Compute Futures: Who Is Left Exposed?*

Rubén Fernández-Fuertes[1], Vincent Grégoire[1]

[1]HEC Montréal

August 3, 2026

CME Group and Intercontinental Exchange have announced cash-settled compute futures on rival GPU rental-rate indices before any trading has revealed which benchmark hedgers will use. A *locked* seller earns one configuration's rental rate; a *flexible* buyer faces a basket of rates. That asymmetry limits any benchmark: an index built for one side leaves the other exposed. Using illustrative inputs, a benchmark aligned with the buyer's basket hedges 99.7% of that buyer's variance but 28.2% of seller variance. Two H100 indices from different publishers correlate too weakly to hedge each other, so which one settles is part of the contract.

JEL Classification: G13, G12, G32, Q41.
Keywords: compute futures, commodity derivatives, basis risk, hedging effectiveness, contract design, financial innovation.

* This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program and from the Social Sciences and Humanities Research Council.

1. Introduction

Training and operating artificial intelligence models requires compute capacity whose financing now extends beyond technology firms' internal cash flows (Aldasoro, Doerr, and Rees, 2026). Higher rental rates raise the cost of marginal capacity for an AI laboratory; lower rates cut a capacity provider's revenue. To transfer that risk, CME Group and Intercontinental Exchange have announced competing cash-settled compute futures contracts on indices from Silicon Data and Ornn, respectively.[1] Each contract settles on a benchmark of *graphics processing unit* (GPU) rental rates, quoted per GPU-hour, that must be specified prior to trading. Yet neither market convention nor empirical precedent constrains this selection, leaving the choice to the issuer's discretion. That discretion takes on economic urgency because firms financing marginal capacity remain exposed to spot rates they cannot fix. Nominally identical devices differ in effective performance, and new generations change what is available to rent. When market participants rent different hardware, no single settlement benchmark can track all of them equally well. Benchmark choice therefore allocates basis risk across participants. We ask how differences between buyer and seller exposures restrict that choice.

Consider a "neocloud", i.e., a specialized provider renting out GPU capacity. It is built around a particular tuple of GPU generation, interconnect, region, and sales channel, which we call a *configuration*. Its revenue depends on the rental rate of that configuration. An AI laboratory buying marginal spot capacity may instead shift expenditure across available configurations, so its cost can behave like a *basket*, an expenditure-weighted combination of configuration rental rates. We propose a model that formalizes this contrast: the buyer holds a basket; the seller remains tied to one configuration. Under the model's separation of common and configuration-specific shocks, and

[1] See the CME Group press release of May 12, 2026, and the ICE announcement of the ICE–Ornn contracts.

with buyer expenditure dispersed across configurations, configuration-specific shocks wash out of the buyer's expenditure basket but remain in the seller's revenue exposure.

That asymmetry restricts contract design: each contract settles on one weighted combination of configuration prices. Doubling those weights merely halves the position a hedger must take, so a contract matters only through its relative weights; we call this scale-free combination its *direction*. The buyer side supplies at most one direction; any further directions must come from heterogeneous seller exposures. Covering another direction takes another contract, which splits hedging interest across a thinner book. What bounds the exchange's menu is that liquidity cost rather than its fee for listing: the menu stops where a further contract no longer removes enough variance to pay for the interest it divides. In equilibrium, each contract's premium equals its net hedging pressure, the summed covariance of its payoff with hedgers' endowments, divided by aggregate risk tolerance.

One candidate benchmark is always available. Buyer expenditure-share weights define a feasible index, one an exchange can construct from configuration prices, whatever share of the hedger population sellers make up. When buyers are the only hedgers that index is optimal: no other exposure, feasible or not, removes more variance. Once sellers hedge too it remains feasible but need not minimize population loss, and the optimum computed without the feasibility restriction bounds how much variance any feasible index can remove.

Our results are conditional. Buyer diversification applies only to uncommitted spot spending; reserved capacity, workload constraints, and switching frictions concentrate buyer exposure. The compute application does not establish these assumptions, and the theory rules candidate designs out without picking one.

Using illustrative baseline inputs, an external benchmark aligned with the modeled buyer's basket hedges 99.7% of that buyer's variance but 28.2% of the seller population's variance under the same inputs. Market data are required to turn that model comparison into expected hedging

performance or a criterion for listing a contract. Two publishers' rental-rate indices for the H100, a single widely deployed Nvidia GPU, correlate at 0.17 weekly. That degree of disagreement makes publisher identity and index construction explicit contract terms, although the data do not identify methodology as its cause. Because we observe neither buyer shares nor seller positions, we cannot rank candidate settlement benchmarks by how well they hedge market participants.

1.1. Related literature

The first relevant literature runs from the classical theory of storage to derivatives on non-storable commodities. In the theory of storage a commodity's forward price is tied to its spot by cost-of-carry arbitrage and the convenience yield on inventory (Kaldor, 1939; Working, 1949; Telser, 1958; Gibson and Schwartz, 1990; Schwartz, 1997); for compute, with no inventory to carry, the relation is vacuous, which is why a compute contract's value is governed by basis risk rather than a no-arbitrage restriction (Section 3.1).

Compute belongs instead with the non-storable commodities, whose forward prices are set in equilibrium. Four of them bear on the design question here. Electricity shows how location-specific indices manage a heterogeneous service (Bessembinder and Lemmon, 2002; Routledge, Seppi, and Spatt, 2000). Shipping freight, settled on the composite Baltic Dry Index, is the closest analogue of combining generation-specific GPU rates into one benchmark (Kavussanos and Nomikos, 2003). Weather shows that a non-traded underlying can support an exchange-traded market (Weagley, 2019), and live cattle that storability is not a precondition (Paulson, 2025). Bandwidth is the family's failure (Kenyon and Cheliotis, 2001).

Compute has been priced before, toward a different end. An older literature on grid and cloud computing economics asked how to allocate a shared resource efficiently (Sutherland, 1968; Nielsen, 1970; Buyya et al., 2002; Wolski, Spring, and Hayes, 1999). The present wave, and this paper, ask

instead how to transfer the price risk of that resource, which puts benchmark design, basis risk, and hedging demand at the center. That is the margin on which the analogues divide: electricity and freight found settlement indices their participants would trade against, and bandwidth did not.

A second literature, on the optimal design of futures contracts and benchmarks, supplies our central analytical lens. A successful contract balances matching hedgers' exposures against concentrating trading interest (Duffie and Jackson, 1989; Tashjian, 1995); a recurring response to heterogeneity is a reference-grade contract with deviations trading as priced basis, an approach industry participants have proposed adapting to compute (Friedman, 2026). Because no-arbitrage is silent on a compute contract's value, our tools are those of minimum-variance hedging and the optimal cross-hedge (Johnson, 1960; Ederington, 1979; Anderson and Danthine, 1981), recast in the factor form of arbitrage pricing (Ross, 1976); the settlement index is itself a benchmark: Duffie, Dworczak, and Zhu (2017) study how publishing one changes trade, matching, and participation in a search market, while Duffie and Dworczak (2021) design transaction-based fixings around manipulation incentives and characterize a linear rule robust to a specified form of collusion. That literature motivates the averaging windows we consider, though Section 4 evaluates those windows on tracking and lag rather than on manipulation resistance.

By asking when a new instrument can complete a missing market, the paper also engages the literature on financial innovation in incomplete markets (Allen and Gale, 1994; Duffie and Rahi, 1995). The case for a compute market rests on hedging demand, which ties it to the real effects of corporate risk management (Smith and Stulz, 1985; Froot, Scharfstein, and Stein, 1993; Pérez-González and Yun, 2013). The published record does not yet let us measure the compute counterpart of that demand.

The closest work studies compute derivatives directly. Closest in data, Cha (2026) uses the same Bloomberg A100, H100 and B200 rental indices to ask whether existing assets already hedge

compute price risk, and finds no proxy hedge that reduces variance out of sample. That negative result is the premise this paper starts from: it motivates a dedicated contract without saying which benchmark the contract should settle on. Xing (2026) defines a standardized inference-token unit and specifies settlement, margins, and market making, then evaluates hedging under an assumed mean-reverting jump-diffusion. Assody (2026) models the GPU forward curve and volatility with scheduled generational jumps and examines basis between a blended settlement index and a specific compute exposure. Our object is complementary: we endogenize how basket-holding buyers and configuration-locked sellers map into benchmark weights, contract-menu size, and signed hedging-pressure premia. The distinctive claim is therefore not that compute is non-storable or depreciating, nor that a compute future can exist. It is that asymmetric exposures create a benchmark-design problem whose answer differs by side of the market.

2. Institutional Details

2.1. Institutional Setting

We treat compute as rented GPU capacity, quoted as a rental rate per GPU-hour. The GPUs are Nvidia's and AMD's data-center parts, and the four indices we use all track Nvidia generations. GPUs and competing accelerators such as Google's *tensor processing units* (TPUs) and Amazon's Trainium differ across generations in throughput, memory, interconnect, and delivered performance. Workloads written for particular capabilities make them imperfect substitutes, so their relative prices turn on current supply and on how fast each part is expected to be superseded. Capacity is sold through hyperscalers, specialized "neoclouds", and rental marketplaces. Shared data-center plant, power, and cooling do not erase configuration-specific service quality, so a settlement benchmark must aggregate prices across distinguishable services and publishers rather than measure one homogeneous

commodity.

Those prices come in two forms. Spot rates price immediate use, while reserved contracts commit a buyer for one to three years on take-or-pay terms: the buyer pays whether or not it uses the capacity. Only spot and marketplace prices are published, so a cash-settled contract must settle on the spot tier. Much of what buyers spend and providers earn is committed at the reserved tier, and a spot-settled hedge covers it only to the extent that the two tiers move together. The exchanges have already named their benchmarks (Section 1), so this is not a choice still waiting on evidence: it was made before anyone could measure how tightly the two tiers move. Whether trading across both tiers will pull the two prices together, and by how much, cannot be tested without access to reserved prices.

An idle GPU-hour cannot enter inventory, so no stock links today's price to tomorrow's. That removes the inventory trade behind cash-and-carry arbitrage, which otherwise pins futures to spot when a good can be stored; it neither mandates cash settlement nor prevents delivery of capacity produced at maturity. What would be delivered is not a homogeneous hour: hardware, cluster, provider, geography, utilization, and delivered performance each separate one rental from another. Generations turn over roughly once a year, and news that a configuration is falling behind is a first-order risk rather than a slow trend. Electricity and freight illustrate different responses to heterogeneity: electricity contracts trade at several locations, while BIFFEX settled one freight future on a composite Baltic index (Bessembinder and Lemmon, 2002; Kavussanos and Nomikos, 2003). Which response applies to compute depends on whether buyers substitute freely enough to make one broad index the hub, with the rest trading as basis; the model takes up that question, and the data cannot yet answer it.

2.2. Market Participants and Hedging Demand

Compute is a large and volatile input cost for AI laboratories and inference businesses, so rising rental rates raise their costs; a neocloud's revenue is the rental rate itself, so falling rates cut it. The two sides therefore need opposite hedges, with users buying the contract and operators selling it. Which of them actually shows up decides how much risk a contract can shift and whether it trades at all. A direct equity test of that supplier prediction is unfavorable, returning an estimate that is negative or too imprecise to distinguish from zero; Internet Appendix Section IA.VIII reports it. Integrated hyperscalers net both exposures internally and pass much of the residual to customers, so natural demand may concentrate among the thinnest and most credit-constrained participants. Lenders exposed through GPU collateral hold related asset-value risk rather than the same operating cash-flow exposure (Aldasoro, Doerr, and Rees, 2026).

A neocloud built around one generation, interconnect, region, and sales channel earns exactly that configuration's rental rate and cannot diversify that exposure in ordinary operations. A buyer needing compute in general can spread its spending across configurations, so its cost behaves like an expenditure-weighted basket of configuration prices. Vertical integration does not remove that exposure: a frontier laboratory facing risky growth optimally under-builds its own data centers relative to its prospective needs (Grégoire, 2026), so even the largest self-hosting labs remain buyers of marginal capacity at market rates. The asymmetry that matters is therefore between one configuration's price and a basket of many, and it decides which side can diversify configuration-specific risk.

That flexibility is strongest within one accelerator vendor's line: software and workload requirements make switching vendors costly, and a switch is a procurement and architecture decision measured in quarters rather than a response to weekly rental rates. Even there, shortages restrict substitution to whatever is available. A take-or-pay reserved commitment removes the flexibility

for its term. Buyers that have entered into reserved contracts to secure access face risk closer to a neocloud's: they are tied to one configuration, at a tier a spot-settled index does not price. The flexibility that bears on hedging therefore sits at the spot margin, across provider, channel, region, and generation inside that line.

Firms manage these exposures today by reserving capacity, integrating vertically, and passing costs through where market power allows. Risk sharing stays incomplete under those arrangements: the contracts are lumpy, credit-intensive, and hard to mark, which leaves them illiquid and closed to third parties. A centrally cleared future would add a standardized instrument, marked to market and open to those third parties, with clearing and margining that reduce and mutualize bilateral counterparty exposure without fully eliminating credit or liquidity risk.

Counterparty exposure is itself a constraint on how fast capacity gets built. Vendor backstop arrangements, under which the chip vendor supports a neocloud's lease revenue so that lenders will finance the hardware, are one response to it.[2] A cleared future would address a different margin, the price risk on capacity once it exists, so it would complement such arrangements rather than replace them.

2.3. Data

The empirical record consists of four published GPU-rental indices with unequal histories of at most about two years and substantial daily staleness. Because it contains no transaction-level configurations, buyer invoices, seller revenues, hardware positions, or traded compute futures, the data describe published-index moments and cross-index tracking, and cannot identify economic basis, exposure weights, causal shocks, hedging demand, or futures premia.

Each index is a publisher's estimate of the prevailing rental rate for one hardware generation,

[2] See https://newsletter.semianalysis.com/p/nvidia-gpu-debt-backstop-unleashes.

Table 1: Compute-price index series: coverage and staleness

Ticker	SDA100RT	SDH100RT	SDB200RT	ORNNH100	SDLLMTK	ORNNTOKA
Provider	Silicon Data	Silicon Data	Silicon Data	Ornn	Silicon Data	Ornn
Subject	A100	H100	B200	H100	LLM token	LLM token
Unit	GPU-hour	GPU-hour	GPU-hour	GPU-hour	token	token
Start	2024-09	2024-09	2025-08	2024-10	2025-12	2026-06
End	2026-07	2026-07	2026-07	2026-07	2026-07	2026-07
Daily obs.	480	480	242	459	155	22
Zero daily changes (%)	61.7	45.6	19.8	19.4	0.6	13.6
Longest flat run (days)	22	5	4	4	1	2

Notes. The table reports each published series' subject, unit, sample span, daily count, unchanged-day fraction, and longest unchanged run. The empty Ornn A100 series is absent; the one-month Ornn token series is excluded from the analysis. Weekly Friday closes are the empirical workhorse.

quoted per GPU-hour and published as a daily level. Bloomberg supplies Silicon Data A100, H100, and B200 indices and an Ornn H100 index (Table 1). The subject of each series is a generation rather than a configuration: one number stands for rentals that differ in provider, region, cluster, contract tier, and delivered performance. The rule that collapses them belongs to the publisher, as do the choices it embeds: which providers are sampled, whether quotes or transactions enter, how they are weighted, and when they are revised.

That rule is the object we do not observe. We see published index levels rather than the underlying transaction panels, and only the July 2026 vintage of each series. Two series labelled H100 need not share venue, observation type, weighting, revision policy, or refresh schedule, so a gap between them combines methodology with economic dispersion and the data cannot decompose the two.

Daily prices are stale: unchanged-day fractions range from 19.4% to 61.7% across the four series (Table 1). We therefore use Friday-close weekly observations and construct log price changes $\Delta \log p_{i,t}$ for the decomposition (2) and factor model (5). Weekly aggregation reduces zero changes but cannot synchronize latent transactions or unequal overlaps.

Bloomberg also supplies a Silicon Data LLM-token index, which prices inference output per token rather than hardware per GPU-hour. Its history begins in December 2025, giving about seven months of overlap with the rental indices through the July 2026 vintage. Figure 1 plots it alongside

the four rental series, but because it settles on a different unit, Internet Appendix Section IA.VII treats it as a short supplementary comparison rather than folding it into the rental-index analysis sample.

Vendor documentation supplies fixed generation-level performance scalars for the hardware, which rescale a series' level without touching any second moment of its log changes and reveal no GPU inventories, procurement contracts, configuration revenues, or hedge positions. Internet Appendix Section IA.III records that adjustment and the time-varying performance series a genuine substitution test would need.

2.4. Compute-Price Dynamics and Comovement

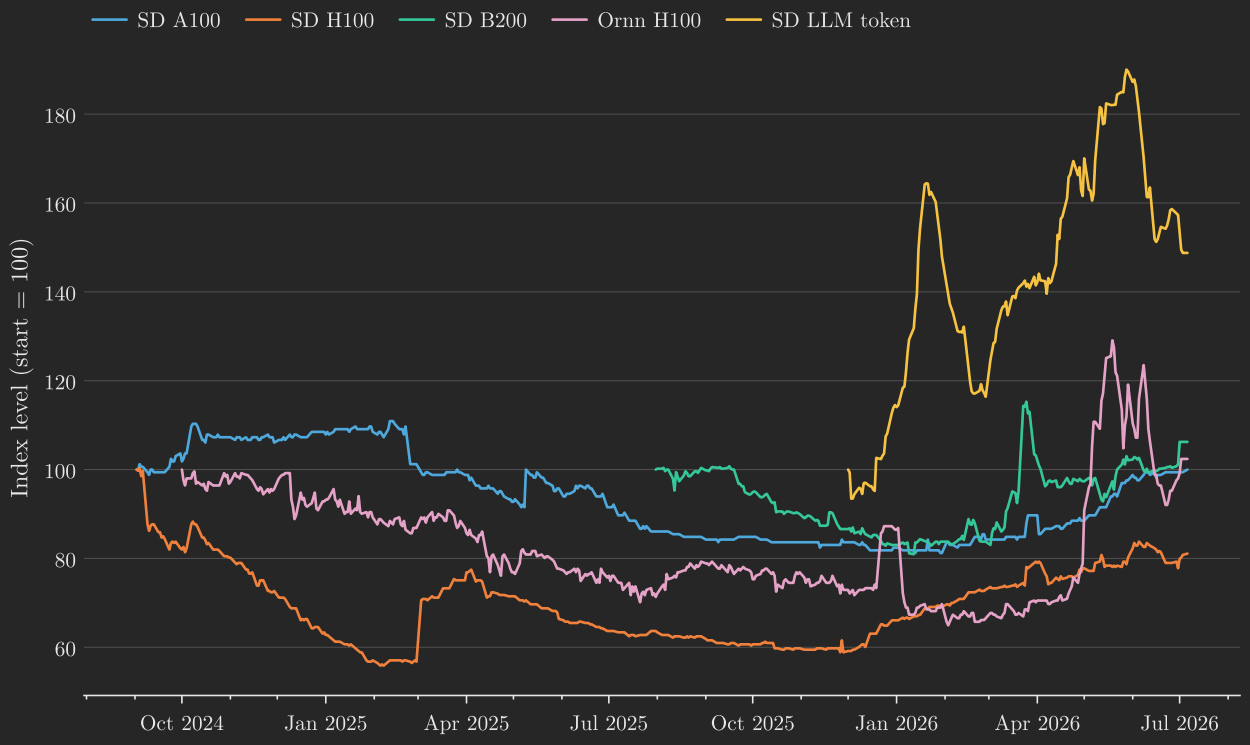
The short published-index record describes co-movement; it does not test the factor model or identify configuration shocks, generation effects, or firms' economic basis. Summary statistics, normalized levels, and drift estimates are reported in Internet Appendix Section IA.V.

2.4.1. Co-movement

Figure 1 plots the five series, each normalized to 100 at its first observation. The two H100 series, one per publisher, diverge; the B200 rises, the A100 is near flat, and the token index follows a path of its own. The series begin at different dates and do not move as one, which is the fact the rest of this section quantifies.

Table 2 reports pairwise overlaps and the common 49-week sample. Co-movement is modest: the two H100 publisher indices correlate at 0.17 pairwise and 0.24 when balanced; the largest estimate, 0.62, links Silicon Data A100 and B200. Balanced-sample moving-block-bootstrap intervals include zero for three of six pairs. Publisher, coverage, methodology, and economics remain confounded. The leading principal component explains 45.9% of balanced variance, with a 95% interval of

Figure 1: Compute-price index levels, normalized to 100 at the start of each series



Notes. Daily observations. Series begin on different dates and are normalized to their own first observation, so vertical distance between series reflects cumulative change since each start date rather than a comparison of price levels. The two H100 series are different publishers of the same generation.

33.8%–56.9%; with four published series, this is a covariance summary, not a factor-count or feasible-menu test.

Pairwise overlaps use each available history; balanced estimates trade sample length for comparability. Their difference is not convergence evidence. The same-H100 gap holds the hardware label fixed, but still combines publisher coverage and methodology with economic dispersion and remains descriptive rather than causal.

That same-hardware gap is provider disagreement, and it identifies none of its possible sources. Settlement-benchmark risk compares a contract benchmark with a hedger-relevant price, while firm economic basis compares that hedger's realized price with the benchmark; only provider disagreement is observed here, and publisher error could affect either clientele. The evidence makes publisher identity and index construction explicit contract terms, but selects no empirical hub

Table 2: Weekly correlations of price changes across indices

	SD A100	SD H100	SD B200	Ornn H100
Panel A: pairwise overlaps				
SD A100	1.00			
SD H100	0.11	1.00		
SD B200	0.62	0.37	1.00	
Ornn H100	0.09	0.17	0.07	1.00
Panel B: balanced sample (49 weeks)				
SD A100	1.00			
SD H100	0.13	1.00		
SD B200	0.62	0.37	1.00	
Ornn H100	0.08	0.24	0.07	1.00

Notes. Pearson correlations use weekly Friday-close log changes. Panel A uses pairwise overlap; Panel B uses the common 49-week sample. "SD" is Silicon Data; the two H100 entries are different publishers of the same hardware.

or menu. Internet Appendix Section IA.V reports the discrepancy exhibits and the relative-price persistence exercise, and Section IA.VI the cross-index hedging diagnostics.

3. The Model

Exchanges are naming settlement indices for compute before any trading history exists. Because participants run different hardware, no index tracks every rental cost equally, and the choice does not remove basis risk but decides who carries it: the seller tied to one configuration, or the buyer holding a basket. What an exchange can do about that turns on which factor direction its index tracks, and on whether a tradeable basket reaches that direction at all.[3]

3.1. Nonstorability and Settlement

For a storable commodity, cash-and-carry arbitrage tethers the forward price to spot through the cost of carrying inventory: the interest, storage cost, and convenience yield that separate the two (Kaldor, 1939; Working, 1949; Telser, 1958). Compute has no inventory trade to impose that restriction, so its forward price is set in equilibrium (Routledge, Seppi, and Spatt, 2000; Bessembinder and

[3] Proofs are in Internet Appendix Section IA.I, which also records what the accompanying Lean 4 package checks and what it leaves to the paper.

Lemmon, 2002):

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}[I_T] = \mathbb{E}_t[I_T] - \pi_{t,T}, \tag{1}$$

the physical expectation of a future settlement index I_T less a risk premium $\pi_{t,T}$. That leaves the benchmark a design choice, because equation (1) restricts the contract level and not its co-movement with the prices hedgers face. Hedge value therefore depends on basis risk, which the loss below isolates (Duffie and Jackson, 1989; Tashjian, 1995), subject to the separate implementation constraints of contractibility and liquidity.

3.2. Benchmark, Basis, and Hedging Effectiveness

Let $p_{i,t}$ denote the rental price of configuration i at time t, where i collects the attributes that distinguish one unit of compute from another (GPU generation, provider, geography, contract type, and performance category). A buyer or seller of i bears price risk with the conceptual decomposition

$$\underbrace{\Delta \log p_{i,t}}_{\text{compute price risk}} = \underbrace{D_t}_{\substack{\text{demand}\\ \text{shocks}}} + \underbrace{S_t}_{\substack{\text{supply}\\ \text{shocks}}} + \underbrace{O_{i,t}}_{\substack{\text{technological}\\ \text{obsolescence}}} + \underbrace{\xi_{i,t}}_{\substack{\text{configuration-}\\ \text{specific risk}}}, \tag{2}$$

where D_t and S_t are common demand and supply shocks, $O_{i,t}$ is configuration-specific obsolescence, and $\xi_{i,t}$ is configuration-specific risk. A benchmark turns it into basis risk: equation (3) defines the wedge a given index leaves.[4]

Let I_t be a contractible aggregator of configuration prices, such as $\log I_t = \sum_i w_i \log p_{i,t}$. Configuration i has basis

$$b_{i,t} \stackrel{\text{def}}{=} \log p_{i,t} - \log I_t, \tag{3}$$

[4] We reserve *basis risk* for the post-benchmark object. Equation (2) contains no index, and basis is defined against one, so the same price series has as many bases as there are candidate benchmarks. Naming $\xi_{i,t}$ basis risk here would presume the choice whose consequences Sections 3.5 and 3.7 study.

the price wedge left by the benchmark.

For exposure $\Delta \log p_{i,t}$ and an index position h, hedged variance is $\mathrm{Var}\big(\Delta \log p_{i,t} - h\,\Delta \log I_t\big)$, minimized at the hedge ratio

$$h_i^* = \frac{\mathrm{Cov}\big(\Delta \log p_{i,t},\, \Delta \log I_t\big)}{\mathrm{Var}\big(\Delta \log I_t\big)}, \qquad \frac{\mathrm{Var}\big(\Delta \log p_{i,t} - h_i^*\,\Delta \log I_t\big)}{\mathrm{Var}\big(\Delta \log p_{i,t}\big)} = 1 - \rho_i^2, \tag{4}$$

where ρ_i is the price-index correlation. Hedging effectiveness is ρ_i^2 (Ederington, 1979). A contract is useful to a hedger only insofar as it removes variance from the configuration she actually holds, which is what the benchmark-design problem below optimizes.

3.3. A Factor Model of Compute Prices

We model the change in the log rental price of configuration i, one of N configurations distinguished by the attributes of Section 3.2, as the linear factor structure

$$\Delta \log p_{i,t} \;=\; \alpha_i \;+\; \beta_i' f_t \;+\; \varepsilon_{i,t}. \tag{5}$$

Here $f_t \in \mathbb{R}^r$ contains r demeaned common compute-price factors with positive-definite covariance $\Sigma_f \stackrel{\text{def}}{=} \mathrm{Var}(f_t)$. The vector $\beta_i \in \mathbb{R}^r$ holds configuration-specific loadings, and $\varepsilon_{i,t}$ is an idiosyncratic disturbance, mean zero with variance σ_i^2, uncorrelated with the factors and across configurations. The factors gather the systematic demand, supply, and obsolescence forces of Section 3.2. They refine the conceptual decomposition (2) in the arbitrage-pricing tradition of Ross (1976); Internet Appendix Section IA.II spells out the mapping.

Write $\alpha_i = \mu - \delta_i$, where $\delta_i \geq 0$ is expected technological depreciation. This known trend affects the forward level, not second moments. Only stochastic obsolescence news enters f_t and hedge risk (Internet Appendix Section IA.II).

With $B \stackrel{\text{def}}{=} [\beta_1, \dots, \beta_N]'$ stacking the loadings, the exposures a benchmark design can reach form the feasible set

$$\mathcal{K} \stackrel{\text{def}}{=} \{B'w : \ w \in \mathcal{W}\} = \text{conv}\{\beta_1, \dots, \beta_N\}, \tag{6}$$

where $\mathcal{W} \stackrel{\text{def}}{=} \{w \in \mathbb{R}^N : \ \mathbf{1}'w = 1, \ w_i \geq 0 \ \forall i\}$ is the baseline set of economically feasible weights. The simplex rules out short or levered constituents; concentration, liquidity, and manipulation screens can shrink it further. A design $w^k \in \mathcal{W}$ defines the settlement index $\log I_t^k = \sum_j w_j^k \log p_{j,t}$, as in Section 3.2. Aggregating equation (5),

$$\Delta \log I_t^k \ = \ \alpha^k \ + \ (\beta^k)' f_t \ + \ u_t^k, \qquad \beta^k \stackrel{\text{def}}{=} \sum_j w_j^k \beta_j = B'w^k, \quad u_t^k \stackrel{\text{def}}{=} \sum_j w_j^k \varepsilon_{j,t}, \tag{7}$$

where $\alpha^k = (w^k)'\alpha$ and $\tau_k^2 = \text{Var}(u_t^k) = \sum_j (w_j^k)^2 \sigma_j^2$. A benchmark is *well diversified* when $\tau_k^2 \approx 0$ and each $w_i^k \approx 0$.

Feasible loadings lie in the convex hull $\mathcal{K}$. A direction can therefore belong to the linear span of the rows of B, so that some signed combination of configurations would deliver it, without being generated by any contractible price index: the combination that reaches it may require short or levered weights, which $\mathcal{W}$ forbids.

First-differencing equation (3) and substituting equations (5)–(7) gives the basis dynamics

$$\Delta b_{i,t}^k \ = \ (\alpha_i - \alpha^k) \ + \ (\beta_i - \beta^k)' f_t \ + \ (\varepsilon_{i,t} - u_t^k), \tag{8}$$

Thus basis risk is factor-loading mismatch plus idiosyncratic noise.

3.4. Minimum-Variance Hedging and Hedging Effectiveness

For the cash-settled design analyzed here, consider a hedger with unit exposure to configuration i taking a position of h units in the futures on I^k; the minimum-variance hedge of the return $\Delta \log p_{i,t} - h\,\Delta \log I_t^k$ generalizes equation (4) to the factor setting.

Proposition 1 (Minimum-variance hedge and hedging effectiveness)**.** *The position minimizing* $\mathrm{Var}(\Delta \log p_{i,t} - h\,\Delta \log I_t^k)$ *is*

$$h_i^k = \frac{\beta_i' \Sigma_f \beta^k + w_i^k \sigma_i^2}{(\beta^k)' \Sigma_f \beta^k + \tau_k^2}, \tag{9}$$

and the hedging effectiveness, the fraction of return variance it eliminates, is

$$HE_i^k \overset{\text{def}}{=} 1 - \frac{\mathrm{Var}(\Delta \log p_{i,t} - h_i^k\,\Delta \log I_t^k)}{\mathrm{Var}(\Delta \log p_{i,t})} = \rho_{i,k}^2 = \frac{(\beta_i' \Sigma_f \beta^k + w_i^k \sigma_i^2)^2}{V_i\left[(\beta^k)' \Sigma_f \beta^k + \tau_k^2\right]}, \tag{10}$$

where $V_i \overset{\text{def}}{=} \mathrm{Var}(\Delta \log p_{i,t}) = \beta_i' \Sigma_f \beta_i + \sigma_i^2$ *and* $\rho_{i,k}$ *is the correlation between* $\Delta \log p_{i,t}$ *and* $\Delta \log I_t^k$*. For a well-diversified benchmark* ($\tau_k^2 \to 0$, $w_i^k \to 0$)*,*

$$h_i^k = \frac{\beta_i' \Sigma_f \beta^k}{(\beta^k)' \Sigma_f \beta^k}, \qquad HE_i^k = \frac{\left\|\Pi_{\beta^k}\,\tilde{\beta}_i\right\|^2}{\|\tilde{\beta}_i\|^2 + \sigma_i^2}, \tag{11}$$

where $\tilde{\beta}_i \overset{\text{def}}{=} \Sigma_f^{1/2} \beta_i$ *and* Π_{β^k} *is the orthogonal projection onto* $\mathrm{span}(\Sigma_f^{1/2} \beta^k)$*.*

A single index spans one factor direction, so effectiveness rises with loading alignment and falls with idiosyncratic variance. At $h = 1$, the residual is the basis change in equation (8), the standard cross-hedging decomposition in factor form (Johnson, 1960; Anderson and Danthine, 1981).

3.5. Optimal Benchmark Design

Let $\eta_i \geq 0$, $\sum_i \eta_i = 1$, denote configuration i's share of aggregate hedging demand. The designer minimizes the exposure-weighted residual hedging variance

$$L(k) \;=\; \sum_i \eta_i \,\mathrm{Var}(\Delta \log p_{i,t} - h_i^k \,\Delta \log I_t^k) \;=\; \sum_i \eta_i \, V_i \,(1 - HE_i^k), \tag{12}$$

where each hedger uses her own optimal ratio h_i^k from Proposition 1. Finding the factor exposure that minimizes $L(k)$ over unrestricted directions is a spectral problem; equation (6) determines whether a price index can deliver that exposure. Hedging effectiveness is a squared correlation, so rescaling the index leaves it unchanged and each hedger absorbs the difference through h_i^k: the designer picks a direction, not a magnitude. The objective is therefore a ratio of quadratic forms, and its maximum is an eigenvalue (Horn and Johnson, 2013, Chapter 4).

Proposition 2 (Relaxed bound and feasible benchmark weights)**.** *In the well-diversified factor approximation:*

(a) The unrestricted problem. *Minimizing $L(k)$ is equivalent to maximizing the scale-free Rayleigh quotient*

$$G(\beta) = \frac{\beta' \Sigma_f \, M_\beta \, \Sigma_f \beta}{\beta' \Sigma_f \beta}, \qquad M_\beta \stackrel{\text{def}}{=} \sum_i \eta_i \, \beta_i \beta_i', \tag{13}$$

over nonzero loading directions β. Let $\widetilde{M} \stackrel{\text{def}}{=} \Sigma_f^{1/2} M_\beta \Sigma_f^{1/2}$, with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r \geq 0$, and let E_1 denote its leading eigenspace. The unrestricted maximum is λ_1, attained when $\Sigma_f^{1/2}\beta \in E_1 \setminus \{0\}$, and the corresponding relaxed loss is

$$L^*_{\text{rel}} \;=\; \sum_i \eta_i \sigma_i^2 \;+\; (\,\mathrm{tr}\,\widetilde{M} - \lambda_1) \;=\; \sum_i \eta_i \sigma_i^2 \;+\; \sum_{l \geq 2} \lambda_l. \tag{14}$$

(b) The feasible index. *Define*

$$G^*_{\mathcal{W}} \stackrel{\text{def}}{=} \max_{w \in \mathcal{W}:\, B'w \neq 0} G(B'w) \leq \lambda_1, \qquad L^*_{\mathcal{W}} = L^*_{\text{rel}} + (\lambda_1 - G^*_{\mathcal{W}}). \tag{15}$$

The maximum exists whenever $0 \notin \mathcal{K}$. Equality with the relaxation holds if and only if some $w \in \mathcal{W}$ satisfies $\Sigma_f^{1/2} B'w \in E_1 \setminus \{0\}$. If no such weight exists, equation (15) is the constrained fractional-quadratic program and the feasibility gap $\lambda_1 - G^*_{\mathcal{W}}$ is strictly positive.

A publisher unwilling to rely on $0 \notin \mathcal{K}$ can instead impose a positive benchmark-variance floor. The gap closes in one economically recognizable case, and rank is not enough to close it.

Corollary 1 (When the feasible index attains the relaxed direction)**.** *If exposures are codirectional, $\beta_i = c_i \bar{\beta}$, and $\sum_i \eta_i c_i \neq 0$, the feasible choice $w = \eta$ attains the relaxed direction. By contrast,* $\text{rank}(B) = r$ *guarantees representation of arbitrary loadings only with signed, unnormalized weights; rank alone does not imply $w \in \mathcal{W}$.*

The largest eigenvalue bounds what a single feasible index can deliver, but identifies no settlement rule. The calibration in Section 4.1 reaches the best available direction inside the model, and that says nothing about which index an exchange should name, because reaching the best direction in a model whose shares are assumed does not show that any real index reaches it. At the buyer-only limit, Divisia weights are what the model recommends once buyer expenditure shares are known, and those shares are exactly what we do not observe. For finite indices, own-weight covariance and τ_k^2 re-enter equation (10), so even that feasible relaxed direction need not dominate under the exact objective.

3.6. Spanning and the Precision–Liquidity Tradeoff

One index leaves the residual $\sum_{l\geq 2}\lambda_l$ in equation (14). The factors still drive that residual rather than idiosyncratic noise, so a contract loading on a direction the first index misses would hedge part of it. Carrying another contract is costly, so the designer covers directions only as far as the variance saved outweighs that cost. Suppose contracts $1,\dots,m$ have loadings $\beta^{(1)},\dots,\beta^{(m)}$ spanning the subspace $\mathcal{S}_m = \operatorname{span}(\Sigma_f^{1/2}\beta^{(1)},\dots,\Sigma_f^{1/2}\beta^{(m)})$ of whitened factor space. A hedger combining the m futures replicates the projection of her factor exposure onto $\mathcal{S}_m$. Feasible menus restrict the span to $\mathfrak{S}_m(\mathcal{W}) = \{\operatorname{span}(\Sigma_f^{1/2}B'w^1,\dots,\Sigma_f^{1/2}B'w^m) : w^j \in \mathcal{W}\}$.

Proposition 3 (Spanning with few contracts and the precision–liquidity tradeoff)**.** *In the well-diversified factor approximation:*

(a) Over unrestricted loading directions, the top-m eigenspace E_m minimizes aggregate residual hedging variance, yielding the Ky Fan lower bound (Horn and Johnson, 2013, Chapter 4)

$$L^*_{m,\mathrm{rel}} \;=\; \sum_i \eta_i\sigma_i^2 \;+\; \sum_{l>m}\lambda_l. \tag{16}$$

Every feasible menu satisfies $L^*_{m,\mathcal{W}} \geq L^*_{m,\mathrm{rel}}$, with equality when $E_m \in \mathfrak{S}_m(\mathcal{W})$, allowing any basis within a tied eigenspace.

(b) If the relevant top-m spaces are feasibly attainable, let Φ, with $\Phi(0) = 0$, be an assumed *reduced-form cost of carrying a menu of m contracts, increasing and discretely convex, so $\Delta\Phi(m) = \Phi(m) - \Phi(m-1)$ is nonnegative and nondecreasing. The designer minimizes*

$$\mathcal{L}(m) \;=\; \sum_i \eta_i\sigma_i^2 \;+\; \sum_{l>m}\lambda_l \;+\; \Phi(m). \tag{17}$$

An interior optimum m^* satisfies $\lambda_{m^*} \geq \Delta\Phi(m^*)$ and $\lambda_{m^*+1} \leq \Delta\Phi(m^*+1)$, with boundary

and equality cases handled by the same one-step comparisons. Under linear costs $\Phi(m) = \gamma m$, the optimum is the threshold rule $m^* = \#\{l : \lambda_l > \gamma\}$. If E_m is not attainable, the designer instead compares the constrained residual reductions; no generic eigenvalue threshold follows.

What Φ measures is mostly the cost of splitting hedging interest, not the fee an exchange charges to list, which we take to be second order beside what a thin book costs the hedgers who must trade in it. Section 3.8 gives that cost an equilibrium source in the compensation speculators require to absorb a narrow contract's net flow.

Whether feasible indices span the relevant directions is the binding question: under the baseline simplex the constituent weights e_i are feasible, so $\operatorname{rank}(B) = r$ lets r single-configuration indices span factor space, but those concentrated indices need not satisfy the diversified approximation or tighter publisher constraints. Fast eigenvalue decay therefore supports a small menu only when feasible indices span the relevant directions. Spanning all of them still leaves $\sum_i \eta_i \sigma_i^2$ in equation (16), so factor contracts leave the idiosyncratic part of configuration-specific risk untouched however long the menu runs. Where the menu stops depends further on the shape of Φ, which the equilibrium below motivates in one component without deriving convexity; Sections 2.4 and 4.1 supply the decay and the cost the count needs.

3.7. Two Populations: Locked-In Sellers and a Flexible Buyer

The model so far treats every hedger symmetrically: each exposure in equation (12) is a single configuration, distinguished only by its weight. Section 2.2 argued that the two sides of the compute market do not fit this description equally well, the seller's exposure being a single point in the space of configurations and the buyer's a diversified index over it. Assumption 1 writes that split into equation (12) through two primitives: the buyer's expenditure shares, which build her basket, and the sell-side share ϕ of aggregate hedging demand, which sets how much configuration heterogeneity

the design must serve.

Assumption 1 (Two-population hedger structure)**.** The hedger population of equation (12) divides into two groups.

(i) *Sellers.* A seller of configuration i has unit exposure to $\Delta \log p_{i,t}$, with loading β_i and idiosyncratic variance σ_i^2 as in equation (5). Sellers have exposure weights $\eta_i^S \geq 0$, $\sum_i \eta_i^S = 1$, and account collectively for a share $\phi \in [0, 1]$ of aggregate hedging demand.

(ii) *The buyer.* The remaining share $1 - \phi$ belongs to a representative buyer who produces effective compute through a constant-elasticity-of-substitution (CES) aggregator over the N configurations with elasticity θ. Expenditure minimization gives unit cost $c(p)$ and Shephard expenditure shares $s_i(p) \geq 0$, $\sum_i s_i(p) = 1$.

(iii) *Bounded flexibility.* The substitution the aggregator grants operates only at the spot margin, across providers, venues, regions, and adjacent generations within a vendor's line; capacity committed under reserved take-or-pay contracts is outside the aggregator, since for the commitment's duration it gives the buyer a seller-like locked exposure (Section 2.2). The aggregator also runs over the N GPU configurations alone, so substitution toward non-GPU accelerators sits outside it in the opposite way: that margin is real flexibility, but it moves expenditure onto capacity no index over these configurations spans.

The buyer's unit cost updates its weights as prices move: the exact chain-linked Divisia differential is (Hulten, 1973)

$$d \log c(p_t) = \sum_i s_i(p_t)\, d \log p_{i,t}. \tag{18}$$

A futures hedge cannot update this way; its position is set at the start of a hedge horizon H. Freezing shares at their baseline values $s_i \stackrel{\text{def}}{=} s_i(p_0)$ gives the static exposure $X_D^H \stackrel{\text{def}}{=} \sum_i s_i\, \Delta_H \log p_{i,t}$, a

first-order approximation to the change in the buyer's cost, and the wedge it leaves is the substitution bias a fixed-weight index carries against a true cost index (Diewert, 1976; Hausman, 2003). The omitted substitution term is second order in price shocks and carries the sign of $1-\theta$; the expansion and its higher-order terms are in Internet Appendix Section IA.I.1. Applying equation (5) to the static exposure gives

$$\beta^D \stackrel{\text{def}}{=} \sum_i s_i\,\beta_i \;=\; B's, \qquad \sigma_D^2 \stackrel{\text{def}}{=} \sum_i s_i^2\,\sigma_i^2, \tag{19}$$

the index structure of equation (7). Its residual variance is of order $1/N$ only when baseline spending is dispersed. If switching costs or reserved commitments concentrate spending, this diversification weakens, and a buyer whose spending sits in a single configuration becomes seller-like. The asymmetry between the two sides is therefore local to the horizon H and to the uncommitted spot margin of Assumption 1, not a claim of frictionless substitution.

Buyers need not share one basket. Give type a population weight ν_a, $\sum_a \nu_a = 1$, and baseline-share loading $\beta^{D,a} = B's^a$. The buyer second moment is then $\bar\beta^D(\bar\beta^D)' + H_D$, where $\bar\beta^D = \sum_a \nu_a\beta^{D,a}$ is the mean loading and $H_D = \sum_a \nu_a(\beta^{D,a} - \bar\beta^D)(\beta^{D,a} - \bar\beta^D)'$ its dispersion across types. One hub spans every type exactly only when the type loadings are codirectional, approximately only when H_D is small; otherwise buyer dispersion adds contract directions of its own. The closed-form propositions below set $H_D = 0$, a single representative buyer standing in for demand that is in fact heterogeneous.

In that representative case the design problem of Section 3.5 places mass $\phi\,\eta_i^S$ on each seller and $1-\phi$ on the buyer, and the design matrix specializes to

$$\widetilde M(\phi) \;=\; \phi\,\widetilde M_S \;+\; (1-\phi)\,\tilde\beta^D(\tilde\beta^D)', \qquad \widetilde M_S \stackrel{\text{def}}{=} \Sigma_f^{1/2}\Big(\textstyle\sum_i \eta_i^S\beta_i\beta_i'\Big)\Sigma_f^{1/2}, \quad \tilde\beta^D \stackrel{\text{def}}{=} \Sigma_f^{1/2}\beta^D, \tag{20}$$

a seller block carrying the full configuration heterogeneity plus a single buyer block of rank one:

each seller enters locked to her own configuration, so the cross-configuration heterogeneity is all theirs, while the buyer spends across every configuration yet enters through a single basket exposure. Propositions 2 and 3 hold as stated with $\widetilde{M}(\phi)$ in place of $\widetilde{M}$.

Proposition 4 (One-sided economic basis risk)**.** *For the representative buyer's static local exposure over H:*

(a) Against a benchmark with weights w^k, exact hedging effectiveness within the local linear exposure is

$$HE^k_{D,\text{exact}} = \frac{\left((\beta^D)'\Sigma_f\beta^k + \sum_i s_i w_i^k \sigma_i^2\right)^2}{V_D\left((\beta^k)'\Sigma_f\beta^k + \tau_k^2\right)}, \qquad V_D \stackrel{\text{def}}{=} \|\tilde{\beta}^D\|^2 + \sigma_D^2. \tag{21}$$

If $w^k = s$ and the benchmark uses the same constituent prices, the exposure is the index itself and $HE^k_{D,\text{exact}} = 1$.

(b) For an independently measured, well-diversified benchmark, the factor-only approximation drops own-weight covariance and τ_k^2. Its shortfall is

$$1 - HE_D^{k,0} = \frac{\left\|(I - \Pi_{\beta^k})\tilde{\beta}^D\right\|^2}{V_D} + \frac{\sigma_D^2}{V_D}. \tag{22}$$

Thus a loading match gives $HE_D^{k,0} = 1 - \sigma_D^2/V_D$, not the exact self-index result in part (a).

(c) The representative-buyer population's relaxed spectral loss splits as

$$L^*_{\text{rel}}(\phi) = \phi\sum_i \eta_i^S\sigma_i^2 + (1-\phi)\sigma_D^2 + \sum_{l\geq 2}\lambda_l(\widetilde{M}(\phi)). \tag{23}$$

The proposition isolates economic configuration risk. Dispersed buyer shares make the factor-loading mismatch and basket residual small against a buyer-aligned hub, whereas a locked seller

retains her configuration loading and idiosyncratic variance. This one-sidedness is conditional on the local, representative-buyer structure: concentrated commitments and heterogeneous buyer loadings weaken it, and the proposition supplies no empirical magnitude. Publisher measurement error is a separate channel, outside Proposition 4 and distinct from economic basis: classical error raises benchmark variance and lowers hedge quality for both clienteles, and nonclassical methodology error can also move the covariance itself.

Proposition 5 (The feasible buyer-share benchmark)**.** *Suppose $\beta^D \neq 0$. The expenditure-share weights $w^k = s$ belong to $\mathcal{W}$ and attain the buyer loading exactly, $\beta^k = B's = \beta^D$. At $\phi = 0$, this loading spans the sole nonzero direction of $\widetilde{M}(0)$, so the buyer-share index attains both the relaxed bound and the constrained objective in Proposition 2. For $\phi > 0$, it remains feasible but need not minimize population loss. If all configuration exposures are codirectional, a feasible index with a nonzero loading in their common direction attains the relaxation for every ϕ.*

The result names a feasible static baseline-share target over H, not a selected settlement product. Its chain-linked counterpart updates shares and is a different rule. Both require utilization, procurement, or marketplace data that the current panel does not observe. The empirical record therefore leaves the hub unselected; equal weight and published indices remain candidates rather than defaults.

Proposition 6 (Menu size is a sell-side object)**.** *The buyer block in equation (20) is positive semidefinite of rank at most one, so the eigenvalues of $\widetilde{M}(\phi)$ interlace those of the scaled seller block $\phi\widetilde{M}_S$: for $l \geq 2$,*

$$\phi\,\lambda_l(\widetilde{M}_S) \;\leq\; \lambda_l(\widetilde{M}(\phi)) \;\leq\; \phi\,\lambda_{l-1}(\widetilde{M}_S). \tag{24}$$

Consequently:

(a) Every eigenvalue of $\widetilde{M}(\phi)$ beyond the first is $O(\phi)$ and vanishes as the sell-side share goes to

zero. For a fixed $l \geq 2$, it is $\Theta(\phi)$, meaning it shrinks at the same rate as that share rather than merely being bounded by it, if $\lambda_l(\widetilde{M}_S) > 0$. If instead $\lambda_{l-1}(\widetilde{M}_S) = 0$, equation (24) forces $\lambda_l(\widetilde{M}(\phi)) = 0$; intermediate rank-deficient cases retain only the $O(\phi)$ statement.

(b) Under the linear threshold rule of Proposition 3(b), if $\phi > 0$ the relaxed threshold count is sandwiched by seller-driven counts,

$$n^S(\gamma/\phi) \;\leq\; m^*_{\mathrm{rel}}(\phi) \;\leq\; 1 + n^S(\gamma/\phi), \qquad n^S(x) \stackrel{\mathrm{def}}{=} \#\{\, l : \lambda_l(\widetilde{M}_S) > x \,\}, \tag{25}$$

and if $\gamma > 0$ and the hub eigenvalue clears the per-contract cost, $\lambda_1(\widetilde{M}(0)) > \gamma$, there exists $\bar{\phi} > 0$ such that the relaxed threshold count is one, $m^*_{\mathrm{rel}}(\phi) = 1$, whenever $0 \leq \phi < \bar{\phi}$. Without feasible attainment, these counts describe the relaxed spectral menu, not the constrained contract optimum.

The representative buyer contributes at most one direction. Positive trailing directions must come from the seller block, but rank deficiency matters: a zero seller direction need not become positive, and the combined matrix has rank at most one plus the seller rank. The substitution elasticity θ rotates the representative buyer direction without adding rank. Neither fact makes menu size globally monotone in ϕ; the Internet Appendix gives a two-configuration counterexample. The count sandwich in part (b) is what survives that non-monotonicity: for every positive sell-side share it bounds the relaxed count through the seller spectrum alone, and it pins the count to one only at small ϕ, provided the single best contract removes more variance than it costs to carry. In plain terms: when sellers are a small enough part of hedging demand, one broad contract is the whole menu, and it is worth listing at all only if it clears that cost.

3.8. A Hedging-Pressure Equilibrium

The design results take $\Phi(m)$ and the premium $\pi_{t,T}$ of equation (1) as primitives. A one-period equilibrium with constant-absolute-risk-aversion agents and jointly normal payoffs derives the premium and motivates one component of Φ, whose shape remains reduced form. A menu of m contracts has index changes stacked in $x = (x^1, \dots, x^m)'$, covariance matrix Ω, and expected payoff to the long $\mathbb{E}[x^k] = \pi^k$. Sellers of configuration i hold endowment $q_i^S \Delta \log p_{i,t}$, the flexible buyer holds $-q^D \Delta \log c_t$, and speculators hold none. Define $Q_S = \sum_i q_i^S$, $q_i^S = Q_S \eta_i^S$, $Q_D = q^D$, and aggregate risk tolerance $T = \sum_a \gamma_a^{-1}$.

Proposition 7 (Hedging-pressure premium)**.** *Suppose Ω is positive definite and every CARA coefficient γ_a is positive. Suppose also $Q_S, Q_D \geq 0$, $Q_S + Q_D > 0$, and $(\beta^D)' \Sigma_f \beta^D > 0$. Each agent's optimal position is $h^a = \gamma_a^{-1} \Omega^{-1} \pi - \Omega^{-1} \operatorname{Cov}(x, e^a)$, and zero-net-supply clearing yields*

$$\pi = \frac{1}{T} \sum_a \operatorname{Cov}(x, e^a), \qquad \pi^k = \frac{P^k}{T}, \quad P^k \stackrel{\text{def}}{=} \sum_a \operatorname{Cov}(x^k, e^a), \tag{26}$$

signed net hedging pressure over aggregate risk tolerance. For a well-diversified broad benchmark aligned with the buyer ($\beta^k = \beta^D$ and $w_i^k \to 0$), let $\bar{\beta}^S = \sum_i \eta_i^S \beta_i$ and $\rho_S = (\beta^D)' \Sigma_f \bar{\beta}^S / ((\beta^D)' \Sigma_f \beta^D)$. With $\phi = Q_S/(Q_S + Q_D)$, its premium has sign

$$\operatorname{sgn} \pi^{\text{broad}} = \operatorname{sgn} \left(\phi \rho_S - (1 - \phi) \right). \tag{27}$$

For a narrow contract on configuration i, $w^k = e_i$, the exact premium is

$$\pi^i = \frac{\beta_i' \Sigma_f (Q_S \bar{\beta}^S - Q_D \beta^D) + \sigma_i^2 (Q_S \eta_i^S - Q_D s_i)}{T}. \tag{28}$$

The residual term is seller-sided only if $Q_S \eta_i^S > Q_D s_i$ and buyer-sided if the inequality reverses.

The premium is net endowment covariance over risk-bearing capital (Bessembinder and Lemmon, 2002). Opposite-signed clientele pressures make either broad-premium sign possible; a narrow premium also depends on quantities, loadings, and risk tolerance. The exact narrow premium, equation (28), is derived in the Internet Appendix.

The same mechanism motivates a risk-bearing component of the liquidity cost. If speculators absorb a narrow contract's net flow F_m and its settlement variance is V_m, its premium bill is

$$c_m^{\mathrm{RP}} = \pi_m F_m \approx \frac{F_m^2 V_m}{T}. \tag{29}$$

This component is nonnegative and falls with aggregate risk tolerance, but the equilibrium does not make its sequence increasing or convex: flow can fall as a fixed hedging pool fragments, and V_m can move either way. Increasing, discretely convex Φ therefore remains a reduced-form assumption. If one identifies $\Delta\Phi(m)$ with c_m^{RP}, the additional restriction $F_{m+1}^2 V_{m+1} \geq F_m^2 V_m$ is needed for nondecreasing marginal cost; only then does the conditional threshold comparison inherit the form of Proposition 3(b). Otherwise the designer compares global values of equation (17). A fixed participation cost can be added as a separate reduced-form viability condition, but its launch threshold remains a scenario quantity until flows, variances, risk tolerance, and entry cost are estimated.

The equilibrium also isolates the stochastic part of obsolescence. Let f_{obs} be obsolescence news after projection on the other common factors, with variance $\sigma_{\mathrm{obs}}^2 > 0$. The loading restrictions below apply to this orthogonalized innovation; they are additional assumptions, not consequences of CES demand. Write $\bar{\beta}_{\mathrm{obs}}^S$ for the seller-exposure-weighted loading and β_{obs}^D for the buyer's loading.

Proposition 8 (One-sided obsolescence pressure)**.** *Suppose* $Q_S, Q_D \geq 0$, $\bar{\beta}_{\mathrm{obs}}^S < 0$, *and* $\beta_{\mathrm{obs}}^D \geq 0$. *The factor's contribution to net pressure on contract* k *is* $P_{\mathrm{obs}}^k = \beta_{\mathrm{obs}}^k \sigma_{\mathrm{obs}}^2 (Q_S \bar{\beta}_{\mathrm{obs}}^S - Q_D \beta_{\mathrm{obs}}^D)$, *and*

the clientele terms reinforce:

$$|P^k_{\text{obs}}| = |\beta^k_{\text{obs}}|\sigma^2_{\text{obs}}(Q_S|\bar{\beta}^S_{\text{obs}}| + Q_D|\beta^D_{\text{obs}}|). \tag{30}$$

Holding fixed gross exposure, factor variance, absolute contract loading, and risk tolerance, this reinforcement weakly raises premium magnitude relative to a factor on which the two clienteles co-load. Without that normalization, sign homogeneity implies no cross-factor ranking. Other factor and finite-index residual components can also offset this contribution, so it does not rank total contract premia.

Anticipated depreciation remains part of the forward level and not the covariance premium. Proposition 8 concerns only stochastic obsolescence news and a conditional sign comparison. A seller mix tilted toward frontier hardware can weaken or reverse the aggregate seller sign, and a locked buyer can restore cancellation. The proposition does not rank obsolescence against factors with different quantities, variances, or loadings; Internet Appendix Section IA.I records why that stronger ranking fails.

The buyer's flexibility is assumed, and only within limits: Assumption 1's CES demander substitutes at the uncommitted spot margin, and reserved contracts lock in capacity (Section 2.2). The one-sidedness of Proposition 4 follows from the two-population structure: it says how that structure allocates residual risk, not who bears it in today's market. The magnitudes sit outside the propositions altogether. Section 2.4 measures the comovement of configuration-level prices and the hedging effectiveness a settlement index delivers against single-configuration exposures. The model's own quantities, per-side hedging effectiveness, menu size, and equilibrium premia, are illustrative and belong to Section 4.1, which Section 4 applies to contract design.

4. Contract Design

4.1. Illustrative Quantitative Laboratory

The theory compares benchmark designs in population objects that a four-index, two-year record cannot estimate. We therefore use an illustrative laboratory whose data-generating process is the factor model of Section 3.3: selected published-index moments discipline its scale, while market structure and contract-design inputs are explicit scenario assumptions. The exercise reports model quantities, not empirical estimates, and agreement with selected moments is not validation.

4.1.1. Data-Generating Process and Input Authority

The laboratory has $N = 32$ configurations, four GPU generations crossed with eight provider-region variants, over $T = 504$ simulated trading days. Log rental-price changes obey equation (5) with three common factors labeled demand, scarcity, and obsolescence, and seller weights η_i enter the loss in equation (12). The labels attach economic content to assumed loading directions without identifying the factors, and matching aggregate scale to index moments does not identify the factor/residual decomposition: different structures reproduce the same volatility yet imply different hedging results.

Input authority differs sharply: published A100, H100, and B200 drift and volatility are estimated; aggregate factor and residual variance levels are matched to selected weekly moments; H200 drift and idiosyncratic volatility are interpolated because no H200 index exists; and the configuration counts, loading profiles and dispersion, residual independence, seller weights, H200 loadings and exposure, and all demand and liquidity parameters are assumed. Only estimated inputs are direct sample statistics; the rest remain scenario choices however plausible their magnitudes (Tables IA.1 and IA.2, Internet Appendix Section IA.IV.1).

The two-population layer adds the seller share of hedging demand ϕ, the buyer CES elasticity θ, and the per-contract liquidity threshold γ; buyer shares come from assumed quality-adjusted unit

prices (Table IA.3, Internet Appendix Section IA.IV.2). Dispersed shares put the constructed buyer idiosyncratic variance far below a typical configuration's and make the buyer block close to rank one by design, a property of the modeled exposure vector, not procurement evidence: commitments, workload lock-in, heterogeneity, or concentrated shares can make an actual buyer more seller-like.

4.1.2. Candidate Benchmarks and Per-Side Hedging Effectiveness

We compare four designs: a broad equal-weighted index; an index with the assumed seller weights; a *relaxed factor-direction index* on the leading direction of the exposure-weighted design matrix; and one equal-weighted index per generation. We call the third *relaxed* because the Rayleigh solution is an unrestricted systematic-fit direction, not automatically a nonnegative, sum-to-one settlement index; its simplex attainability here is scenario-specific and selects no empirical product. Per-design in-sample and exact population values (Table IA.5) are in Internet Appendix Section IA.IV.2.

The matched population comparison is modest: the broad index hedges 0.282 of the exposure-weighted locked-in seller variance, and the generation menu hedges 0.369 of the same population estimand. The larger unweighted in-sample generation number is not the seller-population headline, and neither value is empirical HE. The gain also mixes two channels rather than measuring factor alignment alone. A generation index raises each locked seller's own-constituent weight while concentrating settlement residual noise among fewer names; the theorem's diversified limit drops both terms, and at finite N they rival small loading improvements. This is also why the relaxed factor-direction candidate can slightly underperform simpler indices in the exact finite menu. The leave-one-out decomposition (Internet Appendix Section IA.IV.3) does not authorize a generation menu as the empirical design.

Table 3 separates the modeled seller and buyer exposures.

The broad candidate's buyer population HE is 0.997 by construction: dispersed CES shares

Table 3: Per-side HE in the illustrative two-population scenario

	Seller mean HE	Seller wtd HE	Seller wtd HE (pop.)	Buyer HE	Buyer HE (pop.)
Broad equal-weighted	0.297	0.286	0.282	0.997	0.997
Exposure-weighted	0.291	0.286	0.283	0.993	0.993
Relaxed factor direction	0.289	0.282	0.279	0.998	0.997
Generation-segmented	0.395	0.370	0.369	1.000	1.000

Notes. Seller columns use locked single-configuration exposures and report simulated and exact population summaries. Buyer columns use the modeled CES baseline-share exposure. All entries are outputs under the assumed DGP and shares, not observed hedge performance.

align the buyer's factor loading with the broad benchmark and shrink her share-squared residual variance. The number neither validates buyer flexibility or equal weights independently nor measures realized cost hedging: it rests on assumed unit prices and elasticity, a static share vector, and the factor structure used to build the comparison, and it excludes contract execution and expenditure-measurement basis. Against the exact chain-linked CES cost path, the seeded Divisia diagnostic (Internet Appendix Section IA.IV.2) is tight under symmetric small shocks but deteriorates under skewed ones. The scenario illustrates the conditional asymmetry, not its empirical prevalence: concentrating buyer shares or adding commitment and measurement risk narrows the gap, and no hub is selected because contractible buyer expenditure weights do not exist in the current data.

4.1.3. Spanning, Menu Size, Sensitivity, and Premia

Proposition 3 ranks systematic directions by the eigenvalues of $\widetilde{M}$; the associated Ky Fan curve is an unrestricted ceiling unless feasible nonnegative, sum-to-one indices span the relevant directions. Feasibility constrains settlement weights, not hedge positions. Signed positions across feasible contracts can span directions no single feasible weight vector equals, so the generation indices jointly span the three-factor space here while trailing eigenvectors remain infeasible standalone indices. The relaxed curve saturates after three directions yet plateaus low because factor contracts do not span configuration residuals. A coarse index can exceed the ceiling for its own constituents through the finite own-weight term; that is a different estimand, not a contradiction.

Liquidity converts the eigenvalue sequence into a scenario menu count. Trailing eigenvalues are universally $O(\phi)$ and $\Theta(\phi)$ only when the seller residual block is nondegenerate, and γ implements an internal cost-benefit rule, not an order-book or listing inference. The spanning ceiling and menu-count grid (Figures IA.1 and IA.2, Internet Appendix Section IA.IV.2) support directional comparisons only.

The sensitivity grid varies buyer concentration and heterogeneity, substitution, economic residual correlation, seller share, liquidity cost, seller weights, and measurement-error allocation. Across the finite joint grid, buyer HE spans 0.90 to 1.00, seller HE 0.28 to 0.73, and the menu count one to two, a scenario envelope, not a confidence interval, empirical range, or global bound; Table IA.6 in Internet Appendix Section IA.IV.2 reports the full grid. Correlated economic residuals can raise seller HE; publisher measurement error moves HE for both clienteles and is not seller economic basis; concentrated or heterogeneous buyer shares weaken the constructed alignment behind near-unit buyer HE.

In Proposition 7, the premium is signed net hedging pressure divided by speculative risk tolerance, so seller and buyer systematic terms may offset or reinforce. The baseline produces a smaller absolute broad premium than several narrow premia, but neither positivity nor that ranking is general: a large co-loading can reverse a factor ranking, and normalization matters. The CARA equilibrium compares signed pressure under a common normalization without estimating speculative capacity, and γ is a menu-design assumption, not the risk-tolerance denominator; the exercise shows how pressure can change sign and rank and forecasts no price level. The baseline decomposition (Table IA.7), risk-capacity curves (Figure IA.3), and ranking counterexample are in Internet Appendix Section IA.IV.2.

4.1.4. Authority Boundary

Every conclusion above is conditional on a short published-index record and an assumed market. H200 inputs are partly interpolated; loadings, residual independence, exposure weights, buyer structure, ϕ, θ, and γ are not estimated. The factor/residual split is matched from publisher indices that confound economics and methodology. Finite-menu own-weight effects differ from the diversified theoretical limit, and static buyer shares differ from a chain-linked cost path. Sensitivity explores a finite design, not all plausible data-generating processes. These boundaries allow the laboratory to explain mechanisms and estimands, but not to validate the theory, select a settlement hub, infer a market menu, or forecast traded futures premia.

We combine the theory of Section 3, the illustrative simulation of Section 4.1, and the descriptive evidence of Section 2 into an ex ante design problem that does not select a settlement product: the published record lacks the configuration-level exposures and buyer expenditure shares needed to evaluate the social loss in equation (12).

4.2. Candidate Settlements

The observable candidates are single-name rental indices for the A100, H100, and B200, two of them provider-specific, plus a short token series; the weak same-hardware comovement of Section IA.VI shows these are not interchangeable measurements. A multi-name basket would require a public charter fixing constituents, weights, rebalancing, governance, and revisions, and no such validated charter exists. The data contain no regional or configuration-level transactions, and the only performance adjustment available leaves every second-moment comparison unchanged (Internet Appendix Section IA.III, equation (IA.4)). The model clarifies two targets without supplying that charter. At the buyer-only limit, Divisia weights are what the model recommends once buyer expenditure shares are known, and those shares are not observed. The leading Rayleigh direction is

a ceiling on what any single index can achieve; the calibration's weights reach it, which shows the ceiling is attainable in the model and not that any published index attains it.

4.3. What the Spot Record Can Rank

The spot record supports only cross-tracking diagnostics. The per-index leave-one-out hedge effectiveness in Table IA.18 is low, averaging 0.05 across the four indices and 0.09 in the balanced sample (Internet Appendix Table IA.17); these describe cross-tracking among published indices, not the buyer's hedge, and become the exposure-weighted loss $L(k)$ only with the economic sizes of the hedged positions. Single-name candidates include their own series and so carry a mechanical unit-HE term that leaves them incomparable to the broad rows, and principal-components analysis chooses neither a feasible contract count nor settlement weights. Settlement timing is a separate margin: averaging trades tracking against lag but does not measure manipulation resistance, which requires order-level data (Internet Appendix Section IA.V.4).

4.4. A Conditional Architecture

The theory motivates a hub-and-basis architecture only as a conditional design. If a future publisher validates a multi-name settlement with weights in the simplex $\mathcal{W}$, that hub could pool the common component of buyer and seller demand, and a sparse basis layer could address residual provider or configuration exposures where clientele demand supports listing. The simulation illustrates these mechanisms but does not select them: its menu and premium results are scenario outcomes, so neither a universal premium sign nor a narrow-versus-broad ordering follows from the theory. The evidence therefore supports design requirements rather than a selection: a future hub would need a reproducible charter, auditable inputs, stable cross-publisher units, out-of-sample cross-tracking, and direct exposure evidence; until then, equal weight, Divisia weights, single-name indices, and

published baskets remain candidates. Reserved take-or-pay agreements still hedge some physical capacity risk at a delivery tier the spot evidence cannot rank, and futures data, when trading begins, will be needed to estimate liquidity, premia, and whether a basis layer is viable.

5. Conclusion

A compute future must name its settlement benchmark before any evidence identifies which one hedgers need. A locked seller earns one configuration's rental rate; a flexible buyer faces a basket of rates. That asymmetry limits any benchmark: an index built for one side leaves the other exposed. It governs how far buyer demand alone supports a menu (Section 3.7) and how premia reflect net hedging pressure (Section 3.8). Under dispersed buyer shares and the model's factor structure, an aligned benchmark diversifies the buyer's residual while leaving sellers exposed to their own configurations. These restrictions hold inside the model, on assumptions the compute application does not establish.

Under those assumptions, an external benchmark aligned with the modeled buyer's basket hedges 99.7% of that buyer's variance but 28.2% of calibrated seller variance, and the relaxed menu contains one or two directions across the joint sensitivity envelope (Section 4.1). The inputs are illustrative, so these values show how the mechanism can operate without validating the calibration or identifying an empirical hub.

The descriptive record leaves settlement design open. Two publishers' H100 indices correlate 0.17 weekly (Section 2.4), so publisher identity and index construction are explicit contract terms rather than details.

Moving from restrictions to selection requires evidence the present record does not contain. Buyer expenditure shares and direct buyer and seller positions would test the predicted incidence and make the relevant social loss measurable. A reproducible index charter, prespecified out-of-sample

and leave-one-out tracking, and richer provider coverage would test settlement performance without own-series mechanics. Intraday transaction data paired with a specified counterfactual attack would be needed to study manipulation resistance; averaging-window tracking alone does not identify it. An exchange-traded history would reveal forward premia, participation, and the liquidity conditions that govern whether additional basis contracts are worth listing. Until those objects are observed, the paper restricts candidate designs but cannot rank them.

References

Aldasoro, Iñaki, Sebastian Doerr, and Daniel Rees, 2026, Financing the AI boom: From cash flows to debt, BIS Bulletin No. 120, Bank for International Settlements.

Allen, Franklin, and Douglas Gale, 1994, *Financial Innovation and Risk Sharing* (MIT Press, Cambridge, MA).

Anderson, Ronald W., and Jean-Pierre Danthine, 1981, Cross hedging, *Journal of Political Economy* 89, 1182–1196.

Assody, Amine, 2026, Pricing compute futures: Forward curve and volatility for a non-storable, depreciating commodity, SSRN Working paper 6926798.

Bessembinder, Hendrik, and Michael L. Lemmon, 2002, Equilibrium pricing and optimal hedging in electricity forward markets, *Journal of Finance* 57, 1347–1382.

Buyya, Rajkumar, David Abramson, Jonathan Giddy, and Heinz Stockinger, 2002, Economic models for resource management and scheduling in Grid computing, *Concurrency and Computation: Practice and Experience* 14, 1507–1542.

Cha, Doohwi, 2026, Is compute the new oil? Cross-asset risk and the case for a dedicated market, SSRN Working paper 6829278.

Diewert, W. Erwin, 1976, Exact and superlative index numbers, *Journal of Econometrics* 4, 115–145.

Duffie, Darrell, and Piotr Dworczak, 2021, Robust benchmark design, *Journal of Financial Economics* 142, 775–802.

Duffie, Darrell, Piotr Dworczak, and Haoxiang Zhu, 2017, Benchmarks in search markets, *Journal of Finance* 72, 1983–2044.

Duffie, Darrell, and Matthew O. Jackson, 1989, Optimal innovation of futures contracts, *Review of Financial Studies* 2, 275–296.

Duffie, Darrell, and Rohit Rahi, 1995, Financial market innovation and security design: An introduction, *Journal of Economic Theory* 65, 1–42.

Ederington, Louis H., 1979, The hedging performance of the new futures markets, *Journal of Finance* 34, 157–170.

Friedman, Dave, 2026, Can compute commoditize if it's not fungible?, Substack, https://davefriedman.substack.com/p/can-compute-commoditize-if-its-not.

Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein, 1993, Risk management: Coordinating corporate investment and financing policies, *Journal of Finance* 48, 1629–1658.

Gibson, Rajna, and Eduardo S. Schwartz, 1990, Stochastic convenience yield and the pricing of oil contingent claims, *Journal of Finance* 45, 959–976.

Grégoire, Vincent, 2026, Investing in artificial general intelligence, Working paper, SSRN.

Hausman, Jerry, 2003, Sources of bias and solutions to bias in the consumer price index, *Journal of Economic Perspectives* 17, 23–44.

Horn, Roger A., and Charles R. Johnson, 2013, *Matrix Analysis*, 2nd ed. (Cambridge University Press, Cambridge).

Hulten, Charles R., 1973, Divisia index numbers, *Econometrica* 41, 1017–1025.

Johnson, Leland L., 1960, The theory of hedging and speculation in commodity futures, *Review of Economic Studies* 27, 139–151.

Kaldor, Nicholas, 1939, Speculation and economic stability, *Review of Economic Studies* 7, 1–27.

Kavussanos, Manolis G., and Nikos K. Nomikos, 2003, Price discovery, causality and forecasting in the freight futures market, *Review of Derivatives Research* 6, 203–230.

Kenyon, Chris, and Giorgos Cheliotis, 2001, Stochastic models for telecom commodity prices, *Computer Networks* 36, 533–555.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Nielsen, Norman R., 1970, The allocation of computer resources—is pricing the answer?, *Communications of the ACM* 13, 467–474.

Paulson, Tim, 2025, Paper steaks: Live cattle futures markets and the financial revolution of 1964, *Enterprise & Society* 26, 619–650.

Pérez-González, Francisco, and Hayong Yun, 2013, Risk management and firm value: Evidence from weather derivatives, *Journal of Finance* 68, 2143–2176.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Routledge, Bryan R., Duane J. Seppi, and Chester S. Spatt, 2000, Equilibrium forward curves for commodities, *Journal of Finance* 55, 1297–1338.

Schwartz, Eduardo S., 1997, The stochastic behavior of commodity prices: Implications for valuation and hedging, *Journal of Finance* 52, 923–973.

Smith, Clifford W., and René M. Stulz, 1985, The determinants of firms' hedging policies, *Journal of Financial and Quantitative Analysis* 20, 391–405.

Sutherland, Ivan E., 1968, A futures market in computer time, *Communications of the ACM* 11, 449–451.

Tashjian, Elizabeth, 1995, Optimal futures contract design, *Quarterly Review of Economics and Finance* 35, 153–162.

Telser, Lester G., 1958, Futures trading and the storage of cotton and wheat, *Journal of Political Economy* 66, 233–255.

Weagley, Daniel, 2019, Financial sector stress and risk sharing: Evidence from the weather derivatives market, *Review of Financial Studies* 32, 2456–2497.

Wolski, Rich, Neil T. Spring, and Jim Hayes, 1999, The Network Weather Service: A distributed resource performance forecasting service for metacomputing, *Future Generation Computer Systems* 15, 757–768.

Working, Holbrook, 1949, The theory of price of storage, *American Economic Review* 39, 1254–1262.

Xing, Yicai, 2026, AI token futures market: Commoditization of compute and derivatives contract design, Working paper, arXiv:2603.21690.

Internet Appendix for

"Compute Futures: Who Is Left Exposed?"

Vincent Grégoire, Rubén Fernández-Fuertes

HEC Montréal

August 3, 2026

This Internet Appendix contains supplementary material for "Compute Futures: Who Is Left Exposed?" Section IA.I collects the proofs of Propositions 1–8, the CES derivations behind the two-population model, the contract-level premium cases, and the scope notes for the machine-checked Lean verification. Section IA.II collects model discussion demoted from the main text: the factor-to-institution mapping, the drift/obsolescence separation, and the non-storability argument in full. Section IA.III documents data construction and the boundary between observed, adjusted, and unobserved inputs. Section IA.IV reports the simulation's parameter table, its cross-configuration distribution of hedging effectiveness, and the sample-length experiment that quantifies estimation noise. Section IA.V contains spot-price dynamics, uncertainty, publisher-discrepancy, and relative-price-persistence exhibits. Section IA.VII examines the settlement unit (GPU-hours versus tokens). Section IA.VIII reports additional firm-exposure exhibits. Section IA.IX collects the robustness checks.

IA.I. Proofs of the Propositions

The accompanying Lean 4 package, developed against *Mathlib*, checks the specified finite-dimensional components of every proposition in the main theory section. It carries no probability space: second moments enter as real matrices. The package README gives the proposition-by-proposition dictionary; this scope note records the boundary needed to read the proofs below. Two limits apply throughout. Lean does not infer simplex attainability from matrix rank, so feasible-index inequalities enter with their hypotheses stated explicitly. The CES buyer price index and the CARA-normal reduction are modeling assumptions and are not formalized at all.

Proposition 1 is mechanized in full. For Proposition 2, Lean proves the unrestricted Rayleigh optimum, the simplex upper bound, and equality under normalized or scale-free feasible leading-direction attainment. The convex-hull program and its existence conditions remain paper arguments. For Proposition 3, Lean proves the Ky Fan relaxation, the feasible-menu lower bound, equality conditional on attainable selected directions, and the linear-cost threshold under an assumed cost sequence. The per-hedger multivariate-regression reduction is the paper-level bridge to that residual objective.

For Proposition 4, Lean distinguishes exact shared-basket hedging from the factor-only Pythagorean shortfall and proves the augmented population loss split. The CES path, local-horizon approximation, switching and reservation restrictions, heterogeneous buyers, measurement error, and the $O(1/N)$ reading are outside the package. For Proposition 5, Lean checks simplex feasibility and loading attainment at $w = s$, plus the codirectional relaxed result; it makes no empirical contractibility, hub selection, or eigenvector-limit claim.

For Proposition 6, Lean proves the representative-buyer structural facts and derives the count sandwich and small-ϕ limit from the Weyl bounds as explicit hypotheses. It does not formalize matrix Weyl inequalities, eigenvector rotation, or feasible attainment of the selected eigenspaces.

For Proposition 7, Lean proves agent optimality, clearing, and that a positive prefactor preserves the sign of $\phi\rho_S - (1-\phi)$. The covariance aggregation that yields the broad-hub premium, including the vanishing of idiosyncratic cross terms in the well-diversified limit, and the exact narrow specialization are derived here without a machine check. For Proposition 8, Lean proves the scalar reinforcement identity and supplies a counterexample to unrestricted cross-factor ranking; the factor normalization and aggregation used below remain paper-level.

Proof of Proposition 1. Equation (9) is the population regression coefficient of $\Delta \log p_{i,t}$ on $\Delta \log I_t^k$, obtained from the first-order condition of the variance, using $\mathrm{Cov}(\Delta \log p_{i,t}, \Delta \log I_t^k) = \beta_i' \Sigma_f \beta^k + w_i^k \sigma_i^2$ and $\mathrm{Var}(\Delta \log I_t^k) = (\beta^k)' \Sigma_f \beta^k + \tau_k^2$ from equations (5)–(7). The minimized residual variance is $V_i(1-\rho_{i,k}^2)$, which gives equation (10). Setting $\tau_k^2 = 0$ and $w_i^k = 0$ and writing the covariance as the inner product $\tilde{\beta}_i' g$ with $g = \Sigma_f^{1/2} \beta^k$ yields $(\tilde{\beta}_i' g)^2/(g'g) = \|\Pi_g \tilde{\beta}_i\|^2$, which is equation (11). □

Proof of Proposition 2. By Proposition 1, $V_i(1 - HE_i^k) = V_i - (\beta_i' \Sigma_f \beta^k)^2/[(\beta^k)' \Sigma_f \beta^k]$ under diversification, so $L(k) = \sum_i \eta_i V_i - G(\beta^k)$ with G as in equation (13), since $\sum_i \eta_i (\beta_i' \Sigma_f \beta^k)^2 = (\beta^k)' \Sigma_f M_\beta \Sigma_f \beta^k$. Writing $g = \Sigma_f^{1/2} \beta^k$ gives $G = g' \widetilde{M} g/(g'g)$, a Rayleigh quotient maximized at λ_1 by g equal to the leading eigenvector. Because $\sum_i \eta_i V_i = \operatorname{tr} \widetilde{M} + \sum_i \eta_i \sigma_i^2$, equation (14) follows.

Feasible index loadings are restricted to $\mathcal{K} = \{B'w : w \in \mathcal{W}\} = \mathrm{conv}\{\beta_1, \dots, \beta_N\}$. When $0 \notin \mathcal{K}$, continuity of G on this compact set gives the maximum $G_{\mathcal{W}}^*$ in equation (15). If zero belongs to the hull, imposing a strictly positive benchmark-variance floor gives the same compactness argument on the restricted set whenever that set is nonempty. The Rayleigh bound gives $G_{\mathcal{W}}^* \le \lambda_1$, with equality exactly when a nonzero feasible whitened loading lies in E_1. Subtracting this constrained captured variance from $\sum_i \eta_i V_i$ yields the feasibility gap in equation (15).

If exposures are codirectional and $\sum_i \eta_i c_i \neq 0$, the simplex weight $w = \eta$ produces a nonzero loading in the common direction and attains the relaxation. Full row rank of B' instead guarantees

representation only by an unrestricted signed vector. It does not place that vector in $\mathcal{W}$ and therefore does not establish a contractible index. □

Proof of Proposition 3. (a) For a fixed m-dimensional subspace $\mathcal{S}$, the best combination of the m futures neutralizes the component of the hedger's whitened exposure $\tilde{\beta}_i$ lying in $\mathcal{S}$, leaving systematic residual $\|\Pi_{\mathcal{S}^\perp}\tilde{\beta}_i\|^2$ and the unhedgeable σ_i^2. Summing, $\sum_i \eta_i\sigma_i^2 + \mathrm{tr}(\Pi_{\mathcal{S}^\perp}\widetilde{M})$, since $\widetilde{M} = \sum_i \eta_i\tilde{\beta}_i\tilde{\beta}_i'$. By the Ky Fan / Rayleigh–Ritz theorem (Horn and Johnson, 2013, Chapter 4), $\mathrm{tr}(\Pi_{\mathcal{S}^\perp}\widetilde{M})$ is minimized over m-dimensional $\mathcal{S}$ by the span of the top m eigenvectors, with value $\sum_{l>m}\lambda_l$, giving equation (16). Setting $m = \mathrm{rank}(\widetilde{M})$ drives the systematic term to zero.

For a feasible price-index menu, its whitened loading span belongs to $\mathfrak{S}_m(\mathcal{W})$. The unrestricted Ky Fan value is consequently a lower bound on residual loss. Equality holds when the selected top-m space belongs to that family, including any basis choice within a tied eigenspace. Rank alone does not give that equality: under the baseline simplex, constituent weights e_i are feasible, and if $\mathrm{rank}(B) = r$, a possibly concentrated menu of r constituent indices spans factor space without inheriting the well-diversified approximation.

(b) Conditional on feasible attainment, increasing $m-1$ to m lowers the relaxed residual by λ_m. The marginal cost $\Delta\Phi(m) = \Phi(m) - \Phi(m-1)$ is nondecreasing only because increasing, discretely convex Φ is assumed. One-step comparisons give the stated crossing; with $\Phi(m) = \gamma m$, the smallest minimizer lists exactly the directions with $\lambda_l > \gamma$. If the selected spaces are unattainable, the designer instead compares constrained residual reductions and no generic eigenvalue threshold follows. □

IA.I.1. The Buyer's Cost Function: CES Derivations

This subsection derives the demand-side objects that Assumption 1 and equation (19) use as primitives. Let the buyer produce effective compute from quantities $x_1, \ldots, x_N$ through the CES

aggregator $X = \left(\sum_i (a_i x_i)^{(\theta-1)/\theta}\right)^{\theta/(\theta-1)}$, where $a_i > 0$ converts configuration i's GPU-hours into effective compute and $\theta > 0$ is the elasticity of substitution. Minimizing expenditure $\sum_i p_i x_i$ subject to $X \geq 1$ gives the unit cost function and, by Shephard's lemma, the expenditure shares,

$$c(p) \;=\; \Big(\sum_i (p_i/a_i)^{1-\theta}\Big)^{1/(1-\theta)}, \qquad s_i \;=\; \frac{\partial \log c}{\partial \log p_i} \;=\; \frac{(p_i/a_i)^{1-\theta}}{\sum_j (p_j/a_j)^{1-\theta}}, \tag{IA.1}$$

with the Cobb–Douglas case $\theta = 1$ read as the limit of the shares, $s_i \to 1/N$: the aggregator carries no distribution weights, so the unit cost itself has no $\theta \to 1$ limit, while the log-cost changes and the Hessian below do. The shares are nonnegative and sum to one.

Along a price path, Shephard's lemma gives the exact chain-linked differential $d\log c(p_t) = \sum_i s_i(p_t) d\log p_{i,t}$. Integrating requires updating shares along the path. Fixing instead a hedge horizon H and baseline shares $s_i = s_i(p_0)$ gives a static first-order exposure. Because $\log c$ is smooth in log prices with Hessian $(1-\theta)(\operatorname{diag}(s) - ss')$, its second-order expansion is

$$\Delta \log c_t \;=\; \sum_i s_i \,\Delta \log p_{i,t} \;+\; \frac{1-\theta}{2}\Big[\sum_i s_i \,(\Delta \log p_{i,t})^2 - \Big(\sum_i s_i \,\Delta \log p_{i,t}\Big)^2\Big] \;+\; O(\|\Delta \log p_t\|^3), \tag{IA.2}$$

where the price changes run over H. The bracket is a share-weighted variance, so the substitution term has the sign of $1-\theta$. Under a common shock-scale rescaling, the linear exposure is first order and the substitution term second order. Their covariance is third order and the quadratic term's own variance is fourth order; symmetry can eliminate the third-order contribution, but skewness need not. The main propositions retain only the local linear exposure, which under equation (5) has the loading and residual variance in equation (19). They do not claim static shares are exact away from that local horizon. Nor do these economic price paths model publisher measurement error, which can lower hedge quality for either clientele.

Two limits delineate what the elasticity does and does not govern. As $\theta \to \infty$ the shares of

equation (IA.1) concentrate on the configuration with the lowest quality-adjusted price p_i/a_i: the buyer rides the frontier, her idiosyncratic variance moves from the dispersed-share level $\sum_i s_i^2\sigma_i^2$ to a single configuration's σ_i^2, and her generation-factor loading becomes the frontier configuration's rather than a share-weighted average, so she loses her diversification and becomes locally seller-like. In the opposite, Leontief limit $\theta \to 0$ the quantity mix loses its price responsiveness: the buyer is left holding a fixed *bundle* whose dispersed shares still constitute a diversified index, so zero elasticity does not collapse the two-population structure. What recovers the single-configuration symmetric model of Sections 3.4–3.6 is a degenerate share vector, all expenditure pinned on one configuration, as a prohibitive within-horizon switching cost would force (Section 3.7): it is infinite switching cost, not the Leontief limit, that makes the buyer a seller.

IA.I.2. Proofs of Propositions 4–7

Proof of Proposition 4. Under Assumption 1 and the Divisia approximation of Section IA.I.1, the buyer's cost change aggregates equation (5) to $\Delta \log c_t = \alpha^D + (\beta^D)' f_t + u_t^D$ with $\beta^D = B's$, $u_t^D = \sum_i s_i \varepsilon_{i,t}$, and $\mathrm{Var}(u_t^D) = \sigma_D^2 = \sum_i s_i^2\sigma_i^2$, so her total variance is $V_D = \|\tilde{\beta}^D\|^2 + \sigma_D^2$ with $\tilde{\beta}^D = \Sigma_f^{1/2}\beta^D$.

(a) Applying the covariance calculation of Proposition 1 to the static basket gives $\mathrm{Cov}(X_D^H, \Delta_H \log I^k) = (\beta^D)'\Sigma_f\beta^k + \sum_i s_i w_i^k \sigma_i^2$ and the exact expression in equation (21). If $w^k = s$ and the benchmark uses the same constituent price series, then $\Delta_H \log I^k = X_D^H$ within the maintained static exposure, so their squared correlation and exact HE equal one.

(b) For an independently measured, well-diversified benchmark, the own-weight term and τ_k^2 drop. Writing $g = \Sigma_f^{1/2}\beta^k$ gives $HE_D^{k,0} = \|\Pi_g \tilde{\beta}^D\|^2/V_D$. The Pythagorean identity then yields equation (22). A loading match removes only the projection mismatch, leaving $HE_D^{k,0} = 1 - \sigma_D^2/V_D$; it is not the exact self-basket result in part (a).

(c) Append the buyer to the hedger population as an $(N+1)$st exposure with loading β^D, idiosyncratic variance σ_D^2, and exposure weight $1-\phi$, the sellers carrying weights $\phi\,\eta_i^S$. The design matrix of equation (13) for the augmented population is then exactly $\widetilde{M}(\phi)$ of equation (20), and Proposition 2 applied to the augmented population turns the relaxed loss in equation (14) into $L^*_{\mathrm{rel}}(\phi) = [\phi \sum_i \eta_i^S \sigma_i^2 + (1-\phi)\,\sigma_D^2] + \sum_{l\geq 2} \lambda_l(\widetilde{M}(\phi))$, which is equation (23). Under dispersed shares ($\max_i s_i = O(1/N)$), $\sigma_D^2 \leq (\max_i s_i) \max_i \sigma_i^2 = O(1/N)$, the rate quoted in the text. □

For heterogeneous buyer types, expanding around $\bar{\beta}^D = \sum_a \nu_a \beta^{D,a}$ gives $\sum_a \nu_a \beta^{D,a}(\beta^{D,a})' = \bar{\beta}^D(\bar{\beta}^D)' + H_D$. Thus H_D contributes buyer-side directions unless type loadings are codirectional. This is a second-moment statement about economic baskets. It neither restores exact static shares over long horizons nor assigns settlement-series measurement error to one side of the market.

Proof of Proposition 5. The index loading is linear in the weights, $\beta^k = B' w^k = \sum_i w_i^k \beta_i$, so setting $w^k = s$ attains the buyer's loading exactly, $\beta^k = B's = \beta^D$. The shares are nonnegative and sum to one, hence $s \in \mathcal{W}$. At $\phi = 0$ the augmented design matrix has the sole nonzero direction $\tilde{\beta}^D$, so this feasible loading attains both the Rayleigh relaxation and the constrained objective. At positive ϕ, it remains feasible but need not minimize population loss.

If $\beta_i = c_i \bar{\beta}$ for all i, then $\beta^D = (\sum_i s_i c_i)\bar{\beta}$ is codirectional as well, so the buyer-augmented population of the previous proof is codirectional with coefficients $(c_1, \ldots, c_N, \sum_i s_i c_i)$. The codirectional case of Proposition 2 applied to the augmented population makes $\widetilde{M}(\phi)$ rank at most one in the common direction. Any feasible index with a nonzero loading in that direction therefore attains the relaxation for every ϕ. No eigenvector-continuity limit or empirical hub-selection claim is needed. □

Proof of Proposition 6. The two ingredients are structural facts about equation (20) and Weyl's eigenvalue inequalities, and we state the latter explicitly as the inputs they are. The buyer block

$(1-\phi)\,\tilde{\beta}^D(\tilde{\beta}^D)'$ is positive semidefinite of rank at most one, and $\widetilde{M}(\phi)$ is positive semidefinite with $\operatorname{tr}\widetilde{M}(\phi) = \phi \operatorname{tr}\widetilde{M}_S + (1-\phi)\|\tilde{\beta}^D\|^2$. For a Hermitian matrix H and a positive-semidefinite perturbation P of rank at most one, Weyl's inequalities give, for the descending eigenvalues, $\lambda_l(H) \le \lambda_l(H+P)$ for every l and $\lambda_l(H+P) \le \lambda_{l-1}(H)$ for $l \ge 2$; these interlacing bounds are classical, and they enter the machine-checked counting arguments below as hypotheses rather than as formalized theorems (see the scope note). Applying them with $H = \phi\widetilde{M}_S$ and $P = (1-\phi)\,\tilde{\beta}^D(\tilde{\beta}^D)'$ and using the homogeneity $\lambda_l(\phi\widetilde{M}_S) = \phi\,\lambda_l(\widetilde{M}_S)$ yields equation (24). (a) For $l \ge 2$ the sandwich pins $\lambda_l(\widetilde{M}(\phi))$ between $\phi\,\lambda_l(\widetilde{M}_S)$ and $\phi\,\lambda_{l-1}(\widetilde{M}_S)$. It is always $O(\phi)$. It is $\Theta(\phi)$ when $\lambda_l(\widetilde{M}_S) > 0$, while $\lambda_{l-1}(\widetilde{M}_S) = 0$ forces it to zero. Intermediate rank-deficient cases retain only the upper-order statement. (b) Under the linear threshold rule of Proposition 3(b), $m^*_{\mathrm{rel}}(\phi) = \#\{l : \lambda_l(\widetilde{M}(\phi)) > \gamma\}$. For the lower half of equation (25), if $\lambda_l(\widetilde{M}_S) > \gamma/\phi$ then $\lambda_l(\widetilde{M}(\phi)) \ge \phi\lambda_l(\widetilde{M}_S) > \gamma$, so every seller survivor survives on the menu and $n^S(\gamma/\phi) \le m^*_{\mathrm{rel}}(\phi)$. For the upper half, if $\lambda_l(\widetilde{M}(\phi)) > \gamma$ for some $l \ge 2$ then $\phi\lambda_{l-1}(\widetilde{M}_S) > \gamma$, so the map $l \mapsto l-1$ injects the menu survivors beyond the first into the seller survivors and $m^*_{\mathrm{rel}}(\phi) \le 1 + n^S(\gamma/\phi)$. For the single-hub limit, the trailing eigenvalues obey $\lambda_l(\widetilde{M}(\phi)) \le \phi\,\lambda_1(\widetilde{M}_S)$ for $l \ge 2$. Because $\gamma > 0$, if $\lambda_1(\widetilde{M}_S) = 0$ the seller block vanishes; otherwise the trailing eigenvalues fall below γ whenever $\phi < \gamma/\lambda_1(\widetilde{M}_S)$. Moreover, evaluating the Rayleigh quotient along $\tilde{\beta}^D$ gives $\lambda_1(\widetilde{M}(\phi)) \ge (1-\phi)\|\tilde{\beta}^D\|^2$. The maintained condition $\lambda_1(\widetilde{M}(0)) = \|\tilde{\beta}^D\|^2 > \gamma$ therefore keeps the leading eigenvalue above γ for all sufficiently small ϕ. Combining the two bounds produces a positive $\bar{\phi}$ below which exactly one contract survives: $m^*_{\mathrm{rel}}(\phi) = 1$. These are relaxed counts; they describe a constrained price-index menu only when the selected spaces are feasibly attainable. □

The sandwich of equation (25) is also the strongest statement available: $m^*_{\mathrm{rel}}(\phi)$ need not be globally monotone in ϕ, and a two-configuration example makes the point. Work with the reduced objects the design problem depends on. Take two whitened factor directions, a seller block with

eigenvalues $\lambda_1(\widetilde{M}_S) = 2\gamma$ and $\lambda_2(\widetilde{M}_S) = 0$, and a buyer direction along the seller block's null eigenvector with $\|\tilde{\beta}^D\|^2 = 3\gamma$; both are realized by two configurations whose whitened loadings are linearly independent, with sell-side demand concentrated on the first and the buyer's shares chosen so that her whitened loading is orthogonal to it. Then $\widetilde{M}(\phi)$ is diagonal in this basis with eigenvalues $3\gamma(1-\phi)$ and $2\gamma\phi$, so $m^*_{\text{rel}}(\phi) = \mathbf{1}\{\phi < 2/3\} + \mathbf{1}\{\phi > 1/2\}$: a single contract for $\phi < 1/2$, two contracts on $(1/2,\, 2/3)$, and a single contract again for $\phi > 2/3$, as the growing seller direction crosses the threshold while the shrinking buyer-weighted hub eigenvalue drops out. The count rises and then falls while the sandwich (25) holds throughout. The provable statement is the relaxed sandwich and the threshold-conditional small-ϕ limit. Feasible attainment is separate.

Proof of Proposition 7. Let x denote the vector of settlement values of the m listed contracts, $\Omega \stackrel{\text{def}}{=} \text{Var}(x)$ its positive definite covariance matrix, and $\pi \stackrel{\text{def}}{=} \mathbb{E}[x] - F$ the premium vector, as in equation (1). Agent a holds endowment payoff e^a (configuration revenue for a seller, the negative of realized cost for the buyer, zero for a speculator), jointly normal with x, and chooses a futures position $h \in \mathbb{R}^m$ to maximize CARA expected utility of terminal wealth $W = e^a + h'(x - F)$. Under joint normality this is the mean-variance program: up to terms not involving h, $\mathbb{E}[W] - \frac{\gamma_a}{2}\,\text{Var}(W)$ equals $U(h) = h'\pi - \frac{\gamma_a}{2}(h'\Omega h + 2\,h'c_a)$, where $c_a \stackrel{\text{def}}{=} \text{Cov}(x, e^a)$. This CARA-normal reduction, and the identification of c_a with the settlement-endowment covariance vector, are the modeling step; everything after it is linear algebra. Completing the square gives the exact identity $U(h) - U(h^*) = -\frac{\gamma_a}{2}\,(h - h^*)'\,\Omega\,(h - h^*)$ with $h^* = \gamma_a^{-1}\Omega^{-1}\pi - \Omega^{-1}c_a$, verified by expanding the quadratic form. Since Ω is positive definite and $\gamma_a > 0$, h^* is the unique maximizer: each agent has a unique mean-variance optimal position, linear in π and in c_a (Lemma D0 of the machine-checked package). Zero net supply imposes $\sum_a h^*_a = 0$, that is, $\big(\sum_a \gamma_a^{-1}\big)\,\Omega^{-1}\pi = \Omega^{-1}\sum_a c_a$; multiplying by Ω yields $T\pi = \sum_a c_a$ with $T = \sum_a \gamma_a^{-1}$, which is equation (26): the premium vector is net hedging pressure over aggregate risk tolerance, sellers entering the sum positively and the flexible buyer

negatively. Speculators contribute risk tolerance to T and nothing to the pressure. Scale the two hedging clienteles by Q_S and Q_D with $Q \stackrel{\text{def}}{=} Q_S + Q_D > 0$ and $\phi = Q_S/Q$: the sell side's aggregate endowment is $Q_S \sum_i \eta_i^S \, \Delta \log p_{i,t}$ and the buyer's is $-Q_D \, \Delta \log c_t$. For the broad hub set to the buyers' price ($w^k = s$, well diversified, $\beta^k = \beta^D$), the component of equation (26) on the hub is $T\pi^{\text{hub}} = Q\, D\, [\phi\rho_S - (1-\phi)]$, where $D \stackrel{\text{def}}{=} (\beta^D)'\Sigma_f \beta^D > 0$, $\bar{\beta}^S \stackrel{\text{def}}{=} \sum_i \eta_i^S \beta_i$ is the sell side's exposure-weighted loading, and $\rho_S \stackrel{\text{def}}{=} (\bar{\beta}^S)'\Sigma_f \beta^D / D$ measures the sellers' hedging pressure on the hub per unit of demand share relative to the buyer's; the idiosyncratic cross terms vanish in the well-diversified limit. Since $Q\, D/T > 0$, the hub premium is a positive multiple of the right-hand side of equation (27) and carries its sign; either sign can obtain. □

IA.I.3. Signed Premium, Liquidity, and Obsolescence Cases

The premium identity is component-wise. For a well-diversified contract with loading β^k, its systematic pressure is

$$P^k = (\beta^k)'\Sigma_f(Q_S\bar{\beta}^S - Q_D\beta^D), \tag{IA.3}$$

and $\pi^k = P^k/T$. Setting $\beta^k = \beta^D$ gives the broad sign in equation (27). Neither netting nor segmentation fixes that sign.

For a raw configuration-i contract, the diversified approximation is inappropriate because the contract carries its own residual. Direct covariance aggregation gives

$$T\pi^i = \beta_i'\Sigma_f(Q_S\bar{\beta}^S - Q_D\beta^D) + \sigma_i^2(Q_S\eta_i^S - Q_D s_i),$$

which is equation (28). The residual component is seller-sided, zero, or buyer-sided according to the signed quantity $Q_S\eta_i^S - Q_D s_i$. The systematic component is also net pressure and does not disappear merely because a hub trades alongside the narrow contract.

A basis-spread settlement $x^{b_i} = \Delta \log p_i - h_i x^{\text{broad}}$ obeys $\pi^{b_i} = \pi^i - h_i \pi^{\text{broad}}$ by linearity. A buyer-aligned broad contract removes the buyer's systematic covariance from the orthogonalized spread, but finite-index residuals, seller cross-loadings, and the buyer's own constituent covariance remain. Thus neither the spread premium's sign nor a narrow-versus-broad premium ordering follows without additional restrictions. Any numerical crossing is a scenario calculation, not a theorem or an estimated market threshold.

Liquidity-cost component. In an isolated scalar contract with net flow F_m, settlement variance V_m, and aggregate risk tolerance T, absorption implies $\pi_m \approx F_m V_m / T$. The associated premium bill is $F_m^2 V_m / T$, equation (29). It is nonnegative but need not rise with the menu index: flow and variance can each fall. Identifying it with $\Delta\Phi(m)$ therefore requires the additional sequence restriction $F_{m+1}^2 V_{m+1} \geq F_m^2 V_m$. Without that restriction, or when feasible eigenspaces are unattainable, the designer compares global constrained values. A fixed participation cost is a separate reduced-form scenario input.

Proof of Proposition 8. After projecting obsolescence news on the other factors, orthogonality makes its contribution to contract-k pressure $P^k_{\text{obs}} = \beta^k_{\text{obs}} \sigma^2_{\text{obs}} (Q_S \bar{\beta}^S_{\text{obs}} - Q_D \beta^D_{\text{obs}})$. Under $\bar{\beta}^S_{\text{obs}} < 0$ and $\beta^D_{\text{obs}} \geq 0$, the two clientele terms have the same sign after the buyer endowment enters negatively. Taking absolute values gives equation (30).

For an otherwise comparable factor on which the clienteles co-load, the same gross absolute contributions enter as a difference rather than a sum. Holding gross exposure, factor variance, absolute contract loading, and risk tolerance fixed therefore weakly raises the opposite-sign factor's premium contribution. Changing any of those quantities can reverse a cross-factor ranking, and other factor or finite-index residual terms can offset the contribution in total premia. Anticipated depreciation remains in the forward level and is absent from this covariance argument. Lean proves the scalar sum-versus-difference identity and contains a counterexample to unrestricted ranking; the

orthogonalization and factor aggregation here are paper assumptions. □

The simulation reports signed premia and menu outcomes only for its finite scenario grid. Those quantities illustrate these formulas; they neither estimate the equilibrium objects nor validate a universal premium ranking or per-contract cost.

IA.II. Model Details: Factor Interpretation, Drift, and Non-Storability

This section collects model discussion compressed out of Section 3 of the main text. The propositions it supports are stated there and proved in Section IA.I. Lean 4 checks Proposition 1 in full and the specified relaxed or hypothesis-conditional components of Propositions 2 and 3 (see the scope note opening Section IA.I). The results of Section 3 are comparative statics that hold independently of the length of any price sample, stated as relationships among the factor loadings, covariances, idiosyncratic variances, and exposure weights, objects to be estimated or scenario-specified, and containing no data placeholders.

IA.II.1. The Factor-to-Institution Mapping

The factors of equation (5) collect the systematic forces of Section 3.2: a demand factor, a supply factor, and a generation or obsolescence factor capturing the arrival of new hardware and the repricing of older devices. Equation (5) thus refines the conceptual decomposition (2): the common shocks D_t and S_t and the systematic part of obsolescence $O_{i,t}$ are gathered into $\beta_i' f_t$, with loadings that differ across configurations so that an older generation loads more negatively on the generation factor than a frontier one, while configuration-specific risk $\xi_{i,t}$ splits into a component still driven by the common factors and a purely idiosyncratic residual $\varepsilon_{i,t}$. For a unit hedge ($h = 1$), equation (8)

implies residual variance $(\beta_i - \beta^k)'\Sigma_f(\beta_i - \beta^k) + \sigma_i^2 + \tau_k^2 - 2w_i^k\sigma_i^2$, which increases with the loading mismatch.

IA.II.2. Why the Deterministic Drift Does Not Enter the Hedging Analysis

The depreciation rate δ_i in $\alpha_i = \mu - \delta_i$ is the predictable component of the obsolescence term $O_{i,t}$ in equation (2); because each new generation erodes the value of its predecessors, δ_i is larger for configurations further from the frontier, so obsolescence enters as a downward drift whose magnitude is configuration-specific. The component δ_i is priced, not hedged: it enters the level of the forward through the equilibrium pricing relation of equation (1), where a configuration expected to depreciate faster simply carries a lower forward price, and it drops out of the covariances, the hedge ratios, and the hedging effectiveness, because a constant contributes nothing to a variance. The realized drift estimates are short-sample and imprecise (Section 2.4); their magnitudes are not needed for this algebraic separation. This is why the paper's central friction appears in Propositions 1–3 only through loadings and factor variances and never through α_i: the deterministic part of obsolescence belongs to the price level, the stochastic part to the risk, and separating them is a feature of the modeling rather than an omission.

IA.II.3. Non-Storability and the Absence of No-Arbitrage Discipline

Nonstorability removes the inventory trade that enforces cash-and-carry. It does not eliminate every delivery promise: a forward or reservation can commit future capacity or service even though today's GPU-hour cannot be stored. Standardized cash settlement is instead a design choice that makes heterogeneous obligations fungible and avoids defining a transferable physical-delivery bundle.

For the cash-settled design analyzed in Section 3.1, the forward price is the risk-neutral expectation

of the settlement index,

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}[I_T] = \mathbb{E}_t[I_T] - \pi_{t,T}, \qquad \text{(1 revisited)}$$

the physical expectation less a risk premium for bearing non-diversifiable supply and demand risk. This equilibrium relation replaces the absent storage restriction; nonstorability alone does not determine the premium. Two designs k and k' with the same expected level $F_{t,T}$ can deliver entirely different hedging value, because their loadings β^k and $\beta^{k'}$ span different directions of factor space and imply different HE_i in equation (10).

IA.III. Data Construction and Authority

The empirical panel contains four usable rental indices with unequal histories: Silicon Data A100, H100, and B200, and Ornn H100. The token series is retained only for a short supplementary comparison. The July 2026 vintage incorporates the publishers' then-current restatements. These are published index levels, not underlying transaction panels, invoices, capacity positions, firm revenues, or futures prices.

Repeated daily levels and asynchronous publisher updates attenuate high-frequency comovement. Weekly changes are the workhorse, but aggregation cannot remove methodology differences or create missing overlap. Every empirical table is generated through the package entry point `uv run compute-futures`; manual spreadsheet refreshes are outside the replication path. Publisher methodology versions and licensed-source access remain data-governance inputs rather than economic estimates.

An hour on one generation delivers more compute than an hour on another, so the vendor scalars support a quality adjustment. Let $p_{i,t}$ be the nominal rental rate per GPU-hour and a_i the same scalar of equation (IA.1), here read as configuration i's fixed delivered performance. The adjusted

price is

$$\tilde{p}_{i,t} = \frac{p_{i,t}}{a_i}. \tag{IA.4}$$

Performance is fixed within a generation over the observed sample, so $\log \tilde{p}_{i,t} = \log p_{i,t} - \log a_i$: dividing a price index by that constant shifts its log level but leaves log changes, covariances, hedge effectiveness, and persistence slopes unchanged. The adjustment is a change of units and enters no proposition in Section 3. A future test of performance-adjusted substitution would require a time-varying, workload-specific performance series and an auditable mapping to effective compute. The current fixed scalar supplies no such test.

IA.IV. Simulation: Parameters, the Distribution of Hedging Effectiveness, and Estimation Noise

IA.IV.1. Construction, Provenance, and Input Authority

This subsection records which simulation inputs are estimated, matched, interpolated, or assumed. The laboratory is disciplined by selected moments of the short published-index record; it is not an estimated structural model.

The demand factor is common to the market. Scarcity loadings rise toward the frontier and obsolescence loadings toward older hardware by assumption. The simulation also assigns configuration counts, within-generation variants, seller exposure weights, buyer demand parameters, and the per-contract cost. These choices define scenarios rather than empirical estimates. The H200 has no published price series: its drift and residual volatility are interpolated, while its loadings and exposure are assumed.

Table IA.1 records the available published-index moments used to discipline the exercise; it remains subject to the data scope of main text Section 2.3.

Table IA.1: Published-index moments used in the illustrative laboratory

Provider	A100 Silicon Data	H100 Silicon Data	B200 Silicon Data	H100 Ornn	LLM token Silicon Data
Start date	2024-09-03	2024-09-03	2025-08-01	2024-10-02	2025-12-02
Daily observations	480	480	242	459	155
Zero daily changes (%)	61.7	45.6	19.8	19.4	0.6
Drift (% per year)	−0.3	−11.2	6.2	2.5	76.2
Volatility, daily (%, ann.)	15.0	24.7	29.0	37.0	37.2
Volatility, weekly (%, ann.)	13.6	24.3	29.2	39.2	49.9
Weekly corr. with SD H100	0.11	–	0.37	0.17	0.03

Notes. Drift is the annualized mean weekly log change; volatility is the annualized standard deviation at the stated frequency. The zero-change share measures daily staleness. These are published-index estimates through July 2026, not configuration-market moments. The table is generated directly from the data.

Table IA.2: Authority status of simulation inputs

Input	Baseline implementation	Status
Published-index drift and volatility	Available A100/H100/B200 series	Estimated
Factor and residual variance levels	Selected weekly index moments	Matched
Configurations N	32	Assumed
Variants per generation	8	Assumed
Demand/scarcity/obsolescence loadings	Generation profiles; 0.10 loading s.d.	Assumed
Residual independence ρ_ε	0	Assumed
Seller exposure weights	Generation totals 1:3:2:2	Assumed
Seller share ϕ	0.50	Assumed
Substitution elasticity θ	3.0	Assumed
Liquidity threshold γ	10^{-6}	Assumed
Buyer quality-adjusted unit prices	$1.25, 1.10, 1.05, 1.00$	Assumed
H200 drift and residual volatility	intermediate; idio. vol. 22.2% annualized	Interpolated
H200 factor loadings / exposure	$(1.0, 0.8, 0.2)$; weight 2	Assumed
Measurement-error share q	0 (varied to 0.75)	Assumed

Notes. Estimated inputs come directly from published-index moments; matched inputs are simulation parameters chosen to reproduce selected moments; interpolated inputs fill an unobserved generation; assumed inputs define the illustrative design. A matched or interpolated entry is not an independently estimated structural parameter.

The source series are Bloomberg tickers SDA100RT, SDH100RT, SDB200RT, ORNNH100, and SDLLMTK. Unequal histories, repeated daily levels, and publisher methodology limit the moments they can discipline. Weekly changes reduce, but do not eliminate, staleness. The empirical drifts are individually imprecise and non-monotone in hardware age; cash-cost-floor and frontier-scarcity stories are interpretations, not validated assignments. Table IA.2 makes the resulting authority split explicit.

Measurement subtleties. Table IA.5 in Section IA.IV.2 reports each design both in sample and as an exact population value, the gap being the estimation noise quantified in Section IA.IV.3 below.

The propositions of main text Section 3 characterize designs in the well-diversified limit, whereas the population column prices the finite menu exactly; the difference surfaces where the relaxed Rayleigh direction associated with Proposition 2 does not strictly dominate the exposure-weighted heuristic and where the generation-segmented menu (0.369) exceeds the spanning plateau of Figure IA.1 (0.282). Both reflect the same finite-menu term, coarse indices hedging part of their constituents' idiosyncratic risk through the constituents' own weights, which the diversified limit sets to zero and main text Section 4.1.2 quantifies directly.

IA.IV.2. The Two-Population Layer and the Buyer Variance Identity

The two-population exercises of Sections 4.1 and 4 add a demand-side scenario layer, the sell-side share of hedging demand ϕ, the substitution elasticity θ, and the per-contract liquidity threshold γ, on top of the data-generating process of Section 4.1.1, which is left untouched; Table IA.3 below reports the layer's parameters, generated from the same constants that drive the simulation code. The buyer's baseline expenditure shares are produced by CES demand (equation (IA.1)) over stylized quality-adjusted unit prices under which hardware behind the frontier is more expensive per unit of effective compute, so the shares tilt toward the frontier while remaining dispersed across all configurations.

The buyer's diversification can be read off the parameters symbolically, without reference to any simulated output. Her idiosyncratic variance is the share-squared-weighted identity of equation (19), $\sigma_D^2 = \sum_i s_i^2 \sigma_i^2$, so $\sigma_D^2 \leq (\max_i s_i) \max_i \sigma_i^2$, and under equal shares across the N configurations it equals $1/N$ times the average configuration-level idiosyncratic variance exactly. With the simulation's $N = 32$ configurations and baseline shares dispersed enough that the largest share remains of order $1/N$, the implied σ_D^2 sits more than an order of magnitude below a typical configuration's σ_i^2, the sense in which the modeled buyer is well diversified; Table IA.3 reports the scenario value alongside

Table IA.3: Baseline two-population scenario

Parameter	Value
Sell-side share of hedging demand ϕ	0.50
CES elasticity of substitution θ	3.0
Per-contract liquidity threshold γ	1.0e-06
Buyer idiosyncratic variance σ_D^2	5.83e-06
Mean configuration idiosyncratic variance	1.63e-04
Quality-adjusted unit price p/a; baseline buyer share (per generation):	
A100	1.25; 0.190
H100	1.10; 0.245
H200	1.05; 0.269
B200	1.00; 0.296

Notes. Every entry in this demand-side layer is assumed or implied by assumed buyer prices and CES shares. The buyer variance is the share-squared- weighted idiosyncratic variance; the comparison value is the mean single- configuration idiosyncratic variance. These are scenario parameters, not estimated demand or liquidity primitives.

Table IA.4: Finite-scenario diagnostic for the local Divisia approximation

Shock distribution	Var. order	Resid. order	Rel. var. error	HE shortfall (bp)
Gaussian	4.00	4.00	0.08%	7.83
Student-t(5)	4.01	4.02	0.51%	27.37
Centered lognormal	2.90	3.75	17.14%	109.86

Notes. Variance and residual orders are slopes under a common shock-scale rescaling. Relative variance error compares the exact CES cost variance with the local linear approximation; HE shortfall is in basis points. The seeded Gaussian, Student-$t(5)$, and centered-lognormal experiments form a finite envelope, not a distribution-free guarantee.

the mean configuration-level idiosyncratic variance. The comparison is a property of the CES demand structure and the dispersion of the shares, not an output of the simulated paths, and it is the quantitative counterpart of the $O(1/N)$ reading of equation (19) in the text.

The local Divisia approximation is tight under the symmetric finite shock designs in Table IA.4, whereas the centered-lognormal case has materially larger finite-horizon error. The table is a diagnostic, not a global error bound: symmetric shocks suppress the third-order covariance term, while skewness need not.

Table IA.5 reports each candidate design in sample and as an exact population value under the assumed DGP. The population comparison weights each locked configuration by the assumed seller exposure shares before averaging its exact HE. This differs from the unweighted mean across simulated regressions and from an eigenvalue objective that targets systematic fit: exposure weighting

Table IA.5: Candidate benchmark designs in the illustrative laboratory

	Mean HE	Weighted HE	Median HE	Weighted HE (pop.)	Contracts
Broad equal-weighted	0.297	0.286	0.283	0.282	1
Exposure-weighted	0.291	0.286	0.279	0.283	1
Relaxed factor direction	0.289	0.282	0.282	0.279	1
Generation-segmented	0.395	0.370	0.361	0.369	4

Notes. HE is the fraction of variance eliminated by the minimum-variance hedge in equation (10). Mean, median, and weighted HE are single- regressor estimates on the simulated panel. Weighted HE (pop.) is the exact exposure-weighted population value under the assumed DGP. The relaxed factor direction is a model candidate, not a selected settlement rule.

changes which basis risks matter, and finite benchmark weights change the covariance between each constituent and its own settlement index. These are different estimands, not alternative estimates of one common quantity.

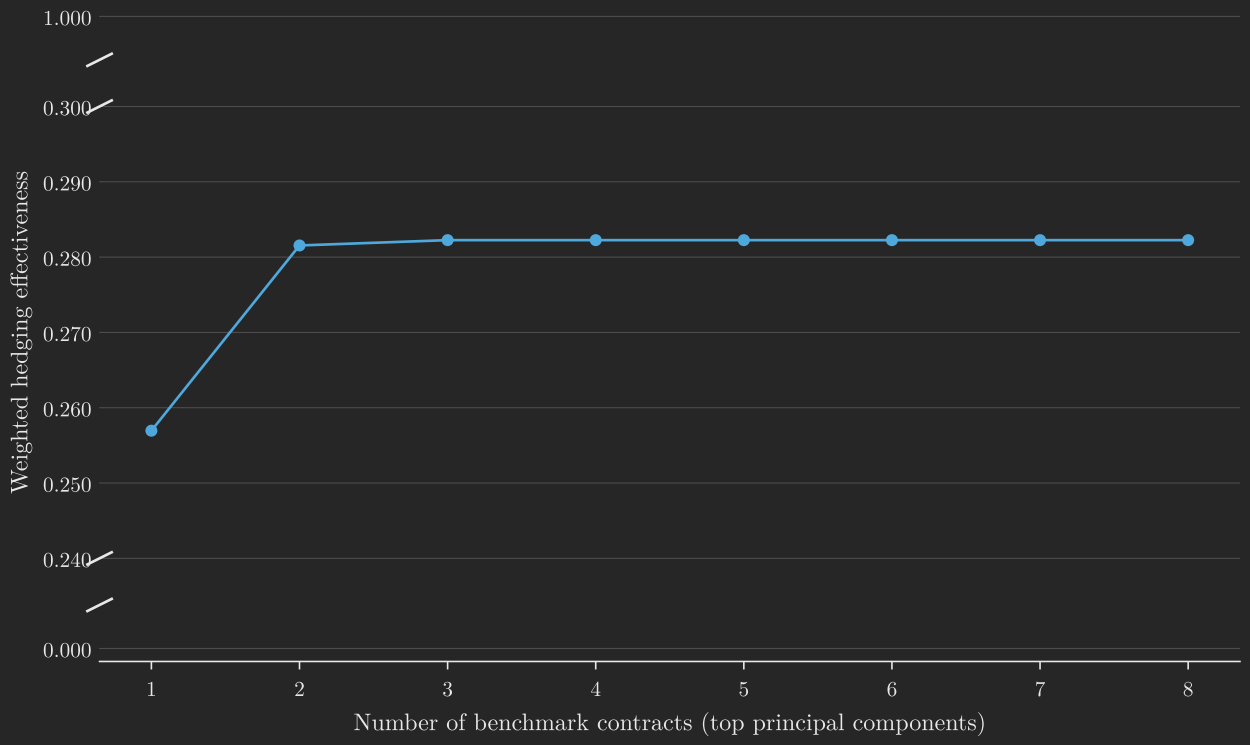
Figure IA.1 plots the relaxed spanning ceiling implied by the design comparison above. Its attainment is conditional on a feasible menu: hedge positions across several nonnegative, sum-to-one indices may be signed, which is why the feasible generation menu can span the factor space even though a standalone trailing eigenvector is not itself a settlement index. Reading the curve as a list of eigen-contracts would mistake the relaxed spectral calculation for a contract recipe.

Table IA.6 varies economic concentration, substitution, residual dependence, seller share, liquidity cost, exposure weights, and measurement error. Buyer HE, seller HE, and the relaxed menu count are scenario outputs. In particular, the count does not certify feasible eigen-contracts or an empirical launch threshold, and publisher measurement error can affect either clientele.

Figure IA.2 translates the sensitivity rows that move ϕ, γ, and θ into the implied contract count. More seller weight can support additional directions because locked exposures make the seller block richer, while changing θ mainly rotates the representative buyer direction and need not create trailing rank. Step changes are finite-grid crossings of the internal cost-benefit rule; they do not estimate where an actual market would add a contract.

Table IA.7 reports the baseline signed premium decomposition for the same scenario; it supplies

Figure IA.1: Relaxed systematic spanning ceiling



Notes. The curve uses the leading factor directions of the exposure-weighted model. In this DGP a feasible generation menu spans the factor space, but unrestricted eigenvectors are not generally settlement indices. The vertical axis is broken to magnify the relevant range.

Table IA.6: Sensitivity of the two-population scenarios

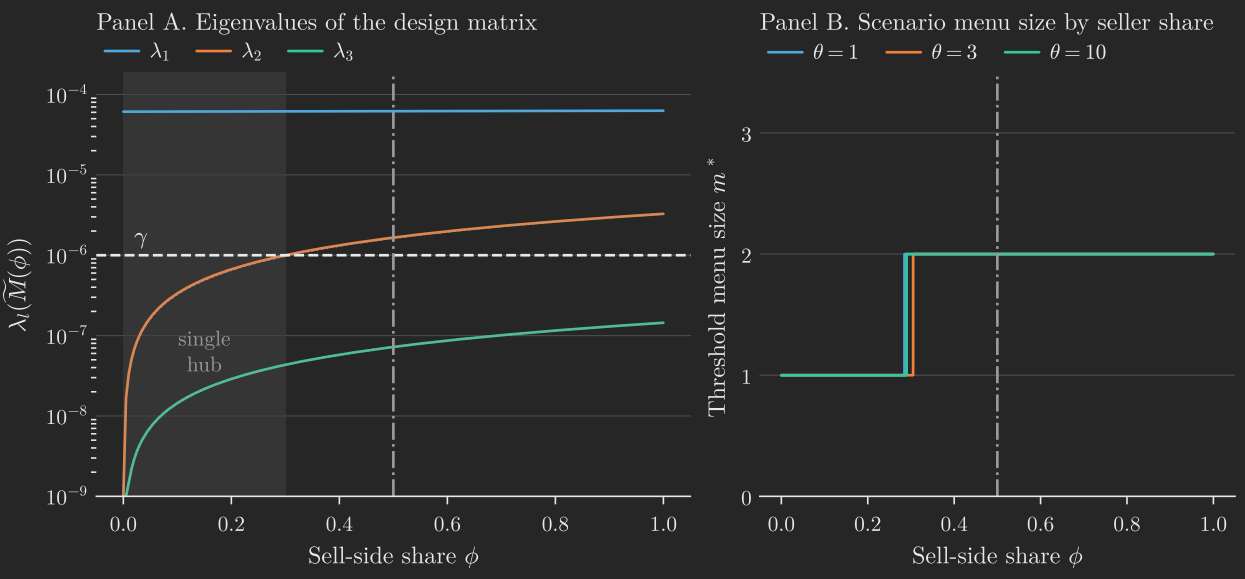
Input varied	Scenario values	Buyer HE	Seller HE	m^*
Buyer concentration / heterogeneity	$c = 0.5$–2; $h = 0$–0.8	0.98–1.00	0.28–0.28	2–2
Switching / substitution	$\theta = 1$–8	0.98–1.00	0.28–0.28	2–2
Economic residual correlation	$\rho_\varepsilon = 0$–0.5	1.00–1.00	0.28–0.63	2–2
Seller share	$\phi = 0.1$–0.9	1.00–1.00	0.28–0.28	1–2
Liquidity threshold	$\gamma = 2 \times 10^{-7}$–$5 \times 10^{-6}$	1.00–1.00	0.28–0.28	1–2
Seller exposure weights	equal / baseline / frontier / legacy	1.00–1.00	0.28–0.30	2–2
Measurement-error share	$q = 0$–0.75	0.94–1.00	0.28–0.54	2–2
Joint scenario envelope	all ranges jointly	0.90–1.00	0.28–0.73	1–2

Notes. Each row varies the listed input over the displayed scenario range; the joint envelope varies all inputs together. The reported m^* is the relaxed threshold count. It describes a feasible price-index menu only when the selected spaces are attainable under the settlement-weight constraints.

the baseline against which the ranking counterexample below is read.

Finally, Table IA.8 gives the counterexample behind the premium boundary in Proposition 8. Opposite-signed clientele loadings reinforce pressure when gross exposure, factor variance, contract loading, and risk tolerance are held fixed. Without that normalization, a large co-loaded factor can

Figure IA.2: Scenario eigenvalues and menu count



Notes. The left panel varies the assumed seller share ϕ against the assumed per-contract threshold γ; the right varies the assumed CES elasticity. Shading and crossings are outputs of this finite grid, not estimated market thresholds or a selected hub.

Table IA.7: Signed premium decomposition in the baseline scenario

	Systematic	Idiosyncratic	Premium
Broad hub	-0.037	0.100	0.063
A100 narrow	-0.034	-0.198	-0.232
H100 narrow	-0.036	1.481	1.445
H200 narrow	-0.037	-0.231	-0.268
B200 narrow	-0.041	-0.653	-0.694

Notes. Systematic and idiosyncratic columns are signed contributions to net hedging pressure in the assumed one-period CARA equilibrium. Their sum is the premium. Magnitudes and rankings depend on scenario loadings, clientele shares, measurement allocation, and speculative risk tolerance.

Table IA.8: Premium-pressure ranking counterexample

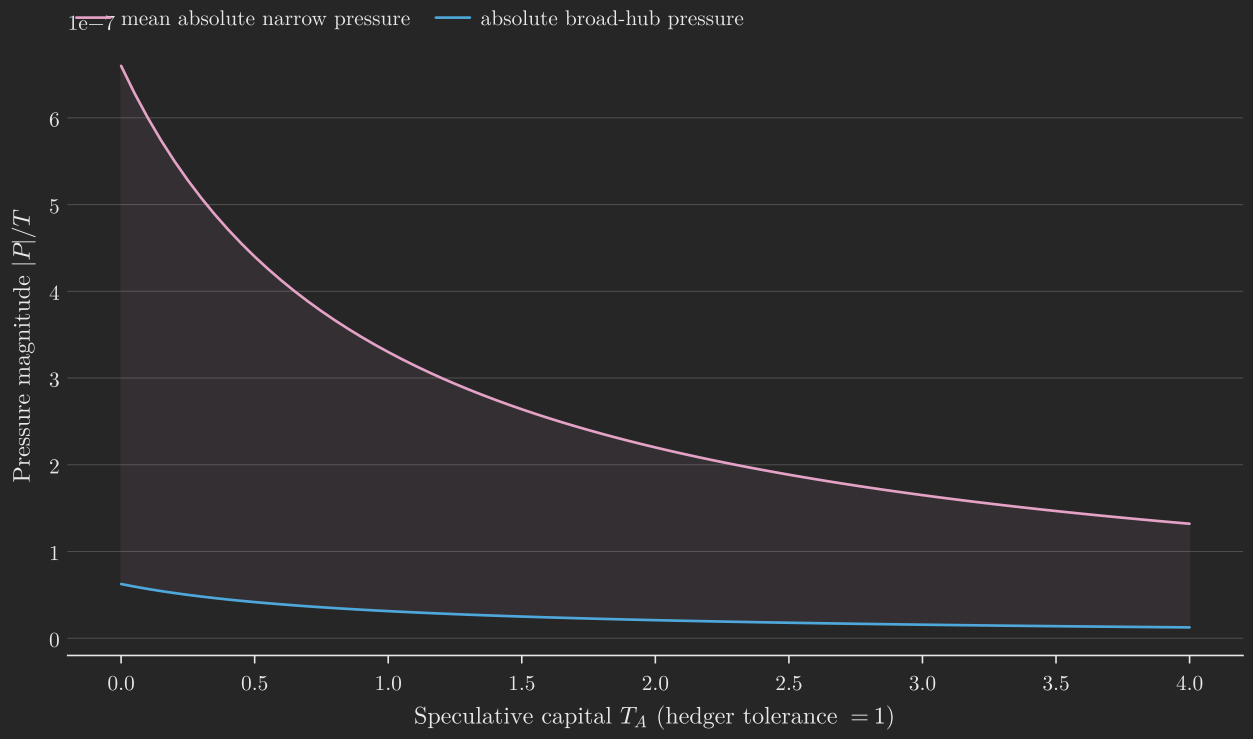
Scenario	Gross seller	Buyer term	|Net pressure|
Unrestricted obsolescence	0.050	0.050	0.100
Unrestricted large co-loading	1.000	0.250	0.750
Normalized sign-reinforced	0.500	0.250	0.750
Normalized co-loading	0.500	0.250	0.250

Notes. Entries are stylized pressure components. The unrestricted rows change gross exposure as well as sign structure; the normalized rows hold gross seller and buyer contributions fixed. The table illustrates why the sum-versus-difference result does not imply an unrestricted ranking.

have greater net pressure. The exercise therefore establishes no universal factor or contract ranking.

Figure IA.3 traces how the baseline decomposition of Table IA.7 scales with assumed speculative

Figure IA.3: Absolute premium magnitudes as assumed speculative risk-bearing capacity varies



Notes. The displayed broad and representative narrow curves are baseline scenario outputs, not estimated premia or universal rankings.

risk-bearing capacity: because the premium is net pressure divided by tolerance, magnitudes decline as capacity grows, while the broad-versus-narrow ordering persists only because the scenario's loadings and normalization are held fixed.

IA.IV.3. Parameters, Distribution, and Estimation Noise

Figure IA.4 shows the full cross-configuration distribution of hedging effectiveness behind the design averages of Table IA.5. Under any of the three single-index designs, effectiveness is low throughout, ranging from roughly 0.20 to 0.50 across configurations, with the upper tail belonging not to the frontier hardware but to the A100: the generation whose idiosyncratic share is smallest is the best hedged, exactly the inversion of prominence one expects when the idiosyncratic component, rather than loading mismatch, sets the level. The generation-segmented menu shifts the whole distribution

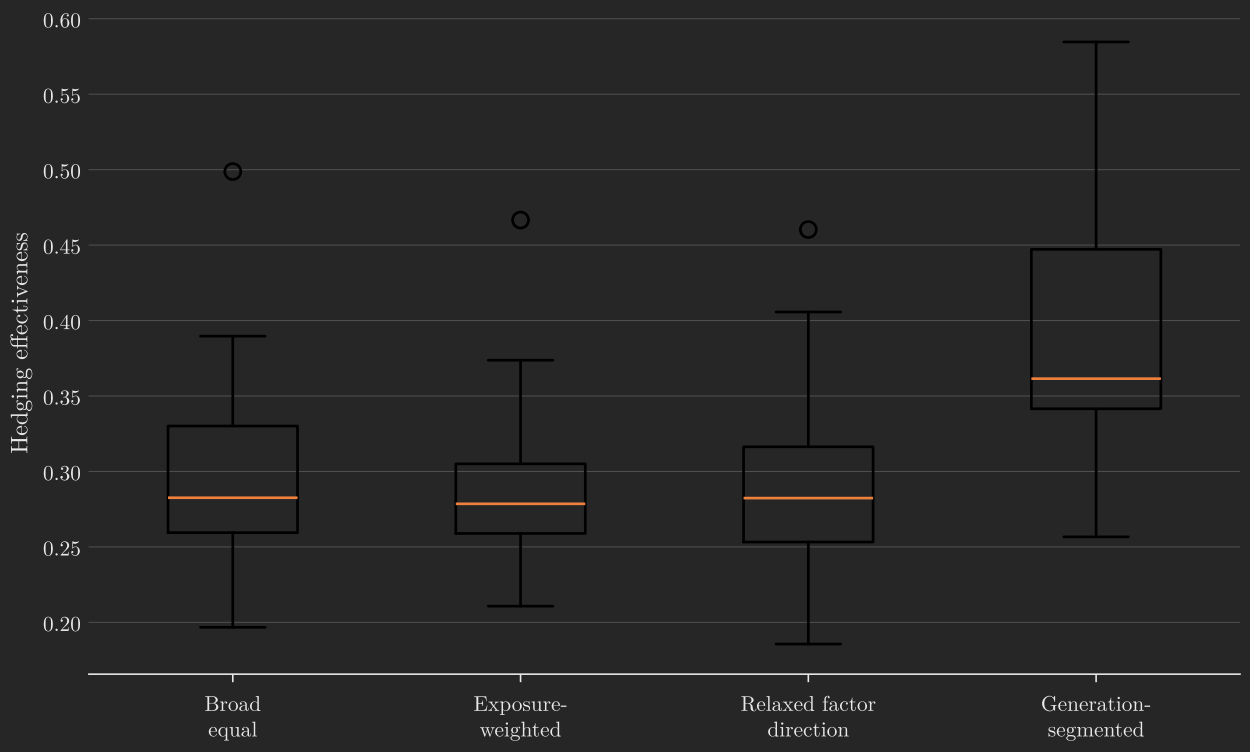
Table IA.9: Parameters of the simulated data-generating process

	A100	H100	H200	B200
Demand-factor loading	1.00	1.00	1.00	1.00
Scarcity-factor loading	0.20	0.60	0.80	1.00
Obsolescence-factor loading	1.00	0.50	0.20	0.00
Drift (% per year)	0.0	−11.1	−5.0	6.3
Idiosyncratic vol. (%, annualized)	11.1	21.4	22.2	23.8
Exposure-weight share	0.125	0.375	0.250	0.250

Factor volatilities (daily; annualized):
demand 0.0055 (8.7%), scarcity 0.0080 (12.7%),
obsolescence 0.0030 (4.8%).
Loading dispersion: 0.10 (s.d.).
$N = 32$ (4 generations × 8 variants); $T = 504$ days.

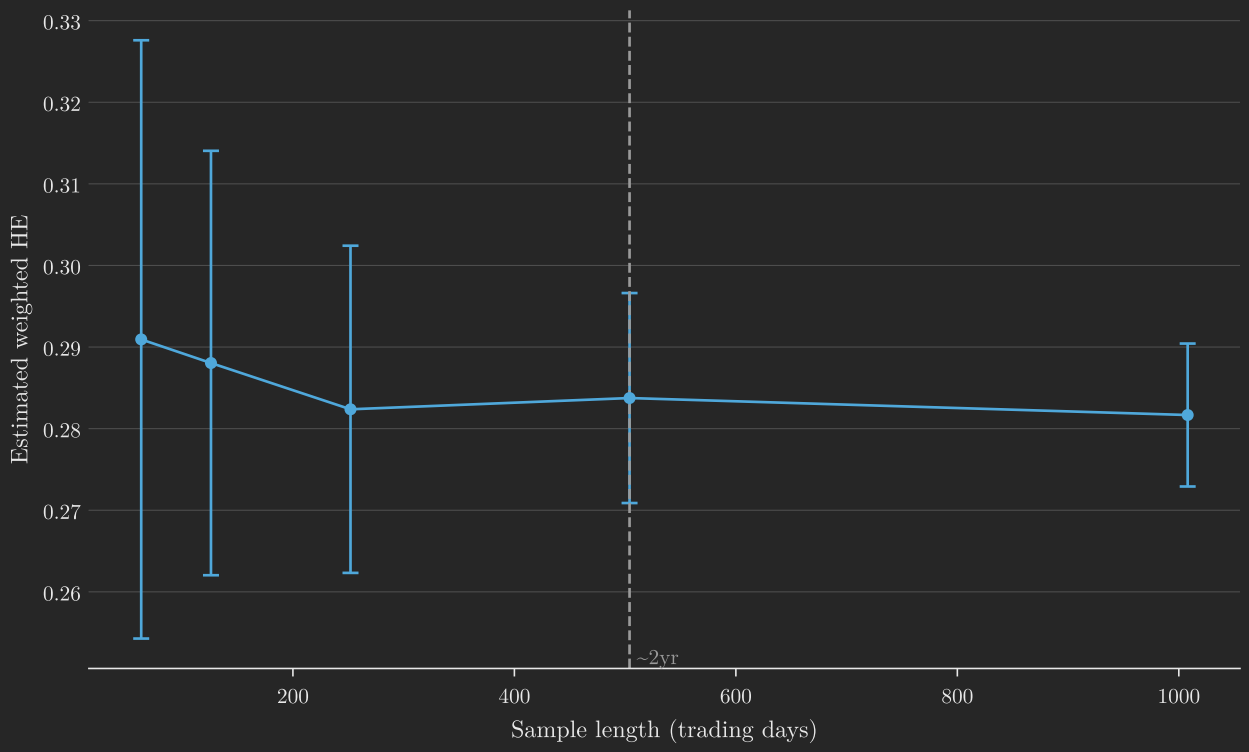
Notes. The table reports the parameters of the illustrative factor model of Section 4.1.1: the per-generation loadings on the demand, scarcity, and obsolescence factors; the deterministic drift $\alpha_i = \mu - \delta_i$ expressed in percent per year; the per-generation idiosyncratic volatility; the exposure-weight share η of each generation, spread equally across the eight variants within it; and the common factor volatilities. Each configuration draws its loadings from a normal distribution centered on its generation value with the stated dispersion. Drifts, volatilities, and the idiosyncratic shares are matched to selected published-index moments in Table IA.1 (Section 4.1.1); the loading pattern across generations is assumed. No H200 index is published, so its drift and residual volatility are interpolated, while its loading and exposure are assumed. Table IA.2 gives the full authority ledger.

Figure IA.4: Distribution of per-configuration hedging effectiveness under each benchmark design



Notes. Each box summarizes hedging effectiveness across the 32 simulated configurations for a given design; the line marks the median, the box the interquartile range, the whiskers the range excluding outliers, and the circles individual outliers.

Figure IA.5: Estimated weighted hedging effectiveness by sample length



Notes. Markers show the mean estimate across simulated samples of each length; error bars show one cross-sample standard deviation on either side. The dashed line marks the simulation's baseline horizon of 504 trading days (approximately two years), not the shorter realized published-index samples of Table IA.1.

upward rather than compressing its lower tail, consistent with the decomposition above: the lift is the own-weight component, which accrues to every constituent of a coarse index, not an alignment gain concentrated where the single-index basis is largest.

Only Weighted HE (pop.) in the design comparisons is the exact population value under the assumed DGP; mean, median, and weighted HE are single-regressor estimates on the simulated panel. In practice these parameters must be estimated, and the compute-price history is short. To gauge how much estimation noise a short sample injects into the hedging-effectiveness estimates, we repeatedly draw samples of varying length from the same data-generating process, estimate the weighted hedging effectiveness of the exposure-weighted index in each draw, and record the mean and the cross-sample standard deviation. Figure IA.5 reports the results.

Within this seeded data-generating process, the mean estimate is approximately flat and its

cross-sample dispersion falls with sample length. The standard deviation declines from 0.037 at 63 trading days to 0.013 at 504 days and 0.009 at 1,008 days. These values describe repeated samples from the assumed laboratory; they are not empirical confidence intervals or a global sampling bound. They show that the equal-versus-exposure-weighted scenario gap (0.282 versus 0.283) is small relative to laboratory sampling noise at the available horizon. The larger raw segmentation gap includes own-constituent mechanics and is not a pure loading-alignment comparison. The short published record can discipline selected moments, but it cannot select a settlement hub or resolve fine design rankings on its own.

IA.V. Spot-Price Dynamics: Additional Exhibits

Figures IA.6 and IA.7 display the sample second and first moments from Section 2.4. The volatility paths are descriptively different, but generation, publisher, unequal histories, and the short B200 record are confounded. The drift estimates are non-monotone in hardware age and individually imprecise.

Table IA.10 reports weekly summary statistics for the four rental indices and, for reference, the token index. None of the drift estimates is individually distinguishable from zero under Newey–West inference, and their ordering is not monotone in hardware age. Volatility, skewness, tails, and drawdowns describe this short publisher vintage rather than population moments or identified generation effects.

Figure 1, in the main text, plots each level normalized to 100 at its first observation.

To describe the cross-generation drift estimates, we pool the weekly log changes of the three single-provider Silicon Data rental series and estimate

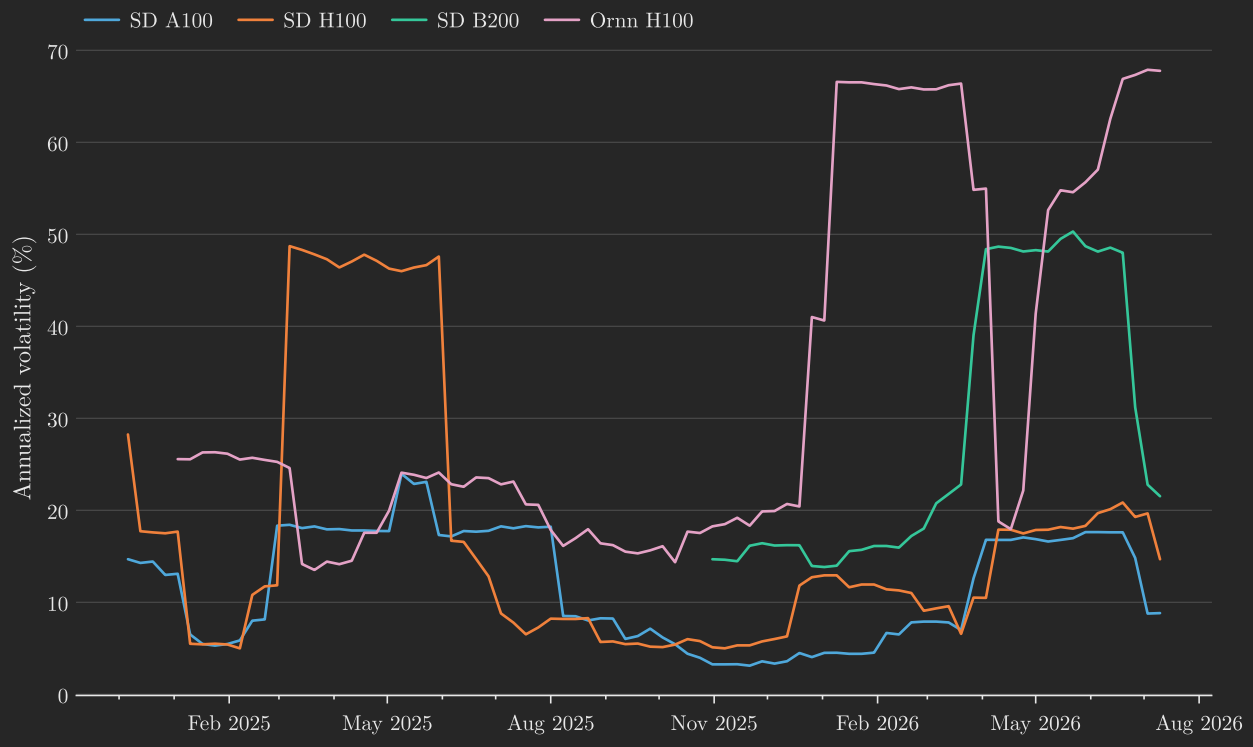
$$\Delta \log p_{g,t} \;=\; \sum_{g'} \alpha_{g'}\, \mathbf{1}[g = g'] \;+\; \varepsilon_{g,t}, \tag{IA.5}$$

Table IA.10: Summary statistics of compute price changes

Provider	A100 Silicon Data	H100 Silicon Data	B200 Silicon Data	H100 Ornn	LLM token Silicon Data
Drift (%/yr)	−0.3	−11.2	6.2	2.5	76.2
Volatility (%/yr)	13.6	24.3	29.2	39.2	49.9
Skewness	0.23	2.47	1.11	0.12	0.06
Excess kurtosis	5.61	19.96	6.96	8.14	0.03
AR(1)	0.03	0.20	−0.20	0.21	0.46
Max drawdown (%)	−26.8	−44.3	−19.7	−35.1	−29.2
Weeks	96	96	49	92	31

Notes. The table reports weekly Friday-close log-change statistics over each index's sample. Drift and volatility are annualized; AR(1) is weekly autocorrelation; maximum drawdown is the largest peak-to-trough level decline. The two H100 columns are different publishers of the same generation.

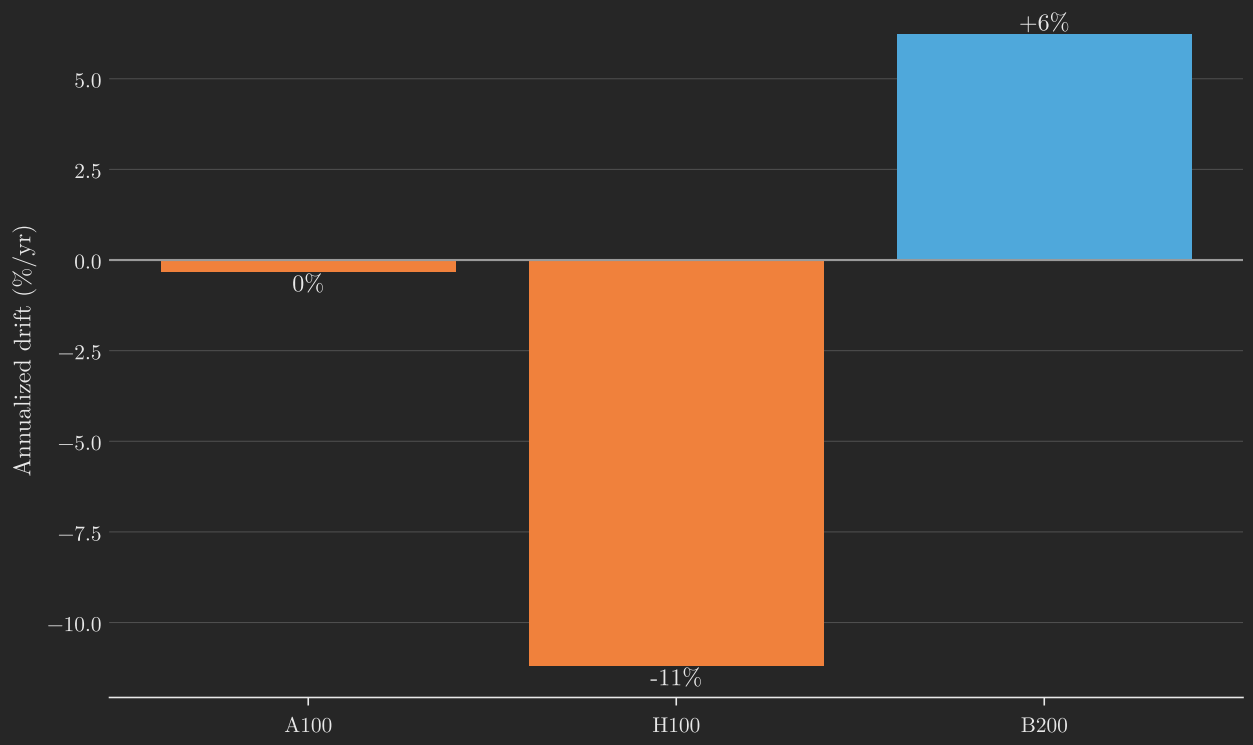
Figure IA.6: Rolling thirteen-week annualized volatility of weekly compute-price changes



so that each coefficient α_g is the mean weekly drift of generation g, with Newey–West (Newey and West, 1987) standard errors that accommodate serial correlation and heteroskedasticity. Equation (IA.5) shares notation with the factor-model drift $\alpha_i = \mu - \delta_i$, but the three cross-sectional means do not identify a depreciation schedule. Table IA.11 reports the estimates.

None of the three drifts is statistically distinguishable from zero under the reported Newey–West

Figure IA.7: Annualized sample drift by Silicon Data GPU generation



Notes. Error bars show Newey–West uncertainty; cash-cost-floor, obsolescence, and scarcity interpretations are hypotheses rather than identified mechanisms.

Table IA.11: Per-generation weekly drift of Silicon Data rental indices

	A100	H100	B200
Drift (%/yr)	−0.3	−11.2	6.2
(HAC t)	−0.03	−0.49	0.26
Volatility (%/yr)	13.6	24.3	29.2
Weeks	96	96	49

Notes. The table reports the per-generation mean weekly drift α_g of equation (IA.5), annualized, estimated by pooling the three single-provider Silicon Data rental series with generation dummies and no common intercept. Newey and West (1987) t-statistics (four lags) appear in the second row. *Volatility* is the annualized weekly volatility and *Weeks* the generation's observation count. The table is generated directly from the data file.

inference. Their point estimates are also non-monotone in hardware age: the H100, not the oldest A100, has the most negative estimate. Cash-cost-floor, obsolescence, and frontier-scarcity stories are possible interpretations of those signs, not findings. The simulation uses them as scenario assignments and does not treat the cross-generation comparison as identification of δ_i.

Table IA.12: Dependent-series uncertainty for correlations, HE, and PCA

Statistic	Estimate	95% block-bootstrap CI	Weeks
Panel A: balanced-sample correlations			
SD H100–SD A100	0.13	[-0.27, 0.46]	49
SD B200–SD A100	0.62	[0.05, 0.83]	49
SD B200–SD H100	0.37	[0.10, 0.58]	49
Ornn H100–SD A100	0.08	[-0.18, 0.24]	49
Ornn H100–SD H100	0.24	[0.07, 0.43]	49
Ornn H100–SD B200	0.07	[-0.14, 0.28]	49
Panel B: leave-one-out broad-basket HE			
SD A100	0.13	[0.00, 0.41]	49
SD H100	0.15	[0.03, 0.38]	49
SD B200	0.07	[0.00, 0.34]	49
Ornn H100	0.02	[0.00, 0.12]	49
Panel C: descriptive PCA shares			
PC1	45.9	[33.8, 56.9]	49
PC2	27.1	[23.5, 33.2]	49
PC3	19.1	[12.7, 24.4]	49
PC4	7.9	[3.0, 16.2]	49

Notes. Intervals are moving-block-bootstrap percentile intervals on the balanced 49-week sample. They do not form an interval for the unequal-overlap mean HE of 0.05, and no aggregate interval is manufactured from row-specific samples.

Table IA.13: Pre-specified out-of-sample design and current status

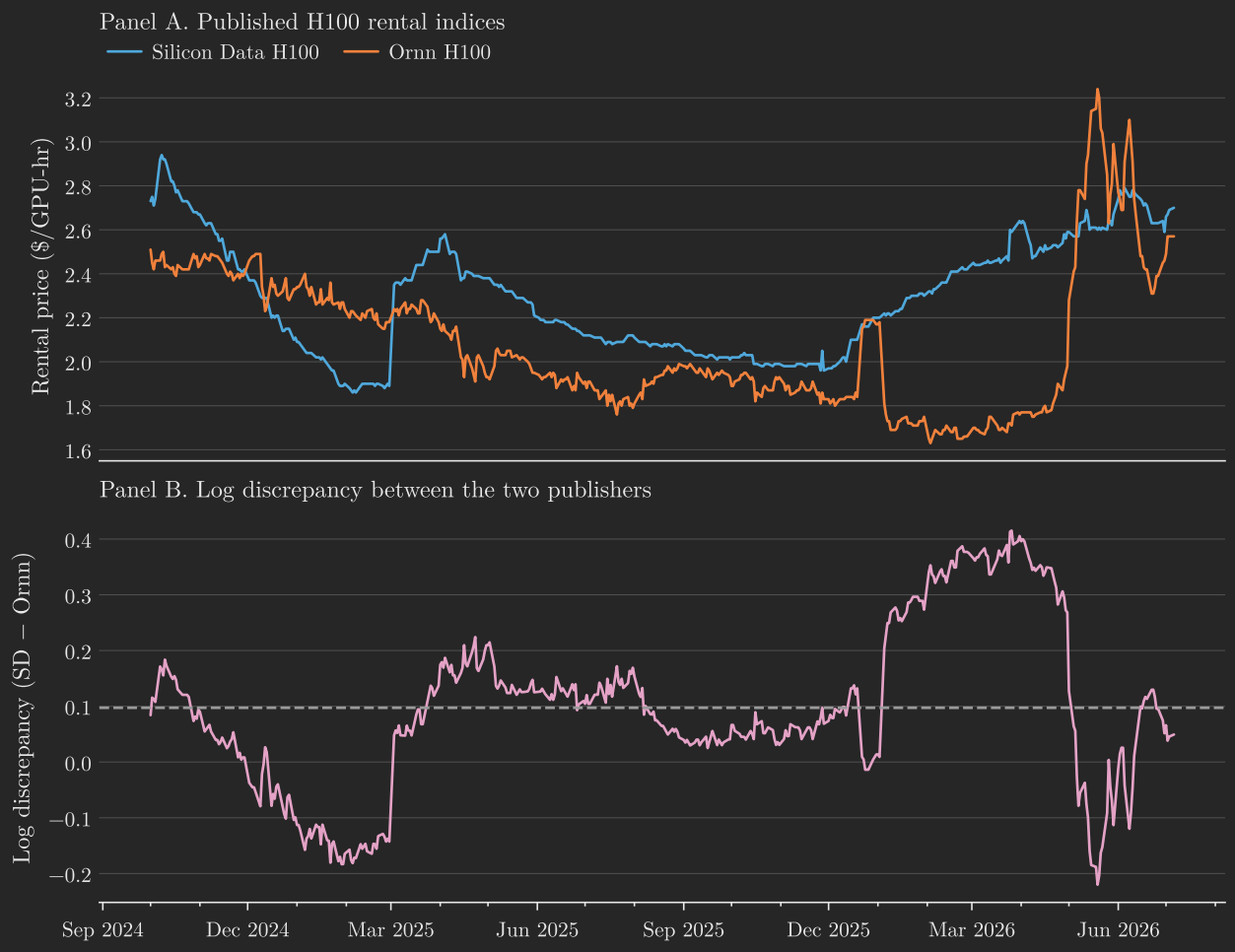
Design element	Pre-specified choice
Target	SD H100 published index change
Benchmark	Equal-weighted rental basket excluding target
Minimum training sample	104 weeks
Frozen quantities	Basket weights and training hedge ratio
Evaluation blocks	Nonoverlapping 26-week windows
Metrics	Out-of-sample HE and residual tracking error
Current status	Not run: insufficient sample (49/130 weeks)

IA.V.1. Dependent-Series Uncertainty and OOS Status

Table IA.12 uses moving-block resampling on the common 49-week panel. The intervals are row-specific dependent-data uncertainty summaries. Three of six balanced correlation intervals include zero, and the leave-one-out HE intervals are wide. The PC shares summarize this four-series panel; they are not a test of factor count or a feasible contract menu.

The pre-specified out-of-sample exercise in Table IA.13 has not been run. It requires 104 training weeks plus a 26-week evaluation block; only 49 common weeks are available. The table records an execution design, not a robustness result.

Figure IA.8: Two H100 rental indices and their log published-index discrepancy.



IA.V.2. Published-Index Discrepancy

Figure IA.8 plots the Silicon Data and Ornn H100 indices and their log discrepancy, $\hat{d}_t = \log \hat{p}_t^{\mathrm{SD}} - \log \hat{p}_t^{\mathrm{Ornn}}$. The discrepancy spans −0.22 to 0.42 log points; its sample mean is 0.10, its standard deviation is 0.15, and its weekly changes have 41.9% annualized sample volatility. Same-hardware cross-index HE is 0.01 daily, 0.03 weekly, and 0.04 monthly. A position defined by one publisher's index and hedged with the other would therefore have carried substantial tracking error, but this is not a firm-basis estimate because firm transaction prices are unobserved.

The cross-correlogram peaks contemporaneously and has no systematic one-sided decay, but venue coverage, observation type, composition, weighting, revision policy, staleness, and genuine

Table IA.14: Cross-provider H100 discrepancy and descriptive panel covariance

H100 cross-provider discrepancy, $\log(SD) - \log(Ornn)$	
Mean (log pts)	0.10
Std. dev. (log pts)	0.15
Range (log pts)	−0.22 to 0.42
Change vol. (%/yr, weekly)	41.9
Same-GPU HE (SD vs Ornn H100)	
Daily	0.01
Weekly	0.03
Monthly	0.04
Weekly correlation-matrix eigenvalue shares (4 rentals)	
$\lambda_1, \ldots, \lambda_4$ (%)	45, 26, 21, 8

Notes. The upper blocks report the published H100 log discrepancy and same-GPU cross-index HE. The lower block reports eigenvalue shares of the weekly pairwise-overlap correlation matrix. The balanced-sample PCA estimate and its uncertainty appear in Table IA.12. Both are descriptive covariance summaries. The table is generated directly from the data file.

segmentation remain observationally confounded. Thus only provider disagreement is observed here; neither contract tracking error against a hedger-relevant price nor firm economic basis is measured. Publisher error can affect either clientele and is not automatically diversified by buyer flexibility. Likewise, the leading component's 45.9% balanced-sample share, with block-bootstrap interval $[33.8\%, 56.9\%]$, neither identifies factor count nor selects a feasible menu.

IA.V.3. Descriptive Relative-Price Persistence

For each generation pair, define $q_{gg',t} = \log \hat{p}_{g,t} - \log \hat{p}_{g',t} - \log(a_g/a_{g'})$, where a_g is the fixed performance scalar of equation (IA.4). Subtracting this time-invariant scalar changes only the level and regression intercept; it leaves Δq, the error-correction slope λ, and the reported half-life unchanged. The exercise is therefore not a test of performance-adjusted substitution or a law of one effective price.

All four slopes are negative. Point half-lives are 11.7 weeks for H100/A100, 9.1 for B200/H100, and 3.5 for B200/A100, against 8.2 for the same-hardware publisher pair. Only B200/A100 is conventionally significant. No ADF test rejects a unit root, KPSS rejects stationarity for two pairs, and short-sample autoregressive bias tends to make reversion look too fast. Similar persistence in the

Table IA.15: Descriptive persistence of cross-generation relative prices

Pair	Weeks	Disp. sd(q)	λ	(NW t)	Half-life (wk)	ADF p	KPSS
Cross-generation relative prices (fixed level adjustment)							
H100/A100	97	0.137	−0.058	−1.02	11.7	0.17	0.68
B200/H100	50	0.112	−0.074	−1.59	9.1	0.51	0.83
B200/A100	50	0.058	−0.179	−2.74	3.5	0.18	0.32
Same hardware (no adjustment): provider comparison							
SD/Ornn H100	93	0.143	−0.081	−1.83	8.2	0.16	0.36

Notes. λ and its Newey–West t-statistic come from the error-correction regression; half-life is $\ln 2/[-\ln(1+\lambda)]$ weeks. ADF p is the augmented Dickey–Fuller unit-root-test p-value; KPSS tests level stationarity against a five-percent critical value of 0.463. Fixed performance adjustment changes only levels and the unreported intercept. The table is generated directly from the data file.

same-hardware comparison bounds any substitution reading. Estimating the CES elasticity, venue frictions, or buyer flexibility requires time-varying workload performance and a longer overlapping panel.

IA.V.4. Settlement-Window Tracking

Table IA.16 hedges a one-month change in each index's month-end value with point, last-week-average, or full-month-average settlement on that same index. Point HE equals one mechanically. Week-average HE ranges from 0.95 to 0.98 and month-average HE from 0.53 to 0.83 across the four indices. These are cross-index ranges, not confidence intervals. Averaging trades tracking against lag; the exercise contains no order-level evidence on manipulation resistance.

IA.V.5. Published-Index Cross-Tracking

Table IA.17 reports the cross-tracking diagnostics that the observable candidates of Section 4 support. Each entry is the equal-weighted mean, across the four rental indices, of the hedging effectiveness (equation (10)) of hedging each index with the listed candidate underlying. The broad rows use a leave-one-out basket that excludes the hedged index, so they are purely cross-index; the balanced row restricts every constituent to the common sample. The single-generation rows average

Table IA.16: Tracking a month-end exposure across settlement windows

Index / settlement window	Hedge ratio	HE	Residual vol. (%/yr)	Months
SD A100: Point (last day)	1.00	1.00	0.0	22
SD A100: Week average	0.98	0.95	3.2	22
SD A100: Month average	0.81	0.53	9.6	22
SD H100: Point (last day)	1.00	1.00	0.0	22
SD H100: Week average	1.02	0.98	4.3	22
SD H100: Month average	1.01	0.83	12.3	22
SD B200: Point (last day)	1.00	1.00	0.0	12
SD B200: Week average	0.77	0.97	4.7	12
SD B200: Month average	1.21	0.68	14.5	12
Ornn H100: Point (last day)	1.00	1.00	0.0	21
Ornn H100: Week average	1.07	0.98	5.6	21
Ornn H100: Month average	1.06	0.82	19.2	21

Notes. The hedge ratio minimizes residual variance for the stated month-end exposure. Residual volatility is annualized and Months gives the available monthly observations. The table measures tracking only; it does not estimate strategic manipulation, liquidity, or launch viability.

Table IA.17: Published-index cross-tracking diagnostics

Candidate published underlying	Mean cross-index HE	Contracts
Broad rental basket (leave-one-out)	0.05	1
Broad rental basket (LOO, balanced)	0.09	1
SD H100 single-generation	0.30	1
SD B200 single-generation	0.38	1
SD A100 single-generation	0.35	1
Ornn H100 single-generation	0.26	1
Token-settled (supplementary)	0.01	1

Notes. Hedging effectiveness follows equation (10). Each entry is the equal-weighted mean, across the four rental indices, of the effectiveness of hedging each index with the listed candidate underlying. The broad rows exclude the hedged series (leave-one-out), with the balanced row restricted to the common sample; the single-generation rows include their own index, contributing a mechanical unit-HE term, and are therefore not directly comparable to the leave-one-out broad rows. The token row uses a seven-month overlap. The Contracts column records that each candidate is a single contract. No exposure weighting is applied, because the available firm evidence does not identify configuration-level weights.

over the same four rental targets and therefore include the candidate's own index, whose mechanical unit HE inflates the reported mean and makes these rows not comparable to the leave-one-out broad rows. The token row uses a seven-month overlap. Because no exposure weighting is applied, the table cannot be read as the exposure-weighted objective of the theory or as a ranking of settlement products.

Table IA.18: Descriptive cross-index hedging effectiveness

Settled index	Broad incl.	Broad LOO	LOO bal.	Other prov.	Other gen.	Token	(token n)
SD A100	0.22	0.05	0.13	–	0.39	0.02	31
SD H100	0.45	0.05	0.15	0.03	0.14	0.00	31
SD B200	0.39	0.07	0.07	–	0.14	0.00	31
Ornn H100	0.53	0.02	0.02	0.03	–	0.00	31

Notes. Each cell is weekly $HE = \rho^2$. "Broad incl." includes the settled row index; "Broad LOO" excludes it on the available overlap; "LOO bal." uses the common 49-week sample. Other columns use the other publisher's same-GPU index, the nearest other Silicon Data generation, or the token index. The table is generated directly from the data file.

IA.VI. Cross-Index Hedging Effectiveness

We measure how one published index tracks another, not the firm economic basis of Proposition 1, for which neither realized firm prices nor expenditures are observed. For settlement index k and published target i, $HE_i^k = R^2 = \rho_{i,k}^2$ in a weekly hedge regression. Table IA.18 separates a broad average that mechanically includes its target from leave-one-out (LOO) baskets on each available overlap and on the common balanced sample.

The unequal-overlap LOO mean is 0.05, with rows from 0.02 to 0.07; the balanced mean is 0.09, with rows from 0.02 to 0.15. The inclusive range, 0.22–0.53, is higher partly because each target enters its own benchmark. Balanced block-bootstrap intervals are wide: the H100 LOO estimate is 0.15 with interval $[0.03, 0.38]$, while the A100 interval is $[0.00, 0.41]$. No aggregate interval exists for the unequal-overlap mean, and the pre-specified out-of-sample design cannot run with only 49 common weeks. Pairwise histories use more observations but change across targets; the balanced panel buys comparability by discarding history. The future design requires 104 training weeks, a frozen LOO basket and hedge ratio, and a nonoverlapping 26-week test block. These are therefore in-sample published-index diagnostics, not firm hedge estimates or benchmark selection.

The two H100 publishers have weekly cross-index $HE = 0.03$, and the token index has little sample co-movement with the rental series. The H100 gap may combine economic dispersion, coverage, staleness, measurement, and methodology; it identifies none of them. The observed gap is

provider disagreement. Settlement-benchmark risk instead compares a contract benchmark with a hedger-relevant price, while firm economic basis compares that hedger's realized price with the benchmark. Only provider disagreement is observed here, and publisher error could affect either clientele. Internet Appendix Section IA.V reports the discrepancy exhibits and the relative-price persistence exercise. Because the performance adjustment is a fixed scalar, persistence is descriptive rather than a substitution or law-of-one-price test. Only one of three cross-generation slopes is significant, no ADF test rejects a unit root, and the same-H100 placebo reverts at least as fast as two generation pairs. The evidence makes publisher and methodology explicit contract terms, but selects no empirical hub or menu.

IA.VII. Settlement Unit: GPU-Hours versus Tokens

A second, more tentative margin concerns the *unit* on which a contract settles. The Silicon Data token index prices inference output per token rather than hardware per GPU-hour, and over its seven-month overlap with the rental indices it behaves as a different asset: its weekly changes correlate with the Silicon Data H100 rental index at 0.03 (daily 0.05), and each hedges essentially none of the other's variance. The token index is also more volatile (50% annualized) and carried a strong drift over its short life. Token prices embed model efficiency and inference-market competition on top of hardware cost, so a $/token contract and a $/GPU-hour contract would serve different hedgers—inference buyers versus capacity lessors—and would have hedged almost none of each other's risk in this sample. We flag this as suggestive only: seven months is too short to bear weight, and the main results do not rest on it.

IA.VIII. Firm-Level Compute Betas: Additional Exhibits

IA.VIII.1. Exploratory Firm Exposure

The direct equity test is unfavorable to the model's side-specific prediction. In the joint weekly regression on a broad rental index and a residual H100 component, no group shows the positive and significant configuration beta predicted for suppliers. The supplier beta instead has the wrong sign, and the configuration component contributes little explanatory power. This is an informative negative result, not evidence that compute hedging demand is absent. Cha (2026) reaches a compatible conclusion from the other direction, finding on the same rental indices that no minimum-variance hedge built from equity proxies reduces variance out of sample.

The 28-firm sample uses weekly equal-weighted group returns, Dimson sums of contemporaneous and lagged coefficients, and four-lag Newey–West inference. The broad regressor averages the four rental series; the configuration regressor orthogonalizes the H100 change on that average. It is a residualized component, not a fleet-specific shock. A second panel first residualizes both indices on equity proxies.

For raw indices, the supplier configuration beta is -0.016 ($t = -2.22$, BH $q = 0.066$); after residualizing the indices on equity proxies, it is -0.013 ($t = -1.78$, $q = 0.188$). No group's beta is positive and significant, and the largest ΔR^2 is 0.012. The supplier 80% MDE is 0.020 in both panels, larger than either wrong-sign estimate in absolute value, so modest positive betas could go undetected. Low power limits claims about smaller effects; it does not turn negative coefficients into support for a positive prediction.

Equity returns are an indirect operating-exposure proxy. These regressions neither test unspanned price variance nor identify physical positions, configuration weights, natural-hedger status, or contract demand. Those objects require invoices, hardware holdings, and contract positions.

Table IA.19: Broad index versus residualized configuration component

Group	Broad β	Config β	BH q	80% MDE	ΔR^2	Weeks
Panel A: raw compute-price indices						
Users	0.003	−0.007	0.000	0.005	0.008	95
	(1.52)	(−4.22)				
Suppliers	−0.016	−0.016	0.066	0.020	0.012	95
	(−1.73)	(−2.22)				
Hyperscalers	−0.003	0.002	0.611	0.008	0.002	95
	(−1.19)	(0.69)				
Upstream (semis)	0.002	−0.001	0.754	0.008	0.009	95
	(0.59)	(−0.31)				
Infrastructure	−0.012	−0.003	0.487	0.008	0.005	95
	(−2.94)	(−1.05)				
Panel B: indices residualized on equity proxies						
Users	0.002	−0.008	0.000	0.004	0.010	
	(1.41)	(−4.95)				
Suppliers	−0.015	−0.013	0.188	0.020	0.007	
	(−1.66)	(−1.78)				
Hyperscalers	−0.003	0.002	0.599	0.008	0.002	
	(−1.27)	(0.76)				
Upstream (semis)	0.002	0.001	0.812	0.008	0.006	
	(0.53)	(0.24)				
Infrastructure	−0.012	−0.002	0.599	0.009	0.006	
	(−2.98)	(−0.71)				

Notes. Broad and configuration betas are Dimson sums from the joint weekly regression. "BH q" adjusts the configuration-loading test within each panel's five-group family. The 80% MDE is the two-sided five-percent minimum detectable configuration loading. ΔR^2 is relative to the broad-only regression. The table is generated directly from the data file.

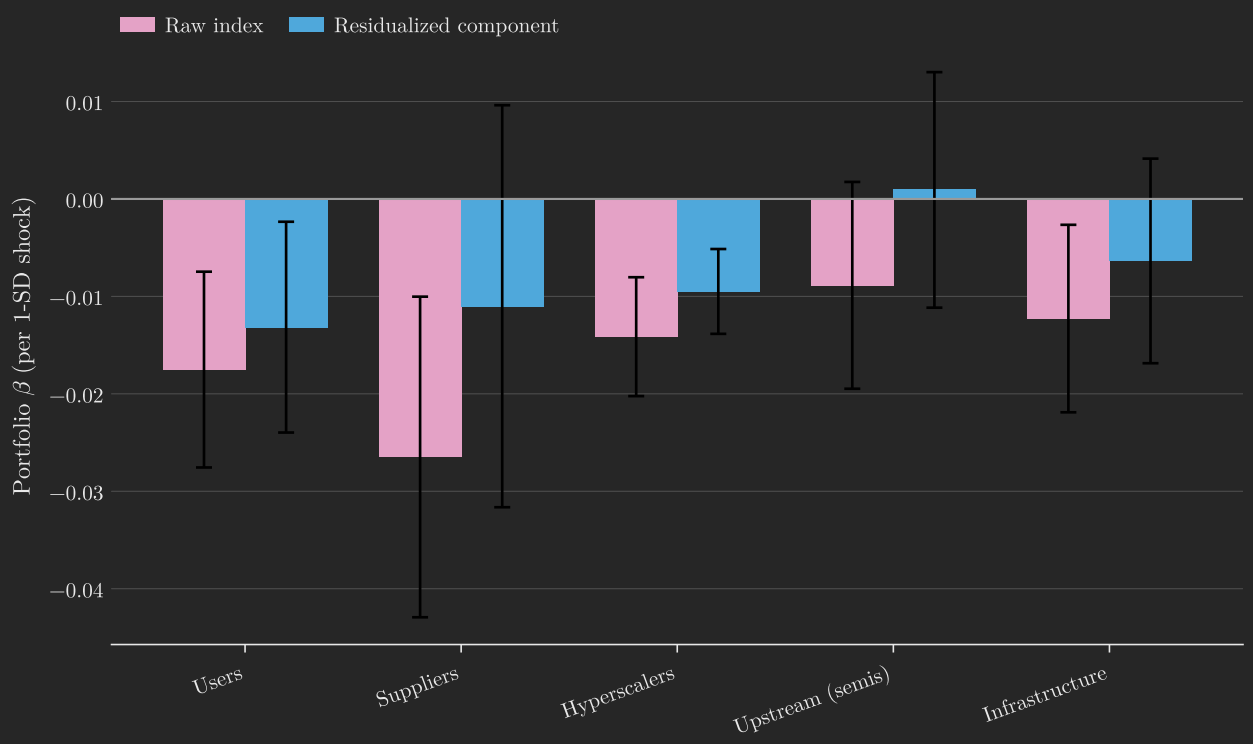
IA.VIII.2. Construction and Group Specifications

The weekly regressions above use Dimson sums of contemporaneous and one-lag coefficients with HAC inference. Coefficients are standardized to a one-standard-deviation change in the compute index. The second specification residualizes that index on contemporaneous equity proxies; this is a statistical projection, not an identified supply shock. The unbalanced panel and controls do not reveal configuration-level physical exposures.

Figure IA.9 plots the group estimates; Figure IA.10 shows the firm-level residualized-component estimates. Their dispersion and wide intervals counsel against reading any individual loading structurally.

Tables IA.20 and IA.21 report the full group specifications. Benjamini–Hochberg q-values are

Figure IA.9: Group compute betas for the raw index and the component residualized on equity proxies, with 95% HAC confidence intervals

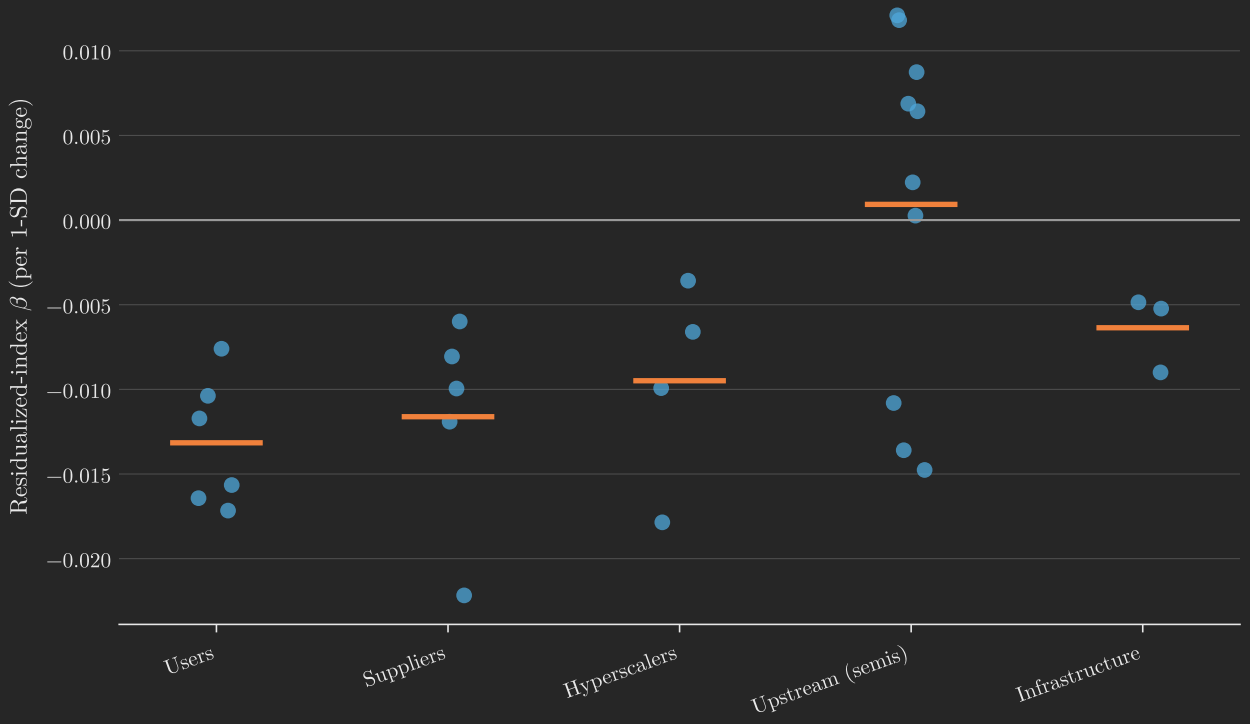


computed within the pre-specified families shown in each panel. The horse race leads with the unfavorable result: supplier configuration exposure is negative or null rather than the predicted positive hedge loading. Multiplicity and minimum detectable effects qualify precision; they do not turn that sign into favorable evidence or establish an absence of hedging demand.

The cross-sectional specification in Table IA.22 regresses firm betas on the group-prior exposure score. Its residualized-index coefficient is exploratory and benchmark-sensitive; it does not identify configuration weights for the design problem.

A future event design would require events classified independently of the equity response, uncontaminated windows, a pre-specified exposure mapping, and multiplicity-aware inference. The current short record offers too few such windows, so no event-study result is reported.

Figure IA.10: Distribution of firm-level residualized-component compute betas by group



Notes. Red bars mark group means.

IA.IX. Robustness Checks

Figure IA.11 is a descriptive covariance summary of four published series. It neither estimates factor count in the configuration universe nor selects a feasible contract menu.

This section varies sampling frequency, winsorization, and sample split for the published-index second moments. The variants are descriptive sensitivity checks. They do not remove unequal histories, publisher methodology, or the short-record limitation, and they do not validate a settlement product.

Table IA.23 reports, for each variant, four quantities that summarize the two headline findings: the annualized volatility of the H100 index, the mean cross-index correlation, the leading eigenvalue share of the correlation matrix, and the same-hardware Silicon Data–Ornn H100 hedging effectiveness.

Same-hardware cross-provider HE remains between 0.01 and 0.07 across these variants. The

Table IA.20: Group compute-price exposure specifications

Group	Univ.	$+X$	+FF3+UMD	+FF5+UMD	Weeks	Weeks (FF)
Panel A: raw compute-price index						
Users	−0.018	−0.004	−0.004	−0.006	95	89
	(−3.42)	(−1.91)	(−0.89)	(−2.14)		
[BH q]	0.006	0.111	0.442	0.073		
Suppliers	−0.026	−0.022	−0.010	−0.015	95	89
	(−3.15)	(−2.74)	(−1.30)	(−1.87)		
[BH q]	0.009	0.021	0.258	0.111		
Hyperscalers	−0.014	−0.000	−0.003	−0.002	95	89
	(−4.54)	(−0.11)	(−1.59)	(−0.90)		
[BH q]	0.000	0.936	0.171	0.442		
Upstream (semis)	−0.009	0.000	0.008	0.006	95	89
	(−1.64)	(0.08)	(2.84)	(2.43)		
[BH q]	0.169	0.936	0.018	0.038		
Infrastructure	−0.012	−0.010	−0.003	−0.005	95	89
	(−2.50)	(−3.11)	(−0.78)	(−1.31)		
[BH q]	0.035	0.009	0.483	0.258		
Panel B: index component residualized on equity proxies						
Users	−0.013	−0.004	−0.005	−0.007		
	(−2.38)	(−2.41)	(−1.03)	(−2.40)		
[BH q]	0.049	0.049	0.468	0.049		
Suppliers	−0.011	−0.018	−0.003	−0.009		
	(−1.05)	(−2.18)	(−0.44)	(−1.19)		
[BH q]	0.468	0.072	0.776	0.427		
Hyperscalers	−0.009	−0.000	−0.003	−0.001		
	(−4.27)	(−0.10)	(−1.58)	(−0.72)		
[BH q]	0.000	0.923	0.256	0.629		
Upstream (semis)	0.001	0.001	0.010	0.008		
	(0.15)	(0.51)	(3.30)	(2.89)		
[BH q]	0.923	0.760	0.010	0.020		
Infrastructure	−0.006	−0.010	−0.001	−0.003		
	(−1.19)	(−2.88)	(−0.29)	(−0.95)		
[BH q]	0.427	0.020	0.862	0.492		

Notes. Columns add the market, cross-exposure, and factor controls named in the main text. Parentheses contain HAC t-statistics; bracketed rows contain BH q-values within each panel's displayed test family. Panel B uses the index component residualized on equity proxies and carries no structural shock label.

leading PCA share is a minority of panel variance at weekly frequency (45.4%) and a majority at monthly (50.9%), while daily staleness lowers comovement further. H100 volatility also changes across splits and after winsorization. These patterns show that the point estimates are sample-sensitive even when the qualitative publisher discrepancy remains. They do not establish a stable market dimension or a maturing price process.

Re-estimating with the fixed performance scalar leaves every second moment unchanged by construction; Section IA.III records the future time-varying-performance design needed for a genuine substitution test. The equity-proxy residualization is already reported as a statistical control

Table IA.21: Broad-versus-configuration exposure horse race

Group	Broad β	Config β	BH q	80% MDE	ΔR^2	Weeks
Panel A: raw compute-price indices						
Users	0.003	−0.007	0.000	0.005	0.008	95
	(1.52)	(−4.22)				
Suppliers	−0.016	−0.016	0.066	0.020	0.012	95
	(−1.73)	(−2.22)				
Hyperscalers	−0.003	0.002	0.611	0.008	0.002	95
	(−1.19)	(0.69)				
Upstream (semis)	0.002	−0.001	0.754	0.008	0.009	95
	(0.59)	(−0.31)				
Infrastructure	−0.012	−0.003	0.487	0.008	0.005	95
	(−2.94)	(−1.05)				
Panel B: indices residualized on equity proxies						
Users	0.002	−0.008	0.000	0.004	0.010	
	(1.41)	(−4.95)				
Suppliers	−0.015	−0.013	0.188	0.020	0.007	
	(−1.66)	(−1.78)				
Hyperscalers	−0.003	0.002	0.599	0.008	0.002	
	(−1.27)	(0.76)				
Upstream (semis)	0.002	0.001	0.812	0.008	0.006	
	(0.53)	(0.24)				
Infrastructure	−0.012	−0.002	0.599	0.009	0.006	
	(−2.98)	(−0.71)				

Notes. The 80% MDE is the coefficient magnitude detectable at 5% size under the fitted design. It is a power diagnostic, not an estimated lower bound. BH q-values adjust the configuration-coefficient family within each panel. The supplier coefficient is wrong-signed or null in both panels.

Table IA.22: Exploratory cross-section of firm compute betas

	(1) Raw index	(2) Residualized comp.
Predicted exposure	0.001	0.005
	(0.40)	(3.24)
Intercept	−0.016	−0.009
	(−9.93)	(−6.91)
R^2	0.003	0.187
Firms	28	28

Notes. Parentheses contain cross-sectional t-statistics. The exposure score is a group prior rather than a measured configuration position, and the result is fragile to the compute benchmark.

in Section IA.VIII, not as supply-shock identification. The future event design likewise remains unexecuted.

Figure IA.11: Cumulative share of weekly rental-panel variance captured by the principal components of the four-series balanced panel

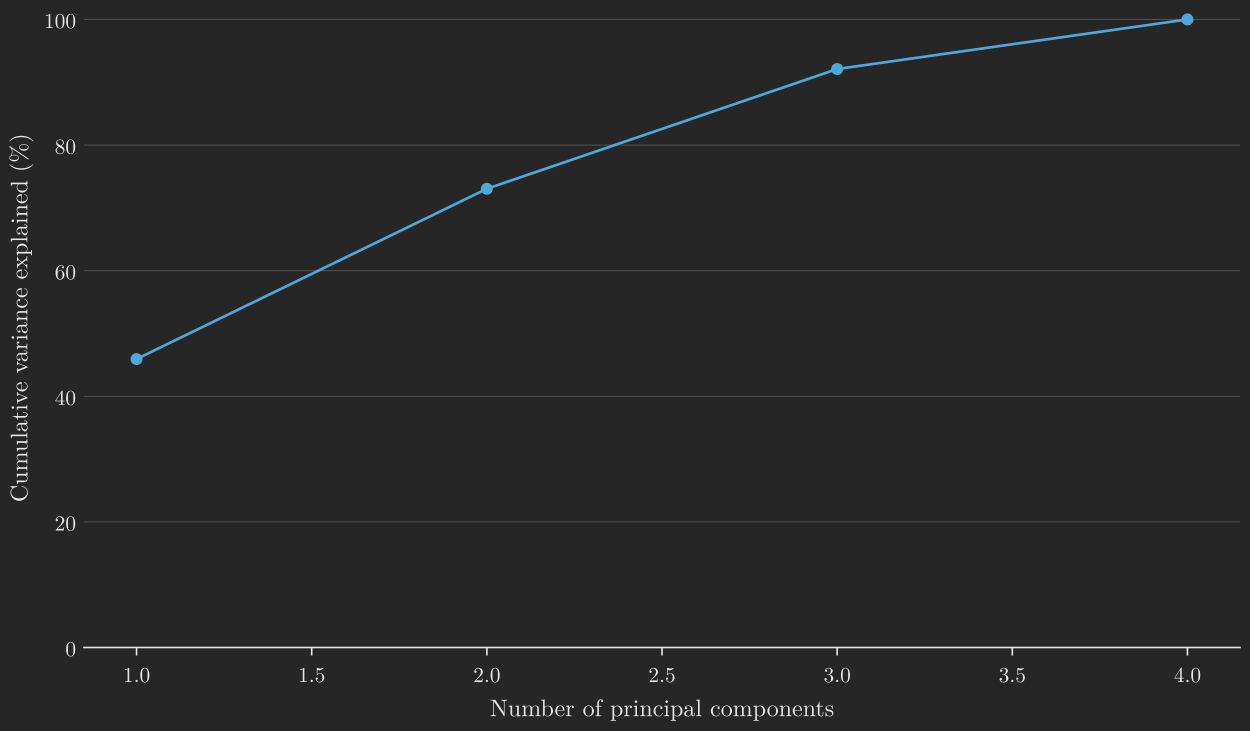


Table IA.23: Robustness of the headline second-moment estimates

Specification	H100 vol.	Mean corr.	λ_1	SD–Ornn HE
Weekly (baseline)	24.3	0.24	45.4	0.03
Daily	24.7	0.04	28.6	0.01
Monthly	30.1	0.33	50.9	0.04
Winsorized (weekly, 5/95)	14.3	0.20	41.0	0.04
First half (weekly)	30.7	0.09	39.5	0.02
Second half (weekly)	13.9	0.26	46.0	0.07

Notes. *H100 vol.* is the annualized volatility of the Silicon Data H100 index; *Mean corr.* the mean off-diagonal correlation of the rental panel; λ_1 the leading eigenvalue share (percent) of the correlation matrix; and *SD–Ornn HE* the same-hardware cross-provider hedging effectiveness. Rows vary the sampling frequency, winsorize weekly changes at the 5th/95th percentiles, and split each series at its midpoint. The table is generated directly from the data file.