In Search of Distress Risk in Emerging Markets*

Gonzalo Asis† Anusha Chari‡ Adam Haas§

April 20, 2019

Abstract

This paper employs a novel multi-country dataset of corporate defaults to develop a model of distress risk specific to emerging markets. The data suggest that global financial variables such as US interest rates and shifts in global liquidity and risk aversion have significant predictive power for forecasting corporate distress risk in emerging markets. We document a positive distress risk premium in emerging market equities and show that the impact of a global "risk-off" environment on default risk is greater for firms whose returns are more sensitive to a composite global factor.

Keywords: emerging markets, distress risk, corporate debt, global factors, default probabilities, asset pricing implications

JEL Classifications: F3, G12, G15, G33

---

*This paper was written under the auspices of the Thematic Fellowship at the Hong Kong Monetary Authority. This research was in part also supported by a NUS-RMI-CRI grant and partially conducted during a visit by Asis to RMI. We thank Sun Wei, Zhifeng Zhang and the team at RMI-CRI for their efforts. We thank Lillian Cheung, Matthew Yiu, Giorgio Valente, Duan Jin-Chuan and seminar participants at the HKIMR and UNC-Chapel Hill for thoughtful comments and suggestions. Special thanks are due to Eric Ghysels, Peter Hansen, Christian Lundblad and Mike Aguilar.

†Graduate Student, Department of Economics, University of North Carolina at Chapel Hill.

‡Professor of Economics, Department of Economics & Professor of Finance, Kenan-Flagler Business School, University of North Carolina at Chapel Hill & NBER. Email: achari@unc.edu. Address: Department of Economics, CB #3305, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

§Graduate Student, Department of Economics, University of North Carolina at Chapel Hill.

1 Introduction

Non-financial corporate debt in emerging markets surged from $4 trillion in 2004 to over $25 trillion in 2016 (IIF, 2017). In view of heightened levels of leverage and worsening solvency positions, there is rising concern about deteriorating health of emerging market firms —recent evidence suggests that the share of debt held by troubled firms is the highest in over a decade (IMF, 2015).[1] Whether through links with the global financial system or through macroeconomic effects, a wave of corporate defaults in emerging markets could trigger broader financial stress (Shin, 2013; Acharya et al., 2015). For example, Fed Chair Jerome Powell highlights concerns that global debt paired with other macro conditions, such as the risk of asset price drops and currency depreciation, could damage the ability of emerging market firms to repay their debts.[2]

Yet there is little systematic research on the determinants of corporate distress specific to emerging markets.[3] An exception is Altman (2005), who adapts a longstanding bankruptcy risk model (Altman, 1968) to the idiosyncrasies of emerging market firms. Recent approaches principally focused on US data have made significant strides to further develop the methodologies to measure probabilities of default. Notable examples are the frailty factor models introduced by Duffie et al. (2009); and the logit models put forth by Shumway (2001) and refined by Campbell, Hilscher, and Szilagyi (2008). However, we find that logit model specifications developed using US data have low predictive power when applied to the emerging market context.

This paper uses a novel multi-country dataset on corporate defaults to study factors that drive corporate distress in emerging markets. We note that extant models do not account for emerging market vulnerabilities to global macro shocks such as advanced economy monetary policy changes, US dollar movements, or shifts in global liquidity and risk aversion. Given the surge in "search for yield" flows from advanced economies to emerging markets during the unconventional monetary policy period, concerns about reversals in these flows during monetary policy normalization in advanced economies could exacerbate corporate distress risk in emerging markets. The objective is to develop a model of distress risk that allows us to quantify the importance of global shocks on corporate distress in emerging markets as a class of assets.

[1]"IMF Flashes Warning Lights for $18 Trillion in Emerging-Market Corporate Debt," Wall Street Journal, September 25, 2015.

[2]"Prospects for Emerging Market Economies in a Normalizing Global Economy," Speech by Jerome Powell, October 12 2017.

[3]We use "default risk" and "distress risk" interchangeably throughout the paper.

Our paper makes three contributions to the literature on corporate defaults. First, it determines precisely which accounting, market, and macroeconomic variables are associated with corporate distress risk in emerging markets – and compares them to those in advanced economies. A number of fundamental idiosyncrasies suggest a modified approach to analyze corporate vulnerabilities in an emerging market setting. Given the documented spillover effects of advanced economy monetary shocks (Fratzscher, Lo Duca, and Straub, 2016) and the impact of changes in international investor risk tolerance on emerging market capital flows (Rey, 2015; Chari, Dilts, and Lundblad, 2017), we suggest that a set of global financial variables can play a key role in predicting corporate distress in emerging markets.[4]

Second, the paper improves current tools to predict corporate distress in emerging markets. Instead of simply estimating US-based models using emerging market data, our specification includes a set of explanatory variables that maximizes predictive power for emerging markets. Additionally, the introduction of stock returns' sensitivities to global factors adds a new dimension to our understanding of how distress risk operates through financial markets. Third, we find a positive distress risk premium in emerging market stocks by examining the pricing of financially distressed firms. We use the probability of default measure developed in the first part of the paper to explore the performance of distressed stocks between 2002 and 2015.

Our dataset consists of corporate default events and a set of firm-specific and macroeconomic explanatory variables. A majority of the data come from the CRI database.[5] The database contains detailed default, accounting, and market data for over 60,000 exchange-listed firms in 119 countries between 1990 and 2016. The dataset is novel in its broad coverage of balance sheet variables and, especially, of default events in emerging markets.

We begin by estimating a logit model of probability of corporate default on a set of firm-specific accounting and market variables, as well as variables reflecting global financial conditions. Our findings are as follows. First, the evidence suggests that the five-year US Treasury rate, the Fed funds rate, and the VIX are correlated with distress risk, even after controlling for firm-specific variables and country fixed effects. Second, introducing a dummy variable indicating whether a firm has defaulted in the past has a very positive impact on the model's predictive power – a novel result in the literature, to the best of our knowledge. A model that includes both types of variables and the prior-default dummy yields a much higher explanatory power for emerging market firms than model specifications

[4]The currency denomination of emerging market corporate debt is a particular source of concern.

[5]CRI is the Credit Research Initiative of the National University of Singapore.

that focus exclusively on accounting and market variables.

Third, computing marginal effects of our probability of default model allows us to examine the economic significance of our coefficients. Leverage and cash have the largest average marginal effects on the probability of default – a one-standard-deviation increase in the predictor is associated with 0.4 and -0.52 percentage point changes in the probability of default, respectively. From the set of global variables, the change in the bilateral exchange rate with the US dollar and five-year US Treasury rates have the largest average marginal effects. Plots of predicted probabilities – for all values of each explanatory variable while keeping all other predictors constant at their mean – reveal variations in the range and curvature of these marginal effects.

Next, we focus on firms whose returns are most sensitive to global financial conditions in order to explore whether stock returns carry information about the impact of the global financial environment on default risk. We label these sensitivities "global betas", and they are extracted from firm-specific time series regressions of stock returns on a global variable, controlling for market returns. Introducing dummies for the tercile of firms with most negative global betas (i.e., firms most negatively affected by increases in the US dollar, sovereign spreads, US interest rates, VIX, and TED spread) reveals that, for five-year Treasury rates, VIX, and TED spread, the effect of increases in the global variable on the probability of default is larger for firms with most negative betas. Further, a composite global beta measure helps us show that the effect of a global risk-off environment on distress risk is greater for firms whose returns are more sensitive to global factors.

Finally, we explore the asset pricing implications of our measure of distress risk. Asset pricing theory suggests investors should demand a premium for holding stocks at risk of default. However, prior literature using US data finds an inverse correlation between distress risk and future stock returns (Campbell, Hilscher, and Szilagyi, 2008). We construct ten portfolios sorted by firms' predicted probability of default and find strong evidence of the presence of a distress risk premium in emerging market stocks. Future twelve-month stock returns are monotonically increasing in the probability of corporate default, a trend that is robust to the inclusion of controls for the Fama-French three factors, momentum, short-term reversal, and long-term reversal.

Related Literature: Shumway (2001) introduces a multiple logit model that combines accounting data with a set of market variables comprised of market size, past stock returns, idiosyncratic standard deviation of stock returns, net income to total assets, and total liabilities to total assets. Chava and Jarrow (2004) improve forecasting by shortening the observation intervals to monthly frequency and find the existence of an industry effect. Campbell et al. (2008) build on the work of Shumway (2001). Their paper uses US data to show that firms are more likely to enter distress if they have higher leverage, lower profitability, lower market capitalization, lower past stock returns, more volatile past stock returns, lower cash holdings, higher market-book ratios, and lower prices per share.[6] An important asset pricing implication of Campbell et al. (2008) is that stocks of distressed companies experience abnormally low returns.

Das et al. (2007) prove that the models mentioned above don't explain all the systematic risk that contributes to firms' probability of default. Duffie et al. (2009) introduce frailty factors – latent common factors – that help explain these shared risks, though at the expense of computational ease. In an attempt to reduce the computational burden, Duan and Fulop (2013) combine frailty factors with a forward intensity approach first developed in Duan et al. (2012).[7]

A small set of papers develop bankruptcy models for emerging markets. Notably, to adjust the Z-Score to the different environment in emerging markets Altman (2005) introduces the modified Z-score.[8] Subsequent studies focus on expanding the types of variables included in the predictive model (Hernandez-Tinoco and Wilson, 2013) and applying US-specific determinants of bankruptcy to other countries (e.g. NUS-RMI, 2016).

Other related research focuses on specific financial sheet variables to identify country-wide corporate distress risk. Alfaro et al. (2017) use firm-level data to show that the correlation between leverage and corporate fragility is time-varying and strongest for large firms and times of local currency devaluations. Bruno and Shin (2016) also focus on firms' balance sheets, as they point out the increase in cash holdings among non-financial corporations in emerging markets. The papers argue that firms are not accumulating cash as a precautionary measure, but to engage in cross-border speculative activities; i.e., to take advantage of interest rate spreads between advanced and emerging economies.

[6] The authors define distress as either filing for bankruptcy, getting delisted, or receiving a D rating. The authors use Shumway's (2001) specification as base and make modifications that improve the model's predictive power. First, they divide net income and leverage (both explanatory variables) by market value of assets instead of book value. Second, they add corporate cash holdings, Tobin's Q, and price per share to the set of explanatory variables. Third, they study default forecasts at different horizons, finding market capitalization, market-book ratio, and equity volatility the most consistently predictive characteristics of corporate distress, and demonstrating the increased importance of balance sheet versus market variables as the horizon increases.

[7] Although not as closely related to this paper, a branch of the literature has developed structural models of default risk. The most influential structural approach was pioneered by Merton (1974) and improved by Oldrich Vasicek and Stephen Kealhofer (Vasicek, 1984; Kealhofer, 2003a; Kealhofer, 2003b) in what became known as the Vasicek-Kealhofer (VK) model. Crosbie and Bohn (2003) base on these prior works their Distance to Default (DTD) measure: DTD = Firm Net Worth / (Market Value of Assets · Asset Volatility). Because a firm will default when its net worth reaches zero, its distance to default will also equal zero at that point. The authors then translate this structure into a probability of default by empirically mapping DTD and historical US default data.

[8] More information on the specifics of the modified Z-score model derivation can be found in Altman (2005).

Hence, the traditional belief that cash increases a firm's repaying ability may not hold in the current environment.

There has been limited research on the drivers and consequences of high currency exposure due to the shortage of reliable data on currency composition of debt.[9] However, the view most widely held is that foreign-currency liabilities are a concern for emerging market non-financial corporations and particularly troubling for firms that do not have natural currency hedges in place (e.g. firms in non-tradable industries). Harvey and Roper (1999) show that high foreign currency-denominated leverage and low profitability were important factors spreading the Asian Financial Crisis.

To the best of our knowledge, ours is the first paper that estimates emerging market-specific probabilities of corporate default and quantifies how the global macroeconomic environment they operate in can affect their ability to remain solvent. Additionally, having a reliable measure of corporate default risk allows us to explore the behavior of distressed stocks in emerging markets.

The rest of the paper is organized as follows. Section 2 explains the methodology. Section 3 describes the data. Section 4 presents the results of logit regressions of the probability of default and introduces global betas as predictors of default. Section 5 shows the asset pricing implications of our measure of distress risk. Section 6 concludes.

2 Methodology

Although leverage levels receive substantial attention in the corporate default literature, several studies show the importance of other accounting and market variables in forecasting corporate bankruptcies. Earlier static bankruptcy prediction models used accounting ratios to forecast default (See Altman, 1968; Ohlson, 1980; Zmijewski, 1984). Shumway (2001) points out that static models effectively require arbitrary choices about how long ahead of bankruptcy to observe the firms' characteristics – adding selection bias to the process. In contrast, dynamic forecasting using hazard or dynamic logit models use all available information to determine each firm's bankruptcy risk at each point in time. By including each firm-year as a separate observation, the data used for estimation is much larger and controls for the "period at risk," namely that some firms fail after being at risk for many years and others go from healthy to bankrupt much faster. In addition to accounting for duration depen-

[9] The two major issues compiling accurate data on debt currency composition are: (a) Many corporate reports present the currency composition of their liabilities in the notes of the reports and not in hard data, and (b) the use of offshore intermediaries to borrow funds makes it difficult to establish the residence of the ultimate debt-holder-a problem documented in Avdjiev et al. (2014).

dence, hazard models include time-varying covariates, which provide a changing picture of a firm's health. Campbell et al. (2008) build on the work of Shumway (2001) and improve the set of variables used to predict distress. The authors run a logit model on US data, putting more emphasis on market variables as predictors of distress.

Similar to Shumway (2001) and Campbell et al. (2008), we estimate a model of probability of default using a logit specification augmented by domestic and global macroeconomic factors that have particular relevance to emerging market firms. We assume a logistic distribution for the marginal probability of default over the next period, which is given by:

$$P_{t-1}(Y_{i,t} = 1) = \frac{1}{1 + exp(-\alpha - \beta x_{i,t-1})} \tag{1}$$

where $Y_{i,t} = 1$ in the month t prior to firm i defaulting and $Y_{i,t} = 0$ in all periods when the firm does not default the following month. Firms disappear from the sample only after they experience a bankruptcy event. Firms that do not default at any point in the sample have $Y_{i,t} = 0$ throughout the entire period, including in the month of their departure if they leave the sample for reasons other than default (e.g. merger). The vector of explanatory variables, $x_{i,t-1}$, is known at the end of the previous period. A higher level of $\alpha + \beta x_{i,t-1}$ implies a higher probability of default.

We suggest that the domestic macroeconomic environment may affect the financial health of emerging market firms through demand for their goods and services, wage and borrowing costs, and other input costs. Evidence from the credit risk literature suggests that the incidence of firm failures rises during recessions (Altman and Brady, 2001). Further, inflation risk affects economic growth and uncertainty about the domestic economy. For example, Hernandez-Tinoco et al. (2013) find a significant relationship between default risk and both domestic inflation and interest rates in UK firms. To control for the impact of the domestic economic conditions in the probability of default of emerging market firms, we include a number of domestic macroeconomic indicators and country fixed effects in different specifications of the model.

Furthermore, the globalization and increased interconnectedness of financial markets propagates the transmission of financial and economic conditions from developed to emerging markets. For instance, a 2015 report by the IMF shows that the increase in corporate debt in emerging markets was driven by global factors. Shin (2013) argues that global liquidity increased in response to the Global Financial Crisis, while Jotikasthira et al. (2012) report that "global funds substantially alter portfolio allocations in emerging markets in response to funding shocks from their investor base." Due to their high reliance on international markets for funding, the listed firms that make up our dataset are likely affected by these changes in global conditions. For this reason, we also include a number of global variables that may influence the distress risk of emerging market firms. Section 4.2 discusses in detail the methodology to compute global betas as a measure of emerging market risk exposure to a range of global factors.

2.1 Model Performance

The existing literature uses a number of measures of a model's predictive power, most of which involve ranking firms by their estimated probability of default. However, studies differ in the number of firms and defaults, the size of the quantiles to group firms, and the allocation of distressed firms across quantiles, making comparisons across models difficult. Chava and Jarrow (2004) improve comparability by relying on the Receiver Operating Characteristics (ROC) score. The ROC score, also known as "area under the power" or "area under the curve" (AUC), uses the cumulative fraction of defaults as a function of the ordered population of firms from most to least likely to fail as predicted by the model. The ROC curve plots the true positive rate (e.g., firms that actually default) against the false positive rate (i.e., firms that are predicted to default but do not in fact default) for different threshold settings. In machine learning, the true-positive rate is referred to as the sensitivity, or the probability of detection.

Figure 1 presents an example. Point A on the "High-Sensitivity Model" curve tells us that the 20% of firms that a particular model identifies as most likely to default include 70% of the firms that go on to default the next month. Point B in the "Low-Sensitivity Model" curve signals that it takes 50% of firms ordered from most to least likely to default for the model to identify 70% of defaulting companies. We compare the two models by computing the area under each of the curves (AUC). A larger area indicates that the model is correctly predicting more distressed firms as being likely to fail. An AUC of 0.5 indicates no discriminatory power, and the closer the score gets to 1 the better the model identifies distressed firms.[10] Contributing to the interpretation of the AUC, Hanley and McNeil (1982) show that the score obtained by ranking observations by estimated likelihood of failure represents the probability that a failed subject will be ranked ahead of a randomly chosen healthy subject.

[10]See Sobehart and Keenan (2001) for more details on the ROC score.

To measure goodness of fit, we use McFadden's pseudo-R^2, which compares the model's likelihood (L) to that of a model consisting of only a constant (L_0), i.e. the average default rate in the sample. Specifically, it is computed as $1 - \frac{log(L)}{log(L_0)}$ and can be interpreted in the same manner as the standard R^2 (between 0 and 1, increasing in model fit).

2.2 Variable Selection

Given the large number of default predictors found by the literature and the lack of studies specific to emerging markets, we want to make sure we include only the most relevant set of explanatory variables in our specification. Hence, in a robustness exercise we add the Least Absolute Shrinkage and Selection Operator (LASSO) routine to our estimation. This procedure allows us to select, from a large set of explanatory variables, the subset with highest predictive power. The LASSO constrains the sum of the absolute value of the coefficients during the maximum likelihood estimation, forcing some coefficients to equal zero.[11] Specifically, it minimizes the following (negative) likelihood function, which includes a constraint on the sum of the coefficients:

$$\sum_{i=1}^{n}(-Y_{i,t+1}(\alpha + \beta x_{i,t}) + log(1 + exp\{-\alpha - \beta x_{i,t-1}\})) + \lambda(\sum_{k=1}^{p} |\beta_k|) \tag{2}$$

We use cross-validation to determine the level of λ that gives the best model fit. Next we choose the set of variables within one standard error of the optimal λ that maximizes the in-sample ROC. The result is enhanced prediction accuracy and ease of interpretation of the coefficients. Tian, Yu, and Guo (2015) employ this routine on US data, and their resulting set of accounting and market variables achieves higher in- and out-of-sample predictability than Campbell et al. (2008). Starting from theoretical arguments and existing advanced-economy specifications, the LASSO routine allows us to add statistical rigor to the variable-selection component of the exercise.

3 Data

Our dataset consists of corporate default events and a set of firm-specific and macroeconomic explanatory variables. A majority of the data come from the CRI database, the Credit Research Initiative

[11] The variables that enter the LASSO procedure are standardized in advance in order for LASSO to accurately compare the importance of each variable. Dummy variables are standardized as well, which prevents us from interpreting their LASSO coefficient in the usual manner – this is not problematic since we re-run our model using the set of variables selected by LASSO.

of the National University of Singapore, accessed on December 1, 2016. The CRI database contains detailed default, accounting, and market data for over 60,000 exchange-listed firms in 119 countries between 1990 and 2016. The countries in our analysis are those classified as Emerging Markets by MSCI during the majority of our sample period (1995-2016): Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Jordan, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Russia, South Africa, South Korea, Taiwan, Thailand, Turkey, and Vietnam.

The dataset is novel in its broad coverage of balance sheet variables and, especially, of default events in emerging markets.[12] As Table 1 shows, the CRI database contains information on firms' bankruptcies and other corporate default actions. This is important because countries differ in their definitions of default. To construct our measure of financial distress, we define a default to be any of the events in the "Bankruptcy Filing" (excluding "Petitions Withdrawn"), the "Delisting", and the "Default Corporate Action" (excluding "Buyback options") groups.[13] Delayed payments made within a grace period are not counted as defaults.

Table B1 in Appendix B shows our distress indicator over time for firms with sufficient data to replicate benchmark specifications from existing US studies. The first column shows the number of firm-months of data in each year, the second column the number of default events per year, and the third column the corresponding percentage of firms that experienced a default event. The average default rate in the sample is close to 0.1% per year, with some variation within years. Importantly, there is no strong clustering across time, as the distress indicator displays considerable cross-time variation in the distribution of corporate defaults. The two years with highest share of defaults coincide with the depth of the Asian Financial Crisis. Coverage of accounting variables varies. The number of firm-months and defaults with data for *any* of the variables in Campbell et al.'s (2008) specification is 2,724,716 and 2,150, respectively. However, in order to run the logit model we require every observation have data for *all* explanatory variables included in the regression specification. Due to missing observations and the sparsity of some accounting data, the final sample includes 671,762 observations and 590 default events. This data serves as the basis for our benchmark regression specification.

As seen in Table B2 in Appendix B, the data coverage varies substantially by country, possibly influencing the lack of a clear pattern in the percentage of defaults by year. Comparing our sample against prior studies using US firms, we find that the ratio of defaults to firms is lower in emerging

[12] Market data from emerging markets on stock prices and related variables are fairly accessible from sources such as Datastream, Bloomberg, etc.

[13] The number of Default Corporate Action events is lower than the sum of its sub-components because some events include multiple concurrent actions (e.g. Missed Loan Payment and Missed Coupon Payment)

markets than in the United States. This could be due to a few reasons. First, governments own a percentage of many listed firms in emerging markets and might be more inclined to bail out or recapitalize struggling companies. Second, politically connected large firms may benefit from government bail-outs. Legal system inefficiencies and lengthy court delays, common institutional features in many emerging market economies, may also lead to lower bankruptcy rates.

The set of covariates consists of three types of variables: firm-specific accounting and market variables; domestic macroeconomic variables; and global variables, i.e. variables from outside the emerging market countries and usually related to financial conditions in advanced economies. Consistent with Campbell et. al. (2008), the monthly firm-specific market variables are: log excess stock returns relative to the country's main index (EXRET), log of price per share (PRICE), volatility of daily returns over the prior month (VOL), and the log ratio of market cap relative to the total market cap of all listed firms in the country (RELSIZE). The accounting variables have quarterly frequency and include the ratio of net income to the market value of total assets (NIMTA), the ratio of total liabilities to the market value of total assets (TLMTA), the ratio of cash and short-term assets to the market value of total assets (CASHMTA), and the market-to-book ratio (MB).[14] In some of our specifications we include a dummy variable that equals one if the firm has experienced a default event in the past. Although we would have liked to include a variable indicating the firm's age or listing date, unfortunately good quality data are not available for the firms in our sample

To control for large outliers and possible errors in the balance sheet and market data, we winsorize the firm-specific variables at the 1^{st} and 99^{th} percentile of their distributions[15]. We also lag the accounting ratios (TLMTA, NIMTA, CASHMTA, and MB) by three months to ensure the balance sheet data was publicly available at the time we predict default.

To capture the domestic macro environment in which firms operate, we incorporate four domestic macro variables for each country. The first is the unemployment rate to capture slack in the economy, retrieved from the World Bank. Inflation is the monthly change in CPI from the Bank for International settlements, which reflects pricing pressures in the local economy. Real interest rates come also from the World Bank, and we include them as a proxy for local borrowing costs and liquidity. Lastly, the JP Morgan Emerging Markets Bonds Spread, which measures the average spread on US

[14]Campbell et al. (2008) include time-weighted averages of NIMTA over the previous four quarters and EXRET over the previous twelve months. Due to the sparsity of emerging market data, we would lose too many observations if we required one consecutive year of data for those two variables. We use the single-period definition instead.

[15]Market-to-book ratio is winsorized at the 5^{th} and 95^{th} percentiles in order to deal with firm-months with very small or negative book-to-equity values, which in turn make MB very large.

dollar-denominated bonds issued by sovereign entities over US Treasuries, incorporates international investors' perception of the government's credit risk.

The set of global macro variables includes the monthly change in a country's exchange rate against the US dollar, since it is a major determinant of firms' revenues from abroad and their ability to repay debts denominated in dollars.[16] We also include the monthly change in the sovereign spread measures the change in the country's perceived credit quality compared to the United States, often driven by increases or decreases in capital flows to the emerging country's financial markets.

Moving on to variables computed only with developed-market data, the CBOE Volatility Index, known commonly as the VIX, measures the market's expectation for 30-day volatility in the S&P 500. A higher VIX typically denotes a general increase in the risk premium and, consequently, an increase in borrowing costs of emerging market firms. Rey (2015) finds one global factor correlated with the VIX that drives the price of risky assets around the world . The effect of changes in US rates on capital flows to emerging markets has also been established in the literature (Chari, Dilts and Lundblad, 2017), and Bruno and Shin (2015) introduce bank leverage as a mechanism through which changes in US monetary policy impact international capital flows. To address the interest rate effect, we include both the US federal funds rate and the five-year US Treasury rate. The federal funds rate is indicative of monetary conditions and changes in monetary policy in the United States, whereas the five-year Treasury rate serves as the risk-free rate against which investors in advanced economies evaluate the payoffs of all other assets of similar maturities. Lastly, the TED spread is a proxy for perceived credit risk in the US economy, and it is computed by subtracting the 3-month Treasury bill rate from the 3-month LIBOR rate. Due to the correlation between TED spread and VIX, we use the orthogonal component of the two, i.e. the residual of a regression of the TED spread on VIX .

The global variables have monthly frequency and are common to all firms in the sample. Appendix A defines variables and their sources in greater detail.

3.1 Summary Statistics

Table 2 reports simple equally-weighted means of the explanatory variables, as well as t-tests for means. The first column presents means for the full sample, the second column for the Default group, and the third for the Bankrupt group – a subset of the Default group. The fourth and fifth columns

[16] The percentage of corporate debt denominated in US dollars has increased dramatically since the Global Financial Crisis, as shown by IMF (2015) and others.

show whether there is a statistically significant difference in means between the full sample and the Default and Bankrupt groups, respectively.

The firm-specific covariates show that firms in the Default group exhibit lower excess returns, stock prices, and volatility. Firms under duress are also smaller, the average firm in the group comprising 0.01% of the country's market cap, compared to over 0.04% for the average firm in the full sample.

Looking at firm balance sheets, firms one month away from default differ from the full sample in the expected direction – and the difference in all four mean accounting ratios is larger for firms in the Bankrupt group. Distressed firms have lower profitability and are on average making losses the month before failing to pay their obligations, compared to an average net income to total assets of 0.004 in the full sample. These firms also have higher leverage (0.578 and 0.759 for Default and Bankrupt groups, respectively) than the overall population (0.366), as well as lower cash holdings over total assets: 0.045 and 0.024 for the Default and Bankrupt groups, compared to a full sample average of 0.082. Both ratios are suggestive of firms' diminishing ability to repay their upcoming liabilities. Lastly, troubled firms have low book value of equity relative to their market capitalization, resulting in higher market-to-book ratios of 2.673 (Default) and 4.400 (Bankrupt), compared to 2.121 for the full sample. All summary statistics described so far are consistent with those in Shumway (2001) and Campbell et al. (2008), except for the fact that volatility of stock returns is lower for firms one month away from default.

We also introduce a variable that, to the best of our knowledge, has not been used in the literature: an indicator of whether a firm has defaulted in the past. Comparing the means of distressed firms and the full sample, we find in the Default and Bankrupt groups a much higher percentage of firms which have already suffered a default event.

The interpretation of the differences in the means of the domestic macroeconomic variables is less clear, given that some countries will have structurally higher levels of interest rates, inflation, unemployment or sovereign spreads than others throughout the sample. In any case, we find that domestic macroeconomic environment for the Default group is characterized by lower unemployment and real interest rates.

On the other hand, the direction of the effect of global variables on corporate distress is intuitive based on how they affect firms' ability to roll over or pay off their financial obligations to avoid default. We would expect an environment of high interest rates in the US to lower the search for yield and corresponding demand for riskier emerging market debt instruments. The summary statistics support this hypothesis, with firms defaulting in times of higher five-year Treasury and Fed Funds rates: 2.890% and 1.869%, respectively, compared to 2.363% and 1.235% in the full sample. Also as expected, defaults occur on average during times when a country's sovereign spread is increasing more than on average during our sample period. Lastly, the Default group is characterized by having a higher TED spread; that is, higher global liquidity risk. VIX levels and exchange rate dynamics are not significantly different between distressed firms and the full sample.

4 Results

4.1 A General Model of Default Risk

As a first step to tailoring the default risk model to emerging market firms, we run a variable selection exercise using the Least Absolute Shrinkage and Selection Operator (LASSO) to choose from a set of accounting and market variables the combination with highest predictive power for emerging markets.[17] While Campbell et al. (2008) (referred to as CHS intermittently hereafter) show their model outperforms other prior specifications in the US-based literature, we examine which combination of firm-specific variables will achieve the highest prediction power for emerging market firms. Therefore, we construct nine other accounting ratios that show some explanatory power in the existing literature, and we add them to the eight accounting and market variables in CHS. The results of the LASSO procedure do not show strong evidence that any subset of accounting and market variables specific to emerging markets outperforms those used by CHS. We therefore use the CHS specification with accounting and market variables as a baseline and examine whether including domestic and global macro variables enhances model performance.

Before moving on to our general model of default risk, we address multicollinearity concerns associated with our multivariate framework. Table B3 in Appendix B shows the correlation matrix of the variables in our model and, in the last two rows, two popular measures of multicollinearlity, the Tolerance value (TOL) and its reciprocal Variance Inflation Factor (VIF), for each of the regressors. VIF is computed as $1/(1 - R_k^2)$, where R_k^2 is the R^2 value of a regression of factor k on all others. TOL is simply $1 - R_k^2$. VIF values larger than 10 are typically considered suggestive of multicollinearity in a model. In our specification, no variable has VIF $\geq$ 10, and only the Fed Funds Rate and 5-year

[17] The methodology section provides details.

Treasury rate have VIF > 5, presumably due to the high correlation between the two. The correlation between Fed Funds Rate and 5-year Treasury rates is 0.888, the only pairwise correlation larger than 0.6 in absolute value among all our variables

In order to estimate the model with country fixed effects we must drop any country with no defaults during the period being used for estimation. The effect on our sample is small: the sample size only falls from 671,762 to 589,224 firm-months, and we are left with firms from Argentina, Brazil, China, India, Indonesia, Malaysia, Mexico, Philippines, Poland, South Korea, and Thailand.

Table 3 shows the results of six iterations of the multivariate logistic regression. Column 1 estimates a baseline specification, using the same set of accounting and market variables in Campbell et. al. (2008). The specification yields a pseudo-R^2 of 0.124 and an AUC of 0.865. The results in Column 1 suggest that while profitability and cash are inversely correlated with the probability of default, leverage is positively correlated with the probability of default. With respect the the market variables, excess returns and the firm's stock price are inversely correlated with the probability of default while higher valuations associated with the market to book ratio are positively correlated. These coefficients are statistically significant at the 1% level and are consistent with the findings in Campbell et al. (2008). The exceptions in our pattern of findings is that the volatility of returns is inversely correlated with the probability of default and the market capitalization variable is not significant. In sum, our baseline results imply that a firm is more likely to default next month if it has lower excess stock returns, a lower stock price, lower volatility of returns, lower profitability, higher leverage, less cash, and a higher market-to-book ratio.

Next, we add a dummy variable signaling whether a firm has defaulted in the past, and we find that it greatly increases explained variation and predictive power (Column 2). The pseudo-R^2 goes up to 0.2 and the AUC to 0.907. We keep the prior default event dummy in the set of firm-specific variables moving forward. To the best of our knowledge ours is the first paper to include this explanatory variable that is remarkably robust across specifications. Including a wider subset of events in our definition of "Default" rather than being restricted to outright bankruptcy, allows us to examine the impact of prior distress states on the current probability of default.

In the third column we add the domestic macro variables – unemployment, inflation, real interest rates, and sovereign spreads – to the regression. The pseudo-R^2 increases to 0.235, but the AUC falls to 0.888, suggesting a better model fit but not better predictive power. We find that default is associated with higher unemployment, lower real interest rates, and lower sovereign spreads, after controlling for firm-specific accounting and market variables. The inclusion of the domestic macro environment makes firm size negatively correlated with the probability of default.

Column 4 presents the results of a model that consists of the basline accounting and market variables, the prior default dummy, and global variables. The higher AUC than in Column 3 suggests that global variables contribute more predictive power than domestic variables after controlling for firm-specific covariates.The coefficients suggest that default risk is associated with higher five-year Treasury rates, lower Fed Funds rates – likely an adjustment for the five-year rates, since Fed Fund rates are unconditionally positively correlated with default – and a higher TED spread. In other words, after controlling for firm-specific accounting and market variables, emerging market firms are more likely to default when US five-year Treasury rates are high, Fed Funds rates are low, and credit risk in the US is more prevalent.

A specification that includes both domestic and global variables is presented in Column 5. Notable results are the significantly positive coefficient on VIX and the fact that real interest rates are the only domestic macro variable to remain (negatively) associated with default risk.

Before introducing country fixed effects, we again rely on LASSO to verify that no subset of variables would deliver a model with better fit. Figure C1 shows the path of the coefficients during the LASSO estimation, with decreasing λ (from left to right) loosening the constraint on the absolute value of the standardized coefficients and allowing more variables to enter the regression. The coefficients the procedure returns for the λ that yields the best fit are shown in Column 2 of Table C1. LASSO eliminates all global variables but the 5-year Treasury rate, as well as volatility of returns, firm size, and cash. When we run a logit regression using the explanatory variables selected by LASSO (Column 3), the sign and significance levels of the coefficients match those of our benchmark model (Column 1). The AUC of the LASSO model is only marginally larger than that of the full model, suggesting that the entire set of explanatory variables is almost as good at explaining and predicting default as the best subset as selected by LASSO.

Finally, a specification that includes country fixed effects yields the best predictive power, with an AUC of 0.914. Including country fixed effects allows us to control for country-specific differences in characteristics like legal system, bankruptcy laws, and state intervention, all of which are difficult to quantify. The ROC curve associated with this and the baseline model is shown in Figure 2. Figure 3 plots, for each quarter, the number of defaults predicted by this model against the number of actual defaults. As in the baseline specification, we find that a firm is more likely to default next month if it has low excess returns, price, profitability, and cash; as well as high leverage and market-to-book ratio. Adding country fixed effects causes relative size (a firm's market cap divided by the total market cap of all listed firms in the country) to be positively correlated with the probability of default. While the corporate default literature finds the opposite relationship for firms in advanced economies (e.g. Campbell et al., 2008; Hernandez-Tinoco et al., 2013), Alfaro et al. (2017) find that firm size is positively correlated with corporate fragility, measured by Altman's Z-score. Volatility of returns is no longer significant, and VIX becomes statistically significant and positively correlated with default, at the expense of the TED spread, which loses its significance.

Being the specification with highest predictive power, we use Column 6 as our measure of probability of default in the remainder of the paper. We test the robustness of our estimates by running two out-of-sample tests of predictive power. First, we estimate the probability of default model one time using data from the earliest 70% of our sample and use the estimates to compute the AUC for each month in the remaining 30% of the sample. Second, we estimate the model in a recursive manner (increasing the estimation window every month, starting with the earliest 60% of data) and predict default on the following month. Both methods yield an AUC of 0.88, compared to an in-sample AUC of 0.914.

Lastly, we compute average marginal effects of each of the regressors in Column 6 of Table 3. Marginal effects allow us to speak about the economic significance of the coefficients; i.e., the effect on the probability of default of changes in a specific predictor variable while keeping all other predictors constant. We present these results in Table 5, sorted from largest negative effect to largest positive effect. We find that leverage and cash have the largest average marginal effects, such that a one-standard-deviation increase in the predictor is associated with 0.45 and -0.4 percentage point changes in the probability of default, respectively. From the set of global variables, ΔFX and five-year Treasury Rate have the largest average marginal effects. Figure 4 shows the vectors of predicted probabilities for the entire range of each explanatory variable in our model, while keeping all the other predictors constant at their means. The red lines are 95% confidence intervals, computed using the 95% confidence intervals of each variable's coefficient in the logit regression. The plots allow us to compare the probability of default at various levels of each variable for the average firm. The convexity in the pattern of stock price and market capitalization predicted probabilities stands out against the more linear excess returns and volatility of returns. The patterns suggest that the tails of the two former variables are relatively more harmful to firms than that of the latter. The flatness of the volatility of returns plot is consistent with its coefficient's lack of statistical significance in the logit regression (Column 6 in Table 3). The leverage plot reveals a large spread in the probability of default between the two extremes, which increases faster at high leverage levels. Since Prior default is a dummy, the line in the plot is simply connecting firms' probability of default when the variable equals zero (no prior default) and one (the firm has defaulted in the past). Consistent with the logit regression coefficients, predicted probabilities for the range of ΔSovereign spread and ΔFX are flat, while the 5-year Treasury plot shows increasing predicted probabilities. The 5-year Treasury rate is also the only global variable with visible convexity. Lastly, the plots show positive and negative slopes for VIX and Fed Funds rate, respectively, but a flatter curve for the TED spread.

4.2 Global Betas

If stock returns accurately carry information about the impact of global factors on firms, we may expect the default risk of corporations with returns more sensitive to global factors to be more correlated with such variables. However, the global variables in the logistic regression may mask the fact that some emerging market firms are more dependent on or exposed to global markets than others. In other words, when we include global variables in our baseline model of probability of default, the average effect of these factors on our entire sample might hide stronger coefficients and predictive power for the more globally-exposed firms.

In order to test the hypothesis that stock returns may contain information that varies across firms by the degree of their global exposure, we compute firm-specific betas of stock returns to each of the global factors in our model. Specifically, we run a time series regression for each firm and global factor, conditional on having at least two years of data on returns and the global variable. The dependent variable is the firm's stock returns, and the explanatory variables are the global factor and the returns of the country's main stock index. We take the resulting coefficient on each global factor to represent the sensitivity of the firm's returns to the global factor, after controlling for the country's returns. Having computed betas for each of the global factors, we select the tercile of firms with most negative betas, i.e. whose returns fall most with increases in the global factor. In the case of the change in the exchange rate, we choose the tercile of firms with most positive betas; that is, whose returns fall most with increases in the rate of change of the US dollar relative to the local currency. Once our firms are sorted by betas, we create a dummy variable that indicates whether a firm belongs to the top tercile.

Panel A in Table 4 reports the results of logit regressions of probability of default where the explanatory variables are the global variable and the interaction of that global variable with the top-tercile beta dummy. The coefficient on the interaction term tells us whether the magnitude of the impact of each global factor on the probability of default differs for the subset of firms with most sensitive returns to that factor. We find positive, statistically significant coefficients in the top-third dummy interactions for five-year Treasury rates, VIX, and Fed funds rate (Columns 3, 4, and 5). This implies that the harmful effect on the probability of default of increases in these variables is larger for the stocks which fall most during increases in those variables. For instance, the risk of default increases more with VIX for firms with most negative VIX betas. To verify that this result is robust to the inclusion of firm characteristics, in Panel B we control for the eight accounting and market variables in our baseline regressions. We find that the interaction coefficient is significant for the same three variables - five-year Treasury rates, VIX, and Fed funds rate. In addition, the unconditional effects of increases in sovereign spreads and the TED spread also become positive and significant once firm-controls are included. We can therefore conclude that the difference in the impact of the global factors between firms with more or less sensitive returns remains robust to the inclusion of firm characteristics between the two groups. Further, the firm-specific variables retain the same signs and levels of statistical significance as in Table 3.

Combining all global variables into one global factor yields further evidence that the sensitivity of returns to global financial conditions is related to the effect those global conditions have on firms' probability of default. We construct an index of return sensitivity to the global environment – which we call the Global Beta Z score – by combining the betas of the six global variables in our model. We standardize the beta for each global factor by subtracting the mean beta across firms and dividing by the standard deviation. We then add together the resulting values of the six factors.[18] The result is a combined measure that gives equal weight to each beta and serves as proxy for how much a firm's returns respond to global financial conditions. A lower Global Beta Z score implies that a firm's returns are more negatively affected by increases in the global variables. We compute a Global Variable Z in the same manner, using the global variables as inputs instead of the betas. A higher Global Variable Z score is associated with a more difficult environment for emerging markets to finance themselves (what is often known as a "risk-off" environment).

In Table 6 we show the results of a logit regression of the probability of default on Global Beta Z, Global Variable Z, and the interaction of the two. We control for firm-specific and domestic macro

[18]We subtract the change in the exchange rate since we want an increase in the US dollar to impact the Global Beta Z score in the same direction as an increase in rates, VIX, sovereign spread, and TED spread.

variables (Column 3) and find that the coefficient on Global Beta Z is not statistically significant, implying that exposure to global financial conditions per se is not a predictor of default. On the other hand, Global Variable Z is positively correlated with default risk; i.e. a firm is more likely to default in global risk-off conditions. Additionally, the interaction of the two returns a significant, negative coefficient. This tells us that the effect of a risk-off environment on default risk is larger for firms whose returns respond more negatively to such global conditions, all else equal. It is worth noting that the coefficient of the interaction term is not significant when computing the global betas using local currency returns rather than US-dollar returns.

We can therefore conclude that, for some global factors like the five-year Treasury rates and for a composite global factor, the sensitivity of a firm's returns to the factor(s) affects the extent to which its solvency depends on the level of such factor(s). There are at least two possible explanations behind this connection between default risk and market betas. First, the stock market captures the effect of global conditions on the firms' probability of default, and the price responds more sharply than for other firms. Second, the fact that returns respond more strongly to the global environment increases the firm's probability of default. In other words, the larger response of returns in some firms accentuates the direct impact of the global conditions on the firm's ability to remain solvent. Should the first explanation hold, it would suggest a distress risk premium exists in emerging market stock returns. We explore this and other asset pricing implications of our measure of probability of default in the next section.

5 Asset Pricing

We use our estimated probability of default (Column 6 in Table 3) to study the stock returns of distressed firms in emerging markets. As was the case with the distress risk measure, research on the distress risk premium has been mostly focused on US equities (e.g. Fama and French, 1996; Vassalou and Xing, 2004; Campbell et al., 2008). Asset pricing theory suggests investors should demand a premium for holding stocks at risk of default. However, Vassalou and Xing (2004) and Campbell et al. (2008), among others, find the opposite: stocks of firms with a high probability of default yield lower returns than their safer or more solvent counterparts. Campbell et al. (2008) show this result holds even after controlling for Fama-French factors and a momentum factor. The findings have important implications for the understanding of risk factors in asset prices, since distress risk is often argued to be the reason behind the small cap and value premia (Chan and Chen, 1991; Fama and French, 1996).

We test whether a distress risk premium puzzle exists in emerging market stocks. Every month between January 2002 and December 2015 we estimate our measure of next-month probability of default using all prior data in the sample to prevent look-ahead bias. In the first month, we sort all stocks based on this predicted probability of default and construct ten portfolios of equal size, placing the stocks with lowest distress risk in Portfolio 1 and those most likely to default in Portfolio 10. We re-balance the portfolios every month thereafter based on the stocks' updated distress risk, again placing the least and most likely to default in Portfolios 1 and 10, respectively. As a proxy for expected returns, we use average realized returns over the 12 months after distress risk is computed. Next-month realized returns leave little room for information surprises to cancel out, and, by looking at returns over a longer horizon, we average out temporary over- and under-performance due to idiosyncratic events unrelated to firm health.

Panel A of Table 7 shows each portfolio's average estimated probability of default, its twelve-month average monthly returns, and the results of a regression of returns on six common factors. The spread in the probability of default across portfolios is large: the average firm in the portfolio with lowest default risk has just a 0.005% probability of failing next month (Column 1), compared to 1.17% for the average firm in the riskiest decile (Column 10). The average twelve-month returns reported in the second row are monotonically increasing in probability of default, consistent with a positive risk premium associated with distress. The safest and riskiest portfolios return 0.5% and 1.4% per month, respectively. These results don't necessarily imply the existence of a distress risk premium in emerging market stocks, since our measure of distress risk may be associated with other factors that demand premia of their own.

To address this, we control for six common factors from the literature in order to separate the distress risk premium from other sources of risk premia that may be present in our sample. The first three factors we control for are the three factors from Fama and French (1993): excess market returns (RM), Small minus Big (SMB), and High value minus Low value (HML). The other three are momentum, short-term reversal, and long-term reversal. Below we describe the computation of these factors for the emerging market firms in our sample in greater detail.

We compute three factors following Fama and French (1993). The market factor, RM, is the return on the market minus the risk-free rate. To account for the various countries in our sample we construct a weighted average of returns of the main index in each country, where the weights are based on the number of stocks from each country in our sample. The Fed Funds rate serves as proxy for the risk-free rate.

The other two factors are size and book-to-market. Each January, we sort all stocks by market capitalization and divide the sample in two groups of equal size: Small and Big. We also sort all firms by book-to-market and use the 30th and 70th percentiles to divide the cross-section into three groups: Low, Neutral, and High. We then construct 6 portfolios from the intersection of the two sorting results and compute their simple monthly average returns: SL, SN, SH, BL, BN, and BH. Lastly, we find the returns of portfolios mimicking the Size (Small-Minus-Big, or SMB) and Book-to-Market (High-Minus-Low, or HML) factors as in Fama and French (1993):

$$SMB = \frac{1}{3}(SL + SN + SH) - \frac{1}{3}(BL + BN + BH)$$

$$HML = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL)$$

We rebalance the portfolios every January and end up with a monthly time series of returns for each factor-mimicking portfolio. For the remaining three factors, momentum is computed as the return in the prior year excluding prior month, short-term reversal as the return in prior month, and long-term reversal as the return in prior five years excluding prior year.

A regression of 12-month average returns on the six factors returns the following results. We find that high failure risk portfolios have higher betas on RM, SMB, and HML. While the HML coefficients are largest on average across deciles, RM and SMB experience larger relative increases between safest and riskiest portfolios. The positive sign on these coefficients implies that the larger returns of riskier portfolios can be partly explained by Fama and French's three factors; correcting for them reduces the outperformance of distressed stocks that can be attributed to distress risk. The constant, or alpha, of this regression can be interpreted as the portion of returns not explained by the factors. We observe alphas that are increasing in default risk. This allows us to conclude that, even after correcting for the sources of risk captured by the factors, investors can expect a higher return on portfolios comprised of stocks at high risk of default. In other words, we find a positive distress risk premium in emerging market stocks.

Figure 5 graphically depicts the returns, alphas, and factor loadings from the 6-factor regression. Figure 6 plots the 6-factor alphas and their 95% confidence intervals, which show that the alpha on Portfolio 10 (highest probability of default) is significantly larger than the alphas on Portfolios 5 and lower, and the alpha on Portfolio 1 (lowest probability of default) is significantly smaller than all other portfolios' alphas.

We run a number of different exercises in order to test the statistical significance of these results. First, we form two "long-short" portfolios, LS90-10 and LS80-20 – the first long the most distressed portfolio (Portfolio 10) and short the portfolio with least distressed stocks (Portfolio 1), and the second long the two most distressed portfolios and short the two with least distressed stocks. We run a 6-factor regression using the 12-month average return of LS90-10 and LS80-20, and the results are shown in the first two columns of Table 8. The first row tells us that average returns are 1.3 percentage points larger in the most distressed portfolio than in the least distressed (Column 1), and 1.1 percentage points larger for the 20% riskiest stocks than the 20% safest (Column 2). Both differences are statistically significant. The coefficients on the Fama-French factors confirm that the exposure to SMB and HML is larger for the portfolios with more distressed stocks. Most importantly, the positive, statistically significant alphas reveal that the factor-adjusted compensation is in fact larger for the portfolios with higher distress risk, validating our finding of a positive distress risk premium.

An alternative method to compute 6-factor alphas using returns over a 12-month period is to run, for every month of the year following portfolio formation, a regression of that month's returns on the six factors computed in that same month. By doing this, the timing of the returns corresponds with the timing of the factors, instead of using the factors computed only on the month following portfolio formation as controls for the average 12-month returns. The two main results described above also hold when running the factor regressions in this manner. Panel B of Table 7 shows monotonically increasing alphas in distress risk, and the long-short regression coefficients in Columns 3 and 4 of Table 8 include positive, statistically significant alphas.

To further test of the robustness of our findings, we run firm-level Fama-Macbeth regressions of 12-month average returns on firms' probability of default. We find a positive, statistically significant coefficient in the second stage cross-sectional regression. Specifically, the average gamma across months if 0.0064, with a t-test p-value of 0.023, also confirming the presence of a distress risk premium in emerging market stocks.

Lastly, we use valuation ratios instead of realized returns to extract the distress risk premium. While realized returns are unbiased estimates of expected returns, their use as proxies for expected returns in the short and medium term has been questioned in the literature (e.g. Elton, 1999; Lundblad, 2007). Valuation measures like implied cost of capital, dividend yield, and earnings-to-price ratio are commonly suggested alternatives (e.g. Pastor et al., 2008), on the basis that they are a better reflection of investors' expectations of future stock performance. Panel C in Table 7 does not show a clear pattern in the relationship between the earnings-to-price ratio (using net-income as proxy for earnings) and our measure of probability of default.[19]

6 Conclusion

There is a dearth of rigorous research on the determinants of corporate distress in emerging markets. The goal of this paper is to shed light on factors that adversely impact the solvency of emerging market firms and explore whether investors are compensated for taking on distress risk. We believe that developing a framework that allows policymakers to anticipate corporate defaults in emerging markets may inform efforts to mitigate their regional and global impact.

We argue that, while existing models proposed for US firms yield reasonable forecasting power, they do not account for vulnerabilities specific to emerging market companies, such as advanced economy monetary policy changes, US dollar movements, or shifts in global liquidity and risk aversion. A novel multi-country dataset of corporate defaults allows us to develop a model of distress risk specific to emerging markets, as well as quantify the importance of global shocks on emerging market corporate distress.

We find that, controlling for firm-specific variables and country fixed effects, the 5-year US Treasury rate, the Fed funds rate, and the VIX are correlated with distress risk. Furthermore, introducing a dummy variable indicating whether a firm has defaulted in the past has a very positive impact on the model's predictive power. To the best of our knowledge this is a novel result. A model that includes accounting, market, and global macro variables along with country fixed effects and the prior-default dummy yields a much higher explanatory power for emerging market firms than Campbell et al.'s (2008) specification.

We also explore whether information about default risk is embedded in stock returns. We first do so by focusing on firms whose returns are most sensitive to global financial conditions. Analysis of these global betas reveals that the effect of the global variable on the probability of default is larger for firms with most negative betas. Furthermore, a composite global beta measure we call the Global Beta Z helps us show that the effect of a global risk-off environment on distress risk is greater for firms

[19] We do not have data to compute dividend yield or implied cost of capital.

whose returns respond more negatively to such global conditions.

Finally, we explore the asset pricing implications of our probability of default measure. Previous studies using reduced-form measures of default risk have struggled to identify a positive distress risk premium in US equities. We, on the other hand, find strong evidence of the presence of a distress risk premium in emerging market stocks. Future 12-month stock returns are monotonically increasing in the probability of corporate default, a trend that holds true after controlling for six popular factors. A number of robustness tests confirm the statistical significance of our findings.

References

[1] Acharya, Viral, Stephen Cecchetti, José de Gregorio, Sebnem Kalemli-Ozcan, Philip Lane, and Ugo Panizza (2015). Corporate Debt in Emerging Economies: A Threat to Financial Stability? *Brookings Institution Report*, September.

[2] Alfaro, Laura, Gonzalo Asis, Anusha Chari, and Ugo Panizza (2017). Lessons Unlearned? Corporate Debt in Emerging Markets. Working Paper (October).

[3] Altman, Edward (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23 (September):589-609.

[4] Altman, Edward (2005). An Emerging Market Credit Scoring System for Bonds. *Emerging Market Review*, 6:3011-323.

[5] Altman, Edward and Brooks Brady (2001). Explaining Aggregate Recovery Rates on Corporate Bond Defaults. Salomon Center Working Paper, November.

[6] Avdjiev, Stefan, Michael Chui, and Hyun Song Shin (2014). Non-financial Corporations from Emerging Market Economies and Capital Flows. *BIS Quarterly Review*, December:67-77.

[7] Bruno, Valentina and Hyun Song Shin (2015). Capital Flows and the Risk-Taking Channel of Monetary Policy. *Journal of Monetary Economics*, 71, 119-132.

[8] Bruno, Valentina and Hyun Song Shin (2016). Global Dollar Credit and Carry Trades: A Firm-Level Analysis. BIS Working Paper no. 510, August.

[9] Campbell, John, Jens Hilscher, and Jan Szilagyi (2008). In Search of Distress Risk. *Journal of Finance*, 63, 6, 2899-2939.

[10] Chan, K.C. and Nai-Fu Chen (1991). Structural and Return Characteristics of Small and Large Firms. *Journal of Finance*, 46, 4, 1467-1484.

[11] Chari, Anusha, Karlye Dilts Stedman, and Christian Lundblad (2017). Taper Tantrums: QE, its Aftermath, and Emerging Market Capital Flows. NBER Working Paper 23474.

[12] Chava, Sudheer and Robert Jarrow (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, 8, 537-569.

[13] Crosbie, Peter and Jeff Bohn (2003). Modeling Default Risk. Moody's KMV White Paper, (San Francisco: Moody's Investor Service) December 18.

[14] Das, Sanjiv, Darrell Duffie, Nikunj Kapadia, and Leandro Saita (2007). Common Failings: How Corporate Defaults Are Correlated. *Journal of Finance*, 62, 1, 93-117.

[15] Duan, Jin-Chuan, Jie Sun, and Tao Wang (2012) Multiperiod Corporate Default Prediction – A Forward Intensity Approach. *Journal of Econometrics*, 170, 1, 191-209.

[16] Duan, Jin-Chuan and Andras Fulop (2013) Multiperiod Corporate Default Prediction with the Partially-Conditioned Forward Intensity. Working Paper (August 21, 2013).

[17] Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita (2009) Frailty Correlated Default. *Journal of Finance*, 64, 5, 2089-2123.

[18] Elton, Edwin (1999). Expected Return, Realized Return, and Asset Pricing Tests. *Journal of Finance*, 54, 4, 1199-1220.

[19] Fama, Eugene and Kenneth French (1993). Common Risk Factors in the Returns of Stocks and Bonds. *Journal of Financial Economics*, 33, 1, 3-56.

[20] Fama, Eugene and Kenneth French (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, 51, 1, 55-84.

[21] Fratzscher, Marcel, Marco Lo Duca, and Roland Straub (2018). On the International Spillovers of US Quantitative Easing. The Economic Journal, 128, 608, 330-377.

[22] Hanley, James and Barbara McNeil (1982). The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 143, 29-36.

[23] Harvey, Campbell and Andrew Roper (1999). The Asian Bet. *The Crisis in Emerging Financial Markets*, ed. by Harwood, A., Litan, R.E., Pomerleano, M., Brookings Institution Press, Washington, DC, 29-115.

[24] Hernandez-Tinoco, Mario and Nick Wilson (2013). Financial Distress and Bankruptcy Prediction among Listed Companies using Accounting, Market and Macroeconomic Variables. *International Review of Financial Analysis*, 30, 394-419.

[25] Institute of International Finance (2017). Global Debt Monitor, June 2017. url: https://www.iif.com/publication/global-debt-monitor/global-debt-monitor-june-2017.

[26] International Monetary Fund (2015). Corporate Leverage in Emerging Markets-A Concern?. *Global Financial Stability Report*, October, 83-114.

[27] Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai (2012). Asset Fire Sales and Purchases and the International Transmission of Funding Shocks. *Journal of Finance*, 67, 6, 2015-2050.

[28] Kealhofer, Stephen (2003). Quantifying Credit Risk I: Default Prediction. *Financial Analysts Journal*, 59, 1, 30-44.

[29] Kealhofer, Stephen (2003). Quantifying Credit Risk II: Debt Valuation. *Financial Analysts Journal*, 59, 3, 78-92.

[30] Lundblad, Christian (2007). The Risk-Return Tradeoff in the Long Run: 1836-2003. *Journal of Financial Economics*, 85, 1, 123-150.

[31] Merton, Robert (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29, 2, 449-470.

[32] NUS-RMI Credit Research Initiative (2016). Technical Report, Version: 2016 Update 1. *Global Credit Review*, 6, 49-132.

[33] Ohlson, James (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 19, Spring, 109-131.

[34] Pastor, Lubos, Meenakshi Sinha, and Bhaskaran Swaminathan (2008). Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital. *Journal of Finance*, 63, 6, 2859-2897.

[35] Powell, Andrew (2014). Global Recovery and Monetary Normalization: Escaping a Chronicle Foretold?. 2014 Latin American and Caribbean Macroeconomic Report, Inter-American Development Bank.

[36] Rey, Helene (2015). Dilemma Not Trilemma: The Global Financial Cycle and Monetary Policy Independence. NBER Working Paper No. 21162.

[37] Shin, Hyun Song (2013). The Second Phase of Global Liquidity and Its Impact on Emerging Economies. *Proceedings of the Asia Economic Policy Conference*, Federal Reserve Bank of San Francisco.

[38] Shumway, Tyler (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74, 1, 101-124.

[39] Tian, Shaonan, Yan Yu, and Hui Guo (2015). Variable Selection and Corporate Bankruptcy Forecasts. *Journal of Banking & Finance*, 52, 89-100.

[40] Vasicek, Oldrich (1984). Credit Valuation. Unpublished paper, KMV Corporation.

[41] Vassalou, Maria and Yuhang Xing (2004). Default Risk in Equity Returns. *Journal of Finance*, 59, 2, 831-868.

[42] Zmijewski, Mark (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research Supplement*, 59-86.

Figure 1: Example of Receiver Operating Characteristics Curve

Point A in the "High-Sensitivity Model" ROC curve shows that the 20% of firms with highest probability of default include 70% of the firms that default the following month. Point B in the "Low-Sensitivity Model" curve indicates that to capture 70% of firms that default next month one needs to include the top 50% firms with highest probability of default.

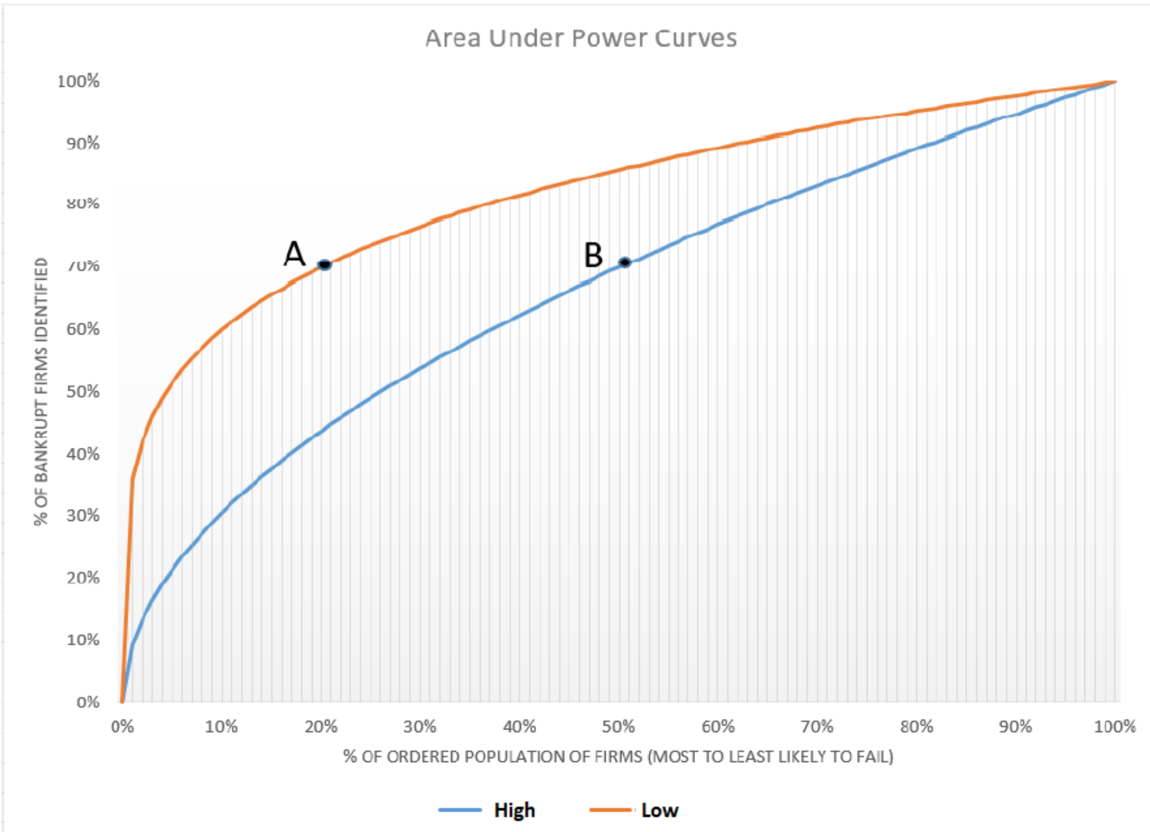

Figure 2: In-sample Predictive Power

This figure shows the Receiver Operating Characteristics (ROC) curve for our best model of distress risk and for the specification in Campbell et al. (2008) - our benchmark. The curves shown are the average of the ROC curves in each month in the sample.

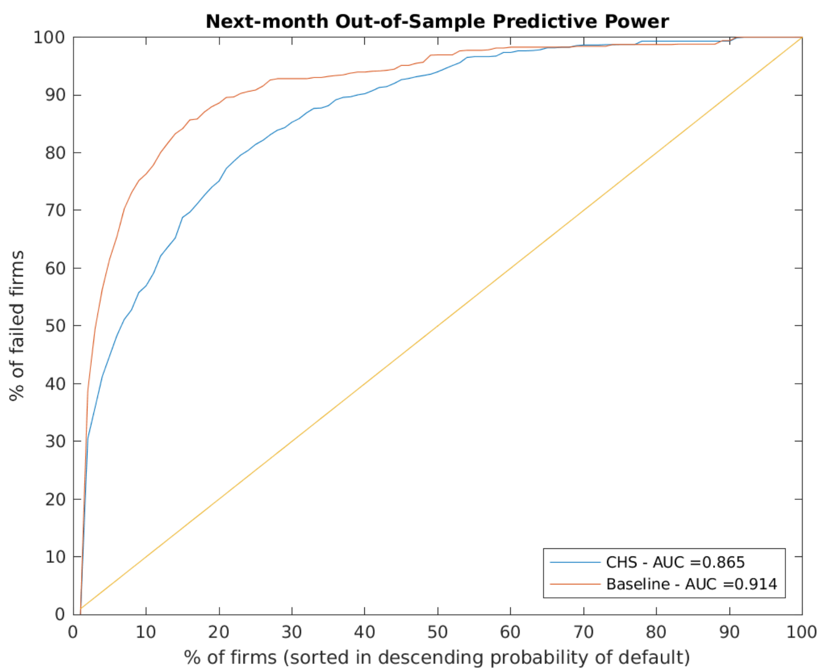

Figure 3: Time Series of Actual and Predicted Defaults

This figure shows the number of actual defaults per month (averaged by quarter) and number of defaults predicted by our model. The number of predicted defaults in a month is the sum of the estimated probabilities of default of all firms.

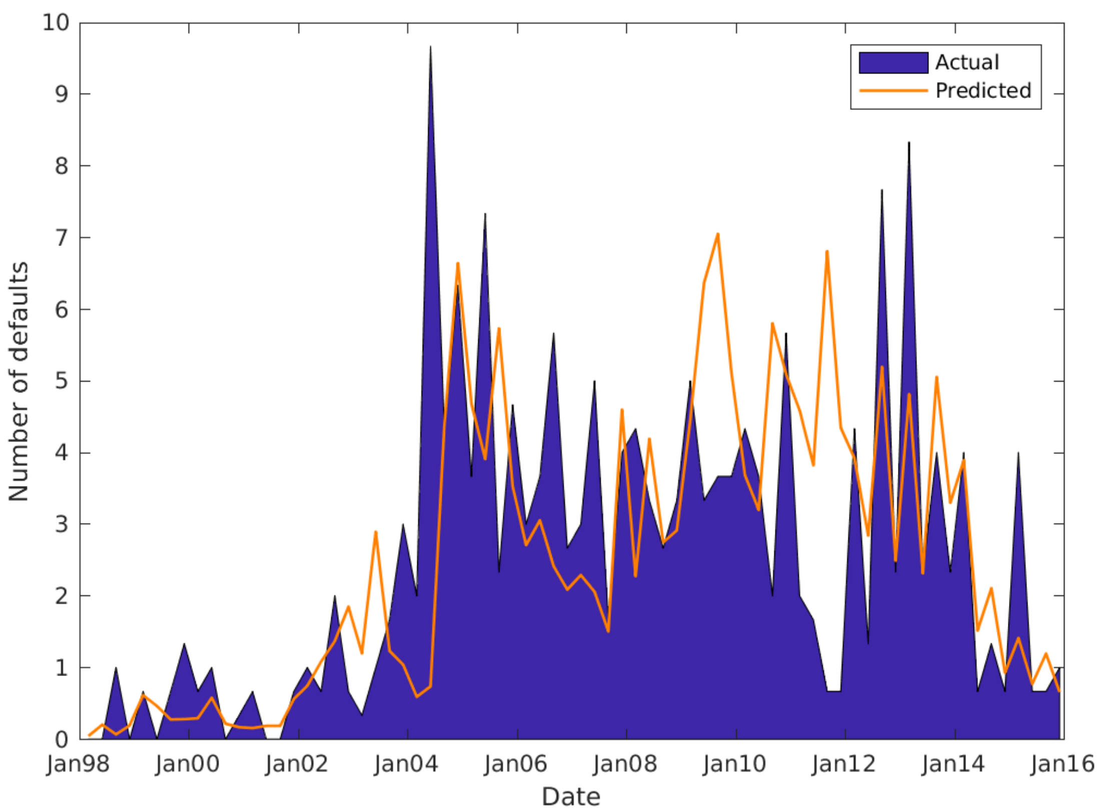

Figure 4: Predicted Probabilities

This figure shows predicted probabilities (in blue) for all values of each variable, keeping all other predictors constant at their mean. The red lines are 95% confidence intervals, computed using the 95% confidence intervals of each variable's coefficient in the logit regression.

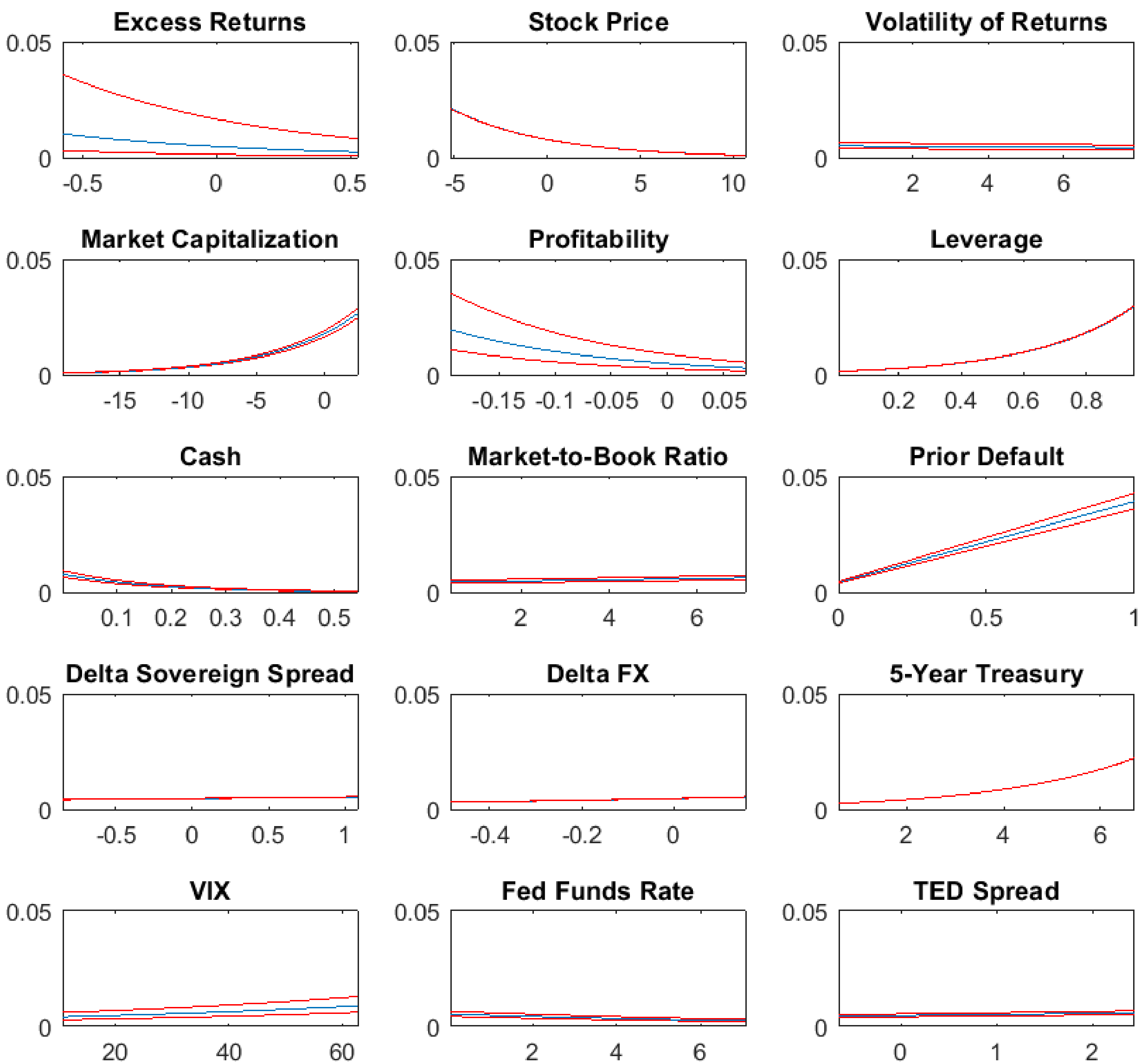

Figure 5: Portfolio Returns, Alphas, and Factor Loadings

For each portfolio ordered from least to most distressed, the first figure (Figure 5a) shows 12-month average future returns and alphas from the 6-factor regression on market, size, value, momentum, short-term reversal, and long-term reversal factors. The second figure (Figure 5b) shows the coefficients on the Fama-French factors from the 6-factor regression.

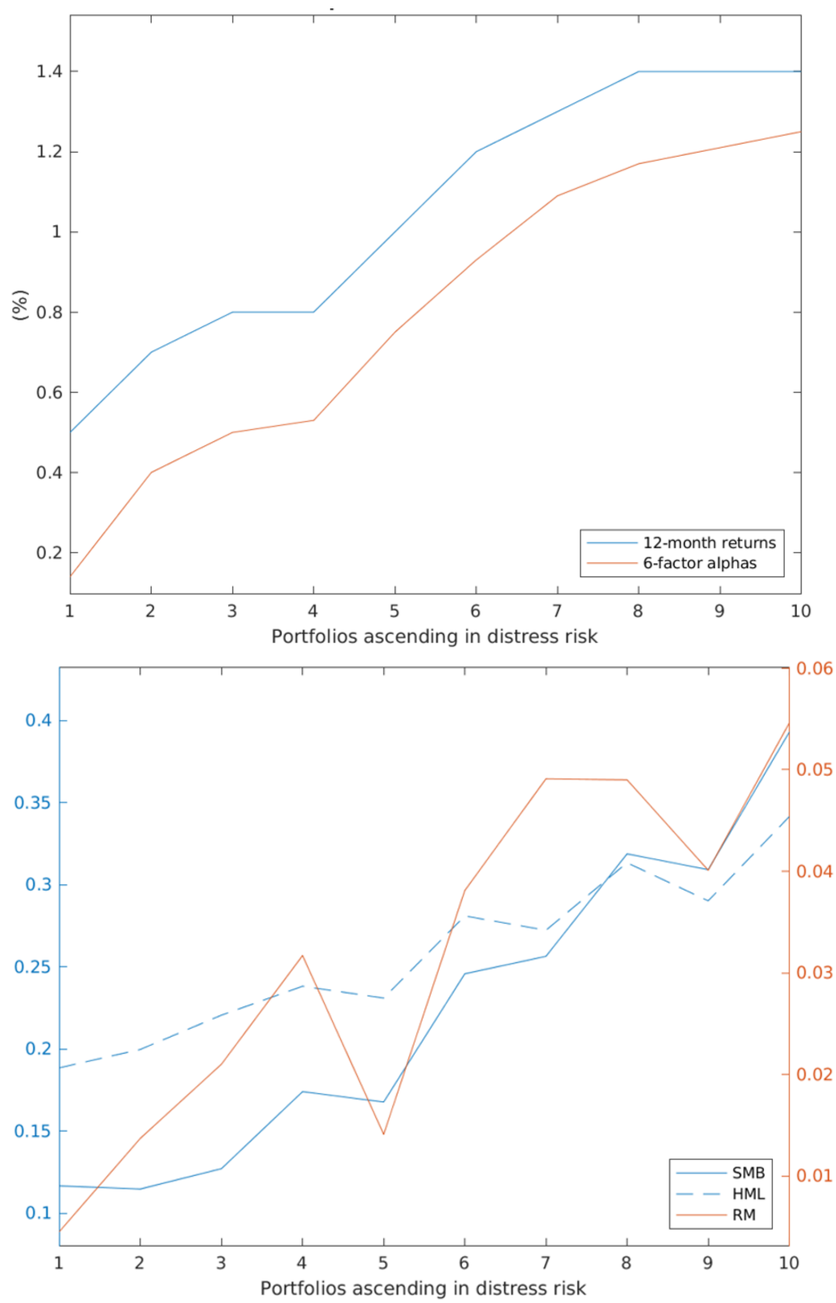

Figure 6: Six-Factor Alphas

This figure plots six-factor alphas and their respective 95% confidence intervals for 10 portfolios sorted by distress risk.

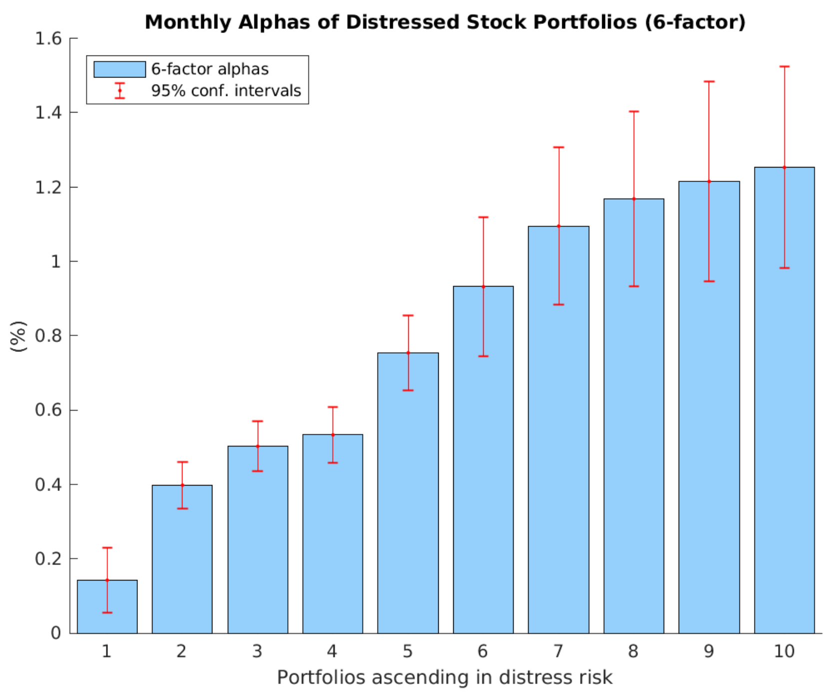

Table 1: Types of Default Events

Panel A presents the types of default events covered in the CRI database and their classification into Bankruptcy, Delisting, and Corporate Default Action categories, as CRI does in its database's technical report (NUS-RMI Technical Report 2016, Table A.9, p. 106). Panel B counts the number of each type of event in our final sample; i.e. the sample of firm-months with data on all accounting and market variables.

PANEL A

Action Type	Subcategory
Bankruptcy Filing	Administration, Arrangement, Canadian CCAA, Chapter 7, Chapter 11, Chapter 15, Conservatorship, Insolvency, Japanese CRL, Judicial Management, Liquidation, Pre-Negotation Chapter 11, Protection, Receivership, Rehabilitation, Rehabilitation (Thailand 1997), Reorganization, Restructuring, Section 304, Supreme court declaration, Winding up, Work out, Sued by creditor, Petition Withdrawn, Other
Delisting	Bankruptcy
Default Corporate Action	Bankruptcy, Coupon & Principal Payment, Coupon Payment Only, Debt Restructuring, Interest Payment, Loan Payment, Principal Payment, ADR (Japan only), Declared Sick (India only), Regulatory Action (Taiwan only), Financial Difficulty and Shutdown (Taiwan only), Buyback option, Other

PANEL B

Action Type	Count
Bankruptcy	74
Delisting	3
Default Corporate Action	509
Bankruptcy Corporate Action	11
Coupon & Principal Payment	19
Coupon Payment	19
Restructuring	133
Interest Payment	10
Loan Payment	320
Principal Payment	10
Other	2
Unknown	12

Table 2: Summary Statistics

Summary statistics for all firm-months, for the group of firm-months that experience any default event, and for the group that experiences a bankruptcy next month. The last two columns show the results of a two-sample t-test for equal means, where the "Default" and "Bankrupt" columns refer to the tests of whether the mean for the full sample is different from the default group or the bankrupt group, respectively. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.10$.

	Means			t-Tests	
	(1) Full Sample	(2) Default	(3) Bankrupt	(4) Default	(5) Bankrupt
Excess returns	-0.008	-0.050	-0.06	***	***
Stock price	2.636	1.305	0.156	***	***
Volatility of returns	1.580	0.714	0.727	**	
Market capitalization	-7.735	-9.130	-9.233	***	***
Profitability	0.004	-0.017	-0.049	***	***
Leverage	0.366	0.578	0.759	***	***
Cash	0.082	0.045	0.024	***	***
Market-to-book ratio	2.121	2.673	4.400	***	***
Prior default	0.058	0.609	0.429	***	***
Unemployment rate	4.683	4.306	5.471	***	**
Inflation	0.036	0.035	0.031		
Real interest rate	4.123	1.929	8.113	***	***
Sovereign spread	2.541	2.503	1.965		*
ΔSovereign spread	0.009	0.022	-0.003	**	
ΔFX	0	0	-0.003		
5-year Treasury	2.363	2.890	2.785	***	**
VIX	19.52	19.59	21.31		*
Fed funds rate	1.235	1.869	1.547	***	
TED spread	-0.067	0.035	-0.111	***	

Table 3: Logit Regressions of Probability of Default Next Month

Results of logit regression combining accounting and market variables with local and global macro variables to explain the probability of default next month. Column 1 replicates Campbell et al.'s (2008) specification, which uses only firm-specific accounting and market variables. Column 2 adds a dummy indicating whether a firm has experienced a default event in the past. Column 3 adds domestic macro variables, Column 4 includes global variables, and Column 5 has both domestic and global. Column 6 is our baseline specification, which incorporates country fixed effects to the model in Column 4. The dependent variable in all specifications is binary, indicating whether a firm experienced a distress event the following month. Pseudo-R^2 refers to McFadden's Pseudo-R^2, and AUC is the area under the ROC curve. ***, **, and * indicate three levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-7.470***	-8.108***	-8.297***	-9.360***	-9.478***	-9.973***
Excess returns	-1.354***	-1.493***	-1.309***	-1.195***	-1.222***	-1.344***
Stock price	-0.224***	-0.168***	-0.079***	-0.080***	-0.054**	-0.193***
Volatility of returns	-0.074**	-0.071**	-0.058*	-0.077**	-0.065*	-0.022
Market capitalization	0.027	-0.002	-0.045*	-0.063***	-0.082***	0.164***
Profitability	-6.543***	-5.997***	-6.416***	-6.330***	-6.570***	-7.141***
Leverage	2.583***	2.102***	2.197***	1.904***	2.041***	3.136***
Cash	-4.837***	-3.235***	-4.195***	-3.451***	-3.717***	-5.993***
Market-to-book ratio	0.206***	0.119***	0.089***	0.082***	0.087***	0.051***
Prior default		2.502***	2.491***	2.560***	2.515***	2.234**
Unemployment rate			0.054**		0.032	
Inflation			-3.488		-2.587	
Real interest rate			-0.048***		-0.038***	
Sovereign spread			-0.037**		-0.018	
ΔSovereign spread				0.193	0.093	0.127
ΔFX				1.194	-1.129	0.878
5-year Treasury				0.351***	0.320***	0.349***
VIX				0.007	0.009*	0.015***
Fed funds rate				-0.119**	-0.110*	-0.109*
TED spread				0.270**	0.218*	0.088
Pseudo-R^2	0.124	0.200	0.235	0.232	0.241	0.221
AUC	0.865	0.907	0.888	0.899	0.893	0.914
Observations	589,224	589,224	372,673	402,253	372,158	402,253
Defaults	589	589	524	544	522	544
Country FE						✓

Table 4: Top Tercile Betas by Global Variable

Results of logit regression of probability of default on each global factor, controlling for firm-specific variables. The explanatory variables in Panel A are the global variable of the same name as each column and the interaction of that variable with a dummy indicating whether a firm's returns are among the top-third most sensitive to the global factor. Panel B also includes accounting and market variables as controls. The dependent variable is binary, indicating whether a firm experienced a distress event the following month.Pseudo-R^2 refers to McFadden's Pseudo-R^2. ***, **, and * indicate three levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

PANEL A

	(1) **ΔSov. spread**	(2) **ΔFX**	(3) **5-year Treasury**	(4) **VIX**	(5) **Fed funds**	(6) **TED spread**
Constant	-6.662***	-7.020***	-7.708***	-7.136***	-7.229***	-7.015***
Global variable	0.469	-0.378	0.245***	-0.002	0.129***	0.244**
Global variable * Top-tercile	0.376	1.063	0.082***	0.020***	0.060**	0.210
Pseudo-R^2	0.003	0	0.01	0.003	0.006	0.001
Observations	479,438	774,705	774,705	774,705	774,705	774,705
Defaults	617	692	692	692	692	692

PANEL B

	(1) **ΔSov. spread**	(2) **ΔFX**	(3) **5-year Treasury**	(4) **VIX**	(5) **Fed funds**	(6) **TED spread**
Constant	-7.417	-7.447***	-8.363***	-7.529***	-7.824***	-7.503***
Excess returns	-1.094***	-1.317***	-1.295***	-1.334***	-1.304***	-1.264***
Stock price	-0.160***	-0.224***	-0.198***	-0.222***	-0.207***	-0.219***
Volatility of returns	-0.100***	-0.083**	-0.088***	-0.078**	-0.088**	-0.084***
Market capitalization	0.009	0.027	-0.016	0.024	0.003	0.021
Profitability	-7.019***	-6.539***	-6.631***	-6.489***	-6.661***	-6.660***
Leverage	2.416***	2.559***	2.473***	2.651***	2.501***	2.560***
Cash	-5.725***	-4.966***	-4.751***	-5.191***	-4.796***	-4.941***
Market-to-book ratio	0.184***	0.204***	0.204***	0.200***	0.205***	0.205***
Global variable	0.612*	1.434	0.182***	-0.009*	0.092***	0.381***
Global variable * Top-tercile	-0.201	0.890	0.075***	0.026***	0.059*	-0.013
Pseudo-R^2	0.102	0.123	0.129	0.128	0.126	0.125
Observations	398,601	586,985	586,985	586,985	586,985	586,985
Defaults	536	586	586	586	586	586

Table 5: Marginal Effects

This table reports marginal effects of each individual regressor in the logit model, sorted from smallest to largest. The AME column shows average marginal effects. The MEM column presents marginal effects at the mean; i.e. the marginal effect of each regressor when all other regressors are kept at their mean.

	(1) MEM	(2) AME
Cash	-0.111	-0.403
Profitability	-0.067	-0.242
Excess returns	-0.056	-0.204
Stock price	-0.009	-0.032
Fed funds rate	-0.003	-0.009
Volatility of returns	-0.001	-0.003
VIX	0.001	0.002
ΔSovereign spread	0.001	0.002
TED spread	0.002	0.006
Market-to-book ratio	0.003	0.013
Market capitalization	0.005	0.019
5-year Treasury	0.010	0.038
ΔFX	0.037	0.132
Prior default	0.084	0.304
Leverage	0.125	0.450

Table 6: Composite Global Beta Z Score as Predictor of Default

Results of logit regression of probability of default on a composite global factor. Beta Z and Variable Z are the sum of the standardized global betas and global variables, respectively. We control for accounting, market, and domestic macro variables. The dependent variable is binary, indicating whether a firm experienced a distress event the following month.Pseudo-R^2 refers to McFadden's Pseudo-R^2. ***, **, *, and † indicate three levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, $p < 0.10$, and $p < 0.15$, respectively.

	(1)	(2)	(3)
Constant	-8.212***	-8.507***	-8.386***
Excess returns	-1.048***	-1.120***	-1.187***
Stock price	-0.094***	-0.076***	-0.063**
Volatility of returns	-0.109***	-0.090**	-0.066*
Market capitalization	-0.072***	-0.066***	-0.068***
Profitability	-7.370***	-6.280***	-6.515***
Leverage	2.350***	1.865***	2.113***
Cash	-5.025***	-3.385***	-3.947***
Market-to-book ratio	0.182***	0.074***	0.082***
Prior default		2.700***	2.567***
Unemployment rate			0.034†
Inflation			-4.259*
Real interest rate			-0.041***
Sovereign spread			-0.027
Beta Z	0.055***	0.016	-0.013
Variable Z	0.096***	0.086***	0.073***
Beta Z * Variable Z	0.001	-0.004	-0.011†
Pseudo-R^2	0.110	0.201	0.202
AUC	0.857	0.906	0.893
Observations	461,977	461,977	386,884
Defaults	536	536	515

Table 7: Returns on Portfolios Sorted by Distress Risk

Stocks sorted monthly based on our predicted probability of default and placed in ten portfolios of equal size. Portfolio 1 contains the firms with the lowest probability of default and Portfolio 10 those with highest predicted distress risk. We rebalance the portfolios every month from January 2002 to December 2015 based on the stocks' updated distress risk. Panel A shows, for each portfolio, average estimated probability of default, average monthly returns for the 12 months following portfolio formation, and coefficients from a 6-factor regression: RM equals the return of a weighted average of country index returns minus the risk-free rate, and SMB and HML are the returns of factor-mimicking portfolios constructed as in Fama and French (1993). The other factors (coefficients not shown) are momentum, short-term reversal, and long-term reversal. Panel B shows the average coefficients of 12 monthly regressions of future returns on the factors computed for each respective month. In Panel C, we run the same regression as in Panel A, but using as dependent variable the net-income-to-price ratio as a proxy for expected returns.

PANEL A: Single Regression

Portfolios	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean P(default) (%)	0.005	0.010	0.016	0.025	0.038	0.058	0.079	0.110	0.171	1.169
Mean 12-month returns	0.005	0.007	0.008	0.008	0.010	0.012	0.013	0.014	0.014	0.014
6-factor alpha	0.0014	0.0040	0.0050	0.0053	0.0075	0.0093	0.0109	0.0117	0.0121	0.0125
RM	0.0045	0.0137	0.0210	0.0317	0.0141	0.0381	0.0491	0.0490	0.0401	0.0546
SMB	0.1166	0.1147	0.1271	0.174	0.1677	0.2458	0.2564	0.3188	0.3092	0.3930
HML	0.1885	0.1997	0.2206	0.2382	0.231	0.281	0.2723	0.3134	0.2902	0.3415

PANEL B: Average Monthly Regressions

Portfolios	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
6-factor alpha	0.003	0.005	0.007	0.007	0.010	0.012	0.014	0.015	0.015	0.017
RM	-0.0487	-0.0357	-0.0344	-0.0281	-0.0511	-0.0393	-0.0306	-0.0268	-0.0366	-0.0491
SMB	0.0609	0.064	0.0733	0.0985	0.0553	0.0807	0.0837	0.132	0.1335	0.1902
HML	0.0837	0.1238	0.1344	0.1472	0.0943	0.1254	0.1304	0.1604	0.165	0.1836

PANEL C: Earnings/Price Ratios

Portfolios	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean E/P ratio	0.029	0.018	0.014	0.013	-0.012	0	0.055	0.066	-0.013	-0.216
6-factor alpha	0.0281	0.0191	0.0151	0.0144	-0.0112	-0.0031	0.1723	0.2337	-0.0189	-0.1908
RM	0.041	-0.0034	-0.0297	0.0002	0.1795	-0.1216	-2.7919	-3.5138	0.1448	0.2616
SMB	-0.035	0.0395	0.018	0.0269	-0.2537	0.0886	-1.1453	-11.9838	-0.2831	-0.1556
HML	-0.0265	0.009	-0.0337	-0.0816	-0.6162	0.3208	-0.2434	0.5291	0.0081	-1.2924

Table 8: Returns on Long-Short Portfolios

For long-short portfolios LS90-10 and LS80-20 (long the riskiest and short the safest one and two deciles, respectively), this table shows average 12-month future returns and alphas and Fama-French factors from a six-factor regression. The first two columns present the results of a regression of average future 12-month returns on first-month factors, while the last two show the average coefficients of 12 monthly regressions of future returns on the factors computed for each respective month. ***, **, and * indicate three levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. ‡ and † indicate p< 0.01 and p< 0.1 in a t-test of means.

	Single		**Average**	
	(1)	**(2)**	**(3)**	**(4)**
	LS90-10	**LS80-20**	**LS90-10**	**LS80-20**
Mean 12-month returns	0.013‡	0.011‡	0.016†	0.0127
6-factor alpha	0.011***	0.010***	0.0139**	0.0115**
RM	0.050	0.038	-0.001	-0.002
SMB	0.276***	0.235***	0.132	0.102
HML	0.153**	0.122*	0.104	0.074

Appendix A: Variable and Factor Definitions

Variable Name	**Variable Definition**
Excess returns	Log (1 + firm returns) - log (1 + country (market) index returns).
Stock price	Log price per share.
Volatility of returns	Standard deviation of daily returns over the previous month.
Market capitalization	Log (Firm market cap) - log (country market cap). The market capitalization of listed domestic companies comes from the World Bank.
Profitability	Ratio of net income to the market value of total assets, where the market value of assets is equal to the sum of the firm's market capitalization and total liabilities.
Leverage	Ratio of total liabilities to the market value of total assets.
Cash	Ratio of cash and cash equivalents to the market value of total assets.
Market-to-book ratio	Ratio of market capitalization to book value of equity, where book value of equity is total assets minus total liabilities. Following Campbell et al. (2008), if a firm has a negative book value of equity, we set its book value of equity equal to $1 in order to place that firm's market-to-book ratio in the right-hand side of the distribution (Large positive MB instead of a negative MB).
ΔFX	Monthly percentage change in the exchange rate between the local currency and the US dollar, quoted as local currency units per dollar and retrieved from Bloomberg.
5-year Treasury rate	Interest rate on US 5-year Treasury notes.
VIX	CBOE Volatility Index.
Fed funds rate	Federal Funds Rate, retrieved from FRED, Federal Reserve Bank of St. Louis.
TED spread	Component of the TED spread orthogonal to VIX. The TED spread is the spread between 3-month LIBOR RATES and 3-month T-bill rates, often used as a measure of liquidity risk in bond markets. Due to collinearity between VIX and the TED spread, we regress the TED spread on the VIX and keep the residual.

Sources: Default data and all accounting and market variables come from the CRI database, the Credit Research Initiative of the National University of Singapore, accessed on December 1, 2016.

Appendix B: Additional Tables and Figures

Table B1: Number of Defaults and Observations per Year

This table lists the number of defaults and observations per year of our sample, aggregated across countries, for the observations with all accounting and market data.

Year	Firm-Months	Defaults	%
1995	16	0	0
1996	370	0	0
1997	692	0	0
1998	1,725	5	0.29
1999	2,561	8	0.31
2000	5,808	6	0.10
2001	7,688	5	0.07
2002	15,406	15	0.10
2003	23,829	19	0.08
2004	30,882	67	0.22
2005	33,640	55	0.16
2006	35,724	46	0.13
2007	40,844	42	0.10
2008	43,601	47	0.11
2009	45,812	54	0.12
2010	50,627	52	0.10
2011	60,405	18	0.03
2012	58,599	47	0.08
2013	72,280	58	0.08
2014	72,730	25	0.03
2015	68,523	21	0.03
Total	671,762	590	0.09

Table B2: Number of Observations per Country and Year

This table lists the number of firm-months with all accounting and market data per country and year of our sample.

	Argentina	Brazil	Chile	China	Colombia	Czech Republic	Hungary	India	Indonesia	Malaysia	Mexico	Pakistan	Peru	Philippines	Poland	South Africa	South Korea	Thailand	Turkey	Vietnam	Total
1995	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	16	.	.	16
1996	.	.	.	.	.	.	.	.	.	.	123	.	.	.	.	.	.	247	.	.	370
1997	27	.	8	.	.	.	.	.	.	.	225	.	.	.	.	.	.	417	15	.	692
1998	42	.	324	.	.	.	.	.	.	.	255	.	.	.	.	.	.	1072	32	.	1725
1999	61	.	459	.	.	.	.	.	.	173	265	.	.	129	.	12	.	1384	78	.	2561
2000	165	373	492	.	.	.	.	.	33	1817	431	.	.	225	57	36	6	1363	810	.	5808
2001	182	457	472	.	.	.	.	.	308	2270	468	.	.	218	301	45	23	1642	1302	.	7688
2002	124	493	442	.	.	15	15	.	454	2469	448	.	7	223	402	57	7077	1806	1374	.	15406
2003	97	553	441	4933	.	43	30	27	513	2447	455	.	8	235	438	74	10204	2001	1330	.	23829
2004	166	609	448	7720	.	94	94	12	732	4761	507	.	21	217	913	67	11057	2122	1342	.	30882
2005	304	709	621	7846	104	130	136	15	869	5556	483	12	159	244	1316	57	11273	2130	1676	.	33640
2006	359	803	921	7539	119	109	131	18	948	6099	525	195	217	265	1469	54	11620	2296	2037	.	35724
2007	422	1188	992	9682	115	76	116	42	1206	6492	559	675	307	320	1732	58	12341	2470	2051	.	40844
2008	401	1350	1084	10647	131	76	129	78	1285	5096	521	478	270	690	2194	71	12956	2474	2048	1622	43601
2009	380	1361	1117	11484	120	.	168	105	1329	4872	577	1152	240	835	2386	63	13147	2464	2026	1986	45812
2010	376	1444	1065	12991	170	.	175	455	1602	5717	590	1563	270	912	2571	93	12892	2785	2071	2885	50627
2011	396	1520	1187	16071	176	.	130	4305	1919	5633	567	1316	287	1004	2787	93	13863	2786	2176	4189	60405
2012	311	1441	1200	17919	159	.	136	9626	2079	5536	595	.	245	1164	2788	87	10013	2957	2343	.	58599
2013	320	1600	1227	18298	127	.	162	12246	2256	5508	610	.	197	1124	2814	78	15942	3162	2418	4191	72280
2014	357	1656	1158	17449	164	.	169	11647	2315	5706	575	.	181	1265	2807	106	16418	3366	2400	4991	72730
2015	371	1545	1108	16958	159	.	181	6791	2341	5472	584	.	188	1364	2854	71	17607	3519	2432	4978	68523
Total	4861	17102	14766	159537	1544	543	1772	45367	20189	75624	9363	5391	2597	10434	27829	1122	176439	42479	29961	24842	671762

Table B3: Correlation Matrix and Multicollinearity Analysis

The last two rows of this table show the Tolerance Value (TOL) and its reciprocal Variance Inflation Factor (VIF). VIF is computed as $1/(1-R_k^2)$, where R_k^2 is the R^2 value of a regression of factor k on all others. TOL is simply $1-R_k^2$. VIF values larger than 10 are typically considered suggestive of multicollinearity in a model. The rest of the matrix presents pairwise correlations between all variables in our various specifications.

	EXRET	**PRICE**	**VOL**	**RELSIZE**	**NIMTA**	**TLMTA**	**CASHMTA**	**MB**	**Unemployment**	**Inflation**	**Real rate**	**sovSpread**	**ΔSovSpread**	**ΔFX**	**5YEAR**	**VIX**	**Fed Funds**	**TED**
EXRET	1																	
PRICE	0.024	1																
VOL	0.002	0.057	1															
RELSIZE	0.022	0.589	0.036	1														
NIMTA	0.049	0.098	0.007	0.109	1													
TLMTA	-0.067	-0.139	-0.024	-0.155	-0.142	1												
CASHMTA	-0.023	-0.188	-0.033	-0.120	0.093	-0.061	1											
MB	0.065	0.165	0.052	0.124	-0.091	-0.466	-0.272	1										
Unemployment	0.010	0.211	0.075	0.351	-0.006	0.003	-0.132	0.084	1									
Inflation	-0.025	0.234	0.039	0.089	0.007	0.137	-0.134	0.004	0.219	1								
Real rate	0.010	0.037	0.065	0.074	-0.011	0.110	-0.110	-0.005	0.334	0.113	1							
SovSpread	-0.016	0.143	0.033	-0.087	-0.008	0.164	-0.142	-0.035	0.193	0.532	0.151	1						
ΔSovSpread	-0.016	-0.027	-0.006	-0.021	0	-0.017	0.010	0.027	-0.014	0.109	-0.078	0.02	1					
ΔFX	-0.007	0.017	0.001	-0.026	-0.003	-0.025	-0.009	0.017	-0.016	-0.057	-0.054	-0.067	-0.181	1				
5YEAR	-0.017	-0.031	-0.001	0.181	0.015	0.099	-0.057	-0.105	0.110	-0.089	-0.021	-0.218	-0.009	0.052	1			
VIX	0.019	0.035	-0.010	0.090	0.010	0.031	0.020	-0.036	0.019	0.022	0.006	0.042	0.212	-0.080	-0.056	1		
Fed Funds	-0.009	-0.084	0.003	0.118	0.018	0.078	-0.044	-0.077	0.096	-0.010	-0.032	-0.179	0.011	0.064	0.888	-0.122	1	
TED spread	-0.016	-0.141	0	-0.004	0.022	-0.018	-0.010	0.029	0.046	0.179	-0.077	-0.043	0.140	-0.016	0.336	0.062	0.477	1
TOL	0.989	0.552	0.988	0.521	0.929	0.670	0.828	0.658	0.727	0.609	0.848	0.621	0.893	0.949	0.184	0.887	0.165	0.653
VIF	1.011	1.810	1.012	1.918	1.077	1.493	1.207	1.519	1.376	1.643	1.179	1.612	1.12	1.054	5.425	1.127	6.064	1.531

Appendix C: LASSO Estimation

Table C1: Robustness Checks Using LASSO for Variable Selection

Column 1 presents coefficients returned by a simple logit estimation of firms' probability of default on our full set of explanatory variables, while Column 2 shows the coefficients returned by the LASSO procedure. Running a logit regression only on the variables with nonzero coefficients in Column 2 yields the coefficients and statistics in Column 3. In all cases the dependent variable is binary, indicating whether a firm experienced a distress event the following month. Pseudo-R^2 refers to McFadden's Pseudo-R^2, and AUC is the area under the ROC curve. ***, **, and * indicate three levels of statistical significance of the coefficients: $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively.

	(1)	(2)	(3)
Constant			-9.213***
Excess returns	-1.222	-0.159	-1.592***
Stock price	-0.054	-0.040	-0.166***
Volatility of returns	-0.065		
Market capitalization	-0.082		
Profitability	-6.57	-7.372	-6.579***
Leverage	2.041	1.438	2.333***
Cash	-3.717		
Market-to-book ratio	0.087	0.029	0.142***
Prior default	2.515	2.630	2.563***
Unemployment rate	0.032		
Inflation	-2.587		
Real interest rate	-0.038		
Sovereign spread	-0.018		
ΔSovereign spread	0.093		
ΔFX	-1.129		
5-year Treasury	0.32	0.058	0.211***
VIX	0.009		
Fed funds rate	-0.11		
TED spread	0.218		
Pseudo-R^2	0.241		0.218
AUC	0.893		0.911
Observations	372,158		744,197
Defaults	522		617

Figure C1: LASSO Coefficient Path

Coefficient path using LASSO for variable selection on our full set of explanatory variables. The vertical axis reports the value of the coefficients of the standardized explanatory variables. The lower horizontal axis shows the level of λ decreasing from left to right, where a lower λ implies a loosening of the constraint on the sum of the absolute value of the coefficients. The numbers at the top of the figure indicate the degrees of freedom or number of variables with coefficient different from zero for each level of λ.