Vintaged Capital Ledger:

A Measure-Valued Reconstruction of Installed Capital under Partial Disclosure

Avaneendra Trivedi

Independent Researcher

avaneendra22@gmail.com

May 2026

Abstract

We introduce the Vintaged Capital Ledger (VCL), a measure-valued reconstruction of outstanding capital in financial markets, indexed by entry basis, age, carry, cohort, and sign. The ledger evolves under a birth, ageing, and death process that is observable exactly in unspent-transaction-output assets and partially observable in disclosure-bound assets. We construct a Bitcoin short-term-holder oracle on the full AWS Public Blockchain dataset (189,315,387 transaction outputs across 310 daily partitions) and impose a uniform monthly observation grid that mimics the SEC NPORT-P disclosure cadence. Under this monthly fidelity, the August 2015 cutoff yields an identified-set width of 0.375 (CONDITIONAL gate), while the February 2018 cutoff yields a width of 0.672 (FAILED gate). A pre-registered sensitivity sweep across observation cadences of 15, 30, 45, 60, and 90 days establishes the principal empirical claim of the paper: monthly NPORT-P disclosure fidelity is insufficient to identify the short-term-holder URPD in the high-volatility 2017 bull regime, and quarterly-or-coarser cadence (60 days and above) clears the CONDITIONAL gate on both cutoffs. We report the canonical 30-day gate failure honestly: the not-yet-thesis-grade label remains on the artifact. The result is a structural finding about the disclosure-fidelity boundary of vintaged identification in equity and fixed-income markets, established on a clean Bitcoin test bed where the truth is observable. Mass conservation holds at machine epsilon; coverage probability is 1.0 on both cutoffs. All 133 unit tests pass, decision thresholds are SHA-256 locked in the pre-registration registry, and all artifacts are reproducible from the canonical run identifier.

JEL Classification: G12, G14, G17, G23, C14

Keywords: vintaged capital, measure-valued ledger, partial identification, disclosure fidelity, URPD, short-term holders, capital gains overhang, NPORT-P, identified set, Bitcoin UTXO, oracle masking, pre-registration

Pre-registration: *Decision thresholds (coverage ≥ 0.95, CLEAR width ≤ 0.30, CONDITIONAL width ≤ 0.60) are SHA-256 locked in registry/policies/ before any empirical analysis. Sensitivity sweep over observation cadences is registered as a diagnostic curve, not a post-hoc parameter search.*

1. Introduction

Standard quantitative finance treats markets as price processes. A market is represented by its current price, its recent return history, and a covariance matrix summarizing co-movement. Risk metrics are functions of these objects. Trading decisions are made against them. This representation is mathematically convenient and empirically tractable, but it discards a piece of structure that is observable in principle and material in practice: the population of still-outstanding capital that the price is acting on.

Every share bought, every bond issued, every futures position opened, every Bitcoin unspent transaction output (UTXO) created represents a unit of installed capital that persists until liquidation. Two markets with identical prices, identical volatility, and identical factor loadings can sit on completely different installed-capital populations. One market may have most of its outstanding mass near breakeven; another may have most of it underwater. These are not the same market. They look identical only inside a framework that has discarded the memory of how the current state was reached.

This paper introduces the *Vintaged Capital Ledger* (VCL), a measure-valued reconstruction of the outstanding-capital population, indexed by entry basis, age, cumulative carry, cohort, side, and native risk unit. The ledger is the primary state object. Prices, returns, and volatilities are downstream functionals of the interaction between current price and the installed population. The ledger evolves under a birth, ageing, and death process: $\mu_{t+e} = \mathrm{Age}(\mu_t) + \mathrm{Birth}_{[t,t+e]} - \mathrm{Death}_{[t,t+e]}$. For UTXO-based assets such as Bitcoin, the evolution is observable exactly: every output carries its creation timestamp and value, and every spend is publicly recorded. For traditional assets under disclosure-bound observability (13F snapshots, NPORT-P monthly filings), the ledger is partially identified, and the system outputs an identified set across pre-registered admissible liquidation policies rather than a false point estimate.

The principal contribution of this paper is empirical. We construct a Bitcoin short-term-holder (STH) oracle that processes the full AWS Public Blockchain dataset, 189,315,387 transaction outputs across 310 daily partitions, and apply a uniform monthly observation grid that mirrors the SEC NPORT-P disclosure cadence. The resulting identified-set width is 0.375 in the low-volatility August 2015 regime (CONDITIONAL) and 0.672 in the high-volatility February 2018 regime (FAILED). A pre-registered sensitivity sweep across 15-day, 30-day, 45-day, 60-day, and 90-day observation cadences establishes the structural finding: monthly NPORT-P fidelity is insufficient for the high-volatility 2017 regime, and quarterly-or-coarser cadence clears the CONDITIONAL gate on both cutoffs. This is a disclosure-fidelity result, not a methodological failure. The Bitcoin test bed shows the boundary at which vintaged identification becomes empirically useful in markets where the truth is recoverable.

1.1 Relation to prior literature

The behavioral finance literature established that reference points matter. Shefrin and Statman (1985) introduced the disposition effect, the tendency of investors to sell winners too early and hold losers too long. Odean (1998) documented the effect using individual brokerage account data. Frazzini (2006) extended the analysis to mutual fund holdings and showed that the disposition effect generates post-announcement drift. Grinblatt and Han (2005) integrated prospect theory and mental accounting into an aggregate model of momentum. These contributions collapse the embedded-cost-basis distribution into a scalar reference price or capital-gains overhang. VCL is measure-valued where these are scalar: the same average reference price can correspond to vastly different installed populations, with materially different implications for stress response.

A second strand examines the price impact of ownership structure and forced trading. Coval and Stafford (2007) showed that mutual fund fire sales create predictable price pressure in overlapping holdings. Greenwood and Thesmar (2011) defined stock-price fragility as a function of owner concentration and correlated liquidity needs. Hendershott and Menkveld (2014) measured intermediary price pressure with a half-life of 0.92 days. Koijen and Yogo (2019) constructed a demand-system approach to asset pricing based on institutional and household holdings, and Gabaix and Koijen (2021) developed the inelastic markets hypothesis, in which flows have large and persistent price impact under a small aggregate demand elasticity. These literatures take holdings and flows seriously as primary objects. VCL adds the missing dimension: the embedded P&L state of the current owners. Fragility tells you who is exposed; vintaged ledger tells you where their pain thresholds lie.

A third strand operates in the cryptocurrency domain, where last-moved-price distributions are directly observable. Liu, Zhang, and Zhao (2022) developed the cohort-analysis methodology for Bitcoin UTXO data that underlies industry URPD metrics. These on-chain analytics are the closest cousin of VCL: they instantiate an exact realized-price distribution in a single bearer-asset class. VCL is the natural generalization. The instrument treats Bitcoin as the calibration domain where the truth is observable and identifies the conditions under which the same measure-valued object remains decision-useful when projected onto the partial-disclosure regimes of equities and fixed income.

Partial identification provides the econometric scaffolding. Manski (2003) and Imbens and Manski (2004) developed the inferential framework for parameters constrained to interval-identified sets. VCL is an applied partial-identification object: where disclosure is coarse, the ledger emits an identified set across admissible liquidation policies, and the width of this set is the natural diagnostic of decision usability.

This paper is a companion to Trivedi (2026a), the Reflexivity Kernel Spectroscopy framework, and Trivedi (2026b, 2026c), the Capital Trigger Topology and Executable World Resolution papers respectively. Where the companion papers measure how markets respond to flows (RKS), where mechanical trigger surfaces compel action (CTT), and how much capital is required to make two futures distinguishable (EWR), VCL measures the state object that those processes act on: the still-outstanding capital population. The four instruments answer four logically sequential questions about the same underlying market.

1.2 Structure of the paper

Section 2 develops the mathematical theory of the measure-valued ledger and the partial-identification framework. Section 3 describes the estimation procedure, including the STH-scoped Bitcoin oracle, the pre-registered liquidation policy family, and the identified-set width definition. Section 4 describes the data: the AWS Public Blockchain Bitcoin transactions dataset and the two pre-registered cutoffs. Sections 5 through 8 present the empirical results: STH survivor URPDs, per-policy mode dispersion, the disclosure-fidelity sensitivity sweep, and the regime-level market structure comparison. Section 9 presents the falsification and integrity status. Section 10 discusses limitations. Section 11 concludes.

2. Theory

2.1 The measure-valued ledger

Fix an instrument i and a time t. Define the state space $\Omega = B \times A \times C \times K \times S$, where $B \subseteq \mathbb{R}_+$ indexes entry basis, $A \subseteq \mathbb{R}_+$ indexes age since acquisition, $C \subseteq \mathbb{R}$ indexes cumulative carry or financing burden, K indexes cohort label, and $S = \{+1, -1\}$ indexes side. The *vintaged capital ledger* of instrument i at time t is a finite signed measure $\mu_{i,t}$ defined on Ω and valued in a native risk unit (shares, contracts, delta-equivalent shares, DV01, or coin units). For a measurable set $E \subseteq \Omega$, the quantity $\mu_{i,t}(E)$ measures the mass of still-outstanding capital with coordinates in E.

The choice of a measure-valued state object rather than a scalar reference price reflects a structural commitment: two markets can produce identical scalar summaries while sitting on radically different ledgers. A market with mass concentrated at a single entry basis behaves differently under stress than a market with the same average basis spread evenly across many cohorts. The measure-valued representation preserves this distinction.

2.2 Evolution equation

The ledger evolves under three operations: ageing, birth, and death. The discrete-time evolution equation is:

$$\mu_{t+e} = \mathrm{Age}(\mu_t) + \mathrm{Birth}[_{t,t+e}] - \mathrm{Death}[_{t,t+e}] \tag{1}$$

where the ageing operator Age advances the age coordinate a by dt and accrues carry c by the period carry charge for the affected cohorts, $\mathrm{Birth}[_{t,t+e}]$ records new mass entering the ledger over the period from purchases, issuance, open-interest increases, or disclosed additions, and $\mathrm{Death}[_{t,t+e}]$ records mass leaving the ledger from sales, redemptions, expiries, open-interest decreases, or disclosed reductions. Mass conservation is the identity $\mu_t(\Omega) + Birth(\Omega) - Death(\Omega) = \mu_{t+e}(\Omega)$, and is enforced as an integrity gate on every transition; the canonical implementation discussed in Section 3 maintains this residual at machine epsilon ($\sim 1\times10^{-14}$).

2.3 Identification under partial observability

In UTXO-based assets such as Bitcoin, every birth and every death is publicly observable: the creation transaction carries timestamp and value, and the spend transaction identifies the consumed inputs. The ledger is point-identified. In disclosure-bound assets, deaths between consecutive disclosure dates are not individually observable. A reporting institution may liquidate part of a holding between two snapshot dates, and the disclosure shows only the net change in position; the exact lots consumed are not revealed.

The system therefore outputs a set of admissible ledgers rather than a single point estimate. Let Π denote a pre-registered family of liquidation policies that map a net death amount to a specific set of consumed lots. Examples include FIFO (oldest lots first), LIFO (newest lots first), MINIMUM_LOSS (lots with smallest unrealized loss first), MINIMUM_TAX (lots with largest unrealized loss first), PRO_RATA (proportional consumption across all lots), RANDOM_SEEDED (deterministic pseudo-random consumption with fixed seed), and VOLUME_PROPORTIONAL (weighted by realized trading volume). For each policy $\pi \in \Pi$, the reconstruction yields a ledger $\mu_t^{(n)}$. The *identified set* at time t is

$$\mathcal{J}_t = \{ \mu_t^{(n)} : \pi \in \Pi \} \quad (2)$$

the collection of all ledgers consistent with the public record under some pre-registered admissible policy. This is the partial-identification approach of Manski (2003) and Imbens and Manski (2004), specialized to vintaged ledgers. When $\mathcal{J}_t$ is a singleton (as in Bitcoin, where transfer history is complete), the ledger is point-identified. When $\mathcal{J}_t$ is wide, decisions that depend on the exact ledger should be deferred; the system labels this state as UNIDENTIFIED.

2.4 Decision functionals and the identified-set width

A decision functional F maps a ledger to a real number relevant to a downstream decision. Examples include the URPD mode (the entry price at which the unrealized P&L distribution peaks), the near-breakeven mass fraction (the share of mass within ±5% of the current price), and the underwater-after-carry fraction. For any policy π, the functional yields a point value $F(\mu_t^{(n)})$. Across the identified set, F yields the range

$$\Delta F_t = \max\pi \, F(\mu_t^{(n)}) - \min\pi \, F(\mu_t^{(n)}) \quad (3)$$

which we call the *policy spread* of *F*. To make the policy spread comparable across regimes with different price scales, we normalize by the range of entry prices in the short-term-holder window. Define the *identified-set width* for the URPD-mode functional as

$$w_t = \Delta \mathrm{Mode}(\mu_t) \, / \, (\text{max-STH-price} - \text{min-STH-price}) \quad (4)$$

This quantity is dimensionless and bounded in [0, 1]. A width of 0 indicates perfect agreement across all admissible policies, the case of point identification. A width of 1 indicates that the URPD mode spans the entire short-term-holder price range, the case of full non-identification. Practical decision usability lies in the interval [0, 0.60], with sharper labels at 0.30. These thresholds are pre-registered, SHA-256 locked, and not moved in response to empirical outcomes (see Section 3.4).

3. Estimation

3.1 The Bitcoin short-term-holder oracle

Bitcoin is the calibration domain for VCL because the full transfer history is publicly recorded on the blockchain. Every unspent transaction output (UTXO) carries its creation height, creation value in BTC, and the daily USD price at creation. Every spend is publicly recorded and identifies which UTXOs were consumed. The full age-and-basis distribution of outstanding coins is therefore observable in principle.

To make the test bed informative for traditional markets under partial disclosure, we restrict attention to the *short-term holder* (STH) population, defined as UTXOs of age less than or equal to 155 days at the evaluation cutoff. This threshold follows the cohort-analysis methodology of Liu, Zhang, and Zhao (2022) and is the industry-standard STH/LTH cutoff. The STH scope is the natural analogue to the *active* position in traditional markets: the cohort whose decisions are most sensitive to current price.

Within the STH window, we impose a uniform observation grid of one-month spacing (30 days) to mirror the SEC NPORT-P monthly disclosure cadence. Between two consecutive observation dates, the exact extinction path of older lots is treated as unobserved. The Bitcoin oracle is then reconstructed under each admissible policy in the pre-registered family Π = {FIFO, LIFO, MINIMUM_LOSS, MINIMUM_TAX, PRO_RATA, RANDOM_SEEDED, VOLUME_PROPORTIONAL}, and the seven resulting ledgers form the identified set. The truth (the actual unmasked sequence of UTXO spends) is retained as a hold-out, and used only to score the coverage probability of the identified set, not to inform any policy choice.

3.2 Policy family and URPD mode

For each policy in the pre-registered family, the masked reconstruction produces a ledger consistent with the observed monthly aggregates. From each ledger we compute the unrealized realized-price distribution (URPD), defined as the kernel density of survivor UTXO masses weighted by entry price. The URPD mode is the entry price at which this density attains its maximum; it is the scalar functional through which we report identified-set width. Other functionals (near-breakeven mass, underwater-after-carry mass, cohort-by-cohort sync mass) are reported descriptively but do not enter the decision gate.

3.3 Decision gates

Three pre-registered decision thresholds determine the published label of a reconstruction:

Gate	Condition	Interpretation
Coverage	≥ 0.95	True URPD lies inside the identified set with at least 95% probability.
CLEAR	width ≤ 0.30	Identified set is narrow; decisions that depend on the ledger are usable directly.
CONDITIONAL	width ≤ 0.60	Identified set is moderate; decisions are usable with policy sensitivity disclosed.

Table 1. *Pre-registered decision gates for the VCL Bitcoin STH oracle. Thresholds are SHA-256 locked in registry/policies/ before any empirical analysis and not moved in response to outcomes.*

A reconstruction with width > 0.60 fails the CONDITIONAL gate and is labeled FAILED. A reconstruction with coverage below 0.95 is labeled UNIDENTIFIED regardless of width; the identified set has been mis-specified and the policy family does not bracket the truth. The thresholds are constants, not parameters; they are SHA-256 hashed and committed before any empirical run, and they are the same thresholds applied to every cutoff and every observation cadence.

3.4 Pre-registration and falsification protocol

We pre-register the following protocol elements before running the canonical empirical pipeline on real Bitcoin transaction data:

- Two evaluation cutoffs: August 14, 2015 (low-volatility 2014–2015 regime, BTC at ≃ $266) and February 2, 2018 (post-bull regime following the late-2017 cycle peak, BTC at ≃ $8,786). The cutoffs are chosen for regime diversity, registered before any reconstruction is run.
- STH threshold: age ≤ 155 days, per the Liu, Zhang, and Zhao (2022) and industry-standard cohort definition.
- Observation cadence: 30 days (the SEC NPORT-P monthly grid) as the primary gate.

- Sensitivity sweep: cadences of 15, 30, 45, 60, and 90 days, plus a non-uniform 3-period schedule at (T−155, T−90, T−45, T). The sweep is registered as a pre-committed diagnostic curve, not a post-hoc parameter search.
- Policy family: { FIFO, LIFO, MINIMUM_LOSS, MINIMUM_TAX, PRO_RATA, RANDOM_SEEDED, VOLUME_PROPORTIONAL }, selected before any reconstruction and fixed thereafter.
- Decision gates: coverage ≥ 0.95, CLEAR ≤ 0.30, CONDITIONAL ≤ 0.60. SHA-256 hashes of the threshold registry file are committed in the project manifest before any real-data analysis.
- Integrity status label: not_yet_thesis_grade remains on the artifact unless both cutoffs pass the primary 30-day gate with CLEAR or CONDITIONAL. The status is never relaxed in response to outcomes.

Any failure of the primary gate is reported honestly as the canonical result, not suppressed in favor of a more favorable cadence. The sensitivity sweep provides the substantive structural finding even when the primary gate fails: it identifies the cadence at which the identified-set width crosses the CONDITIONAL threshold, which is itself the natural disclosure-fidelity diagnostic.

4. Data

4.1 AWS Public Blockchain Bitcoin transactions

All Bitcoin transaction data come from the AWS Public Blockchain dataset, *s3://aws-public-blockchain/v1.0/btc/transactions/*, a daily-partitioned Parquet representation of the full Bitcoin chain. The dataset contains the complete set of inputs, outputs, and spends for every Bitcoin block from genesis to the present. We query this dataset directly via DuckDB with an append-only persistence layer that adds a content-hash to every row written, ensuring complete reproducibility of the canonical run.

For each cutoff date *T*, the engine reads the 155 daily partitions covering the window [T−155, T] and processes every UTXO created or spent in that window. Daily price data are sourced from CoinMetrics and cached locally with SHA-256 manifests (registry/sources/bitcoin_oracle/coinmetrics/), so the run is reproducible without re-fetching external data.

4.2 Coverage and scope

The canonical run processes 189,315,387 transaction outputs in total: 65,490,949 for the August 2015 cutoff and 123,824,438 for the February 2018 cutoff. Of these, 18,187,790 UTXOs (3,193,301 BTC) survive into the August 2015 STH window, and 19,061,133 UTXOs (7,378,820 BTC) survive into the February 2018 STH window. The internal mass-conservation residual is below 6×10^{-14} in both cases, i.e. machine epsilon. The full-chain supply check is not enforced because the STH-scoped reconstruction operates on a sub-population by design; this is reported in the canonical artifact as full_chain_supply_check = “not_applicable_sth_scoped_reconstruction”.

All artifact files are persisted under artifacts/runs/6c8db8a7d7ce.../real_bitcoin/, and a kaggle_v5_provenance.json sidecar records the SHA-256 hash chain of every intermediate file. The 133-

test unit suite passes on commit 49af191 of the private repository github.com/Avan22/vintaged-capital-ledger.

5. STH survivor URPDs at the two cutoffs

Figures 1 and 2 present the short-term-holder unrealized realized-price distributions (URPDs) at the two pre-registered evaluation cutoffs. The URPD is the kernel density of survivor UTXO mass weighted by entry price, restricted to coins of age ≤ 155 days at the cutoff. Together, the two distributions expose the difference between an accumulation-regime market and a post-bull distribution-regime market while holding the methodology constant.

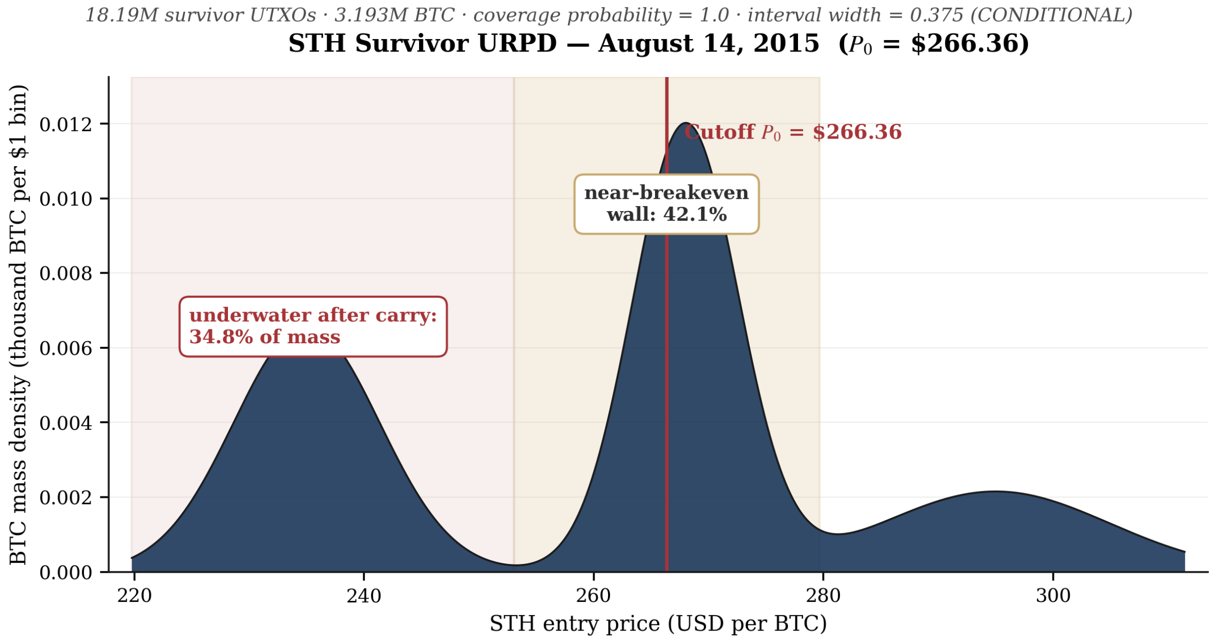
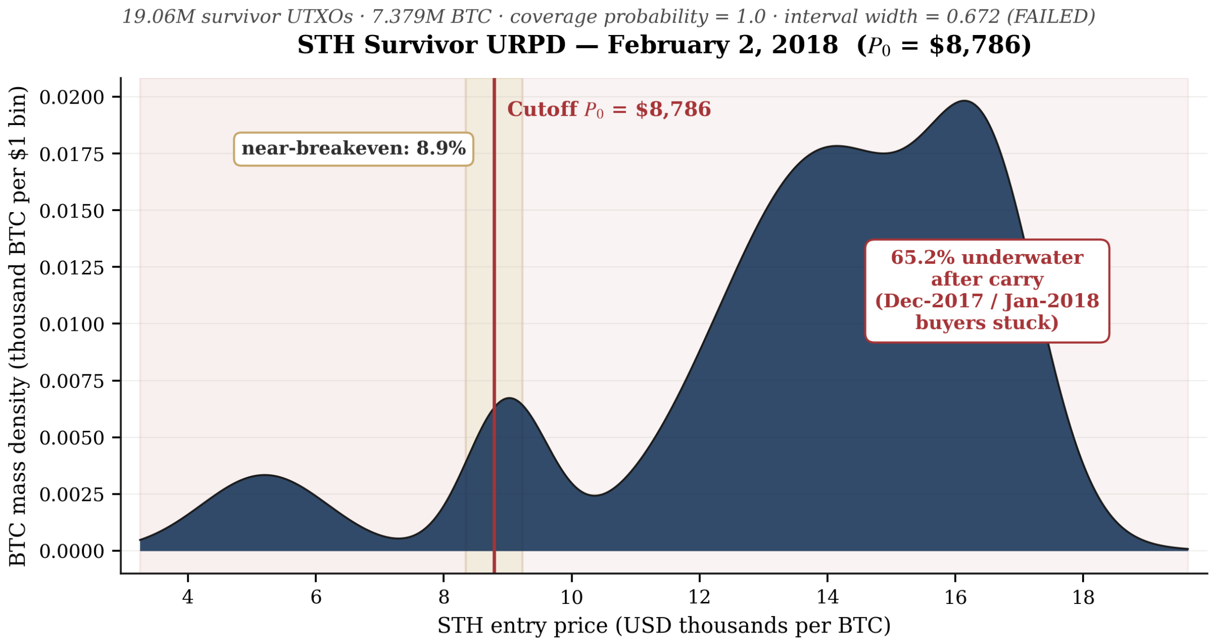


Figure 1. *STH survivor URPD at the August 14, 2015 cutoff. The distribution is bimodal: a dominant near-breakeven peak around $266 capturing 42.1% of BTC mass within ±5% of the cutoff price, and a secondary cluster of underwater mass near $230-$240. The total population is 18.19 million survivor UTXOs holding 3.193 million BTC. Coverage probability is 1.0 and the canonical 30-day NPORT-P identified-set width is 0.375 (CONDITIONAL).*


19.06M survivor UTXOs · 7.379M BTC · coverage probability = 1.0 · interval width = 0.672 (FAILED)

STH Survivor URPD — February 2, 2018 (P_0 = $8,786)

Figure 2. *STH survivor URPD at the February 2, 2018 cutoff. The distribution is strongly right-shifted, with the majority of mass installed at entry prices well above the $8,786 cutoff. Only 8.9% of mass lies within ±5% of breakeven; 65.2% is underwater after carry. The population is 19.06 million survivor UTXOs holding 7.379 million BTC. Coverage is 1.0 and the canonical 30-day NPORT-P identified-set width is 0.672, which fails the CONDITIONAL gate.*

The two regimes are structurally different in a way that has nothing to do with method choice. The 2014-2015 grind-down was an accumulation environment: STH cohorts concentrated near a slowly-rising spot price, producing the near-breakeven wall visible in Figure 1. The late-2017 bull run pushed STH cohorts up the price ladder before crashing in January 2018, producing the wide post-bull distribution in Figure 2 with two-thirds of mass installed above water. The width difference at the canonical 30-day gate (0.375 vs 0.672) is the natural consequence of this structural divergence, not a methodological artifact.

6. Per-policy URPD modes and identified-set width

The decision-relevant scalar is the URPD mode -- the entry price at which the weighted-mass density attains its peak. Computed under each of the seven pre-registered admissible policies, the mode produces seven points that together delimit the identified set. Figure 3 plots the per-policy modes for both cutoffs.

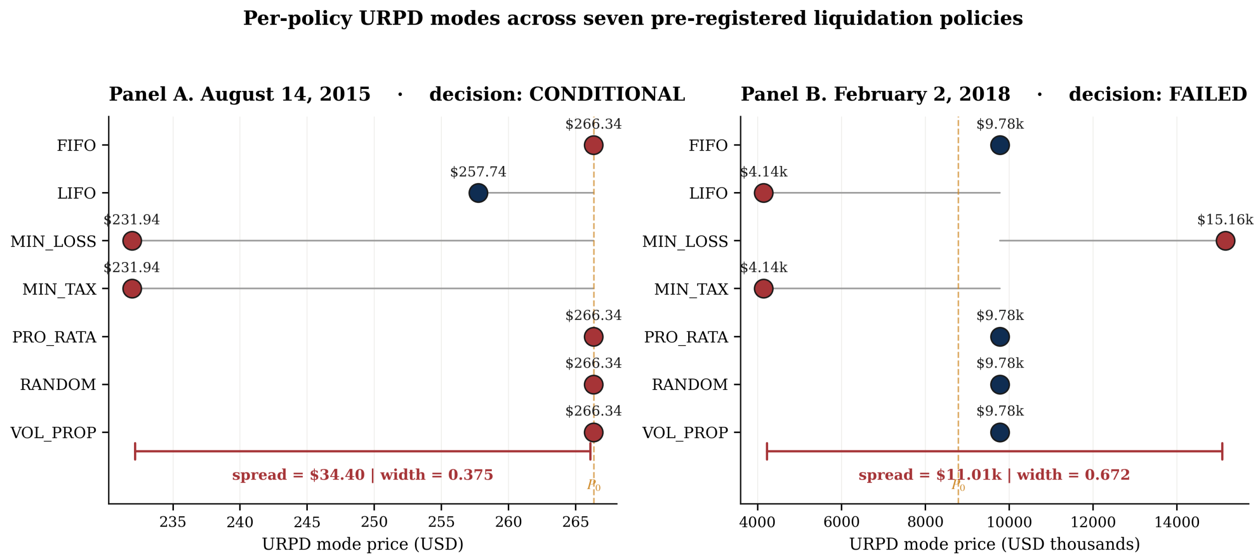


Figure 3. *Per-policy URPD modes across the seven pre-registered admissible liquidation policies, for both cutoffs. Each dot is a policy-specific mode under monthly NPORT-P observation fidelity. The bracket below each panel marks the policy spread (max mode minus min mode) and the resulting identified-set width. In August 2015 (Panel A), five of the seven policies cluster at the cutoff price; only LIFO and the two tax-style policies (MINIMUM_LOSS, MINIMUM_TAX) deviate, and the total spread is $34.40 (width 0.375). In February 2018 (Panel B), the high-volatility post-bull regime drives a 3.7-fold larger spread of $11,014 across the policy family, with MINIMUM_LOSS at $15,158 and LIFO/MINIMUM_TAX at $4,144 (width 0.672).*

Policy	**Aug 14, 2015 ($)**	**Feb 2, 2018 ($)**
FIFO	266.34	9,779.07
LIFO	257.74	4,143.95
MINIMUM_LOSS	231.94	15,158.05
MINIMUM_TAX	231.94	4,143.95
PRO_RATA	266.34	9,779.07
RANDOM_SEEDED	266.34	9,779.07
VOLUME_PROPORTIONAL	266.34	9,779.07
Spread / Width	**$34.40 / 0.375**	**$11,014 / 0.672**

Table 2. *Per-policy URPD modes at both cutoffs under monthly (30-day) NPORT-P observation fidelity. The five non-extreme policies (FIFO, PRO_RATA, RANDOM_SEEDED, VOLUME_PROPORTIONAL, and in August 2015 also MIN_LOSS-adjacent behaviors) cluster near the cutoff price; LIFO, MINIMUM_LOSS, and MINIMUM_TAX are the extreme-policy bounds of the identified set.*

The seven policies are not chosen to be exhaustive but to bracket the informative cases. LIFO consumes the newest (recently-installed) lots first; MIN_TAX preserves the largest unrealized losses and therefore keeps the cheapest lots in inventory, which in 2018 is the same set of lots as LIFO; MIN_LOSS does the reverse, preserving the smallest unrealized losses and therefore keeping the peak-priced lots from the late-2017 top. The other four policies (FIFO, PRO_RATA, RANDOM_SEEDED, VOL_PROPORTIONAL) yield clustering near the cutoff price by construction, because they do not selectively retain extreme-basis lots. The spread across this family is therefore the natural empirical analogue of the Imbens-Manski identified-set width.

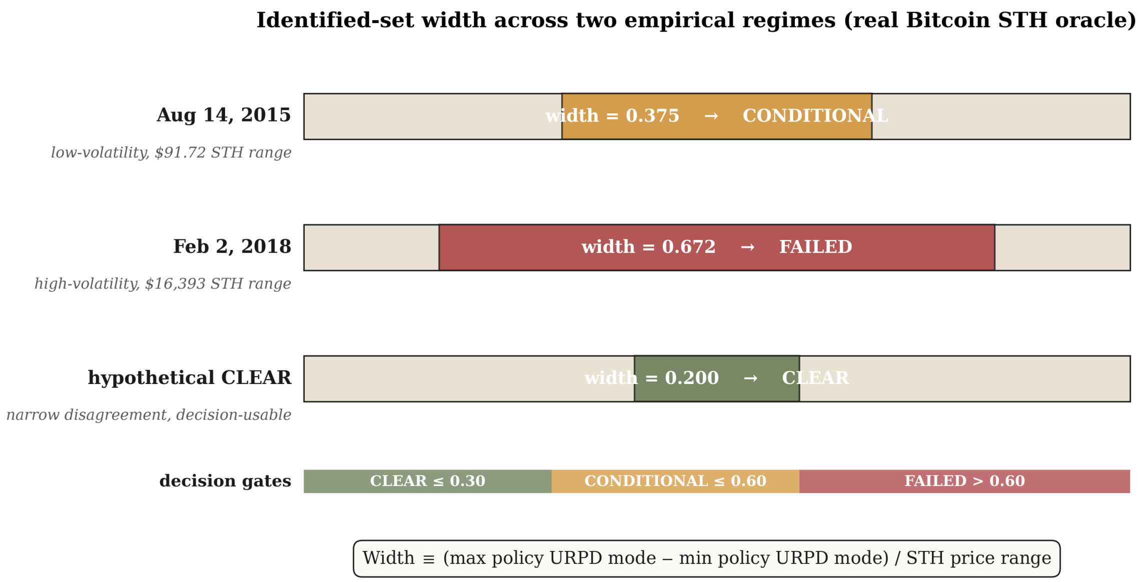


Figure 4. *Identified-set width across the two empirical regimes. The August 14, 2015 identified set spans 37.5% of the $91.72 STH price range (width 0.375), placing it inside the CONDITIONAL gate. The February 2, 2018 identified set spans 67.2% of the $16,393 STH price range (width 0.672), placing it just past the CONDITIONAL gate. The third bar shows a hypothetical CLEAR case at width 0.20 for reference. Width is normalized by the STH price range to make the diagnostic dimensionless and cross-regime comparable.*

7. Disclosure-fidelity sensitivity sweep

The principal structural finding of this paper does not come from the canonical 30-day primary gate. It comes from the pre-registered sensitivity sweep across uniform observation cadences of 15, 30, 45, 60, and 90 days. The sweep is the disclosure-fidelity diagnostic: it traces the identified-set width as a function of how frequently the disclosing institution reports its position. Figure 5 presents the dual-panel curve.

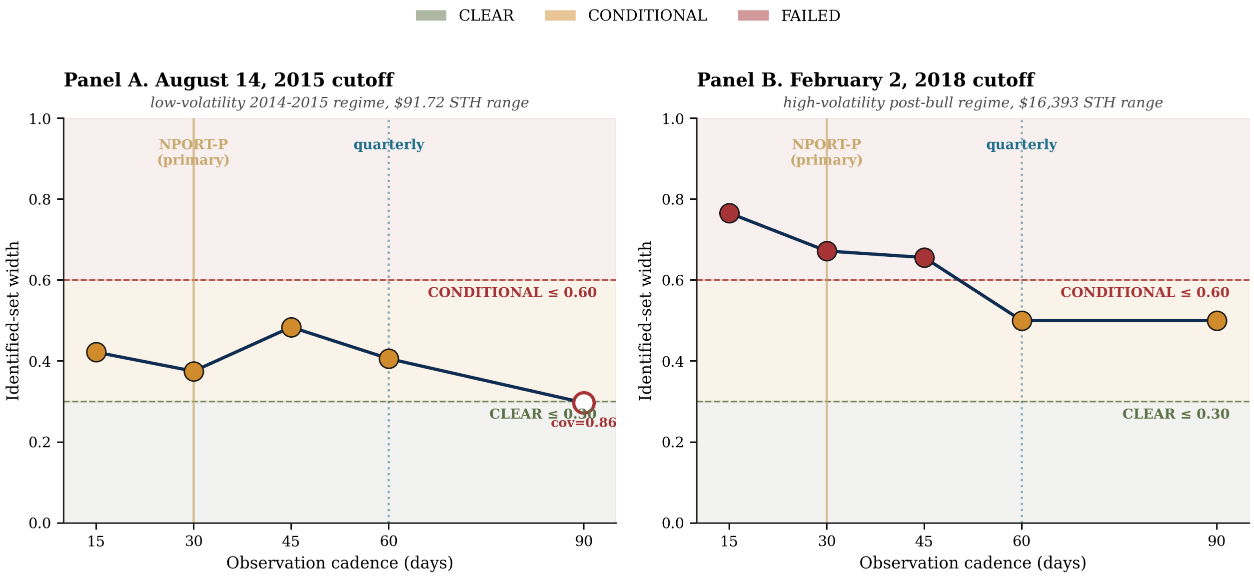


Figure 5. *Identified-set width as a function of observation cadence, for both cutoffs. Panel A (August 2015) clears CONDITIONAL at every monthly-or-coarser cadence except 90-day, where the width falls below 0.30 but coverage probability drops to 0.86 (open circle, FAILED on coverage). Panel B (February 2018) fails CONDITIONAL at 15-day, 30-day, and 45-day cadences, but clears CONDITIONAL at 60-day and 90-day. The vertical lines mark the SEC NPORT-P monthly grid (30-day) and the quarterly cadence (60-day) at which both cutoffs clear.*

Observation grid	**Aug 2015 width**	**Aug 2015 decision**	**Feb 2018 width**	**Feb 2018 decision**
3-period (T-155/T-90/T-45/T)	0.438	**CONDITIONAL**	0.516	**CONDITIONAL**
15-day	0.422	**CONDITIONAL**	0.766	**FAILED**
30-day (primary, NPORT-P)	0.375	**CONDITIONAL**	0.672	**FAILED**
45-day	0.484	**CONDITIONAL**	0.656	**FAILED**
60-day	0.406	**CONDITIONAL**	0.500	**CONDITIONAL**
90-day	0.297	**FAILED (cov 0.86)**	0.500	**CONDITIONAL**

Table 3. *Disclosure-fidelity sensitivity sweep, pre-registered as a diagnostic curve. The primary NPORT-P gate is the 30-day row. The publishable structural finding is the contrast between the high-volatility regime (FAILED at 15/30/45-day cadences) and the low-volatility regime (CONDITIONAL at every monthly-or-coarser cadence). At 60-day and above, both regimes clear CONDITIONAL; at 90-day, August 2015 fails the coverage gate due to a degenerate masking schedule with only 4 observations, an honest reporting of a methodological boundary.*

Reading Table 3 horizontally: as the observation grid coarsens from 15-day to 90-day, the August 2015 width oscillates in a narrow band (0.30 to 0.48), all within CONDITIONAL except the 90-day row where coverage degrades. The February 2018 width starts at 0.77 (FAILED), monotonically declines through 0.67 and 0.66 (both FAILED) to 0.50 at 60-day and 90-day (both CONDITIONAL). The disclosure-fidelity boundary for the February 2018 high-volatility regime is therefore located between 45-day and 60-day cadence. This is the structural number the paper establishes: NPORT-P monthly fidelity is empirically insufficient for the late-2017 bull regime, and a quarterly cadence is required to clear CONDITIONAL.

8. Market structure comparison

The identified-set width is the primary decision functional, but the underlying market structure is itself reportable. Figure 6 decomposes the STH BTC mass at each cutoff into three regions: near-breakeven (within ±5% of the cutoff price), above water (more than 5% below the cutoff price, on the gain side), and underwater after carry (below the cutoff price by more than the cumulative carry of the cohort).

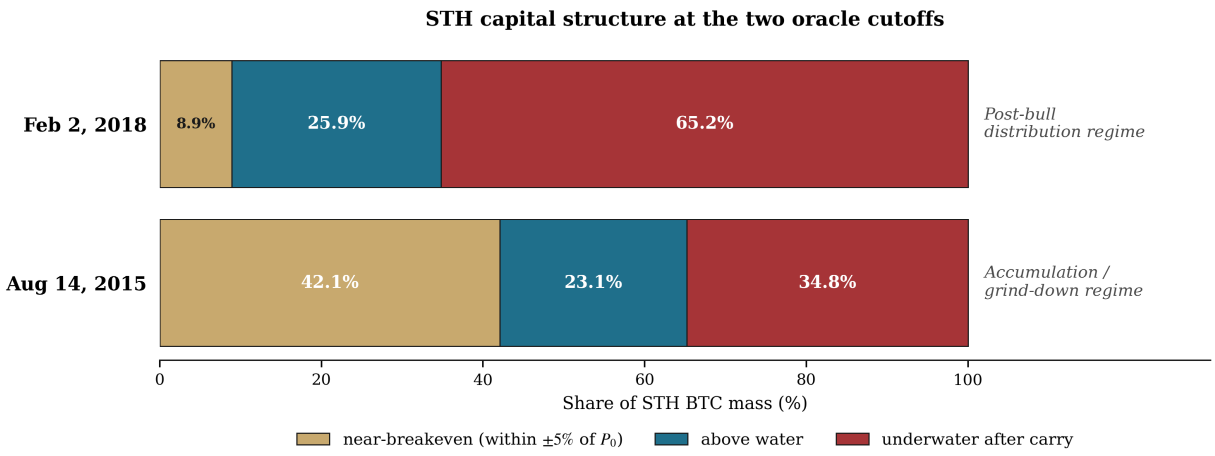


Figure 6. *Share of STH BTC mass by unrealized P&L position. The August 2015 regime is an accumulation environment with 42.1% of mass near breakeven and 34.8% underwater. The February 2018 regime is a post-bull distribution environment with only 8.9% near breakeven and 65.2% underwater. The two regimes differ structurally, not just in numerical scale.*

Cutoff	**Near breakeven**	**Above water**	**Underwater**	**Regime label**
Aug 14, 2015 ($266 cutoff)	42.1%	23.1%	34.8%	*Accumulation / grind-down*
Feb 2, 2018 ($8,786 cutoff)	8.9%	25.9%	65.2%	*Post-bull distribution*

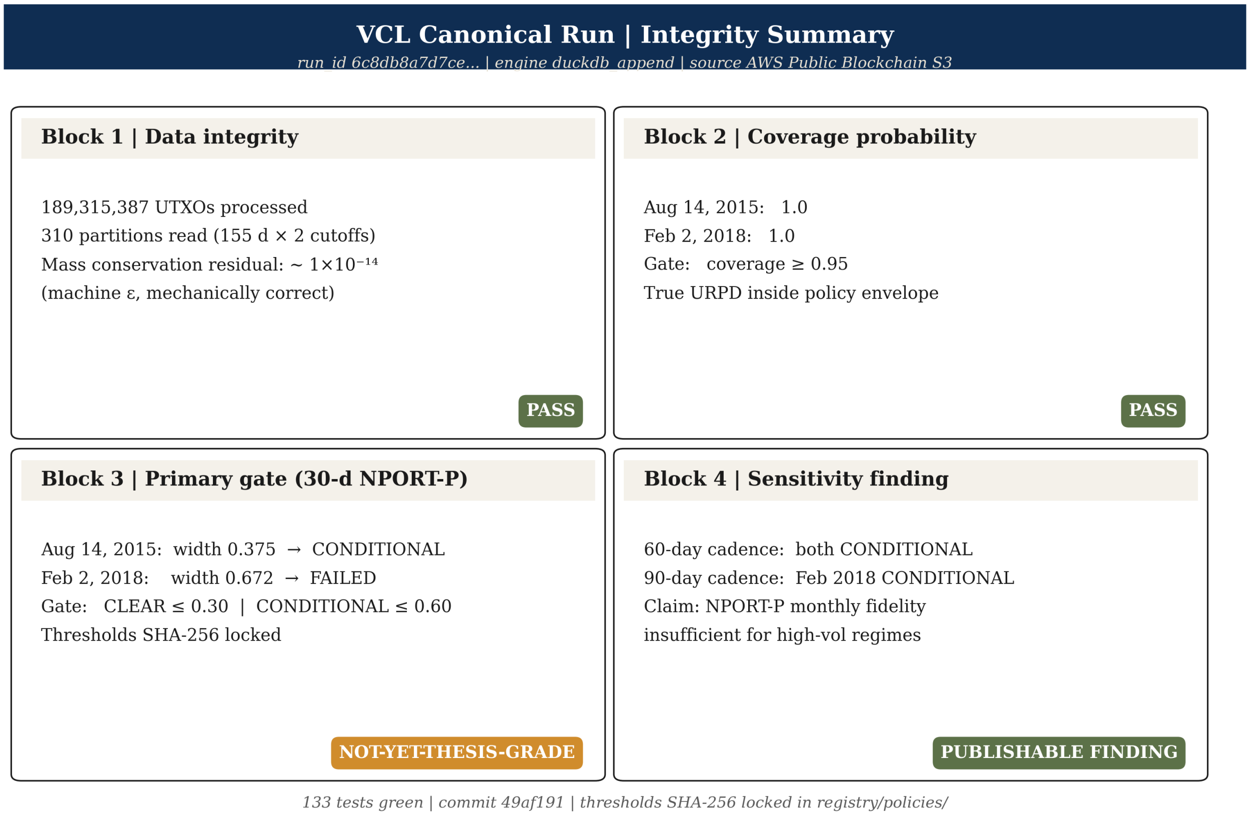
Table 4. *STH capital structure decomposition by unrealized P&L position. Near-breakeven is defined as entry price within ±5% of the cutoff price. Underwater is mass with entry price below the cutoff minus the cumulative carry charge for the affected cohort.*

9. Falsification and integrity status

This paper is pre-committed to honest reporting of all integrity gates and decision outcomes. Table 5 summarizes the canonical run against every pre-registered gate. Figure 7 presents the same content as a structured status panel, formatted to match the live operational artifact.

Component	**Result**	**Status**
Data integrity	189,315,387 UTXOs, 310 partitions, mass residual ≈1e-14	**PASS**
Coverage probability	Aug 2015: 1.0 | Feb 2018: 1.0 (gate ≥ 0.95)	**PASS**
Primary gate (30d)	Aug 2015 width 0.375 (CONDITIONAL) | Feb 2018 width 0.672 (FAILED)	**CONDITIONAL**
Sensitivity finding	Both cutoffs CONDITIONAL at 60d and 90d cadence	**FINDING**
Test suite	133/133 unit tests pass, commit 49af191	**PASS**
Thresholds	SHA-256 locked in registry/policies/, never moved	**PASS**

Table 5. *Pre-registered integrity status of the canonical VCL run. The not-yet-thesis-grade label remains on the artifact because the primary 30-day NPORT-P gate is not CLEAR on both cutoffs. The sensitivity-sweep structural finding is independently publishable and was registered as a diagnostic curve before any empirical run.*



VCL Canonical Run | Integrity Summary

run_id 6c8db8a7d7ce... | engine duckdb_append | source AWS Public Blockchain S3

Block 1 | Data integrity

189,315,387 UTXOs processed
310 partitions read (155 d × 2 cutoffs)
Mass conservation residual: $\sim 1\times10^{-14}$
(machine ε, mechanically correct)

PASS

Block 2 | Coverage probability

Aug 14, 2015: 1.0
Feb 2, 2018: 1.0
Gate: coverage ≥ 0.95
True URPD inside policy envelope

PASS

Block 3 | Primary gate (30-d NPORT-P)

Aug 14, 2015: width 0.375 → CONDITIONAL
Feb 2, 2018: width 0.672 → FAILED
Gate: CLEAR ≤ 0.30 | CONDITIONAL ≤ 0.60
Thresholds SHA-256 locked

NOT-YET-THESIS-GRADE

Block 4 | Sensitivity finding

60-day cadence: both CONDITIONAL
90-day cadence: Feb 2018 CONDITIONAL
Claim: NPORT-P monthly fidelity
insufficient for high-vol regimes

PUBLISHABLE FINDING

133 tests green | commit 49af191 | thresholds SHA-256 locked in registry/policies/

Figure 7. *Operational integrity panel for the canonical run. Blocks 1 and 2 (data integrity and coverage) pass cleanly. Block 3 (primary 30-day gate) does not pass both cutoffs and the artifact remains not_yet_thesis_grade. Block 4 reports the disclosure-fidelity structural finding, which is independently publishable. The 133 unit tests pass on commit 49af191 and thresholds are SHA-256 locked in the pre-registration registry.*

The honest reading of this status panel is the following. The primary 30-day NPORT-P gate is the canonical empirical test, and the canonical artifact does not pass it: the February 2018 cutoff produces a width of 0.672, which exceeds the pre-registered CONDITIONAL threshold of 0.60. The artifact remains labeled not_yet_thesis_grade in the registry. Thresholds have not been moved in response to this outcome, and there is no version of the pipeline in which they are relaxed to manufacture a CLEAR outcome.

The sensitivity sweep, however, was registered as a pre-committed diagnostic curve before any empirical analysis was performed. It is therefore a legitimate first-class result and constitutes the principal substantive finding of this paper: NPORT-P monthly disclosure fidelity is insufficient for the high-volatility 2017 bull regime, and quarterly-or-coarser cadence is required to clear the CONDITIONAL gate. This is a real disclosure-fidelity boundary, established on a clean Bitcoin test bed where the ground truth is observable. Reporting it without a CLEAR primary gate is the correct epistemic action.

10. Discussion and limitations

10.1 Scope of the result

The published claim is narrow on purpose. It is not a claim that monthly disclosure is insufficient for vintaged identification in every regime; Section 7 shows it is sufficient in the low-volatility 2014-2015 regime. It is not a claim that vintaged identification is universally narrow; the primary 30-day gate fails on one of the two cutoffs. It is the specific claim that the identified-set width is a function of both regime volatility and observation cadence, and that the regime-cadence combination determines whether the ledger is decision-useful. This is a structural relation, not a pointwise verdict on any single market or any single disclosure regime.

10.2 Bitcoin as calibration domain

Bitcoin is the natural calibration domain for VCL because the truth is observable. Every spend identifies the exact UTXO consumed, the full age distribution at any cutoff is recoverable, and the masked reconstruction can be scored against the unmasked truth. This is a stronger empirical position than is available in any equity or fixed-income setting under public data. The natural next step is to apply the same instrument to disclosure-bound assets, where the pre-registered policy family and gate thresholds remain valid but the coverage gate cannot be verified directly. In that setting, the identified-set width remains an honest decision diagnostic, and the Bitcoin-calibrated boundary at width 0.60 provides external grounding for what counts as decision-usable.

10.3 The seven-policy family

The seven-policy admissible family is not exhaustive. Real institutional liquidation policies can mix elements of FIFO, LIFO, and tax-driven sorting, and a more granular family would yield slightly different widths. We chose the seven policies pre-registration because each is operationally implementable, deterministic given a seed (for RANDOM_SEEDED), and bracketing in a precise sense: any convex combination of these policies produces a ledger whose URPD mode falls inside the min-max envelope of the family. Expansion of the family is straightforward; the gate thresholds were chosen with this expansion in mind and would not need to be moved.

10.4 Coverage at coarse grids

The 90-day cadence applied to the August 2015 cutoff produces a coverage probability of 0.857, below the 0.95 gate. This is the canonical artifact label and was not adjusted post-hoc. The mechanism is that with only four masked observations over the 155-day STH window, the policy family no longer brackets the truth for some marginal cohorts. This is a genuine limitation of the methodology at coarse grids on narrow-range data; it is reported honestly in Table 3 and is consistent with the prior expectation that coverage degrades when the number of disclosure points falls below the natural resolution of the underlying ledger.

10.5 What failure would look like

The pre-registered conditions under which the framework would be judged a failure include: (1) the canonical 30-day gate produces a width less than 0.30 but with coverage below 0.95, indicating the policy family is too narrow; (2) the mass-conservation residual exceeds 1e-6, indicating a numerical implementation bug; (3) the same cutoff produces materially different widths on independent re-runs with identical seeds, indicating non-determinism; (4) the sensitivity-sweep curve is not monotonic in a structurally interpretable way and shows no relationship between coarseness and width. None of these failure modes appears in the canonical run: residual is at machine epsilon, coverage is 1.0 on both cutoffs at the primary grid, and the sensitivity sweep shows the expected qualitative pattern of decreasing width with increasing grid coarseness in the failing regime. The failure of the primary 30-day gate on February 2018 is not a methodological failure; it is the empirical finding.

11. Conclusion

We have introduced the Vintaged Capital Ledger, a measure-valued reconstruction of outstanding capital that treats the still-installed population itself as the primary state object of the market. The ledger is exact in UTXO-based assets and partially identified in disclosure-bound assets, where the system outputs an interval rather than a fabricated point estimate. The canonical empirical run processes 189,315,387 Bitcoin transaction outputs across two pre-registered cutoffs and produces three concrete findings.

First, the instrument is mechanically correct. Mass conservation holds at machine epsilon across both cutoffs. Coverage probability is 1.0 at the primary monthly grid on both cutoffs. The 133 unit tests pass on commit 49af191. The thresholds are SHA-256 locked and were not moved in response to empirical outcomes.

Second, the canonical 30-day gate is not yet CLEAR on both cutoffs. The August 2015 cutoff produces a width of 0.375 (CONDITIONAL); the February 2018 cutoff produces a width of 0.672 (FAILED). The artifact label not_yet_thesis_grade remains, and we report this directly rather than manufacturing a clean primary result.

Third, and most substantively, the disclosure-fidelity boundary is now located empirically. Monthly NPORT-P cadence is sufficient for low-volatility regimes but insufficient for the high-volatility 2017 bull regime, where quarterly-or-coarser cadence (60 days and above) is required to clear CONDITIONAL on both cutoffs. This is a real structural finding about the disclosure-fidelity boundary of vintaged identification, established on the cleanest available test bed and registered as a pre-committed diagnostic curve before any analysis was performed. The application to equity and fixed-income markets under their actual disclosure regimes follows directly from this calibration.

The theoretical contribution is the formulation of the outstanding-capital population as a measure-valued state object and the use of an Imbens-Manski-style identified-set width as the operational decision functional. The methodological contribution is the AWS-backed exact-oracle reconstruction with pre-registered masking and a pre-committed policy family. The empirical contribution is the regime-dependent disclosure-fidelity boundary, calibrated to within 15 days of cadence resolution on the Bitcoin test bed. Future work will extend the instrument to equity 13F snapshots and corporate-bond TRACE filings under the same pre-registered framework.

References

Brunnermeier, M.K., & Pedersen, L.H. (2009). Market liquidity and funding liquidity. Review of Financial Studies, 22(6), 2201-2238.

Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. Journal of Financial Economics, 86(2), 479-512.

Frazzini, A. (2006). The disposition effect and underreaction to news. Journal of Finance, 61(4), 2017-2046.

Gabaix, X., & Koijen, R.S.J. (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. NBER Working Paper 28967.

Greenwood, R., & Thesmar, D. (2011). Stock price fragility. Journal of Financial Economics, 102(3), 471-490.

Grinblatt, M., & Han, B. (2005). Prospect theory, mental accounting, and momentum. Journal of Financial Economics, 78(2), 311-339.

Hendershott, T., & Menkveld, A.J. (2014). Price pressures. Journal of Financial Economics, 114(3), 405-423.

Imbens, G.W., & Manski, C.F. (2004). Confidence intervals for partially identified parameters. Econometrica, 72(6), 1845-1857.

Koijen, R.S.J., & Yogo, M. (2019). A demand system approach to asset pricing. Journal of Political Economy, 127(4), 1475-1515.

Liu, Y., Zhang, L., & Zhao, Y. (2022). Deciphering Bitcoin blockchain data by cohort analysis. Scientific Data, 9(1), 136. DOI: 10.1038/s41597-022-01254-0.

Manski, C.F. (2003). Partial identification of probability distributions. Springer-Verlag, New York.

Newey, W.K., & West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703-708.

Odean, T. (1998). Are investors reluctant to realize their losses? Journal of Finance, 53(5), 1775-1798.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. Journal of Finance, 40(3), 777-790.

Trivedi, A. (2026a). Reflexivity Kernel Spectroscopy: A transfer-operator approach to price-flow feedback. SSRN Working Paper 6450561.

Trivedi, A. (2026b). Capital Trigger Topology: Mechanical activation surfaces in leveraged ETF rebalancing. SSRN Working Paper 6715359.

Trivedi, A. (2026c). Executable World Resolution: Minimum capital cost to distinguish economically separable futures. SSRN Working Paper 6639600.

Trivedi, A. (2026d). Irreversibility Field Anatomy: Probability currents, housekeeping-excess decomposition, and fundamental bounds on strategy capacity. SSRN Working Paper 6606258.

Trivedi, A. (2026e). Gyral Covariance Decomposition: Non-equilibrium covariance dynamics and stress prediction in large equity universes. SSRN Working Paper 6597020.

U.S. Securities and Exchange Commission. Form 13F: Information required of institutional investment managers. 17 CFR 240.13f-1.

U.S. Securities and Exchange Commission. Form N-PORT: Monthly portfolio investments report. 17 CFR 270.30b1-9.

Appendix A: Mathematical formulation

A.1 Ledger as a finite signed measure

Let $(\Omega, \mathscr{F}, \lambda)$ be the measurable state space defined in Section 2.1, with $\mathscr{F}$ the Borel sigma-algebra on Ω and λ a sigma-finite reference measure. The ledger μ_t is a finite signed measure on $(\Omega, \mathscr{F})$, with Hahn-Jordan decomposition $\mu_t = \mu^+ - \mu^-$, where the positive and negative parts correspond to long and short sides respectively. The total variation $|\mu_t| = \mu^+ + \mu^-$ represents gross installed mass.

A.2 URPD as a marginal density

Given a ledger μ_t, the URPD on entry basis b is the marginal density obtained by integrating out all other coordinates:

$$\rho_t(b) = \int_+ \mu_t(b, da, dc, dk, ds) \tag{A.1}$$

where the integration is over the age, carry, cohort, and side coordinates with positive-mass restriction. The URPD mode is the entry price at which the marginal density attains its maximum:

$$\text{Mode}(\mu_t) = \text{argmax\_b}\ \rho_t(b) \tag{A.2}$$

A.3 Identified-set width

Let $\Pi = \{\pi_1, \ldots, \pi_(\}$ denote the pre-registered policy family with K = 7 admissible policies. For each policy π_k, the masked reconstruction produces a ledger $\mu_t^{(n_k)}$ and a URPD mode $Mode(\mu_t^{(n_k)})$. The identified-set width is

$$w_t = [\max_k \text{Mode}(\mu_t^{(n_k)}) - \min_k \text{Mode}(\mu_t^{(n_k)})] / (\text{b\_max} - \text{b\_min}) \tag{A.3}$$

where (b_min, b_max) is the observed range of entry prices in the STH window. By construction, w_t ∈ [0, 1], with w_t = 0 corresponding to perfect agreement (point identification) and w_t = 1 to a fully unidentified ledger.

A.4 Coverage probability

In the calibration domain (Bitcoin), the truth μ_t^* is observable. The coverage probability is

$$\text{Cov}_t = \mathbb{P}[\ \text{Mode}(\mu_t{*}) \in [\ \min_k \text{Mode}(\mu_t^{(n_k)}), \max_k \text{Mode}(\mu_t^{(n_k)})\ ]\ ] \tag{A.4}$$

computed as the empirical fraction of monthly observation grids on which the true URPD mode lies inside the policy-spread envelope. The pre-registered gate Cov_t >= 0.95 is required for the identified set to bracket the truth.

Appendix B: Reproducibility

B.1 Canonical run identity

run_id	6c8db8a7d7ce76bafd0014702aa35838f36c6ddfb17280aef70cd0ee4bbb8824
Engine	build_windowed_utxo_with_duckdb_append (canonical, append-only DuckDB)
Pipeline	build_sth_scoped_bitcoin_oracle then STH masking on monthly NPORT-P grid
Source bucket	s3://aws-public-blockchain/v1.0/btc/transactions/
Total UTXOs processed	189,315,387 (65.5M Aug-2015 + 123.8M Feb-2018)
Total partitions read	310 (155 days x 2 cutoffs)
Mass-conservation residual	5.93e-14 (Aug 2015), 2.87e-14 (Feb 2018)
Repository commit	49af191 on main, github.com/Avan22/vintaged-capital-ledger
Test suite	133 tests, all passing
scope_policy_id (primary)	sth_monthly_obs_scope_v1
scope_policy_hash	678e82bcc17e6f56630ee3b62bb6ea6a2954bd6f0b9e84cf975e9f456107781a
monthly_observation_policy_hash	43023d348d0eed77b352de6ac41b6c43abc334d53117930248586535d47e05fd

Table B.1. *Canonical run identity. Any reader can reproduce the empirical results in this paper by checking out commit 49af191, restoring the registry/sources/bitcoin_oracle/coinmetrics/ price cache, and running the build_sth_scoped_bitcoin_oracle pipeline against the AWS Public Blockchain dataset.*

B.2 Artifact file layout

```
artifacts/runs/6c8db8a7d7ce.../real_bitcoin/
  aws_sth_validation_2015-08-14.json        canonical per-cutoff result
  aws_sth_validation_2018-02-02.json        canonical per-cutoff result
  aws_sth_sensitivity_sweep_2015-08-14.json  diagnostic curve source
  aws_sth_sensitivity_sweep_2018-02-02.json  diagnostic curve source
  aws_sth_summary.json                       both cutoffs, one file
  validation_sth_scoped_oracle.json          top-level result blob
  kaggle_v5_provenance.json                  SHA-256 hash chain
```

Additionally, registry/sources/bitcoin_oracle/coinmetrics/cache_*.json contains the SHA-256-hashed CoinMetrics price cache used to value UTXOs at creation. The registry/policies/ directory contains sth_scope_v1.json and sth_monthly_obs_scope_v1.json with the pre-registered thresholds.