Discussion of "Answering the Queen: Machine learning and financial crises"

by Jérémy Fouliard, Michael Howell and Hélène Rey

Luiz Awazu Pereira da Silva, Boris Hofmann and Marco Lombardi

Bank for International Settlements

BIS Annual Conference, Zurich, 28 June 2019

We thank Leonardo Gambacorta for suggestions. The views expressed here are not necessarily those of the Bank for International Settlements.

Great paper, novel approach… What the paper does

- Objective: assess financial crisis prediction models for 6 advanced economies over the period 1985-2018 (quarterly data)

- Novel method: focusing on online machine learning technique (sequential prediction) adapted to forecasting financial crisis

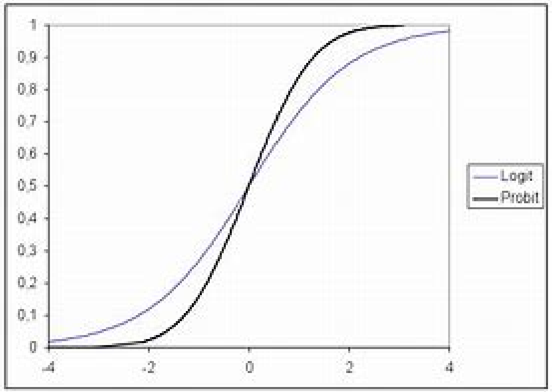
- Finding: Online machine learning considerably improves out-of-sample predictions

- Discussion: (1) What types of answers to H.M. the Queen? (2) Different machine learning techniques; (3) Policy implications

Pre-history of "prediction" of financial crises

- Textbook literature (eg., Minsky, Kindleberger, etc) focuses on big picture intuitions, and "information" of one "expert" using specific variables (eg., bubbles of asset prices, etc) not aiming at predictions, warning about regimes…
  - Answer (1) to H.M. the Queen: *a financial crisis will come, "<u>at some point</u>"*

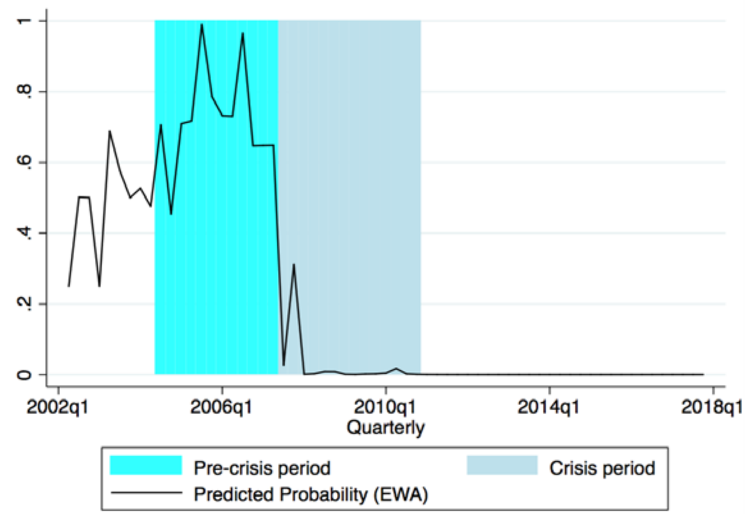
- More modern literature on early warning indicators: several classes of models (Probit, logit, etc.) based on well-known financial variables
  - Answer (2) to H.M. the Queen: *the "<u>probability of a crisis is higher/lower</u>", but "mechanically" linked to the model's few indicator variables and no specific prediction for crisis date*



Novel approach: combine models & machine learning real time

- Combine several, potentially "millions of experts" (models) not just "one"
- Assess predictive performance and select best sequential combination of models (on line machine learning)
- Re-assess based on new information and improve combination of models
- <u>Better</u> out-of-sample prediction not tied to any specific model
  - *Answer (3) to H.M. the Queen: "there is a xx% chance of a financial crisis in the next 12 quarters.. Nobody saw this Global Financial Crisis coming because nobody used machine learning"*

Predicted probability of crisis UK



Congratulations to Hélène, Jérémy and Michael, very promising new path!

Answers to H.M. The Queen but helps policy-makers: accurate timing prediction and variable identification allows "preventive" policy action

- Paper also identifies relevant variables for best predictions at each point in time → better insights into variables that increase financial risk (usual suspects)

- Policymakers then can select and deploy best tools to "mitigate and/or reduce probability" of financial crises (more on that later)

- With early interventions and more targeted policies → less financial crises → lower output losses → Social Welfare improves

- Learning & adapting by incorporating new information? Parallel with other "scientific" progress?

"My mind seems to have become a kind of machine for grinding general laws out of large collections of facts" Origin of Species, Charles Darwin

Two different machine learning techniques and “big data”

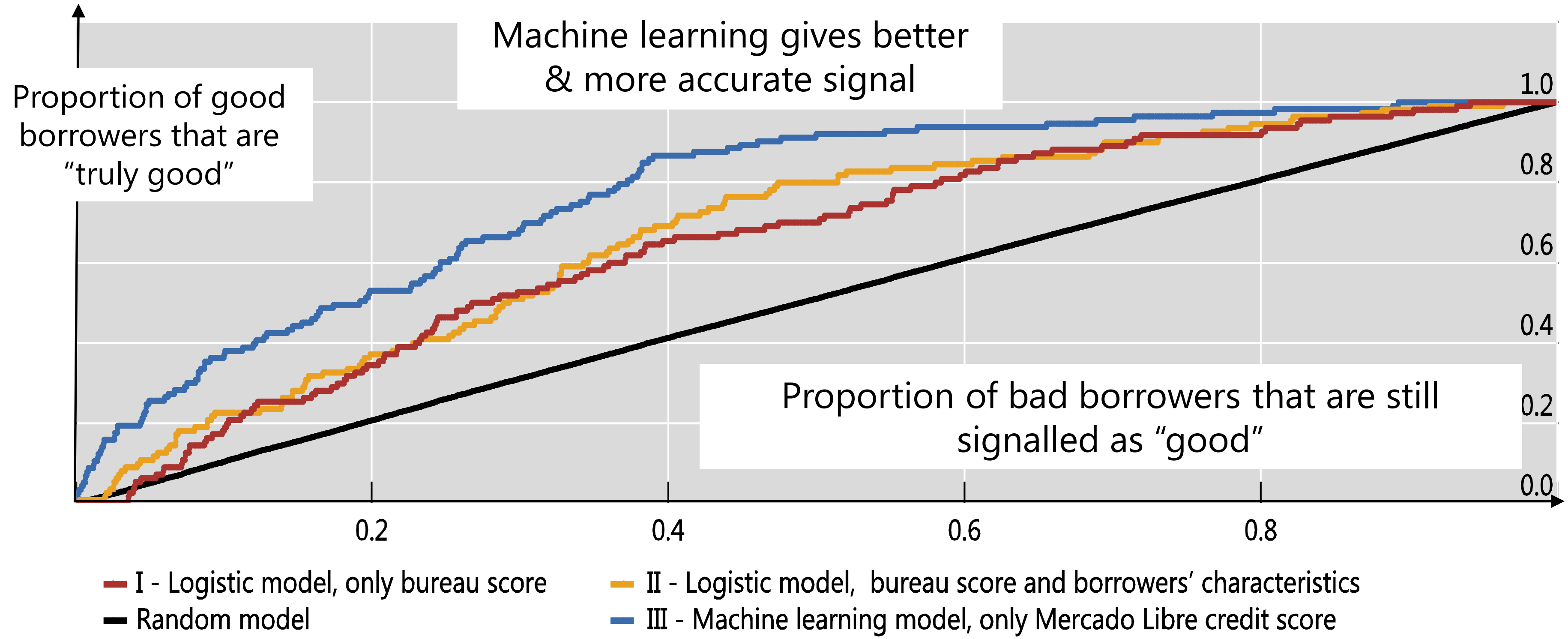
- Obviously no “big data” set for systemic crises
  - Systemic crises are rare events; only one crisis (early 1990s) to forecast the second crisis in the sample period (GFC)

- Therefore paper considers 18 “experts” (models) but can be extended to “millions” of experts (models)
  - Experts (models) will use “big data” set for all variables that are associated with systemic crises; but “learning” comes with the better performance of the models

- Other machine learning applications are used to predict frequent “specific events” (eg., defaults based on credit scores, etc.); they rely on much larger big data sets
  - “Learning” comes from non-traditional data sources; see recent BIS research (Frost, Gambacorta, Huang, Shin and Zbinden (2019))

Performance of machine learning with credit scoring models

Machine learning to differentiate "good" from "bad" borrowers (credit score) compared with two different models and random signal



The credit-scoring algorithm used by Mercado Libre (e-commerce firm in Latin America). Ideally, there should only be good signals for good borrowers and bad for bad borrowers. The random model is the 50-50 chance good-bad, pictured here by the 45 degree line. The better is the credit scoring model at signalling credit risk, the higher is the curve above the 45 degree line.

Sources: Mercado Libre; J Frost, L Gambacorta, Y Huang, H S Shin and P Zbinden, "BigTech and the changing structure of financial intermediation", BIS Working Papers, no 779, April 2019.

The machine learning … complemented by other models

- Machine learning designed and capable of producing good predictions, but is agnostic about "theory", i.e., underlying causations
  - This is ok if the goal is credit scoring to measure credit risk
  - Better performance for measuring probability of financial crisis

- But once "probable timing" of crisis is known, need other approaches and different classes of models to analyse trade-offs between various instruments (eg LAW with MP, MaPs, FP, FXI, everything?)

- Because sign and size of these correlations are crucial inputs for the calibration of instruments to be used, for ex, MP and MaP policies
  - How do central banks map crisis risk signals into policy decisions?
  - How do the drivers of the signals respond to policy measures?
  - How to take into account the effects of other policies?

Policy implication of this new research path: selection of accurate "timing" and best combination of "instruments"

- Earlier prediction can improve "timing" of treatment, when to begin administrating countervailing measures (policy instruments)
  - Analogy with medical science: early cancer detection → less invasive surgical treatment → greater likelihood of survival

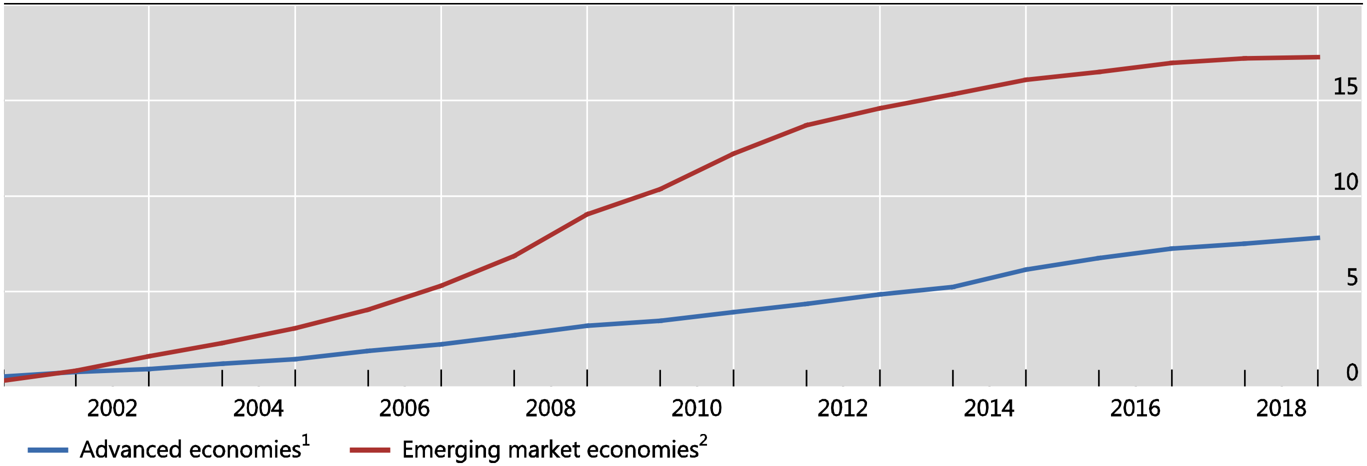
But needs to be complemented with…

- Knowledge of relative effectiveness of instruments / policies to be administered: what is the best combination and dosage of treatment?
  - Same analogy: using broad spectrum medicine (chemotherapy) and/or more targeted ones (laser therapy)

Example: Macroprudential policies are increasingly being used for crisis prevention

Use of macroprudential tools

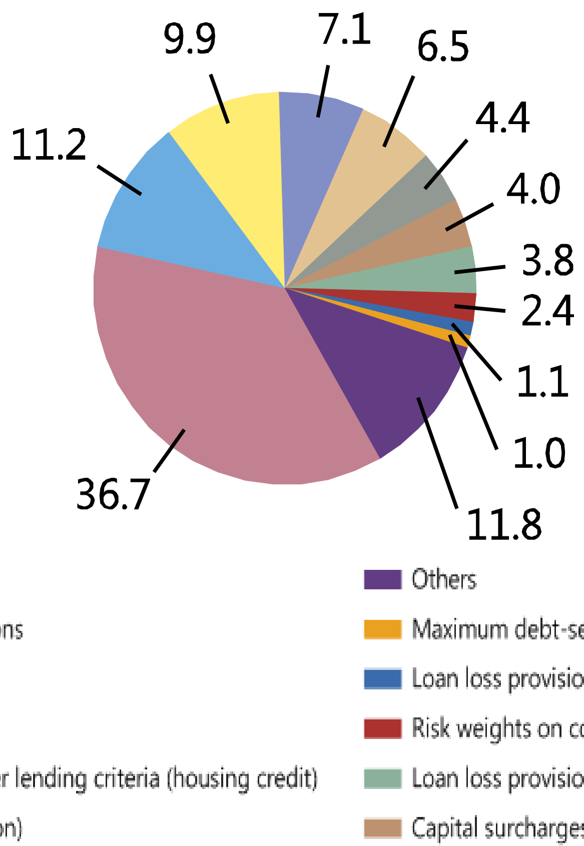
Number of measures



[1] AT, AU, BE, CA, CH, DE, DK, EE, ES, FI, FR, GR, GB, IE, IS, IT, LT, LU, LV, NL, NO, NZ, PT, SE , SI, SK and US; simple average across countries. [2] AE, AR, BG, BR, CL, CN, CO, CZ, HK, HR, HU, ID, IL, IN, KR, MX, MY, PE, PH, PL, RO, RS, RU, SA, SG, TH, TR and ZA; simple average across countries.

Sources: I Shim, B Bogdanova, J Shek and A Subelyte, "Database for policy actions on housing markets", *BIS Quarterly Review*, September 2013, pp 83–95; BIS calculations

But what MaP should be used? policy actions in the EU 95-2014



[1] Out of 1361 actions.

Sources: Budnik, K and J Kleibl (2018): "Macroprudential regulation in the European Union in 1995–2014: introducing a new data set on policy actions of a macroprudential nature", ECB Working Papers,no 2123, January; Reinhardt, D and R Sowerbutts (2015): "Regulatory arbitrage in action: evidence from banking flows and macroprudential policy", Bank of England, Staff Working Papers, no 546, September; Shim, I, B Bogdanova, J Shek and A Subelyte (2013): "Database for policy actions on housing markets", BIS Quarterly Review, September, pp 83–91; national data; BIS calculations.

So next steps? Machine learning to inform policies to avert financial crises

- There is more analytical (DSGE) and empirical evidence that MaP policies:
  - Moderate macroeconomic and financial volatility (Agenor et al 2018)
  - Restrain excessive credit growth (Richter et al 2018)
  - Result (A): relative "dosage" of MaP instruments to be used...
- Possible extension of this paper, showing that MaP affects probability of crises, improve "learning"; could confirm MaPs "preventive power",
  - One could try introducing "experts" who avail of indices of MaP policies deployed;
  - Result (B): machine learning gives best "timing" for their usage...
- Combining Results (A) & (B): (1) what is the best "timing" and threshold for policy intervention? (2) with what dosage and instrument do you intervene?
  - A task for financial stability committees? What models are needed?

Win-win for Social Welfare? Knowledge of "timing" and of "effectiveness" of policy interventions

- This paper → early detection of high probability of financial crisis gives "window" for policy intervention

- MaP effectiveness literature + extensions of this paper gives best possible "combination" of instruments to deploy in the window:
  - MaP1 (eg. CCyCB)
  - MaP2 (eg. LTV)
  - MaP3
  - Etc.