Information Noise and Credit Risk: Evidence from Corporate Bankruptcy

Credit Scoring and Credit Control Conference XV

Xin Xu
Unisys Advanced Data Analytics
xin.xu@unisys.com
Paper In A Nutshell

• Current practice in bankruptcy/default prediction
  \[\log(\lambda_t) = \beta'X_t \quad \text{or} \quad \logit(\lambda_t) = \beta'X_t\]
  where \(\lambda_t\) is hazard rate, \(X_t\) is time-varying covariates, \(\beta\) is coefficients

• When there is information noise in input data, theory of Duffie and Lando [2001] (DL) implies RHS should be nonlinear and approximated by interaction effects

• Empirical implications of those noise-induced interactions
  - Existence of the interactions and improvements in Goodness-of-Fit
  - Improved out-of-sample forecasting accuracy of models augmented by the interactions
  - Insights into “Are larger, or more transparent, firms less likely to fail? Are they more predictable?”
  - Reconciliation of puzzling empirical results in the literature

• Using over two million firm-months panel of North American firms during 1979-2012, we demonstrate the existence of the interactions, and strong evidence supporting their implications
Theory

Problem
• There is a noise in the accounting reports of a debt issuer (firm), which is
  – Associated with log(asset)
  – Normally distributed, $\sim N(u, a^2)$ *
• Only noisy value of assets is observed by creditors/modellers
• Firm files bankruptcy (or defaults) when true log(asset) first falls below a low boundary, $v$
• $a$ is “a measure of the degree of noise” (DL)

Solution
• A filtering problem: Conditional on the observed (noisy) asset value, what is the conditional Probability of Bankruptcy/Default (PB/PD)?
• DL provided an analytical solution to this problem
• What does the theoretical PB/PD look like?

* We assume unbiased accounting reports ($u = -\frac{a^2}{2}$) throughout the paper
Two Dimensions, \( PB-\alpha \)

The plot when the observed (noisy) asset growth rate is 0

- Is this the whole picture? What happens if the observed (noisy) asset growth rate is non-zero?
- Is (1-year) PB/PD always increasing in \( \alpha \)?

Source: Figure 4 in Duffie and Lando [2001, Econometrica], with permissions from Econometrica and the authors
Three Dimensions, $PB-a-r_N$

**Bayes rule:** the higher degree of noise, the less responsiveness of observed asset growth rate to PB/PD, and vice versa

Similar patterns for any monotonic transformation (e.g. log or logit) of PB/PD

Plot on $PB-a-r_N$ and projections on $PB-r_N$
Testable Hypotheses (1/2)

PB can be approximated by hazard rate (see DL for existence of the limit)

$$\lambda_t = \lim_{{\Delta t \to 0}} \frac{PB(t, t + \Delta t)}{\Delta t}$$

- **Hypothesis 1** (Existence of the noise-induced interaction effects):
  (i) We can explicitly predict the sign of coefficients on interaction effects, $\tilde{\gamma}_i$
  (ii) Improved Goodness-of-Fit

$$\log(\lambda_t) = \tilde{\beta}'X_t + \tilde{\gamma}_0\tilde{a} + \sum_{i=1}^{l} \tilde{\gamma}_i (\tilde{a} \ast X^i)$$

- Without loss of generality, $\tilde{a}$ is a proxy for the degree of noise, **such that $\tilde{a}$ is decreasing in $a$**
- If $X^i$ is increasing in $\lambda_t$, then $\tilde{\gamma}_i > 0$
- If $X^i$ is decreasing in $\lambda_t$, then $\tilde{\gamma}_i < 0$

- **Example:**
  - **Observed (noisy) asset return** is decreasing in $\lambda_t$
  - When interacting it with **firm size** (a proxy that is decreasing in $a$), the interaction effect should have a **negative** coefficient
Testable Hypotheses (2/2)

• **Hypothesis 2** (Out-of-Sample Forecasting Accuracy): Out-of-sample forecasting accuracy is improved by models augmented by the interactions, compared to those without

• **Hypothesis 3** (Relation of the degree of noise to credit risk and predictability):
  – Are larger, or more transparent, firms less likely to fail? NO
  – Are they more predictable? YES

• **Hypothesis 4** (Reconciliation of the puzzling empirical findings about RSIZE and NI/TA): In hazard models estimated from different types of firms, the partial effects of RSIZE may have different signs, and the significance of NI/TA’s partial effects may vary, all else equal.
Empirical Design

- This paper uses Cox [1972] Proportional Hazard Model to model hazard rate $\lambda_t$

- Identify time-varying covariates, $X_t$
  - From well-known, well-accepted models (reference models) in the literature
  - Main results: Shumway [2001, JoB] (S01 Model)
  - Robustness checks
    - Chava and Jarrow [2004, RoF] (CJ04 Model)
    - Simplified version of Duffie, Saita and Wang [2007, JFE] (DSW07-S Model)
    - Bharath and Shumway [2008, RFS] (BS08 Model)

- Proxy for the degree of noise, $\bar{\alpha}$, that is decreasing in $\alpha$
  - Well-accepted proxies from the Finance literature
  - Absolute firm size: log[Total Assets] (log(TA)), log[Equity] (logE)
  - Relative firm size: log[Asset Rank] (log(AR)), RISZE from S01 Model
  - Analyst coverage (AC)
  - Analysts’ forecast variation (-log(CV))
Our Bankruptcy Panel Dataset (1979-2012)

- 2.15M firm-months, 2,112 bankruptcies
- from a total of 20,180 North American public firms

Bankruptcy (Response Variables)
- New Generation
  - www.BankruptcyData.com

Independent Variables
- UCLA-LoPucki
- Mergent FISD
- Compustat deletion reason “2”
- IBES (analyst forecasts)
- Datastream (3m T-rate)
- Compustat
  - Quarterly / Annual
- CRSP
  - Monthly
- CRSP
  - Quarterly / Annual

Quarterly / Annual

Monthly

Empirical Results (1/3)


Conclusions (details in paper)

• Hypothesis 1: Existence of interactions by in-sample tests
  – The signs of coefficients on the interaction effects are consistent with the hypothesis
  – Likelihood Ratio tests and AIC demonstrate significant improvements in Goodness-of-Fit

• Hypothesis 2: Out-of-sample forecasting accuracy in the 10-year holdout periods
  – Main results: measured by Area Under ROC Curve (AUC) and using all the proxies for a, the interactions bring notable AUC uplifts, in both statistical significance and economic magnitude
    ▪ Scenario 1: data is not winsorized
    ▪ Scenario 2: data is winsorized
    ▪ Scenario 3: information is noisier, i.e., less frequently updated financial reporting, no stock market information, fewer covariates, more outliers
  – Robustness checks: confirmed improvements by another measure, “captured bankruptcies within deciles”
Empirical Results (2/3)

• Hypothesis 3: Relations of the degree of noise to hazard rate and predictability
  – When firms’ return-related variables (NI/TA or EXRET) are “substantially negative”, higher firm size robustly entails higher hazard rate, contrasting with common intuitions (e.g. Shumway [2001])
  
  Similar results are obtained using other proxies for the degree of noise, implying more transparent firms entail higher credit risk. This is in contrast with discretionary disclosure theories (e.g. references in Yu [2005])

  – The above relations are inverted when firms’ return-related variables become less negative or positive

  – Demonstrated using all the proxies, a single model has higher discriminative power (out-of-sample AUC) in sub-sample with lower degree of noise

• Hypothesis 4: Reconciliation of puzzling empirical results
  – Contradictory empirical findings in the literature about the sign of partial effects of RSIZE, and significance of NI/TA

  – Controlling for sample periods and model specifications, we replicate the findings in different MECE segments of a single dataset, and attribute them to the interactions
Empirical Results (3/3)

• Robustness checks
  – Find very similar results in **sub-periods**
  – Re-run all the tests based on **three alternative reference models**, and obtain similar results, both quantitatively and qualitatively
  – Confirm that the out-of-sample results are robust to **different sample sizes**
  – Test all the hypotheses based on **non-financial firms**, and draw the same conclusions
  – Examine and confirm out-of-sample results based on an **alternative measure of forecasting accuracy**
    ▪ Augmented models capture more bankruptcies in top 2 or 3 deciles, with CAP curve being no lower than that of reference models in all deciles
    ▪ Within the low-risk deciles, augmented models capture notably less bankruptcies, leading to less potential misclassification
  – Re-run the empirical tests and find qualitatively similar results, relating to the proxies AC and -log(CV), by **imputing the missing values** and including firms with no IBES information
Contributions

• New tests on DL theory from the perspective of PD/PB, which are cleaner and more straightforward

• Contributions to the literature of corporate bankruptcy/default prediction
  – Introduce new and simple ways of taking into account information noise, which improve hazard models’ empirical performance
  – Shed light on the causal effects of the covariates that lead to inconsistent empirical findings

• Highlighting the importance of the noise-induced interactions in empirical research, by demonstrating that intuitions without considering the interactions’ implications may lead to spurious conclusions

• New insights into the predictability of credit risk, offering an interesting perspective and deeper understanding on the causal effects of predictability

• Industry applications on credit rating models: applicable to improving the accuracy and robustness of PD models
Future Work

This paper opens a number of opportunities for future empirical research

• A starting point to study the complicated relations of credit risk to asset pricing, induced by information quality

• Approaches to investigate the degree of information transparency and asymmetry, between creditors and firms, and its impact on credit spreads and credit markets

• Study of different forms of incomplete information, like biased accounting reports or delayed information, or different proxies for the degree of noise

• Tests using default events and for general firm types (e.g. non-public firms)
Acknowledgement

I Thank Rodrigo Fontecilla and Unisys Analytics for support and generous sponsorship for this presentation.

I also thank discussant and attendees of Credit Scoring and Credit Control Conference XV for their helpful comments.

Unisys Analytics offers Advanced Data Analytics using Big Data Ecosystem (Hadoop, Spark, …)

- Machine Learning as a Service
- Advanced Analytics Consulting
- Data Scientist on Demand
- Artificial Intelligence applications
- Expertise in Credit Modelling and Financial Risk/Compliance Modelling
- Global expertise in Financial Services industry