Answering the Queen: Online Machine Learning and Financial Crises

Jérémy Fouliard, Michael Howell, Hélène Rey

BIS Annual Conference
This does not reflect the views of the French Macroprudential Authority

June 28, 2019

Forecasting the financial crisis of 2008

Visiting the LSE and being shown how terrible the situation was and had been, the Queen asked: "Why did nobody notice it?"

Motivation

- The Queen got a terse answer. The economic and finance profession have a bad record at predicting crises.

Motivation

- The Queen got a terse answer. The economic and finance profession have a bad record at predicting crises.
- After several months, the Economic Section of the British Academy wrote a three-page missive to Her Majesty blaming the lack of foresight of the crisis on the "failure of the collective imagination of many bright people".

Motivation

- The Queen got a terse answer. The economic and finance profession have a bad record at predicting crises.
- After several months, the Economic Section of the British Academy wrote a three-page missive to Her Majesty blaming the lack of foresight of the crisis on the "failure of the collective imagination of many bright people".
- This paper aims at predicting systemic crises well in advance (12 quarters ahead) using cutting-edge machine learning tools.

Literature Review

- **The Signal Approach** : Kaminsky [1998], Kaminsky and Reinhart [1999], Borio and Lowe [2002], Gourinchas and Obstfeld (2012), Drehmann and Juselius [2013]
- **Logit Specifications** : Demirgüç-Kunt and Detragiache [1998], Eichengreen and Rose [1997], Bongini et al. [2000], Frankel and Saravelos [2012], Schularick and Taylor [2012], Coudert and Idier [2017]
- **Machine Learning** : Davis and Karim [2008], Duttagupta and Cashin [2012], Ward [2014], Joy et al. [2017], Alessi and Detken [2018].
- **Online Learning**: Cesa-Bianchi and Lugosi [2006]; Stoltz and al. [2013].

New Framework: Online learning

This framework is very suitable for crisis prediction:

- **Multivariate** : Which variables cause a financial crisis?
- **Time-varying weights** : Causes of financial crises may be different over time.
- **Statistically robust** : overfitting is a problem in the literature.
- **Not "black-box"** : assess the role each model plays to predict the pre-crisis.
- **Theoretically grounded** : asymptotic properties of our aggregation rules ensure convergence (Cesa-Bianchi and Lugosi[2006]; Stoltz and al.[2013]).
- **More general than Bayesian Model Averaging**

New Framework: Online learning

This framework is very suitable for crisis prediction:

- **Multivariate** : Which variables cause a financial crisis?
- **Time-varying weights** : Causes of financial crises may be different over time.
- **Statistically robust** : overfitting is a problem in the literature.
- **Not "black-box"** : assess the role each model plays to predict the pre-crisis.
- **Theoretically grounded** : asymptotic properties of our aggregation rules ensure convergence (Cesa-Bianchi and Lugosi[2006]; Stoltz and al.[2013]).
- **More general than Bayesian Model Averaging**
- This framework has not been applied in economics. Used to predict French electricity load (EDF); the tracking of climate models; the network traffic demand.

Sequential prediction with expert advice

Sequential prediction

Online learning is performed in a sequence of consecutive rounds where at time instance t the forecaster:

1. Receives a question.
2. Uses expert advice $\{f_{j,t} \in \mathcal{D} : j \in \mathcal{E}\}$
3. Predicts $\hat{y}_t \in \mathcal{Y}$
4. Receives true answer $y_t \in \mathcal{Y}$
5. Suffers a loss $\ell(\hat{y}_t, y_t)$.

Sequential prediction with expert advice

To combine experts' advice, the forecaster chooses a sequential aggregation rule $\mathcal{S}$ which consists in setting a time-varying weight vector $(p_{1,t}, ..., p_{N,t}) \in \mathcal{P}$:

$$\hat{y}_t = \sum_{j=0}^{N} p_{j,t} f_{j,t}$$

The forecaster incurs a cumulative loss defined by :

$$L_T(\mathcal{S}) = \sum_{t=1}^{T} \ell(\sum_{j=0}^{N} p_{j,t} f_{j,t}) = \sum_{t=1}^{T} (\hat{y}_t - y_t)^2$$

Sequential prediction with expert advice

- How can we measure the performance of a sequential aggregation rule ?

Sequential prediction with expert advice

- How can we measure the performance of a sequential aggregation rule ?
- The difficulty is that we do not have any ideas about the generating process of the observations.
- Forecaster's performance is relative. We define the regret :

$$R_{j,T} = \sum_{t=1}^{T} (\ell(\hat{y}_t, y_t) - \ell(f_{j,t}, y_t)) = \hat{L_T} - L_{j,T}$$

where $\hat{L_T} = \sum_{t=1}^{T} \ell(\hat{y}_t, y_t)$ denotes the forecaster's cumulative loss and $L_{j,T} = \sum_{t=1}^{T} \ell(f_{j,t}, y_t)$ is the cumulative loss of expert j.

Sequential prediction with expert advice

We **minimize the regret** with respect to the best combination of experts:

$$R(\mathcal{S}) = \hat{L}_T(\mathcal{S}) - \inf_{q \in \mathcal{P}} L_T(q)$$

We only select aggregation rules with a "vanishing per-round regret" (regret goes to zero asymptotically).

For each selected aggregation rule, the regret is bounded by the smallest possible quantity which depends on T, on the learning rate and on the log(number of experts).

Sequential prediction with expert advice

This approach is a **meta-statistic approach**: the aim is to find the best sequential combination of experts (who can be any economic models or judgement).

$$\hat{L}_T(\mathcal{S}) = \inf_{q \in \mathcal{P}} L_T(q) + R(\mathcal{S})$$

Forecaster's cumulative loss is the sum of :

- **An approximation error** : given by the cumulative loss of the best combination of experts.
- **An estimation error** : given by the regret and which measures the difficulty to approach the best combination of experts.

The rule of the game is to minimize the regret and to combine best possible experts: garbage in, garbage out!.

Sequential prediction with expert advice and delayed feedback

- We adjust the approach to take into account the fact that we only know whether we are in a pre crisis after 12 quarters.
- **Experts** at t are estimated using information available until **t-1**.
- **Aggregation rules**: The forecaster receives information on expert's performance with a delay of **12 quarters**.

Data of systemic crisis episodes: off-the-shelf

- The ECB provides an official database of systemic crisis episodes [Lo Duca et al., 2017q] and additional smaller crises. Judgement of national authorities is involved.

- Our sample starts in 1985q1 and ends in 2018q1.
- Our sample includes 6 countries (so far): France, Germany, Italy, Spain, Sweden, UK.

Macroeconomic variables

Our database contains commonly used Early Warning Indicators (n=101) with transformations (1-y, 2-y, 3-y change and gap-to-trend). It could be enriched. They are not vintage data.

- **Macroeconomic indicators** : Current account, Consumer Price Index, GDP, M3, Unemployement rate, Cross-border flows, Total Liquidity Index.
- **Credit indicators** : Bank (or Total) credit, Household debt, Debt Service Ratios (household, non-financial corporations, non-financial sector)
- **Interest rates indicators** : 3-month rate, 10 years rate, slope of the yield curve (10y-3m), gap of 10-y rate to real GDP.
- **Real estate indicators** : Loans for House purchase, Residential real estate prices, Price-to-income ratio, Price-to-rent ratio.
- **Market indicators** : Real effective exchange rate, Stock prices, Financial Conditions Index, Risk Appetite Index.

Oecumenical Choice of Experts : Horse Race

Horse race of existing models

We first take as given the models developed by some European central banks to predict pre-crisis period and summarized by the Macro-prudential Research Network :

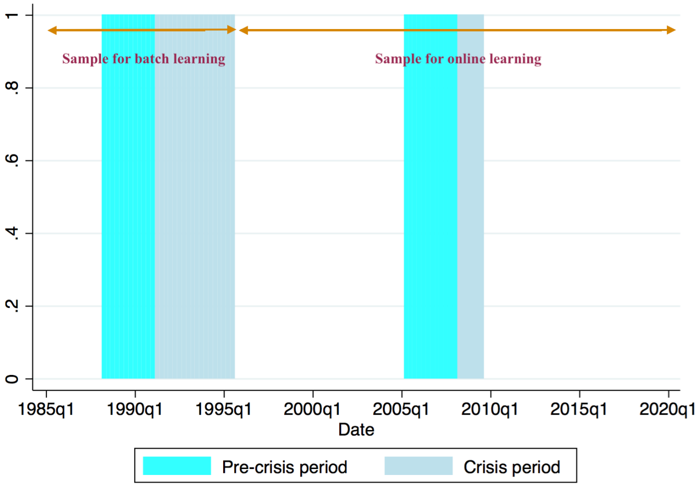
- Bank of England : Panel logit - fixed effect - PCA (Bush et al.)
- Bank of Portugal : Dynamic probit model (Antunes et al.)
- ECB : Panel logit - fixed effect
- Oesterreichische Nationalbank : Bayesian Random Coefficient Logit (Neudorfer and Sigmund)
- Czech National Bank : Bayesian Random Coefficient Logit Panel model (Babecky and al.)
- Bank of England : Binary Classification Tree (Joy et al)
- ECB : Random forests (Alessi and Detken)
- Bank of Finland : Bubble models (Taipalus et al.)

Horse race

We also add two other classes of experts :

- Logit with elastic net penalty for 5 groups of variables :
  - Housing : Price-to-rent, Price-to-income, Real estate prices.
  - Real Economy : GDP, Multi-factor productivity, Current account, Unemployment rate.
  - Bank variables : Total Assets, Total Equity, Leverage.
  - CrossBorder Capital's variables : Risk Appetite, Financial Condition Index, Total Liquidity Index, Cross-border flows.
  - Monetary/Financial : M3, Short-term interest rate, Long-term interest rate.
- Two "lazy" experts : Bashful and Lazy.
  - Bashful : simple logit regression with three best Principal Component scores.
  - Lazy: simple logit regression with the 10 best predictive variables selected using an AUROC analysis (mainly Housing variables)

France



1
.8
.6
.4
.2
0
Sample for batch learning
Sample for online learning
1985q1
1990q1
1995q1
2000q1
2005q1
2010q1
2015q1
2020q1
Date
Pre-crisis period
Crisis period

Forecasting the pre-crisis period out-of-sample

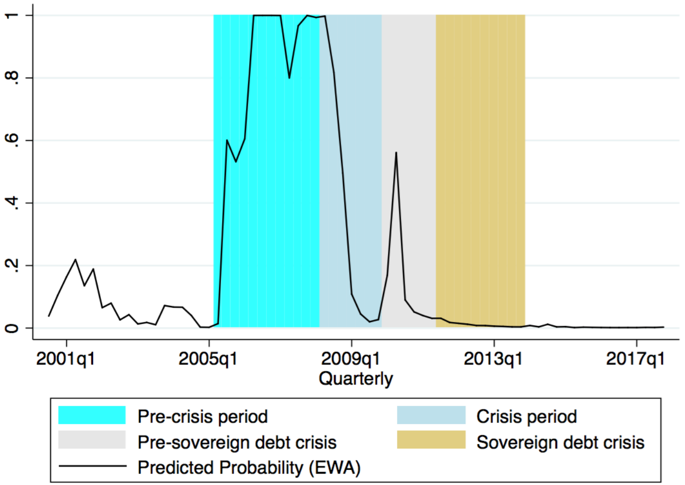


Figure 1: Predicted probability - EWA

Forecasting the pre-crisis period out-of-sample

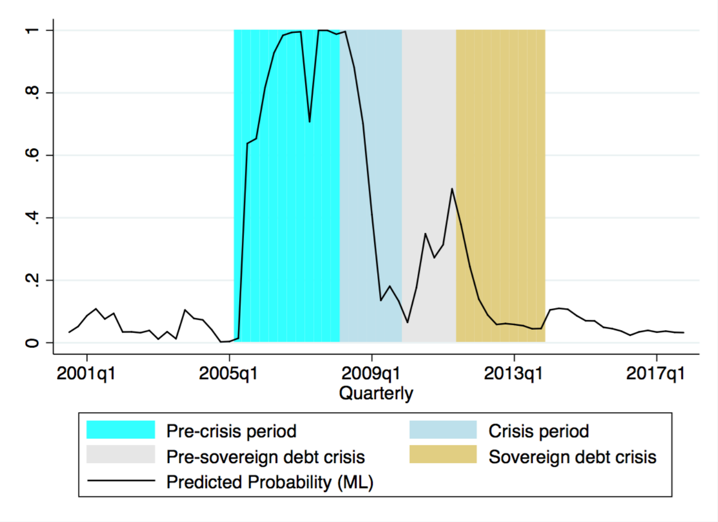


Figure 2: Predicted probability - ML

Forecasting the pre-crisis period out-of-sample

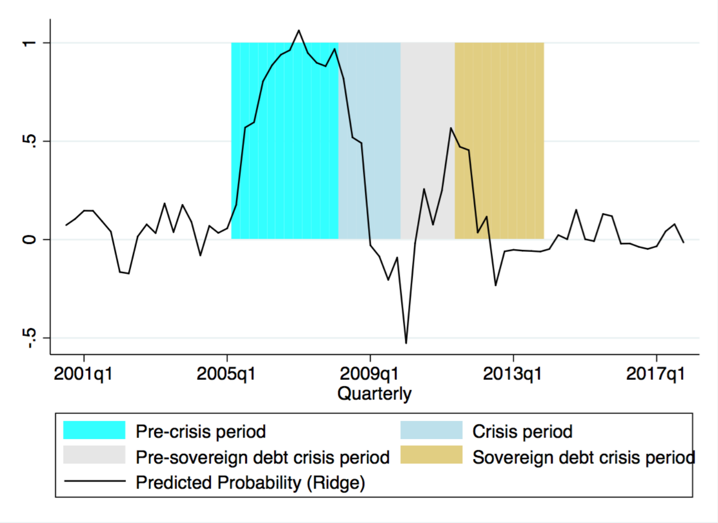


Figure 3: Predicted probability - Ridge

France

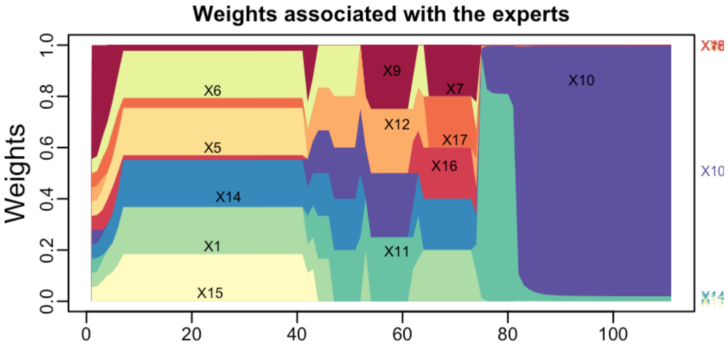


Figure 4: Weights - EWA

France: Sequence of Experts

- Period [1985q1-1996q1]: estimation in-sample. Winners are models with many variables (X1 total credit, price to rent, to income , consumer prices; X14 -bashful; X15 -lazy; X5 binary classification tree panel; X6 binary classification tree country). Probably overfitting.
- Period [1996q1-onwards]: online learning with delayed feedback (frozen weights for the first 12 quarters).

France: Out-of-sample: Endogenous reweighting

- Period [1996q1-2005q1]: calm. Broad models.
- Period [2005q2-2007q4]: pre- systemic crisis. Models dominating are X9 (credit BIS), X10 (risk appetite, financial conditions, share price) X11 (monetary: interest rate, M3); X12 (bubble).
- Period [2008q1-2009q4]: systemic crisis. Models dominating are X14 (bashful); X1; X16 (housing; real economy; X17 (housing, credit BIS); X7 (housing).
- Period [2010q1-2010q4]: pre euro-crisis. Models dominating are X11 (monetary) and X10 (risk appetite, financial conditions, share price).
- Period [2011q1-2013q4]: euro-crisis. Model dominating is X10 (risk appetite, financial conditions, share price)
- Period [2013q4-]: tranquil. X10.

Root Mean Square Errors

Online aggregation Rule	RMSE
EWA	.0899
Uniform	0.317
Best expert (ex post)	0.0093
Best convex combination (ex post)	0.0073

Table 1: RMSE France

AUROC

- The ROC curve represents the ability of a binary classifier by plotting the true positive rate against the false positive rate for all thresholds.
- The AUROC is the area under the ROC curve :

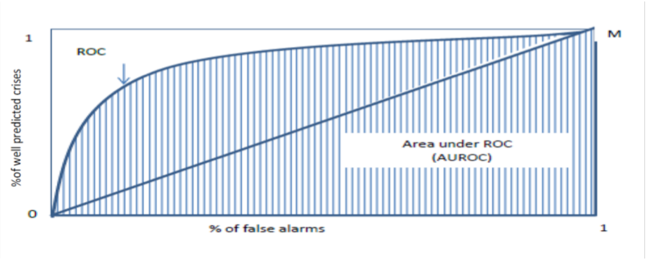


Figure 5: ROC curve

AUROC: France, out-of-sample

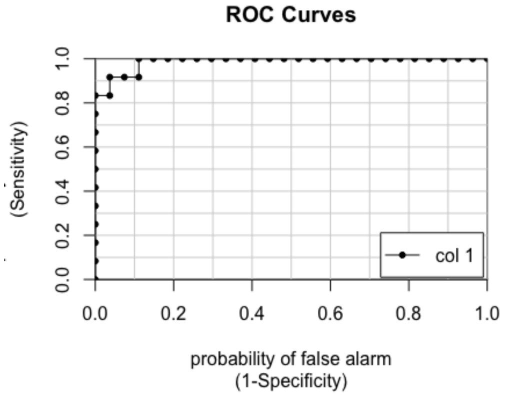


Figure 6: AUROC- EWA

$AUROC(EWA) = 0.988$

Germany: pre-crisis period out-of-sample

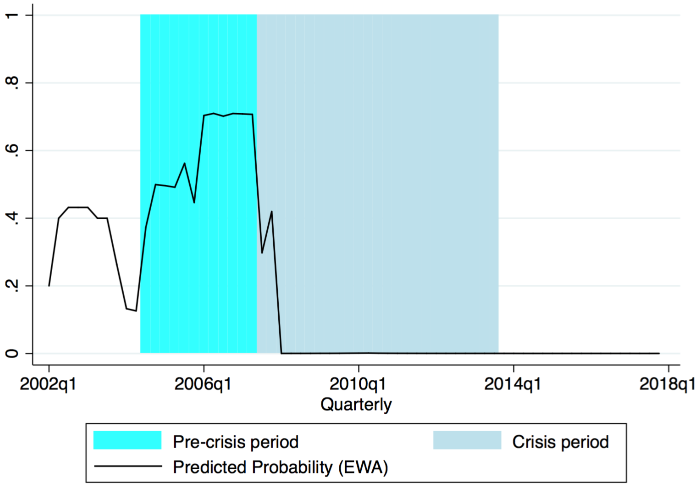


Figure 7: Predicted probability - EWA

Germany: pre-crisis period out-of-sample

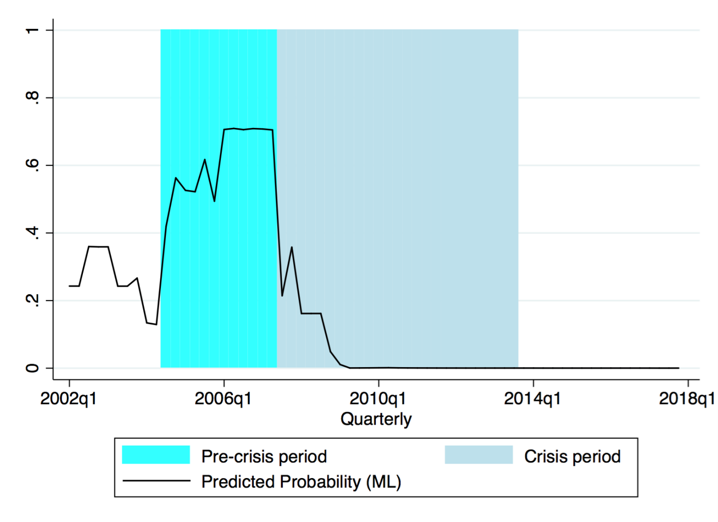


Figure 8: Predicted probability - ML

Germany: pre-crisis period out-of-sample

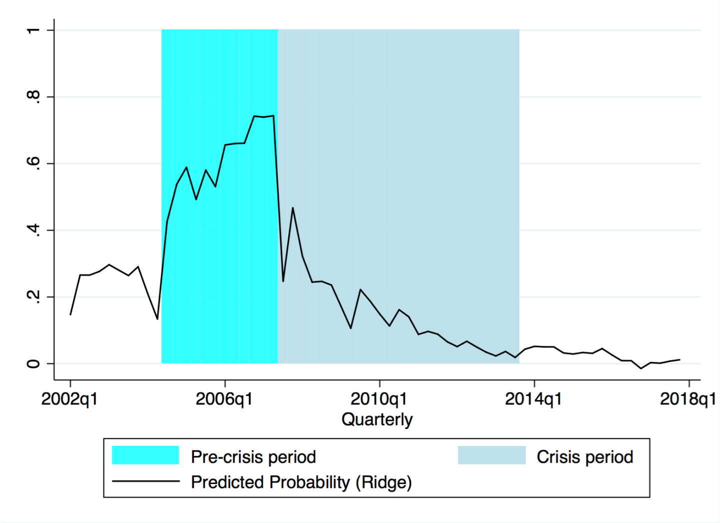


Figure 9: Predicted probability - Ridge

Germany

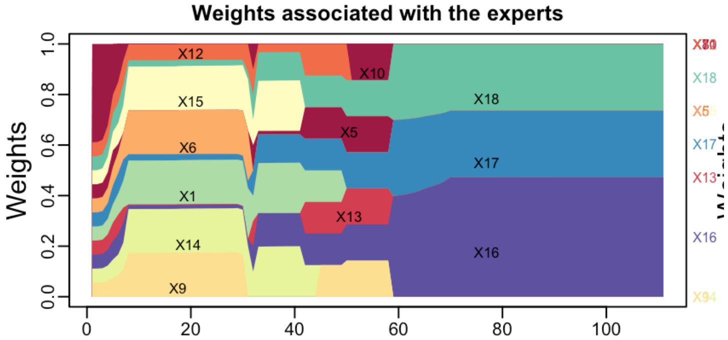


Figure 10: Weights - EWA

Root Mean Square Errors

Online aggregation Rule	RMSE
EWA	.0738
Uniform	0.288
Best expert (ex post)	0.00169
Best convex combination (ex post)	0.00168

Table 2: RMSE Germany

AUROC: Germany, out-of-sample

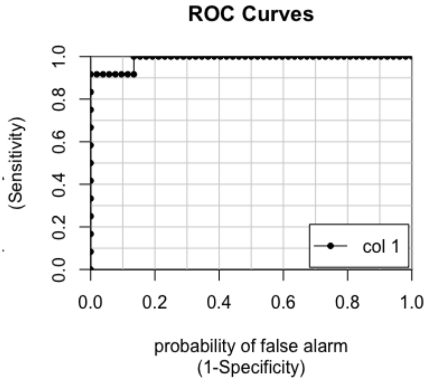


Figure 11: AUROC- EWA

$AUROC(EWA) = 0.99$

Italy: pre-crisis period out-of-sample

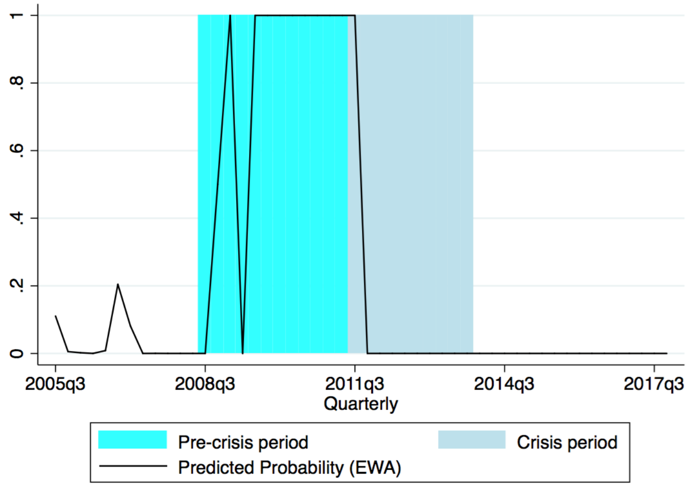


Figure 12: Predicted probability - EWA

Italy: pre-crisis period out-of-sample

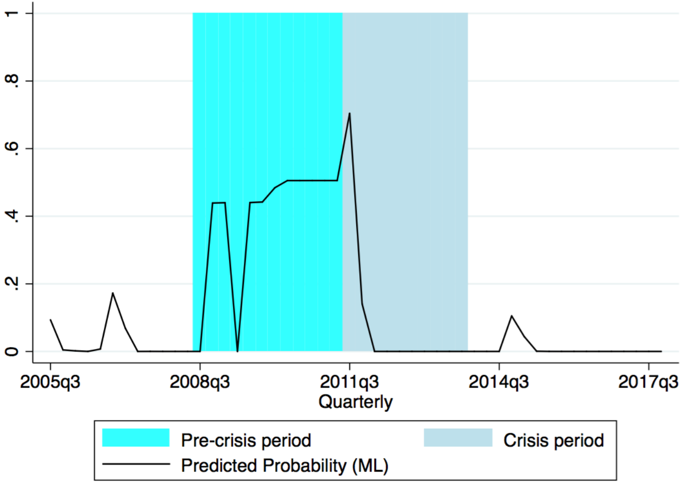


Figure 13: Predicted probability - ML

Italy: pre-crisis period out-of-sample

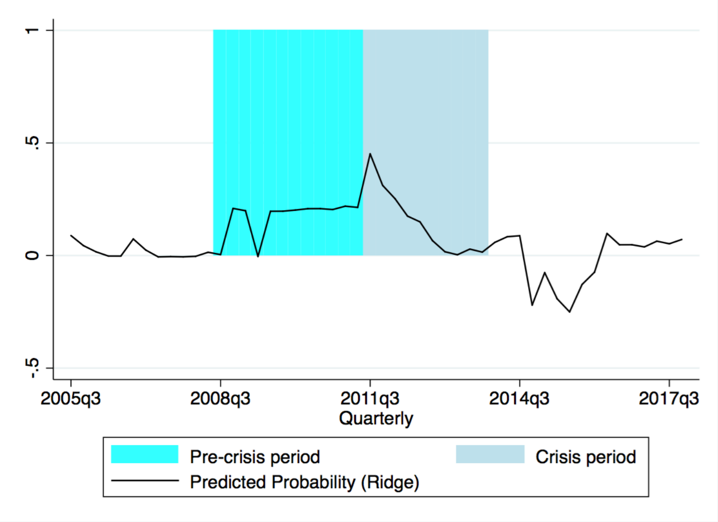


Figure 14: Predicted probability - Ridge

Spain: pre-crisis period out-of-sample

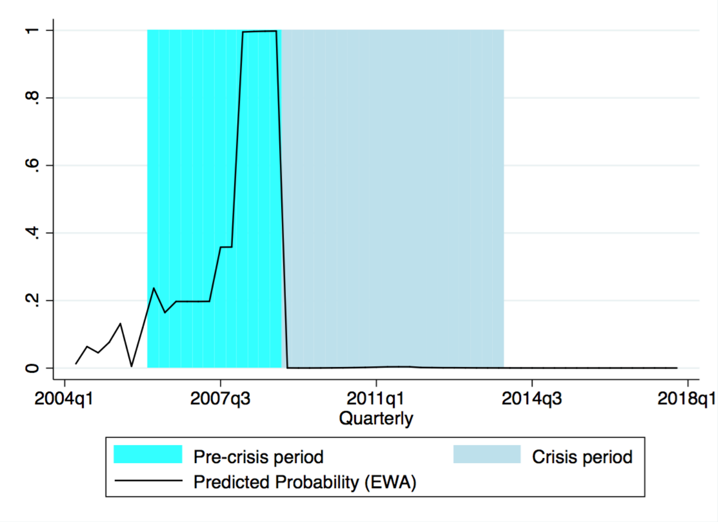


Figure 15: Predicted probability - EWA

Spain: pre-crisis period out-of-sample

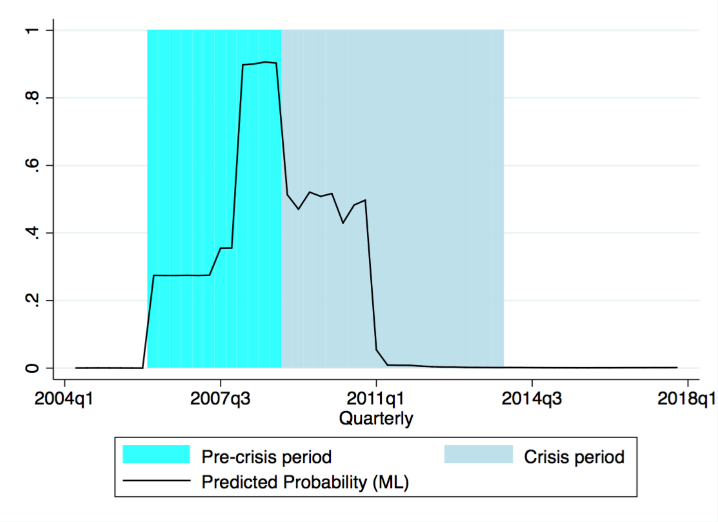


Figure 16: Predicted probability - ML

Spain: pre-crisis period out-of-sample

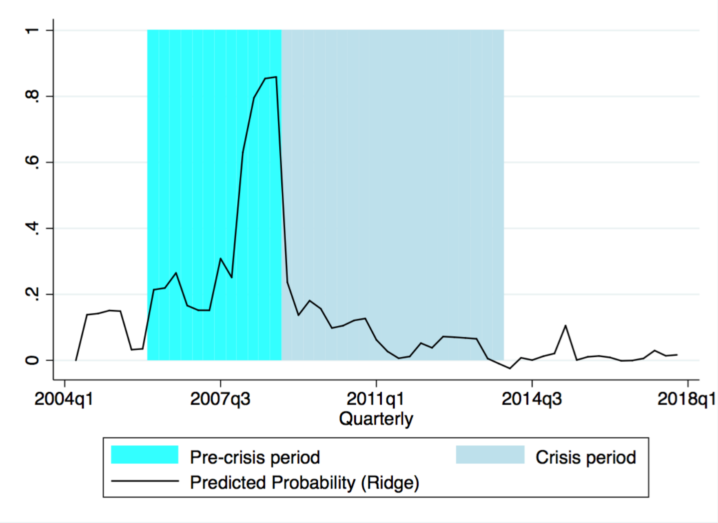


Figure 17: Predicted probability - Ridge

Sweden: pre-crisis period out-of-sample

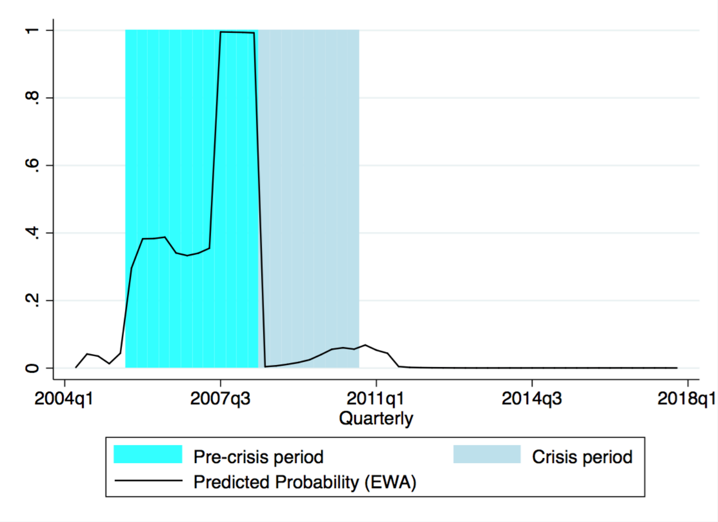


Figure 18: Predicted probability - EWA

UK: pre-crisis period out-of-sample

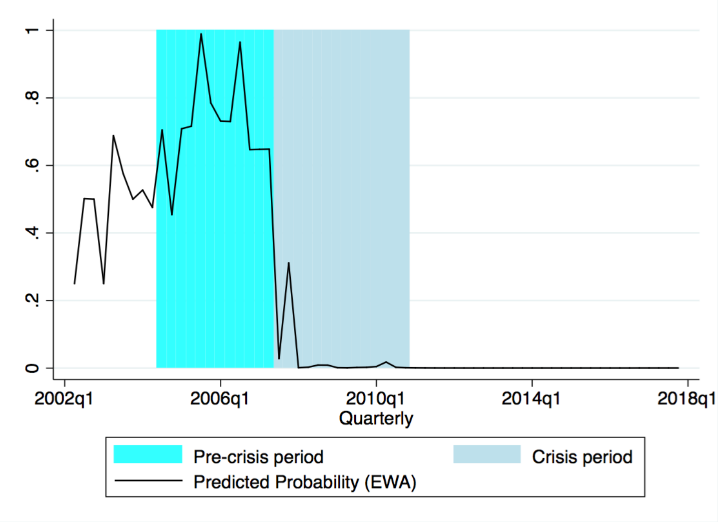


Figure 19: Predicted probability - EWA

Conclusions

- This approach gives strong out-of-sample forecasting results to predict financial crises. Also gives interesting information on relevant variables.

- We now work on the historical database as well (we predict the Great Depression out-of-sample). We can investigate which models are better for Great Depression versus Lehman Brothers.
- We can apply same framework to predicting recessions.
- We can construct millions of experts and investigate performance.
- Open questions :
  - Using more microeconomic data from bank databases?
  - Revisions and lags in the data? Approach is now quasi real time.
  - Causality?
  - How to test for the effect of macroprudential policies on crisis probabilities?