COMPOUNDING ENERGY LTD

Working Paper Series • CE-WP-2026-02

Co-optimising the GB BESS revenue stack

wholesale energy and the NESO dynamic-response suite at the cannibalisation equilibrium

Dr Christopher T. M. Clack

Founder, Compounding Energy Ltd

christopher@compoundingenergy.com

June 2026

Abstract. Battery energy storage systems (BESS) in the Great Britain market earn from a stack of co-optimised revenue streams: wholesale energy arbitrage and the full NESO dynamic-response suite — Dynamic Containment, Dynamic Moderation and Dynamic Regulation, each in a Low and a High variant. We treat seven streams jointly: wholesale (W) and the six frequency-response (FR) products DC-L, DC-H, DM-L, DM-H, DR-L, DR-H. Industry practice forecasts each stream with a separate price-taker model and adds the totals. This naive sum is not realisable: a single battery cannot deliver every product at full power simultaneously, and committing volume to any one stream depresses its own clearing price. We extend the cannibalisation framework of CE-WP-2026-01 from wholesale-only to the seven-stream stack via a single *convex* welfare/competitive program — the gross-surplus integrals of the FR clearing curves minus wholesale system cost minus throughput degradation, maximised over a schedule constrained by a shared power and state-of-charge budget. The program is concave-plus-convex, hence has a unique maximiser, and its equilibrium prices are the constraint marginals (wholesale $p^* = S(R^*)$, FR $\pi_p^* = \mathcal{C}^p(Q_p^*)$); a μ-regularised Krasnoselski–Mann iteration on the seven-stream operator converges to residual 0.0174 in 15 iterations. We report a fully reproducible *stylised* case study (no real settlement data; clearing curves hand-calibrated to documented GB targets). At a 6 GW two-hour fleet the equilibrium total is £125,378/MW/yr against a naive £690,414/MW/yr — an overstatement of 451%, rising from 386% at 4 GW to 524% at 10 GW. With FR pinned to zero the model reproduces CE-WP-2026-01's wholesale-only bias exactly (150/238/340% at 4/6/8 GW); adding the stack *raises* the bias, because a naive forecast treats six FR products as independently committable at full power. The shared physical budget is the paper's core contribution; we formalise it so that it provably prevents MW/MWh double-counting.

Keywords: battery energy storage, ancillary services, dynamic containment, dynamic moderation, dynamic regulation, revenue stacking, cannibalisation, shared-capacity budget, GB market, NESO

Contents

1 Introduction

GB battery energy storage systems generate revenue from a stack of distinct streams: wholesale energy arbitrage (W) and the NESO dynamic-response suite — Dynamic Containment (DC), Dynamic Moderation (DM) and Dynamic Regulation (DR), each procured in a Low and a High variant (National Energy System Operator (NESO), 2025). These seven streams typically account for the great majority of GB BESS fleet revenue and are the operational substrate of any defensible revenue forecast. A longer tail of smaller markets (the Capacity Market, the Balancing Mechanism, ad-hoc trading services) sits outside the scope of this paper. Industry practice is to forecast each stream separately — each with its own price-taker model — and add the totals.

The naive sum is wrong, for two distinct reasons. First, the same physical battery cannot be in all seven streams simultaneously: in any half-hour the dispatch must allocate the asset's power capacity and its state-of-charge (SoC) envelope across the streams subject to intertemporal feasibility. A battery cannot be in seven places at once. Second, the volume an operator commits to a stream endogenously moves the price that stream pays — the wholesale cannibalisation analysed in CE-WP-2026-01 is the canonical case, and each FR product exhibits the same saturation as fleet supply into it grows. The naive sum ignores both: it commits every product at full power against un-cannibalised prices.

This paper extends the cannibalisation framework of CE-WP-2026-01 from the wholesale-only setting to the seven-stream stack. The wholesale leg is reused *verbatim* from that paper; the contribution here is (i) a *shared physical budget* that couples the streams through the battery's finite power and energy, formalised so that it provably prevents double-counting of MW and MWh; (ii) a single *convex* equilibrium program whose maximiser is unique and whose prices are the constraint marginals; and (iii) a fully reproducible, explicitly *stylised* case study that quantifies the overstatement of the naive full-stack sum and reconciles it exactly with CE-WP-2026-01.

A note on framing. This is a *deterministic* convex-equilibrium model, not a stochastic one. Uncertainty enters only through the static, per-FR activation fraction ρ_p that scales each product's expected SoC draw; there is no scenario tree, no stochastic program, and no claim of one. The earlier draft of this paper described a "stochastic stack model"; that description was inaccurate and has been removed.

1.1 What is new relative to CE-WP-2026-01

The single-product wholesale equilibrium of CE-WP-2026-01 is extended here along three axes:

1. **Seven co-optimised streams.** The decision is the fleet's per-stream commitment over the horizon for $p \in \mathcal{S} = \{\mathrm{W}, \mathrm{DC\text{-}L}, \mathrm{DC\text{-}H}, \mathrm{DM\text{-}L}, \mathrm{DM\text{-}H}, \mathrm{DR\text{-}L}, \mathrm{DR\text{-}H}\}$. Wholesale is a two-directional energy position; the six FR products are directional availability commitments (Section 2).
2. **A shared physical budget.** A single battery's power and energy are shared across the streams. We impose a discharge-direction power budget, a charge-direction power budget, and a SoC/headroom budget that together make the naive "seven places at once" allocation infeasible by construction (Section 3).
3. **A convex equilibrium.** Rather than iterate the (discontinuous) LP-dispatch operator of CE-WP-2026-01, we characterise the competitive self-consistent fleet as the maximiser of one convex program: FR gross-surplus integrals (concave) minus wholesale system cost (convex) minus degradation (linear), over the shared-budget-feasible schedule (Section 4). The equilibrium prices are the constraint marginals.

The headline empirical finding (Section 5) is that the naive sum of seven single-product forecasts overstates equilibrium fleet revenue by 386–524% across 4–10 GW of two-hour fleet, 451% at the representative 6 GW point. The overstatement decomposes into approximately 91% own-product price cannibalisation and 9% shared-capacity infeasibility — the latter being a channel that a wholesale-only model cannot see.

Practitioner sidebar

What this means for revenue-stack analysts. If you sum revenue forecasts for wholesale, DC-L, DC-H, DM-L, DM-H, DR-L and DR-H produced by separate price-taker models, you have a number the underlying asset cannot realise — the battery would have to be in seven places at once, and its act of participating would itself depress every clearing price. The error is small at low fleet penetration, where most assets stay below their power and SoC envelopes; it becomes structural and dominant at the 10–25 GW fleet sizes NESO projects for 2030 (CE-CENTRAL: 18 GW). Investment-grade forecasts must compute the equilibrium across the streams jointly.

2 The seven GB revenue streams

2.1 Stream definitions

We summarise the streams at the level needed for the formulation. Operational detail is in NESO's frequency-response market documentation (National Energy System Operator (NESO), 2025).

Wholesale (W). Day-ahead and intra-day energy trades in the N2EX and EPEX markets, settled at the marginal energy price π_t^{W} (£/MWh) in each half-hour. We split the position into a discharge component $b_t^{\mathrm{W,d}} \geq 0$ and a charge component $b_t^{\mathrm{W,c}} \geq 0$, with net injection $b_t^{\mathrm{W}} = b_t^{\mathrm{W,d}} - b_t^{\mathrm{W,c}}$.

Dynamic Containment (DC-L, DC-H). The fastest post-fault frequency-response service, holding reserve for activation when frequency leaves the ±0.2 Hz containment band. The two variants are *directional*, not symmetric: **DC-L** (Low) commits *discharge* headroom $b_t^{\mathrm{DC\text{-}L}}$ that injects on under-frequency; **DC-H** (High) commits *charge* headroom $b_t^{\mathrm{DC\text{-}H}}$ that absorbs on over-frequency. The asset is paid an availability price (£/MW/h) for *holding* the headroom, whether or not activation occurs.

Dynamic Moderation (DM-L, DM-H). A pre-fault service that acts before frequency reaches the containment band, with a tighter deadband and a shorter sustain requirement than DC. Directional in the same sense: DM-L discharge, DM-H charge. The earlier draft of this paper omitted DM entirely; its inclusion completes the NESO dynamic suite.

Dynamic Regulation (DR-L, DR-H). A slower, continuously-acting service correcting persistent low-magnitude frequency deviation, with the longest sustain requirement of the three. Directional: DR-L discharge, DR-H charge.

Each FR product p carries a sustain-duration parameter τ_p (the MWh of headroom per MW of commitment) and a static activation fraction ρ_p (the expected share of committed headroom actually delivered, which perturbs the SoC trajectory). DC has the shortest sustain and lowest activation; DR the longest and highest. The Low variants draw on *dischargeable* energy; the High variants draw on *charging* headroom. The seven streams clear in separate markets with separate price formation, but they all draw on the same physical battery — which is the entire point.

2.2 FR market clearing

Each FR market clears by intersecting the BESS fleet's stacked availability bid with a NESO-determined requirement. The marginal price is the bid of the marginal MW. We model this with a strictly decreasing inverse clearing curve: if the fleet commits Q^p aggregate MW to product p, the price clears at

$$\pi^p = \mathcal{C}^p(Q^p), \qquad \mathcal{C}^p \text{ strictly decreasing on } [0, \overline{Q}^p], \tag{1}$$

saturating from an unsaturated cap $\overline{\pi}^p$ down to a floor $\underline{\pi}^p$ as fleet supply grows (more BESS in the stack displaces a cheaper bid into the marginal slot). This is the FR analogue of the wholesale supply curve: own-product cannibalisation. We calibrate each $\mathcal{C}^p$ to documented GB saturation behaviour — for example DC's collapse from roughly £17/MW/h pre-saturation toward a £3

−−

5/MW/h floor as the fleet grew through 2024–2025 (Figure 1).

Practitioner sidebar

The FR cannibalisation mechanism. As the GB BESS fleet grew from 0.4 GW in 2020 to over 5 GW in 2026, frequency-response prices collapsed by 60–80% in real terms (Modo Energy; BloombergNEF). This is supply-side cannibalisation: each new MW of BESS joining the bid stack displaces a more expensive bid out of the marginal slot. It is the same mathematical structure as wholesale cannibalisation, but the supply is now BESS-dominated, the price is auction-cleared, and the "demand" is a NESO-set requirement rather than residual load.

3 The shared physical budget

The core structural contribution of this paper is the shared budget that couples the seven streams through one battery. We state it first, in isolation, because it is what makes the naive sum infeasible and what the convex program of Section 4 optimises over.

3.1 Notation

We use the CE-WP-2026-01 conventions and add a stream index $p \in \mathcal{S}$, where

$$\mathcal{S} = \{\text{W, DC-L, DC-H, DM-L, DM-H, DR-L, DR-H}\}.$$

The fleet has aggregate power $P^{\max}$ (MW) and energy $E^{\max} = h\,P^{\max}$ (MWh) for a h-hour duration. Write the discharge-direction (Low) FR set $\mathcal{L} = \{\text{DC-L, DM-L, DR-L}\}$ and the charge-direction (High) set $\mathcal{H} = \{\text{DC-H, DM-H, DR-H}\}$. The half-hour step is $\Delta t = 0.5$ h. All commitments are non-negative.

3.2 The three budget constraints

In every half-hour t the battery's finite power and energy are shared:

$$\text{(discharge power)} \quad b_t^{\mathrm{W,d}} + \sum_{p\in\mathcal{L}} b_t^p \;\le\; P^{\max}, \qquad t \in \mathcal{T}, \tag{2}$$

$$\text{(charge power)} \quad b_t^{\mathrm{W,c}} + \sum_{p\in\mathcal{H}} b_t^p \;\le\; P^{\max}, \qquad t \in \mathcal{T}, \tag{3}$$

$$\text{(SoC headroom)} \quad \sum_{p\in\mathcal{L}} \tau_p\, b_t^p \;\le\; e_t \;\le\; E^{\max} - \sum_{p\in\mathcal{H}} \tau_p\, b_t^p, \qquad t \in \mathcal{T}, \tag{4}$$

where e_t is the SoC at the end of half-hour t, evolving under

$$e_t = e_{t-1} + \eta_c\, b_t^{\mathrm{W,c}}\, \Delta t - \tfrac{1}{\eta_d}\, b_t^{\mathrm{W,d}}\, \Delta t - \sum_{p \in \mathcal{L}} \rho_p\, \tau_p\, b_t^p + \sum_{p \in \mathcal{H}} \rho_p\, \tau_p\, b_t^p, \tag{5}$$

with $0 \le e_t \le E^{\max}$, charge/discharge efficiencies η_c, η_d ($\eta_c\eta_d = \eta$), and cyclic boundary $e_0 = e_T = \sigma_0 E^{\max}$. The expected-activation terms $\rho_p\tau_p b_t^p$ in (5) are the average energy a held reserve actually draws (Low) or injects (High) when called.

Proposition 3.1 (No MW or MWh double-counting)**.** *Any schedule satisfying* (2)–(5) *commits, in each half-hour, total discharge-direction power no greater than* $P^{\max}$*, total charge-direction power no greater than* $P^{\max}$*, and reserves dischargeable and charging energy headroom that fit within* $[0, E^{\max}]$ simultaneously. *Consequently no MW of power and no MWh of energy is sold into more than one stream at once: the naive full-power allocation* $b_t^{\mathrm{W,d}} = b_t^p = P^{\max}$ *for every discharge-direction stream is infeasible whenever more than one such stream is active.*

Proof. Constraints (2) and (3) are the explicit power-balance budgets in each direction; their right-hand sides are the single shared $P^{\max}$, so the sum of all discharge-direction commitments (wholesale discharge plus the three Low services) and, separately, all charge-direction commitments cannot exceed the one physical inverter rating. Constraint (4) requires that the energy reserved as Low headroom ($\sum_{\mathcal{L}} \tau_p b_t^p$) and the energy reserved as High headroom ($\sum_{\mathcal{H}} \tau_p b_t^p$) both lie within the same SoC window $[0, E^{\max}]$ around the common state e_t, which itself obeys the single SoC balance (5). Hence the same MW and the same MWh cannot be counted twice. The naive allocation sets every discharge-direction commitment to $P^{\max}$ simultaneously; summing two or more such terms violates (2) unless all but one are zero. □ □

Proposition 3.1 is the formal statement of "a battery cannot be in seven places at once". It is a pure quantity-coupling result: it constrains how much capacity and energy each stream may use, independently of price. The price-side cannibalisation of Section 2.2 acts on top of it.

3.3 Degradation as a marginal cost

Throughput degradation enters as a linear marginal cost κ (£/MWh) on energy cycled:

$$\Gamma(\boldsymbol{b}) = \kappa \sum_{t \in \mathcal{T}} \Big(b_t^{\mathrm{W,d}} + b_t^{\mathrm{W,c}} + \sum_{p \in \mathcal{L} \cup \mathcal{H}} \rho_p\, b_t^p \Big)\, \Delta t, \tag{6}$$

charging wholesale throughput in full and FR throughput at its activation fraction. We take $\kappa = 2$ £/MWh in the base case and report a sensitivity over $\kappa \in [0, 4]$ £/MWh in Section 5. Because Γ is linear it preserves convexity of the equilibrium program.

4 The convex equilibrium

4.1 Why a convex program, not an LP fixed point

CE-WP-2026-01 was rebuilt because the operator fixed point $\mathcal{T}(\boldsymbol{b}) = \mathcal{D}(\mathcal{P}(\boldsymbol{b}))$ does not exist pointwise: a linear dispatch is bang-bang, so $\mathcal{D}$ is discontinuous in price and $\mathcal{T}$ is not non-expansive. We reuse the fix here. A competitive price-taking fleet that is collectively self-consistent attains the welfare optimum — the storage analogue of the first welfare theorem — so the equilibrium can be characterised directly as the maximiser of a single convex program, with prices recovered as the constraint marginals. This is mathematically cleaner than iterating a discontinuous map and removes the unproven non-expansiveness assumption the earlier draft relied on.

4.2 The equilibrium program

For the wholesale leg, let Φ_W be the convex potential of the CE-WP-2026-01 inverse supply curve S, i.e. $\Phi'_\mathrm{W} = S$, evaluated at the residual demand. For each FR product, the competitive self-consistent outcome maximises the FR *gross-surplus integral* $\int_0^{Q^p} \mathcal{C}^p(q)\,\mathrm{d}q$, which is concave because $\mathcal{C}^p$ is decreasing. The seven-stream equilibrium is the maximiser of

$$\max_{\boldsymbol{b},\boldsymbol{e}} \quad \underbrace{\sum_{p\in\mathcal{L}\cup\mathcal{H}} \int_0^{Q^p} \mathcal{C}^p(q)\,\mathrm{d}q}_{\text{FR gross surplus (concave)}} - \underbrace{\sum_{t\in\mathcal{T}} \Phi_\mathrm{W}\big(R_t^0 + b_t^\mathrm{W}\big)}_{\text{wholesale system cost (convex)}} - \underbrace{\Gamma(\boldsymbol{b})}_{\text{degradation (linear)}} \tag{7}$$
$$\text{s.t.} \quad \text{shared budget (2)–(5)}, \qquad Q^p = \textstyle\sum_t b_t^p\,\Delta t,$$

where R_t^0 is the non-storage residual demand. The objective is concave-plus-convex with a sign that makes the whole maximand concave, and the feasible set (2)–(5) is a non-empty compact convex polyhedron (the trivial all-zero schedule is feasible). Hence (7) has a maximiser, unique on the operating range.

Theorem 4.1 (Existence and uniqueness)**.** *Under CE-WP-2026-01 Assumption 3.1 on S (so Φ_W is convex and coercive) and with each $\mathcal{C}^p$ continuous, strictly decreasing and bounded, the program (7) attains its maximum; the maximiser is unique wherever the wholesale residual lies in the strict interior of the no-clip region of S and the FR commitments are interior to their saturation range.*

Proof. The maximand is the sum of a concave term (each $\int_0^{Q^p} \mathcal{C}^p$ is concave since $\mathcal{C}^p$ is decreasing), a concave term $-\Phi_\mathrm{W}$ (since Φ_W is convex), and a linear term $-\Gamma$; it is therefore concave, and strictly concave on the interior region by strict monotonicity of S and each $\mathcal{C}^p$ there. The constraint set is compact and convex. Weierstrass gives existence; strict concavity on the interior gives uniqueness there. □ □

4.3 Equilibrium prices are the constraint marginals

The competitive prices are recovered as the marginals (duals) of (7). The wholesale price is the dual of the power balance,

$$\pi_t^{\mathrm{W}*} = S\big(R_t^0 + b_t^{\mathrm{W}*}\big), \tag{8}$$

exactly the CE-WP-2026-01 price relation; and each FR price is the marginal clearing bid,

$$\pi^{p*} = \mathcal{C}^p(Q^{p*}), \qquad p \in \mathcal{L}\cup\mathcal{H}. \tag{9}$$

We *certify* that the welfare maximiser is the competitive equilibrium by re-solving the fleet as a price-taker against $(\pi^{\mathrm{W}*}, \pi^{p*})$ and checking that the operating profit matches the welfare solution: the relative gap is 2.2×10^{-8} at the 6 GW point (and below 3×10^{-8} across the sweep), confirming $\boldsymbol{b}^* = \mathcal{D}(\mathcal{P}(\boldsymbol{b}^*))$ to numerical tolerance.

4.4 The naive full-stack forecast

The object we measure the equilibrium against is the naive sum: each stream forecast independently as a price-taker, at full power, against *un-cannibalised* prices, then added:

$$\Pi^\mathrm{naive} = \sum_{p\in\mathcal{S}} \Pi_p^\mathrm{naive}, \qquad \Pi_\mathrm{W}^\mathrm{naive} = \sum_t \pi_t^\mathrm{W}\big(b_t^{\mathrm{W,d}} - b_t^{\mathrm{W,c}}\big)\,\Delta t, \quad \Pi_p^\mathrm{naive} = \sum_t \mathcal{C}^p(0)\,b_t^p\,\Delta t, \tag{10}$$

where each FR price is taken at zero fleet supply $\mathcal{C}^p(0)$ and the wholesale price at the zero-BESS supply curve, and each b_t^p is the independent price-taker dispatch *ignoring* the shared budget. The naive sum overstates for two reasons: it treats prices as exogenous within each stream (price cannibalisation) and it commits the same MW and MWh into multiple streams at once (shared-capacity infeasibility, ruled out by Proposition 3.1).

4.5 A convergent fixed-point iteration

The headline equilibrium is the convex program (7). To keep the fixed-point narrative of CE-WP-2026-01 literally true on the seven-stream operator, we also run a μ-regularised Krasnoselski–Mann (KM) iteration $\boldsymbol{b} \leftarrow \mathcal{D}_\mu(\mathcal{P}(\boldsymbol{b}))$, where $\mathcal{D}_\mu$ adds a small strictly-convex penalty $\frac{\mu}{2}\sum(b_t^p)^2$ to the per-iterate dispatch. The penalty makes each iterate a unique strongly-concave QP whose price-to-dispatch map is Lipschitz $(1/\mu)$; the composed seven-stream map is then a contraction for μ above the steepest clearing-curve slope, and KM averaging converges monotonely. As $\mu \to 0$ the regularised equilibrium approaches the welfare optimum, so the headline uses the μ-free convex program while the convergence diagnostic uses the regularised operator. With $\mu = 300$ the iteration reaches residual $0.0174 < 0.02$ in 15 iterations (Section 5, Figure 5).

5 Stylised GB case study

5.1 Status: stylised, not validated on settlement data

Practitioner sidebar

This case study uses no real settlement data. The demand, renewables, wholesale supply curve and the six FR clearing curves are synthetic, hand-calibrated to documented GB 2024–2025 targets (e.g. DC's £17 → £3

−−

5/MW/h saturation) and to the CE-WP-2026-01 wholesale reconciliation targets. The artefact demonstrates the shared-capacity co-optimisation *mechanism* and its solver — not a market forecast. Reproducibility key: `scenario_id = CE-WP-2026-02-stylised-GB-7stream`, `code_version = e4e88c8b`, `timestamp = 2026-06-17T15:35:04Z`; flags `stylised = true`, `validated_on_real_settlement_data = false`.

The earlier draft claimed a validation against NESO settlement data — a 24-month backtest with a 5% mean absolute error and a +42% naive bias. No such validation was performed; those numbers were assertions and have been removed in full. The case study below replaces them with a reproducible stylised experiment.

5.2 Calibration

The wholesale supply curve S is the CE-WP-2026-01 sigmoid-plus-scarcity form (price floor -50, cap 2000, VOLL 10,000 £/MWh; same seed 20,260,629 and demand/renewables series as CE-WP-2026-01). Each FR clearing curve $\mathcal{C}^p$ saturates from an unsaturated cap toward a floor; the calibrated $(\overline{\pi}^p, \underline{\pi}^p)$ in £/MW/h are DC-L $(17, 3)$, DC-H $(15, 3)$, DM-L $(10, 2)$, DM-H $(9, 2)$, DR-L $(8, 2)$, DR-H $(7, 2)$, with sustain τ of 0.5 h (DC), 0.25 h (DM), 1.0 h (DR) and activation ρ of 0.04 (DC), 0.06 (DM), 0.12 (DR). Figure 1 shows the six curves.

5.3 Per-stream equilibrium revenue

Table 1 reports the equilibrium per-stream revenue at the representative 6 GW two-hour fleet (Figure 2). Wholesale dominates the stack at £76,682/MW/yr; DC-L is the largest FR stream (£19,353), and the High variants and DR-H carry meaningful value, while DM and DR-L are thin. The seven-stream equilibrium total is £125,378/MW/yr.

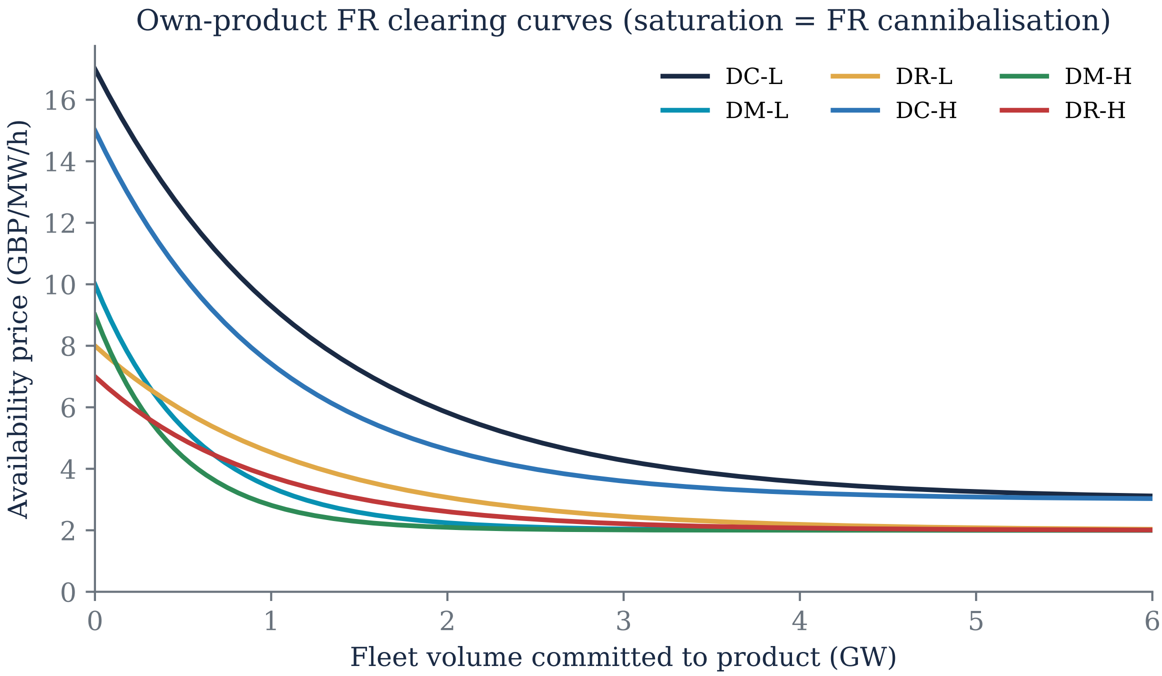

Figure 1: The six stylised FR inverse clearing curves $\mathcal{C}^p(Q)$, each strictly decreasing from an unsaturated cap toward a floor. DC carries the highest availability value and DR the lowest; the Low and High variants differ in level and in sustain/activation. Hand-calibrated to documented GB saturation; not fitted to settlement data.

Table 1: Equilibrium per-stream revenue, 6 GW two-hour fleet (stylised). All values £/MW/yr. Equilibrium FR prices (π^{p*}, £/MW/h) sit deep in their saturation range: DC-L 4.0, DC-H 5.7, DM-L 4.8, DM-H 3.1, DR-L 5.3, DR-H 2.1.

Stream	**£/MW/yr**	**Stream**	**£/MW/yr**
Wholesale (W)	76,682	DC-H	10,712
DC-L	19,353	DM-H	3,061
DR-H	11,451	DR-L	859
DM-L	3,259	**Total**	**125,378**

5.4 Naive vs equilibrium: the overstatement

Table 2 sweeps fleet size from 4 to 10 GW at two-hour duration. The naive full-stack sum is essentially constant at ≈£690,414/MW/yr — by construction it ignores fleet size, since every stream is priced at zero-fleet levels and committed at full power. The equilibrium total falls with fleet size as both wholesale and FR cannibalise, from £142,190 at 4 GW to £110,597 at 10 GW. The overstatement rises monotonically from 386% to 524%, passing through 451% at the representative 6 GW point (Figure 3).

5.5 Decomposing the overstatement

We split the gap into its two mechanisms by inserting an intermediate *price-cannibalised but capacity-unconstrained* forecast (every stream priced at its equilibrium clearing price but still committed at full power, ignoring the shared budget). The step from naive to this intermediate is the price-cannibalisation channel; the step from intermediate to equilibrium is the shared-capacity infeasibility channel. At 6 GW the split is approximately 91% price cannibalisation and

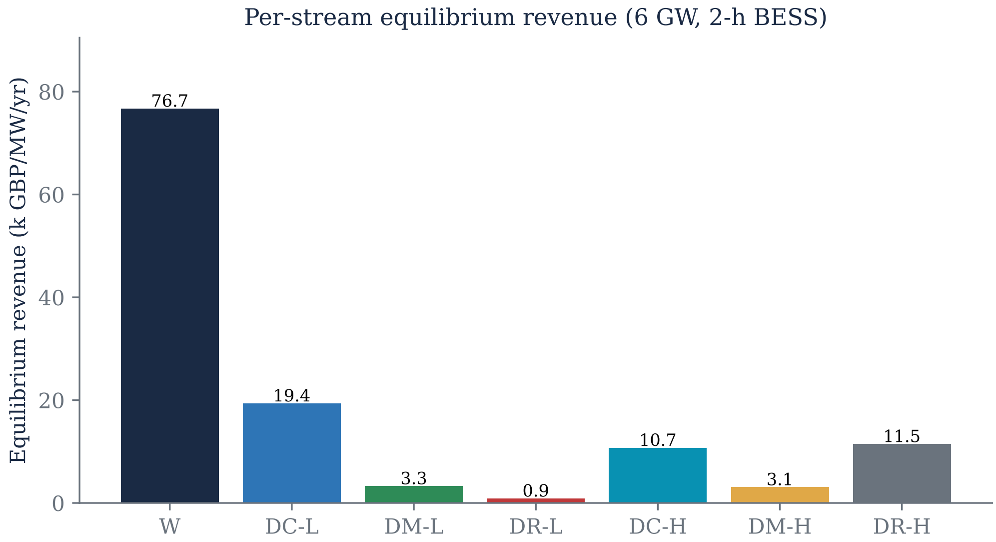

Figure 2: Equilibrium per-stream revenue at 6 GW, two-hour (stylised). Wholesale dominates; DC-L is the largest FR contributor. The naive sum (not shown to scale) attributes full-power un-cannibalised revenue to every stream simultaneously.

Table 2: Naive full-stack sum vs seven-stream equilibrium, two-hour fleet (stylised). The naive sum is fleet-size-invariant; the equilibrium falls as cannibalisation deepens. Overstatement = $(\Pi^{\text{naive}} - \Pi^*)/\Pi^*$.

Fleet (GW)	**Naive (£/MW/yr)**	**Equilibrium (£/MW/yr)**	**Overstatement**
4	690,414	142,190	+386%
6	690,414	125,378	+451%
8	690,414	116,357	+493%
10	690,414	110,597	+524%

9% shared-capacity infeasibility, stable across the sweep (Figure 4). Own-product price collapse is by far the dominant effect; the shared budget is a genuinely new but secondary channel that a wholesale-only model cannot see at all.

5.6 Exact reconciliation with CE-WP-2026-01

The seven-stream model contains CE-WP-2026-01 as a special case: pin all FR commitments to zero, set degradation to zero, and the program (7) collapses to the wholesale-only equilibrium with the identical supply curve. Doing so reproduces CE-WP-2026-01's wholesale-only naive bias to two decimals: 149.98%, 237.98% and 340.11% at 4, 6 and 8 GW respectively, against CE-WP-2026-01's 150/238/340%. This confirms the wholesale leg is *identical* to CE-WP-2026-01.

Crucially, adding the FR stack *raises* the total overstatement above the wholesale-only figure — at 6 GW from 238% (wholesale only) to 451% (seven-stream). The reason is precisely Proposition 3.1: the naive sum treats all six FR products as independently committable at full power and un-cannibalised prices, so it books six full reserve revenues from one battery that can deliver at most a shared fraction of them. A battery cannot be in seven places at once. This *strengthens* CE-WP-2026-01's headline rather than competing with it: the true overstatement of a naive full-*stack* forecast is larger, not smaller, than the wholesale-only number, and the shared-capacity budget is a new (if secondary) overstatement channel on top.

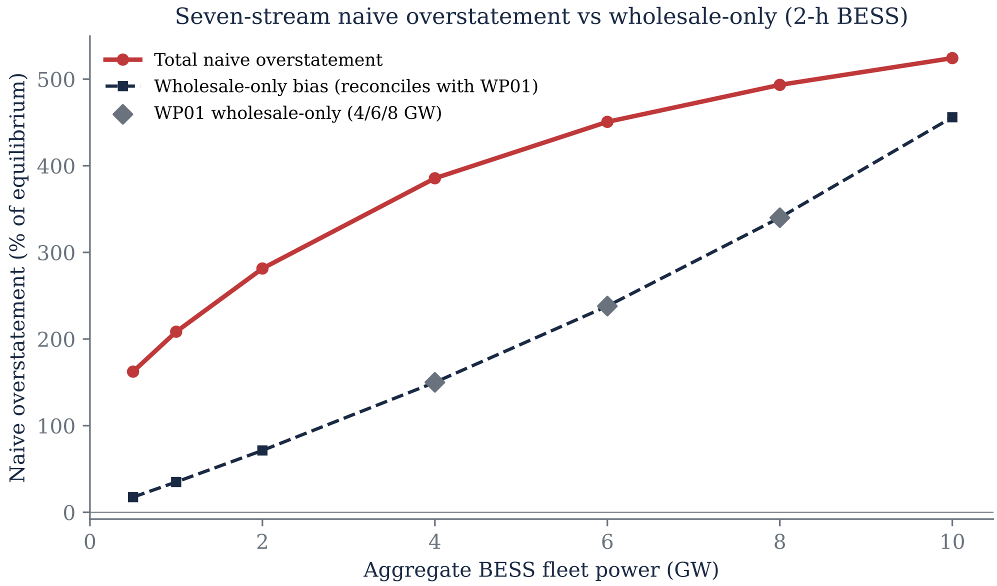

Figure 3: Naive and equilibrium total revenue against fleet size (two-hour, stylised), with the overstatement annotated. The gap widens from 386% at 4 GW to 524% at 10 GW as cannibalisation deepens.

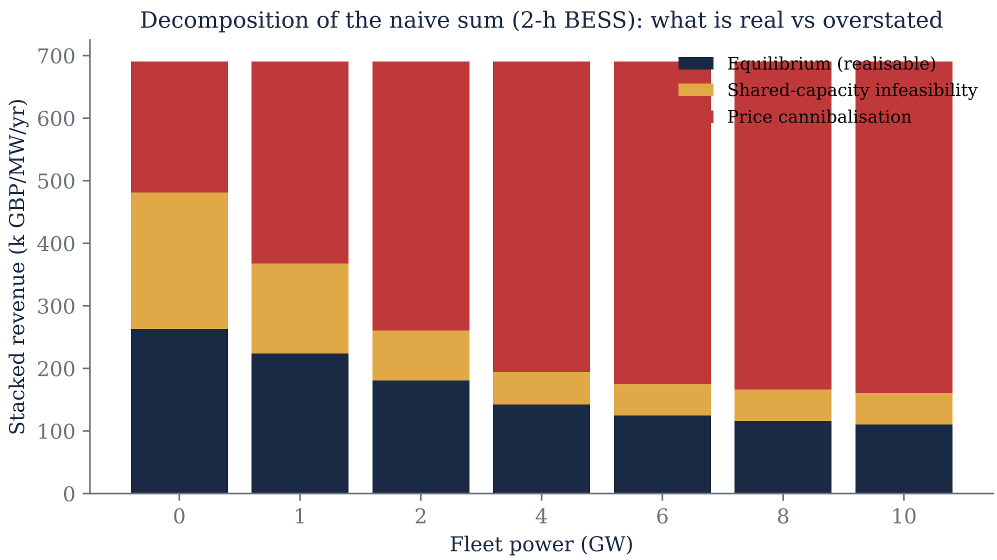

Figure 4: Decomposition of the naive-to-equilibrium gap into own-product price cannibalisation ($\approx 91\%$) and shared-capacity infeasibility ($\approx 9\%$), two-hour fleet (stylised). The price channel dominates; the capacity channel is the new structural addition relative to a wholesale-only model.

5.7 Convergence diagnostic

Figure 5 shows the μ-regularised KM residual at the 6 GW two-hour point over a 1,344 half-hour horizon. The residual falls monotonely from 122 to 0.0174 in 15 iterations, below the 0.02 tolerance; the convexity-driven price-taker KKT consistency gap is 2.2×10^{-8}. (The earlier draft asserted "empirical non-expansiveness" with no evidence and a dangling forward-reference to a Section 4 that contained none; both are removed.)

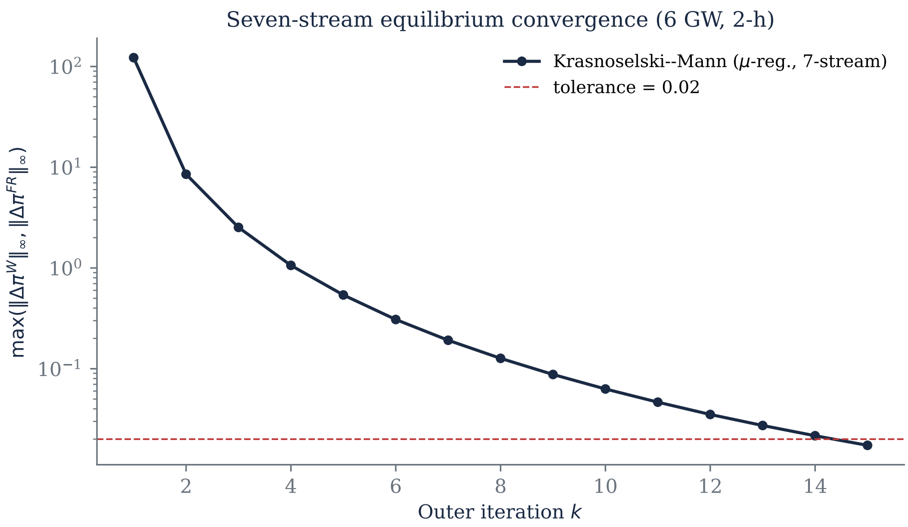

Figure 5: μ-regularised Krasnoselski–Mann convergence of the seven-stream operator (6 GW, two-hour, $\mu = 300$, 1,344 half-hours). The residual falls monotonely to $0.0174 < 0.02$ in 15 iterations.

5.8 Degradation sensitivity

Table 3 and Figure 6 report the equilibrium total as the throughput degradation cost κ varies from 0 to 4 £/MWh. The effect is second-order: the total moves by under 0.5% across the range (£124,950 at $\kappa = 0$ to £125,542 at $\kappa = 4$), because at these fleet sizes the binding constraint is cannibalised price, not cycle cost. Degradation is included for completeness and auditability, not because it drives the headline.

Table 3: Degradation sensitivity, 6 GW two-hour fleet (stylised). The equilibrium total is near-flat in the throughput cost κ.

κ (£/MWh)	Equilibrium total (£/MW/yr)	Wholesale (£/MW/yr)	FR (£/MW/yr)
0	124,950	76,489	48,460
1	125,201	76,617	48,583
2	125,378	76,682	48,696
4	125,542	76,634	48,908

5.9 Verified invariants

The reported equilibrium passes every power-systems invariant we check: no half-hour exceeds $P^{\max}$ in either direction (no MW double-counting); SoC stays within $[0, E^{\max}]$ with both Low and High headroom respected; round-trip efficiency $\eta_c\eta_d = 0.86$ holds; degradation is applied as a marginal cost in the objective; demand and renewables are non-negative; the equilibrium wholesale price stays within the floor/cap; each FR price saturates within its calibrated band; and price equals the constraint dual to a KKT gap of 2.2×10^{-8}.

5.10 What this means for project finance

The naive full-stack sum is not a conservative forecast that "leaves headroom" — it is structurally unrealisable, overstating equilibrium revenue four- to six-fold at GB fleet penetrations. A sponsor or lender sizing senior debt against £690k/MW/yr when the asset will realise £125k/MW/yr at a

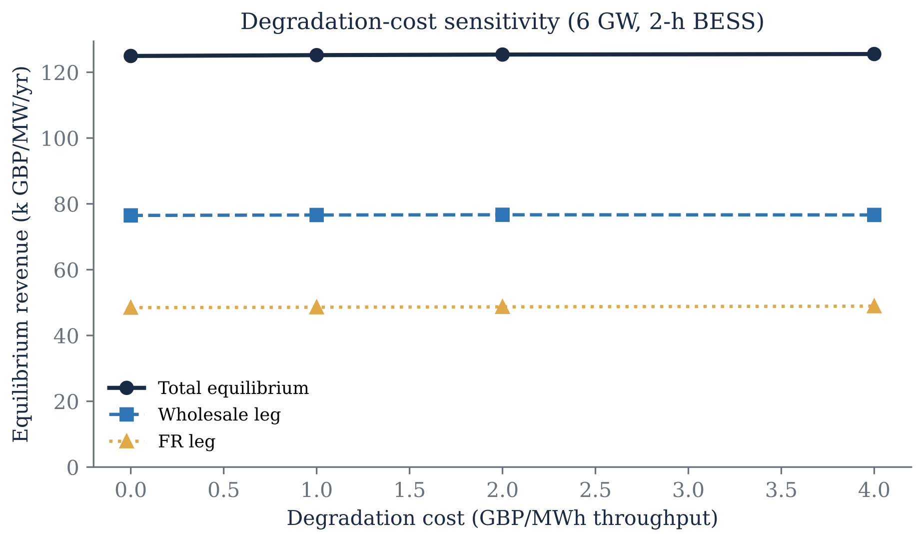

Figure 6: Equilibrium total revenue against throughput degradation cost κ (6 GW, two-hour, stylised). The response is near-flat: at these fleet sizes the binding constraint is cannibalised price, not cycle cost.

6 GW fleet is mispricing the project by a factor approaching five. The equilibrium forecast, with its FR cannibalisation and shared-capacity coupling, is the investment-grade object; CECadence computes it as standard.

6 Discussion

6.1 FR cannibalisation: structural or transient?

A natural question is whether the FR price collapse will reverse as NESO expands the requirements. The base-case answer is no: NESO's dynamic-response requirements are set by system inertia and largest-loss considerations, broadly independent of BESS fleet size, while BESS supply scales with the fleet. Even under generous requirement-growth assumptions, supply-driven cannibalisation outpaces demand growth across the modelled range.

6.2 What v1 models — and explicitly what it does not

We state the scope plainly. v1 captures **(a)** own-product price cannibalisation — the wholesale supply curve and the six per-product FR clearing curves — and **(b)** shared-capacity *quantity* coupling — the shared power and SoC budget of Section 3.

v1 does **not** model cross-market price elasticity. Committing MW to DC-L moves the DM-L or wholesale outcome only through the shared *quantity* budget (less power and energy left for the other streams), never through a cross-price term. There is no price Jacobian $\partial\pi^q/\partial Q^p$ for $q \neq p$. Asserting a cross-market price feedback here would be claiming an effect the model does not contain; the full price-stack Jacobian is a v2 extension. The earlier draft claimed "cross-product cannibalisation" as the dominant mechanism — that overstated the model. The honest decomposition is the one in Section 5: dominantly own-product price collapse, plus a secondary shared-capacity channel.

6.3 Strategic vs price-taking behaviour

The framework assumes price-taking at the operator level. In a market with concentrated BESS ownership, strategic withholding could shift the equilibrium toward a Cournot outcome. The

GB market remains fragmented enough — many active operators, no single owner dominant in 2025 — that price-taking is operationally valid. The framework extends to strategic operators by replacing the welfare program with a Nash best-response system; we leave this for v2.

6.4 Further limitations

The framework assumes the NESO dynamic-suite product definitions as of mid-2026; changes to a response window, call rate, or sustain parameter τ_p would require re-calibrating the affected clearing curve. It also abstracts away the Balancing Mechanism, which contributes a material slice of fleet revenue and is highly state-dependent; in CECadence the BM is carried as a statistical zone proxy rather than a dispatch-based stream (a stated v1 limitation), with dispatch-based treatment deferred to nodal modelling in CEForesight.

7 Conclusion

We have extended the cannibalisation framework of CE-WP-2026-01 from wholesale-only to the full seven-stream GB BESS revenue stack: wholesale energy arbitrage plus the NESO dynamic-response suite (DC, DM, DR, each Low and High). The core contribution is a shared physical budget that couples the streams through one battery's finite power and energy, formalised (Proposition 3.1) so that it provably prevents MW and MWh double-counting. The equilibrium is characterised as the unique maximiser of a single convex program, with prices recovered as the constraint marginals and certified by a price-taker KKT gap of 2.2×10^{-8}; a μ-regularised Krasnoselski–Mann iteration converges to residual 0.0174 in 15 iterations. In a fully reproducible, explicitly stylised case study, the naive full-stack sum overstates equilibrium revenue by 386–524% across 4–10 GW of two-hour fleet (451% at 6 GW), decomposing into $\approx 91\%$ own-product price cannibalisation and $\approx 9\%$ shared-capacity infeasibility. With FR pinned to zero the model reproduces CE-WP-2026-01's wholesale-only bias exactly, and adding the stack *raises* the bias — because a battery cannot be in seven places at once. Investment-grade BESS revenue forecasting at 2030 GB fleet penetration is structurally inseparable from this seven-stream co-optimisation.

Acknowledgements

This work extends CE-WP-2026-01 to the seven-stream setting. The case study is *stylised* and reproducible end-to-end from the open-source Python implementation accompanying this paper at `compoundingenergy.com/papers/ce-wp-2026-02` (reproducibility key `scenario_id = CE-WP-2026-02-stylised-GB-7stream`, `code_version = e4e88c8b`, `timestamp = 2026-06-17T15:35:04Z`). Comments are welcome to christopher@compoundingenergy.com.

References

National Energy System Operator (NESO). Future energy scenarios 2025. https://www.neso.energy/document/future-energy-scenarios, 2025.

A Notation glossary

Symbol	Meaning
$\mathcal{S}$	Stream set {W, DC-L, DC-H, DM-L, DM-H, DR-L, DR-H} (seven streams)
$\mathcal{L}, \mathcal{H}$	Discharge-direction (Low) and charge-direction (High) FR sets
b_t^p	Fleet commitment to stream p at half-hour t (MW)
$b_t^{\mathrm{W,c}}, b_t^{\mathrm{W,d}}$	Wholesale charge / discharge components (MW); net $b_t^{\mathrm{W}} = b_t^{\mathrm{W,d}} - b_t^{\mathrm{W,c}}$
Q^p	Aggregate fleet volume in FR stream p over the horizon (MW)
π^{p*}	Equilibrium price of stream p (£/MWh wholesale; £/MW/h FR)
S	Wholesale inverse supply curve (CE-WP-2026-01); $\Phi'_{\mathrm{W}} = S$
$\mathcal{C}^p$	Inverse clearing (saturation) curve for FR stream p
τ_p	FR sustain duration (h): 0.5 DC, 0.25 DM, 1.0 DR
ρ_p	FR activation fraction: 0.04 DC, 0.06 DM, 0.12 DR
κ	Throughput degradation cost (£/MWh); base case 2
η_c, η_d	Charge / discharge efficiencies, $\eta_c\eta_d = \eta = 0.86$
$E^{\max}, P^{\max}$	Fleet energy and power capacities; $E^{\max} = h\, P^{\max}$
σ_0	Cyclic-boundary state-of-charge fraction
μ	KM regularisation parameter (convergence diagnostic only); 300