ASEAN +3 MACROECONOMIC RESEARCH OFFICE

Working Paper (WP/24-06)

Monitoring Privately-held Firms' Default Risk in Real Time: A Signal-Knowledge Transfer Learning Model

Jorge A. Chan-Lau, Ruogei Hu, Luca Mungo, Ritong Qu, Weining Xin, and Cheng Zhong

June 2024

[This page is intentionally left blank]

Monitoring Privately-held Firms' Default Risk in Real Time: A Signal-Knowledge Transfer Learning Model

Prepared by Jorge A. Chan-Lau, Ruofei Hu, Luca Mungo, Ritong Qu, Weining Xin, and Cheng Zhong[1 2 3]

Authorized by Li Lian Ong

June 2024

Abstract

We develop a mixed-frequency, tree-based, gradient-boosting model designed to assess the default risk of privately held firms in real time. The model uses data from publicly-traded companies to construct a probability of default (PD) function. This function integrates high-frequency, market-based, aggregate distress signals with low-frequency, firm-level financial ratios, and macroeconomic indicators. When provided with private firms' financial ratios, the model, which we name signal-knowledge transfer learning model (SKTL), transfers insights gained from 35 thousand publicly-traded firms to more than 4 million private-held ones and performs well as an ordinal measure of privately-held firms' default risk.

JEL classification: E43, E47, G33

Keywords: Default risk, corporate sector, privately held firm, gradient boosting, transfer learning

[1] Authors' e-mails: Chan-Lau, jchanlau@gmail.com and Jorge.Chan-Lau@amro-asia.org; Hu, hurfevan@gmail.com; Mungo, mungo@maths.ox.ac.uk; Qu, rqu@imf.org; Xin, WXin@imf.org; Zhong, CZhong@imf.org.

[2] Luca Mungo is affiliated with the INET, Oxford University; and Ritong Qu, Weining Xin, and Cheng Zhong are affiliated with the International Monetary Fund. The views expressed in the paper are those of the authors and do not necessarily reflect the views of the AMRO, INET or IMF.

[3] This paper is also published as IMF Working Paper WP/24/115. The authors would like to thank Alexander Copestake, Borja Gracia, Marco Gross, Srobona Mitra, and Laura Valderrama for useful comments. All errors and omissions are solely the authors' responsibility.

Abbreviations

AUROC	area under the receiver operating characteristic
NUS CRI	National University of Singapore Credit Risk Initiative
PD	probability of default
SKTL	Signals-Knowledge Transfer Learning

Contents

Abbreviations ii

I. Introduction 1
2 Literature Review 3
3 Data: Sources and Transformations 5
4 From Public Firms' PDs to Private Firms' Financial Conditions 9
5 SKTL: A Mixed Frequency Gradient-Boosting Tree PD Model- 11
6 Case Studies 19
7 Conclusion 21

References 22

Figures

Figure 1. Training, validation and testing sets in cross-validation and out-of-sample examination of model performance 31
Figure 2. MSE of main and benchmark model at different lags since publicatoin of annual reports 32
Figure 3. AUC of predicting non-listed firm bankruptcy..33
Figure 4. Distribution of Shapley values of main models applied to privately held firms.......34
Figure 5. Distribution of Shapley values of main models applied to privately held firms (continued)..35
Figure 6:. Case suty 1: German gas and electricity firms during the post-covid energy crisis 36
Figure 7:.Case study 2: UK commercial real estate firms 37

Appendix Figures

Figure A.1: MSE of main and benchmark model at different lags since publication of annual reports..40
Figure A.2: AUC of predicting non-listed firm bankruptcy (15 month publication lag)...........41

Tables

Table 1: List of predictors 26
Table 2: Summary statistics 27
Table 3: Orbis bankruptcy events and NUS-CRI credit events 28
Table 4: Intra-year prediction regression of sectoral financial ratios against sectoral PD.....28
Table 5: R^2 of Logit PD, assumiing balance sheets are published 3 months after closing dates 29
Table 6:Loss diferential decomposition of predicting Logit PD, conditional level of average PD 30

Appendix Tables

Table A.1: R^2 of Logit PD, assumiing balance sheets are published 15 months after closing dates 38
Table A.2: Loss differential decomposition, conditional level of average PD (15 months publication lag) 39

[This page is intentionally left blank]

1 Introduction

The reliable assessment of a firm's default risk is an important task of credit analysts, loan officers, auditors, and risk managers. The evaluation determines loan terms, screens insolvent and high-risk borrowers, and helps to manage credit risks adequately. Real-time default risk evaluation becomes critical when market conditions deteriorate rapidly, as during crisis episodes and economic recessions. Even in the absence of a crisis or recession, financial conditions can deteriorate rapidly. For instance, in the post-COVID-19 pandemic period, the threat of high inflation has prompted central banks to raise rates rapidly, increasing funding and borrowing costs. This situation threatens the viability of the non-financial corporate sector, where debt levels have doubled as borrowers benefited from the low-interest rate environment post the 2008 global financial crisis through the pandemic (Adasoro et al., 2021).

In isolation, corporate default crises do not impact the real economy as severely as banking crises (Giesecke et al., 2014). Nevertheless, widespread defaults in the corporate sector may distress the banking sector and pose risks to financial stability. Therefore, timely surveillance of default risk at the firm level should be a key element in macro-financial surveillance.

Assessing the default risk of exchange-listed and publicly traded firms is easy due to the abundance of market data. Currently, several commercial firms provide probability of default (PD) estimates for these firms using the market prices of their traded securities and balance sheet information from their corporate reports. In contrast, it is costly and difficult to get credit ratings or credit risk assessments for privately-held, non-listed firms. Furthermore, when ratings and assessments are available, they are updated infrequently and with significant delays. This makes them unsuitable for monitoring risk in real-time.

This is an unfortunate situation, as the importance of privately held firms dwarfs that of publicly traded firms. Over 99% of firms in Europe and the US are not listed in stock markets but account for a significant fraction of the economy (Anderson, 2009; Kalemli-Ozcan et al., 2020). There is also evidence that exchange-listed firms are becoming less important to the US economy, with shrinking contributions to employment and GDP (Schlingemann and Stulz, 2022), a trend likely present in other economies. In the context of corporate credit risk and its macroeconomic and financial stability implications, the risks posed by privately-held firms are comparable to those posed by publicly-traded firms .[1]

Two questions arise naturally. Can market signals and models fitted to a small set of publicly-traded firms offer insightful information about the default risk of the larger segment of privately-held firms? And if this is the case, are these models suitable for real-time risk monitoring?

This paper addresses both questions by introducing the Signals-Knowledge Transfer Learning (SKTL) Model. Within this model, "signals" refers to the real-time financial condition of publicly

[1]For 13 out of 20 economies covered in the paper, the aggregate total assets of privately-held firms exceed that of publicly-listed firms.

traded companies, as summarized by their market-price-based probability of default. “Knowledge” indicates the functional form that maps publicly traded firms’ features to their PDs. “Transfer learning” is about applying this method to privately held firms. Thus, the paper contributes to the literature on estimating PDs for privately held firms in two significant ways. First, the approach computes an “average” probability of default for each industry sector, using the market-based PDs from firms within that sector. This average is then used to predict the default probability of privately held firms, enhancing the real-time evaluation of their financial distress. Second, it transfers the functional mapping from publicly traded firms’ features to PD to the domain of privately held firms, allowing the estimation of PD for privately held firms without the need for their default data.

The SKTL model is founded on the premise that, despite the differences between publicly traded and privately held firms, despite differences across publicly traded and privately held firms, the functional form connecting firms’ features to their default probabilities is likely to be similar. This similarity suggests that changes in financial condition indicators, such as balance sheet variables, affect the risk of default in a comparable manner for both types of firms. For example, an increase in leverage signals rising financial distress for both publicly traded and privately held firms. Understanding the PD function for publicly traded firms will, therefore, help us understand privately held firms’ PD. This concept draws parallels with transfer learning in machine learning, where knowledge acquired from addressing one problem is repurposed for solving a different, albeit related, problem (Weiss et al., 2016).

The SKTL model includes as covariates: macroeconomic variables; “average”, sector-specific PDs for publicly traded firms (from now on, sectoral PDs); and firm-specific features available for publicly and privately held firms. These firm-specific features, usually derived from balance sheet data and income statements, are standard in models predicting corporate default. However, they are often reported with a delay and only updated annually or semi-annually. By including both high-frequency macroeconomic variables and aggregate sector PDs, the model allows for a more timely assessment of a firm’s default risk. This is because firms’ financial health is influenced by broader economic conditions and sector-wide risk trends, making timely updates crucial for accurate risk evaluation. The importance of including sectoral PDs cannot be overstated, as mid-year observations serve to “nowcast” year-on-year changes of the financial ratios of privately held firms.[2] Different data frequency requires setting up the PD model as a mixed-frequency tree-based gradient boosting model. We select the LightGBM model (Ke et al., 2017a) because of its robust performance in the panel setting (Barboza et al., 2017).

Arguably, the PDs of privately held firms generated by the model might be biased vis-a-vis publicly traded firms since they have restricted access to equity financing. The PDs, however, can still serve as an ordinal measure of firms’ financial distress, and allow us to compare the level of

[2]Industry or sectoral effects are important for predicting bankruptcy - see, e.g., Chava and Jarrow (2004).

financial distress across firms and periods. Also, the SKTL model could be adversely affected if a distribution shift is present in the data – conditional on a firm's features, the target conditional distribution of the PDs of privately held firms is different from that of publicly traded firms.[3]

To assess whether distribution shift is a problem, we construct a dataset of privately held firm default events covering eleven advanced economies – France, Germany, Hong Kong, Italy, Japan, Korea, Netherlands, Singapore, Spain, the UK, and the US – and nine emerging market economies – Brazil, China, India, Malaysia, Philippines, Poland, Thailand, Russia and Vietnam.[4] The good performance of the SKTL model on this dataset enables us to generate real-time PD estimates for more than three million privately held firms, up from the thirty thousand publicly traded firms used to estimate the model.

The remainder of the paper is organized as follows. Section 2 reviews the relevant academic literature, providing a context for this paper's contribution. Section 3 introduces data used for the analysis. Section 4 motivates sectoral PDs as predictors by examining how market-based publicly-traded firm PD can help predict key measures of privately-traded firms' profitability, liquidity, and solvency. Section 5 describes the model estimation and evaluation. Section 6 shows two case studies and evaluates the model's real-time performance. Section 7 concludes.

2 Literature Review

The paper is related to the academic literature covering corporate default prediction and risk estimation. Modern modeling approaches can be grouped into two categories: structural risk models and reduced-form models.[5] The structural risk approach builds on the insight, first noted by Merton (1974) and Black and Scholes (1973), that the two main components of the capital structure of a firm, debt and equity, are derivative securities of the underlying asset value of the firm. One key risk measure derived in structural credit risk models is the Distance-to-Default (DD), a volatility-adjusted leverage ratio that quantifies the buffer available to a firm over what it owes to its creditors.[6]

The reduced-form, or intensity-based, approach models a firm's default rate as a function of both firm-specific and broader economic variables. Early examples of these models include Jarrow and Turnbull (1995), Madan and Unal (1998), Lando (1998), and Duffie and Singleton (1999). Reduced-form models are numerically tractable since they resemble dynamic term structure models closely.

[3]See Storkey (2009) and Ben-David et al. (2010) for a discussion of distribution shifts and their implications for model performance.

[4]The number of default events is not large enough to estimate a model specifically for privately held firms.

[5]Bakshi et al. (2022) provide an up-to-date survey of modeling approaches and estimation methods.

[6]The DD is widely used by several risk providers, such as Moody Analytics (Crosbie and Bohn, 2002) and the National University of Singapore's Credit Research Initiative (Duan and Wang, 1994). Since the economic asset value of the firm is a latent variable, several approaches have been proposed to estimate it using equity prices and balance-sheet data, such as maximum likelihood (Duan, 1994), and iterative methods (Vassalou and Xing, 2004). Once asset values are estimated, it becomes possible to calculate the DD.

Compared to structural models, they can incorporate a larger number of covariates. However, they often lack economic interpretability as, by design, these models prioritize statistical fit over theoretical grounding.

Early work on corporate failure prediction, such as the seminal papers of Beaver (1966) and Altman (1968), proposed models relying on accounting-based financial ratios to evaluate the likelihood of failure of a firm. In contrast, modern default risk models, either structural or reduced, are estimated using both accounting-based data and the market prices of the securities issued by the firm.[7] Several studies, among others Hillegeist et al. (2004), Bharath and Shumway (2008), Duffie et al. (2007) and Campbell et al. (2008), have shown that models that use both market and accounting data outperform those that only use the latter.

Advances in credit risk modeling have predominantly centered on publicly traded firms, either because models had to be calibrated using market-based data or because the goal was to replicate the observed prices of the securities issued by these entities. Research in modeling the default risk of privately held firms has not progressed at the same pace. The two main modeling approaches discussed earlier are not feasible without market-based information unless modelers adopt some market price proxies. Moreover, the application of accounting-based models faces its own set of hurdles. Data from privately-held firms tend to be reported less frequently, suffer from significant delays, and lack the breadth and depth of information compared to their publicly traded counterparts.

To compensate for the data shortcomings, several academic studies and industry reports covering privately-held firms rely on supervisory or proprietary datasets. For example, Falkenstein et al. (2000) introduced Moody's RiskCalc default model and calibrated it using a proprietary financial statement dataset covering firms in Canada and the US. One major finding was that the relationship between financial variables and default risk could be substantially different between publicly traded and privately held firms, suggesting that a distribution shift might invalidate extending models first calibrated to publicly traded firms to privately-held firms. Zhou et al. (2006), building on and extending earlier work by Cangemi et al. (2003) for French firms, estimated private firm default probabilities for North American firms using maximum expected utility models derived from an ℓ_1 regularized maximum likelihood problem and calibrated using financial ratios, economic indicators, and industry market prices from a Standard and Poor's dataset.

Bhimani et al. (2010) used data collected by the central bank of Portugal to estimate a logit default prediction model. In addition to accounting ratios, the model included as covariates the industry classification, geographic location, and firm size. The results suggest that the information gathered only from accounting ratios is insufficient to capture default risk. Supporting this finding, Dierkes et al. (2013) found that, for German privately-held firms, the accuracy of default risk

[7] While possible, incorporating market-based data in structural models is more difficult than in reduced-form models. See, for instance, Miao et al. (2018), where the authors use forward-looking option prices to estimate DD measures.

prediction models improved substantially if covariates included credit registry information on the creditworthiness, order book, and payment history of the firm. Charalambakis and Garrett (2019) found that a multi-period logit model including financial statement and income-based covariates, real GDP growth, and export dummies could predict well the probability of financial distress of a large dataset of Greek firms over short and long-time horizons. Altman (2013) showed that the original Altman Z-score model, presented in Altman (1968), could be extended to privately-held firms by estimating the model after substituting the book value of the firm for its market value.

Finally, this paper's transfer-knowledge approach follows the spirit of Duan et al. (2018). The authors estimated a mapping from financial ratios to the Distance-to-Default in the former set of firms before applying to the latter using a proprietary dataset of Korean privately held firms. The map allowed them to estimate the default probability's term structure using a forward intensity model including balance sheet variables and macro-risk factors. The forward intensity model outperformed alternative models including Logit, Probit, and Altman's Z-score. The accuracy ratio, at 0.5, was slightly better than random guessing.

3 Data: Sources and Transformations

Estimating the SKTL model requires firm-level balance sheet data and publicly traded firms' PDs. We collected balance sheet data from the Orbis dataset and PDs from the Credit Research Initiative, National University of Singapore (NUS-CRI). To account for the impact of business cycles on the financial conditions of the firms we include GDP growth rates and inter-bank interest rates from Datastream and Federal Reserve Economic Data (FRED). To assess the SKTL model's performance in the domain of privately held firms, we construct a bankruptcy dataset based on Orbis firm status data. The dataset, described below, covers firms in eleven advanced economies – France, Germany, Hong Kong, Italy, Japan, Korea, Netherlands, Singapore, Spain, the UK, and the US – and nine emerging market economies – Brazil, China, India, Malaysia, Philippines, Poland, Thailand, Russia, and Vietnam.

3.1 Orbis Balance Sheet Data

The firm-level balance sheet data, reported at annual frequency, are sourced from the historical financial dataset of the Orbis database, which links several data vintages. For advanced economies, we use data from 1995 onwards to match the data availability in the CRI-PD dataset. For emerging markets, the Orbis data start at later dates, ranging from 1997 to 2005. Firms with total assets below USD 1 million are dropped for three reasons: first, Bajgar et al. (2020) shows that Orbis' smaller firms coverage is not representative of the actual small firms' distribution; second, Altman et al. (2017) shows that financial ratios of small firms tend to be unstable and not suitable for financial distress prediction; third, the SKTL model's calibration is based on data from publicly-traded firms, which are, on average, significantly larger than privately-held firms. This discrepancy suggests that the model may be more accurately applicable or effective among larger privately-held firms.

We clean the balance-sheet observations following the procedure proposed by Kalemli-Ozcan et al. (2015):

1. We exclude firms if any of the following conditions are satisfied:
   (a) any of their recorded sales, total assets, or total fixed assets are negative;
   (b) any of the following balance sheet items are missing: total assets, shareholders' funds, current liabilities, noncurrent liabilities, operating revenue, and net income;
   (c) the observation is an outlier outside the 99.9% percentile of sales-to-asset ratio, number of employee-to-asset ratio, and number of employee-to-sales ratio.
2. We drop duplicate records of a firm by eliminating observations from the unconsolidated account when the consolidated account is available at the year;
3. To further remove duplicated data, for each unique Orbis identifier (BvD ID) we select only the last observation in each report year. The report year of a balance sheet is its fiscal year when the closing date is after June. Otherwise, the report year of a balance sheet is the preceding fiscal year.

Table 1 (panels A, B, and C) presents the list of firm characteristics and financial ratios retrieved and computed using Orbis data. The financial ratios are grouped into three broad categories: liquidity, profitability, and solvency. In addition to financial ratios, we also include the number of employees, total assets in constant price (2015 USD), and industry sector (NACE Rev. 2 main section) as predictors. To reduce the number of outliers, only firms' observations for which the denominator in their financial ratio exceeds 100 USD are kept; missing values are assigned if the condition is violated. This measure also removes cases where the denominator is negative, which makes the resulting ratio uninterpretable, e.g., profit margin computed from negative operating revenue as the denominator.

Table 2 presents the summary statistics of the key firm's features collected from the Orbis dataset. After cleaning and filtering data, the sample includes more than 4 million firms, of which around 35 thousand are publicly traded. Despite its widespread usage in empirical studies, the representativeness of the dataset varies across countries due to differences in reporting requirements and firm coverage. For example, European countries require mandatory reporting to the national business registers, while the US does not. The coverage of privately held companies is visibly low in Brazil, Hong Kong, Netherlands, and the US, as demonstrated by significantly fewer privately held firms and significantly higher total assets, with medians exceeding 20 million USD. In these countries the sample might not be representative since, in general, privately held firms typically have a smaller scale compared to their publicly traded counterparts, with their median solvency ratios often being lower, indicating lesser equity financing and reduced capital buffers. However, regarding return on assets and liquidity ratios, there are no consistent patterns across countries to suggest whether privately held firms' balance sheets are stronger or weaker relative to publicly traded firms.

3.2 NUS-CRI Probability of Default dataset

NUS-CRI calculates PDs using the forward-intensity model of Duan et al. (2012). The model considers two independent doubly stochastic Poisson processes, one for delisting (exits) due to default events and the other for delisting due to other reasons. Model inputs include, at the economy-wide level, the stock index return, the short-term risk-free rate, and the median distance-to-default (DD) of the firms listed in the economy's stock exchange. At the firm level, the model inputs include the firm's distance-to-default, the firm's idiosyncratic volatility, the following ratios: cash to total assets, current assets to current liabilities, net income to total assets; and the relative size and market-to-book ratio of the firm with respect to the corresponding median market capitalization and market-to-book ratio in the economy.

The calibration of the NUS-CRI PD models requires data on credit default events, which NUS-CRI gathers from various international sources like Thomson Reuters Datastream and Bloomberg Backoffice License.[8] Due to limited credit default events in some economies with few listed firms (Table 3, column 3), calibrating models for individual economies is not statistically meaningful. Therefore, public companies worldwide are categorized into six calibration groups based on similarities in economic development stage and primary exchange location: North America, Europe, Asia-developed economies, Emerging Markets, China, and India.

Our study uses 1-year probabilities of default (PDs) sourced from the NUS-CRI Probability of Default (NUS-CRI PD) database, which provides time series of PDs for more than 85,000 publicly traded firms, both live and dead (delisted), in 134 countries. The 1-year NUS-CRI PDs are appropriate for the calibration of the SKTL model, as the prediction accuracy of the NUS-CRI PD model, measured as the area under the receiver operating characteristic (AUROC), is comparable to that of other commercial PD providers (Table 3, last column).[9] The SKTL model uses as sectoral predictors the 20, 50 (median), and 80 percentile of the cross-sectional distribution of the Logit PD

[8] The credit events comprise Bankruptcy filing, receivership, administration, liquidation or any other legal impasse to the timely settlement of interest and/or principal payments; A missed or delayed payment of interest and/or principal, excluding delayed payments made within a grace period; Debt restructuring/distressed exchange, in which debt holders are offered a new security or package or securities that result in a diminished financial obligation (e.g. a conversion of debt to equity, debt with lower coupon or par amount, debt with lower seniority, debt with longer maturity).

[9] The prediction performance is based on in-sample tests, using data until the end of the selected data sample and then comparing the forecasts to actual defaults occuring during the forecasting horizon starting immediately after the end point of the data sample.

in each of the sectors considered in the analysis (Table 1, Panel E). To reduce the noise induced by the PDs of individual firms, sectoral PDs are computed only for sectors with more than ten firms.

3.3 Orbis Bankruptcy Data

The SKTL model transfers a functional form, estimated for publicly traded firms, to privately held firms. It is natural to ask whether this transfer might be undermined by a distribution shift or inadequate representativeness of the data used in the model estimation. The quality of the bankruptcy data in Orbis, while not adequate for evaluating predictive default models directly as explained below, is nevertheless good enough to validate models calibrated using other data sources, such as the SKTL model presented here. Below, we describe how the bankruptcy data is constructed and explain in detail why the Orbis bankruptcy data should not be used to estimate a default predictive model directly.

Orbis collects data on the status of the firms it covers. The six major status categories are: *Active*, *Bankruptcy*, *Dissolved*, *In liquidation*, *Inactive* and *Unknown*. In some cases, status labels like *Active* and *Dissolved* are followed by more detailed descriptions, e.g., *Active (default of payment)* and *Dissolved (bankruptcy)*. Table 3, first column, shows the number of bankrupt firms, defined as firms with status labels "Bankruptcy", "Dissolved (bankruptcy)", "Active (insolvency proceedings)', or"Active (default of payment)' in our sample.

The Orbis bankruptcy data does not include the dates of bankruptcy or defaults. To create the model validation dataset, we assume that a firm becomes bankrupt or default during the year following its last observation in Orbis. To alleviate potential bias from firms that exit with "Active" status, we discard the last observations of firms with "Active" status.

There are several reasons to use Orbis' bankruptcy dataset only for validation purposes rather than to estimate directly a default prediction model. First, the data quality in certain countries is significantly scarce. For instance, in Brazil, out of 15,000 firms, only one bankruptcy is recorded, while in China, with 880,000 firms, there are merely 38 recorded bankruptcies. To assess the reliability of Orbis bankruptcy data, a comparison with NUS-CRI records of credit events among publicly traded firms is conducted (refer to Table 3, columns 2 and 3). Across most countries, Orbis documents fewer bankruptcies than NUS-CRI credit events among listed firms, indicating substantial omissions in Orbis bankruptcy records.

Second, Orbis is inadequate for studying firm entries and exits, including bankruptcy, because it is hard to tell if a firm has exited the market or simply left the dataset (Bajgar et al., 2020). Notably, a significant portion of firms for which Orbis has ceased updating their data, regardless of the reason for exiting, are still labeled as "Active".

Third, the historical bankruptcy data within Orbis are incomplete as bankruptcy information is deleted after five years of inactivity, resulting in a truncated data sample. Consequently, predictive models utilizing these incomplete data as input fail to capture the sectoral and cyclical factors responsible for the observed pattern of clustered defaults in the real world (Das et al., 2007).

3.4 Macroeconomic Variables

Besides accounting ratios, domestic macroeconomic variables have proved useful for predicting corporate defaults as in Carling et al. (2007), Duffie et al. (2007) and Koopman et al. (2012). We include two macroeconomic variables, quarterly GDP growth and monthly interbank interest rates. Quarterly GDP growth data are sourced from the World Economic Outlook database and interbank interest rate data from the Federal Reserve Bank of Saint Louis FRED database. Since we focus on real-time prediction of PDs, we assume that the quarterly GDP growth is published one month after the quarter ends and that interbank interest rates are available at the end of each month.

4 From Public Firms' PDs to Private Firms' Financial Conditions

One of the contributions of this paper is to exploit the information embedded in the market-based PDs of publicly traded firms, which is available at a high frequency, e.g. daily, weekly, and monthly, to evaluate the PDs of privately held firms. The underlying intuition is simple: sectoral shocks should affect the majority of firms in the sector regardless of whether they are privately held or publicly traded. On the other hand, it is possible that publicly traded firms, with easier access to equity financing and much larger size, may be so different from privately held firms that the market-based information of the former is irrelevant for the latter.

One could test this empirical question by running a logit regression of privately-held firms' credit incidents against lagged listed firms' PD. However, this exercise is not feasible due to the deficiency of Orbis bankruptcy data discussed in Sec. 3.3 so an indirect approach is required to assess whether the PDs of publicly traded firms are informative enough about the financial conditions of privately held firms. Specifically, we evaluate whether changes in selected financial conditions of the latter can be explained by an intra-year predictive regression model that includes changes in the sectoral PDs of publicly traded firms as covariates. The financial conditions selected are solvency, profitability, and liquidity, which are proxied as shareholders' fund-to-total assets ratio (solvency ratio), net income-to-total assets ratio (return on assets), and current assets minus inventory to current liability ratio, or current ratio (liquidity ratio).[10]

We estimate panel regressions with economy-sector fixed effects. Let $\Delta y_{cs,t}$ denote the median change in privately held firms' financial ratio of year t relative to the previous year, within sector s of economy c; let $\Delta PD_{cs,t+0.25k}$ denotes the median change in publicly traded firms' PD observed at end of quarter k of year $t+1$, relative to the end of year t. To facilitate interpretation, $\Delta PD_{cs,t+0.25k}$ is scaled to unit variance within each combination of economy, sector, and quarter. At the end of quarter k in year $t+1$, we evaluate whether the latest sectoral PD available can forecast or

[10]The three financial ratios are similar to the ones used by Zmijewski (1984) and Shumway (2001).

nowcast changes in end-of-year financial ratios of privately held firms in the sector. For each quarter $k \in \{1, 2, 3, 4\}$, the panel regression is:

$$\Delta y_{cs,t+1} = \alpha_{csk} + \beta_k \Delta PD_{cs,t+0.25k} + \boldsymbol{\gamma_k} \cdot \text{controls} + \varepsilon_{cs,t+1|t+0.25k} \quad (t = 1, 2, ..., T), \tag{1}$$

where the control variables include country c's real GDP growth rate, the average intra-bank interest rates in years t and $t+1$, and the end-of-year sectoral PD and sectoral median of the financial ratio in year t.

The panel regressions allow examining the sectoral-specific signals from publicly traded firms' PDs after controlling for the cyclical information contained in the macroeconomic variables. Note also that the macroeconomic variables, GDP growth, and interest rates, enjoy a forward-looking advantage over the sectoral PDs since their current values, not available at the time of the prediction, are included in the regression. For comparison, we also examine the ability of sectoral PDs of publicly traded firms to predict changes in their financial conditions. To achieve this, we substitute the dependent variable in equation (1) with the median changes of sectoral PDs of publicly traded firms.

Table 4 presents estimates of the slope coefficients, β_k, in equation (1). For each financial ratio, the table reports the slope coefficients corresponding to the regression equations of the financial ratios of the publicly traded firm and privately held firm financial ratios. To interpret the economic significance of coefficients, we scaled the independent variables to unit variance at the country-sectoral level. The coefficients can therefore be interpreted as the response to a one standard deviation shock of sectoral PD.

The analysis uncovers several noteworthy patterns. First, a rise in the sectoral PD of publicly traded firms predicts decreased solvency ratios, return on assets (ROA), and liquidity ratios of all firms, regardless of their listing status. Second, the impact of sectoral PDs as information from latter quarters is used, i.e. β_k increases with k. For example, a one-standard-deviation shock to end-of-year sectoral PD, $k = 3, 4$, would predict lower expectation of current-year nonlisted-firm ROA by about 0.3 percent whereas the same shock for earlier quarters, $k = 1, 2$, would only lower it by 0.2 percent. Third, the magnitudes of slope coefficients are generally higher among privately held firms than publicly traded firms. This is in line with the fact that privately held firms, due to their smaller size, are more susceptible to macroeconomic shocks as observed in Fama and French (1993). The sectoral PDs of the listed firms, therefore, are informative for both groups of firms, which validates the signal transfer approach of the SKTL model.

In summary, the sectoral PD of publicly traded firms contains useful and timely information for monitoring privately held firms' financial conditions, whose reaction to PD innovations is larger than that of publicly traded firms. Including this high-frequency, real-time data in predictive models could significantly improve their performance.

5 SKTL: A Mixed-Frequency Gradient-Boosting Tree PD Model

The absence of reliable publicly available default and bankruptcy data for privately held firms required earlier studies to use proprietary datasets. Drawing from the findings outlined in Section 4, we implement a Signal-Knowledge Transfer Learning (SKTL) model. This model uses information obtained from the PDs of publicly traded firms, addressing the issues posed by data scarcity. In this section, we begin with a discussion of the model setup, followed by an overview of the machine learning algorithm the model employs. Subsequently, we detail the steps involved in model estimation and evaluation. Finally, we present our results.

5.1 Mixed-frequency Model Setup

A combination of mixed-frequency data is essential to monitor real-time financial conditions. Firms' balance sheet data are typically published annually, sectoral PD and interest rates are available at the end of each month, and GDP figures are released quarterly, with a one-month lag. We assume privately held firms' balance sheets are released three months after the closing date of each fiscal year. The assumption mirrors the timeline an investor or market analyst might encounter when seeking balance sheet data online.[11]

Before delving into the model, we introduce some notations. Let i denote firm index; $s(i)$ denote the sector that firm i belongs to; t denote the end of the month observed; l denote the number of months that have passed since the release date of most recent annual balance sheet. Let $Firm_{i,t-l}$, $Macro_t$ and $Sectoral_{s(i),t}$ denote the most recent set of firm features, macroeconomic variables, and sectoral PD predictors available at time t, which are defined in Table 1. We adopt a modular approach, consistent with that used in the NUS-CRI PD model, and calibrate models for each group of economies consistent with the group specification in NUS-CRI PD.

We aim to generate end-of-month PD using the available predictors at that time. This implies that the frequency of predictors is either equal to or lower than the target variable. Following Foroni et al. (2018), the forecasting problem is formulated as:

$$\text{Logit}\left(PD_{it}\right) = F_{cl}\left(Macro_t, Sectoral_{s(i),t}, Firm_{i,t-l}\right) + \varepsilon_{i,t|t-l}. \tag{2}$$

Hence, for each l in $\{0,1,2,...,11\}$, we have a different model F_{cl} to account for the degree of staleness of the balance sheet data. As time elapses from the release date, balance sheets become outdated to reflect the firm's financial condition. The functional form of F_{cl} can account for the data staleness by placing less weight on the balance sheet variables. To simplify the model estimation, F_{cl} only accommodates the release frequency of balance sheets but not GDP growth. The latter aligns more closely with the PD data, lagging only three months, thus mitigating the issue of staleness.

[11]Note, however, that Orbis might take one additional year to add and report the balance sheet data for these firms. The case studies in Section 6 report results from models that assume a 15-month publication lag.

The choice of the logit transformation of the PD as the target of F_{cl} ensures that the implied PD, once the logit is reversed, falls in the $[0,1]$ interval. Another desirable feature is that it also magnifies the sensitivity of the squared loss function for small values of the PD, which is consistent with the fact PD ranges from one credit rating to the next are wider as the credit rating of a firm deteriorates (Duan and Li, 2021). For example, the range of AA+ to AA- rating is 0.40 bps wide, in contrast, the range of BB+ to BB- rating is 72 bps wide. On the other hand, when the Minimum Squared Error (MSE) loss function is used, the logit transformation might exaggerate the economic significance of errors when the target PD is too small, as the logit function tends to infinity. Since Duan and Li (2021) suggest that the AAA rating corresponds to PD below $3.5 \cdot 10^{-5}$, we truncate the sample one-year PD by assigning 10^{-5} to all values below that threshold to minimize this problem.

The remainder of this section explains the next steps in the implementation of the SKTL model. First, the choice of the Gradient-Boosting Tree as the functional form for F_{cl}. Second, the training of the model to predict the default probabilities of publicly traded firms using predictors also available for privately held firms, namely balance sheet data, market-based sectoral PDs, and macroeconomic indicators. Third, the evaluation of the model performance is conditional on the systemic credit cycle and the time lags since the release of balance sheet data. Finally, the extension of the model to predict the PDs of privately held firms, and its validation with the bankruptcy data from Orbis.

5.2 Gradient-boosting Tree and its Hyper-parameters

We use a gradient-boosting tree as the functional form of F_{cl}. Tree-based ensemble methods, i.e., ensemble methods that use regression trees or classification trees (Breiman et al., 1984; Sutton, 2005) as base or weak learners, are among the best algorithms for regression and classification on tabular data (Grinsztajn et al., 2022). The Gradient Boosting (Friedman, 2001) algorithm is one of the most popular tree-based ensemble methods and the one chosen as the functional form for F_{cl}.

The fundamental idea in boosting is to train a series of base learners sequentially, where each learner attempts to rectify the mistakes of those preceding it. Following Mungo et al. (2023), we start with a dataset of n samples and m features $\mathcal{D} = \{(\mathbf{x_i}, y_i)\}$ $(|\mathcal{D}| = n, \mathbf{x}_i \in \mathcal{R}^m, y_i \in \mathcal{R})$. A mapping function $y_i = \phi(\mathbf{x}_i)$ links inputs to the outputs. In Gradient Boosting, we aim to build an approximation $\phi_K(\mathbf{x}_i)$, expressed as the cumulative sum of K functions,

$$\hat{y}_i = \phi_K(\mathbf{x}_i) = \sum_{k=1}^{K} \rho_k f_k, \tag{3}$$

with $f_k = f(\mathbf{x}_i, \boldsymbol{\theta}_k)$ representing the ensemble's base learners, parametrized by $\boldsymbol{\theta}_k$. The construction of ϕ_K seeks to minimize the expected value of a loss function $\mathcal{L}(y_i, \hat{y}_i)$ and is built in K steps.

Initially, we compute a constant approximation as

$$\phi_0^* = \arg\min_{\alpha} \sum_{i=1}^{n} \mathcal{L}\left(y_i, \alpha\right). \tag{4}$$

Subsequent models are formulated in a sequence,

$$\phi_k = \phi_{k-1} + \rho_k f_k, \quad k = 1, \ldots, K \tag{5}$$

where ρ_k and f_k minimize

$$\{\rho_k, f_k\} = \arg\min_{\rho, f} \sum_{i=1}^{n} \mathcal{L}\left(y_i, \phi_{k-1} + \rho f\left(\mathbf{x}_i, \boldsymbol{\theta}\right)\right). \tag{6}$$

In an ideal scenario, to solve the minimization problem in equation (6), we would choose f_k as the negative gradient of the loss function,

$$f_k\left(\mathbf{x}_i\right) = -g_k\left(\mathbf{x}_i\right) = -\left[\frac{\partial \mathcal{L}\left(y_i, \phi\left(\mathbf{x}_i\right)\right)}{\partial \phi\left(\mathbf{x}_i\right)}\right]_{\phi(\mathbf{x}_i)=\phi_{m-1}(\mathbf{x}_i)}, \tag{7}$$

and find the value of ρ_k with a line search,

$$\rho_k = \arg\min_{\rho} \sum_{i=1}^{n} \mathcal{L}\left(y_i, \phi_{k-1}\left(\mathbf{x}_i\right) + \rho f_k\left(\mathbf{x}_i\right)\right). \tag{8}$$

However, the condition set in equation (7) is not always feasible. Consequently, we employ the learner $f_k\left(\mathbf{x}_i\right) = f\left(\mathbf{x}_i, \boldsymbol{\theta}_k\right)$, which most closely aligns with g_k across the data distribution. This is the solution to the problem

$$\boldsymbol{\theta}_k = \arg\min_{\beta, \boldsymbol{\theta}} \sum_{i=1}^{n} \left[-g_k\left(\mathbf{x}_i\right) - \rho_k f\left(\mathbf{x}_i, \boldsymbol{\theta}\right)\right]^2. \tag{9}$$

In tree-based models, the weak learners are regression trees. A tree consists of branches originating from a common node, with each branch being a sequence of internal nodes culminating in a leaf. The internal nodes contain decision-making criteria; starting at the tree's root and following the decision rules, each data point can be allocated to one of the leaves, or a set of scores can be assigned to each leaf and later combined into a single prediction. The goal is to create a model that predicts a target variable's value by learning the correct decision rules inferred from the data features. In this framework, optimizing the parameters in equation ((9)) involves determining the ideal tree structure and leaf weights, a computationally intensive task. The conventional "greedy" method requires evaluating every possible split point for each feature in the training set.

Recently, different algorithms and engineering solutions have been proposed to train gradient boosting models more efficiently (see, e.g., Tyree et al. (2011); Chen and Guestrin (2016); Ke et al. (2017b)). Among these, LightGBM (Ke et al., 2017b) is notable for its focus on optimizing training time, especially on large datasets. **LightGBM** introduces two techniques to train models more efficiently. First, it trains the trees on a subset of the dataset by keeping the data points (observations) associated with larger gradients and discarding the rest. Second, it bundles features together to reduce the dimensionality of the data. **LightGBM** significantly outperforms the other gradient-boosting implementations in terms of computational speed and memory consumption with minor compromises on predictive performance (Bentéjac et al., 2021).

Tree-based methods tend to overfit the data. Several hyperparameters in the **LightGBM** implementation guard against it. In the SKTL model, only a few of them are tuned while setting the others at their default value. Below is the list of hyperparameters tuned to improve model performance:

- *The number of boosting rounds.* We try to find the optimal number of base learners, K, as defined in equation (3). Setting a lower K would reduce overfitting.

- *The maximum depth of regression tree.* The parameter restricted the maximum depth of base learner $f(\mathbf{x}, \boldsymbol{\theta})$. A lower maximum depth would restrict the complexity of the base learner and reduce overfitting.

- *The learning rate of base learners.* The parameter shrinks each base learner $\rho_k f_k$ in equation (3) towards zero by applying a shrinkage parameter before it. A lower learning rate, coupled with more boosting rounds, would lead to a better fit of the model to the sample.

- *Path smoothing.* The parameter governs how each base learner tree generates outputs to reduce overfitting in leaves with fewer samples. Under default value of zero, the output is solely determined by the samples from each leaf. A positive path-smoothing parameter indicates the tree outputs are the weighted average of the output of current leaf and outputs of past nodes that lead to the leaf. The output of a leaf or node is defined as average sample dependent variables in the leaf or node.

A grid search for different combinations of the hyper-parameters listed above yield those that minimize the cross-validation MSE using data up to 2015. The design of training and validation sets are described next.

5.3 Model Estimation and Evaluation

A distinct feature of our exercise relative to earlier panel forecasting studies is that the domain of model application (privately held firms) is distinct from the domain of model estimation (publicly traded firms). To evaluate the model's performance more accurately, we design an out-of-sample evaluation exercise to avoid data leakage across firms. The testing set includes all observations post-year $T_{threshold}$. We divide the set of firms into M distinct groups. To generate PD for each group of firms in the testing set, we estimate the model with observations in the other firm groups before or equal to year $T_{threshold}$, as demonstrated in Panel (a) of Figure 1. The out-of-sample performance is therefore evaluated by pooling the model-generated PDs for observations after $T_{threshold}$ together. For this exercise, we set M to 5, and $T_{threshold}$ to 2015.

One drawback of solely relying on the out-of-sample evaluation outlined above is that the testing set only covers the later portion of the sample. This portion coincides with a period of relatively little defaults across most economies. While this issue could be addressed by selecting a larger testing set, doing so may result in large forecasting errors owing to the shorter training sample used vis-a-vis those of models estimated using the whole sample.

The cross-validation method is employed for estimating the model performance during sample periods before the out-of-sample testing set. The principle is to minimize the dependence of forecast errors from the validation set and its corresponding training set. Similar to the design of training and estimating sets in the out-of-sample exercise, firms are divided into M groups. We further divide the sample period before $T_{threshold}$ into N blocks, each consisting of a consecutive period. For each validation consisting of firms from group m observed in time block n, we apply the model estimated from firms outside group m and periods outside block n, as shown in Panel (b) of Figure 1. Cross-sectional and serial dependency between the training and validation sets is avoided by ensuring no observation overlaps in time and firms between both sets.

5.3.1 Model Performance among Publicly Traded Firms

The gradient-boosting tree model with all predictors in Table 1 is our "main model" or the SKTL model. We compare its performance against two other models. The first is the prevailing mean model (PM) which sets the PD forecast equal to the sample mean in the training set. The second model, the benchmark model, is also a gradient-boosting tree model that includes all the predictors in Table 1 except for the sectoral logit PD predictors. This model is estimated similarly to the main model. Comparing the results of the main model with the benchmark model helps to evaluate the contribution of the sectoral PDs to improve the PD prediction performance.

The R^2 of model A relative to model B is used to compare the performance of the two models. Let MSE_A and MSE_B denote the mean squared error of model A and model B respectively. The R^2 of model A against model B is defined as

$$R^2_{A/B} = 1 - \frac{MSE_A}{MSE_B}. \tag{10}$$

Table 5 presents the R^2 of the main model against the PM and benchmark models as well as the benchmark model against the PM model.[12] The left panel shows the cross-validation R^2 for observations with balance sheet data up to 2015; the right panel shows the out-of-sample R^2 which includes observations with balance sheet data from 2016 onwards. The cross-validation R^2 results for different economies show that the main model yields a cross-country median R^2 of 55 percent against the PM model, above the benchmark model cross-country median R^2 of 37 percent against the PM model. The median R^2 of the main model against the benchmark model is 30 percent, indicating a significant improvement from including higher frequency sectoral logit PD as predictors. Applying the time-block t-test for panel data (Qu et al., 2023) to compare the MSE difference of the main model and the benchmark model, we find the main model significantly outperforms the benchmark model in all economies, with p-values (not shown here) below 0.01. In the post-2015 period, the main model yields a cross-country median R^2 of 57 percent against the PM model, similar to its performance in data prior to 2015. The benchmark model's performance improves post-2015, yielding an R^2 of 45 percent. The median R^2 of the main model against the benchmark model is at 18 percent, with economy-specific R^2 positive and smaller but still statistically significant.[13] The smaller improvement of the main model on the benchmark model can be attributed to the stability and low default rates observed in the period.

However, the outperformance of the main model should be more pronounced during periods of rapid macroeconomic and financial changes, such as recessions or widespread financial distress, owing to the timely information provided by sectoral PDs. Models relying solely on lagged balance sheet data might not perform as well because past balance sheet data may not accurately capture the most current conditions.

The idea can be illustrated by decomposing MSE into mean squared bias and mean squared idiosyncratic errors at the sector level. Let $e_{it,m}$ denote the forecast error of model m regarding the logit PD of firm i at time t; let $\bar{e}_{st,m}$ denote the average forecast errors of model m about all firms in sector s observed at time t. The MSE of model m can be reformulated as the following:

$$MSE_m = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} e_{it,m}^2}{NT} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} \bar{e}_{s(i)t,m}^2}{NT} + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(e_{it,m} - \bar{e}_{s(i)t,m})^2}{NT}, \tag{11}$$

where the first term, squared bias, captures the loss caused by the bias of forecasts at the sector level, and the second term, squared idiosyncratic errors, captures the variance of forecast errors after correcting for sector-level biases. Similarly, the MSE differentials of the main model relative to the benchmark model can be decomposed into squared loss differentials and squared idiosyncratic error differentials. Table 6 shows the loss differential decomposition, conditional on different credit risk cycles as proxied by the average PD at each period.[14] The values presented in Table 6 are divided

[12] Annex Table A.1 presents analogous result from models which assume 15-month publication lags. The assumption is more consistent with the Orbis data.

[13] The individual economy-level test statistics are not shown here because all of them are statistically significant at 0.01 level.

[14] Annex Table A.2 presents analogous results for models assuming 15-month publication lags.

by the MSE of the benchmark model to make them more interpretable. Negative values indicate the main model outperforms the benchmark model. There are several interesting observations. First, the main model's improvement in squared idiosyncratic errors is small, with a cross-economy median improvement of 3 to 4 percent of benchmark model MSE, regardless of the credit risk cycle. Second, most of the main model's improvement comes from reducing squared sectoral biases, which is intuitive as the main model includes additional sectoral PD predictors. The outperformance is more apparent during periods when the average PD deviates from the historical median: squared bias differential contributes to 33 percent when the average PD is above the 75th percentile, 25 percent when the average PD is below the 25th percentile, and only 13 percent when average PD is within the 25 and 75 percentile.

Timely PD prediction requires using balance sheet data available at the time of the prediction. The data, published annually, could be lagging behind the prediction time by up to 12 months.[15] Therefore, to assess the effect of data staleness on the PD prediction, we estimate twelve submodels, each corresponding to a data release lag ranging from 0 to 11 months.[16] Figure 2 shows the distribution of MSE of the main model (blue line) and the benchmark model (red line) MSE across different countries at each lag $l \in \{0, 1, \ldots, 11\}$, scaled by the main model's MSE at $l = 0$. [17]

In the case of the main model, increasing data staleness, i.e. longer lags, impairs its predictive performance. The MSE ratio for $l = 11$, an 11-month lag, is about 1.1. In other words, suppose we want to estimate the PDs at the end of February 2023 and that balance sheet data are released every March, three months after the end of the fiscal year. The latest available balance sheet data corresponds to fiscal year 2021, published 11 months ago. If fiscal year 2022 data could be accessed earlier, we would improve the current forecast of PD by 12 percent. This is the "value added" of early data access. Similar patterns are observed in the performance of the benchmark model.

In summary, the results here show that our main model, a mixed frequency model that combines different data like balance sheet and macroeconomic information along with sectoral PDs, performs much better than benchmark models that ignore sectoral PDs. This difference is especially clear when aggregate PD level deviates to the tail of sample distribution. Adding real-time sectoral PDs gives us an advantage, as our main model using balance sheet data released eleven months earlier still performs better than the benchmark model using current balance sheet data. This suggests that sectoral PDs offer valuable forward-looking insights alongside the most recent balance sheet data, reinforcing the findings in Section 4.

5.3.2 Model Performance among Privately Held Firms

Given the documented discrepancy in default risk between privately held and publicly traded firms by Altman (2013), it's pertinent to assess the SKTL model's performance specifically within the

[15] See the discussion in Section 3.

[16] In equation (2), the lag corresponds to the variable l.

[17] Annex Figure A.1 presents analogous results from models assuming 15-month publication lags.

domain of privately held firms to ensure its reliability. However, as outlined in Section 3.3, the limitations in the Orbis data quality prevent us from directly estimating PD models tailored for privately held firms. Consequently, we evaluate the efficacy of the model originally calibrated with publicly traded firm data by using current end-of-year sectoral PDs in conjunction with privately held firms' balance sheet data to forecast default events affecting these entities one year in advance.

To justify the use of Orbis bankruptcy data for model evaluation, we examine the NUS-CRI PD performance in predicting publicly traded firms' credit events sourced from Orbis. If Orbis bankruptcy records are not severely biased, the NUS-CRI PD should perform well on Orbis credit events in the domain of listed firms. To alleviate potential bias from firms that exit with "Active" status, we discard the last observations of firms with "Active" status. We compute the AUROC of the end-of-year NUS-CRI PD applied to predict Orbis credit events in a one-year horizon. The results are presented in the fourth column of Table 3. Except for a few economies with very few credit events recorded by Orbis (potentially significant omissions), the NUS-CRI PD yields AUROC above 0.8. The result gives us some assurance that, for most of the economies in our sample, Orbis bankruptcy records can provide a reality check for our main model.

Figure 3 shows the ROC curves and their corresponding AUROC for each calibration group, excluding the emerging markets model, for which only one bankruptcy event has been recorded.[18] The results for China should also be viewed cautiously, as the recorded 38 bankruptcy events are notably low compared to the large sample size of over 800,000 privately held firms. When China results are omitted, AUROCs for the remaining groups range between 0.7 and 0.81. We can conclude that the SKTL model, initially estimated using publicly traded firms' data, can be confidently extended to privately held firms.

5.4 Interpreting Model Forecasts

The SKTL model establishes a statistical relationship between sectoral PDs, macroeconomic variables, and balance sheet data to derive probabilities of default for individual firms. However, a real-world application of the model requires understanding how the model comes to its conclusions. Shapley values may assist in this interpretive process.

Shapley values were originally developed in the game theory literature as a way to divide fairly a game's outcome across a set of cooperative players based on their contribution to the outcome of a game (Shapley, 1953). In statistical and machine learning models, they are used to evaluate the contribution of the predictors, or features, to the outcome of a model.

To gain some intuition on Shapley values, let us consider the following example. Assume a set of n players $P = \{1, \ldots, n\}$ cooperating in a game and a function $\nu(S)$ that, given a subset of players S, computes the outcome of the game when played by the members of S. To compute the

[18] Annex Figure A.2 presents analogous results from models assuming a 15-month publication lag of balance sheet data.

contribution of a specific player i, one could compute the outcome $\nu(S_{-i})$ of the game for a generic set S_{-i}, $i \notin S_{-i}$, and compare it to the outcome $\nu\{S_{-i} \cup \{i\}\}$ of the game when i is present. The Shapley value ϕ_i of player i will be the average difference $\nu(S_{-i} \cup \{i\}) - \nu(S_{-i})$ computed over all the possible sets S_{-i} that do not include player i,

$$\phi_i = \sum_{S_{-i},\, \{i\} \notin S_{-i}} \frac{|S_{-i}|!\,(n - |S_{-i}| - 1)}{n!} \left(\nu(S_{-i} \cup \{i\}) - \nu(S_{-i})\right). \tag{12}$$

In machine learning, evaluating the contribution of a feature requires computing the Shapley value of a feature i as the average difference in the model's outcome computed over all the possible sets of features S_{-i} that excludes i and the respective sets $S_{-i} \cup \{i\}$. A key strength of Shapley values, compared to other methods to compute features' importance, is that they can be estimated for each observation in the dataset, thus allowing researchers to evaluate the contribution of a feature to each point prediction the model made. Another useful property is that adding the Shapley values of a group of features yields their combined effect on the prediction.[19]

We decompose the SKTL model predictions for privately held firms into the Shapley values of the different features. Figures 4 and 5 shows the Shapley values of the top 15 predictors based on their average contribution to the sample predictions, or average absolute Shapley values. Each dot in the figures represents one observation, with the Shapley values plotted along the x-axis. The color of the dots indicates the Shapley value of the features: red hues correspond to higher values, while blue hues represent lower values. Among the balance sheet variables, those making the most significant contributions, in descending order, are the solvency ratio, profit margin, earnings-to-debt ratio, interest coverage ratio, and total assets. The sign of the Shapley values aligns with the expectation that higher leverage, lower profitability, and liquidity contribute to higher PD values. Across all country models, sectoral PD variables consistently rank among the top contributors. In contrast, macroeconomic variables such as GDP growth and interest rates contribute much less, suggesting that much of the macroeconomic information might be already captured by sectoral PD predictors.

6 Case Studies

The following section delves into two case studies exemplifying potential applications of the SKTL model as a real-time monitoring tool for assessing financial conditions in the corporate sector. The initial case scrutinizes the German utility sector, while the subsequent case delves into the UK commercial real estate sector. Comparative analyses are conducted, juxtaposing outcomes obtained through applying the model to publicly traded firms and those resulting from employing

[19]However, there are two problems. First, for a given ML model, it is not possible to "exclude" a feature when making a prediction. Second, the computation of all the possible sets S_{-i} might be time-consuming. Nevertheless, there are numerical strategies to compute the Shapley values efficiently (Lundberg and Lee, 2017).

the benchmark model on privately held firms.

6.1 Case 1: Germany Utility Sector During 2022 Energy Crisis

In 2022, the Russia-Ukraine conflict led to an unprecedented crisis in Europe's energy system, resulting in the depletion of Russian gas supplies to the continent. Wholesale prices for electricity and gas surged dramatically, escalating by as much as 15-fold above their early 2021 levels, causing severe repercussions for households and businesses. Numerous utility companies encountered restrictions on their ability to pass on additional costs to clients. Consequently, governments swiftly intervened to provide liquidity to energy firms to mitigate the crisis (Eckert and Buli, 2022).

For a thorough examination of the energy crisis's impact on the utility sector, it's essential to encompass privately held firms. As of July 2023, there were 821 privately held firms, significantly outnumbering the 13 publicly traded firms. Utilizing the SKTL model allows for real-time analysis while accounting for the 15-month lag in Orbis's publication of balance sheet data for most firms.

Figure 6 (a) illustrates the evolving dynamics of the privately held PD distribution within the utility sector, juxtaposed with that of publicly traded firms (shaded area). Both distributions are generated using the SKTL model, with sectoral PDs serving as the real-time predictor. Additionally, the figure depicts the PD distribution of privately held firms derived from the benchmark model.

Notably, the PD distribution for privately held firms has exhibited a steady rise since early 2021, experiencing a significant spike in September 2022 coinciding with the Nord Stream 2 gas pipeline event, after which it has remained elevated. If real-time sectoral PD information is omitted, as in the benchmark model, the PDs of privately held utilities would have remained stagnant until July 2022, with a gradual increase thereafter.

Incorporating privately held firms into the analysis offers deeper insights into the default risk dynamics within the utility sector. Compared to publicly traded firms, the distribution of PDs for privately held firms exhibits more dispersion particularly evident in the right tail of the distribution, or tail risk. Notably, the top decile of the distribution, averaging around 45 basis points, indicates that approximately one out of ten firms faced the risk of losing its BBB- rating and potentially being downgraded to junk status. Municipal utilities and renewable energy firms were among the most affected by this risk. Conversely, all publicly traded firms remained comfortably within the investment-grade rating.[20]

The Shapley values help us spot the key factors driving higher default risk since 2019. To do this with the real-time SKTL model, we compare the changes in median Shapley values across firms between July 2023 and July 2019. We categorize predictors into groups like liquidity, profitability, solvency, macroeconomic conditions (macro), and sectoral PD, as outlined in Table 1. Figure 6 (b) illustrates the changes in Shapley values for each group, highlighting a worsening of liquidity, profitability, and solvency as the primary drivers behind the increased PDs. Sectoral PD variables,

[20] The mapping between the 1-year PD and credit ratings is based on Duan and Li (2021)

which embed forward-looking information, play a major role, contributing 88 percent to the rise in PDs, overshadowing contributions from other variables.

6.2 Case 2: UK Commercial Real Estate Sector

Falling prices and rising funding costs have put the commercial real estate sector under pressure (Fioretti et al., 2023). We examine the financial condition of privately held commercial real estate companies in the UK from a longer perspective, looking at their default risk dynamics since before the Global Financial Crisis (GFC) in 2008.

As of July 2023, the data sample contains 507 UK privately held commercial real estate firms with assets of at least 10 million USD.[21] Figure 7 (a) shows the median logit PD of these firms, estimated using the real-time SKTL model. In the aftermath of the GFC, the median PD was trending downwards until it rose sharply during the Covid-19 pandemic. The sector's median PD have improved since then but have remained above their pre-pandemic levels, and July 2023 level came close to those observed in 2007.

As in the previous case study, we group the predictors into liquidity, profitability, solvency, macro, and sectoral PDs and find the changes in the Shapley values between July 2023 and July 2007. Figure 7 (b), which is analogous to Figure 6 (b), shows that macroeconomic conditions not only contribute the most to default risk in the sector in 2023, but are relatively more important than in July 2007. On the other hand, current balance sheets provide more liquidity, profitability, and solvency buffers than during the pre-GFC period.

7 Conclusion

Transferring signals and knowledge from publicly traded firms can get us a long way in measuring default risks of privately held firms while sidestepping the need for privately held firms' default events data. Model evaluation results indicate the SKTL model performs decently as an ordinal measure of default risk, leading to AUROC which ranges between 0.7 and 0.8.

Including real-time sectoral PD significantly improves SKTL performance relative to the benchmark model that only relies on balance sheets and macroeconomic variables. As expected, the improvement from real-time sectoral PD concentrates during recessions and periods when systemic default probability is more volatile. The inclusion of real-time sectoral PD buys us time in the sense that the MSE of the SKTL eleven months after the release of balance sheets is still lower than the MSE of the benchmark model in the release months, which highlights sectoral PD contains forward-looking information in addition to the most recent balance sheet data.

[21] The commercial real estate firms are identified by Nace Rev.2 code 6820, and include firms that rent and operate either owned or leased real estate. The minimum asset threshold eliminates firms that do not own real estate and only provide management services.

We further evaluate the model's performance in two real-time case studies, the financial condition of the German utility sector during the recent energy crisis and the recent financial condition of the UK's commercial real estate sector in comparison to 2007. The SKTL model greatly expands the coverage of tail risks and quickly responds to macroeconomic events, e.g., the PD of German utility firms jumped in the month of Nord Stream 2 pipeline explosion and remained elevated afterward.

One proviso for SKTL is that the economy needs to have a relatively large and broad stock market, such that the sectoral PD of listed firms can effectively proxy for shocks affecting the whole sector in general. The constraint can be alleviated by pooling together similar economies. One promising direction for future research is to harness the balance sheet information of publicly traded firms that are more frequently updated to nowcast the balance sheets of privately held firms and inform about their financial condition.

References

Adasoro, I., Hardy, B., and Tarashev, N. (2021). Corporate debt: post-GFC through the pandemic. In *BIS Quarterly Review*. Bank for International Settlements.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609.

Altman, E. I. (2013). Predicting financial distress of companies: revisiting the z-score and zeta® models. In *Handbook of research methods and applications in empirical finance*. Edward Elgar Publishing.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., and Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's z-score model. *Journal of International Financial Management & Accounting*, 28(2):131–171.

Anderson, P. L. (2009). The value of private businesses in the United States. *Business Economics*, 44(2):87–108.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., and Timmis, J. (2020). Coverage and representativeness of Orbis data. Technical report, OECD.

Bakshi, G., Gao, X., and Zong, Z. (2022). Decoding default risk: a review of modeling approaches, findings, and estimation methods. *Annual Review of Financial Economics*, 14:391 – 413.

Barboza, F., Kimura, H., and Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83:405–417.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, pages 71–111.

Ben-David, S., Blitzer, J., Cramer, K., Kuesza, A., Pereira, F., and Vaughan, J. W. (2010). A theory of learning from different domains. *Machine Learning*, 79(1):151–175.

Bentéjac, C., Csörgő, A., and Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3):1937–1967.

Bharath, S. T. and Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies*, 21(3):1339–1369.

Bhimani, A., Gulamhussen, M. A., and Lopes, S. D.-R. (2010). Accounting and non-accounting determinants of default: An analysis of privately-held firms. *Journal of Accounting and Public Policy*, 29(6):517–532.

Black, F. and Scholes, M. J. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Breiman, L., Friedman, J., Stone, C., and Olshen, R. (1984). *Classification and Regression Trees*. Taylor & Francis.

Campbell, J. Y., Hilscher, J., and Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63(6):2899–2939.

Cangemi, B., de Servigny, A., and Friedman, C. (2003). Standard & poor's credit risk tracker for private firms. technical document. Technical report, Standard & Poor's, Risk Solutions.

Carling, K., Jacobson, T., Lindé, J., and Roszbach, K. (2007). Corporate credit risk modeling and the macroeconomy. *Journal of Banking & Finance*, 31(3):845–868.

Charalambakis, E. C. and Garrett, I. (2019). On corporate financial distress prediction: What can we learn from private firms in a developing economy? evidence from greece. *Review of Quantitative Finance and Accounting*, 52(2):467–491.

Chava, S. and Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of finance*, 8(4):537–569.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA. Association for Computing Machinery.

Crosbie, P. J. and Bohn, J. R. (2002). Modeling default risk–kmv. *Moody's KMV*.

Das, S. R., Duffie, D., Kapadia, N., and Saita, L. (2007). Common failings: How corporate defaults are correlated. *The Journal of Finance*, 62(1):93–117.

Dierkes, M., Erner, C., Langer, T., and Norden, L. (2013). Business credit information sharing and default risk of private firms. *Journal of Banking & Finance*, 37(8):2867–2878.

Duan, J. C. (1994). Maximum likelihood estimation using price data of the derivatives contract. *Mathematical Finance*, 4:155 – 167.

Duan, J. C., Kim, B., Kim, W., and Shin, D. (2018). Default probabilities of privately held firms. *Journal of Banking & Finance*, 94:235–250.

Duan, J. C. and Li, S. (2021). Enhanced pd-implied ratings by targeting the credit rating migration matrix. *The Journal of Finance and Data Science*, 7:115–125.

Duan, J. C., Sun, J., and Wang, T. (2012). Multiperiod corporate default prediction. a forward intensity approach. *Journal of Econometrics*, 170(1):191–209.

Duan, J. C. and Wang, T. (1994). Measuring distance-to-default for financial and non-financial firms. *Global Credit Review*, 2:95 – 108.

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

Duffie, D. and Singleton, K. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 52:1287 – 1321.

Eckert, V. and Buli, N. (2022). Europe props up energy firms amid liquidity crunch. *Reuters*.

Falkenstein, E. G., Boral, A., and Carty, L. V. (2000). Riskcalc for private companies: Moody's default model. *Available at SSRN 236011*.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.

Fioretti, P., Skrutkowski, M., and Strauch, R. (2023). Commercial real estate and financial stability this time, itś different. *European Stability Mechanism Blog*.

Foroni, C., Guérin, P., and Marcellino, M. (2018). Using low frequency information for predicting high frequency variables. *International Journal of Forecasting*, 34(4):774–787.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232.

Giesecke, K., Longstaff, F., Schaefer, S., and Strebulaev, I. (2014). Macroeconomic effects of corporate default crisis: a long-term perspective. *Journal of Financial Economics*, 111(2):297–310.

Grinsztajn, L., Oyallon, E., and Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data?

Hillegeist, S. A., Keating, E. K., Cram, D. P., and Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of accounting studies*, 9(1):5–34.

Jarrow, R. and Turnbull, S. (1995). Pricing options on financial securities subject to credit risk. *Journal of Finance*, 50:53 – 85.

Kalemli-Ozcan, S., Gourinchas, P.-O., Penciakova, V., Sander, N., et al. (2020). Covid-19 and sme failures. Technical report, International Monetary Fund Washington, DC, USA.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S. (2015). How to construct nationally representative firm level data from the orbis global database: New facts and aggregate implications. Technical report, National Bureau of Economic Research.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017a). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017b). Lightgbm: A highly efficient gradient boosting decision tree. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Koopman, S. J., Lucas, A., and Schwaab, B. (2012). Dynamic factor models with macro, frailty, and industry effects for us default counts: the credit crisis of 2008. *Journal of Business & Economic Statistics*, 30(4):521–532.

Lando, D. (1998). On cox processes and credit risky securities. *Review of Derivatives Research*, 2:99 – 120.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc.

Madan, D. and Unal, H. (1998). Pricing the risks of default. *Review of Derivatives Research*, 2:121 – 160.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470.

Miao, H., Ramchander, S., Ryan, P., and Wang, T. (2018). Default prediction models: The role of forward-looking measures of returns and volatility. *Journal of Empirical Finance*, 46:146–162.

Mungo, L., Lafond, F., Astudillo-Estévez, P., and Farmer, J. D. (2023). Reconstructing production networks using machine learning. *Journal of Economic Dynamics and Control*, 148:104607.

Qu, R., Timmermann, A., and Zhu, Y. (2023). Comparing forecasting performance with panel data. *International journal of forecasting*.

Schlingemann, F. P. and Stulz, R. M. (2022). Have exchange-listed firms become less important for the economy? *Journal of Financial Economics*, 143(12):927–958.

Shapley, L. S. (1953). A value for n-person games. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games (AM-28)*, volume 2, pages 307–318. Princeton University Press, Princeton.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, 74(1):101–124.

Storkey, A. (2009). When training and test sets are different: characterizing learning transfer. In Quinonero-Candela, J., Sugiyama, M., Schwaighofer, A., and Lawrence, N., editors, *Dataset shift in machine learning*, pages 1–28. MIT Press.

Sutton, C. D. (2005). 11 - classification and regression trees, bagging, and boosting. In Rao, C., Wegman, E., and Solka, J., editors, *Data Mining and Data Visualization*, volume 24 of *Handbook of Statistics*, pages 303–329. Elsevier.

Tyree, S., Weinberger, K. Q., Agrawal, K., and Paykin, J. (2011). Parallel boosted regression trees for web search ranking. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 387–396, New York, NY, USA. Association for Computing Machinery.

Vassalou, M. and Xing, Y. (2004). Default risk in equity returns. *The journal of finance*, 59(2):831–868.

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1):1–40.

Zhou, X., Huang, J., Friedman, C., Cangemi, R., and Sandow, S. (2006). Private firm default probabilities via statistical learning theory and utility maximization. *Journal of Credit Risk*, 2(1):51 – 65.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, pages 59–82.

Table 1: List of predictors

Label	Definition or transformation
Panel A: Liquidity	
Acid test	Cash and cash equivalent / Current liabilities
Cash to operating revenue ratio	Cash flow / Operating revenue turnover
Current ratio	Current assets / Current liabilities
Debt-to-suppliers to operating revenue ratio	Debt to suppliers / Operating revenue
Interest coverage ratio	Operating profit / Interest paid
Liquidity ratio	(Current assets - inventory) / Current Liabilities
Tangible fixed assets to debt ratio	Tangible fixed assets / (Current liability + Noncurrent liability)
Trade receivables to operating revenue ratio	Trade receivables / Operating revenue
Working capital	(Current assets - Current liabilities) / Total assets
Panel B: Profitability and activity	
Asset turnover	Operating revenue / Total assets
Gross profit margin	(Operating revenue - Costs of goods sold) / Operating revenue
Profit margin (earnings before tax based)	Earnings before tax / Operating revenue
Profit margin (net income based)	Net income / Operating revenue
ROA (earnings after tax based)	Earnings after tax / Total assets
ROA (earnings before tax based)	Earnings before tax / Total assets
ROA (EBIT based)	EBIT / Total assets
ROA (EBITDA based)	EBITDA / Total assets
ROA (net income based)	Net income / Total assets
ROCE using net income	(Net income + Interest paid) / (Shareholders' funds + Non current liabilities)
ROCE using profit before tax	(Profit before tax + Interest paid) / (Shareholders' funds + Non current liabilities)
ROE using net income	Net income / Shareholders' funds
Stock turnover	Operating revenue / Inventory
Panel C: Solvency	
Earnings to debt ratio	EBIT / (Current liabilities + Noncurrent liabilities)
Gearing ratio	(Noncurrent liabilities+Loans) / Shareholders' funds
Shareholders liquidity ratio	Shareholders' funds / Noncurrent liabilities
Solvency ratio (asset based)	Shreholders' funds / Total assets
Solvency ratio (liability based)	Shreholders' funds / (Non current liabilities+Current liabilities)
Panel D: Other firm characteristics	
Industry sector	NACE Rev. 2 main section
Number of employee	
Total assets in constant price	In 2015 USD
Panel E: PD and macro economic variables	
Inter-bank overnight rates	End-of-month values
Real quarterly GDP growth	
Sectoral 20 percentile of logit PD	End-of-month values
Sectoral 80 percentile of logit PD	End-of-month values
Sectoral median of logit PD	End-of-month values

Table 2: Summary statistics

	Start	End	Number of firms[1]		Aggregate total asset[2]		Total asset[3]		Solvency Ratio[4]		Return on Assets[5]		Liquidity Ratio[6]	
			Listed	Non-listed	Listed	Non-listed	Listed	Non-listed	Listed	Non-listed	Listed	Non-listed	Listed	Non-listed
BR	1995	2023	405	14,651	840	512	656.59	23.37	0.39	0.51	3.18	3.43	1.14	1.26
CN	1995	2023	2,576	731,608	9,251	2,567	322.15	4.57	0.53	0.39	3.87	1.50	1.08	0.90
DE	1995	2023	1,155	102,087	4,118	3,680	120.56	12.79	0.41	0.29	2.25	2.95	1.24	1.25
ES	1995	2023	297	406,899	1,072	2,829	380.35	2.65	0.38	0.40	2.69	1.41	0.95	1.04
FR	1995	2023	1,078	395,692	4,164	4,912	100.13	2.51	0.40	0.34	2.63	3.76	1.10	1.08
GB	1995	2023	2,918	209,814	4,371	12,079	88.63	9.02	0.53	0.37	1.99	3.09	1.16	1.12
HK	1995	2023	219	1,210	2,161	3,399	608.06	246.74	0.62	0.55	3.34	3.80	1.42	1.50
IN	1995	2023	4,027	91,595	1,604	2,236	21.03	5.22	0.44	0.35	1.98	0.91	1.16	0.99
IT	1995	2023	511	565,962	1,078	4,389	200.02	2.84	0.33	0.19	1.87	0.52	0.99	0.89
JP	1995	2023	5,006	191,487	9,255	3,522	245.63	4.88	0.49	0.30	2.54	1.50	1.33	1.42
KR	1995	2023	2,934	348,991	3,338	2,826	71.92	3.05	0.50	0.37	2.66	3.59	1.11	1.27
MY	1995	2023	1,106	79,106	482	920	83.12	2.84	0.59	0.30	2.81	2.21	1.36	1.22
NL	1995	2023	310	16,843	1,092	1,624	405.67	34.01	0.41	0.34	3.58	3.93	0.99	1.06
PH	1995	2023	222	16,127	388	270	108.36	3.93	0.57	0.35	2.79	1.62	1.32	1.00
PL	1995	2023	673	67,951	230	856	30.89	4.08	0.50	0.44	3.29	3.33	0.97	0.99
RU	1995	2023	206	273,631	1,264	2,135	775.25	2.86	0.51	0.19	3.72	1.26	0.92	0.94
SG	1995	2023	852	32,765	793	1,642	96.24	8.28	0.55	0.37	3.14	2.41	1.33	1.24
TH	1995	2023	687	81,338	609	1,096	74.97	3.23	0.52	0.38	4.32	1.61	0.96	0.99
US	1995	2023	8,819	3,240	23,753	1,548	243.66	26.10	0.48	0.38	1.53	-4.06	1.41	1.06
VN	2001	2023	691	107,805	118	1,047	20.65	2.67	0.48	0.32	4.39	0.12	0.96	0.86

[1] We focus on observations with total assets above 1 million USD and with data available on total assets, sharholders' funds, current liabilities, noncurrent liabilities, operating revenue and net income.

[2] Aggregate total asset of fiscal year 2021 in billion of USD.

[3] Median total asset in million of USD.

[4] Median shareholders' fund to total assets ratio.

[5] Median net income to total assets ratio (%).

[6] Median current asset minus inventory to current liability ratio.

Table 3: Orbis bankruptcy events and NUS-CRI credit events

	Orbis events		CRI events	CRI-PD AUROC[1]	
	Nonlisted firms	Listed firms	Listed firms	Orbis events	CRI events
BR	1	1	46	0.94	0.87
CN	38	0	986		0.83
DE	478	25	175	0.84	0.87
ES	4303	2	20	0.59	0.76
FR	22369	40	57	0.85	0.87
GB	844	14	154	0.79	0.89
HK	1	0	98		0.82
IN	683	21	1453	0.96	0.86
IT	60784	26	27	0.90	0.89
JP	377	11	215	0.94	0.93
KR	5	64	239	0.61	0.90
MY	0	0	237		0.86
NL	344	14	39	0.81	0.88
PH	0	0	41		0.84
PL	1342	2	61	0.97	0.88
RU	4976	0	25		0.66
SG	613	9	61	0.86	0.87
TH	0	1	127	0.78	0.91
US	59	85	1420	0.86	0.93
VN	1	0	2		

[1] The AUROC of CRI PD when applied to predict CRI credit events or Orbis firms' exits in the next twelve months

Table 4: Intra-year prediction regression of sectoral financial ratios against sectoral PD

	Solvency ratio[1]		Return on assets[2]		Liquidity ratio[3]	
Month	Listed	Non-listed	Listed	Non-listed	Listed	Non-listed
March	−0.02	−0.07	−0.11	−0.17**	0.11	0.40***
	(0.04)	(0.05)	(0.05)	(0.07)	(0.09)	(0.15)
June	−0.06	−0.18***	−0.15**	−0.22***	−0.25	−0.14
	(0.05)	(0.05)	(0.06)	(0.06)	(0.17)	(0.16)
September	−0.11***	−0.29***	−0.24**	−0.37***	−0.39***	−0.55***
	(0.03)	(0.04)	(0.10)	(0.07)	(0.13)	(0.17)
December	−0.17***	−0.28***	−0.27***	−0.33***	−0.58***	−0.63***
	(0.05)	(0.04)	(0.10)	(0.06)	(0.16)	(0.17)

Each regression includes 5044 country-sector-year observations. The standard errors are derived from clustered covariance matrix. * indicates pvalues between 5 percent and 10 percent; ** indicates pvalues between 1 percent and 5 percent; *** indicates pvalues below 1 percent.

[1] Shareholders' funds to total assets ratio (%)

[2] Net income to total assets ratio (%)

[3] Current asset minus inventory to current liability ratio (%)

Table 5: R^2 of Logit PD, assuming balance sheets are published 3 months after closing dates

	Cross-validation (Year $\leq$ 2015)			Out of sample (Year > 2015)		
	$\frac{\text{Main}}{\text{PM}}$	$\frac{\text{Benchmark}}{\text{PM}}$	$\frac{\text{Main}}{\text{Benchmark}}$	$\frac{\text{Main}}{\text{PM}}$	$\frac{\text{Benchmark}}{\text{PM}}$	$\frac{\text{Main}}{\text{Benchmark}}$
Median	0.55	0.37	0.30	0.57	0.45	0.18
BR	0.60	0.47	0.25	0.63	0.55	0.19
CN	0.79	0.62	0.45	0.64	0.59	0.11
DE	0.46	0.28	0.25	0.46	0.41	0.08
ES	0.55	0.32	0.34	0.62	0.55	0.15
FR	0.53	0.32	0.31	0.49	0.45	0.09
GB	0.44	0.28	0.22	0.47	0.42	0.09
HK	0.59	0.45	0.25	0.58	0.50	0.14
IN	0.54	0.37	0.27	0.55	0.43	0.21
IT	0.54	0.31	0.33	0.51	0.45	0.11
JP	0.66	0.50	0.32	0.73	0.68	0.17
KR	0.69	0.57	0.27	0.56	0.43	0.23
MY	0.66	0.56	0.22	0.58	0.52	0.13
NL	0.53	0.25	0.38	0.53	0.34	0.29
PH	0.62	0.44	0.32	0.57	0.44	0.23
PL	0.52	0.28	0.33	0.59	0.26	0.44
RU	0.43	-0.10	0.48	0.53	0.21	0.40
SG	0.54	0.37	0.27	0.55	0.47	0.14
TH	0.66	0.49	0.34	0.70	0.47	0.44
US	0.55	0.37	0.29	0.51	0.39	0.19
VN	0.67	0.54	0.28	0.64	0.54	0.20

R^2 of Model A over Model B is defined as $R^2_{A/B} = 1 - MSE_A/MSE_B$. Model "PM" denotes prevailing mean model where the forecasts are sample mean in the training sets.

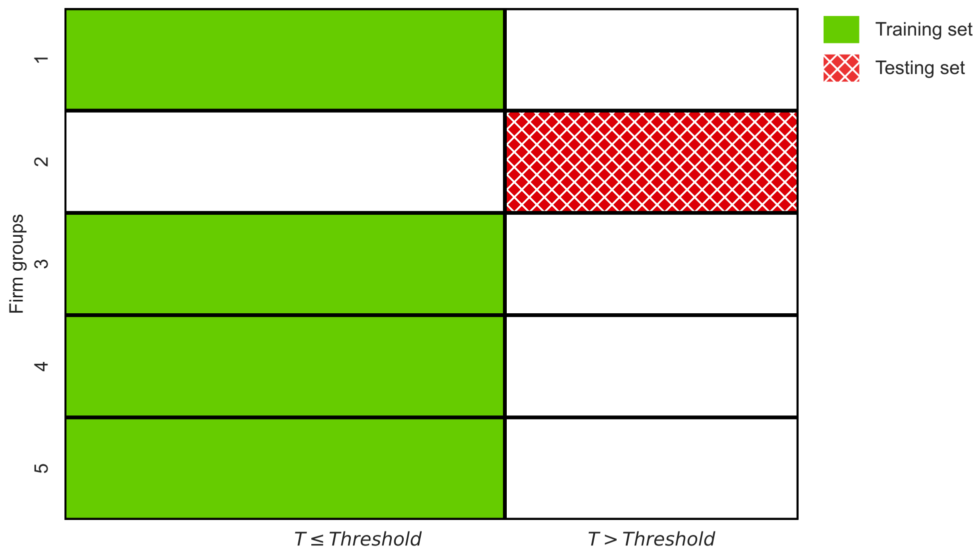
Table 6: Loss differential decomposition of predicting Logit PD, conditional level of average PD

	$\bar{PD}_t \leq qt_{25}$		$qt_{25} \leq \bar{PD}_t \leq qt_{75}$		$\bar{PD}_t \geq qt_{75}$	
	Squared bias	Squared idiosyncratic loss	Squared bias	Squared idiosyncratic loss	Squared bias	Squared idiosyncratic loss
Median	-0.25	-0.03	-0.13	-0.04	-0.33	-0.04
BR	-0.22	-0.04	-0.13	-0.06	-0.28	-0.06
CN	-0.17	-0.16	-0.18	-0.13	-0.31	-0.18
DE	-0.12	-0.02	-0.12	-0.03	-0.35	-0.03
ES	-0.28	-0.02	-0.21	-0.03	-0.34	-0.02
FR	-0.29	-0.03	-0.10	-0.02	-0.46	-0.03
GB	-0.32	-0.04	-0.13	-0.03	-0.14	-0.03
HK	-0.17	-0.03	-0.14	-0.06	-0.18	-0.09
IN	-0.06	-0.16	-0.03	-0.15	-0.04	-0.31
IT	-0.22	-0.05	-0.23	-0.07	-0.20	-0.06
JP	-0.25	-0.03	-0.18	-0.05	-0.42	-0.04
KR	-0.33	-0.02	-0.13	-0.03	-0.47	-0.03
MY	-0.25	0.00	-0.07	-0.03	-0.32	-0.05
NL	-0.49	-0.02	-0.17	-0.03	-0.49	-0.03
PH	-0.43	-0.03	-0.13	-0.04	-0.37	-0.01
PL	-0.12	-0.11	-0.28	-0.05	-0.52	-0.07
RU	-0.15	-0.09	-0.34	-0.10	-0.57	-0.13
SG	-0.23	-0.06	-0.11	-0.05	-0.32	-0.04
TH	-0.43	-0.02	-0.29	-0.03	-0.44	-0.07
US	-0.39	-0.05	-0.11	-0.03	-0.30	-0.04
VN	-0.47	-0.07	-0.13	-0.04	-0.25	-0.02

The table shows the decomposition of loss differentials between the main model and the benchmark model, conditional levels of cross-sectional average PD. The loss differential is decomposed into squared bias differentials within each sector, and the corresponding squared idiosyncratic error differentials. The sample are divided into three groups, periods when average PD is below the 0.25 quantile (left panel), periods when average PD is above the 0.75 quantile (right panel) and the other periods (middle panel).

Figure 1: Training, validation and testing sets in cross-validation and out-of-sample examination of model performance

(a) Out-of-sample training and testing sets



(b) Cross-validation training and validation sets

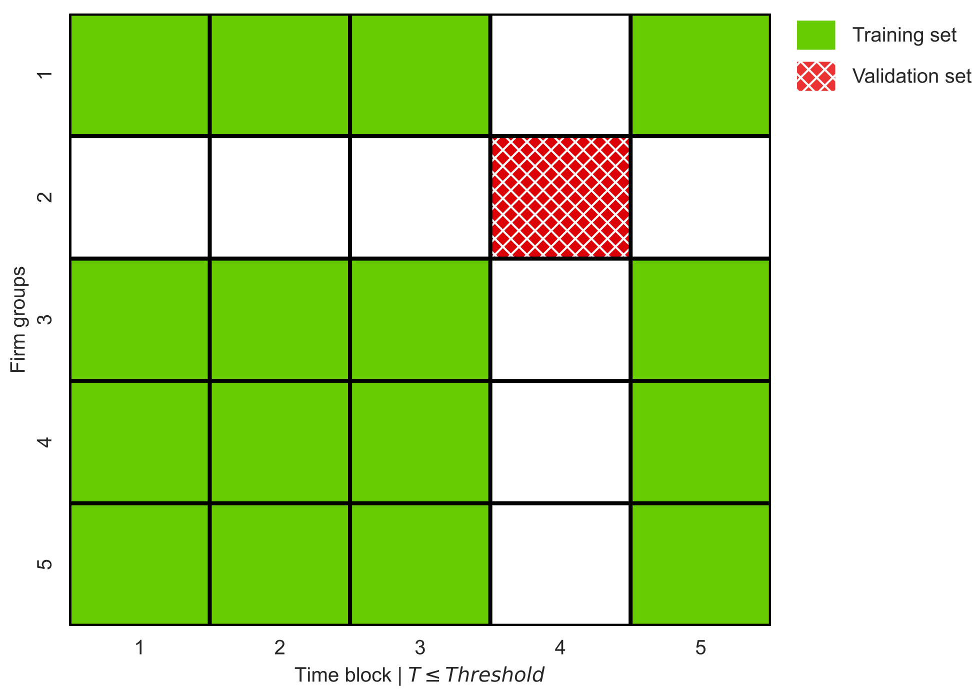


Figure 2: MSE of main and benchmark model at different lags since publication of annual reports
For each country, the MSE at different lags are divided by the zero-lag MSE. The figure shows median, 0.2 and 0.8 quantiles of MSE distribution across countries.

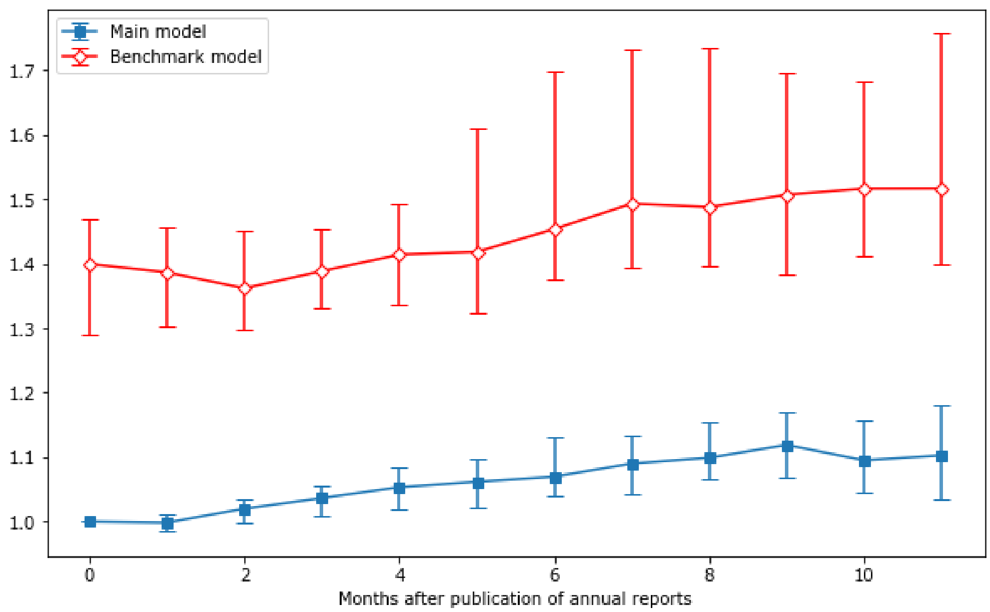
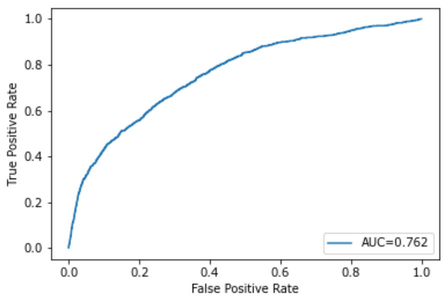
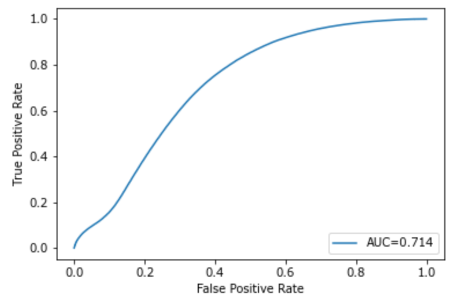


Figure 3: AUC of predicting non-listed firm bankruptcy

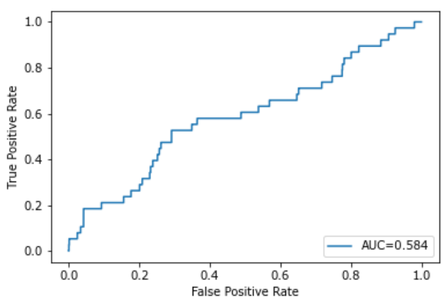
(a) Advanced Asia



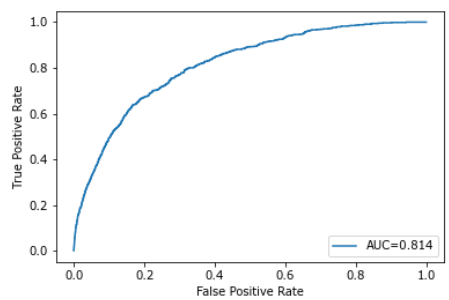
(b) Europe



(c) CN



(d) IN



(e) US

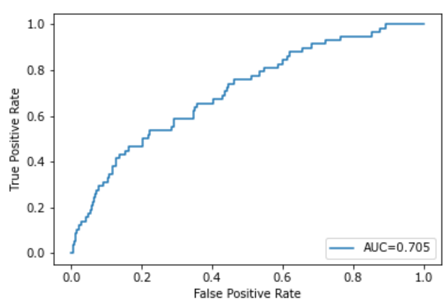
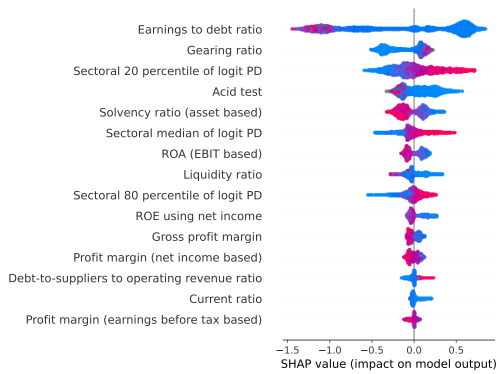
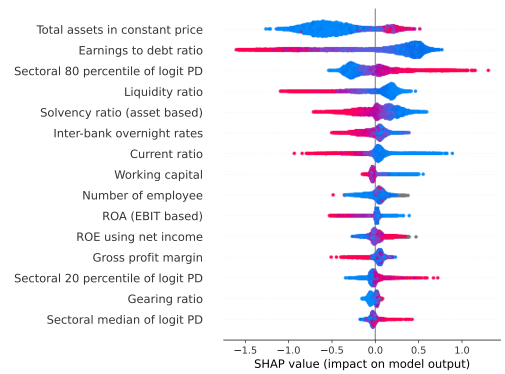


Figure 4: Distribution of Shapley values of main models applied to privately held firms

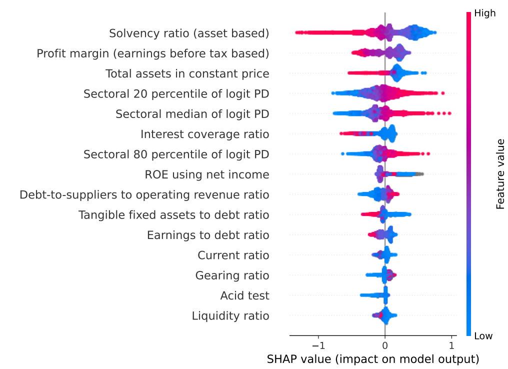
(a) BR



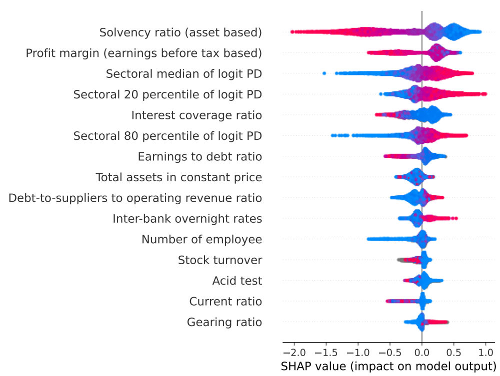
(b) CN



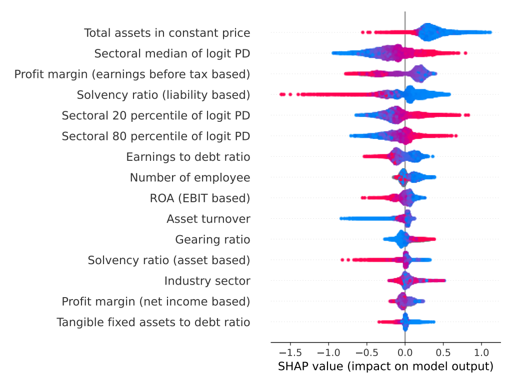
(c) DE



(d) ES



(e) FR



(f) GB

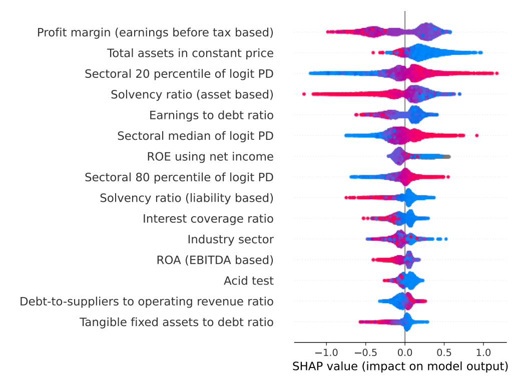


Figure 5: Distribution of Shapley values of main models applied to privately held firms (continued)

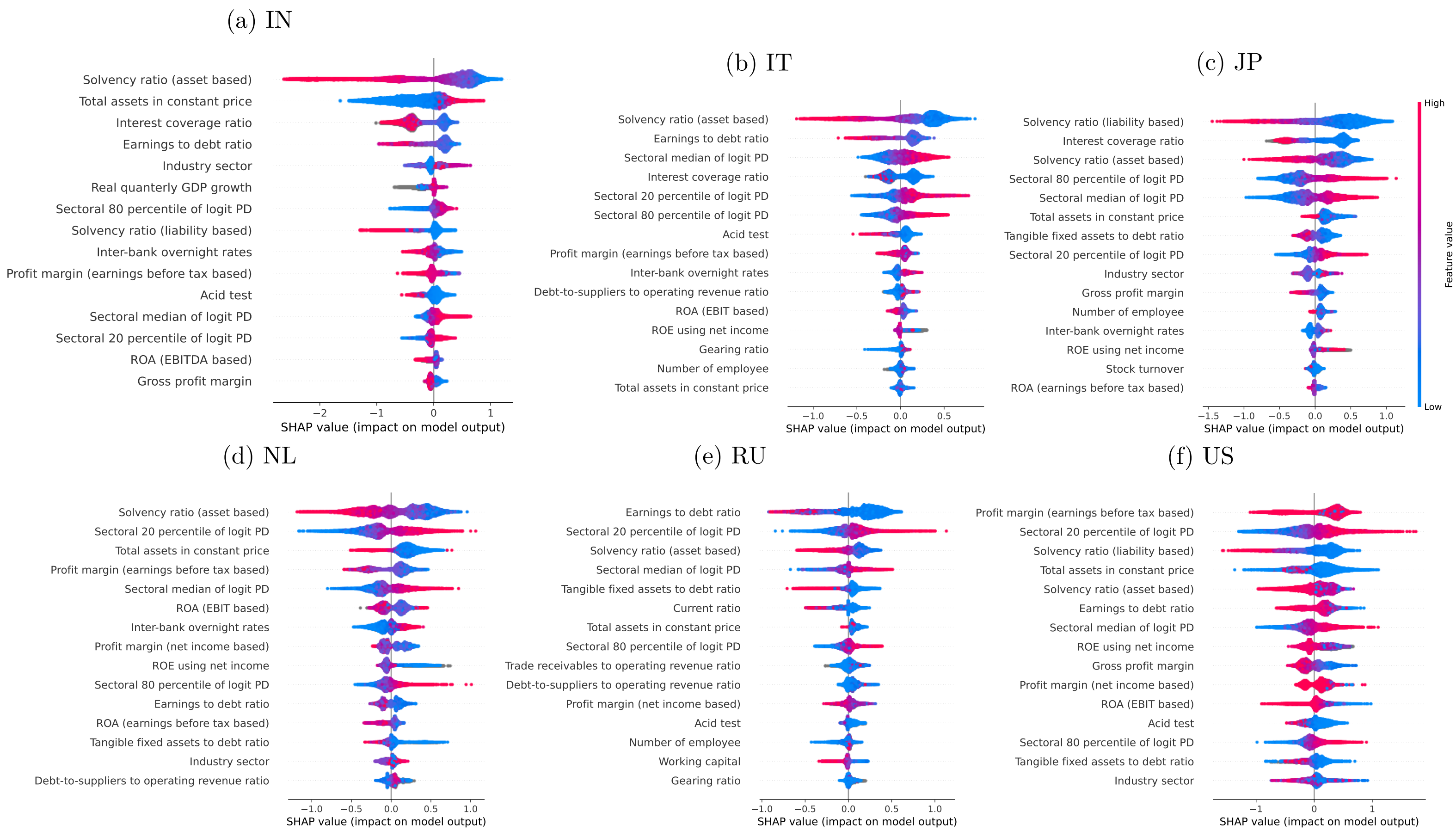
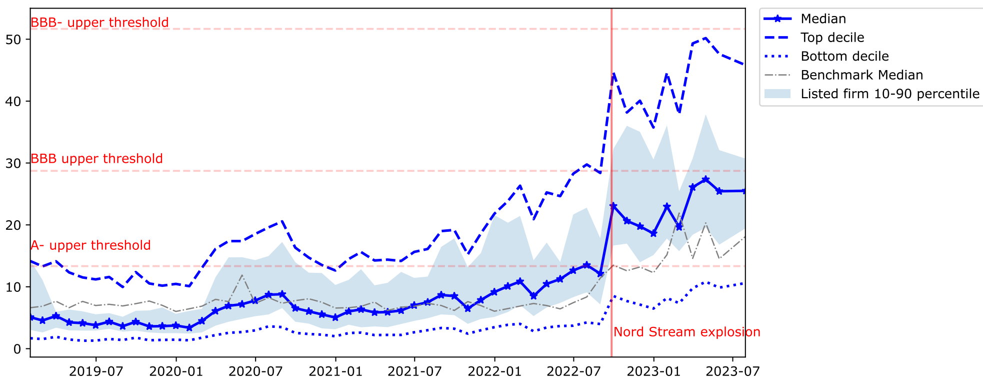


Figure 6: Case study 1: German gas and electricity firms during the post-covid energy crisis

(a) PD of German gas and electricity firms



(b) Main model's Shapley values July 2023 minus July 2019

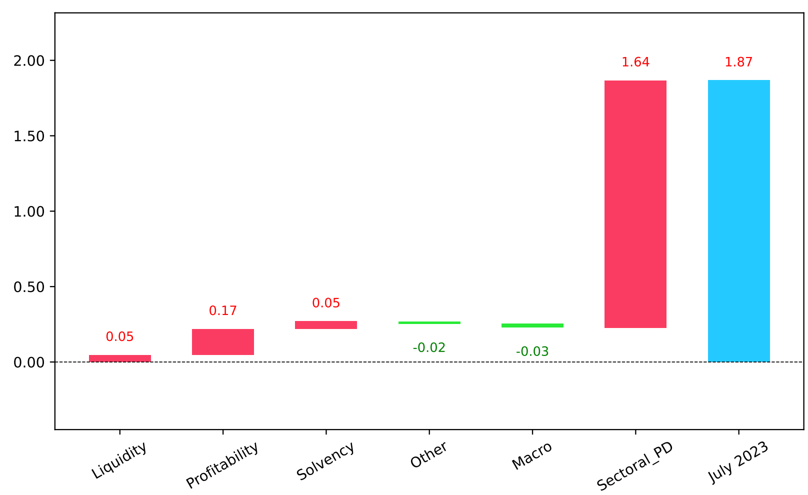
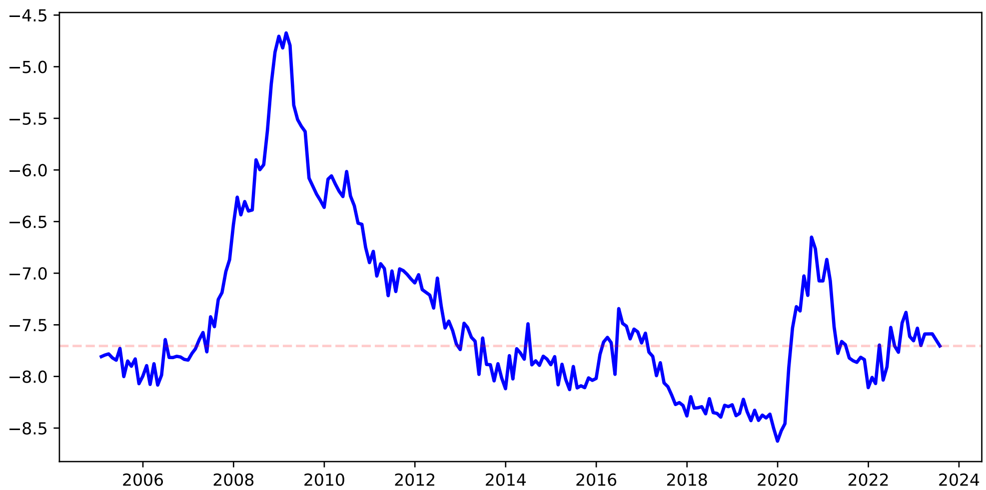


Figure 7: Case study 2: UK commercial real estate firms

(a) Logit PD of UK privately held commercial real estate firms



(b) Main model's Shapley values July 2023 minus July 2007

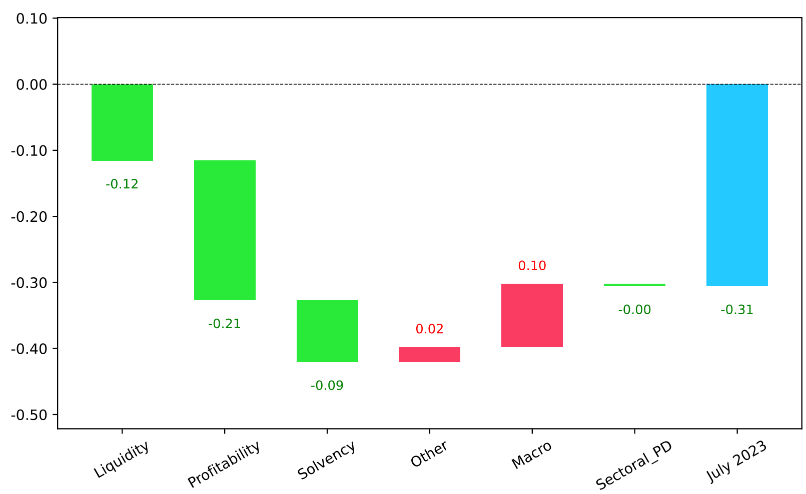


Table A.1: R^2 of Logit PD (%), assuming balance sheets are published 15 months after closing dates

	Cross-validation (Year $\leq$ 2015)			Out of sample (Year > 2015)		
	$\frac{\text{Main}}{\text{PM}}$	$\frac{\text{Benchmark}}{\text{PM}}$	$\frac{\text{Main}}{\text{Benchmark}}$	$\frac{\text{Main}}{\text{PM}}$	$\frac{\text{Benchmark}}{\text{PM}}$	$\frac{\text{Main}}{\text{Benchmark}}$
Median	0.51	0.35	0.29	0.51	0.41	0.15
BR	0.53	0.38	0.25	0.58	0.47	0.21
CN	0.72	0.52	0.42	0.51	0.45	0.12
DE	0.41	0.22	0.24	0.37	0.28	0.13
ES	0.51	0.26	0.34	0.60	0.54	0.14
FR	0.48	0.24	0.32	0.44	0.35	0.14
GB	0.35	0.18	0.21	0.39	0.32	0.12
HK	0.55	0.42	0.22	0.49	0.43	0.10
IN	0.50	0.45	0.09	0.52	0.49	0.05
IT	0.47	0.19	0.34	0.41	0.32	0.13
JP	0.66	0.51	0.32	0.67	0.63	0.11
KR	0.64	0.51	0.27	0.45	0.35	0.16
MY	0.60	0.53	0.16	0.54	0.47	0.12
NL	0.47	0.21	0.33	0.48	0.32	0.25
PH	0.59	0.38	0.33	0.50	0.40	0.17
PL	0.44	0.21	0.29	0.54	0.16	0.45
RU	0.42	0.02	0.41	0.47	0.20	0.34
SG	0.51	0.33	0.26	0.52	0.42	0.18
TH	0.61	0.44	0.31	0.66	0.46	0.36
US	0.48	0.26	0.29	0.41	0.27	0.20
VN	0.58	0.44	0.25	0.59	0.47	0.22

R^2 of Model A over Model B is defined as $R^2_{A/B} = 1 - MSE_A/MSE_B$. Model "PM" denotes prevailing mean model where the forecasts are sample mean in the training sets.

Table A.2: Loss differential decomposition, conditional level of average PD (15 months publication lag)

	$\bar{PD}_t \leq qt_{25}$		$qt_{25} \leq \bar{PD}_t \leq qt_{75}$		$\bar{PD}_t \geq qt_{75}$	
	Squared bias	Squared idiosyncratic loss	Squared bias	Squared idiosyncratic loss	Squared bias	Squared idiosyncratic loss
Median	-0.26	-0.04	-0.14	-0.03	-0.322	-0.04
BR	-0.16	-0.08	-0.12	-0.07	-0.32	-0.10
CN	-0.22	-0.16	-0.12	-0.11	-0.55	-0.12
DE	-0.25	-0.04	-0.10	-0.02	-0.31	-0.03
ES	-0.27	-0.02	-0.25	-0.03	-0.30	-0.03
FR	-0.34	-0.04	-0.15	-0.03	-0.41	-0.03
GB	-0.29	-0.04	-0.14	-0.02	-0.13	-0.03
HK	-0.22	-0.01	-0.12	-0.03	-0.17	-0.06
IN	-0.06	-0.01	-0.05	-0.03	-0.03	-0.03
IT	-0.31	-0.06	-0.23	-0.05	-0.20	-0.04
JP	-0.21	-0.03	-0.18	-0.06	-0.44	-0.05
KR	-0.31	-0.01	-0.13	-0.03	-0.31	-0.05
MY	-0.13	-0.01	-0.08	-0.03	-0.22	-0.06
NL	-0.40	-0.03	-0.15	-0.02	-0.44	-0.03
PH	-0.40	-0.04	-0.15	-0.04	-0.37	-0.03
PL	-0.15	-0.08	-0.25	-0.03	-0.51	-0.06
RU	-0.09	-0.05	-0.32	-0.07	-0.49	-0.08
SG	-0.16	-0.08	-0.11	-0.08	-0.34	-0.04
TH	-0.40	-0.02	-0.20	-0.03	-0.39	-0.10
US	-0.33	-0.05	-0.13	-0.05	-0.33	-0.03
VN	-0.41	-0.07	-0.11	-0.05	-0.28	-0.03

The table shows the decomposition of loss differentials between the main model and the benchmark model, conditional levels of cross-sectional average PD. The loss differential is decomposed into squared bias differentials within each sector, and the corresponding squared idiosyncratic error differentials. The sample are divided into three groups, periods when average PD is below the 0.25 quantile (left panel), periods when average PD is above the 0.75 quantile (right panel) and the other periods (middle panel).

Figure A.1: MSE of Main and Benchmark Model at Different Lags since Publication of Annual Reports
For each country, the MSE at different lags are divided by the zero-lag MSE. The figure shows median, 0.2 and 0.8 quantiles of MSE distribution across countries.

(a) Data released 15 months after closing dates

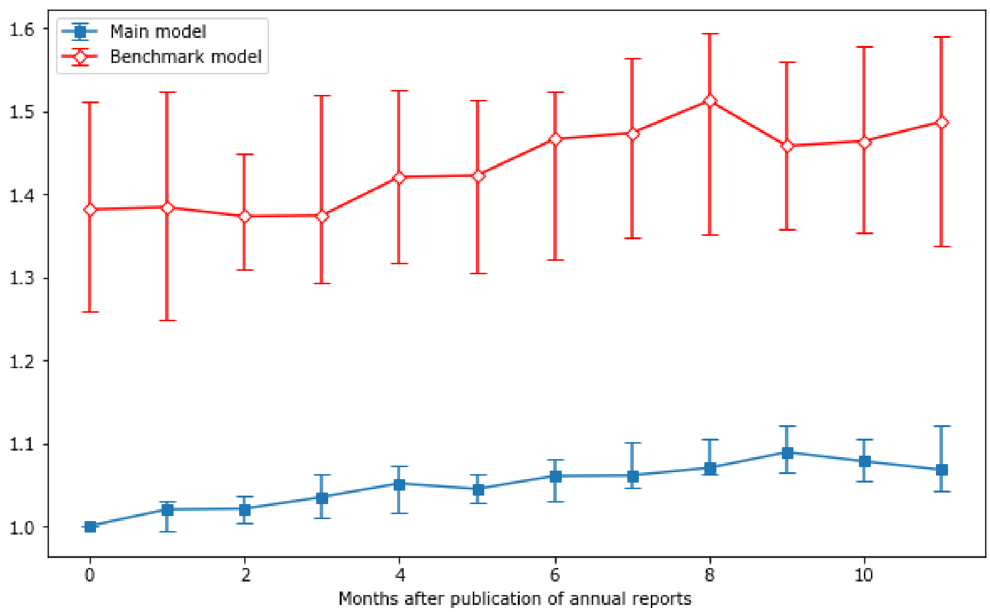
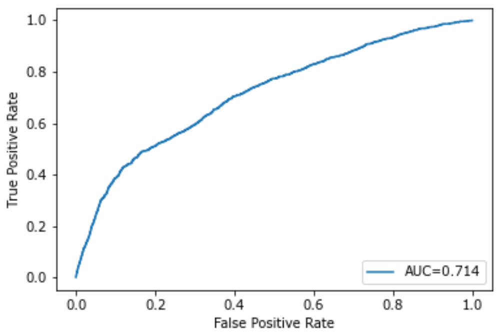
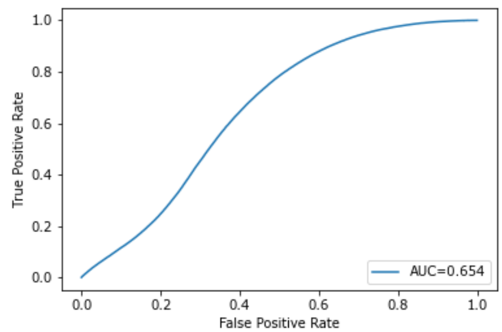


Figure A.2: AUC of predicting non-listed firm bankruptcy (15 month publication lag)

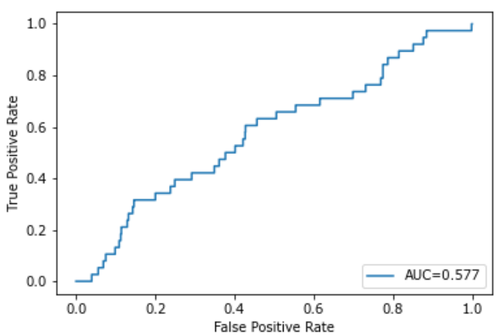
(a) Advanced Asia



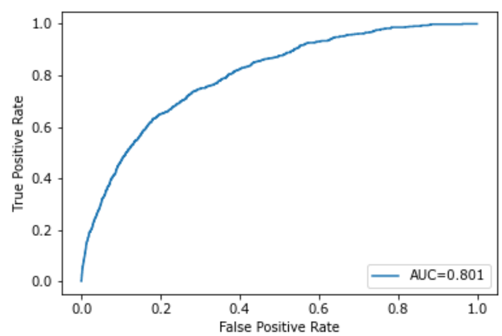
(b) Europe



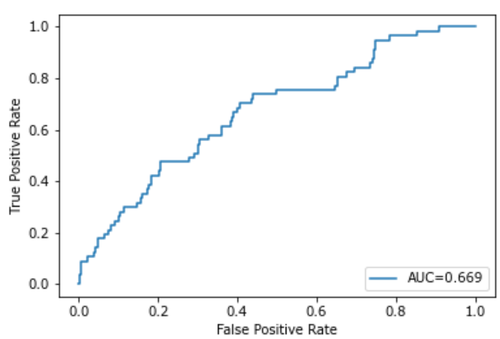
(c) CN



(d) IN



(e) US



Address: 10 Shenton Way, #15-08
MAS Building, Singapore 079117
Website: www.amro-asia.org
Tel: +65 6323 9844
Email: enquiry@amro-asia.org
LinkedIn | Twitter | Facebook | YouTube