# Are you a Zombie? A Supervised Learning Method to Classify Unviable Firms and Identify the Determinants

#### Angela De Martiis<sup>1</sup>

Institute for Financial Management University of Bern

#### Big Data & Machine Learning in Finance Conference

Politecnico di Milano

June 10, 2021

<sup>&</sup>lt;sup>1</sup>Joint with Thomas Heil and Franziska Peter (Zeppelin Universität)

### **Motivation**

UNIVERSITÄT BERN

 $u^{\scriptscriptstyle b}$ 

The Corporate Zombies Stalking Europe *Financial Times*, September 2020



Rising Number of Zombie Companies Threaten Financial Stability Forbes, March 2020





#### Definition

Firms whose enterprise-contributed capital has been lost (Kane 1989)

Insolvent companies with little hope of recovery but avoid failure thanks to support from their banks (Hoshi 2006)

Firms that are unable to cover debt servicing costs from current profits over an extended period (Banerjee and Hofmann 2018)

Cementir CEO "... we have zombie kept alive by banks"2

<sup>&</sup>lt;sup>2</sup> Il Mattino, April 2017

#### UNIVERSITÄT BERN

#### Literature Review

- Zombie Firms:
  - Phenomenon, characteristics (Caballero et al., 2008; Hoshi 2006)
  - Rise of zombie, identification strategy (McGowan et al., 2018; Banerjee and Hofmann 2018; Banerjee and Hofmann 2020; Schivardi et al., 2020)
- Machine Learning:
  - o Alternative zombie definition using Italian firms (Bargagli Stoffi et al., 2020)
- Corporate Finance:
  - Concept, measure and characteristics of distressed firms (Altman 1968; Asquith et al., 1994; Campbell et al., 2008; Kahle and Stulz 2017)

#### Contribution

- Analyze the prevalence of zombie in Europe and the USA
- Examine the firm-specific characteristics of zombie companies based on two large datasets of company financial and market data
- Supervised learning methods, Decision Trees and Support Vector Machines (SVM), used for feature selection
- Contrast firm-specific characteristics of zombie with those of non-zombie, distressed, healthy and recovered zombie
- Ultimately:
  - Better classify a firm as zombie
  - o (Policy-relevant) Tool to make prediction



UNIVERSITÄT BERN

#### Results

- Document differences between zombie firms in the USA and in Europe on a number of firm-specific and industry-specific factors
- Income and leverage variables are decisive in classifying zombie, while shareholders equity is predominant among US distressed
- Zombie and distressed firms show differences and similarities

#### Identification

- We use firm-level data from Compustat Global and Compustat North America Fundamentals Annual from 1990 to 2018
- Following Banerjee and Hofmann (2020), we define a zombie firm by
  - *ICR<sub>it</sub>* < 1 for 3 consecutive years
  - Tobin's q below median within sector and year

#### Prevalence of Zombie Firms

UNIVERSITÄT BERN

 $\boldsymbol{u}^{\scriptscriptstyle b}$ 



#### Trend of Zombie Firms

UNIVERSITÄT BERN

 $u^{\scriptscriptstyle b}$ 



### Data Preprocessing

- Use data on 32 European countries and the USA
- Construct two cross-sections: 2007 and 2016
- Add performance measures capturing leverage, asset tangibility, profitability, liquidity, market value, growth opportunities, and stock market variables from Thomson Reuters Datastream
- Impute missing values based on K-Nearest-Neighbors imputations
- 70 variables per company/year

#### **Descriptive Statistics**

UNIVERSITÄT BERN

 $\boldsymbol{u}^{\scriptscriptstyle b}$ 

	Healthy		Distressed		Zombie		Recovered	
	EU	US	EU	US	EU	US	EU	US
Leverage	0.469	0.471	0.633	0.696	0.631	0.742	0.537	0.535
Net Leverage	0.048	0.101	0.267	0.297	0.222	0.286	0.145	0.162
Asset Tangibility	0.281	0.243	0.375	0.270	0.290	0.224	0.310	0.261
Cash ST Investments	0.115	0.076	0.063	0.064	0.064	0.077	0.083	0.075
Operating Profit	0.104	0.136	0.049	-0.027	-0.015	-0.049	0.047	0.081
Capex	0.038	0.045	0.028	0.032	0.015	0.021	0.028	0.039
Ebit ICR	7.417	5.228	1.115	-1.234	-2.613	-2.524	1.064	1.389
$\Delta$ Tot. Assets	0.069	0.086	0.021	-0.030	-0.039	-0.062	0.020	0.040
Size Log(Tot. Assets)	7.262	4.600	7.723	3.515	6.429	3.426	6.773	4.222

Examples Indus

UNIVERSITÄT BERN

#### **Decision Trees**

- Decision Trees are a non-parametric supervised learning method used for classification and regression
- Create a model predicting the value of a target variable by learning simple decision rules inferred from the data features
- We use them to search through the whole range of explanatory variables and find the variables that can best classify and separate zombie vs. non-zombie, distressed vs. non-distressed, and in multi-class trees healthy and recovered

#### **Decision Trees**

UNIVERSITÄT BERN

 $u^{b}$ 



Source: James, Witten, Hastie, Tibshirani: Introduction to Statistical Learning

#### **Decision Trees**

• Rely on CART algorithm



- Consider a sample of input and output (y, X), y is a discrete variable with classes K and  $X = (x_1, x_2, \ldots, x_p)$  includes the input variables
- The algorithm finds the best input variable and split point s

$$\hat{p}_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k),$$

A standard loss function is the Cross-entropy:

$$L(p) = -\sum_{k=1}^{K} \hat{p}_{mk} log(\hat{p}_{mk}),$$

Consider splitting variable  $X_l$  and split point s, define the pair of regions:

$$R_1(l,s) = \{X | X_l \le s\}$$
 and  $R_2(l,s) = \{X | X_l > s\}.$ 

#### Europe 2016

UNIVERSITÄT BERN

 $\boldsymbol{u}^{\scriptscriptstyle b}$ 



Binary Tree: Zombie and Non-Zombie

*pi* Pretax Income, *at* Total Assets, *nopi* Non-Operating Income, *dd1* Long-Term Debt due in 1 Year, *cstk* Common Stock, *re* Retained Earnings

#### **Selected Variables**



Т<sup>ь</sup>

Europe 2016	Zom	bie	Distressed		
Pretax Income	$\leq$ 0	1	Shareholders Equity	$\leq$ 0.5	1
Total Assets	$\leq$ 1910.4	0.9	Pretax Income	$\leq$ 0.1	0.9
Non-Operating Income	$\leq$ 0.0	0.5	Cash	$\leq$ 0.1	0.8
Long-Term Debt 1y	$\leq$ 0.0	0.9	Total Revenue	$\leq$ 1.4	0.8
Common Stock	$\leq$ 0.0	0.9	Retained Earnings	≤0.2	0.9
Retained Earnings	≤-0.2	0.8	Debt in Current Liabilities	$\leq$ 0.0	0.8
			Sale/Turnover	$\leq$ 0.7	0.2
			Total Liabilities	$\leq$ 0.3	0.9

USA 2016	Zon	nbie	Distressed		
Total Assets	$\leq$ 387.2	1	Retained Earnings	≤-0.1	1
Pretax Income	$\leq$ 0.0	0.9	Shareholders Equity	$\leq$ 0.7	0.7
Operating Activities NCF	$\leq$ 0.3	0.8	Capital Surplus	$\leq$ 0.6	0.5
Debt in Current Liabilities	$\leq$ 0.0	0.8	Total Current Liabilities	≤-0.1	0.9
Total Income Taxes	$\leq$ 0.0	0.7	Pretax Income	$\leq$ 0.1	0.9
Capital Expenditures	$\leq$ 0.0	0.8	Stock Return	≤-0.1	0.4
Total Current Assets	$\leq$ 0.5	0.9	Other Current Assets	$\leq$ 0.0	0.9
Total Long-Term Debt	$\leq$ 0.3	0.1	Total Income Taxes	$\leq$ 0.0	0.3

#### Binary Trees: Selected split variables, split points, and entropy of the nodes

#### Europe 2016



Multi-Class Tree: Zombie, Distressed, Recovered and Healthy

*pi* Pretax Income, *at* Total Assets, *nopi* Non-Operating Income, *It* Total Liabilities, *txdb* Deferred Taxes, *seq* Shareholders Equity, *xopr* Operating Expenses, *acox* Other Current Assets Excluding Prepaid Expenses, *recch* Accounts Receivable, *invfg* Finished Goods Stocks, *ppegt* Total Property, Plant, Equipment, *txdi* Deferred Taxes, *revt* Total Revenue, *cstk* Common Stock

Angela De Martiis (University of Bern)

UNIVERSITÄT BERN

h

#### **Selected Variables**

UNIVERSITÄT BERN

 $\boldsymbol{u}^{\scriptscriptstyle b}$ 

Europe 2016	Zombie	Distressed		Healthy		Recovered	
Total Assets	$\leq$ 1951.802	Pretax Income	$\leq -0.009$	Shareholders Equity	$\leq 0.532$	Deferred Taxes	≤-0.0
Total Liabilities	$\leq 0.371$	Operating Expenses	$\leq$ 1.212	Total Revenue	$\leq 0.723$		
Other Current Assets	$\leq 0.0$	Total Liabilities	$\leq 0.396$	Common Stock	$\leq 0.159$		
Accounts Receivable	$\leq 0.018$			Non-Operating Income	$\leq 0.036$		
Finished Goods Stocks	$\leq 0.002$						
Property, Plant, Equip.	$\leq$ 1.245						
LISA 2016	Zombio	Distances		L La a Marco		Deserves	
004 2010	ZUITIDIE	Distressed		Healthy		Hecovered	
Pretax Income	≤-0.017	Shareholders Equity	≤0.446	Total Assets	≤167.365	Current Income Taxes	≤0.003
Pretax Income Debt in Current Liabilities	≤-0.017 ≤0.005	Shareholders Equity Other Current Assets	≤0.446 ≤0.0	Total Assets Income Before Extr. Items	≤167.365 ≤0.026	Current Income Taxes Operating Activities NCF	≤0.003 ≤0.039
Pretax Income Debt in Current Liabilities Financing Activities NCF	≤-0.017 ≤0.005 ≤0.072	Shareholders Equity Other Current Assets Total Assets	≤0.446 ≤0.0 ≤461.482	Total Assets Income Before Extr. Items Total Liabilities	≤167.365 ≤0.026 ≤0.644	Current Income Taxes Operating Activities NCF Shareholders Equity	≤0.003 ≤0.039 ≤0.813
Pretax Income Debt in Current Liabilities Financing Activities NCF Other Current Liabilities	≤-0.017 ≤0.005 ≤0.072 ≤0.229	Shareholders Equity Other Current Assets Total Assets Capital Surplus	≤0.446 ≤0.0 ≤461.482 ≤0.325	Total Assets Income Before Extr. Items Total Liabilities Current Income Taxes	≤167.365 ≤0.026 ≤0.644 ≤0.024	Current Income Taxes Operating Activities NCF Shareholders Equity	≤0.003 ≤0.039 ≤0.813
Pretax Income Debt in Current Liabilities Financing Activities NCF Other Current Liabilities Depreciation Amortization	≤-0.017 ≤0.005 ≤0.072 ≤0.229 ≤0.109	Shareholders Equity Other Current Assets Total Assets Capital Surplus	≤0.446 ≤0.0 ≤461.482 ≤0.325	Total Assets Income Before Extr. Items Total Liabilities Current Income Taxes Sale/Turnover	$\leq 167.365$ $\leq 0.026$ $\leq 0.644$ $\leq 0.024$ $\leq 0.616$	Gurrent Income Taxes Operating Activities NCF Shareholders Equity	≤0.003 ≤0.039 ≤0.813

Multi-Class Trees: Selected split variables and split points

### Support Vector Machines

- Objective of SVM algorithm is to find a hyperplane in the feature space that distinctly classifies the data points
- Effective in high dimensional spaces
- · Decision based on support vectors, robust to outliers

### Support Vector Machines

Position hyperplane by maximizing distance to closest observation



b

UNIVERSITÄT BERN

h

#### Europe 2016



*invrm* Inventories - Raw Materials, *caps* Capital Surplus, *ib* Income Before Extr. Items, *txp* Taxes Payable, *at* Total Assets, *intan* Intangible Assets, *dp* Depreciation Amortization, *pi* Pretax Income

#### **Selected Variables**



 $\boldsymbol{u}^{\scriptscriptstyle b}$ 

Europe 2016	Zombie	Distressed	
Total Assets	-3.8	Intangible Assets	-4.1
Intangible Assets	-1.5	Cash and Due from Banks	-1.3
Depreciation Amortization	-0.9	Capital Expenditures	-1.2
Pretax Income	-0.8	Long-Term Debt	-1.2
Inventories - Raw Materials	0.3	Notes Payable	0.8
Capital Surplus	0.3	Deferred Taxes	1
Income Before Extr. Items	0.3	Investing Activities NCF	1.2
Taxes Payable	0.5	Inventories - Raw Materials	1.3

USA 2016	Zombie	Distressed	
Total Assets	-6.3	Shareholders Equity	-1.3
Inventories - Work in Progress	-1.2	Current Assets	-1.1
Funds from Operations	-1.1	Total Current Liabilities	-1
Other Current Liabilities	-0.9	Sales/Turnover	-0.9
Inventories - Decrease(Increase)	0.2	Depreciation Amortization	2.2
Intangible Assets	0.2	Inventories - Raw Materials	2.6
Long-Term Debt 1 y	0.4	Depreciation of Fixed Assets	2.7
Inventories - Raw Materials	0.9	Long-Term Debt 1 y	3.7

#### SVM: Top input features and weights

### Summary

- Classification accuracy is high (75%-85 % for Trees, 85%-88% for SVM classifier, higher for zombies than distressed)
- Selected features are mostly constant across time
- Trees report lower total asset (size) and negative income decisive splits for Europe and US
- Stock market information and profitability ratios not key determinants
- SVM confirms shareholders equity to classify US distressed not zombie
- SVM supports the importance of total assets for zombie classification for both regions



UNIVERSITÄT

#### Where to go from here?

- Main drawback: endogeneity due to zombie definition
  - Alternative zombie definitions
- Different legislative backgrounds can explain differing determinants across countries
  - · Zoom in at country level, examine differences across industries
- Predict other than classifying
  - Random Forests, Non-linear Kernels to classify unseen samples
- Deeper analysis of the group of recovered zombies

Angela De Martiis Institute for Financial Management University of Bern www.angelademartiis.com

## Thank you!

UNIVERSITÄT BERN

h