Social Science and Management
Print ISSN: 3007-6854
Online ISSN: 3007-6862
DOI: https://doi.org/10.61784/ssm4096

DISCLOSURE QUALITY AND VALUATION RISK IN LEVEL 3 INVESTMENTS: A DISCLOSURE-IMPLIED RISK FRAMEWORK FOR U.S. BUSINESS DEVELOPMENT COMPANIES

WenJie Ping[1*], HuiJie Fan[2]
[1]*Georgetown University, Washington 20057, DC, USA.*
[2]*Hunan Agricultural University, Changsha 410127, Hunan, China.*
[*]*Corresponding Author: WenJie Ping*

Abstract: Business development companies (BDCs) report most of their investments as Level 3 fair value measurements, so the net asset value (NAV) relied on by investors and auditors rests on unobservable inputs. This paper asks how much of that valuation risk can be recovered from what BDCs publish. We propose a disclosure-implied valuation risk (DIVR) measure that reconstructs each row of the ASC 820 unobservable-input table from its range and weighted average with a mean-matched Beta-PERT law, calibrates ranges for the number of holdings with order statistics, propagates rows to NAV through a Gaussian copula, and adds a benchmark-gap estimate of overstatement from stale or optimistic marks. DIVR is paired with a nine-item Level 3 Disclosure Quality Index (L3-DQI). Both are evaluated on a disclosure-calibrated panel of 48 BDCs over fiscal years 2016–2023 (384 fund-years, 54,618 positions) in which hidden portfolios are fixed while disclosure formats vary counterfactually. Granular input tables raise the out-of-fund R^2 for latent valuation dispersion from 0.71 to 0.80, whereas weighted averages are what make overstatement detectable (rank correlation 0.54 versus 0.09). In price-to-NAV regressions estimated by panel OLS and double machine learning, disclosure-implied risk carries a discount and the benchmark gap is priced only after weighted averages became mandatory, while the price effect of disclosure quality is positive but small relative to price noise in a market of 48 listed funds. For auditors, a back-testing score captures 44% of absolute misstatement when 20% of Level 3 fair value is tested, against 22% for size-based selection. Recoverability and audit results hold in all 10 independent panels.
Keywords: Business development companies; Fair value hierarchy; Level 3 inputs; Disclosure quality; Valuation risk; Double machine learning; Conformal prediction; Audit sampling

1 INTRODUCTION

Business development companies (BDCs) are closed-end investment companies regulated under the Investment Company Act of 1940 that lend to and invest in middle-market firms, including borrowers that bank regulation steers toward nonbank lenders [1]. Their loans and co-investments rarely trade, so BDCs measure most of their portfolios with Level 3 inputs under ASC 820: market yields, EBITDA multiples and discount rates selected by the adviser, often with independent valuation firms, under board oversight. The NAV that determines regulatory leverage, incentive fees, dividend capacity and the reference point for the share price is therefore an estimate whose reliability depends on how those inputs are chosen and disclosed.

Regulation has tightened on three fronts. ASU 2018-13 requires public entities to disclose the range and weighted average of significant unobservable inputs for fiscal years beginning after December 15, 2019 [2]. SEC Rule 2a-5, adopted in December 2020 with a compliance date of September 8, 2022, specifies how the board or a valuation designee of a registered fund or BDC determines fair value in good faith, including valuation risk assessment and the testing of methodologies [3]. PCAOB AS 2501 strengthened the audit of fair value measurements for fiscal years ending on or after December 15, 2020 [4]. Each change presumes that published Level 3 information lets investors and auditors understand valuation risk.

Whether it does is an open question. Level 3 estimates are priced at a discount and carry information risk [5-6], managers exercise discretion over them [7], and auditors find them difficult to verify [8-10]. Empirical work nevertheless tends to proxy opacity with the share of assets classified as Level 3, which is close to one for most BDCs and carries little cross-sectional information. The quantitative input tables, the most detailed part of the disclosure, are rarely used in a structured way, and there is no evidence on which of their elements convey risk.

This paper asks three questions. How much of a BDC's latent valuation dispersion and overstatement can be recovered from its Level 3 disclosures, and which elements matter? Are disclosure-based risk measures and disclosure quality reflected in the price-to-NAV ratio? Can the same information direct audit testing toward the positions most likely to be misstated? The paper makes four contributions:

1) A disclosure-implied valuation risk (DIVR) measure that converts each row of an ASC 820 input table into a mean-matched Beta-PERT distribution, removes the mechanical widening of disclosed ranges with the number of holdings by order-statistic calibration, propagates rows to NAV with a Gaussian copula, and adds a benchmark-gap estimate of overstatement from stale or optimistic marks.

© By the Author(s) 2024, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

2) A nine-item Level 3 Disclosure Quality Index (L3-DQI) that codes the elements of current BDC reporting, including the weighted-average requirement of ASU 2018-13 and Rule 2a-5 governance disclosure.
3) A disclosure-calibrated BDC panel (BDC-Sim) in which every published figure is produced by aggregating holdings as a BDC would, so that measures are validated against known ground truth and disclosure formats are varied counterfactually with holdings fixed.
4) An estimation design built on recent methods: double machine learning with fund-clustered cross-fitting [11-12], conformal prediction beyond exchangeability and gradient-boosted trees [13-16].
The results separate two functions of the input table. Granularity by asset class conveys valuation dispersion beyond portfolio composition, raising the out-of-fund R^2 from 0.71 to 0.80, whereas weighted averages are what reveal overstatement. Disclosure-implied risk is reflected in price-to-NAV and the benchmark gap is priced only after weighted averages became mandatory; the price effect of disclosure quality is positive on average but small relative to price noise, so single-market tests of it have limited power. A back-testing score doubles the misstatement captured by an audit sample of fixed size.

2 INSTITUTIONAL BACKGROUND AND RELATED WORK

2.1 Level 3 Reporting by BDCs

Most BDCs are externally managed and pay base fees on gross assets and incentive fees on income, which ties adviser compensation to reported values. The Small Business Credit Availability Act of 2018 permitted BDCs to lower their minimum asset coverage ratio from 200% to 150%, and sector leverage rose thereafter. ASC 820 requires a roll-forward of Level 3 balances and quantitative information about significant unobservable inputs, which BDCs present as tables listing, for each asset class and valuation technique, the fair value, the input, its range and, since ASU 2018-13, its weighted average [2]. Rule 2a-5 adds documented valuation risk assessment, methodology testing, oversight of pricing services and the designation of a valuation designee [3].

2.2 Fair Value Hierarchy and Information Risk

Level 3 net assets are valued less per dollar than Level 1 and 2 net assets, particularly under weak governance [5], and Level 3 exposure is associated with higher equity betas [6], consistent with models in which poorer information raises the cost of capital [17-18]. Reliability concerns about model-based fair values were central to the post-crisis debate [19]. Discretion is used strategically: insurers' self-estimated Level 3 values exceed the consensus of other holders when incentives to appear healthy are strong [7]. Disclosure mitigates these frictions. SEC comment-letter monitoring reduces fair value uncertainty [20], voluntary disclosures about valuation controls lower information risk for Level 3 estimates [21], and specific disclosures are more informative than generic ones [22]. The incomplete revelation hypothesis predicts that information that is costly to extract, such as dispersed input tables, is only partly impounded in prices [23].

2.3 Valuation Discretion in Funds and Closed-End Fund Discounts

Hedge funds without strong pricing controls report smoother returns [24], private equity funds' interim valuations display discretion around fundraising [25], and illiquid holdings generate serial correlation consistent with stale marks [26]. Closed-end fund discounts have been attributed to investor sentiment and to managerial ability and fees [27-28]. For BDCs, uncertainty about the NAV itself adds a further component.

2.4 Auditing Level 3 Estimates

Measurement uncertainty in fair value estimates can exceed materiality [8]; auditors depend heavily on valuation specialists and management models [9]; and standard setters face difficulty in specifying procedures for complex estimates [10]. Office-level expertise in auditing Level 3 estimates is associated with higher fair value audit quality [29]. AS 2501 requires auditors to address management bias [4], which presupposes a way to rank positions by misstatement risk.

2.5 Methodological Background

Double machine learning estimates a low-dimensional effect while controlling flexibly for confounders [11], and the DoubleML implementation supports cluster-robust cross-fitting [12]. Conformal prediction produces distribution-free prediction intervals [14], and weighted conformal prediction bounds the coverage loss when data are not exchangeable, as in time-ordered panels [13]. Tree ensembles remain strong learners for tabular data [15], with LightGBM as an efficient implementation [16]. Three-point (minimum, most likely, maximum) PERT distributions are a standard device for encoding ranges [30].

3 DISCLOSURE-IMPLIED VALUATION RISK FRAMEWORK

3.1 Disclosure Representation

For BDC i in fiscal year t, the Level 3 input table has rows $r = 1,\ldots,R_{it}$. Each row reports a fair value F_r, a technique k_r (yield analysis, discounted cash flow, market multiple or enterprise-value waterfall), an input type, a range $[a_r, b_r]$ and, when provided, the fair-value-weighted average w_r. The central value is $\mu_r = w_r$ when the weighted average is disclosed and $\mu_r = (a_r + b_r)/2$ otherwise. Portfolio-level disclosures provide the weighted-average contractual yield on performing debt $y_{c,it}$, the number of holdings n_{it}, leverage and the asset-class composition of fair value, see Figure 1.

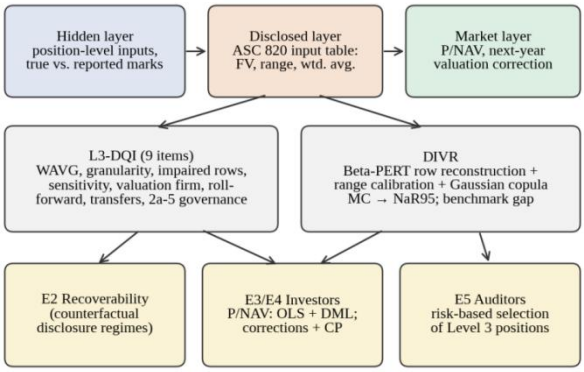


Figure 1 Research Framework

Note: Hidden holdings generate the published ASC 820 table and market outcomes; L3-DQI and DIVR are computed from the published layer only and evaluated for recoverability, pricing, prediction and audit use.

3.2 Level 3 Disclosure Quality Index

L3-DQI averages nine binary items q_{ijt} that code the content of BDC Level 3 reporting (Table 1):

$$DQI_{it} = (1/9)\sum_j q_{ijt}\,. \tag{1}$$

Table 1 L3-DQI Items and Frequencies by Period

Item	Coded 1 if the 10-K …	16–19	20–21	22–23
WAVG	reports weighted averages for every input row	0.40	1.00	1.00
GRAN	separates rows by asset class	0.56	0.69	0.73
SEPIMP	shows impaired loans in separate EV rows	0.83	0.90	0.92
SENS_D	states directional input sensitivity	0.79	0.79	0.79
SENS_Q	quantifies NAV impact of input shifts	0.12	0.12	0.15
TPV	discloses independent valuation coverage	0.65	0.65	0.65
ROLL	presents the roll-forward by asset class	0.65	0.65	0.65
XFER	explains transfers into/out of Level 3	0.60	0.60	0.60
GOVTX	describes valuation designee/2a-5 process	0.27	0.27	0.83
L3-DQI	equal-weighted mean of the nine items	0.54	0.63	0.70

Note: Frequencies are means over fund-years in the BDC-Sim panel.

3.3 Row Reconstruction with Range Calibration

Each row is modeled as a Beta-PERT law on $[a_r, b_r]$ whose mean equals μ_r. With shape $\lambda = 4$ the mode and shape parameters are

$$m_r = \min\{b_r, \max\{a_r, [(\lambda+2)\mu_r - a_r - b_r]/\lambda\}\}, \tag{2}$$

$$\alpha_r = 1 + \lambda(m_r - a_r)/(b_r - a_r),\ \beta_r = 1 + \lambda(b_r - m_r)/(b_r - a_r). \tag{3}$$

A disclosed range is the minimum and maximum over a row's holdings, so it widens mechanically with the number of holdings. Each row is therefore rescaled by

$$c_r = 2\sqrt{7} \,/\, \mathrm{E}[R(n_r)], \tag{4}$$

where $n_r = n_{it}F_r/\Sigma_s F_s$ approximates the holdings in row r, $\mathrm{E}[R(n)] \approx 2\Phi^{-1}((n - 0.375)/(n + 0.25))$ is the expected range of n standard normal draws, and $2\sqrt{7}$ is the range-to-standard-deviation ratio of a symmetric PERT law. The calibration makes row dispersion comparable across BDCs of different size.

3.4 Propagation to NAV

Rows are coupled by a Gaussian copula with correlation 0.6 between rows that share an input type and 0.3 otherwise, representing coordinated re-marking. For draw s,

$$\Delta NAV^{(s)} = NAV_{it}^{-1}\sum_r S_r h c_r \left(x_r^{(s)} - \mathrm{E}x_r\right), \tag{5}$$

where $S_r = -D_k F_r$ for yield-type inputs (spread duration 3.0 for yield analysis, 4.5 for discounted cash flow), $S_r = e_k F_r/\mu_r$ for multiples (elasticity 2.0 for equity, 1.0 for enterprise-value waterfalls), and $h = 0.5$ scales a coordinated shift of half a calibrated dispersion unit. With 2,000 draws we report NaR95 $= -Q_{0.05}(\Delta NAV)$ and the debt and equity components $DIVR_D$ and $DIVR_E$, the standard deviations of the yield and multiple parts of (5).

3.5 Benchmark Gap

Floating-rate loans reprice with credit spreads rather than base rates. If marks track the market, the change in disclosed market yields relative to contractual yields should follow the change in a market spread index S_t. The input-benchmark gap

$$IBG_{it} = (S_t - \bar{S}_t) - (y_{YA,it} - y_{c,it}), \quad (6)$$

where $\bar{S}_t$ is the trailing three-year mean of the index and $y_{YA,it}$ is the fair-value-weighted central value of yield-analysis rows, is positive when reported yields lag the market. Its NAV impact, $Gap_{it} = 3.0 \cdot F_{YA,it} \cdot IBG_{it}/NAV_{it}$, truncated at ±15% of NAV, estimates overstatement from stale or optimistic marks.

3.6 Hypotheses

H1: Granular tables allow DIVR to recover latent valuation dispersion beyond composition. H2: Weighted averages are necessary for the gap to recover overstatement. H3: DIVR and L3-DQI are reflected in price-to-NAV, and the gap is priced only when it is informative. H4: Disclosure- and back-testing-informed selection of Level 3 positions captures more misstatement than size-based selection.

4 DISCLOSURE-CALIBRATED PANEL AND ESTIMATION

No public dataset links the published Level 3 tables of BDCs to the true values of their holdings, which is what validating a risk measure requires. We therefore construct BDC-Sim, a panel calibrated to the structure of BDC 10-K reporting, in which every published figure is produced by aggregating simulated holdings as a BDC would. Ground truth is known by construction, and disclosure formats can be changed while holdings stay fixed. All results refer to this panel.

4.1 Hidden Portfolio Layer

The panel covers $N = 48$ listed BDCs over fiscal years 2016–2023. Fund traits include external management (85% of funds), valuation governance, disclosure propensity (latent correlation 0.4 with governance), borrower-risk orientation, fee rates and an asset mix over first lien, second lien, subordinated, equity and structured/joint-venture positions that drifts toward first lien after 2018. Each fund-year holds 30–350 positions (54,618 in total). A debt position with contractual spread c and spread duration D has true price $P^* = 1 - D(y^* - r_t - c)$, where y^* is the current market yield and r_t the base rate. Management reports

$$\hat{y} = \kappa_{it} y* + (1 - \kappa_{it})(r_t + c) - \varphi_{it}, \quad (7)$$

so marks are anchored to origination with weight $1 - \kappa$, where κ rises with governance and independent-valuation coverage, and shifted by optimism φ, which is larger for externally managed, weakly governed and highly levered funds. Equity is valued from EBITDA multiples, $\hat{m} = \kappa m^* + (1 - \kappa)m_0 + \psi$, net of portfolio-company debt, and impaired loans by enterprise-value waterfalls. Broadly syndicated loans and listed equity are Level 2 or Level 1 at observable prices. Latent overstatement is $OS_{it} = \Sigma_k(\hat{V}_k - V^*_k)/NAV_{it}$, and latent valuation dispersion σ^*_{it} aggregates position-level uncertainties with common correlation 0.25.

4.2 Disclosure and Market Layers

Each fund-year's table is generated from its Level 3 holdings. Granular funds report rows by asset class and technique; others aggregate debt classes and, in half of the cases, fold impaired loans into yield rows at implied yields. Weighted averages are voluntary before fiscal 2020 and universal afterwards, Rule 2a-5 governance text appears from fiscal 2022, and the other items follow sticky fund-level propensities. Price-to-NAV follows

$$\ln(P/NAV)_{it} = a_t + \gamma' z_{it} - 1.0 \cdot OS_{it} - 2.0 \cdot \sigma*_{it} - 0.12(1 - Q_{it}) + u_i + e_{it}, \quad (8)$$

where **z** contains internal management (0.22), net investment income return relative to the year mean (2.0), size (0.02) and non-accruals (−0.35); Q_{it} is the share of position-level input variation resolved by the published row statistics, an information-risk channel [17-18]; $u_i \sim N(0, 0.05^2)$ and e_{it} is AR(1). Year effects a_t set the cross-sectional mean of P/NAV on a stylized path with discounts in 2018, 2020 and 2022. The realized valuation correction in t+1 is $RC_{i,t+1} = 0.6 \cdot OS_{it} + 0.3 \cdot \sigma^*_{it}\varepsilon_{i,t+1} + \nu_{t+1}$, observed through fiscal 2023. Table 2 summarizes the calibration.

Table 2 Calibration of the BDC-Sim Panel

Component	Specification / resulting moment
Panel	48 BDCs × FY2016–2023; 384 fund-years; 54,618 positions
Asset mix	Mean shares: first lien 0.63, second lien 0.13, subordinated 0.06, equity 0.11, structured/JV 0.07
Market states	Base rate 1.00% (2016), 0.21% (2021), 4.66% (2022), 5.33% (2023); loan spread index 5.0–6.7%; EBITDA multiple 9.7–12.0×
Mark formation	Anchoring κ mean 0.65 (0.26–0.97); optimism φ mean 30 bp (max 89 bp); ψ = 60φ turns
Leverage	Debt/NAV 0.60 (2016) to 1.05 (2023)
Outcomes	Level 3 share 0.89; NII/NAV 7.4%; mean P/NAV 0.87–1.04; OS 0.3–4.6% and σ* 3.9–5.4% of NAV (year means)

Note: Market states are stylized year-end values; the base-rate path approximates three-month USD benchmark rates.

4.3 Counterfactual Disclosure Regimes

Holding holdings fixed, every table is regenerated under four regimes: R1 aggregated rows with range midpoints, R2 aggregated rows with weighted averages, R3 granular rows with midpoints, and R4 granular rows with weighted averages, the post-ASU 2018-13 best practice. Regime R0 uses only financial-statement composition: Level 3 assets to NAV, equity and junior-debt shares, non-accruals and leverage.

4.4 Estimation

Recoverability is measured by the mean within-year Spearman correlation and by out-of-fund R^2 from six-fold cross-validation grouped by fund with year effects. Pricing regressions use standardized regressors, year fixed effects and fund-clustered standard errors. The partially linear model $\ln(P/\text{NAV}) = \theta D + g(\mathbf{X}) + \varepsilon$ is estimated with DoubleML [12], LightGBM nuisance learners [16], five folds, fund-level clusters and five repetitions. Correction forecasts use expanding windows with test years 2019–2022. Ninety-percent prediction intervals compare split conformal prediction, a DIVR-adaptive version that scales residuals by a difficulty model fitted on DIVR components, and a weighted version that down-weights older calibration years by a factor of 0.7 per year [13]. Audit selection is evaluated within each fund-year by the share of absolute misstatement captured as Level 3 fair value is tested in score order; the GBDT score is trained on position-level back-testing labels with fund-grouped cross-fitting, mirroring the comparison of marks with realized exit prices that Rule 2a-5 testing produces.

5 RESULTS

5.1 Descriptive Patterns

Table 3 reports means by fiscal year. Level 3 holdings account for 0.89 of investments with little variation over time, which is why the Level 3 share is a weak signal. Disclosed weighted-average market yields rise from 7.5% in 2016 to 11.6% in 2023 as base rates increase, and latent overstatement peaks at 4.6% of NAV in 2022, when spreads widen and anchored marks lag. L3-DQI rises from 0.54 to 0.70 in steps that coincide with the ASU 2018-13 and Rule 2a-5 dates (Figure 2a).

Table 3 Descriptive Statistics by Fiscal Year (Means)

Year	L3 share	DQI	yYA (%)	NaR95 (%)	Gap (%)	OS (%)	σ* (%)	P/NAV
2016	0.89	0.54	7.5	8.5	−4.7	1.5	4.0	0.93
2017	0.89	0.54	7.6	8.2	−4.7	0.8	3.9	0.96
2018	0.89	0.54	9.3	8.7	−3.6	1.8	4.1	0.87
2019	0.89	0.54	8.1	9.4	−4.3	0.9	4.4	0.98
2020	0.89	0.63	6.8	8.9	1.1	2.1	4.8	0.87
2021	0.90	0.63	5.8	9.1	0.2	0.3	4.8	1.04
2022	0.89	0.70	11.5	9.3	2.0	4.6	5.4	0.91
2023	0.88	0.70	11.6	10.2	2.0	4.2	5.4	0.98

Note: yYA is the weighted-average market yield of yield-analysis rows; NaR95 and Gap are from observed disclosures; OS and σ* are latent (% of NAV).

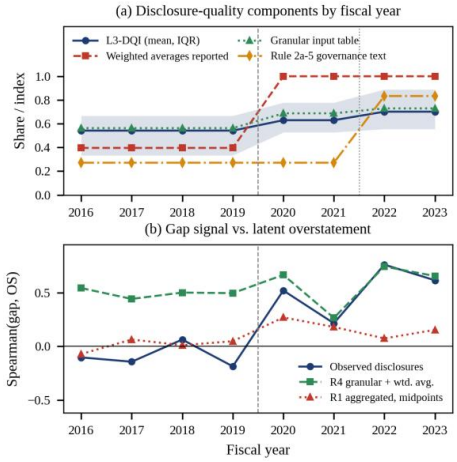


Figure 2 Disclosure Quality over Time (a) and Within-Year Spearman Correlation Between the Benchmark Gap and Latent Overstatement (b)

Note: Dashed line: first fiscal year with mandatory weighted averages; dotted line: first year after the Rule 2a-5 compliance date.

5.2 Recoverability of Valuation Risk (H1, H2)

Table 4 and Figure 3 report recoverability. Composition alone explains 0.71 of the out-of-fund variance of latent dispersion σ*. Adding DIVR from granular tables raises this to 0.81 (R3) and 0.80 (R4), and DIVR alone reaches 0.70–0.73, on a par with composition. DIVR from aggregated tables adds nothing (0.71–0.71). The increment comes from the debt component: $DIVR_D$ from granular tables has a within-year rank correlation of 0.51 with funds' borrower-risk orientation, against 0.18–0.21 under aggregation, because aggregated rows mix senior and junior yields with impaired positions. Range calibration raises the rank correlation between NaR95 and σ* from 0.731 to 0.755. For reference, an oracle that observes borrower risk directly reaches 0.92.

Table 4 Recovery of Latent Dispersion (σ*) and Overstatement (OS)

Disclosure regime	Risk ρs	Risk R^2 alone	Risk R^2 +comp.	Bias ρs	Bias R^2 +comp.
R0 Composition only	0.39	0.71	0.71	0.24	0.44
R1 Aggregated, midpoints	0.41	0.55	0.71	0.09	0.44
R2 Aggregated, wtd. avg.	0.42	0.54	0.71	0.42	0.53
R3 Granular, midpoints	0.77	0.73	0.81	0.31	0.45
R4 Granular, wtd. avg.	0.75	0.70	0.80	0.54	0.59
Observed (mixed)	0.58	0.56	0.74	0.22	0.43

Note: ρs is the mean within-year Spearman correlation of NaR95 (risk) or Gap (bias) with the latent target; R^2 is out-of-fund (fund-grouped six-fold) with year effects. R0 rows use composition variables only.

For overstatement, weighted averages are decisive. With range midpoints the gap's rank correlation with OS is 0.09 (R1) and 0.31 (R3); with weighted averages it is 0.42 (R2) and 0.54 (R4), and the out-of-fund R^2 rises from 0.44 with composition to 0.59 with the R4 gap. In observed disclosures the correlation lies between −0.19 and 0.06 before 2020 and equals 0.52, 0.22, 0.76 and 0.62 in 2020–2023 (Figure 2b): the weighted-average mandate turned the input table into a usable signal of mark staleness. H1 and H2 are supported.

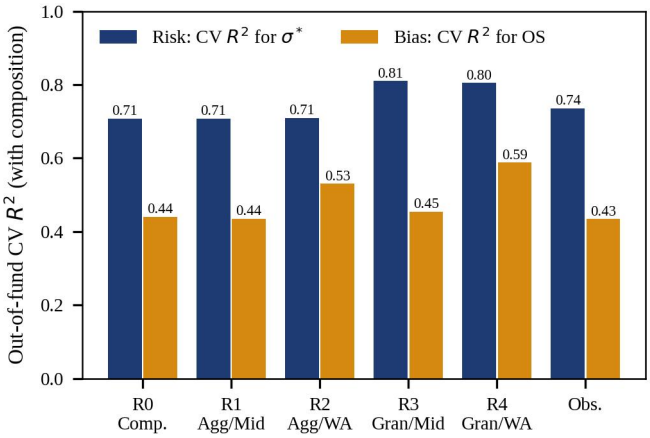


Figure 3 Out-Of-Fund R^2 for Latent Dispersion (risk) and Overstatement (bias) by Disclosure Regime, with Composition Variables Included

5.3 Pricing of Disclosure-Implied Risk and Quality (H3)

Table 5 Price-to-NAV Regressions and Double Machine Learning

	(1)	(2)	(3)	(4)	DML
Level 3 share	−0.013*	−0.005	−0.009	−0.006	
	(0.007)	(0.006)	(0.006)	(0.006)	
DIVR NaR95		−0.033***	−0.027***	−0.030***	−0.028***
		(0.007)	(0.007)	(0.008)	(0.009)
DIVR gap		−0.001	−0.006	−0.005	−0.003
		(0.009)	(0.009)	(0.009)	(0.007)
Gap × Post2020				−0.041*	
				(0.022)	
L3-DQI			0.022***	0.021**	0.019**
			(0.008)	(0.008)	(0.008)
Year FE, controls	Yes	Yes	Yes	Yes	Yes
R^2	0.750	0.784	0.796	0.799	–

Note: Dependent variable ln(P/NAV); 384 fund-years. Regressors standardized; fund-clustered standard errors in parentheses. Controls: NII return, size, leverage, internal management, fee rate, non-accruals, equity and junior-debt shares. DML: partially linear model, LightGBM nuisances, one treatment per row. ***, **, * denote p < 0.01, 0.05, 0.10.

Table 5 reports the pricing tests. The Level 3 share alone has a coefficient of −0.013 per standard deviation (p = 0.063). NaR95 enters at −0.033 (p < 0.001) and raises R^2 from 0.750 to 0.784. L3-DQI enters at 0.022 (p = 0.009) with NaR95 retained. Over the full period the gap is not priced, but its interaction with the post-2020 period is −0.041 (p = 0.060); estimated separately, the gap coefficient is −0.067 after 2020 (p = 0.008) and 0.002 before (p = 0.864). Investors price the bias signal only once weighted averages make it informative.

Double machine learning confirms the estimates without imposing linear controls: 0.019 for L3-DQI (s.e. 0.008, p = 0.020) and −0.028 for NaR95 (p = 0.004). Two checks support the design. When the information-risk channel is switched off in the pricing equation, the L3-DQI estimate falls to 0.011 (OLS, p = 0.18) and 0.009 (DML, p = 0.285), so the index does not pick up spurious correlation through governance. A regression on the latent drivers recovers −1.36 (s.e. 0.26) for OS, −2.53 (0.51) for σ* and −0.14 (0.03) for $1 - Q$, consistent with the structural values. The interaction of L3-DQI and NaR95 is 0.003 (p = 0.697). H3 is supported for disclosure-implied risk and for the post-mandate gap; Section 5.6 shows that the L3-DQI effect, while positive on average, is weakly identified in a single panel.

5.4 Valuation Corrections and Prediction Intervals

Table 6 Rank Validity for Next-Year Valuation Corrections

Measure at year t	ρs 16–21	AUC 16–21	ρs 2022	AUC 2022
Level 3 share	0.02	0.51	0.17	0.54
DIVR NaR95	0.07	0.57	0.04	0.59
DIVR gap (observed)	0.00	0.51	0.45	0.69
DIVR gap (R4 counterfactual)	0.15	0.53	0.40	0.68
DIVR gap (R1 counterfactual)	−0.01	0.49	0.12	0.62
Price-to-NAV discount	0.18	0.57	0.23	0.71
Latent OS (upper bound)	0.32	0.64	0.73	0.88

Note: Within-year Spearman correlation with RCi,t+1 and AUC for the top quartile of corrections. Latent OS is an infeasible upper bound.

Table 6 evaluates measures as forecasts of next-year corrections. In normal years corrections are dominated by noise and weakly predictable even from latent overstatement. In 2022, when spreads widened and anchored marks lagged, the observed gap reaches a rank correlation of 0.45 and an AUC of 0.69, close to the full-disclosure counterfactual (0.68) and above aggregated midpoints (0.62) and the Level 3 share (0.54); the market discount itself reaches 0.71. Forecasters trained on earlier years reach an AUC of at most 0.52, because the informativeness of the gap changed with the 2020 mandate; structural use of disclosed inputs is preferable to models fitted on the pre-mandate history.

Figure 4 shows the prediction intervals. In 2019–2021 all procedures attain coverage of 0.90–0.91 against the 0.90 target. In 2022 coverage falls to 0.65 for split conformal prediction and 0.69 for the DIVR-adaptive versions, since no calibration year contains a comparable level shift. Across NaR95 terciles, DIVR-adaptive intervals equalize coverage (0.84–0.86, against 0.81–0.84 for split conformal prediction) by widening intervals for high-risk funds, and the weighted version attains the same balance with intervals of 5.9% of NAV instead of 6.1%. Averaged over the replication panels, 2022 coverage is 0.69 for split, 0.72 for DIVR-adaptive and 0.74 for weighted intervals, and the coverage gap across terciles falls from 0.15 to 0.09 and 0.08.

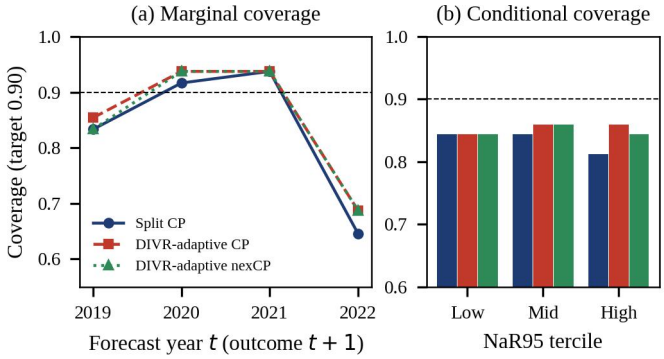


Figure 4 90% Conformal Prediction Intervals for Next-Year Valuation Corrections: Marginal Coverage by Forecast Year (a) and Coverage by NaR95 Tercile (b)

5.5 Audit Allocation (H4)

Figure 5 plots the share of absolute misstatement captured as Level 3 fair value is tested. Testing the largest positions first captures 11% and 22% of misstatement at 10% and 20% coverage, barely above random selection. A staleness heuristic that compares each position's reported spread or multiple change with the market change since origination captures 23% and 36%, and the GBDT back-testing score captures 29% and 44%, close to the oracle (31% and 48%). At 30% coverage the score captures 55%. H4 is supported: the information that Rule 2a-5 back-testing generates is directly usable for risk-based sample selection under AS 2501.

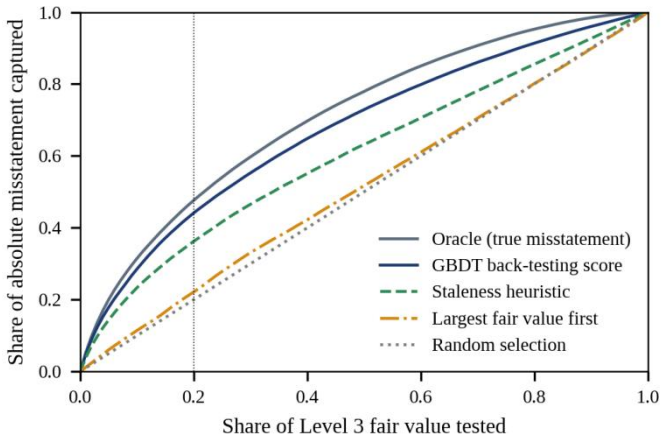


Figure 5 Share of Absolute Level 3 Misstatement Captured as Fair Value Is Tested in Score Order, Averaged over Fund-Years

5.6 Robustness Across Independent Panels

Table 7 repeats the full pipeline on 10 independently simulated panels. Recoverability is stable: adding R4 DIVR to composition raises the out-of-fund R^2 for σ* in every panel (mean 0.85 versus 0.71), the R4 gap dominates the R1 gap in every panel (0.49 versus 0.08), the observed gap is more informative after 2020 than before in every panel, and range calibration improves the rank correlation in 9 of 10 panels. Pricing effects keep their sign but vary in significance. Under DML the NaR95 coefficient is negative in 10 of 10 panels (mean −0.022, across-panel t = −5.8), and the gap × post-2020 interaction is negative in 10 of 10 panels (mean −0.038, t = −6.4) and significant at 5% in 5 of 10. The L3-DQI coefficient is positive on average (0.008 under DML, t = 2.2) and positive in 7 of 10 panels, but significant at 5% in 1 of 10; with the information-risk channel switched off its mean is 0.000 and its rejection rate 0.10. A market of 48 listed BDCs observed over eight years thus identifies the pricing of disclosure-implied risk reliably, whereas the price effect of disclosure quality itself is small relative to price noise and calls for pooled or longer samples. The 2022 AUC of the observed gap exceeds that of the Level 3 share in every panel. Weighted conformal intervals raise mean 2022 coverage from 0.69 to 0.74, and DIVR-adaptive scaling narrows the coverage gap across NaR95 terciles from 0.15 to 0.09 in 9 of 10 panels. GBDT selection beats size-based and heuristic selection in every panel (mean capture at 20% of fair value: 44.8% versus 22.8% and 36.4%).

Table 7 Key Results Across 10 Independent Panels

Statistic	Main	Mean (SD)	Sign / order	p < 0.05
Risk R^2, R0 composition	0.71	0.71 (0.03)	–	–
Risk R^2, R4 + composition	0.80	0.85 (0.02)	1.00	–
Bias ρs, R1	0.09	0.08 (0.06)	–	–
Bias ρs, R4	0.54	0.49 (0.06)	1.00	–
Bias ρs, observed 2016–19	−0.09	0.14 (0.12)	–	–
Bias ρs, observed 2020–23	0.53	0.47 (0.07)	1.00	–
NaR95, OLS (3)	−0.027	−0.016 (0.011)	0.90	0.50
NaR95, DML	−0.028	−0.022 (0.012)	1.00	0.60
Gap × Post2020, OLS (4)	−0.041	−0.038 (0.019)	1.00	0.50
L3-DQI, OLS (3)	0.022	0.011 (0.013)	0.70	0.20
L3-DQI, DML	0.019	0.008 (0.012)	0.70	0.10
L3-DQI, DML, null DGP	0.009	0.000 (0.010)	0.50	0.10
AUC 2022, observed gap	0.69	0.70 (0.11)	1.00	–
Coverage 2022, split CP	0.65	0.69 (0.07)	–	–
Coverage 2022, weighted CP	0.69	0.74 (0.06)	0.70	–
Tercile coverage gap, split	0.03	0.15 (0.05)	–	–
Tercile coverage gap, adaptive	0.02	0.09 (0.05)	0.90	–
Capture@20%, size-based (%)	22	22.8 (0.4)	–	–
Capture@20%, GBDT (%)	44	44.8 (0.9)	1.00	–

Note: Mean (SD) over 10 independently simulated panels. Sign / order: share of panels with the expected sign (coefficients) or ordering (R4 > R0, R4 > R1, post > pre, gap AUC > Level 3 share AUC, weighted > split, adaptive < split, GBDT > size). p < 0.05: share of panels significant at 5% with the expected sign.

6 DISCUSSION AND IMPLICATIONS

For investors, the results imply that Level 3 opacity should be measured from the input table rather than the hierarchy share. A DIVR computation needs only the published ranges, weighted averages and fair values, together with the number of holdings, and it separates two dimensions that a single Level 3 share conflates: dispersion, which granular tables reveal, and staleness, which weighted averages reveal. The finding that the gap is priced only after the weighted-average mandate

is consistent with the incomplete revelation hypothesis [23]: information that is costly to extract is impounded once its extraction becomes reliable.

For auditors, the input table and Rule 2a-5 back-testing records provide a ranking device for AS 2501 procedures. A score trained on differences between marks and subsequent exit prices doubles the misstatement captured by a sample of fixed size, which addresses the concern that fair value audits rely on verifying management's numbers rather than challenging them [9-10].

For standard setters, the counterfactual regimes isolate the value of each element. Weighted averages, mandated by ASU 2018-13, carry the bias signal; granularity by asset class, which remains a preparer choice, carries the dispersion signal; and disclosing the number of holdings per row would allow ranges to be calibrated exactly rather than approximately. Requiring impaired positions to be shown in separate rows would further prevent implied yields from contaminating yield ranges.

For empirical researchers, the replications indicate where real-data tests have power. Recoverability and audit-selection effects are large and stable, and the pricing of disclosure-implied risk is detected in most single panels, whereas the price effect of disclosure quality is small relative to price noise in a market of about fifty listed BDCs; pooling quarterly observations or adding NAV-based outcomes for non-traded BDCs increases the effective sample for such tests.

7 CONCLUSION

This paper develops a disclosure-implied valuation risk framework for the Level 3 investments of BDCs and a nine-item index of Level 3 disclosure quality, and evaluates them on a disclosure-calibrated panel with known ground truth. Granular input tables convey valuation dispersion beyond composition, weighted averages convey overstatement, disclosure-implied risk and the post-mandate gap are reflected in price-to-NAV, and back-testing information sharpens audit selection. The framework applies directly to the inline XBRL financial statements that BDCs now file, and a natural next step is to estimate DIVR on the full population of BDC 10-K filings and relate it to realized NAV revisions and auditors' critical audit matters.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Chernenko S, Erel I, Prilmeier R. Why do firms borrow directly from nonbanks? Review of Financial Studies, 2022, 35(11): 4902-4947.

[2] Financial Accounting Standards Board. Fair value measurement (Topic 820): Disclosure framework—Changes to the disclosure requirements for fair value measurement. Accounting Standards Update No. 2018-13. Norwalk: FASB, 2018.

[3] U.S. Securities and Exchange Commission. Good faith determinations of fair value. Release No. IC-34128. Washington: SEC, 2020.

[4] Public Company Accounting Oversight Board. Auditing accounting estimates, including fair value measurements, and amendments to PCAOB auditing standards. PCAOB Release No. 2018-005. Washington: PCAOB, 2018.

[5] Song C J, Thomas W B, Yi H. Value relevance of FAS No. 157 fair value hierarchy information and the impact of corporate governance mechanisms. The Accounting Review, 2010, 85(4): 1375-1410.

[6] Riedl E J, Serafeim G. Information risk and fair values: An examination of equity betas. Journal of Accounting Research, 2011, 49(4): 1083-1122.

[7] Hanley K W, Jagolinzer A D, Nikolova S. Strategic estimation of asset fair values. Journal of Accounting and Economics, 2018, 66(1): 25-45.

[8] Christensen B E, Glover S M, Wood D A. Extreme estimation uncertainty in fair value estimates: Implications for audit assurance. Auditing: A Journal of Practice & Theory, 2012, 31(1): 127-146.

[9] Cannon N H, Bedard J C. Auditing challenging fair value measurements: Evidence from the field. The Accounting Review, 2017, 92(4): 81-114.

[10] Glover S M, Taylor M H, Wu Y J. Current practices and challenges in auditing fair value measurements and complex estimates: Implications for auditing standards and the academy. Auditing: A Journal of Practice & Theory, 2017, 36(1): 63-84.

[11] Chernozhukov V, Chetverikov D, Demirer M, et al. Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 2018, 21(1): C1-C68.

[12] Bach P, Chernozhukov V, Kurz M S, et al. DoubleML—An object-oriented implementation of double machine learning in Python. Journal of Machine Learning Research, 2022, 23(53): 1-6.

[13] Barber R F, Candès E J, Ramdas A, et al. Conformal prediction beyond exchangeability. The Annals of Statistics, 2023, 51(2): 816-845.

[14] Angelopoulos A N, Bates S. Conformal prediction: A gentle introduction. Foundations and Trends in Machine Learning, 2023, 16(4): 494-591.

[15] Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deep learning on typical tabular data? In: Advances in Neural Information Processing Systems 35 (NeurIPS 2022), Datasets and Benchmarks Track, 2022.
[16] Ke G, Meng Q, Finley T, et al. LightGBM: A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems 30, 2017: 3146-3154.
[17] Lambert R, Leuz C, Verrecchia R E. Accounting information, disclosure, and the cost of capital. Journal of Accounting Research, 2007, 45(2): 385-420.
[18] Easley D, O'Hara M. Information and the cost of capital. The Journal of Finance, 2004, 59(4): 1553-1583.
[19] Laux C, Leuz C. The crisis of fair-value accounting: Making sense of the recent debate. Accounting, Organizations and Society, 2009, 34(6-7): 826-834.
[20] Bens D A, Cheng M, Neamtiu M. The impact of SEC disclosure monitoring on the uncertainty of fair value estimates. The Accounting Review, 2016, 91(2): 349-375.
[21] Chung S G, Goh B W, Ng J, et al. Voluntary fair value disclosures beyond SFAS 157's three-level estimates. Review of Accounting Studies, 2017, 22(1): 430-468.
[22] Hope O K, Hu D, Lu H. The benefits of specific risk-factor disclosures. Review of Accounting Studies, 2016, 21(4): 1005-1045.
[23] Bloomfield R J. The 'incomplete revelation hypothesis' and financial reporting. Accounting Horizons, 2002, 16(3): 233-243.
[24] Cassar G, Gerakos J. Hedge funds: Pricing controls and the smoothing of self-reported returns. Review of Financial Studies, 2011, 24(5): 1698-1734.
[25] Brown G W, Gredil O R, Kaplan S N. Do private equity funds manipulate reported returns? Journal of Financial Economics, 2019, 132(2): 267-297.
[26] Getmansky M, Lo A W, Makarov I. An econometric model of serial correlation and illiquidity in hedge fund returns. Journal of Financial Economics, 2004, 74(3): 529-609.
[27] Lee C M C, Shleifer A, Thaler R H. Investor sentiment and the closed-end fund puzzle. The Journal of Finance, 1991, 46(1): 75-109.
[28] Berk J B, Stanton R. Managerial ability, compensation, and the closed-end fund discount. The Journal of Finance, 2007, 62(2): 529-556.
[29] Ahn J, Hoitash R, Hoitash U. Auditor task-specific expertise: The case of fair value accounting. The Accounting Review, 2020, 95(3): 1-32.
[30] Malcolm D G, Roseboom J H, Clark C E, et al. Application of a technique for research and development program evaluation. Operations Research, 1959, 7(5): 646-669.