Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach

Kristina Bluwstein
Marcus Buckmann
Andreas Joseph
Miao Kang
Bank of England

Sujit Kapadia
European Central Bank

Özgür Şimşek
University of Bath

Use of Machine Learning Algorithms

AEA, San Diego, CA, 3 January 2020

Motivation: Cost and consequences of economic crises

Caption: Migrant mother in the US (left) and bank runs in Berlin (middle) during the Great Depression and in 2007 at a Northern Rock branch in the UK (right).

Sources: Wikipedia & The Guardian.

- Financial crises can have **severe** social, economic and political consequences
- Policy makers would like to **minimise** these costs or avoid them altogether
- **Policy tools**, e.g. macropru, could stabilise system if implemented early enough
- Timely and accurate **prediction** methods needed
- And, **understanding** of the underlying economic mechanisms

Motivation: Cost and consequences of economic crises

Caption: Migrant mother in the US (left) and bank runs in Berlin (middle) during the Great Depression and in 2007 at a Northern Rock branch in the UK (right).

Sources: Wikipedia & The Guardian.

- Financial crises can have **severe** social, economic and political consequences
- Policy makers would like to **minimise** these costs or avoid them altogether
- **Policy tools**, e.g. macropru, could stabilise system if implemented early enough
- Timely and accurate **prediction** methods needed
- And, **understanding** of the underlying economic mechanisms

⇒ Our paper addresses these points using machine learning (ML) for financial crisis prediction

Preview of main results

- **ML models outperform benchmark** logit in out-of-sample prediction and forecasting evaluations
- Shapley value framework enable **well-defined inference** (Joseph, 2019)
- **Small number of factors** explain majority of model output:
    - Credit growth and flat/negative slope of the yield curve at low nominal rates

      Story: search-for-yield in low-interest rate low-returns environment
    - Global factors (also credit growth & slope)

      Story: shared narrative in coupled economic/financial system

    $\Rightarrow$ **Global yield curve slope** new indicator with greatest robustness across long sample

Related literature in financial crisis analysis

- **General/historic:** Minsky (1977); Kindleberger (1978); Bordo et al. (2001); Laeven and Valencia (2008); Reinhart and Rogoff (2009); Cecchetti et al. (2009)
- **Credit:** Borio and Lowe (2002); Drehmann et al. (2011); Schularick and Taylor (2012); Aikman et al. (2013)
- **Yield curve** (not too extensive): Babecký et al. (2014); Joy et al. (2017); Vermeulen et al. (2015)
- **Global factors:** Alessi and Detken (2011); Duca and Peltonen (2013); Cesa-Bianchi et al. (2018)
- **Machine learning:** Ward (2017); Alessi and Detken (2018); Beutel et al. (2018)

Machine Learning (ML) approach

- Statistical toolbox of **non-linear & non-parametric** models mostly originating from computer science with a focus on prediction
- Today **supervised learning**: Universal approximators minimising an error function of the form

$$\mathbb{E}_x\left[||y - \hat{f}(\theta)||_p\right]$$

- Models we **compare**:
    - logistic regression (benchmark)
    - support vector machines (SVM)
    - artificial neural networks
    - tree models (decision tree, random forests & "extreme trees")
- Shapley value and regression framework for **statistical inference**

Pros & Cons of ML relative to econometric approach

Advantages

- Often higher accuracy
- Lower risk of misspecification
- Return richer information set

Pros & Cons of ML relative to econometric approach

Advantages

- Often higher accuracy
- Lower risk of misspecification
- Return richer information set

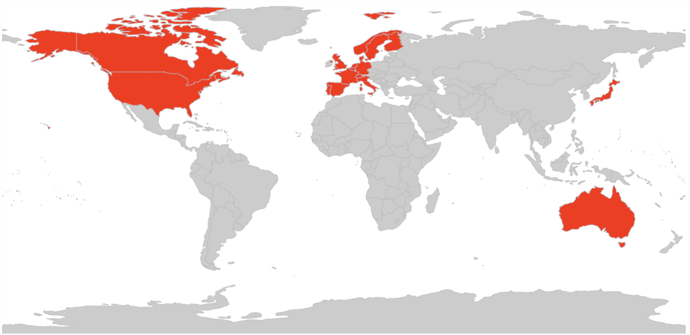
Disadvantages

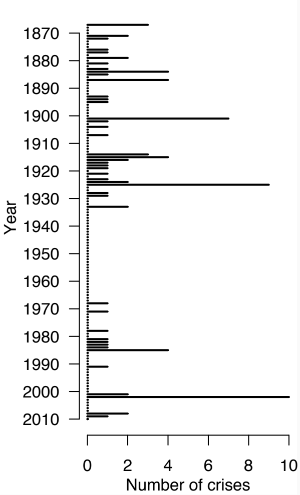
- Higher model complexity ("black box critique")
- Less analytical guarantees, e.g. risk of overfitting
- Often larger data requirement

Jordà-Schularick-Taylor Macrohistory Database

Observations

- 17 developed countries, annual data between 1870 and 2016
- 92 crisis episodes
- 20+ potential indicators

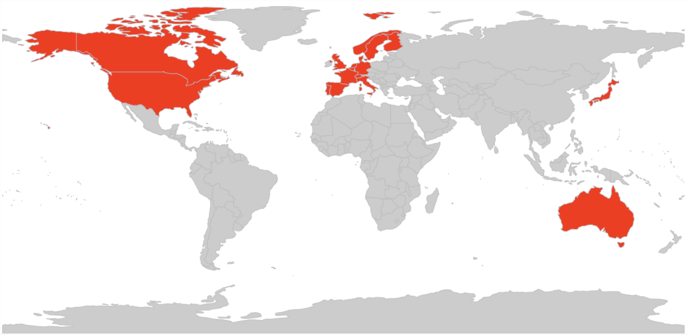




Jordà-Schularick-Taylor Macrohistory Database

Observations

- 17 developed countries, annual data between 1870 and 2016
- 92 crisis episodes
- 20+ potential indicators



Subset of variables we use

- Non-financial credit
- Rates, yield curve
- Debt service ratio
- Current account balance
- Stock Prices
- CPI
- Consumption
- Investment
- Broad money
- Public debt

Empirical approach

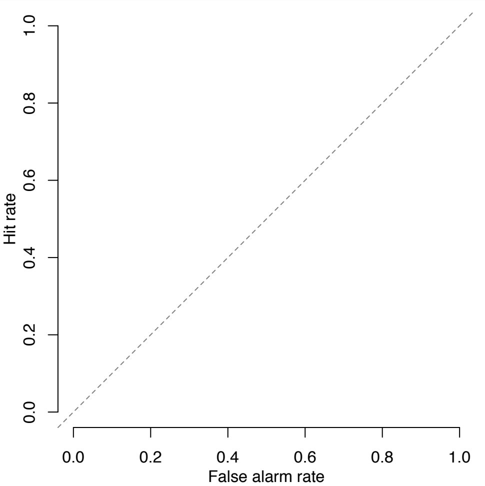
Baseline approach (extensive robustness checks):

- Target: Predict a crisis one and two years in advance (policy space)
- Transformation: 2-year ratio changes or growth rates (sustainability/stationarity)
- Global variables for credit & slope of the yield curve
- Cleaning: Exclude crisis and post-crisis period (5 years), world wars and 1933–1938

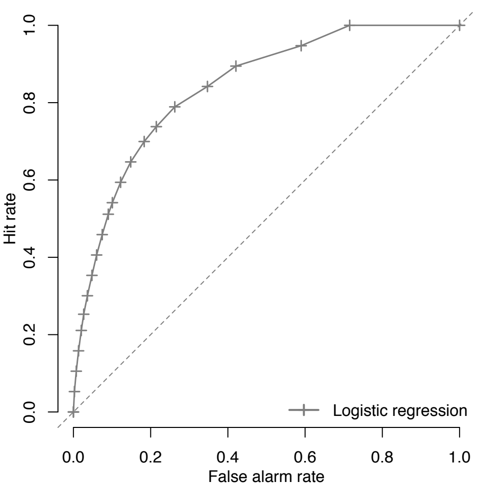
Modelling

- Bootstrapped & averaged models (bagging)
- Out-of-sample evaluation: Nested cross-validation & expanding window forecasting

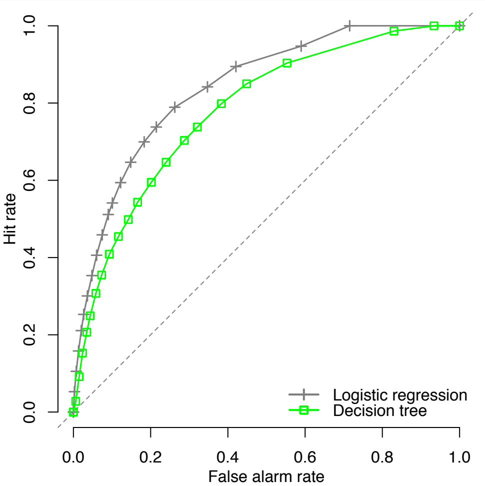
Out-of-sample performance in the ROC space



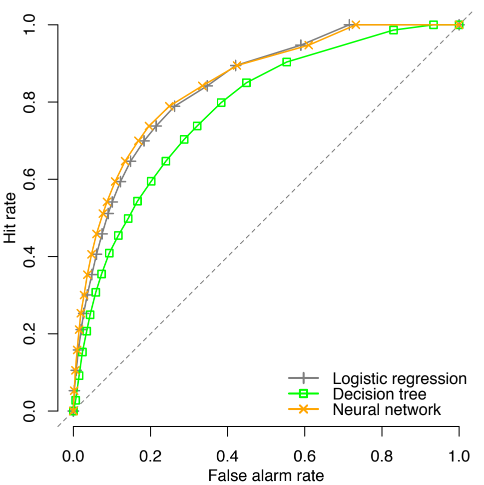
Linear baseline



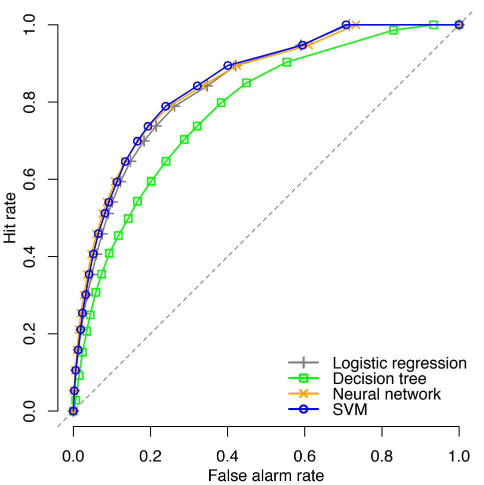
+ Decision trees



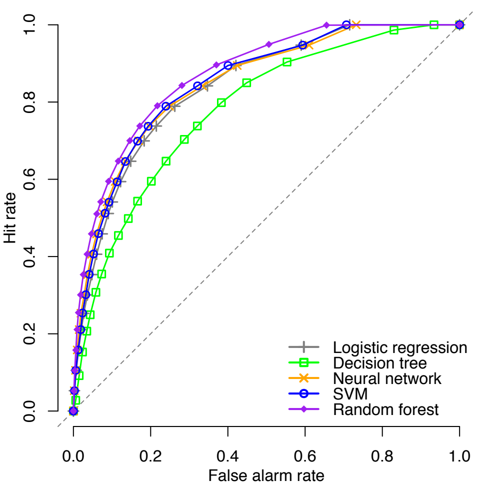
+ Neural network



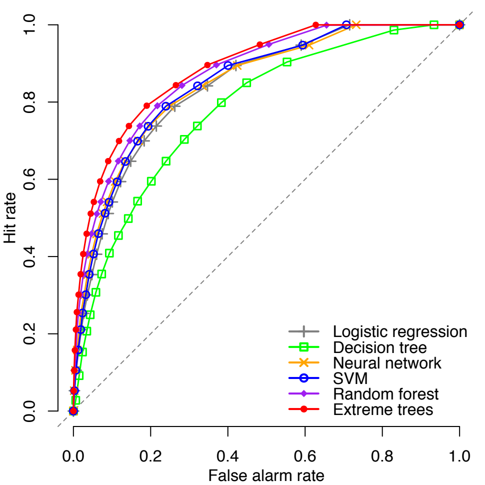
+ SVM


Hit rate
False alarm rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic regression
Decision tree
Neural network
SVM

+ Random forest



The winner is: Extremely randomized trees



Area under the curve (AUC) performance

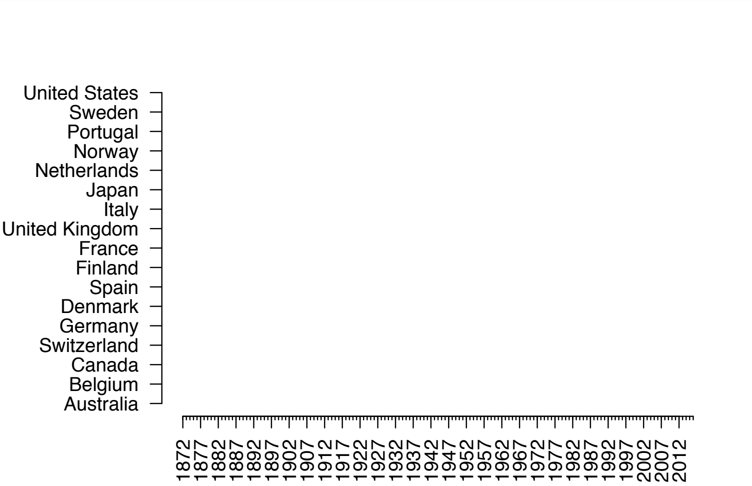
Extreme trees	**0.870**
Random forest	0.855
SVM	0.832
Neural net	0.829
Logistic regression	0.822
Decision tree	0.759

100 replications of 5-fold cross-validation. Standard errors not shown but consistently below 0.002.

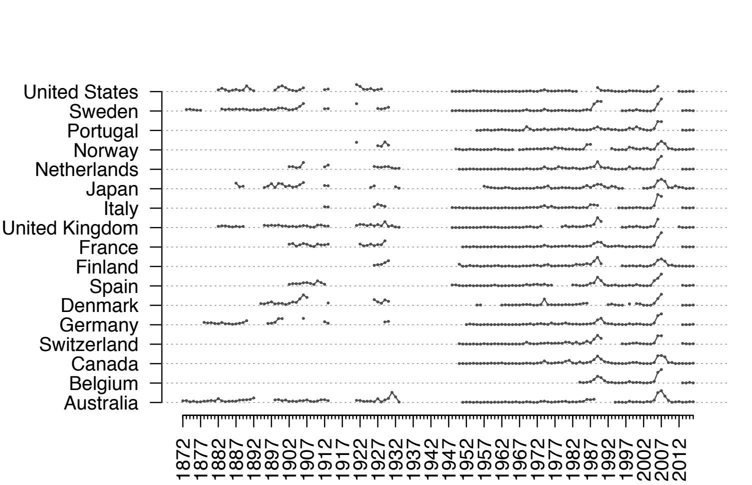
What's the meaning of this differences?

$\Rightarrow$ Aiming at a 80% true positive rate, extreme trees reduce the number of false positives by <u>41%</u> (32%/367 $\rightarrow$ 19%/219) compared to the logistic regression.

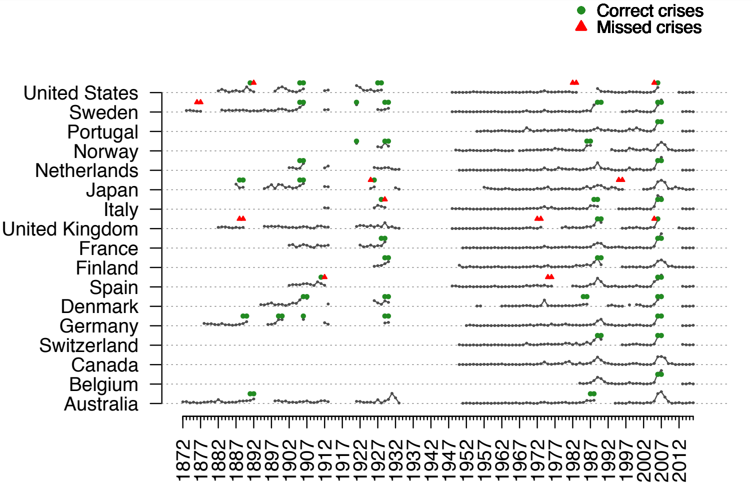
Prediction summary for all countries across time (extreme trees)



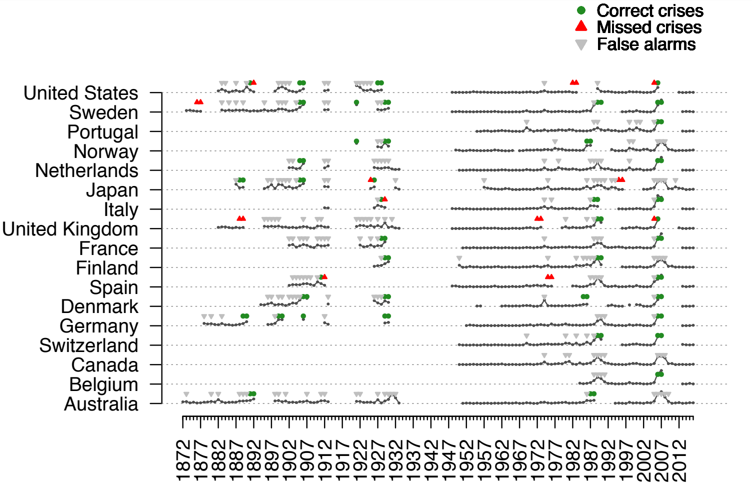
Prediction summary for all countries across time (extreme trees)



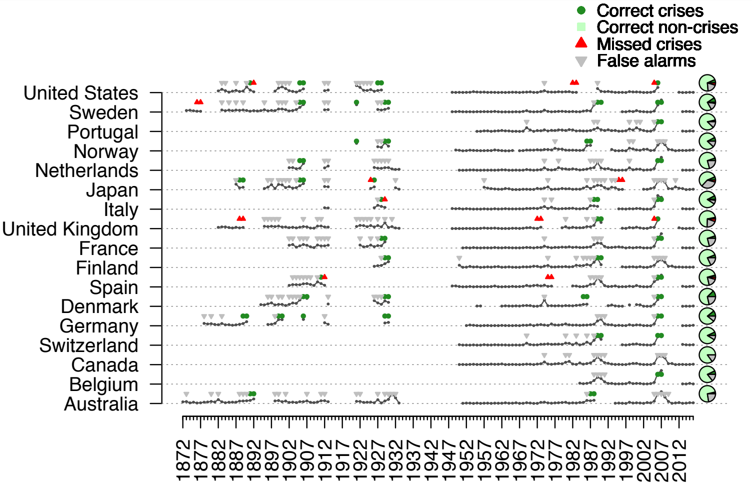
Prediction summary for all countries across time (extreme trees)



Prediction summary for all countries across time (extreme trees)



Prediction summary for all countries across time (extreme trees)



Shapley values for variable importance

	Game Theory	Machine Learning
N	Players	Predictors
$\hat{f}/\hat{y}$	Collective payoff	Predicted value for one observation
S	Coalition	Predictors used for prediction
Source	Shapley (1953)	Strumbelj and Kononenko (2010)
		Lundberg and Lee (2017)

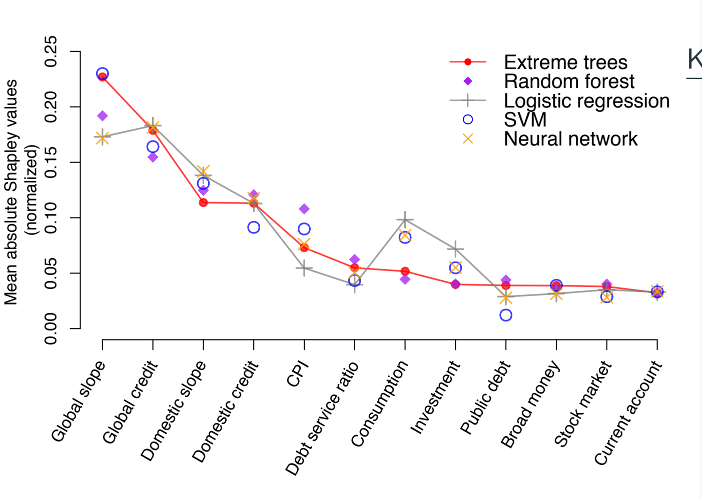
Model Shapley decomposition:

$$\phi_k^S = \sum_{S \subseteq N \setminus k} \frac{|S|!(|N|-|S|-1)!}{|N|!}[\hat{f}(S \cup \{k\} - \hat{f}(S)]$$

$$\Phi^S\left(\hat{f}(x_{ik})\right) = \phi_0 + \sum_{k=1}^{m} \phi_{ik}^S$$

Intuitive example

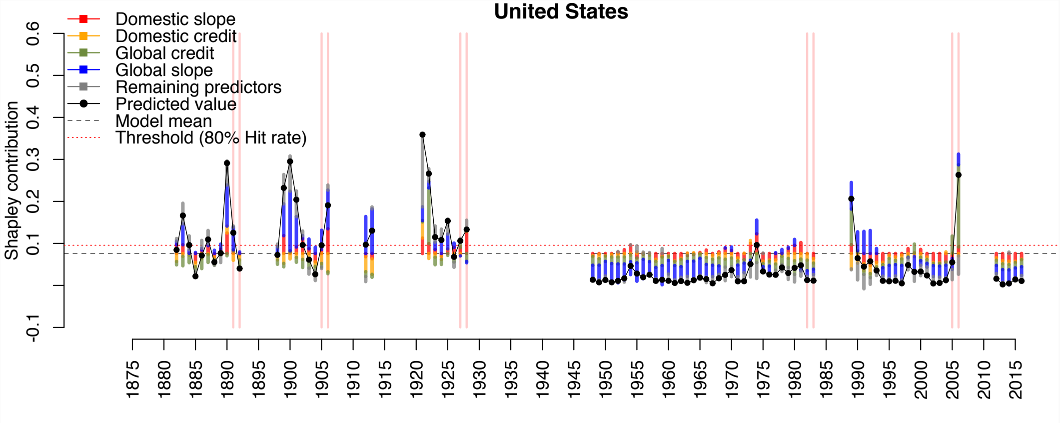
Model explanations using Shapley decompositions: high agreement



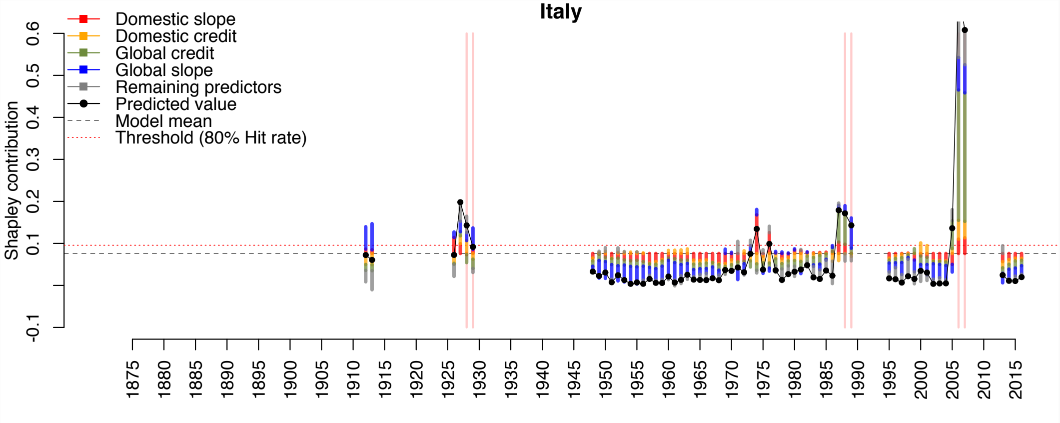
Key indicators:

- Domestic credit (Schularick and Taylor, 2012; Aikman et al., 2013)
- Global credit (Alessi and Detken, 2011; Cesa-Bianchi et al., 2018)
- Domestic slope (Babeckỳ et al., 2014; Joy et al., 2017)
- Global slope (new finding)

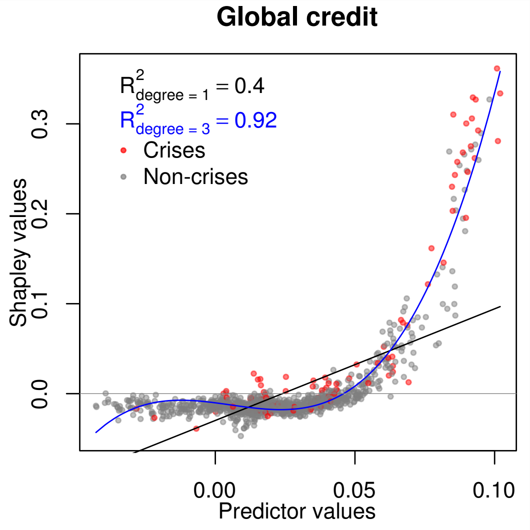
Extreme trees model Shapley value decomposition



Extreme trees model Shapley value decomposition

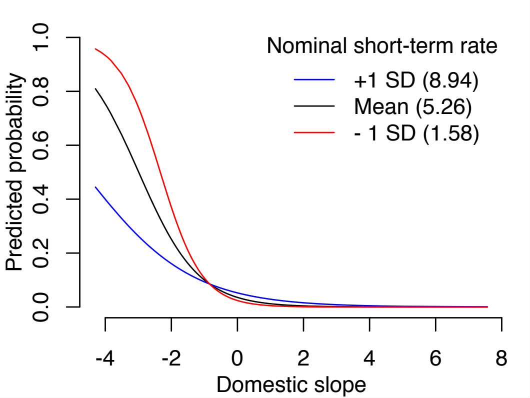


Non-linearity of extreme trees for global credit



- ML models identify strong non-linearities
- Importantly, these are not known a priori
- Directions of associations match those in the linear model

A closer look at the slope of the yield curve



Logit slope interaction with high/low nominal short-term rates.

- Flat or inverted yield curve slope increases predicted crisis probability substantially
- Low nominal short-term rates give stronger interaction effect

  $\Rightarrow$ Likely search-for-yield behaviour
- ML models learn nonlinearity and interactions 'endogenously'

Shapley regression for econometric analysis (Joseph, 2019)

$$\hat{y} \;=\; P[y_{crisis}|x] \;=\; Logit\big(\phi_0 + \hat{\beta}^S\, \Phi^S_{ML}(x)\big) \tag{1}$$

The Shapley values $\Phi_{ML}(x_k)^S$ are interpreted as model-based transformations of variable x_k.

See also: bankunderground.co.uk/opening-the-machine-learning-black-box

(Shapley) regression table for extreme trees

	Shapley regression				Logistic regression		
Name	Direction	Share	α-level	p	Coeff.	α-level	p
Global slope	–	0.23	***	0.000	-0.61	***	0.000
Global credit	+	0.18	***	0.000	0.67	***	0.000
Domestic slope	–	0.11	***	0.000	-0.58	***	0.000
Domestic credit	+	0.11	***	0.000	0.43	***	0.002
CPI	–	0.07	***	0.002	-0.24		0.160
Debt service ratio	+	0.05		0.236	0.16		0.347
Consumption	–	0.05	**	0.029	-0.42	***	0.003
Investment	+	0.04	***	0.005	0.32	**	0.016
other variables	public debt, money, stock prices**, current account						

Table 1: Left: Shapley regression. Direction from logistic regression, p-values against the null hypothesis of neg. or zero regression coefficient (not shown). Right: Coefficients and p-values of a logistic regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Wrap-up

Insights

- **Machine learning models outperform benchmark** logistic regression in out-of-sample financial crisis prediction
- Most important model drivers: **Credit growth & yield curve slope** (domestically & globally)
- ML models learn pronounced **nonlinearities and interactions** from the data
- Especially: global + domestic and slope + low nominal interest rates

Potential policy take-aways

- Yield curve connects monetary policy and financial stability
- System-wide leverage suggests importance of macroprudential tools, e.g. CyCB or LTV/I-ratios
- Global factors suggest importance of international policy coordination

The End: THX - Q & A

Robustness checks (I)

Setup	Crises	Extreme trees	Random forest	Logit regression	SVM	Neural net	Decision tree
Baseline	93	0.84	0.83	0.80	0.79	0.79	0.73
TESTING TRANSFORMATIONS							
Growth rates only	93	0.78	0.77	0.74	0.71	0.72	0.68
Hamilton filter	87	0.82	0.83	0.79	0.78	0.80	0.75
*	87	0.84	0.83	0.80	0.77	0.78	0.76
ADDING VARIABLES							
Nominal rates	93	0.83	0.82	0.80	0.78	0.77	0.73
Real rates	93	0.82	0.82	0.80	0.78	0.79	0.75
Loans by sector	50	0.85	0.84	0.84	0.77	0.82	0.78
*	50	0.87	0.86	0.84	0.76	0.81	0.79
House prices	81	0.86	0.84	0.80	0.78	0.78	0.76
*	81	0.85	0.84	0.80	0.77	0.79	0.76

Robustness checks (II)

Setup	Crises	Extreme trees	Random forest	Logit regression	SVM	Neural net	Decision tree
Baseline	93	0.84	0.83	0.80	0.79	0.79	0.73
CHANGING THE HORIZON							
1 year	93	0.81	0.81	0.80	0.78	0.78	0.71
*	93	0.85	0.83	0.80	0.78	0.79	0.74
3 years	90	0.83	0.83	0.80	0.78	0.77	0.74
*	90	0.84	0.83	0.80	0.79	0.79	0.73
4 years	88	0.86	0.85	0.79	0.80	0.78	0.76
*	88	0.84	0.83	0.80	0.78	0.79	0.75
5 years	87	0.85	0.84	0.79	0.80	0.77	0.75
*	87	0.84	0.83	0.80	0.78	0.79	0.76
PREDICT ONE YEAR BEFORE CRISIS							
	48	0.85	0.81	0.81	0.79	0.80	0.72

Detour: Shapley values in cooperative game theory

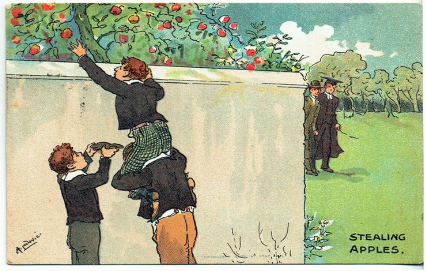
- How much does player A contribute a collective payoff f obtained by a group of n? (Shapley, 1953).
- Observe payoff of the group with and without player A.
- Contribution depends on the other players in the game.
- All possible coalitions S need to be evaluated.

$$\phi_A = \sum_{S \subseteq n \setminus A} \frac{|S|!(|n| - |S| - 1)!}{|n|!}[f(S \cup \{A\} - f(S)] \tag{2}$$

$2^{|n|-1}$ coalitions are evaluated. Computationally complex!

Intuitive example: stealing apples together

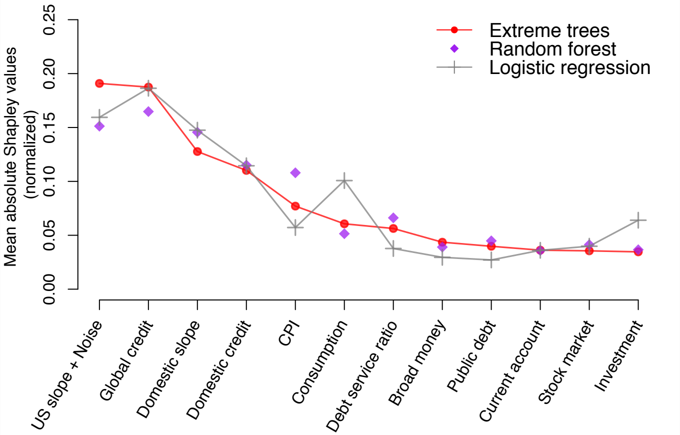
- Three siblings (strong [S], tall [T] & smart [M]) set off to nick some apples A (pay-off) from the neighbour's tree
- For each sibling, sum over marginal contribution to coalitions of one and two
- So, the Shapley value of the strong sibling is then:



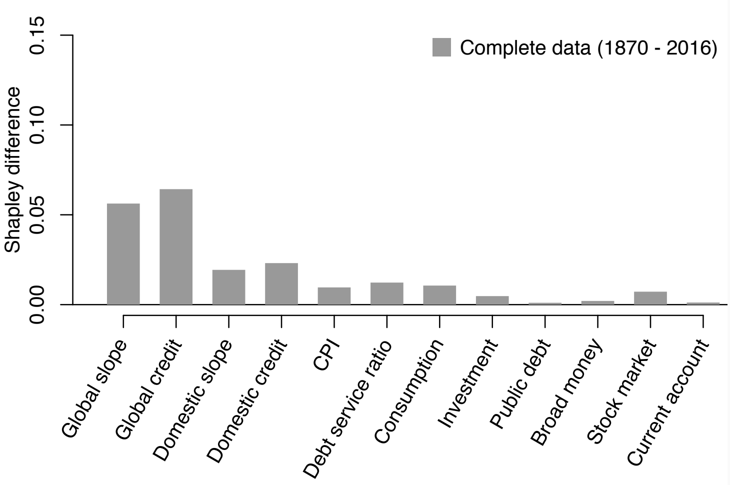
Source: 6oxgangsavenueedinburgh

$$\phi_S = \frac{1}{6}[A(S) - A(\emptyset)] + \frac{1}{6}[A(T,S) - A(T)] + \frac{1}{6}[A(M,S) - A(M)] + \frac{1}{3}[A(T,M,S) - A(T,M)] \tag{3}$$

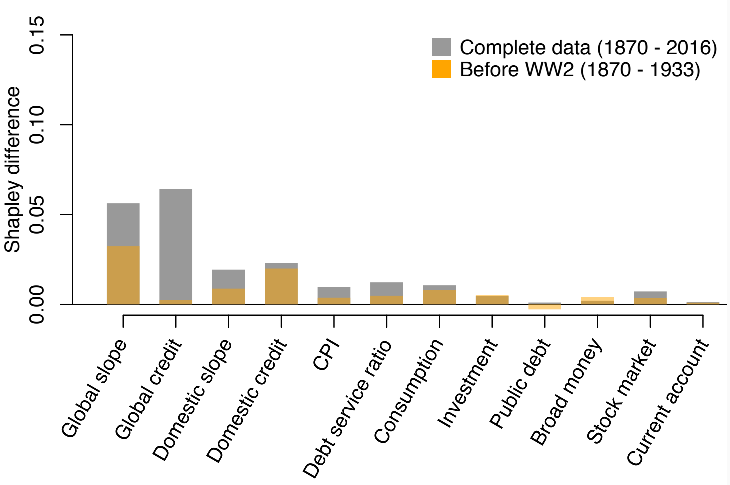
Replacing global slope with US slope



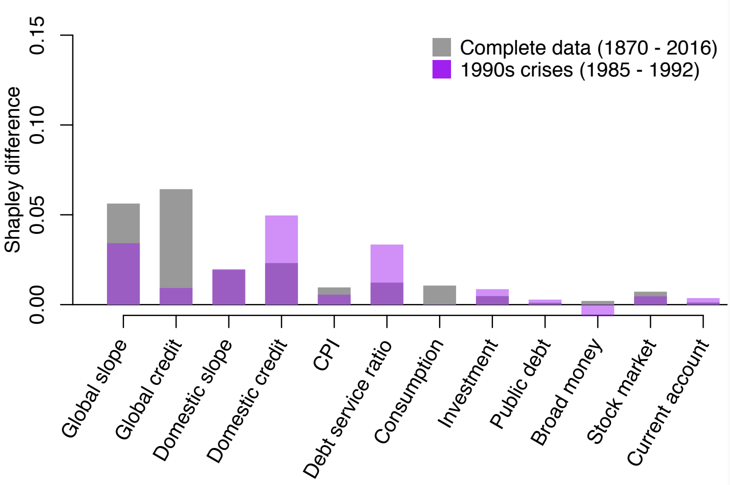
Change of Shapley values over time



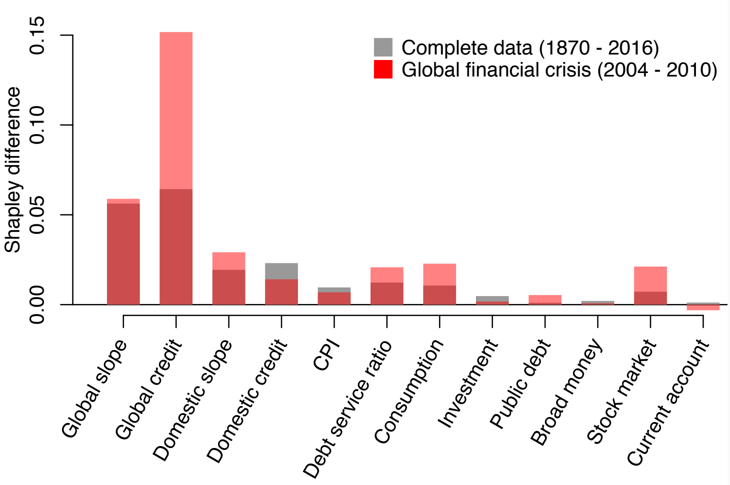
Change of Shapley values over time



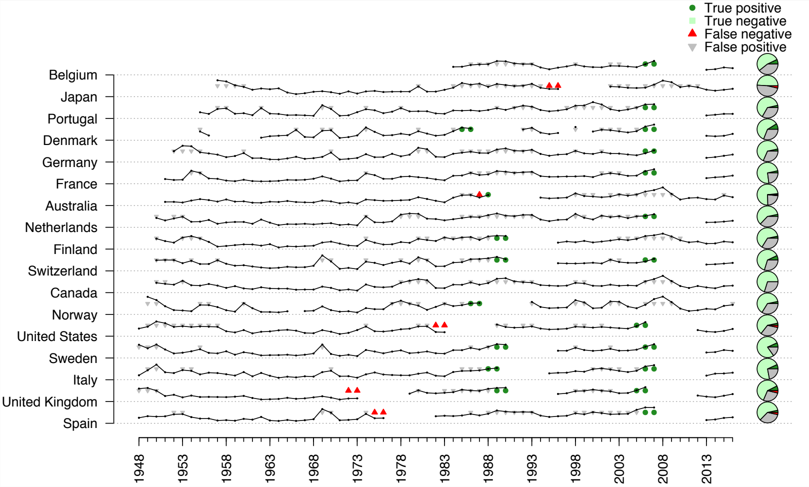
Change of Shapley values over time



Change of Shapley values over time



Neural net forecasting casting evaluation

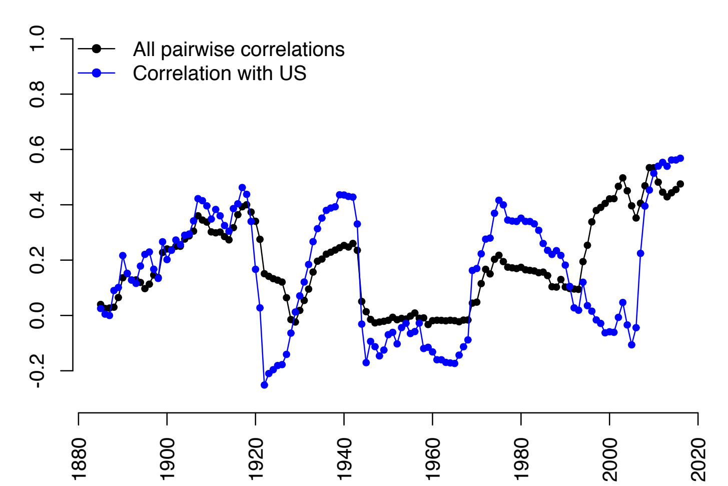


More interactions with *domestic* factors

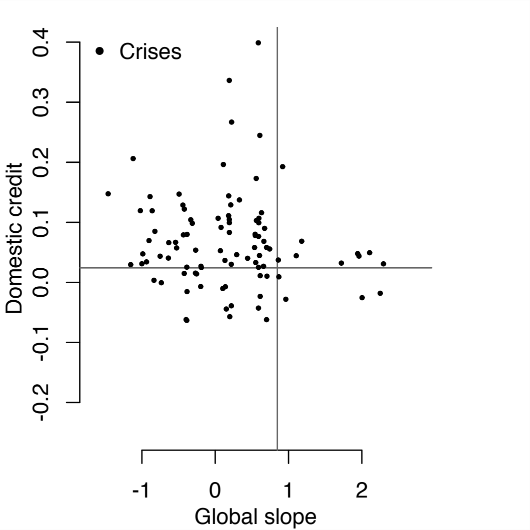
Interaction	Sign	Share	α-lvl	p-values
Domestic slope x Domestic credit	-	0.08		0.154
Domestic slope x Debt service ratio	-	0.15	*	0.051
Domestic slope x Investment	-	0.11	*	0.070
Domestic slope x Consumption	+	0.17	**	0.043
Domestic slope x CPI	+	0.04		0.365
Domestic slope x Stock market	+	0.09		0.109
Domestic credit x Debt service ratio	+	-0.13		0.070
Domestic credit x Investment	+	0.21	***	0.005
Domestic credit x Consumption	-	-0.20		0.005
Domestic credit x CPI	+	0.17	**	0.012
Domestic credit x Stock market	+	-0.17		0.009

Extreme trees interaction terms, α-level: *: 10%, **: 5%, ***: 1%, n-obs: 1249.

Average slope correlations (15yr sliding window)

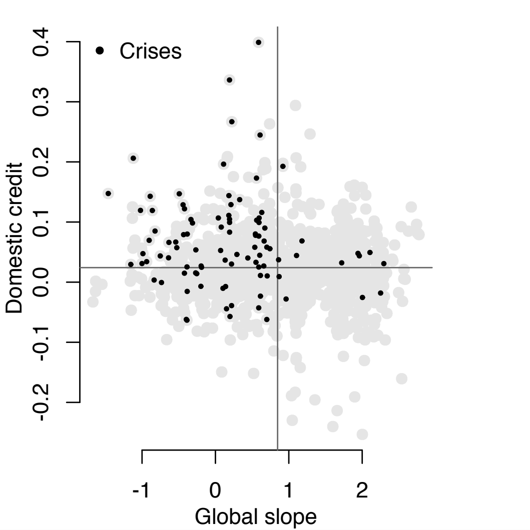


Shapley interactions Effects: E.g. slope and credit



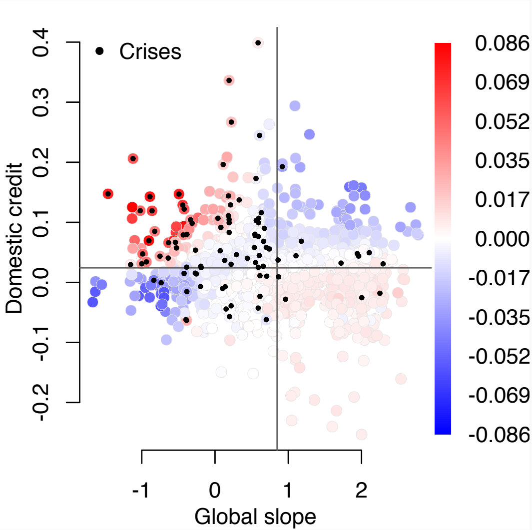
- Many crisis fall into upper left quadrant
- High domestic credit growth *and* flat/negative slope of the global yield curve well separate crisis built-up and normal times.
- Credit booms might be more dangerous in a low/inverted yield curve global environment

Shapley interactions Effects: E.g. slope and credit



- Many crisis fall into upper left quadrant
- High domestic credit growth *and* flat/negative slope of the global yield curve well separate crisis built-up and normal times.
- Credit booms might be more dangerous in a low/inverted yield curve global environment

Shapley interactions Effects: E.g. slope and credit



- Many crisis fall into upper left quadrant
- High domestic credit growth *and* flat/negative slope of the global yield curve well separate crisis built-up and normal times.
- Credit booms might be more dangerous in a low/inverted yield curve global environment

Interaction with global factors important

Interaction	Sign	Share	α-lvl	p-values
Global slope x Global credit	-	0.06	***	0.002
Global slope x Domestic slope	+	0.03		0.169
Global slope x Domestic credit	-	0.07	***	0.004
Global slope x Investment	-	0.04	***	0.000
Global slope x Consumption	+	0.03	*	0.058
Global slope x CPI	+	0.04	***	0.003
Global slope x Stock market	-	0.03		0.185
Global credit x Domestic credit	+	0.03	*	0.083
Global credit x Domestic slope	-	0.03	**	0.027
Global credit x Investment	+	0.02	**	0.036
Global credit x CPI	-	0.04	***	0.001
Global credit x Consumption	-	0.03	***	0.002
Global credit x Stock market	+	0.03	**	0.014

Extreme trees interaction terms, α-level: *: 10%, **: 5%, ***: 1%, n-obs: 1249.

References i

Aikman, D., Haldane, A. G., and Nelson, B. D. (2013). Curbing the Credit Cycle. *The Economic Journal*, 125(585):1072–1109.

Alessi, L. and Detken, C. (2011). Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy*, 27(3):520–533.

Alessi, L. and Detken, C. (2018). Identifying excessive credit growth and leverage. *Journal of Financial Stability*, 35:215–225.

Babeckỳ, J., Havranek, T., Mateju, J., Rusnák, M., Smidkova, K., and Vasicek, B. (2014). Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. *Journal of Financial Stability*, 15:1–17.

Beutel, J., List, S., and von Schweinitz, G. (2018). An evaluation of early warning models for systemic banking crises: Does machine learning improve predictions?

References ii

Bordo, M., Eichengreen, B., Klingebiel, D., and Martinez-Peria, M. S. (2001). Is the crisis problem growing more severe? *Economic policy*, 16(32):52–82.

Borio, C. and Lowe, P. (2002). Asset prices, financial and monetary stability: exploring the nexus. BIS Working Papers 114, Bank for International Settlements.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Cecchetti, S. G., Kohler, M., and Upper, C. (2009). Financial crises and economic activity. Technical report, National Bureau of Economic Research.

Cesa-Bianchi, A., Martin, F. E., and Thwaites, G. (2018). Foreign booms, domestic busts: The global dimension of banking crises. *Journal of Financial Intermediation*.

Drehmann, M., Borio, C., and Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. *International Journal of Central Banking*.

Duca, M. L. and Peltonen, T. A. (2013). Assessing systemic risks and predicting systemic events. *Journal of Banking & Finance*, 37(7):2183–2195.

References iii

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1):3–42.

Joseph, A. (2019). Shapley regressions: A universal framework for statistical inference on machine learning models. *Bank of England Staff Working Paper Series*, (784).

Joy, M., Rusnák, M., Šmídková, K., and Vašíček, B. (2017). Banking and currency crises: Differential diagnostics for developed countries. *International Journal of Finance & Economics*, 22(1):44–67.

Kindleberger, C. P. (1978). *Manias, Panics and Crashes - A History of Financial Crises*. New York: Basic Books.

Laeven, M. L. and Valencia, F. (2008). *Systemic banking crises: a new database*. Number 8-224. International Monetary Fund.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, pages 4765–4774.

Minsky, H. P. (1977). The Financial Instability Hypothesis: An Interpretation of Keynes and an Alternative to"Standard" Theory. *Challenge*, 20(1):20–27.

Reinhart, C. M. and Rogoff, K. S. (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press. Google-Books-ID: ak5fLB24ircC.

Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102(2):1029–61.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.

Strumbelj, E. and Kononenko, I. (2010). An efficient explanation of individual classifications using game theory. *The Journal of Machine Learning Research*, 11:1–18.

Verikas, A., Vaiciukynas, E., Gelzinis, A., Parker, J., and Olsson, M. C. (2016). Electromyographic patterns during golf swing: Activation sequence profiling and prediction of shot effectiveness. *Sensors*, 16(4):592.

References v

Vermeulen, R., Hoeberichts, M., Vašíček, B., Žigraiová, D., Šmídková, K., and de Haan, J. (2015). Financial stress indices and financial crises. *Open Economies Review*, 26(3):383–406.

Ward, F. (2017). Spotting the danger zone: Forecasting financial crises with classification tree ensembles and many predictors. *Journal of Applied Econometrics*, 32(2):359–378.