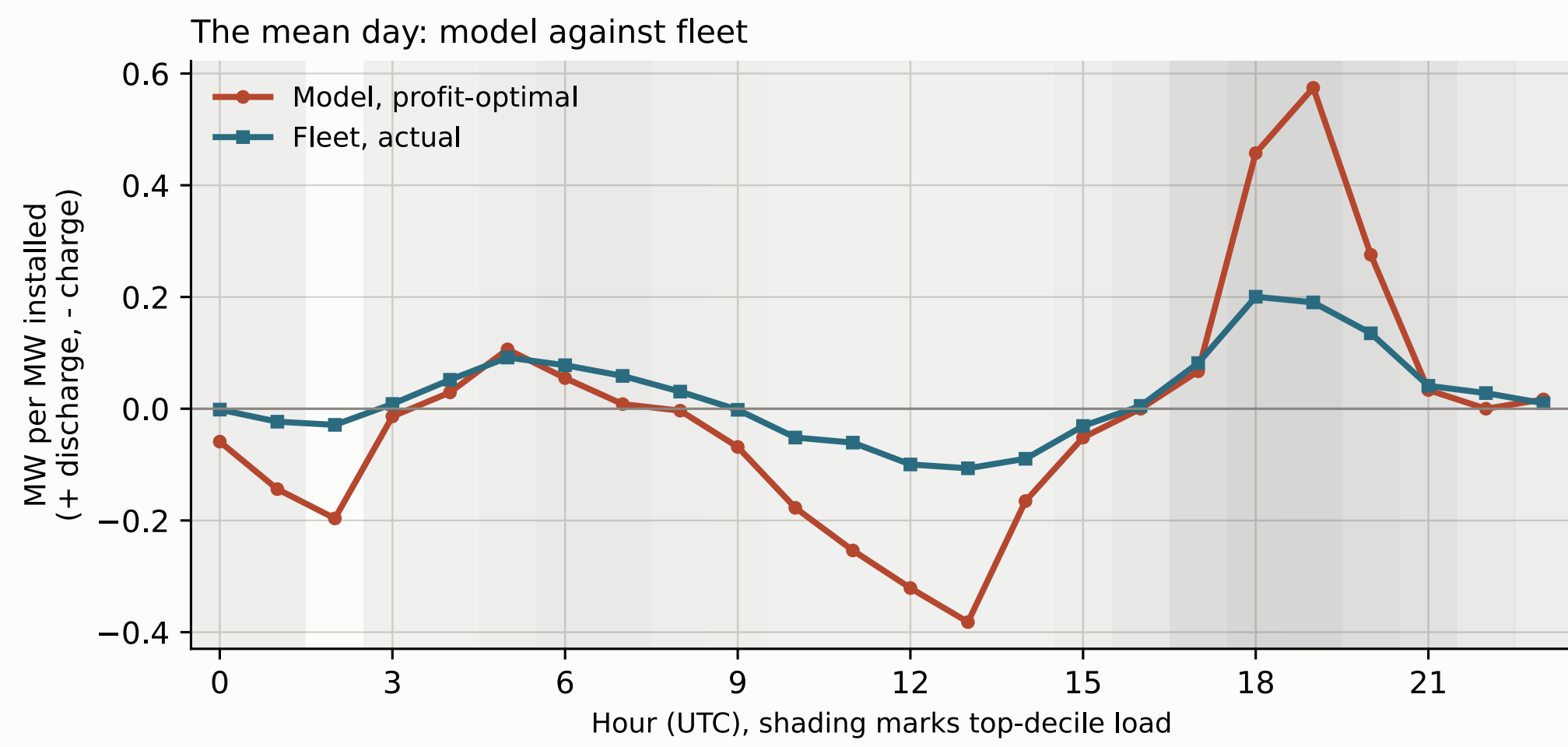


Figure 4. The mean day, per MW installed. The fleet discharges at the same hours as the model, at roughly a third of the depth. Neither series is metered output.

The most-ancillary quartile delivers about half as much

The least-ancillary quartile delivers **83%**, the most-ancillary **42%**: Spearman **-0.47** (p < 0.001, n = 49), surviving a duration control. The relationship is associational, on a summer window. Earnings stand in for committed capacity: auction volume sums a block count differing by an order of magnitude between services, so it tracks service mix.

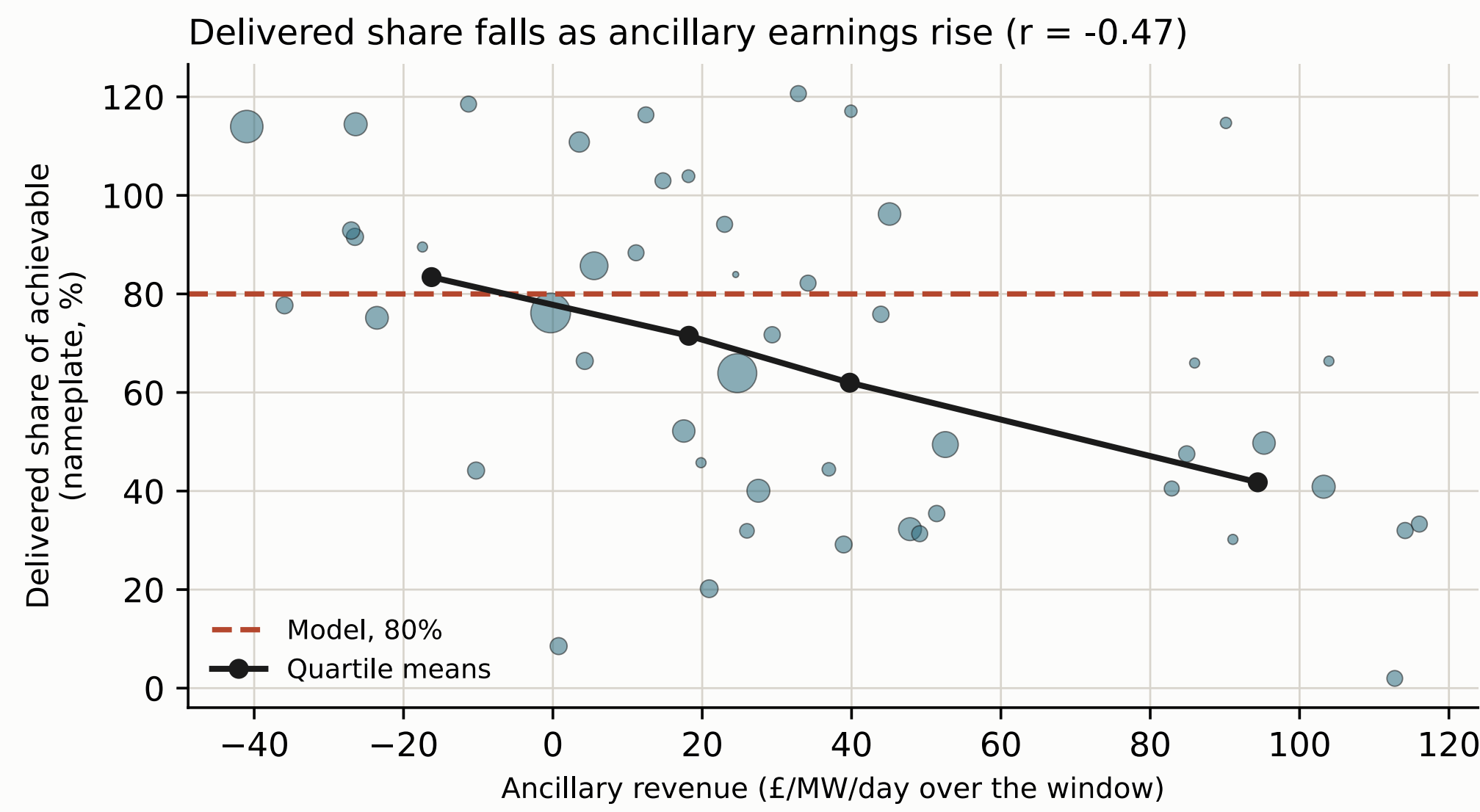


Figure 5. Quartile means run 83% to 42% against the model's 80%. Dot area is site capacity.

4 · Behaviour under operator scarcity

**SCARCITY** · the operator's own instruments · 2018 to 2026

The fleet usually responds during scarcity

In the flagged half-hours the fleet is net discharging 87% of the time, at **+0.060 MW per MW online** against +0.003 overall. Response is stronger from April 2024, **+0.117** against **+0.047** before it, and the difference holds when whole scarcity episodes are resampled rather than individual half-hours.

Readiness has changed

Across 428 events the median inferred charge at onset is **61%**, no fuller than on matched control days, and **48%** of usable energy is still in the fleet at its deepest point. Nothing pays a battery for arriving charged.

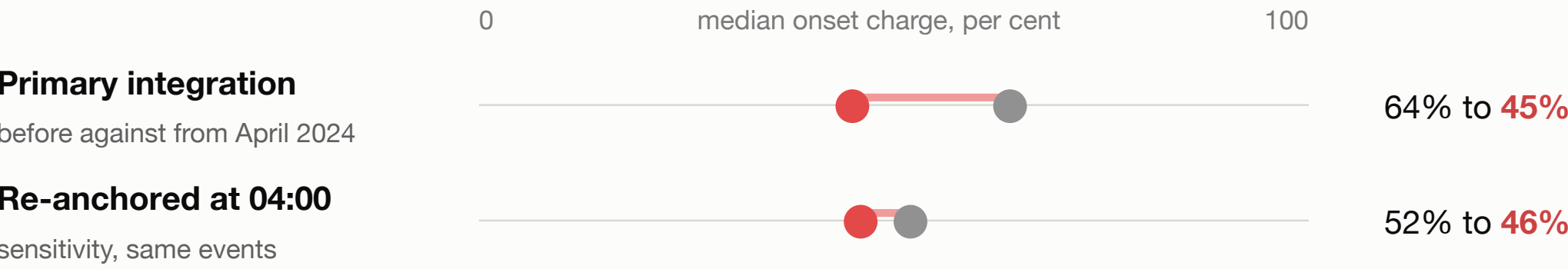


Figure 6. Median charge at onset, before April 2024 (n = 308) against from April 2024 (n = 37). The decline is 19 points under the primary integration, interval [-22, -13], and 6 points re-anchored: direction robust, magnitude scheme-dependent. Measured against arriving full, an upper bound on remediable energy, since a battery providing Response holds headroom in both directions. These 345 events use the loss-of-load rule alone, so both eras are scored on one instrument; the 428 and 45 quoted above bridge the union of all three instruments defined in Section 1.

Acceptances change little on this basis

Replacing notifications with acceptances cuts the scarcity response by **6%** and moves onset charge by two points. The same correction removes **27%** of delivered energy in Section 3, on a different window and stress definition. The instruction effect sits in high-load hours, not in scarcity.

The level change is robust but its timing is not

The rise is not composition: a fixed panel of 35 sites shifts **4.3x**, more than the fleet's **3.5x**. The October 2023 boundary was imposed rather than found, and searching every candidate quarter it ranks **18 of 29** on the Akaike information criterion.

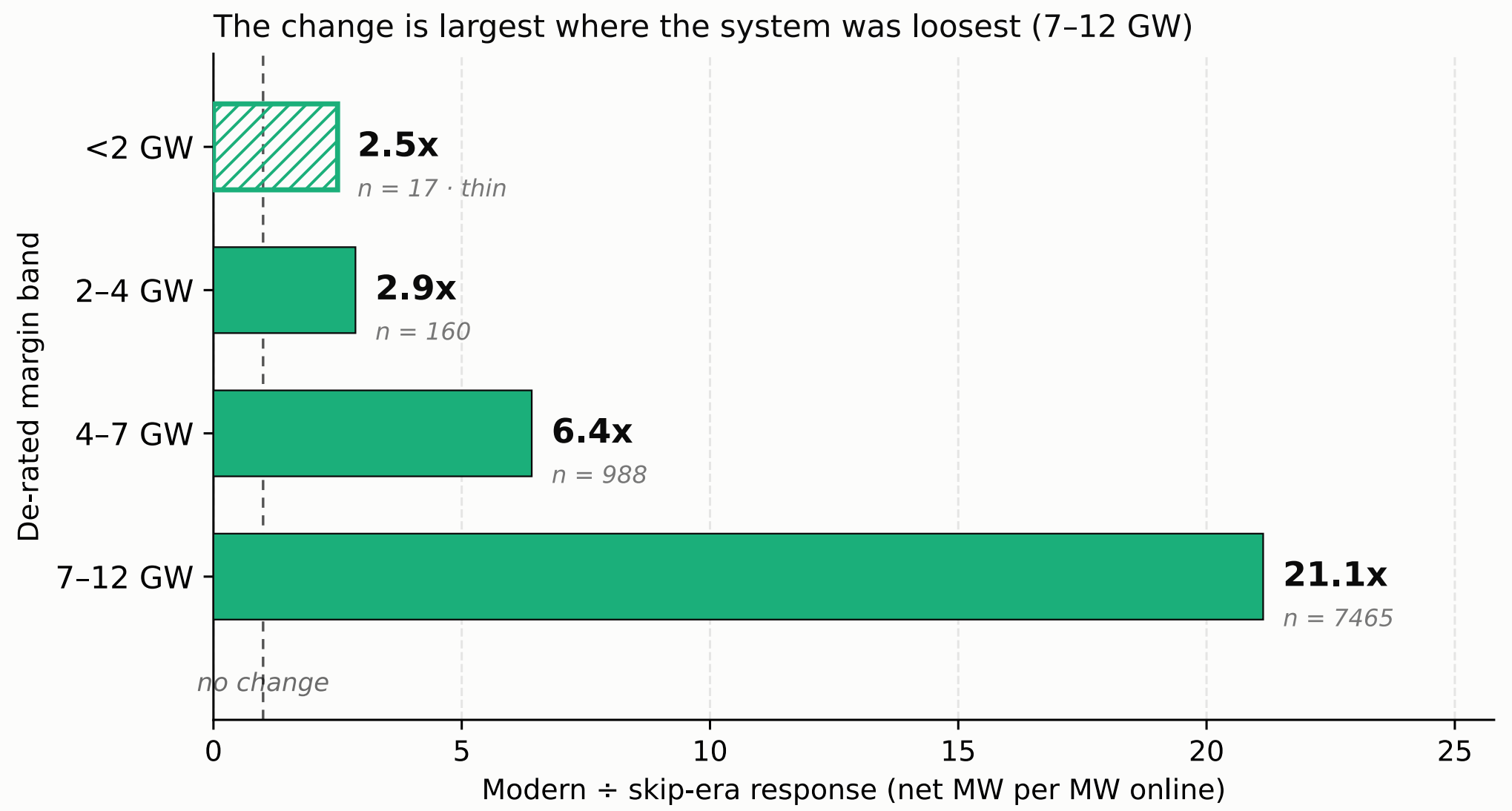


Figure 7. Later against earlier response under the same margin, by band. The change is largest where the system was loosest. Periods are split at an imposed date, so this is a contrast rather than an estimated break.

Sources & data

**Elexon BMRS/Insights:** physical notifications, export limits, bid-offer acceptances, cash-out.

**NESO Data Portal:** loss-of-load probability, de-rated margin, Capacity Market notices, TEC and Embedded registers, ancillary auction results.

**Nord Pool N2EX, ENTSO-E:** day-ahead prices. **PV\_Live:** solar outturn.

**Operator disclosures:** site energy capacity, each carrying its source and read date.

All feeds are public. The census, the dispatch engine and every figure here rebuild from them at [github.com/shaw029/power-trading](https://github.com/shaw029/power-trading), with the full methods and limitations.

**Policy, external, as at September 2026:** Papavasiliou et al., *The Electricity Journal* 33 (2020); CRE, *TURPE 7* (2025); DESNZ, *REMA Summer Update* (10 July 2025).



Live dashboard



Code & notebooks