Revista de Estudios Empresariales. Segunda época.
Número: 1 (2027). Páginas: 1-30 DOI: https://doi.org/10.17561/ree.n1.2027.9888

A HYBRID FORWARD-INTENSITY MODEL FOR PRIVATE FIRM DEFAULT. AN EMPIRICAL TEST OF MINSKY'S TAXONOMY IN AN EMERGING MARKET

UN MODELO HÍBRIDO DE INTENSIDAD FORWARD PARA EL INCUMPLIMIENTO DE EMPRESAS PRIVADAS. UNA PRUEBA EMPÍRICA DE LA TAXONOMÍA DE MINSKY EN UN MERCADO EMERGENTE

Bruno de Jesus Rahmer (Fundación Universitaria Tecnológico Comfenalco, Cartagena de Indias, Colombia)[1]*

Abstract

In financial risk management, predicting default for privately held firms is a critical area of study, and more concretely in emerging markets where comprehensive market data are often unavailable. Traditional accounting-based models are backward-looking and lack a theoretical foundation for how financial fragility develops, while market-based models are, in certain circumstances, infeasible to estimate. To address this analytical gap, this paper develops a hybrid forward-intensity model to estimate default probabilities for privately held firms and to construct a firm-level classification grounded in the Minskyan taxonomy of financial positions. An implied distance-to-default metric for non-listed firms is constructed through a cross-sectional mapping from publicly traded firms that matches accounting fundamentals with market-based risk measures. From this metric, multi-horizon forward default intensities follow the approach of Duan *et al*. (2012) and incorporate firm-level balance sheet variables together with aggregate financial conditions. Firms are then classified endogenously into Hedge, Speculative, and Ponzi regimes through hierarchical clustering. The empirical application draws on a panel of 9,346 Colombian privately held manufacturing firms spanning Q1 2009 to Q1 2018, with annual financial statements and monthly default intensity estimation. A parameter transfer procedure based on propensity score matching and temporal out-of-sample validation matches public-firm parameters to comparable private firms. The hybrid model achieves higher out-of-sample discrimination accuracy than benchmark alternatives. Credit spreads charged by commercial lenders increase monotonically across Hedge, Speculative, and Ponzi clusters. Maturity mismatch is the dominant predictor at medium and long horizons, while liquidity retains predictive power at short horizons.

Keywords: credit risk, financial fragility, private firms, default prediction, forward intensity model.

JEL Codes: C25, G32, G33, O16

[1] ORCID: https://orcid.org/0000-0003-1925-0432
* Correspondence author: brunodejesus.2509@gmail.com

Date received: 06/09/2025. Date accepted: 27/07/2026

CC BY

Resumen

Predecir el incumplimiento de las empresas privadas es crítico en la gestión del riesgo financiero, especialmente en mercados emergentes, donde los datos de mercado suelen ser incompletos. Los modelos contables son retrospectivos y carecen de fundamento teórico sobre el origen de la fragilidad financiera, mientras que los modelos de mercado suelen ser inviables de estimar. Este trabajo desarrolla un modelo híbrido de *forward intensity* que estima probabilidades de incumplimiento de empresas privadas y las clasifica según la taxonomía minskyana de posiciones financieras. Una métrica de *distance-to-default* implícita para empresas no cotizadas, obtenida mediante un mapeo transversal desde empresas que cotizan en bolsa, alimenta las *forward default intensities* en múltiples horizontes, siguiendo a Duan *et al*. (2012), con variables de balance y condiciones financieras agregadas. Mediante *hierarchical clustering*, las firmas se clasifican de forma endógena en los regímenes Hedge, Speculative y Ponzi. La aplicación emplea un panel de 9.346 empresas manufactureras colombianas privadas entre el primer trimestre de 2009 y el primer trimestre de 2018, con transferencia de parámetros desde empresas cotizadas mediante *propensity score matching* y validación temporal fuera de muestra. El modelo híbrido alcanza una capacidad discriminatoria fuera de muestra superior a la de los modelos de referencia. Los diferenciales de crédito aumentan de forma monótona entre los grupos Hedge, Speculative y Ponzi. El descalce de vencimientos es el predictor dominante en los horizontes medio y largo, mientras que la liquidez conserva poder predictivo en el corto plazo.

Palabras clave: riesgo de crédito, fragilidad financiera, empresas privadas, predicción de incumplimiento, *forward intensity model.*

Códigos JEL: C25, G32, G33, O16

1. INTRODUCTION

The analysis of corporate default probability represents a central discipline within financial risk management, yet a formidable methodological impediment arises for the preponderance of firms globally, namely privately held corporations. These entities, devoid of observable equity valuations and market-derived metrics, preclude the application of canonical market-based credit risk models, such as the Merton (1974) formulation. Credit analysis for private firms has historically employed accounting-based statistical models, which, being retrospective, fail to incorporate the dynamic, forward-looking risk perceptions fundamental to robust analysis. The ramifications of this analytical deficiency are especially problematic, as opaque private small and medium-sized enterprises (SMEs) are the foundation of economic activity and the core driver of employment and gross domestic product (Ayyagari *et al.*, 2007). The insolvency of these entities therefore carries systemic macroeconomic consequences rather than remaining an idiosyncratic event, with defaults in this segment associated with bank portfolio stress and regional credit contractions. Methods such as logistic regression or linear discriminant analysis often function as "black boxes" that link a set of financial ratios (such as leverage, liquidity, or profitability) directly to a default outcome.

A methodological shortcoming of a myriad of standard credit risk formulations is their treatment of default probability as a function of exogenous shocks to an otherwise equilibrated system. As a consequence, they fail to account for the endogenous generation of financial fragility. This conceptual lacuna motivates the development of a systematic classification of financial vulnerability; without a formal taxonomy, analysts treat each corporate failure as an idiosyncratic event, which masks systematic causal factors. The theoretical apparatus of Hyman

Minsky addresses this deficiency and his Financial Instability Hypothesis sets out three discrete balance-sheet postures amenable to empirical operationalization at the firm level.

In the Financial Instability Hypothesis, Minsky (1986) states that, for all economic units, whether firms or individuals, there exist various discrete income-debt relationships that explain the genesis of financial crises. Instead of an equilibrium-seeking and self-contained economic system (as neoclassical economists propound), Minsky explains that we are more likely to witness a "deviation-amplifying system", whereby borrowers and lenders gravitate towards more perilous activities in pursuit of profits. Minsky labels these as "speculative financing entities", defined by their cash flows proving sufficient merely for the discharge of mandatory interest payments.

Minsky argued that an economy dominated by speculative and Ponzi units is inherently unstable, as any shock to credit availability or interest rates can trigger a cascade of defaults. A core feature of this fragility is maturity mismatch, which originates in the financing of long-duration assets with short-duration liabilities and thereby creates severe rollover risk. This emphasis on maturity mismatch and sensitivity to credit shocks is salient for the private sector. Privately held corporations and SMEs, by virtue of their restricted access to public capital markets and their consequential dependence on bank-intermediated, short-duration credit facilities, frequently conform to the archetypes of Minsky's speculative and Ponzi financing units.

The application of Minsky's models and quantitative credit risk formulations, such as the Merton model, within the context of emerging economies is a specific area of inquiry that remains underdeveloped in the contemporary financial literature. The preponderance of extant scholarship concentrates principally upon advanced economies, where mature capital markets, robust data infrastructure, and stable institutional settings provide fertile ground for empirical investigation (Bharath & Shumway, 2008; Duffie *et al*., 2007). The Minskyan analytical tradition, while extensively debated in theoretical terms (Keen, 1995; Whalen, 2001), has been operationalized empirically for private firms in emerging markets in limited circumstances. The direct transposition of these theoretical constructs to emerging economies without adaptation poses methodological complications. For instance, the Colombian economy has a distinct financial structure defined by bank-centric credit allocation, severe macroeconomic volatility, and informational frictions.

The prediction of corporate default has evolved through successive methodological generations, each confronting distinct limitations when applied to privately held firms. Early accounting-based discriminant models, exemplified by Altman (1968) and Ohlson (1980), achieved considerable classification accuracy but functioned as atheoretical statistical mappings without economic foundations for the temporal dynamics of financial distress. Hazard-rate models accommodated time-varying covariates and firm exit dynamics (Shumway, 2001; Chava & Jarrow, 2004), yet their adaptation to small and medium enterprises confronted severe data scarcity in emerging markets where disclosure standards remain weak (Altman & Sabato, 2007). Balcaen and Ooghe (2006) documented persistent methodological deficiencies across these approaches such as sample selection bias, stability of coefficient estimates across economic cycles, and limited out-of-sample validation. Recent systematic reviews find that SME default prediction encounters constraints that do not arise in public-firm settings, notably the absence of observable market valuations, ownership concentration that alters incentive structures and governance dynamics, and a financing structure grounded in relationship-based bank lending rather than arm's-length debt issuance (Ciampi *et al.*, 2021; Cheraghali & Molnár, 2024).

A parallel strand of research explored machine learning models to accommodate nonlinear feature interactions and high-dimensional covariate spaces. Comparative studies record superior

CC BY

classification metrics for ensemble methods (random forests, gradient boosting) and neural networks relative to traditional logistic specifications when applied to SME credit datasets (Fernández-Delgado *et al.*, 2014; Lessmann *et al.*, 2015). Recent implementations employ dynamic self-organizing feature maps (du Jardin, 2021). Despite these advances in predictive accuracy, such approaches sacrifice economic interpretability and supervisory transparency. The absence of a mapping between balance-sheet fundamentals and insolvency mechanisms precludes causal inference and limits applicability to stress-testing exercises that need scenario analysis under counterfactual macroeconomic conditions. Rashid *et al.* (2024) identify this interpretability deficit as a principal barrier to regulatory adoption of machine learning credit models, particularly in jurisdictions where Basel III capital requirements mandate theoretically grounded risk-weight functions.

The scarcity of empirical studies for emerging market contexts obstructs a comprehensive appraisal of global financial stability. This circumstance provides the impetus for new lines of investigation and for the construction of context-specific analytical techniques that can systematically map corporate default probabilities and systemic fragility within these pivotal economies.

This study addresses this methodological constraint by formulating and empirically validating an original hybrid methodology to construct a term structure of default probabilities for a sample of privately held Colombian manufacturing corporations. This approach synthesizes three distinct research streams, with the core specification drawing from the contributions of Duan *et al.* (2012). Methodologically, the projection of distance-to-default (DTD) metrics using public-firm regressions finds a corollary in Chan-Lau *et al.* (2024), who utilize transfer-learning ensembles to extrapolate public-firm information to the private-firm domain. The Chan-Lau *et al.* (2024) procedure, however, depends on a machine-learning infrastructure that needs large training datasets and limits interpretability and replicability in computationally constrained supervisory settings. This approach instead relies on transparent parametric projection, which preserves economic interpretability and is directly applicable to regulatory stress-testing exercises.

A proxy Distance-to-Default is imputed for private entities through a mapping between accounting fundamentals and the market-derived DTDs of a peer group of public corporations. This imputed DTD then enters as a principal covariate within a forward-intensity model, consistent with the specification of Duan *et al.* (2012), an econometric technique that estimates a complete term structure of default probabilities across multiple temporal horizons. The resulting risk measures classify firms according to Minsky's taxonomy of financing positions, which permits an empirical operationalization of the Financial Instability Hypothesis through an examination of the predictive determinants of insolvency for the most financially fragile cohorts.

The analytical structure facilitates an empirical operationalization of the central tenets of Minsky's hypothesis. To apply Minsky's taxonomy while avoiding the imposition of arbitrary thresholds, the analysis employs hierarchical clustering to partition the sample firms into discrete risk-based cohorts. This unsupervised learning algorithm permits the endogenous identification of statistical clusters, which the study subsequently maps onto Minsky's Hedge, Speculative, and Ponzi schemes. This procedure examines financial fragility as a predictive determinant of corporate insolvency.

For the purpose of model validation and to determine its incremental predictive content, the research constructs a sequence of additional empirical procedures that extend the baseline specifications. These procedures comprise a temporal out-of-sample evaluation, a balance-sheet transfer protocol between public and private firms through propensity score matching, and a set of benchmark estimations drawn from standard default prediction models.

The study's primary scholarly contribution resides in its synthesis of these distinct research streams. It conjoins Minsky's classification of financial fragility with an advanced forward-intensity

CC BY

model, with a calibration specific to an emerging-market context and an application to private manufacturing corporations. This novel specification supersedes conventional models predicated exclusively on accounting information. The imputation of a public-equivalent Distance-to-Default metric for private firms, when combined with financial and macroeconomic covariates, produces dependable multi-horizon default probability estimates.

2. DATA AND METHODOLOGY

The empirical investigation draws upon a proprietary panel dataset which includes financial statement variables and default records for 9,346 unlisted Colombian manufacturing corporations over the temporal scope of Q1 2009 through Q1 2018. The data originate from mandatory financial disclosures submitted to the Colombian Superintendence of Societies. The analysis defines a default event as a corporation's initiation of formal insolvency proceedings under statutory provision Ley 1116, or its classification by a principal financial intermediary as having obligations with arrearages exceeding 90 days. A supplementary dataset of all non-financial corporations listed on the Colombian stock market (BVC) with concurrent market and accounting data provides the requisite inputs for the primary imputation procedure.

2.1. Parameter transfer and imputation of Distance-to-Default

The non-observability of equity capitalizations for unlisted entities precludes the direct calculation of the asset volatility parameter in the Merton (1974) formulation. The methodology constructs a Public-Firm Equivalent Distance-to-Default under the axiom that, conditional on comparable accounting fundamentals, public and private obligors share a functional mapping to creditworthiness. Selection bias induced by the listing decision contaminates this mapping because listing correlates systematically with size, opacity, and capital structure. To restore counterfactual comparability, the analysis applies propensity score matching using a nearest-neighbor algorithm with a caliper of 0.05, which identifies the closest public-firm counterpart for each private firm in the propensity score space.

The propensity score is estimated from a first-stage logistic classification model:

$$p_i = Pr\,(D_i = 1 \mid X_i) \tag{1}$$

where D_i indexes public firm status and X_i collects the balance-sheet covariates listed above. Each private firm is matched to the nearest public firm in propensity score space subject to a caliper of 0.05, ensuring that only comparably positioned firms enter the parameter transfer. On the matched public-firm sample, the analysis estimates a sequence of twelve discrete OLS regressions indexed by the informational lag m ∈ {0, 1, …, 11} relative to the fiscal closure:

$$DTD_{i,t}^{pub} = \alpha_m + \beta_m X_{i,t} + \gamma_m M_t + \varepsilon_{i,t} \tag{2}$$

where $DTD_{i,t}^{pub}$ denotes the Distance-to-Default of public firm i at time t, derived via the iterative inversion of the Black-Scholes-Merton equations. The vector $X_{i,t}$ comprises log-transformed accounting covariates selected through the propensity-score diagnostics, namely profitability, equity coverage, leverage, operational efficiency, and interest coverage, and the macroeconomic vector M_t captures the COP/USD exchange rate. Parameter identification employs the HC3 covariance estimator to neutralize heteroskedasticity and projection of the estimated β_m onto the private-firm accounting set generates the imputed metric $DTD_{i,t}^{priv}$, indexed by informational staleness m.

CC BY

2.2. Forward default intensity model

The imputed $DTD_{i,t}^{priv}$ enters as the cardinal input of a discrete-time forward-intensity specification (Duan *et al.*, 2012) that treats default as the first jump time τD of a doubly-stochastic Poisson process. The forward default intensity $f_i(t,\tau)$ represents the conditional instantaneous probability of failure at future time t + τ given the filtration at t and assumes an exponential affine form:

$$f_i(t,\tau) = exp\{\theta_0(\tau) + \theta_1(\tau)DTD_{i,t}^{priv} + \theta_2(\tau)'Z_{i,t}\} \quad (3)$$

where $Z_{i,t}$ integrates firm-specific covariates, including the Minskyan maturity mismatch indicator, defined as current liabilities net of cash over total liabilities, the interest coverage ratio, debt over EBITDA, cash over total assets, and total assets adjusted by the GDP deflator. The vector $Z_{i,t}$ also incorporates macroeconomic covariates (the yield on 90-day certificates of deposit, CDTs, as the Colombian equivalent of the commercial paper rate used by Duan *et al.* (2018), and the COLCAP return). Unlike the DTD components, the maturity-mismatch indicator is directly observable from private-firm statements and enters the intensity as a native covariate rather than an imputed metric. The cumulative physical default probability over horizon τ takes the form:

$$PD_i(t,\tau) = 1 - exp\{-\sum_{s=1}^{\tau} f_i(t,s)\Delta t\} \quad (4)$$

Calibration proceeds through maximization of a horizon-decomposed overlapped pseudo-likelihood that accommodates the serial correlation induced by overlapping prediction windows (Duan *et al.*, 2012):

$$L(\Theta) = \sum_{\tau=1}^{K} \sum_{i=1}^{N} \sum_{t=0}^{T-\tau} w_{i,t} [Y_{i,t+\tau} ln(1 - e^{-f_i(t,\tau)\Delta t}) + (1 - Y_{i,t+\tau}) ln(e^{-f_i(t,\tau)\Delta t})] \quad (5)$$

The binary indicator $Y_{i,t+\tau}$ assumes unity upon default realization within the horizon τ and zero otherwise. The weighting scheme $w_{i,t}$ adjusts for sample attrition and non-default exits. The full martingale description of the doubly-stochastic default process and the formal derivation of the horizon-decomposable structure of $L(\Theta)$are reported in Appendix A.

2.3. Endogenous classification

Empirical operationalization of the Financial Instability Hypothesis avoids ex-ante thresholds on leverage or liquidity ratios. The identification of Minsky's three postures (Hedge, Speculative, Ponzi) proceeds through unsupervised agglomerative hierarchical clustering with Ward's linkage applied to the multidimensional risk vector.

The selection of the optimal number of clusters is guided by the silhouette coefficient, the gap statistic, and the Jaccard stability index under B = 1,000 bootstrap resamples, with the stability threshold of 0.70 proposed by Hennig (2007).

Benchmark models

The hybrid specification is evaluated against three benchmarks commonly employed in credit-risk assessment for private firms. The Altman (1968) Z-score takes the form:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5 \quad (6)$$

CC BY

where X_1 is working capital over total assets, X_2 retained earnings over total assets, X_3 EBIT over total assets, X_4 the book value of equity over total liabilities (substituted for the unobservable market value), and X_5 sales over total assets. The Ohlson (1980) O-score combines nine indicators within a logistic regression:

$$O = -1.32 - 0.407\, ln(TA) + 6.03\, TL/TA - 1.43\, WC/TA + 0.0757\, CL/CA$$
$$- 2.37\, NI/TA - 1.83\, FFO/TL + 0.285\, INT - 1.72\, OENEG$$
$$- 0.521\, \Delta NI/|NI| \quad (7)$$

with the probability of default obtained through the logistic transformation $PD_{Ohlson} = \frac{1}{1 + e^{-O}}$. Njeru (2019) advances a Modified KMV-Merton model for unlisted entities that substitutes book values for market capitalizations and employs the volatility of bank account balances as a proxy for asset volatility:

$$DD_{Njeru} = \frac{In(\frac{V_0}{DPT}) + (\mu - 0.5\sigma^2)T}{\sigma\sqrt{T}} \quad (8)$$

where V_0 is the book value of total assets, DPT the default point defined as short-term debt plus half of long-term debt, μ the expected asset return, and σ the annualized standard deviation of logarithmic changes in daily bank balances normalized by annual turnover.

2.4. Out-of-sample validation and performance metrics

The forward-intensity specification is a predictive model calibrated on observational data. No instrumental variable, difference-in-differences, or regression discontinuity strategy is employed to identify causal effects. Accordingly, the estimated coefficients should be interpreted as predictive associations between observable firm attributes and default intensity, not as causal parameters. In particular, covariates such as the imputed Distance-to-Default, maturity mismatch, leverage, and cash-to-assets ratio are endogenous to the firm's financial condition. A firm approaching distress will simultaneously experience deteriorating DTD, shorter effective liability maturities, higher leverage, and lower liquidity. This predicament renders the direction of any causal relationship ambiguous in an observational setting. Similarly, in the market discipline test, higher credit spreads may precede, rather than follow, exacerbation of financial malaise. The estimation of a term structure of default probabilities for privately held firms in an emerging market and the empirical operationalization of Minsky's taxonomy through forward-looking risk measures define descriptive objectives rather than causal claims. The interpretive approach throughout the paper is coherent with this identification boundary.

The empirical design imposes strict temporal separation between estimation and evaluation. All parameters of the Distance-to-Default imputation and the forward-intensity specification derive from observations over 2009–2014, and predictive assessment relies on default realizations observed over 2015–2018. Predictive accuracy is measured through the Accuracy Ratio and the Brier Score:

$$AR = 2 \times (AUC - 0.5) \qquad BS = \frac{1}{N}\sum_{i=1}^{N} (P_i - Y_i)^2 \quad (9)$$

where P_i denotes the predicted probability of default and Y_i the realized default indicator. Marginal contributions of individual covariates are separated through the decrement in the Accuracy Ratio upon exclusion of each variable k at horizons $\tau \in \{3, 6, 12\}$ months.

2.5. Market discipline test

A final specification examines whether commercial lenders incorporate the estimated default probabilities into the pricing of debt. Following Duan *et al.* (2018), the analysis performs a panel regression of the implicit interest-rate spread against the market-based risk metrics, segmented across the three Minsky clusters:

$$r_{i,t} - r_{f,t} = \delta_0 + \delta_1 PD_{i,t} + \delta_2 Mat_{i,t} + \delta_3 Mat_{i,t}^2 + \delta_4 (Collateral/Debt)_{i,t} + \Phi' Control_t + \mu_i + \varepsilon_{i,t} \quad (10)$$

where $r_{i,t} - r_{f,t}$ denotes the firm-specific risk premium over the risk-free rate, $PD_{i,t}$ the annualized probability of default derived from (4), $Mat_{i,t}$ the value-weighted average maturity of outstanding liabilities, and μ_i firm fixed effects that absorb time-invariant unobserved heterogeneity. The coefficient δ_1 captures the extent to which lenders price the predicted default probability, with values that rise monotonically across Hedge, Speculative, and Ponzi clusters.

3. RESULTS

Propensity score matching procedure seeks to neutralize the compositional heterogeneity inherent between listed and unlisted entities prior to parameter transfer. Table 1 reports sample moments for each covariate in the untreated (public) and treated (private) groups, the standardized mean differences prior to matching, and the post-match balances together with variance ratios. This is the primary diagnostic for whether the nearest-neighbor algorithm achieved distributional comparability needed by the transferability assumption and the table functions as a precondition test for all imputations. If post-match standardized differences remain above conventional thresholds, any inference that projects public-firm parameter estimates onto private firms would suffer selection bias.

The matching procedure expeditiously rectified these distributional imbalances and achieved a nearly perfect isomorphism between the treatment (public) and control (matched private) groups. As detailed in the rightmost columns of Table 1, the post-match SMD values converged below the critical 0.10 threshold across all covariates except "Sales Growth".

TABLE 1. COVARIATE BALANCE DIAGNOSTICS

Covariate	Pre-Match SMD	Post-Match SMD	Variance Ratio
Total Assets (GDP-deflated)	1.28***	0.07	1.18
Return on Assets	0.32**	0.04	1.45
Debt-to-Assets	-0.89***	-0.09	0.88
Asset Turnover	-0.48*	-0.08	1.22
Current Ratio	0.43*	0.06	1.09
Sales Growth (excluded)	-0.29	-0.18	1.54

Note: SMD denotes Standardized Mean Difference. Significance levels for t-tests of mean differences pre-matching: * p < .05, ** p < .01, *** p < .001. Post-match variance ratios approaching 1.0 imply distributional convergence. The propensity score was estimated using a logit model including all listed covariates. Matching was performed using the nearest-neighbor algorithm with a caliper of 0.05.

Total Assets (GDP-deflated) registers a post-match standardized difference that falls below the accepted threshold and a variance ratio near unity and this result conforms with

the matching objective to erase scale-related selection on listing status. Following Duan *et al.* (2018), total assets are employed in level form (adjusted by the Colombian GDP deflator) rather than in logarithmic transformation, since the range of the log-transformed variable proves excessively restrictive for capturing sufficient variation across private firms. Theoretical logic links firm scale to decision to list through fixed costs of public listing and monitoring intensity and appropriately balancing on size reduces the risk that market-implied DTD parameters are confounded by scale. This legitimizes the claim that private firms function as "shadow public firms" for the attributes used in the imputation.

Return on Assets (ROA) shows a variance ratio modestly above one as profitability enters the theoretical model as a flow-based solvency metric that markets price with high frequency. In this case, the profitability-driven market perceptions embedded in the imputed metrics do not differ systematically between the matched private sample and their public peers.

Debt-to-Assets shows a corrected post-match standardized difference brought well within the acceptable band; leverage is a persistent determinant of DTD in the Merton tradition, so balance on leverage reduces parameter transfer error. In this case, capital structure heterogeneity will not drive spurious divergence in imputed solvency metrics.

Asset Turnover attains near-zero post-match imbalance and an acceptable variance ratio because turnover functions as an operational efficiency metric that maps into asset value through earnings capacity.

Current Ratio yields post-match balance within the pre-specified threshold and this is consistent with the fact that liquidity metrics govern short-run survival probabilities in the forward-intensity logic. Matched liquidity reduces measurement error when mapping accounting stocks into forward hazards rates and similarity in current ratios means short-term liquidity will not bias the imputed DTD.

Sales-growth dynamics displays an outcome that contradicts the central purpose of the matching procedure. This attribute therefore falls outside the analytical threshold that would allow it to be treated as comparable and must be interpreted as an active source of distortion rather than a neutral conditioning variable.

Sales Growth shows a pre-match imbalance that the matching procedure fails to absorb as its post-match standardized difference remains the largest and its variance ratio the most uneven of all covariates. Within financial logic, the growth trajectory embeds projections of operational capacity, sensitivity to the business cycle, and expansion prospects for economic value. The two populations face fundamentally different dynamics of scale and volatility. Growth prospects therefore fracture the public-private equivalence and convert this covariate into a contaminant for all inferences drawn from the matched sample.

Nonetheless, the overall picture conveys that these results do not invalidate the entire model and it is empirically testable that the matched private control group mimics the fundamental solvency, liquidity, and growth profile of the listed peers with moderate fidelity. The functional relationship between accounting fundamentals and market-implied default risk observed in the public sphere is transferable to this specific subset of private firms, as they are effectively "shadow public firms" without a market ticker.

Given the failure of the PSM procedure to balance Sales Growth, this variable is excluded from all subsequent parameter transfer procedures.

Table 2 examines how the informational content of accounting fundamentals erodes as the lag between the fiscal closing date and the market-implied Distance-to-Default widens, and it operationalizes a set of cross-sectional OLS regressions stratified by informational lag that condition the mapping on temporal staleness. The table functions as an empirical stress test of

CC BY

TABLE 2. DTD IMPUTATION ESTIMATES ACROSS INFORMATIONAL LAGS

Variable	Horizon m = 0 (3 Months Post-Fiscal)	Horizon m = 5 (8 Months Post-Fiscal)	Horizon m = 11 (14 Months Post-Fiscal)
Intercept	0.119 (0.082)	0.158 (0.094)	0.087 (0.113)
Net Income / Assets	0.342 (0.031)***	0.209 (0.054)***	0.076 (0.072)
Book Equity / Liabilities	0.514 (0.309).	0.443 (0.331)	0.118 (0.384)
Total Liabilities / Assets	-0.718 (0.058)***	-0.742 (0.069)***	-0.653 (0.082)***
Sales / Assets	-0.394 (0.021)***	0.311 (0.034)***	0.182 (0.045)***
EBIT/Interest expenses	0.247 (0.119)*	0.208 (0.126).	0.054 (0.151)
FX Rate (COP/USD)	-0.048 (0.061)	-0.092 (0.063)	-0.029 (0.074)
Adjusted R^2	0.683	0.541	0.389
F-Statistic	142.48***	98.17***	45.09***

Note: Coefficients represent OLS estimates with Heteroskedasticity-Consistent (HC3) standard errors in parentheses. Significance levels: . p < .10, * p < .05, ** p < .01, *** p < .001. The dependent variable is the Market-Based Distance-to-Default. The horizon *m* denotes the months elapsed since the publication of financial statements.

the paper's central premise that public-firm market signals can be reconstructed from private-firm accounting data only when the underlying determinants retain relevance throughout the horizon used for parameter transfer. Each regression separates a distinct informational content and the entire structure acts as a dynamic laboratory for observing how flow-based and stock-based accounting variables lose or retain explanatory power as time separates market valuation from financial reporting.

Net Income / Assets retains explanatory content only when the accounting information is recent. The coefficient is positive and significant at m=0 and at m=5, but it becomes non-significant at m=11 (0.076), so the earnings signal deteriorates once the data are stale. Once the data becomes stale, accrual distortions and interim shocks contaminate the earnings signal, so its importance disappears.

Total Liabilities / Assets enters with a negative and highly significant coefficient across all three horizons (-0.718, -0.742, -0.653, all $p < 0.001$), because capital structure imposes a persistent set of contractual obligations whose relevance for the Merton barrier does not depend on reporting recency. An operational efficiency proxy based on sales relative to assets retains explanatory power across horizons. At the shortest informational lag, the coefficient is negative (-0.394, $p < 0.001$) and reproduces the pattern reported by Duan *et al.* (2018) for Korean non-financial firms at equivalent horizon. In the Colombian public-firm sample (concentrated on holdings, utilities, banking and large consumer-goods corporations) asset turnover correlates with exposure to cyclical segments of lower average solvency. As information staleness increases, the coefficient turns positive and remains significant (0.311 at m=5, 0.182 at m=11, both $p < 0.001$), so higher sales-to-assets are associated with lower implied default probability.

CC BY

Book equity relative to total liabilities exerts a positive and theoretically consistent effect on the Distance-to-Default. A more robust equity cushion relative to outstanding liabilities increases the margin separating asset value from the contractual default barrier, reducing the probability of breaching that threshold. This result is coherent with both the credit-risk model of Merton (1974) and the empirical findings of Duan *et al.* (2018, Table 2, Panel A), who report positive coefficients for this variable across all twelve informational horizons in the Korean private-firm sample. The coefficient remains positive across all horizons but is statistically significant only at m=0 (0.514, $p < 0.10$). In other terms book equity's explanatory power is not reliably detected once informational staleness increases.

EBIT/Interest expenses carries a positive coefficient that is significant at m=0 (0.247, $p < 0.05$), marginally significant at m=5 (0.208, $p < 0.10$), and non-significant at m=11 (0.054), because immediate stress on debt service affects asset option value only when the information mirrors current conditions; once temporal distance widens, this flow-based marker decays rapidly. When interest obligations approach current earnings capacity, the margin for contractual compliance narrows immediately, which directly conditions default risk. This informational content remains relevant only insofar as it corresponds to current business conditions; as the forecast horizon extends, the ratio loses explanatory relevance because future solvency depends less on current earnings coverage and more on balance-sheet adjustment, refinancing capacity, and shifts in the liability structure.

The COP/USD exchange rate, included as a control across all horizons, remains non-significant in every specification as currency fluctuations do not systematically shift the DTD in this sample.

Across all horizons, balance-sheet structures with long persistence such as leverage and equity coverage have consistent explanatory force; short-run flow variables only matter when the informational lag is minimal; and accounting measures without a direct connection to market valuation lose relevance regardless of timing. The overall narrative affirms the model's foundational claim that transferable parameters must be anchored either in "slow-moving attributes" or in recent information tied to the evolution of cash flows.

Table 3 reports maximum pseudo-likelihood estimates for the forward-default intensity model across successive forecast horizons and thus implements the horizon-decomposed calibration described in the methodology. This is the principal empirical device that translates accounting and market-implied variables into a term structure of intensities suitable for clustered firms, instead of imposing temporal homogeneity on the default process.

The imputed Distance-to-Default functions as a short-horizon determinant of default intensity insofar as it synthesizes leverage and asset risk into a unified solvency measure. Firms situated further from the contractual default threshold display a systematically lower incidence of failure within the immediate forecast horizon. This effect weakens monotonically as the horizon extends, a pattern that accords with the erosion of accounting-based solvency information once refinancing conditions and contractual timing assume dominance over static balance-sheet variables.

The maturity-mismatch indicator enters as a positive and horizon-amplified determinant of default intensity. A concentration of short-dated obligations relative to liquid resources coincides with cumulative refinancing pressure as maturities approach.

Liquidity, proxied by cash relative to total assets, exerts a negative influence at the shortest horizons through a coverage channel that permits the settlement of immediate obligations without recourse to external funding, but this influence dissipates once the forecast window reaches periods in which liquidity stocks no longer shield firms from refinancing risk.

CC BY

TABLE 3. MAXIMUM PSEUDO-LIKELIHOOD ESTIMATES OF FORWARD DEFAULT INTENSITY COEFFICIENTS ACROSS THE TERM STRUCTURE

Covariate	Horizon $\tau = 1$	Horizon $\tau = 3$	Horizon $\tau = 6$	Horizon $\tau = 9$	Horizon $\tau = 12$
Intercept	-3.819 (0.241)***	-3.412 (0.268)***	-3.118 (0.294)***	-2.975 (0.315)***	-2.764 (0.331)***
Imputed DTD	-0.412 (0.078)***	-0.376 (0.089)***	-0.309 (0.106)**	-0.218 (0.102)*	-0.137 (0.129)
Maturity Mismatch	0.094 (0.113)	0.182 (0.092)*	0.341 (0.084)***	0.468 (0.117)***	0.623 (0.141)***
Cash / Total Assets	-0.841 (0.219)***	-0.518 (0.243)*	-0.209 (0.276)	-0.087 (0.312)	-0.029 (0.358)
CDT 90-day yield	0.043 (0.138)	0.112 (0.156)	0.229 (0.104)*	0.384 (0.129)**	0.441 (0.113)***
COLCAP Return	-0.028 (0.074)	-0.059 (0.091)	0.017 (0.083)	-0.042 (0.108)	0.013 (0.119)
Debt / EBITDA	0.118 (0.059)*	0.142 (0.068)*	0.189 (0.063)**	0.167 (0.081)*	0.109 (0.094)
Total Assets (GDP-deflated)	-0.023 (0.041)	-0.038 (0.052)	0.012 (0.059)	-0.029 (0.073)	0.058 (0.081)
EBITDA / Interest Expense	-0.218 (0.087)*	-0.174 (0.096).	-0.091 (0.108)	-0.047 (0.119)	-0.029 (0.131)
Pseudo R^2	0.274	0.241	0.213	0.182	0.164

Note: Standard errors clustered at the firm level appear in parentheses. Coefficients represent the impact on the log-hazard rate of default. Significance levels: . p < .10, * p < .05, ** p < .01, *** p < .001. The dependent variable is the annualized forward intensity of default. Pseudo R^2 is computed as McFadden's $Pseudo - R^2 = 1 - [log\ L(full)/log\ L(null)]$, where log *L(null)* is the log-likelihood of a model with intercept only.

The CDT yield variable acquires relevance only at longer horizons, which conforms to the cost-of-capital channel in which tighter monetary conditions raise future debt-service burdens and impair refinancing capacity once rate resets and loan renewals occur.

Leverage, measured through debt relative to cash generation from operations, maintains a positive association with default intensity across horizons, consistent with interest-burden mechanics that magnify insolvency risk when earnings volatility interacts with fixed contractual payments.

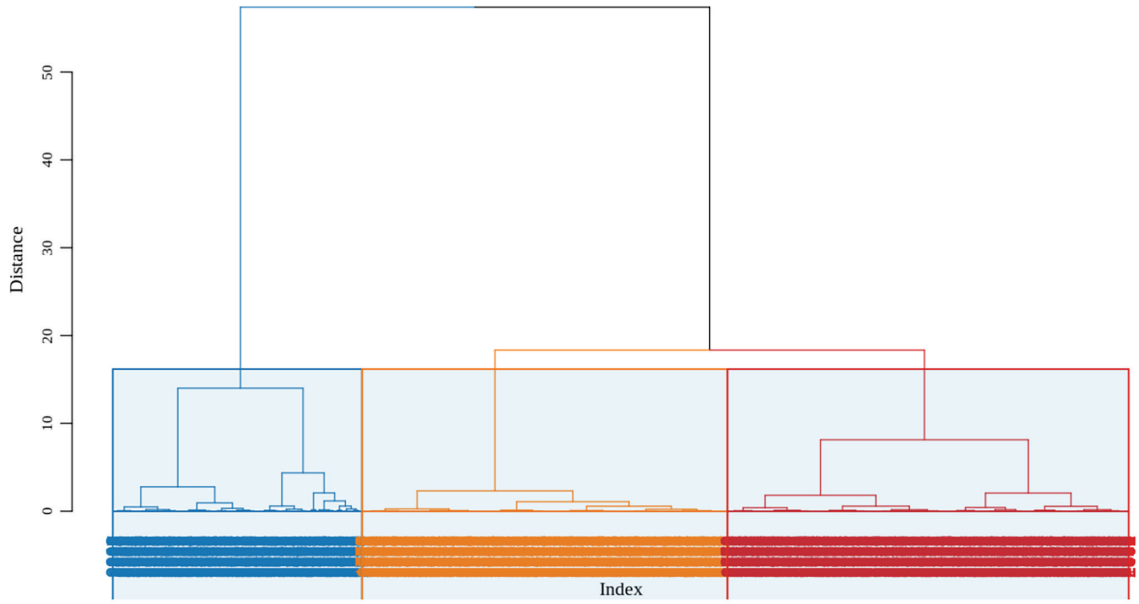
The interest coverage ratio, measured as EBITDA relative to interest expense, enters the forward intensity with a negative coefficient that is statistically significant at the shortest forecasting horizon and decays progressively as the horizon extends. A higher coverage ratio signify that cash generation from operations comfortably exceeds mandatory interest obligations which reduces the probability of imminent contractual breach. This effect works through the debt-service channel, since firms with ample interest coverage retain the capacity to honour current financial commitments without recourse to asset sales or emergency external funding. The progressive attenuation of this coefficient across horizons mirrors the transitory nature of flow-based solvency signals. As the forecast window widens, survival depends less on current earnings coverage and more on the capacity to refinance or restructure the principal stock of liabilities as it matures.

CC BY

Other regressors fail to achieve stable statistical significance, and these null results admit coherent economic explanations rather than ad hoc rationalization. The domestic equity-market return does not exert a systematic effect on private-firm default intensity because of the institutional separation between listed equity valuation and bank-mediated credit allocation in emerging economies, where private firms neither transmit nor absorb public-market sentiment in a mechanically tight manner. Firm size lacks a consistent directional role once solvency buffers, liquidity positions, and liability schedules receive control and this result is compatible with the ambiguous theoretical status of scale in credit risk, since diversification benefits coexist with larger fixed obligations and governance frictions.

Figure 1 consists of a dendrogram, the graphical output of the hierarchical clustering algorithm used to partition the sample corporations into endogenous segments. A hierarchical clustering analysis, using forward-looking probabilities of default, segregates the sample into three distinct groups based on credit risk.

FIGURE 1. DENDROGRAM OF HIERARCHICAL CLUSTERING FOR MINSKYAN CLASSIFICATION



Note: The tripartite partition is supported by the following validation metrics: Silhouette coefficient = 0.312, a modest value that mirrors the continuum of firm-level credit risk rather than discrete categories; Gap Statistic (0.346) maximized at k = 3, which provides distributional justification for the three-cluster solution against a null reference distribution; and Jaccard Stability Index = 0.72 under B = 1,000 bootstrap resamples, above the 0.70 threshold proposed by Hennig (2007) for cluster-wise stability assessment. Cluster labels (Hedge, Speculative, Ponzi) are assigned ex post according to the distributional properties of each cluster's forward default probability, interest coverage ratio, and maturity mismatch indicator.

The twelve-month-ahead forward default probability enters the clustering procedure together with selected balance-sheet indicators. The vertical axis measures the linkage distance, a measure of dissimilarity between clusters, while the horizontal axis arrays the individual firms, which are the terminal nodes of the tree. The structure depicts an agglomerative procedure, wherein individual firms are iteratively merged into larger groups based on a minimum inter-cluster separation metric. The final structure partitions the entire sample into

CC BY

three distinct super-clusters, which are mapped onto Hyman Minsky's theoretical categories of Hedge, Speculative, and Ponzi finance. This unsupervised procedure allows for the empirical identification of these financial postures based on the statistical properties of the data and simultaneously circumvents the need for arbitrary classification rules. The clear vertical separation between the three principal branches of the tree is a patent distinction among the identified subgroups and validates the discrete nature of the Minskyan classifications within this sample.

The first cluster is the most financially precarious segment of the sample. Corporations in this group report exceptionally high forward default probabilities. This financial profile is the empirical signature of the Ponzi category, which comprises corporations approaching insolvency.

The second cluster forms the largest group. It contains corporations with high forward default probabilities and elevated leverage, but whose profitability is marginal. A dependence on continuous refinancing subjects these corporations to rollover risk and severe vulnerability to constricting credit markets or adverse movements in interest rates. This condition is a defining attribute of Speculative units.

The third cluster contains the most financially robust corporations in the sample and these corporations generate sufficient income from operations to service both interest and principal obligations, a condition that insulates them from immediate market dislocations. This group, although a minority within the sample, typifies the stable capital structure of the Hedge category.

The temporal out-of-sample validation, reported in Table 4, corroborates the superior generalization capacity of the Hybrid forward-intensity model relative to conventional benchmark approaches. The discriminatory power, measured through the area under the receiver operating characteristic curve, maintains robust performance when confronted with defaults occurring in periods excluded from model calibration. The degradation from training to testing performance remains modest for the Hybrid specification; the model captures persistent default drivers. By modeling the term structure of default hazard rates conditional on observable covariates, the specification avoids overfitting to transitory correlations and instead learns the relations between solvency, liquidity, maturity mismatches, and insolvency risk that transcend specific macroeconomic regimes.

TABLE 4. TEMPORAL OUT-OF-SAMPLE VALIDATION AGGREGATE PERFORMANCE METRICS (TEST PERIOD: 2015-2018)

Model	Sample	AUC	Precision	Recall	F1-Score	Brier Score
Hybrid Forward	Train	0.878	0.510	0.624	0.561	0.092
	Test	**0.846**	**0.445**	**0.579**	**0.503**	**0.089**
Altman Z-Score	Train	0.745	0.416	0.530	0.466	0.106
	Test	0.669	0.396	0.482	0.435	0.122
Ohlson O-Score	Train	0.816	0.430	0.563	0.487	0.102
	Test	0.772	0.420	0.532	0.470	0.111
Njeru Modified	Train	0.772	0.412	0.535	0.466	0.108
	Test	0.723	0.415	0.494	0.451	0.117

Note: Training sample: Q1 2009 – Q4 2014. Test sample: Q1 2015 – Q1 2018. Precision, Recall, and F1-Score computed at the optimal threshold that maximizes F1 on training data (threshold ≈ 0.18).
Brier Score measures calibration quality; lower values denote closer correspondence between predicted probabilities and observed default frequencies.

CC BY

The precision and recall metrics illuminate critical operational trade-offs inherent in credit risk assessment. Precision measures the fraction of firms flagged as high-risk that actually default, a statistic directly relevant to the cost-benefit calculus of credit denial. Excessive false positives erode profitability by rejecting creditworthy borrowers, while false negatives expose the lender to unexpected losses. The Hybrid model sustains acceptable precision in out-of-sample prediction. Recall, conversely, measures the fraction of actual defaults successfully identified prospectively. The model registered strong recall across both training and testing samples and most firms subsequently experiencing insolvency were assigned elevated risk scores prior to distress.

The F1-score synthesizes precision and recall through their harmonic mean, and provides a balanced assessment particularly valuable when analyzing imbalanced datasets where defaults constitute a small minority of observations. The Hybrid model's F1 performance exceeds benchmark alternatives. The integration of imputed distance-to-default metrics with maturity structure variables and forward-looking intensity specifications produces superior classification accuracy relative to static accounting-based discriminant functions or logistic regressions.

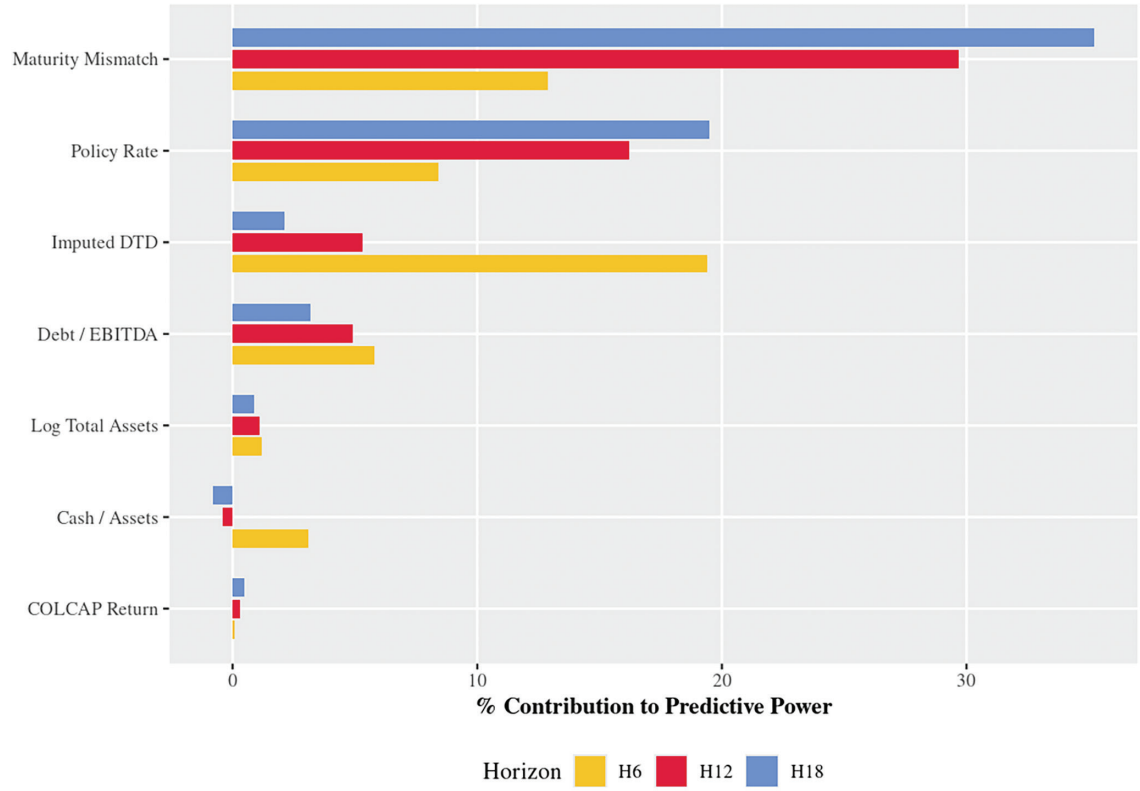
Calibration quality, assessed via the Brier score, corroborates that the Hybrid model's predicted probabilities match closely with realised default frequencies. The predicted default probability can be comprehended as the expected fraction of similar firms that will experience insolvency within the forecast horizon. The Brier score for the Hybrid specification is low in both training and testing periods; notably, the test-period score is marginally below its training-period counterpart, a pattern that mirrors the elevated incidence of financially stressed firms in the 2015–2018 test window relative to the more stable 2009–2014 calibration period. Benchmark models deteriorated progressively in calibration when confronted with out-of-sample data, specifically the Altman Z-score.

Figure 2 reports the marginal contribution decomposition of out-of-sample classification accuracy across forecasting horizons, where each firm-specific attribute is sequentially permuted while the remaining predictors remain intact. This design operationalizes a temporal attribution exercise rather than a model estimation, since the objective is to trace how the informational content of each covariate rotates as the prediction window lengthens. The Accuracy Ratio functions as the summary metric of rank-order performance, and the permutation protocol separates each variable's incremental informational role without alteration of the underlying model structure. By indexing the exercise to multiple horizons, the figure converts the theoretical distinction between short-run liquidity stress and medium-run rollover fragility into an empirical pattern that can be directly evaluated.

The results expose a stark intertemporal evolution of predictive content. At the shortest horizon, the imputed Distance-to-Default is the dominant signal, since its removal causes a sharp deterioration in classification accuracy because near-term failure hinges on proximity to the default barrier once volatility enters the metric. This dominance erodes steadily as the horizon extends, and by the one-year window the contribution collapses to statistical irrelevance. This decay accords with the concept of informational staleness embedded in accounting-derived metrics, since the volatility-adjusted distance contains dense short-horizon information yet loses persistence once balance-sheet configurations evolve. In contrast, the maturity mismatch variable follows an inverted trajectory. Its short-horizon contribution remains dormant, but its informational weight varies monotonically as the window lengthens, and eventually becomes the principal determinant of predictive accuracy at the annual horizon. This regularity embodies the Minskyan logic of financial fragility as immediate survival depends on liquidity buffers, whereas medium-term survival depends on the margin of safety and the feasibility of refinancing obligations as they mature.

CC BY

FIGURE 2. VARIABLE CONTRIBUTION TO OUT-OF-SAMPLE CLASSIFICATION ACCURACY ACROSS FORECAST HORIZONS



Cash relative to assets plays a preponderant role at the shortest horizon, congruent with liquidity absorption of transitory shocks, but its relevance dissipates rapidly as time progresses. The leverage proxy, based on total liabilities, retains a modest and stable contribution across horizons but never dominates the informational hierarchy.

Several variables fail to contribute materially at any horizon, and these null results are theoretically coherent. Firm size, for example, contributes negligibly throughout. Scale per se does not govern failure risk once liquidity, leverage, distance to default and liability schedule enter the information set. The waning contribution of cash at longer horizons is an indication of the transitory nature of liquidity buffers. Cash stocks absorb shocks in the near term. Nonetheless, they do not determine survival once rollover constraints bind.

The attenuation of the Distance-to-Default over time does not contradict theoretical expectations because delineates its temporal boundary condition because accounting-imputed parameters are devoid of stability. The delayed dominance of the maturity mismatch variable is explained by the fact that rollover risk materializes only as debt obligations approach their contractual due dates. Short-horizon variables therefore fail because the forecasting window precedes the period in which financial constraints become binding Defaults associated with liability structure originate from the inability to replace maturing obligations with new funding rather than from immediate insolvency.

Table 5 reports the panel regression of credit spreads across the three endogenously derived Minsky clusters. The econometric specification applies a pooled Ordinary Least Squares estimator with firm fixed effects and firm-level error cluster adjustment to determine the drivers

CC BY

TABLE 5. PANEL REGRESSION OF INTEREST RATE SPREADS BY CLUSTER

Variable	Cluster 1 (Hedge)	Cluster 2 (Speculative)	Cluster 3 (Ponzi)
Predicted PD (12m)	0.041 (0.058)	0.329 (0.142)*	0.941 (0.213)***
Maturity (Years)	0.187 (0.051)***	0.093 (0.054).	-0.022 (0.087)
*Maturity*2	-0.019 (0.008)*	-0.008 (0.009)	0.003 (0.014)
Collateral/Debt	-0.042 (0.048)	-0.167 (0.071)*	-0.284 (0.093)**
Term Spread	0.412 (0.156)**	0.291 (0.183)	0.114 (0.245)
Credit Spread	0.186 (0.192)	0.634 (0.221)**	0.892 (0.288)**
Firm Fixed Effects	Yes	Yes	Yes
Adjusted R^2	0.382	0.224	0.368

Note: Standard errors are clustered at the firm level and reported in parentheses. Significance levels: . p < .10, * p < .05, ** p < .01, *** p < .001. All specifications include firm fixed effects and a set of firm-level characteristics as additional controls; individual coefficients for these controls are not reported for conciseness and are available from the authors upon request.
The risk premium is the difference between the interest rate and the risk-free rate. The interest rate is the weighted average of the interest rates on outstanding debts for each firm year. The risk-free rate is estimated for each maturity using cubic spline interpolation based on the Colombian sovereign yield curve. Specifically, the interpolation is constructed from yields on 90-day certificates of deposit (CDTs) and Colombian government bonds (TES) at maturities of one, three, and five years, as published by the Banco de la República. This procedure generates a smooth, maturity-consistent term structure of risk-free rates applicable to firm-level debt instruments. Default probability is the annualized value of $p_t^i(\tau) = E_t[\int_t^{t+\tau} exp(-\int_t^s (\lambda_u^i + \phi_u^i)du)\lambda_s^i ds]$ obtained from the fitted forward intensities. Maturity is the value-weighted average of the associated maturities for each debt class, where the weights are the relative proportions of each debt class of the estimated maturity for that firm year. The collateral-to-debt ratio is the sum of the maximum credit amounts of the collateral scaled by the sum of long-term borrowing, short-term borrowing, the current portion of long-term debt, and corporate bonds. The term spread is defined as the yield differential between five-year Colombian government bonds (TES) and 90-day certificates of deposit (CDTs). This function is a proxy for the slope of the domestic yield curve and prevailing monetary conditions. The credit spread is measured as the difference between the yield on three-year corporate bonds issued by Colombian firms rated AA– and the yield on Colombian government bonds of equivalent maturity. Firm characteristics are the ratio of cash over total assets, the ratio of net income over total assets, the ratio of retained earnings over total assets, current assets less current liabilities then divided by total assets, the ratio of sales over total assets, earnings before interest, taxes, depreciation and amortization divided by total assets, the ratio of gross profits over current assets, total assets adjusted by the GDP deflator, and current liabilities minus cash then divided by total liabilities. Firm fixed effects denote whether firm fixed effects are included in the specification.

of the implicit interest rate spread. This segmented regression analysis interrogates the validity of the market discipline hypothesis through a test for regime-dependent parameter stability in the pricing of default risk, maturity duration, and collateral coverage. The model ascertains whether commercial lenders enforce a uniform rate determination function or whether the elasticity of the risk premium with respect to financial fragility varies non-linearly across the Hedge, Speculative, and Ponzi taxonomies.

CC BY

Estimates for Hedge schemes validate the dominance of the term structure of interest rates, as the statistical significance of the linear maturity term, coupled with the negative quadratic term, supports the presence of a concave yield curve. Lenders assess the duration of capital for financially robust entities while they assign negligible weight to marginal variations in default probability.

The function for the Speculative and Ponzi segments incorporates the predictive default probability as a distinct and primary determinant of the cost of funds. The relevance of the model-implied hazard rate supports the view that financial intermediaries internalize the insolvency risk when they extend credit to fragile borrowers. The inverse relationship between collateral coverage and the risk premium within these distressed firms reinforces the theoretical expectation that secured contracts moderate agency costs and loss-given-default expectations when the probability of repayment depends on asset liquidation rather than operational cash flow. Among the firm-level controls, liquidity buffers (the cash-to-assets ratio) also appear to moderate spreads for the Speculative group. This result offers support for the Minskyan proposition that the maintenance of payment commitments for this intermediate group depends on liquid assets. Credit spread exerts a discernible influence on Ponzi firms bec entities with unsustainable capital structures are highly sensitive to aggregate liquidity contractions and to shifts in the external availability of credit.

The absence of statistical significance for the predicted default probability within Hedge firms points to the existence of an information-insensitive region for high-quality debt where lenders treat solvency risk as negligible. Market participants discern no material difference in creditworthiness among robust firms, and this condition renders the spread inelastic to marginal fluctuations in the distance-to-default. A distinct theoretical mechanism explains the diminished role of maturity for Speculative firms and its statistical irrelevance for the Ponzi regime. The breakdown of the standard term structure for distressed entities is explained by the fact that loan maturity no longer matters for pricing. As default draws near, lenders focus on the time to default rather than the stated length of the contract. All loans are treated as short-term because default, not maturity, determines the payoff horizon. Firm scale also fails to exert explanatory power within unstable firms and does not neutralize the financial stress inherent to a Ponzi financial structure, despite the diversification advantages ordinarily attributed to larger balance sheets.

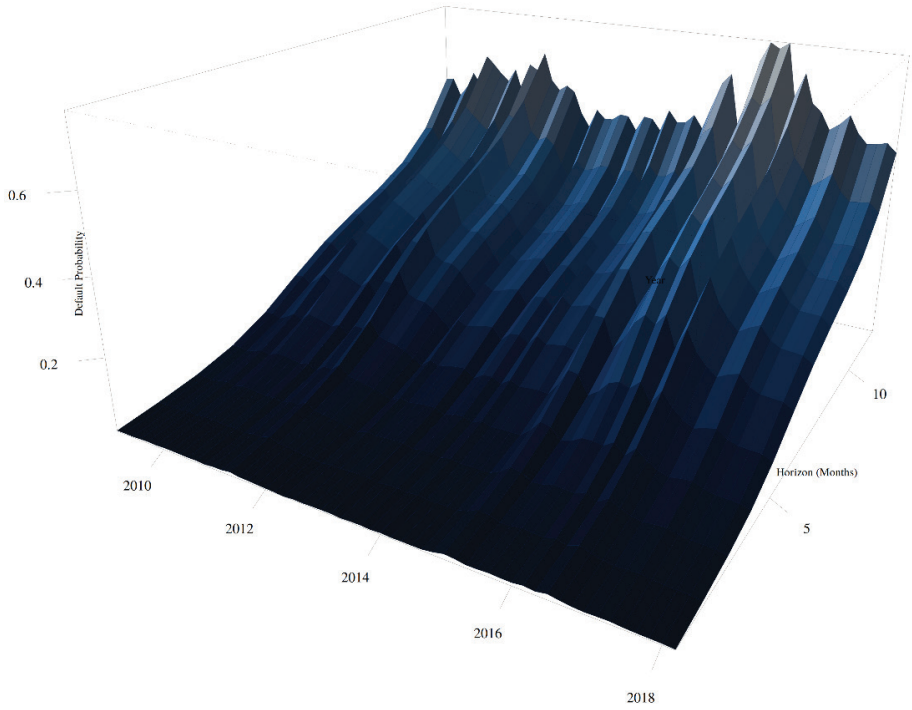
The fitted term-structure surface of median default probabilities (Figure 3), obtained from the panel of Colombian private manufacturing firms, shows a temporal profile whose morphology exceeds the representation of a single cycle around the terms-of-trade shock. The trajectory opens at elevated levels during the first half of 2009, a residual of the lagged propagation of the global financial crisis to the Colombian economy. The decline observed over the subsequent months coincides with the expansionary phase of domestic monetary policy and the recovery of international hydrocarbon prices, which restored profit margins in the tradable goods sector.

The interval between 2010 and 2013 records a low floor of median default probability, interrupted by moderate undulations around international macroeconomic events such as the European sovereign debt crisis during the third quarter of 2011 and the taper tantrum of the second quarter of 2013. Neither episode propagates to the private manufacturing panel and it is observed that the COLCAP return, included in the model as a financial-cycle covariate, carries limited transmission to bank-intermediated corporate credit, leaving the underlying level of credit risk essentially unchanged in either case.

The uneven temporal profile of the 2014 to 2018 window is the distinctive feature of the surface because the median probability rises rapidly from the fourth quarter of 2014, coincident with the rupture in international crude prices. It reaches its maximum over the 2015 to 2016 cycle under the contractionary monetary regime associated with the defense of the peso, and

CC BY

FIGURE 3. FITTED TERM-STRUCTURE SURFACE OF MEDIAN DEFAULT PROBABILITIES



recedes along an extended period of gradual descent whose temporal scale is longer than the propagation phase of the adverse impulse. This uneven profile works as visual evidence of the mechanism of accumulated fragility in the sense of Minsky. The transition of private firms toward Speculative and Ponzi regimes materializes with speed under macroeconomic deterioration, while the rebuilding of balance-sheet buffers and the restoration of safety margins unfold over an extended horizon. Adjustment costs faced by the private manufacturing firm are severe in an economy with shallow capital markets and dependence on relationship-based bank financing.

4. DISCUSSION

The association between DTD and default probability works through the option-theoretic representation introduced by Merton (1974) and extended by Black and Cox (1976) to incorporate time-varying default barriers. Equity holders hold a European call option on firm assets with strike equal to debt face value. The probability that the option expires out of the money, which equals firm default, depends critically on the moneyness measure that DTD approximates. A larger buffer between current asset value and the contractual default threshold is associated with lower instantaneous default intensity and this relation appears robust across the macro-financial regimes covered by the estimation window, and the DTD coefficients show only modest variation between tranquil and stressed periods.

The imputed Distance-to-Default metric preserves explanatory content at short forecasting horizons and decays monotonically as temporal separation between observation and outcome increases. This pattern accords with the incomplete-information models developed by Duffie

CC BY

and Lando (2001), which formalize processes where market participants extract signals from observable accounting fundamentals. In those specifications default intensity contains a Bayesian updating mechanism under which investors refine beliefs about the firm's latent asset value through noisy balance-sheet observations. The informational content of those signals degrades as staleness accumulates and this study operationalizes that property through sequential regression horizons indexed by months elapsed since fiscal-year publication.

Bharath and Shumway (2008) reported that simplified Distance-to-Default measures, constructed without an iterative solution of the Merton equations, achieve discrimination capacity comparable to full maximum-likelihood implementations and this study extends that result. An imputed DTD metric, constructed from accounting variables through parameter transfer from public firms, retains predictive efficacy despite the absence of direct market observables. This outcome expands the methodological frontier opened by Duan *et al.* (2018), who projected market-implied risk measures onto privately held firms in a Korean institutional setting with distinct disclosure rules and ownership structures.

Short-term liabilities net of liquid assets scaled by total liabilities is the dominant predictor of medium- and long-term default intensity and oefficients increase markedly as forecast horizons lengthen. This temporal pattern provides empirical support for Minsky's (1986) Financial Instability Hypothesis, specifically the proposition that financial fragility is associated with accumulated rollover risks that crystallise when credit conditions tighten. The mechanism that links this variable to default intensity works through the coordination-failure logic formalised by He and Xiong (2012). In their dynamic model of debt rollover, creditors face strategic complementarity as each lender's willingness to refinance depends on beliefs that other creditors will also renew claims. When a firm concentrates short-term liabilities that mature simultaneously, creditors find themselves in a collective-action problem. If any subset refuses renewal, the firm faces a liquidity shortfall that can trigger default even when asset present value exceeds liabilities. This coordination friction generates multiple equilibria, namely, a good equilibrium in which creditors roll claims and the firm survives, and a bad equilibrium in which creditor runs become self-fulfilling. This finding reconciles two empirical regularities reported in corporate finance. Gopalan *et al*., (2014) observed that firms with shorter-debt maturities recorded lower default rates during expansion, a result they attributed to disciplinary effects of frequent renegotiation. In another inquiry, Almeida *et al.* (2012) found that short-maturity firms suffered sharply higher default rates during the global financial crisis. The forward-intensity specification accommodates both empirical regularities because it allows the maturity-mismatch coefficient to vary across horizons. It is observed that over short windows rollover risk remains latent and over extended windows accumulated refinancing pressure dominates.

Cash holdings relative to total assets is negatively related with default intensity at the one-month horizon, consistent with precautionary-liquidity theory articulated by Keynes (1936) and formalized in modern corporate finance by Holmström and Tirole (1998). In their treatment, financially constrained firms retain liquid reserves to self-insure against transitory cash-flow shocks and to avoid costly premature liquidation of productive assets. The magnitude of this effect is consistent with estimates from Acharya *et al*. (2012), who found that cash balances reduce near-term default probability for firms in lower credit-quality quintiles.

The negative association between cash holdings and default intensity at short horizons obeys to the fact that firms with temporary revenue shortfalls can draw on cash reserves to meet contractual debt service without activation of acceleration clauses or covenant violations. This buffer function proves especially valuable for Speculative financing units in Minsky's taxonomy, firms whose cash flows from operations cover interest but not principal amortization.

CC BY

For these entities, cash stocks give firms the runway necessary to arrange renewal of credit before illiquidity converts into insolvency.

The insurance role of cash erodes rapidly as forecast horizons lengthen as at intermediate windows coefficient magnitudes attenuate, and at the twelve-month horizon statistical significance vanishes. This decay follows from the distinction between stock and flow variables in credit assessment. Liquidity represents a finite, depletable resource; absent sustained positive cash generation reserves erode. Denis and Sibilkov (2010) attested empirically that the marginal value of cash for constrained firms derives from the flexibility it confers to undertake profitable investments during downturns rather than from the stock itself. Once immediate liquidity shortfalls pass, long-term viability depends on the firm's capacity to generate cash from operations, a feature captured by profitability and leverage ratios rather than balance-sheet liquidity.

The yield on 90-day certificates of deposit (CDT yield), the Colombian equivalent of the commercial paper rate employed by Duan *et al.* (2018), exerts negligible effect on default intensity at short horizons but becomes a material determinant at twelve-month windows. This lagged transmission accords with the credit-channel paradigm developed by Bernanke and Gertler (1989) and tested in emerging-market contexts by Jiménez *et al.* (2014). The central mechanism works through refinancing costs because Colombian firms finance operations primarily through short-duration bank credit with floating rates indexed to the domestic money market. When market rates rise during episodes of peso defense, firms do not default instantaneously; rather, rollover risk escalates as existing obligations come due and need renewal at less favorable terms.

The temporal lag between rate moves and default realizations exposes institutional frictions in credit markets. Kashyap and Stein (2000) reported that monetary transmission aggravates through bank-lending channels when intermediaries face balance-sheet constraints. Colombian banks, after funding-cost increases and borrower-quality deterioration during the commodity collapse, tightened lending standards and shortened loan maturities. Firms with obligations due during this interval faced higher effective rates and reduced credit availability, a double squeeze captured by the CDT yield.

The estimated association between the CDT yield and default intensity is congruent with a financial-accelerator mechanism in which monetary tightening propagates through endogenous borrowing constraints because rate increases produce heterogeneous effects across the firm distribution. Hedge units with ample coverage ratios and access to longer-duration funding absorb rate increases through limited margin compression. By contrast, Speculative and Ponzi units function with narrow interest-coverage margins and depend on continuous renewal of credit; for those entities marginal rate changes can exhaust the residual buffer between income from operations and debt service.

The absence of a firm-size effect in full specifications, after control for model covariates, resolves a long-standing empirical puzzle. Early studies treated size as a default predictor under the assumption that larger firms benefit from diversification and scale economies. Bharath and Shumway (2008) showed that size effects disappear once models explicitly include leverage, profitability, and distance-to-default and these findings extend that result to private firms in emerging markets. Size works through observable channels (access to liquidity, bargaining power with creditors, operational diversification) and does not exert independent causal influence. Size being redundant as a fundamental predictor supports the parsimony of the forward-intensity specification.

Colombian stock-index returns show no systematic relation with private-firm default intensity, an empirical regularity predicted by the market-segmentation thesis for bank-based emerging economies (Levine, 2002). In bank-based financial systems typical of Latin America, equity-market signals carry limited transmission power to corporate credit conditions because

CC BY

firms finance operations through bank relationship lending. The manufacturing firms in this sample lack access to public-equity issuance, obtain no funding relief from favorable market valuations, and face funding costs determined through bilateral negotiations with domestic banks.

The irrelevance of equity returns supports the core assumption that underlies the DTD imputation because private firms, despite absent observable equity prices, conform to the same relation between accounting fundamentals and credit risk as their public counterparts. If private firms reacted systematically to public-equity movements, the imputation would introduce contamination bias. Public and private firms function in distinct informational and financing spheres and the parameter-transfer exercise therefore remains theoretically defensible.

The partitioning of the sample into three risk clusters with financial profiles coherent with Minsky's taxonomy, within the replicability bounds reported by the validation metrics, provides empirical support for the claim that Minsky's financial postures map onto observable firm characteristics. The clustering algorithm, applied without ex ante thresholds to forward default probabilities and financial ratios, recovers discrete groupings that correspond to Minsky's definitions. The progression of default probabilities across clusters measures credit-quality deterioration that accompanies each transition. These magnitudes correspond with financial stratification reported by Altman *et al.* (2010) for rated corporate bonds, where investment-grade issuers default at low single-digit rates, speculative-grade non-distressed issuers default at intermediate rates, and distressed issuers stand on the edge of imminent failure. Minsky's taxonomy, although formulated as a macro theory of systemic instability, contains microeconomic content applicable to firm-level risk assessment.

The regime-dependent pricing exercise gives independent corroboration of the taxonomy's practical relevance. Interest-rate spreads show no sensitivity to estimated default probabilities for Hedge firms. This pattern is compatible with the information-insensitive debt regime of Dang *et al*. (2012), though a low within-cluster dispersion of predicted default probability cannot be ruled out as an alternative reading. By contrast, spreads for Ponzi firms react with acuteness to default-probability estimates and creditors price tail risk actively when they lend to distressed borrowers. The loading of the Ponzi cluster on predicted default probability reaches the same order of magnitude reported by Duan *et al.* (2018, Table 4, Panel B, Models 1 and 2) for the full Korean private-firm sample, which strengthens the external comparability of the segmented pricing exercise.

The identification of regimes through unsupervised algorithms avoids arbitrary cutoffs that afflict traditional credit-score methods. Altman's Z-Score, for example, classifies firms into risk bands via fixed thresholds calibrated to 1960s U.S. manufacturing data. Those thresholds lack theoretical grounding and fail to adapt to changes in industrial composition, accounting standards, or macro conditions. Hierarchical clustering allows the data to determine boundaries. Cluster limits are derived endogenously from the multivariate distribution of risk metrics rather than researcher fiat.

Maturity mismatch and short-term refinancing pressure function as salient predictors of default. The maturity mismatch indicator shows positive and statistically significant loadings across forward horizons from one-quarter ahead onward, with amplification toward the twelve-month window. Minsky (1986) contended that as economies undergo prolonged periods of stability, financial structures become more fragile as speculative and Ponzi arrangements proliferate. Adrian and Brunnermeier (2016) likewise identified maturity mismatch and leverage as primary mechanisms for the propagation of systemic risk.

Hybrid models, such as those developed by Duan *et al.* (2018) and Chan-Lau *et al.* (2024), attest to the utility of commingling market-based and accounting-based signals through transparent and replicable procedures well-suited for computationally constrained settings. The

CC BY

evidence from the Colombian market affirms that a parametric imputation of DTD, performed without direct equity price inputs, raises the accuracy of default forecasts and certain balance-sheet proxies adequately compensate for absent market data.

5. CONCLUSION

This research addressed a foundational problem in corporate credit risk assessment for emerging markets, namely the measurement of default probability for privately held corporations that lack observable equity prices, yet function within financial structures vulnerable to macroeconomic shocks and refinancing constraints. The central contribution resides in the development and empirical validation of a hybrid forward-intensity model that reconstructs market-implied solvency measures for unlisted entities through parameter transfer from public firms.

Prior research predominantly confined Minskyan analysis to aggregate financial stability or applied market-based credit models exclusively to listed corporations with liquid equity markets. The theoretical vacuum concerned the operational translation of Minsky's taxonomy (Hedge, Speculative, and Ponzi financing postures) into empirically testable classifications grounded in observed accounting data and forward-looking probability measures for private firms.

The inquiry begins from the econometric mapping of accounting fundamentals to the market-implied default probabilities observed in publicly listed Colombian corporations, which yields an imputed Distance-to-Default metric for private firms that circumvents the data limitations inherent to unlisted entities. This metric then feeds a forward-intensity model of corporate default that combines the market-implied measures with explicit liability-structure variables across multiple forecast horizons. From the predicted probabilities, unsupervised hierarchical algorithms partition the sample into endogenous risk clusters that correspond to Minsky's financing categories, whose economic relevance the regime-dependent interest-rate pricing equations subsequently corroborate.

The propensity score matching procedure successfully neutralized compositional heterogeneity between listed and unlisted corporations for the majority of conditioning variables and post-match standardized mean differences converged below conventional thresholds for firm size, profitability, leverage, operational efficiency, and liquidity measures. Pre-match disparities remained uncorrected for sales growth, as growth dynamics differ fundamentally between publicly listed and privately held corporations; this variable was consequently excluded from all subsequent parameter transfer procedures.

The horizon-decomposed ordinary least squares regressions exposed systematic temporal decay in the informational content of accounting fundamentals. Profitability measures constructed from net income relative to total assets retained explanatory power exclusively at short informational lags, consistent with the proposition that cash-flow generation provides a transitory buffer against asset-value volatility that markets re-price rapidly as staleness accumulates. Leverage ratios anchored in total liabilities maintained stable influence across all twelve informational horizons, which accorded with the persistent nature of capital structure as a determinant of contractual default barriers in market-based credit models. Operational efficiency proxies based on sales relative to assets showed a sign reversal across informational horizons, with a negative coefficient at the shortest lag that reproduces the pattern documented by Duan *et al.* (2018) at equivalent horizons and a positive coefficient at extended lags that recovers the standard signal of operational efficiency as a determinant of asset value-generating capacity. Interest coverage measured through earnings before interest relative to

CC BY

interest expenses showed short-horizon relevance that dissipated rapidly as temporal separation widened, a pattern that mirrors the flow-based nature of debt-service stress. Book equity relative to total liabilities entered with the expected positive sign on the imputed Distance-to-Default, strongest and marginally significant at the shortest informational lag, with attenuation thereafter, consistent with the credit-risk model and with the positive coefficients reported by Duan *et al.* (2018); a stronger equity cushion is associated with a larger margin separating asset value from the default barrier. The nominal exchange rate showed heterogeneous and unstable effects across firms, attributable to unobserved variation in currency exposure and hedging practices that diluted its detectability in pooled specifications.

Hierarchical clustering applied to predicted default probabilities partitions firms into three financial regimes. One group depends on refinancing or asset liquidation to meet obligations, a second group services interest but not principal and remains exposed to rollover pressure, and a third group honors contractual commitments through internal cash flow with low insolvency frequency.

Comparative out-of-sample evaluation against Altman, Ohlson, and modified market-based benchmarks ranks the hybrid specification above traditional scoring rules. Accounting-only systems misclassify fragile firms, whereas the specification that combines distance-to-default with maturity structure identifies cumulative refinancing pressure preceding failure.

The final spread regression estimated within each cluster shows lenders price refinancing exposure heterogeneously across financial regimes, with borrowing premia concentrated in firms whose cash flow cannot amortize principal and minimal premia for internally funded firms.

Several limitations warrant explicit recognition as the empirical analysis relies on a finite temporal window that restricts observation of full macroeconomic cycles, which constrains inference across differing credit regimes. The reliance on accounting disclosures excludes qualitative dimensions such as governance practices or lender relationships, which bear on credit outcomes but escape formal measurement. The transfer of information from public to private firms rests on maintained comparability assumptions that, despite the matching procedure, remain subject to institutional heterogeneity. The empirical classification of firms into Hedge, Speculative, and Ponzi categories relies on proxy financial indices calculated from annual balance sheet and income statement data. These proxies necessarily simplify Minsky's original concepts, which involved present values of expected future cash flows and commitments over extended periods, rather than solely current or realized values. The analysis is also confined to privately held (unquoted) manufacturing firms in Colombia during a specific period, which limits the generalizability of the findings to other sectors, institutional contexts, or macroeconomic regimes.

This research delineates several avenues warranting further inquiry. Two extensions follow directly. First, replication to other Latin American manufacturing samples with similar bank-centric intermediation would test the stability of the parameter-transfer procedure. Second, incorporation of bank-relationship variables, such as loan concentration and duration of creditor relationships, would separate the maturity-mismatch effect from its relational counterpart. Given the limitations of balance-sheet data in capturing forward-looking risks, the inclusion of market-based proxies like the DTD, even if imputed, maximizes the predictive capacity of default events. Firms function under uncertainty and frequently depend on anticipated funding rounds or refinancing events; this market-implied measure captures shifts in investor assessment more expeditiously than lagged financial statements permit.

Despite these constraints, the study advances a disciplined approach to default analysis under conditions that mirror those faced by practitioners and supervisors in bank-centered

CC BY

economies. By unifying balance sheet information, maturity structure, financing spreads, and financial taxonomy within a single empirical object, the analysis offers a pathway toward interpretable risk assessment beyond market-dependent settings.

6. FUNDING

This research received no external funding.

7. BIBLIOGRAPHY

Acharya, V. V., Davydenko, S. A., & Strebulaev, I. A. (2012). Cash holdings and credit risk. *Review of Financial Studies, 25*(12), 3572–3609. https://doi.org/10.1093/rfs/hhs106

Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review, 106*(7), 1705–1741. https://doi.org/10.1257/aer.20120555

Almeida, H., Campello, M., Laranjeira, B., & Weisbach, M. S. (2012). Corporate debt maturity and the real effects of the 2007 credit crisis. *Critical Finance Review, 1*, 3–58. https://doi.org/10.1561/104.00000001

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I., & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the US market. *Abacus, 43*(3), 332–357. https://doi.org/10.1111/j.1467-6281.2007.00234.x

Altman, E. I., Sabato, G., & Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk, 6*(2), 95–127.

Ayyagari, M., Beck, T., & Demirgüç-Kunt, A. (2007). Small and medium enterprises across the globe. *Small Business Economics, 29*(4), 415–434. https://doi.org/10.1007/s11187-006-9002-5

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *The British Accounting Review, 38*(1), 63–93. https://doi.org/10.1016/j.bar.2005.09.001

Bernanke, B. S., & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review, 79*(1), 14–31.

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies, 21*(3), 1339–1369. https://doi.org/10.1093/rfs/hhn044

Black, F., & Cox, J. C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance, 31*(2), 351–367.

Chan-Lau, J. A., Hu, R., Mungo, L., Qu, R., Xin, W., & Zhong, C. (2024). Monitoring privately-held firms' default risk in real time: A signal-knowledge transfer learning model. *IMF Working Papers, 2024*(115). https://doi.org/10.5089/9798400278396.001

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*(4), 537–569. https://doi.org/10.1093/rof/8.4.537

Cheraghali, H., & Molnár, P. (2024). SME default prediction: A systematic methodology-focused review. *Journal of Small Business Management, 62*(6), 2847–2905. https://doi.org/10.1080/00472778.2023.2277426

Ciampi, F., Giannozzi, A., Marzi, G., & Altman, E. I. (2021). Rethinking SME default prediction: A systematic literature review and future perspectives. *Scientometrics, 126*(3), 2141–2188. https://doi.org/10.1007/s11192-020-03856-0

Dang, T. V., Gorton, G., & Holmström, B. (2012). *Ignorance, debt and financial crises*. https://www.columbia.edu/~td2332/Paper_Ignorance.pdf

Denis, D. J., & Sibilkov, V. (2010). Financial constraints, investment, and the value of cash holdings. *Review of Financial Studies, 23*(1), 247–269. https://doi.org/10.1093/rfs/hhp031

Duan, J. C., Kim, B., Kim, W., & Shin, D. (2018). Default probabilities of privately held firms. *Journal of Banking & Finance, 94*, 235–250. https://doi.org/10.1016/j.jbankfin.2018.08.006

Duan, J. C., Sun, J., & Wang, T. (2012). Multiperiod corporate default prediction: A forward intensity approach. *Journal of Econometrics, 170*(1), 191–209. https://doi.org/10.1016/j.jeconom.2012.04.007

Duffie, D., & Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica, 69*(3), 633–664.

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics, 83*(3), 635–665.

du Jardin, P. (2021). Dynamic self-organizing feature map-based models applied to bankruptcy prediction. *Decision Support Systems, 147*, 113576. https://doi.org/10.1016/j.dss.2021.113576

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research, 15*, 3133–3181. https://jmlr.org/papers/v15/delgado14a.html

Gopalan, R., Song, F., & Yerramilli, V. (2014). Debt maturity structure and credit quality. *Journal of Financial and Quantitative Analysis, 49*(4), 817–842. https://doi.org/10.1017/S0022109014000520

He, Z., & Xiong, W. (2012). Dynamic debt runs. *Review of Financial Studies*, *25*(6), 1799–1843. https://doi.org/10.1093/rfs/hhs004

Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis, 52*(1), 258–271.

Holmström, B., & Tirole, J. (1998). Private and public supply of liquidity. *Journal of Political Economy, 106*(1), 1–40.

Jiménez, G., Ongena, S., Peydró, J. L., & Saurina, J. (2014). Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica, 82*(2), 463–505. https://doi.org/10.3982/ECTA10104

Kashyap, A. K., & Stein, J. C. (2000). What do a million observations on banks say about the transmission of monetary policy? *American Economic Review, 90*(3), 407–428. https://doi.org/10.1257/aer.90.3.407

Keen, S. (1995). Finance and economic breakdown: Modeling Minsky's financial instability hypothesis. *Journal of Post Keynesian Economics, 17*(4), 607–635.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. Macmillan.

Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. *European Journal of Operational Research, 247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Levine, R. (2002). Bank-based or market-based financial systems: Which is better? *Journal of Financial Intermediation, 11*(4), 398–428. https://doi.org/10.1006/jfin.2002.0341

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*(2), 449–470. https://doi.org/10.1111/j.1540-6261.1974.tb03058.x

Minsky, H. P. (1986). *Stabilizing an unstable economy*. Yale University Press.

Njeru, A. (2019). An alternative approach to estimation of the probability of default for commercial entities: The modified KMV Merton model. *Journal of Risk and Financial Management, 12*(3), 134.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131. https://doi.org/10.2307/2490395

Rashid, U., Abdullah, M., Khatib, S. F., Khan, F. M., & Akhter, J. (2024). Unravelling trends, patterns and intellectual structure of research on bankruptcy in SMEs: A bibliometric assessment and visualisation. *Heliyon, 10*(2), e24254. https://doi.org/10.1016/j.heliyon.2024.e24254

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74*(1), 101–124. https://doi.org/10.1086/209665

Whalen, C. J. (2001). The financial instability hypothesis: A restatement. *Public Policy Brief No. 62*. Levy Economics Institute.

APPENDIX A. DERIVATIONS OF THE FORWARD INTENSITY MODEL

The modeling apparatus centers on a Forward Intensity Model and this taxonomy specifies the default of corporation *i*as the first jump of a doubly-stochastic Poisson process, $N_i(t)$. The jump rate of this process, $\lambda_i(t)$, defines the spot default intensity. A distinct spot intensity, $\phi_i(t)$, governs 'other exits' such as mergers or acquisitions.

The model avoids the direct specification of spot intensities, a procedure with known complexities, and instead models forward intensities. The variable $f_i(t,\tau)$ is the forward default intensity at time *t*for a default projected at a future time $t + \tau$. The variable $g_i(t,\tau)$ is the combined forward exit intensity-encompassing both default and other exit types-at time *t*for an event at $t + \tau$.

The functional specification for these forward intensities is an exponential function of a vector of covariates, $X_i(t)$, observable at time t. The coefficients for these covariates, $\alpha_j(\tau)$ and $\beta_j(\tau)$, are themselves functions of the forward horizon τ. The forward default intensity is specified as $f_i(t,\tau) = exp\ (\alpha_0(\tau) + \sum_j \alpha_j(\tau)x_{ij}(t))$, and the combined forward exit intensity is $g_i\ (t,\tau) = f_i\ (t,\tau) + exp\ (\beta_0(\tau) + \sum_j \beta_j(\tau)x_{ij}(t))$.

The covariate vector $x_i(t)$ incorporates macroeconomic factors. Specifically, the yield on 90-day certificates of deposit (CDTs, the Colombian equivalent of the commercial paper rate),

the COLCAP stock market index return and firm-specific attributes. The definition of these forward intensities permits the calculation of the conditional default probability, $p_i(t,T)$, for corporation *i*over the interval $[t,t+T]$. Because the covariates $X_i(t)$ are known at time t, the conditional expectation operator, E_t, is superfluous.

The estimation constructs a pseudo-likelihood function, $L_{i,k\Delta t,\tau}$, for each corporation i at each discrete time point $k\Delta t$ for a given horizon τ. This function calculates the probability of the corporation's state (survival, default, or other exit) over the prediction interval $[k\Delta t, k\Delta t+\tau]$, conditional on its status at the interval's start. Its construction evaluates outcomes for each subinterval $s\Delta t$ (where s ranges from 1 to $\tau/\Delta t$). First, survival, with a probability related to $exp(-g_i(k\Delta t, s\Delta t)\Delta t)$; second, default, with a probability related to $(1-exp(-f_i(k\Delta t, s\Delta t)\Delta t))$; and third, another exit type, with a probability related to $(1-exp(-h_i(k\Delta t, s\Delta t)\Delta t))$, where $h_i = g_i - f_i$. The aggregate pseudo-likelihood for a specific horizon τ, L_τ, is the product of these individual contributions, $L_{i,k\Delta t,\tau}$, across all corporations and all starting points $k\Delta t$.

The 'pseudo' designation is appropriate because firm events may not be strictly independent, and prediction intervals overlap when τexceeds Δt. A central feature of this estimation approach is the separability of the likelihood which allows the independent estimation of coefficients $\alpha(s\Delta t)$ and $\beta(s\Delta t)$ for each sub-horizon $s\Delta t$.

Computation of default probabilities

The default term structure model employed closely follows the approach of Duan *et al.* (2012), which builds upon the foundational approach of Duffie *et al.* (2007) by introducing forward default intensities. Specifically, the default event of the *i*-th private firm is modeled as the occurrence of a jump in a doubly-stochastic Poisson process, denoted by N_t^i, which is driven by a non-negative spot default intensity process λ_t^i. The default time τ_D^i of the firm is defined as the first instance at which N_t^i reaches one.

Mathematically, $N_t^i - \int_0^t \lambda_s^i ds$ forms a martingale relative to F and P and the model's dynamics are considered only up to the stopping time τ_D^i.

The spot default intensity λ_t^i captures the instantaneous probability of default conditional on survival up to time t, such that for a sufficiently small interval $\Delta > 0$, the conditional probability of default within the next period satisfies:

$$s_t^i(\tau) = E_t\left[exp\left(-\int_t^{t+\tau}(\lambda_s^i + \phi_s^i)ds\right)\right]$$

This probabilistic structure permits the derivation of a forward-looking default probability term structure for each firm, and traces the evolution of credit risk over time, based on firm-specific intensity processes.

The model extends beyond default events by incorporating exits for non-default reasons (mergers or acquisitions, for example) to prevent censoring bias. Each firm's non-default exit follows a distinct doubly-stochastic Poisson process M_t^i with a non-negative spot exit intensity ϕ_t^i. $M_t^i - \int_0^t \phi_s^i ds$ is also a martingale relative to F and P.

By construction the condition $\tau_D^i \geq \tau_C^i$ holds, and the instantaneous combined exit intensity is $\lambda_t^i + \phi_t^i$ at time t. Accordingly default probability over $[t,t+\tau]$ assumes the form:

$$p_t^i(\tau) = E_t\left[\int_t^{t+\tau} exp\left(-\int_t^s (\lambda_u^i + \phi_u^i)du\right)\lambda_s^i ds\right]$$

The Duan *et al.* (2012) approach begins to deviate from the spot intensity model by introducing a forward intensity version of the above model as a new tool for default prediction over a range of horizons. We first denote by $f_t^i(\tau)$ the forward default.

The forward default intensity for *i*-th firm, at horizon u, conditional on survival to time *t* is defined by:

$$f_t^i(\tau) = s_t^i(\tau) \cdot \lim_{\Delta \downarrow 0} \frac{P(t+\tau < \tau_D^i \leq t+\tau+\Delta \mid F_t)}{\Delta}$$

In analogous fashion, the forward combined-exit intensity takes the form:

$$g_t^i(\tau) = s_t^i(\tau) \cdot \lim_{\Delta \downarrow 0} \frac{P(t+\tau < \tau_C^i \leq t+\tau+\Delta \mid F_t)}{\Delta}$$

Notice that spot intensity is a special example of forward intensity in that $f_t^i(0) = \lambda_t^i$ and $g_t^i(0) = \lambda_t^i + \phi_t^i$. Equivalently, we can also express $s_t^i(\tau)$ and $p_t^i(\tau)$

$$s_t^i(\tau) = ex\left(-\int_0^\tau g_t^i(s)ds\right)$$

$$p_t^i(\tau) = exp\left(-\int_0^s g_t^i(u)du\right) f_t^i(s)ds$$

The forward-intensity formulation incorporates the aggregate impact of observable covariates at time t into the functions $f_t^i(\tau)$ and $g_t^i(\tau)$. In other terms, this formulation embeds the impact of all future spot-intensity dynamics into the functions previously mentioned. This approach bypasses the need to model high-dimensional stochastic covariate processes and their joint evolution, and it enables direct calculation of default and exit probabilities across multiple horizons.

The forward-intensity model explicitly absorbs into a set of forward intensity functions the effects arising from the evolution of future spot intensities. The forward intensities corresponding to different forward starting times are functions of variables (i.e., stochastic covariates) observable at the time of making predictions. In short, predictions for various future horizons can be made without having to know the dynamics of the stochastic covariates.

In this paper, the methodology proposed by Duan *et al.* (2012) is implemented by specifying the following family of forward intensity functions:

$$f_t^i(\tau) = ex\left(\alpha_0(\tau) + \sum_{j=1}^{k} \alpha_j(\tau) x_t^i(j)\right)$$

$$g_t^i(\tau) = f_t^i(\tau) + exp\left(\beta_0(\tau) + \sum_{j=1}^{k} \beta_j(\tau) x_t^i(j)\right)$$

where $X_t^i = (x_t^i(1), x_t^i(2), \cdots, x_t^i(k))$ collects the set of observable covariates that comprise common risk factors and firm-specific attributes. Coefficients vary with the forward starting time, t, and horizon, u. Empirical implementation proceeds via a discrete-time approximation with a one-month interval and yields a monthly forward-intensity model.

Overlapped pseudo-likelihood function

The calibration of the forward-intensity model proceeds over a discrete monthly grid from time 0 to *T*. The estimation is based on the construction of a pseudo-likelihood function under the doubly stochastic assumption.

Let the monthly time step be denoted by $\Delta t = 1$. The pseudo-likelihood function over the entire sample period for prediction horizon Δt takes the form:

$$L_\tau(\alpha, \beta; \tau_C, \tau_D, X) = \prod_{i=1}^{N} \prod_{k=0}^{\lceil T/\Delta t \rceil - 1} L_{\tau, k\Delta t}^i(\alpha, \beta; \tau_C^i, \tau_D^i, X^i)$$

CC BY

where $\alpha = \{\alpha(0),\ldots,\alpha(\tau-1)\}$, $\beta = \{\beta(0),\ldots,\beta(\tau-1)\}$,$\tau_C = \{\tau_C^1,\ldots,\tau_C^N\}$,$\tau_D = \{\tau_D^1,\ldots,\tau_D^N\}$, $X = \{X^1,\ldots,X^N\}$, and

$$
\begin{aligned}
L^i_{\tau,t}(\alpha,\beta;\tau_C{}^i,\tau_D{}^i,X^i) &= 1_{\{t_0^i \le t,\ \tau_C^i > t+\tau\}}\, L_t(\tau_C^i;\ \tau_C^i > t+\tau) \\
&+1_{\{t_0^i \le t < \tau_C^i \le t+\tau,\ \tau_D^i = \tau_C^i\}}\, L_t(\tau_C^i;\ \tau_C^i \le t+\tau,\ \tau_D^i = \tau_C^i) \\
&+1_{\{t_0^i \le t < \tau_C^i \le t+\tau,\ \tau_D^i \ne \tau_C^i\}}\, L_t(\tau_C^i;\ \tau_C^i \le t+\tau,\ \tau_D^i \ne \tau_C^i)
\end{aligned}
$$

where t_0^i is the first month that it appeared in the data set, and the parameters α and β describe the forward default intensity $f_t^i(\tau)$ and forward other exit intensity $g_t^i(\tau) - f_t^i(\tau)$, respectively. The first factor assigns the survival probability beyond both default and non-default exits for firm i at time t, the second factor gives the instantaneous default probability at t, and the third factor gives the probability of exit for reasons other than default.

This calibration scheme represents the full pseudo-likelihood function as the product of two multiplicative components, one involving the forward-default parameters and the other involving the forward-other-exit parameters α and β.

In addition, the pseudo-likelihood function for α or β can be further decomposed into separate terms involving $\alpha(\tau)$ or $\beta(\tau)$ across different τ's. The pseudo-likelihood function for either the forward-default parameters or the forward-other-exit parameters decompose into factors indexed by prediction horizon. For a fixed horizon, the corresponding pseudo-likelihood takes the form:

$$
\begin{aligned}
L_s^{\alpha}(\alpha;\tau_C,\tau_D,X) &= \prod_{i=1}^{N} \prod_{k=0}^{\lceil T - S/\Delta t\rceil - 1} L^{\alpha,i}_{s,k\Delta t}(\alpha(s);\tau_C^i,\tau_D^i,X^i) \\
L_s^{\beta}(\beta;\tau_C,\tau_D,X) &= \prod_{i=1}^{N} \prod_{k=0}^{\lceil T - S/\Delta t\rceil - 1} L^{\beta,i}_{s,k\Delta t}(\beta(s);\tau_C^i,\tau_D^i,X^i)
\end{aligned}
$$

for $s = 0, \Delta t, 2\Delta t, \cdots, \tau - \Delta t$ as intended prediction horizons measured in months.

The uncertainty is modeled by a complete probability space (Ω, F, P), where Pdenotes the physical probability measure. It represents information flow by the right-continuous, complete filtration $F = (F_t)_{t \ge 0}$. Expectation conditional on F_t is denoted by $E_t(\cdot)$. The processes λ_t^i and ϕ_t^i need not satisfy independence; each adapts to the filtration F. When both are expressed as functions of common stochastic covariates, statistical dependence naturally arises. Conditioning on λ_t^i and ϕ_t^i. restores independence between the intensity processes N_t^i and M_t^i.