Available online at www.sciencedirect.com
ScienceDirect
The Journal of Finance and Data Science 5 (2019) 183–207
http://www.keaipublishing.com/en/journals/jfds/

Detection of rare events: A machine learning toolkit with an application to banking crises[☆]

Jérôme Coffinet [a,*], Jean-Noël Kien [a]

[a] *Banque de France, France*

Received 26 April 2019; revised 5 July 2019; accepted 14 April 2020
Available online 22 April 2020

Abstract

We propose a machine learning toolkit applied to the detection of rare events, namely banking crises. For this purpose, we consider a broad set of macroeconomic series (credit-to-GDP gap, house prices, stock prices, inflation rates, long-term and short-term interest rates, etc.), in combination with their leads and lags, various filtering methodologies, and datascience models that complement time series analysis. The main advantages of the approach are its robustness, its flexibility and its prediction performance. Based on the best model specification, our methodology allows to compute an indicator for the probability of banking crisis along with an alert threshold up to 6 quarters ahead in real time for various developed economies.

JEL classification: C52; C53; E58; G21

Keywords: Artificial neural networks; Random forest; Banking crises; Early warning system; Financial stability; Machine learning

1. Introduction

Understanding the roots of financial crises is a crucial stake for today's economics but predicting them in order to reduce their costs remains a key challenge for economic policy.[32] Following the global financial crisis, a wide bunch of economists has proposed various indicators and numerous methods aiming at a better identification of risks to financial stability.

Among them, the *credit-to-GDP* gap is currently widely used in the euro area and in developed countries as the early warning indicator of risks to financial stability. Precisely, its importance rests on its easy computation. As a matter of fact, many influential papers originating the Bank for International Settlements demonstrate its usefulness in determining whether or not a countercyclical capital buffer should be activated.[12–15] In that context, the Basel Committee on Banking Supervision provided a "*Guidance for National Authorities Operating the Countercyclical*

☆ The results presented in the paper do not necessarily coincide with the views of the Banque de France. All remaining errors remain our sole responsibility.

* Corresponding author.
E-mail addresses: jerome.coffinet@banque-france.fr (J. Coffinet), jean-noel.kien@banque-france.fr (J.-N. Kien).
Peer review under responsibility of China Science Publishing & Media Ltd.

https://doi.org/10.1016/j.jfds.2020.04.001

Capital Buffer",[5] where the computation of the credit-to-GDP was considered as a standard tool, computed under certain circumstances, for implementing the countercyclical capital buffer.

Like many Basel standards, its use is not binding but has to be transposed into national or regional law for that purpose. For instance, in the euro area, the European Systemic Risk Board (ESRB) recommendation of the 18 June 2014 adopted it on the basis of Articles 135 and 136 of Capital Requirements Directive IV. Hence, for instance in France, the *Haut Conseil de Stabilité Financière*, in charge of setting the countercyclical capital buffer, relies on the credit-to-GDP gap defined by the Basel Committee in order to determine when and what countercyclical capital buffer should be applied.

Nonetheless, what is more interesting, the ESRB regulation leaves open the possibility of using alternative methods (‘*When the designated authorities consider that another method of measuring and calculating the credit-to-GDP ratio gap better reflects the specificities of the national economy, it is recommended that they measure and calculate a quarterly difference in the ratio additional credit-GDP, in addition to the difference calculated in accordance with point 1.*‘). As a result, a different method of calculation can be used but the underlying economic variables of interest should remain.

Our paper rests on this open question and proposes a machine learning toolbox so as to anticipate banking crises as accurately as possible, with the aim to deliver unambiguous forward-looking and operationally-relevant signals to measure risks to financial stability. For that purpose, we examine the predictive power of various economic and financial indicators of banking crises. Hence, we provide a model of the probability of banking crisis along with an alert threshold up to 6 quarters ahead in real time and for 15 developed countries over the period 1970Q1-2017Q2.

From the economic point of view, our contribution is twofold: first, we test whether additional variables to the credit-to-GDP gap are likely to improve the detection of risks to financial stability,[25] and second we explicitly consider the possibility that the variables of interest are not synchronized, which improves our predictions. Hence, we show that our strategy enhances the standard methodology.

From a purely methodological point of view, our contribution is twofold. First, we introduce additional filter methodologies to the HP-filter retained in the standard credit-to-GDP methodology, in particular considering Kalman filters. Second, as many authors have decided to tackle this challenge adopting a wide variety of methodologies, from standard econometric models[4] to machine learning methods,[30] we also build a comprehensive toolkit for crisis identification that includes not only basic logit models, but also random forests and artificial neural networks. In fact, these machine learning methods proved promising results in several general public communications[31] and more specialized academic publications.[9,16]

All in all, we demonstrate that additional variables are useful complements to the usual credit-to-GDP gap in order to assess the risks to banking and financial stability. Based on the best model specification, our methodology allows to compute a reworked indicator with the ability to produce the probability of banking crisis along with an alert threshold up to 6 quarters ahead in real time for various developed economies. Using additional variables to the credit-to-GDP improves significantly the prediction of crisis for about two thirds of the countries. As a result, we end up with more reliable probabilities of banking crisis than the ones stemming from the credit-to-GDP gap taken alone. Our results are crucial to the operational implementation and policy-making decisions related to the countercyclical capital buffer.

The remainder of the paper is structured as follows. Section 2 discusses the Basel Committee methodology. Section 3 presents the data at hand and our datascience framework. Section 4 interprets the results. Section 5 concludes.

2. Designing a comprehensive methodology

2.1. The usual credit-to-GDP methodology

In Europe, a countercyclical buffer might be implemented by national authorities that compute and publish every quarter a countercyclical buffer rate of reference that matches the Basel Committee on Banking Supervision guidance. This guidance relies on the usual credit-to-GDP gap defined at each time t (quarter) as follows:

$$basel\_gap_t = \frac{credit_t}{gdp_t} - \tau_t$$

where τ_t is the trend of the credit-to-GDP ratio obtained with a recursive Hodrick-Prescott filter[19] with $\lambda = 400000$ for the smoothing parameter. With this definition, the $basel\_gap_t$ is also called the cycle component. Fig. 1 presents the trend and the usual credit-to-GDP obtained from this standard methodology for France.

Using this standard credit-to-GDP at hand, the Basel Committee provides an *ad hoc* rule[5] to determine when and how a countercyclical buffer should be activated. A buffer is activated when the usual credit-to-GDP gap reaches 2% point. Its rate grows linearly until the gap reaches 10% point (Fig. 2).

2.2. Some improvements to the usual method

The Basel methodology (hereafter 'benchmark' methodology) strongly relies on the credit-to-GDP gap and its computation. In fact, the countercyclical buffer is activated or deactivated based on a threshold and the rate of the buffer depends on the amplitude of the gap. Hence, the amplitude of the usual credit-to-GDP directly determines the countercyclical buffer.

2.2.1. The filtering step

The proposed calculation of the credit-to-GDP gap uses a very high smoothing parameter ($\lambda = 400000$). However, [19] suggest using a smoothing parameter $\lambda = 1600$ for quarterly time-series. This value is still a standard for the analysis of economic cycles in time series. The benchmark credit-to-GDP methodology breaks this principle, on the basis of a result from [27] stating that the smoothing parameter λ should be adjusted depending on the length of the cycles. Then, the proposed methodology includes the *a priori* hypothesis that credit cycles are from 3 to 4 times longer on average than economic cycles. The corresponding values for the smoothing parameter λ are from $3^4 * 1600 \sim 125000$ to $4^4 * 1600 \sim 400000$. The best λ is chosen such that the noise-to-signal ratio $\frac{\%\ type\ 2\ errors}{1-\%\ type\ 2\ errors}$ is minimal.[12] This suggests choosing a value of $\lambda = 400000$ to get the best results.

Nevertheless, some choices underlying the Basel gap can be challenged. The choice of the activation threshold (2%) seems to have been set *a priori*, thus the optimal λ is only valid for this level. A signal of 1 (0) is judged to be correct if a crisis (no crisis) occurs any time within a three-year horizon. Hence, a flexible horizon as in [7] is considered in this methodology.[12] This three-year horizon offers flexibility but it is large and may lead to errors of prediction. The noise-to-signal ratio and consequently the choice of λ can be biased by this flexibility. Finally, "while a simple moving average or a linear time trend could be used to establish the trend, the Hodrick-Prescott filter is used in this regime as it has the advantage that it tends to give higher weights to more recent observations. This is useful as such a feature is likely to be able to deal more effectively with structural breaks".[4] The type of filter used has not been challenged yet and choosing a Hodrick-Prescott filter might not be the best option[18] because, by construction, the Hodrick-Prescott filter produces distortions on borders which limit its use in real-time.

2.2.2. The predictive property

The predictive property of the standard credit-to-GDP gap is questionable. In fact, the way it is built makes it a tool which is more explicative than predictive:

- At time *t*, we have at hand the proposed credit-to-GDP gap at time *t* with the previous values. The only conclusion that can be made is whether or not there will be at least one crisis between *t* and *t+3 years* because λ has been chosen with a 3-year flexible horizon. This is too weak and approximate. Neither information concerning the number of crises, nor their probability of occurrence, nor a precise horizon of occurrence is provided;
- The standard credit-to-GDP gap is not reactive enough (see Fig. 1 (a)). In fact, it takes too much time for the trend to include structural changes like the one between 1994 and 1999. The increasing trend during this period is questionable. More generally, the filtered gap just presents a delay compared to the raw series;

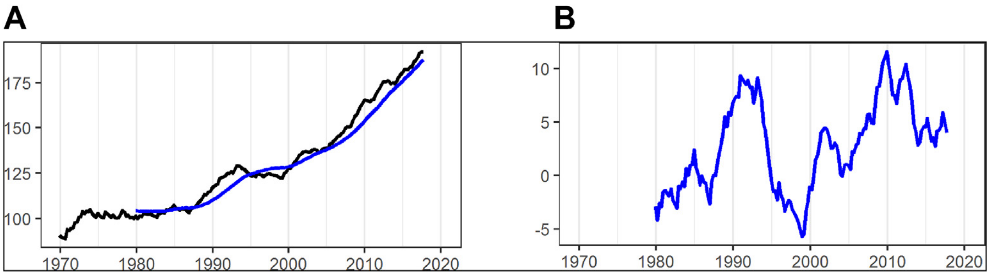


Fig. 1. (a) The credit-to-GDP ratio (in black) and the trend (in blue) computed by a recursive Hodrick-Prescott filter with a smoothing parameter for France; source BIS (cf. 3.1), authors' calculations (cf. 2.1). (b) The cycle component obtained for France with the standard methodology; source BIS (cf. 3.1), authors' calculations (cf. 2.1).

- The magnitude of the proposed credit-to-GDP gap directly determines whether or not a countercyclical capital buffer should be activated and the value of the buffer in percentage. However, depending on the threshold arbitrarily fixed, the methodology could have missed the internet bubble of 2001/2002 (see Fig. 1 (b) with a threshold of 5% point) or in other cases, could have activated an insufficient buffer;
- The methodology is not universal: all the parameters fixed and learnt (λ and the threshold of detection) are attached to the learning sample and there is no guarantee that these parameters will be identical for other countries or other sectors (private sector, public sector, etc.).

2.2.3. The desynchronization of credit and GDP series

The standard methodology considers the ratio of credit-to-GDP for a given economy during a given time span. Yet, there might be, depending of the structural features of the economy, a natural dynamic of series that induces a non-zero phase between credit and GDP series.[10] As a result, when the credit series is forward-looking with respect to the GDP series, the credit-to-GDP ratio estimated at time t_1 (cf. Fig. 3) is higher the one estimated at time t_2, independently from any policy measure.

If credit constantly lags behind GDP, the credit-to-GDP ratio increases and decreases accordingly without signaling any divergence of credit compared to GDP. Our prior is that the credit-to-GDP ratio at the peak of each series taken separately, that corresponds to the natural dynamics of the variables, is more relevant from a policy point of view than the ratio of contemporaneous numerator and denominator. From the example above, the intuition is quite simple: deciding the implementation of a countercyclical capital buffer based on the ratio observed at time t1 would be misleading because the natural dynamics of the series should yield an inferior ratio at time t2. The risk is hence to impend the financing of the economy while being inefficient as regards financial stability purposes, and finally to increase the risks of pro-cyclicality of the policy decision.

We have highlighted the rooms for improvement of the benchmark credit-to-GDP gap methodology. In order to tackle these issues, we will present our datascience framework in the next section.

3. Building our datascience framework

3.1. The data

A systemic banking crisis is when a major disruption in the financial systems occurs. It consists in country's corporate and financial sectors experiencing a large number of defaults and financial institutions and corporations facing great difficulties repaying contracts on time. As a result, non-performing loans increase sharply and a large part of the aggregate banking system capital is exhausted. This situation may be accompanied by depressed asset prices (such as equity and real estate prices) on the heels of run-ups before the crisis, sharp increases in real

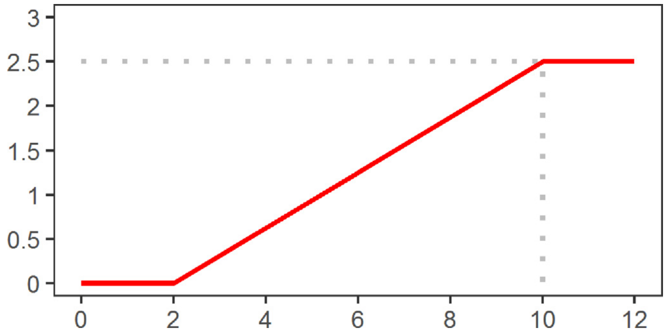


Fig. 2. Description of the buffer activation: when the usual credit-to-GDP gap reaches 2% point. Its rate grows linearly until the gap reaches 10% point Source: authors' computation.

interest rates, and a slowdown or reversal in capital flows. In some cases, the crisis is triggered by depositor runs on banks.

Building a database of crisis occurrences is a complex task because opinions of experts diverge. Many authors over years have been studying such databases.[21−23,28,29,20,8] The data we chose is a quarterly database of systemic banking crisis occurrences for 41 countries and at most from 1970Q1 to 2010Q4. This database[2] has been built with experts from central banks, international institutions and universities on top of previous works cited above, with the objective to identify *"banking crises (…) identified either according to a systemic loss of bank capital, or bank runs, or the size of public intervention in the banking sector"*. This benchmark has been used also in recent working papers such as [11] Over the most recent period (from 2010 to 2017), we add information on European countries from [24] and on other countries from.[3]

The database of explanatory variables is composed of the following series:

- The ratio of credit-to-GDP and the related credit and GDP series come from the database on macroprudential indicators from the Bank for International Settlements, related to IMF-IFS Claims on private sector for the credit series. Credit encompasses the outstanding amount of loans and bonds to the non-financial private sector, which is a broad definition of the private sector financing. Consistently with the Basel definition, the GDP series consists of the nominal gross domestic product. The series are quarterly from 1969Q4 to 2017Q2;
- House prices statistics are available at a quarterly frequency over the whole sample. They encompass for each country the most important house price index as identified by the Bank for International Settlements, which is also the source of the series and allows for some cross-country homogeneity of the variables in that respect. Data

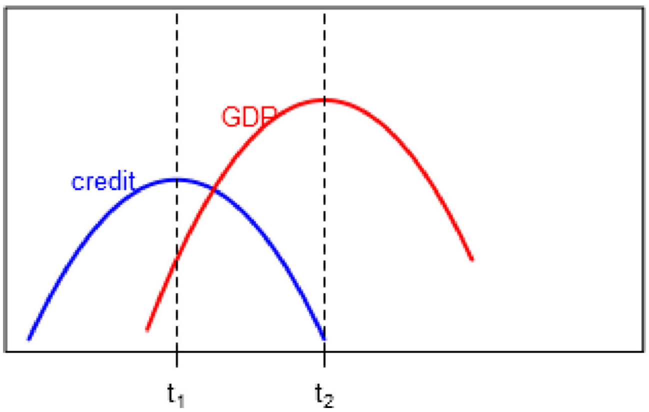


Fig. 3. Example of desynchronization between GDP and credit cycles.

gathered on the BIS website are complemented with national sources (national statistical institutes or national central banks) when needed;

- We also consider the following financial series, that are all taken from the International Financial Statistics database (IFS) of the IMF: the inflation rate is defined as the year-on-year growth rate of the consumer price index; the short-term interest rate is defined as the 3-month interbank money market rate when available or the overnight money market rate when the former is not available; the long-term interest rate is the 10-year government bond yield; share prices correspond to the domestic stock market index. Consistently with the prior that the latest available information is relevant for policy decisions, all daily financial variables are transformed so as to take the end-of-quarter observations in the final dataset.

Data are collected over 32 countries: Brazil, Canada, Chile, China, Czech Republic, Greece, Hungary, India, Indonesia, Korea, Malaysia, Mexico, New Zealand, Poland, Portugal, South Africa, Thailand, as well as Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom and Unites States. Table 1a provides an overview of the data by presenting the starting date, the ending date and the number of quarters under crisis for each country.

3.2. The algorithm

We build a datascience framework to predict the probability of banking crisis along with an alert threshold up to 6 quarters ahead in real time and for various developed economies. It is split in 5 steps:

1. *Importance of lags*: in this part, we present the need to include lags in the credit and the GDP when building the $\frac{credit_t}{gdp_t}$ ratio;
2. *Filtering*: in this part, we explore alternatives to the Hodrick-Prescott filter that could overcome its limitations;
3. *Additional variables*: we investigate whether adding variables improves the forecast of crises;
4. *Predictive model*: in this part, we present the models used to predict crisis occurrences;
5. *Validation*: this part is devoted to the validation of the models and how best models are selected.

3.3. Importance of lags

Based on the elements presented in Section 2.2, our proposal, in the paper framework, is to let the model optimize the best combination of leads and lags between credit and GDP, within the limit of 6 quarters, in order to predict banking crises.

3.4. Filtering

In our proposed methodology, the filtering step is essential. In fact, it enables to remove the trend in time series. The remaining part is what is used to detect financial bubbles. Filtering the original time series has also another advantage. It can be viewed as a way to normalize the underlying phenomenon by focusing only on its dynamics. Hence, a universal methodology across countries makes sense and can be implemented.

We also have qualitatively tested several filters, namely Beveridge-Nelson, Savitzky–Golay, Baxter–King, Christiano-Fitzgerald, Butterworth, trigonometric regression, cubic splines, LOESS regression and different specifications of dynamic linear models. As a result, we decided to choose the class of dynamic linear models. Indeed, this class of models is very flexible, allowing modeling every component of a time series (trend, cycle, seasonal components). Moreover, certain filters can be set up as a specific dynamic linear model (the Beveridge-Nelson filter for instance), as explained in more details in.[26]

In a nutshell, dynamic linear models are a class of models for time series, including multivariate time series. They are particularly popular among Bayesians. Associated with a dynamic linear model, the Kalman filter allows us to

Table 1a
basic statistic on the dataset.

Country	Beginning date	Ending date	Number of quarters under crisis
Austria	12/1/1960	9/1/2017	35
Belgium	12/1/1970	9/1/2017	35
Brazil	3/1/1996	9/1/2017	8
Canada	12/1/1955	9/1/2017	12
Chile	3/1/1983	9/1/2017	12
China	12/1/1985	9/1/2017	32
Czech Republic	3/1/1993	9/1/2017	41
Denmark	12/1/1966	9/1/2017	57
Finland	12/1/1970	9/1/2017	32
France	12/1/1969	9/1/2017	27
Germany	12/1/1960	9/1/2017	51
Greece	12/1/1970	9/1/2017	59
Hungary	12/1/1989	9/1/2017	31
India	6/1/1951	9/1/2017	24
Indonesia	3/1/1976	9/1/2017	32
Ireland	6/1/1971	9/1/2017	29
Italy	12/1/1960	9/1/2017	56
Japan	12/1/1964	9/1/2017	80
Korea	12/1/1962	9/1/2017	16
Luxembourg	3/1/1999	9/1/2017	12
Malaysia	6/1/1964	9/1/2017	36
Mexico	12/1/1980	9/1/2017	64
Netherlands	3/1/1961	9/1/2017	46
New Zealand	12/1/1960	9/1/2017	24
Norway	12/1/1960	9/1/2017	56
Poland	3/1/1992	9/1/2017	34
Portugal	12/1/1960	9/1/2017	41
Russia	6/1/1995	9/1/2017	12
Singapore	12/1/1970	9/1/2017	8
South_africa	3/1/1965	9/1/2017	12
Spain	3/1/1970	9/1/2017	65
Sweden	3/1/1961	9/1/2017	67
Switzerland	12/1/1960	9/1/2017	32
Thailand	12/1/1970	9/1/2017	20
Turkey	3/1/1986	9/1/2017	16
United Kingdom	3/1/1963	9/1/2017	50
United States	3/1/1952	9/1/2017	60

update our beliefs about the current value of the unobserved vector (the trend) each time we incorporate a new observation.

Finally, after testing qualitatively different types of dynamic linear models, we have decided to extract the trend τ_t from the credit-to-GDP ratio by modeling the phenomenon with a dynamic linear model named *local level model* or *random walk + noise model*. This model is defined as the following state-space model:

$$\frac{credit_t}{gdp_t} = \tau_t + v_t, \;\; v_t \sim N(0,\, V_t)$$

$$\tau_t = \tau_{t-1} + w_t, \;\; w_t \sim N(0,\, W_t)$$

where v_t is the noise of the observations and w_t the noise of the model. In fact, at each time t, it is not the true trend that is observed but a noisy version of it. The unobserved true trend is modeled as a random walk with a noise. This noise w_t is called the noise of the system or the noise of the model. In our specification, the noise of the model should be low because only the trend is modeled.

Table 1b
Basic statistics on the retained dataset.

Country	Training period	Test period	Number of quarter under crisis during the training period	Number of quarter under crisis during the test period
Belgium	6/1/1977 to 3/1/2005	3/1/2005 to 6/1/2017	14	21
Denmark	9/1/1989 to 3/1/2005	3/1/2005 to 6/1/2017	23	24
Finland	9/1/1993 to 3/1/2005	3/1/2005 to 6/1/2017	14	8
France	6/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	19	8
Germany	9/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	24	24
Ireland	9/1/1978 to 3/1/2005	3/1/2005 to 6/1/2017	1	28
Italy	9/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	32	24
Japan	6/1/1972 to 3/1/2005	3/1/2005 to 6/1/2017	57	24
Netherlands	9/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	25	21
Norway	9/1/1976 to 3/1/2005	3/1/2005 to 3/1/2017	36	6
Spain	12/1/1983 to 3/1/2005	3/1/2005 to 6/1/2017	14	24
Sweden	9/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	49	12
Switzerland	6/1/1981 to 3/1/2005	3/1/2005 to 6/1/2017	24	8
United Kingdom	9/1/1977 to 3/1/2005	3/1/2005 to 6/1/2017	24	13
United States	9/1/1976 to 3/1/2005	3/1/2005 to 6/1/2017	56	4

In our case, we will rather estimate the signal to noise ratio $\frac{W}{V}$ rather than W and V separately. This ratio will be optimized in the estimation of the models. To avoid being stuck in a local minima like this is the case in our case with a quasi-Newton method, the optimization of the ratio is done by estimating models with the ratio picked from a grid of values (from 400 to 6000 by 200). Then, the value giving the best quality of prediction is selected.

As a result, we will use two filters: the recursive Hodrick-Prescott filter as a benchmark for our proposed credit-to-GDP gap methodology and the local level model. We will also refer to the local level model as *the Kalman filter model* by reference to the algorithm it uses. When using additional variables, a year-on-year variation is computed to the house prices and the share prices time series.

3.5. Wrapping up in predictive models

We aim to model the banking crises in order to predict them. Usually, when the phenomenon is binary, the natural choice is to use a logistic model. We also have decided to use a neural network considering the possibilities to model complex and non-linear interactions between variables and robustness when structural changes occur.

The neural network used is a so-called “multilayer perceptron” from the class of feedforward artificial neural networks. We test 2 configurations: the first configuration is a neural network with 1 hidden layer with 2 neurons and the second is a neural network with 2 hidden layers with sequentially 4 and 2 neurons. All neurons are built with a logistic activation function. This choice enables stable estimations with convergence of optimization steps at stake. The configuration with the best predictions is chosen.

We also consider random forest methods that prove to be less sensitive to overfitting risks and hence especially well adapted to smaller samples such as the ones that are considered in the paper. The Random Forest algorithm used is based on the Breiman's random forest algorithm with 500 trees, a sampling with replacement and a minimum size of the terminal nodes set to 1.

For the variables included in the models, we have chosen 4 possibilities:

- Include only one variable from the set of ratios $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$. As we noticed in Section 3.3, adding lags in both credit and GDP should be considered;
- Include only one variable from the set of ratios $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ with all the additional variables;

- Include all combinations of the $\frac{credit}{gdp}$ simultaneously. This means we include all variables from the set $\left\{ \frac{credit_{t-k}}{gdp_{t-l}},\ k,l\ \in [1,6] \right\}$ The purpose here is not to include a single value but sequences of values that represent the piecewise dynamics;
- Include all combinations of the $\frac{credit}{gdp}$ simultaneously with all the additional variables.

The models have to give the probability of banking crisis up to 6 quarters ahead in real time. This forces the models to have the following form. Assuming we want to predict the probability of banking crisis in h quarters, then the model is:

$$prob_t = f(X_{t-h})$$

where $prob_t$ is the probability of banking crisis at time t and X_{t-h} is all information set available up to time t-h. Once estimated, the predicted probabilities as follows read:

$$\widehat{prob}_t = \widehat{f}(X_{t-h})$$

where $\widehat{prob}_t$ is the probability of banking crisis at time t and $\widehat{f}$ the model estimated. Then, the predicted probability of banking crisis in h quarters from time t can be obtained by shifting the model as follows:

$$\widehat{prob}_{t+h} = \widehat{f}(X_t)$$

As a result, when we want to predict the probability of banking crisis in h quarters, the estimation of such a model consists in learning the crisis at time t knowing the predictors up to time t-h. In fact, the predictors in the interval *]t-h,t]* has to be ignored because they will become unknown once the model is shifted to the interval *]t, t + h]*.As a consequence, our forecast are not recursively computed.

For instance, suppose we have estimated the following model:

$$\widehat{prob}_t = \widehat{f}\left(\frac{credit_{t-4}}{gdp_{t-3}}\right)$$

then, at time t, we shift the model and we can predict the probability of banking crisis in at most $h = \min(|4|,|3|) = 3$ quarters.

We have recapitulated in Table 2 the possible models in competition. The best model will be determined by the quality of prediction on a test sample (see Section 3.6):

3.6. Validation

At this step, we have a group of models in competition to predict the probability of banking crisis for each horizon h. In order to compare them and determine which model is the best, we defined a test protocol that consists in learning of all classes of models on a period ending in 2004Q4, and using the 2005Q1-2017Q2 period as a test sample to validate the models.

Our quantitative assessment of the model performance is based on the ‘Area Under the Curve’ criterion of the ‘ROC (Receiver Operating Characteristic) Curve’. More specifically, Area Under the Curve is a criterion for evaluating classification models by reporting “true positive” signals (the proportion of episodes actually detected) to “false positives” (proportion of episodes without crises identified as episodes of crisis). When the area under the curve (cf. Fig. 4) approaches 1 (green curve), the model is discriminating and makes it possible to identify the periods of crises and non-crises. Conversely, when the area approaches 0.5, the model does not distinguish between true and false signals and corresponds to a ‘random’ detection of episodes of crises. Conditionally to the choices of modeling, the optimization of the probability of detection and the probability of false alarm equally weighted in the out-of-sample

Table 2
Overview of various models considered in the estimation.

(M0) Basel methodology benchmark	- Hodrick-Prescott filter - Logistic model - Single ratio among the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-k}}, \mathrm{k} \in [1,6]\right\}$
(M1) Univariate model with HP filter	- Hodrick-Prescott filter - Logistic model - Any single ratio among the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$
(M2) Univariate model with LLM filter	- Local level model - Logistic model - Any single ratio among the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$
(M3) Univariate model with HP filter with additional variables	- Hodrick-Prescott filter - Logistic model - Any single ratio among the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$ - Additional variables
(M4) Univariate model with LLM filter with additional variables	- Local level model - Logistic model - Any single ratio among the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$ - Additional variables
(M5) Multivariate model with HP filter	- Hodrick-Prescott filter - Logistic model - All ratio from the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$
(M6) Multivariate model with LLM filter	- Local level model - Logistic model - All ratio from the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$
(M7) Multivariate model with HP filter with additional variables	- Hodrick-Prescott filter - Logistic model - All ratio from the set $\left\{\frac{\mathrm{credit}_{t-k}}{\mathrm{gdp}_{t-l}}, \mathrm{k,l} \in [1,6]\right\}$ - Additional variables
(M8) Multivariate model with LLM filter with additional variables	- Local level model - Logisti- Hodrick-Prescott filterc model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ - Additional variables
(M9) Neural network model with HP filter	- Hodrick-Prescott filter - Neural Network model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$
(M10) Neural network model with LLM filter	- Local level model - Neural Network model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$
(M11) Neural network model with HP filter with additional variables	- Hodrick-Prescott filter - Neural Network model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ - Additional variables
(M12) Neural network model with LLM filter with additional variables	- Local level model - Neural Network model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ - Additional variables- Hodrick-Prescott filter
(M13) Random Forest model with HP filter	- Hodrick-Prescott filter - Random Forest model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$
(M14) Random Forest model with LLM filter	- Local level model - Random Forest model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$
(M15) Random Forest model with HP filter with additional variables	- Hodrick-Prescott filter - Random Forest model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ - Additional variables
(M16) Random Forest model with LLM filter with additional variables	- Local level model - Random Forest model - All ratio from the set $\left\{\frac{credit_{t-k}}{gdp_{t-l}},\ k,l \in [1,6]\right\}$ - Additional variables

Table 3
regression table for the AUC estimated over different specifications.

Explained variable: AUC (Nb obs: 510)			
Variable	Coefficient	*p*-value	Significance
Intercept	0,6	0	***
Countries			
Belgium	ref		
Denmark	−0,04	0,45	
Finland	−0,05	0,37	
France	−0,02	0,66	
Germany	−0,11	0,04	*
Ireland	−0,09	0,11	
Italy	0,09	0,1	.
Japan	0,04	0,5	
Netherlands	−0,01	0,82	
Norway	0,06	0,26	
Spain	0,04	0,49	
Sweden	−0,04	0,5	
Switzerland	−0,19	0	***
United Kingdom	0,06	0,29	
United States	−0,13	0,02	*
Country-specific model			
Country-wide model	ref		
Country-specific model	−0,06	0,01	**
Filters			
HP	ref		
Kalman	−0,19	0	***
Models			
Model benchmark individual	ref		
Model benchmark global	−0,1	0,25	
Model LOGIT_MULTIVARIATE	−0,09	0,14	
Model LOGIT_UNIVARIATE	−0,02	0,73	
Model Neural Network	−0,11	0,09	.
Model Random Forest	−0,14	0,03	*
Additional variables			
No additional variable	ref		
Additional variables	0,13	0	***

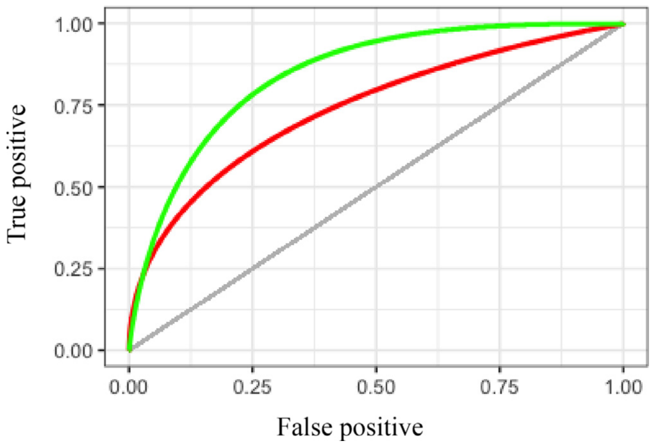


Fig. 4. ROC curve and area under the curve.

prediction exercise makes it possible to determine the thresholds from which the probability of crisis can be interpreted.

In so doing, we gauge the capability of the different procedures to predict the last financial crisis, which acts as an ‘experimental proof’. Nonetheless, we reduce significantly our sample of interest as our specification imposes that, for each country, we can observe at least one crisis in the learning sample (before 2005) and at least one crisis in the test sample (after 2005). As long as we refer to the crisis identification explained in Section 3.1, we are compelled to retain only 15 countries presenting those characteristics: Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom and Unites States. Table 1b provides an overview of the data retained by presenting the learning period, the test period, the number of quarters under crisis during the learning period and the number of quarters under crisis during the test period for each country.

Finally, although our framework is flexible because it allows selecting the best model for one country and one horizon of prediction, we decided to force the model to be the same for all horizons. This guarantees some stability between predictions along different horizon. The flexibility of the model across countries is still conserved.

4. Results

4.1. Overall results

Fig. 5 presents the distribution of obtained AUC over different classes of models, with reference to the benchmark model, which is the strict credit-to-GDP BIS-recommended model, with simultaneous credit and GDP observations.

At the first glance, we obtain the following results:

- On average, the best class of models is the one including a unique credit-to-GDP ratio with a Hodrick-Prescott filter, but with desynchronized credit and GDP as the AUC distribution is significantly over 0. Moreover, the distribution of AUC is quite small, which tends to indicate that the proposed method tends to maximize model homogeneity across countries. Hence, our assessment tends to confirm in general the relevance of the Basel methodology, but underlines the importance of lags and leads in the dynamics of credit and GDP;
- Using multiple ratios does not provide any significant gain on the AUC (Table 3) distribution, while using neural networks seems quite interesting as it scarcely over performs the unique ratio models but is likely to improve the model performance at the top of the distribution: as a result, neural networks appear as a rare but not marginal gain on the prediction performance. The same feature holds true for random forests, all the more than they are used with additional variables.

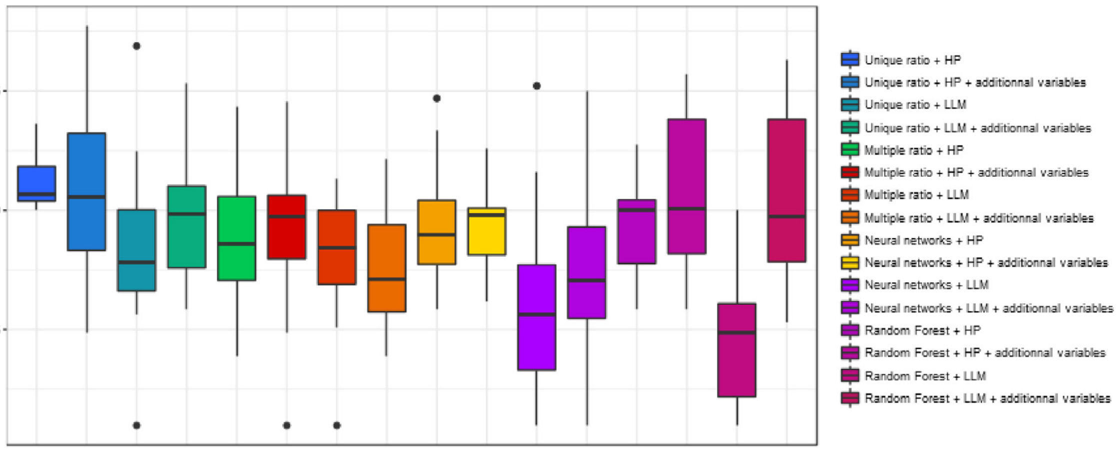


Fig. 5. distribution of the obtained AUC over different classes of models.

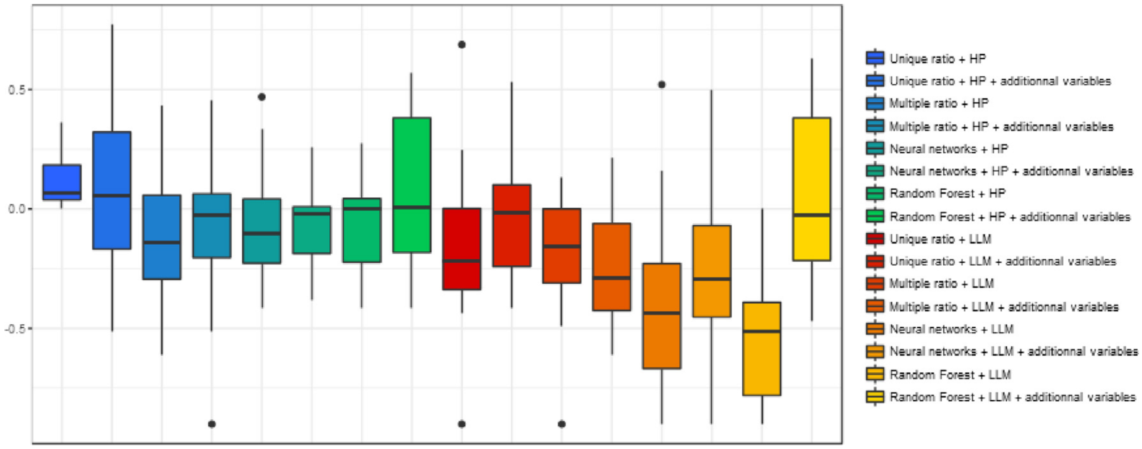


Fig. 6. distribution of the obtained AUC over different filters.

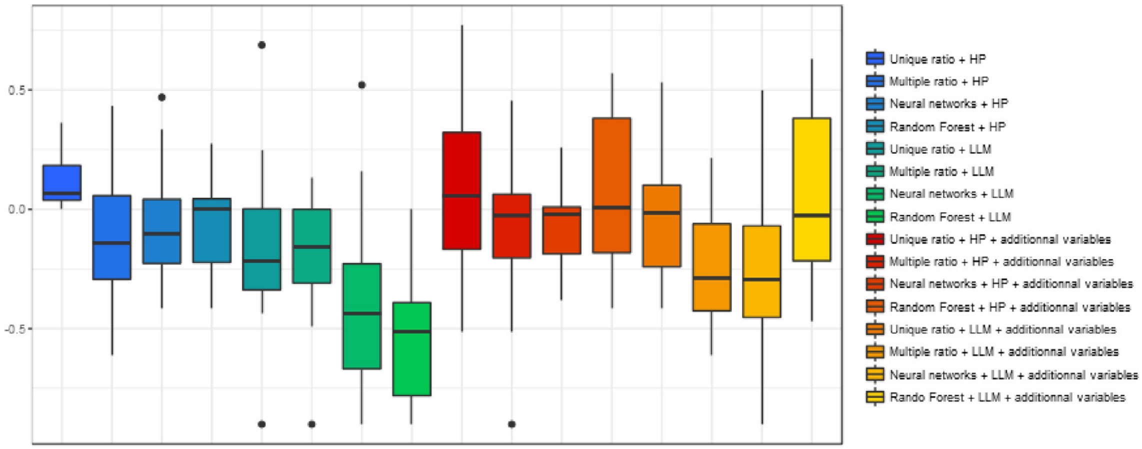


Fig. 7. distribution of the obtained AUC with additional variables.

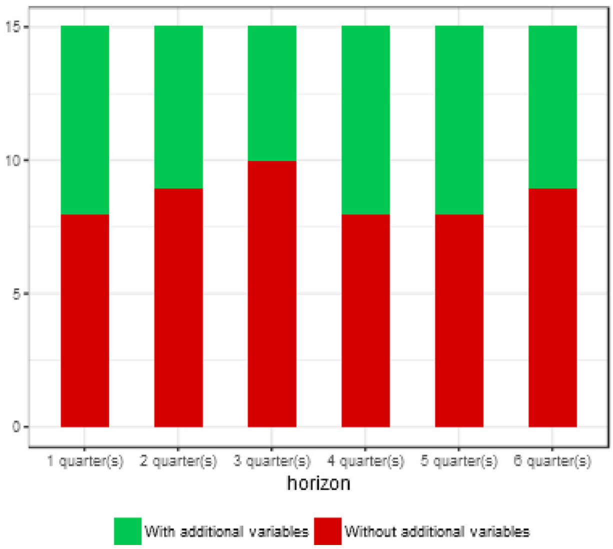


Fig. 8. share of best models by horizon including additional variables.

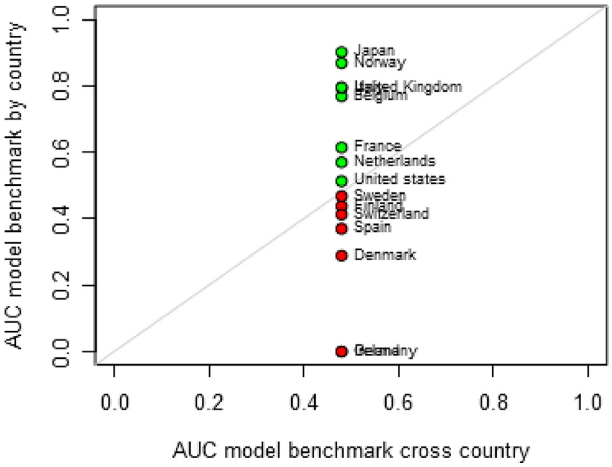


Fig. 9. Model comparison: country-specific vs. cross-country, all benchmark models.

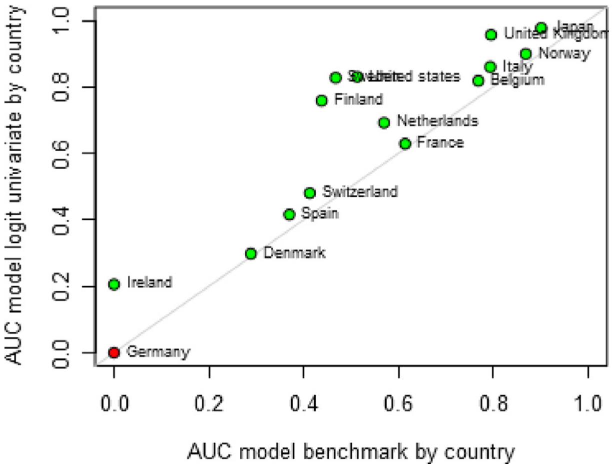


Fig. 10. Model comparison: logit univariate vs. benchmark, all country-specific models.

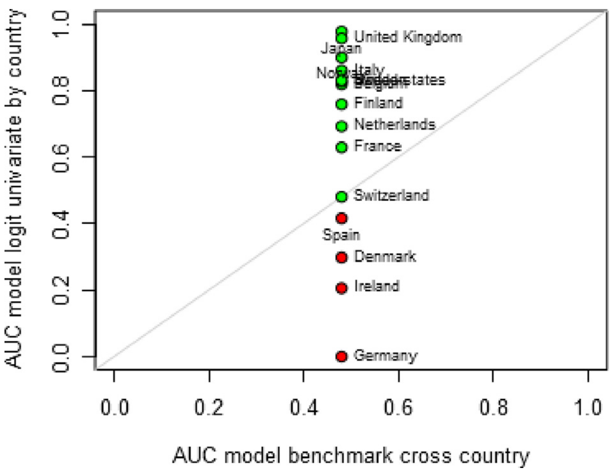


Fig. 11. Model comparison: logit univariate country-specific vs. benchmark cross-country.

Figs. 6 and 7 present the same distributions but re-ordered in order to analyze the effects of filtering choices and additional variables.

Fig. 6 shows that the Kalman filter only slightly improves the average prediction when it is used with multiple ratios without additional variables.

Fig. 7 indicates that introducing additional variables is often beneficial for the different configurations, but the one using a unique ratio with HP filter. In particular, it improves the top of the distribution for some of model classes, especially the ones based on neural networks and even more random forests.

Taking the best models without continuity constraint, 42% retain additional variables from all horizons. Fig. 8 below suggests that, however, there does not seem to have a monotonous relationship between the horizon and the added value of including additional variables.

As it is difficult to isolate simply the marginal gain of using one specification rather than another one, and the contribution of each methodological innovation, we carry out a variance analysis (ANOVA) on our distribution of computed AUC.

Switzerland, Germany and the United States prove atypical countries with a significant negative effect on the quality of forecasting, while banking crises in Italy prove to be better forecasted on average than in other countries. All things being equal, on average, estimating a model specifically for a country decreases the quality of forecasting. This means that the quality of learning is improved by a set of countries that provides useful information for the prediction of crisis for others. The filtering is significant, and all else being equal, the Kalman filter decreases the quality of prediction, on average. On average, using neural network and random forest classes of models deteriorate the goodness-of-fit of the forecast. Finally, the effect of the additional variables is significant with a positive effect on the forecast quality.

4.2. Which model to use for each country?

Fig. 9 presents the AUCs by country obtained on the one hand with the benchmark model without distinction of country and on the other hand with benchmark models estimated by country. It is interesting to note that a country model offers better forecast quality for Japan, Norway, United Kingdom, Belgium, France, Netherlands and the United States. On the other hand, a country can also benefit from the global information provided by other countries, such as the Spain, Switzerland, Italy, Denmark, Ireland and above all Germany.

Fig. 10 presents the AUCs by country obtained on the one hand with benchmark models estimated by country and on the other hand with univariate logistic models by country. Each univariate logistic model is composed of the ratio

Table 4
Best model by country.

Country	Country-specific model	Filter	Model	Additional variables	AUC	Optimal threshold
Belgium	YES	HP	LOGIT_UNIVARIATE	NO	0.82	0.11
Denmark	YES	LLM	RANDOM_FOREST	YES	0.85	0.09
Finland	YES	HP	LOGIT_UNIVARIATE	NO	0.76	0.32
France	NO	HP	LOGIT_UNIVARIATE	YES	0.70	0.24
Germany	NO	HP	LOGIT_UNIVARIATE	YES	0.70	0.24
Ireland	NO	HP	LOGIT_MULTIVARIATE	YES	0.76	0.20
Italy	YES	HP	RANDOM_FOREST	NO	0.91	0.16
Japan	YES	HP	LOGIT_UNIVARIATE	NO	0.98	0.42
Netherlands	NO	HP	LOGIT_UNIVARIATE	YES	0.72	0.24
Norway	YES	HP	LOGIT_UNIVARIATE	YES	0.93	0.06
Spain	NO	HP	LOGIT_MULTIVARIATE	YES	0.87	0.23
Sweden	YES	HP	LOGIT_UNIVARIATE	NO	0.83	0.43
Switzerland	NO	HP	LOGIT_UNIVARIATE	YES	0.70	0.23
United Kingdom	YES	HP	LOGIT_UNIVARIATE	NO	0.96	0.20
United States	YES	HP	LOGIT_UNIVARIATE	NO	0.83	0.49

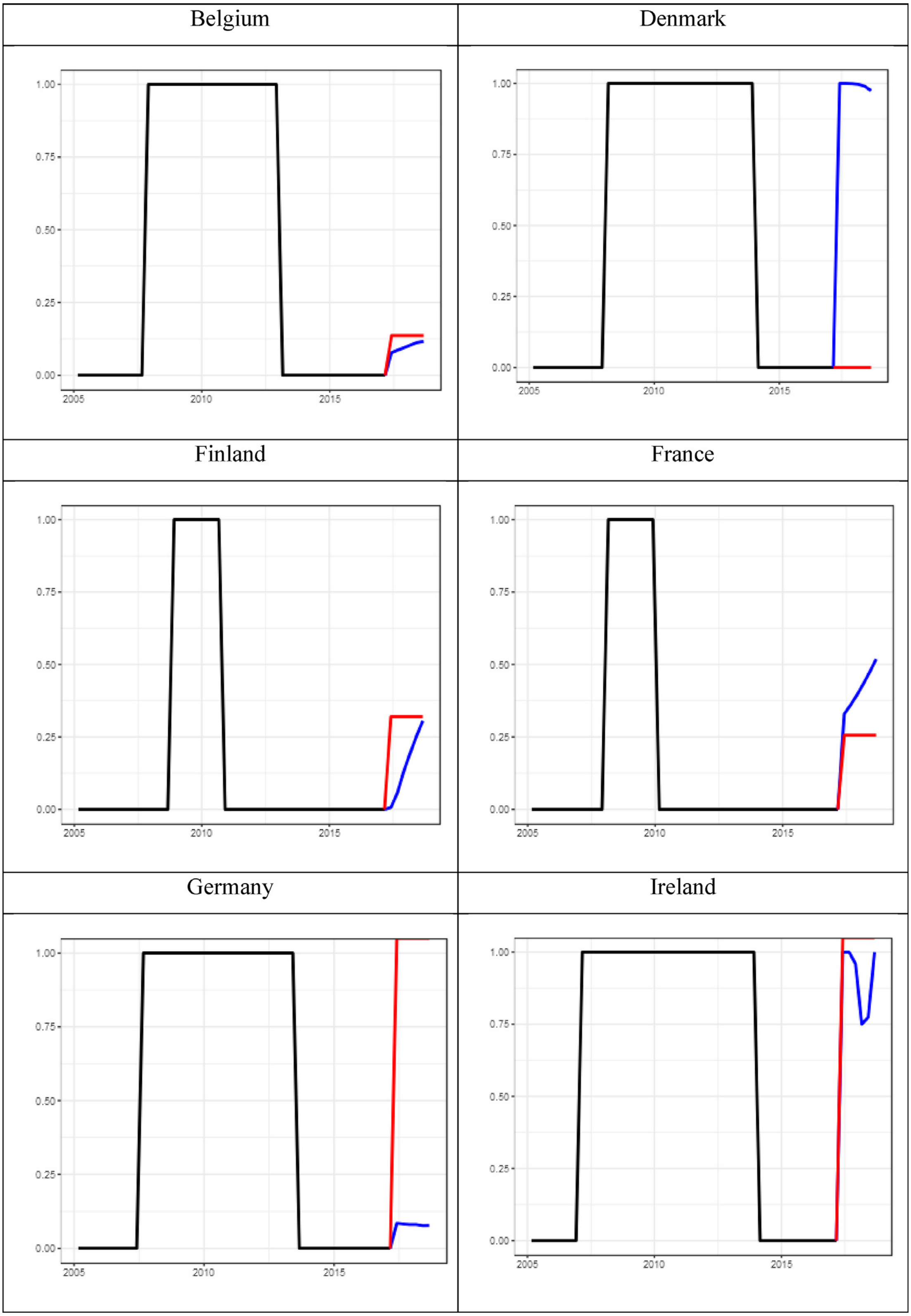


Fig. 12. crisis predictions – benchmark models.

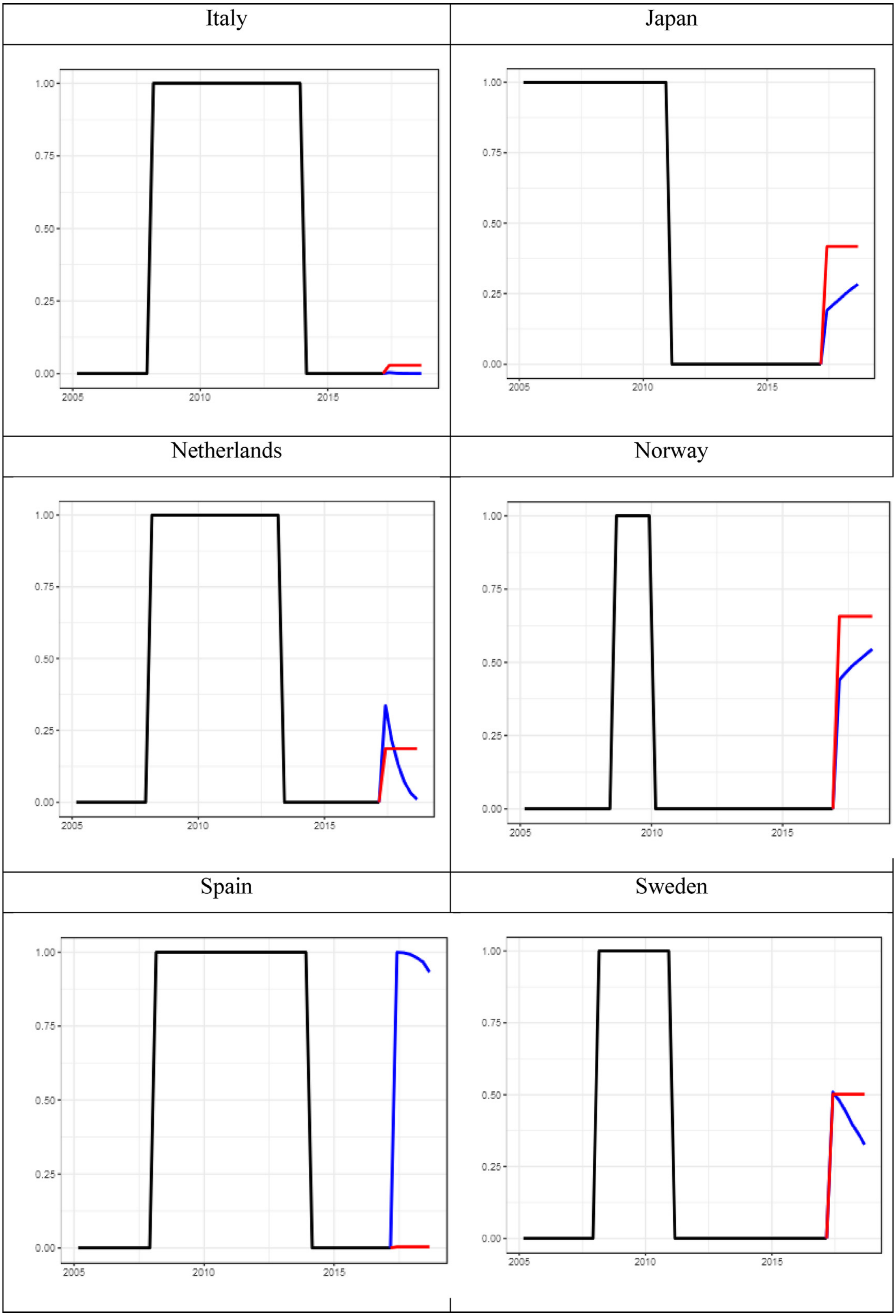


Fig. 12. (*continued*).

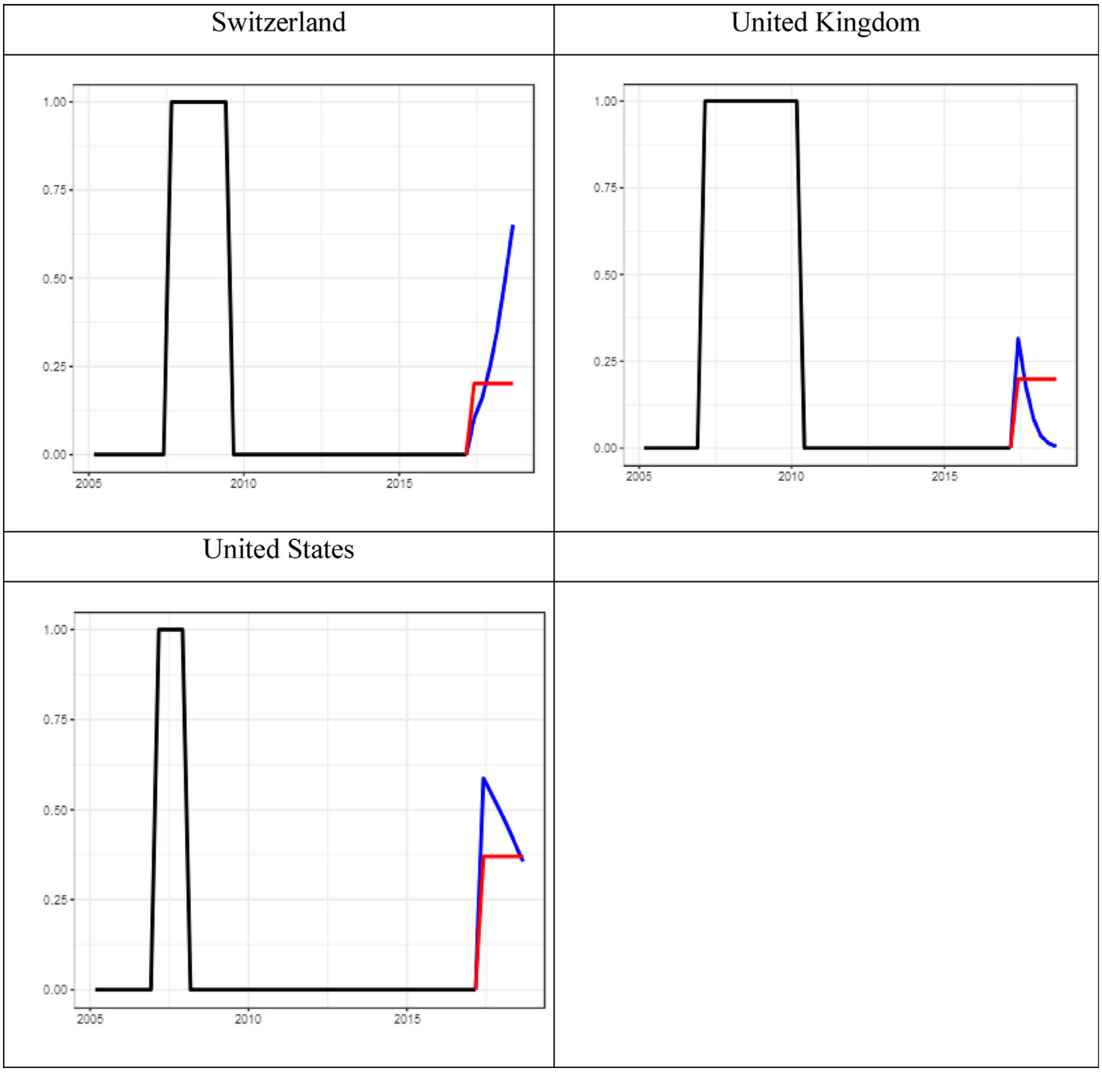


Fig. 12. (*continued*).

lagged credit and lagged GDP by retaining only the most optimal lag combination by country. For all countries but Germany, univariate logistic models provide better forecasting quality.

By comparing (Fig. 11) the performance of univariate logistic models by country with the benchmark model without distinction of countries, we succeed to get a better forecasting performance for Sweden for instance, which was not possible with the benchmark model by country. Other countries (Spain, Denmark, Ireland and Germany) continue to benefit from the global information provided by other countries. The forecast quality for these same countries can be improved but it will then be a more sought-after model.

Here in Table 4 are the models we obtain country by country when imposing the same model by country whatever the horizon:

With additional variables being included in the best model for 8 countries out of 15, our results are in line with the results of [6] and [1] In fact, the later show that housing prices, stock prices, inflation and interest rates improve the predicting power of credit-to-GDP. In that respect, [17] explain that booms increase the probability of a crisis only in relatively fragile financial systems. Our results also prove that including machine learning processes might be relevant for some countries, such as Denmark and Italy, with random forest over-performing other classes of models. Kalman

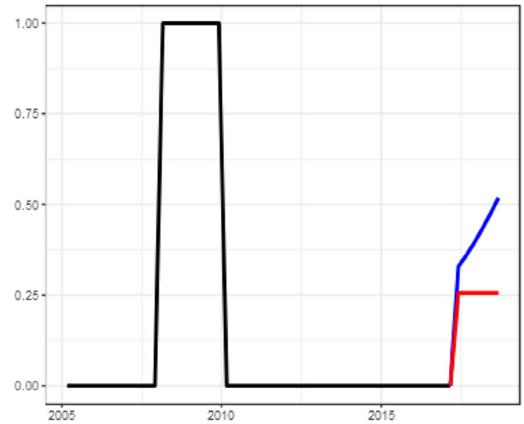
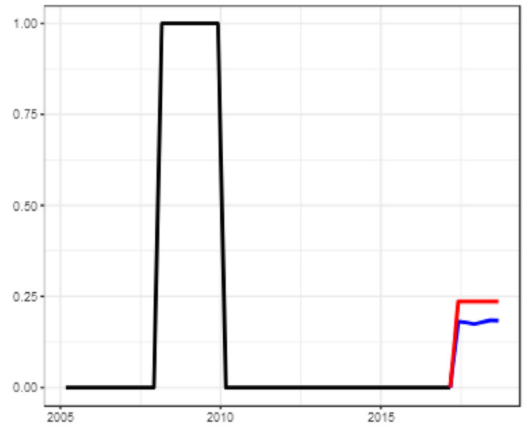
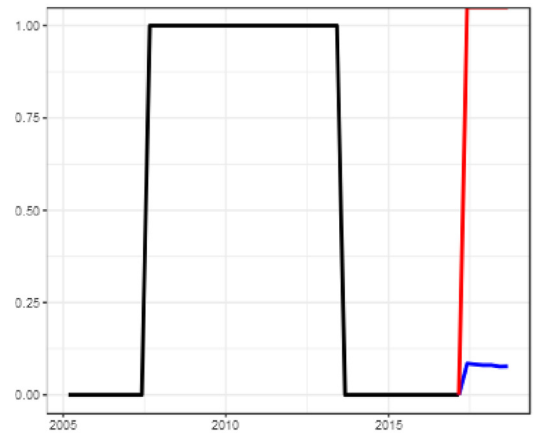
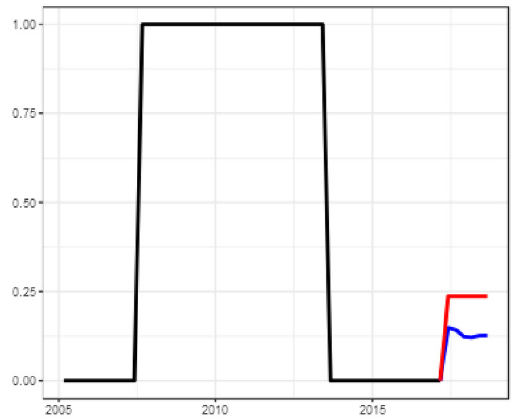
<table>
<tr><th colspan="2">France</th></tr>
<tr><td>Hendrick-Prescott filter
Logistic model
$\frac{credit_{t-k}}{gdp_{t-k}}$</td><td>Best model
Cf. Table 3</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><th colspan="2">Germany</th></tr>
<tr><td>Hendrick-Prescott filter
Logistic model
$\frac{credit_{t-k}}{gdp_{t-k}}$</td><td>Best model
Cf. Table 3</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Fig. 13. Crisis probability in France and Germany: benchmark vs. best model.

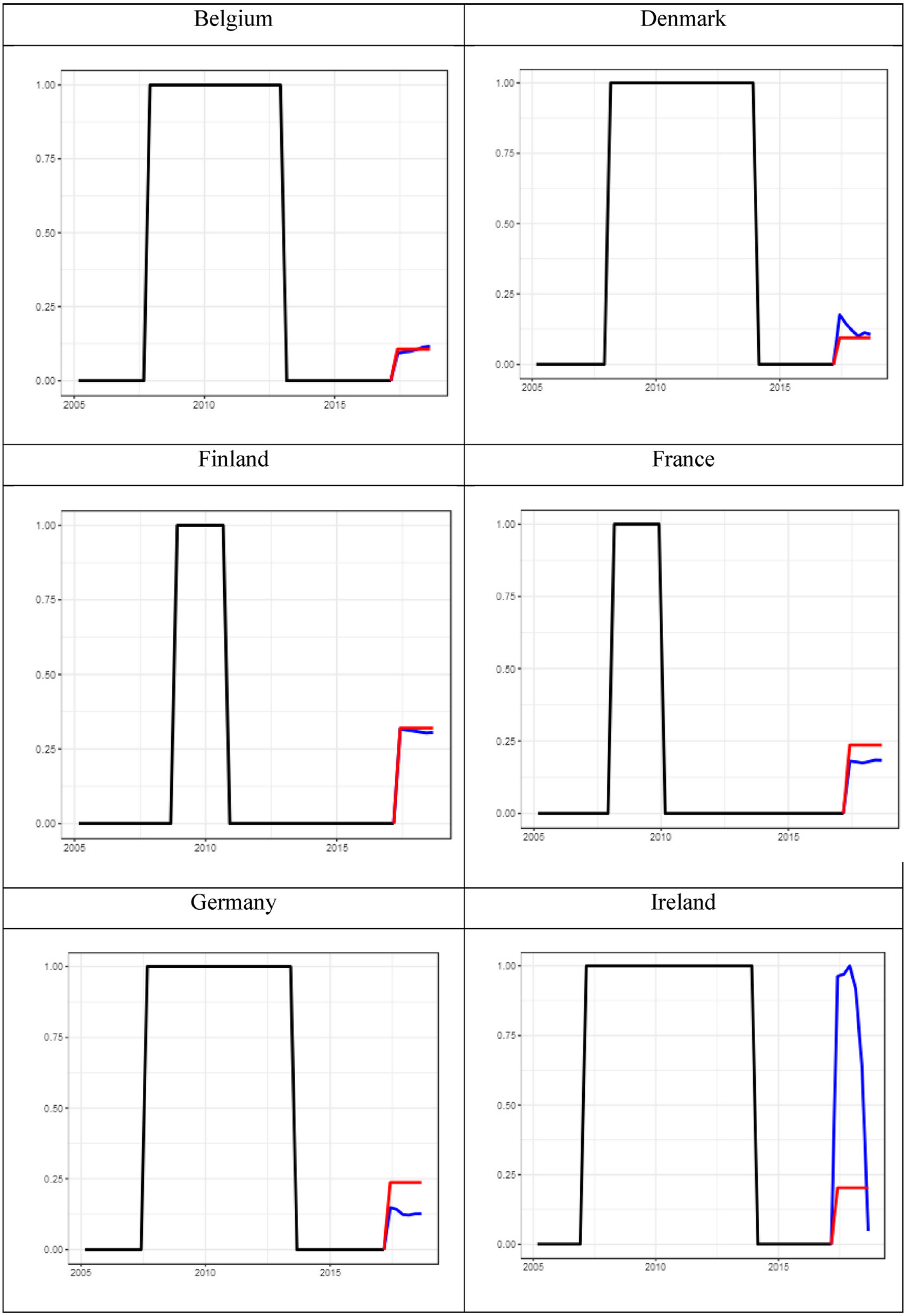


Fig. 14. Expected probability of crisis from the best model of each country.

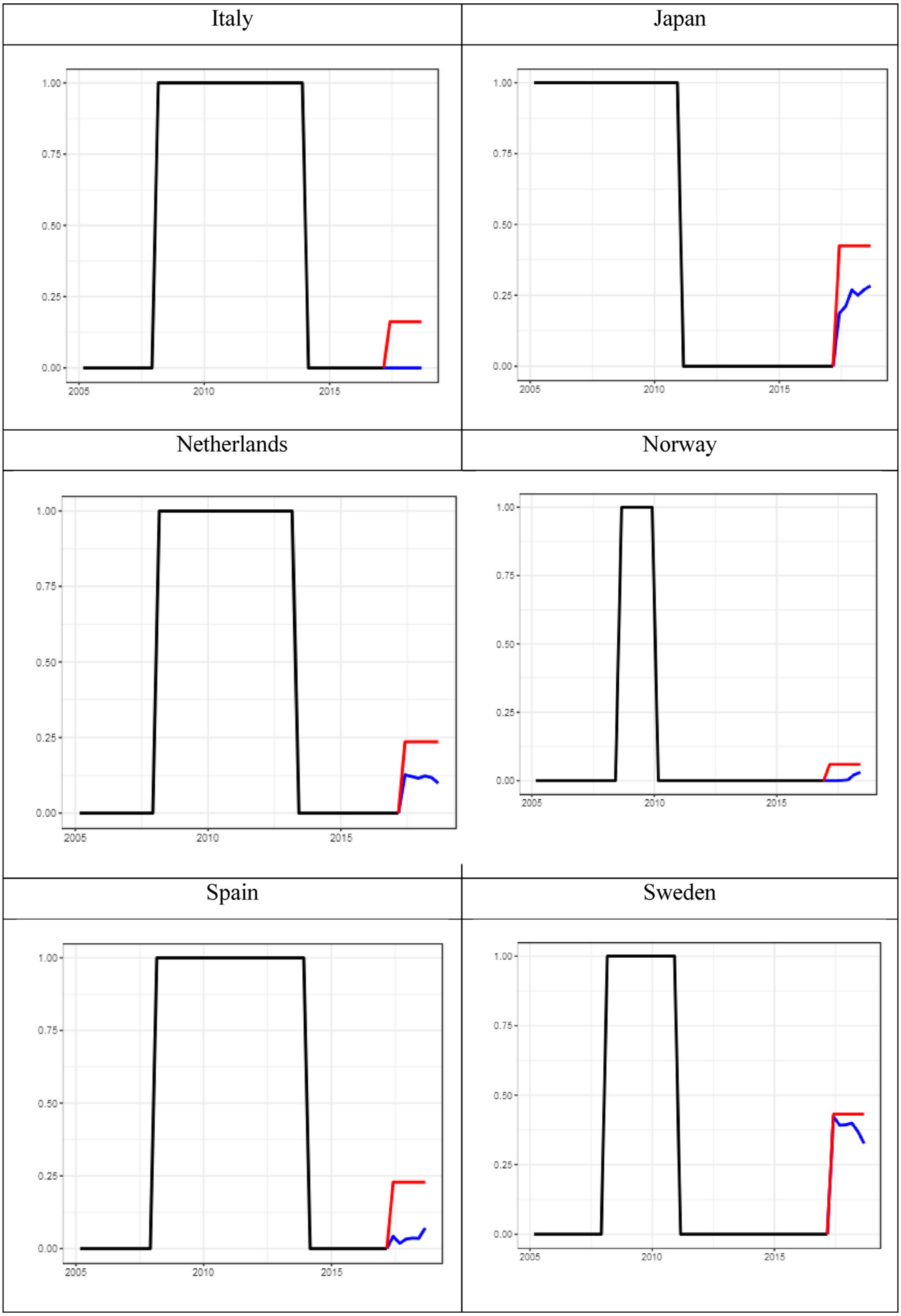


Fig. 14. (*continued*).

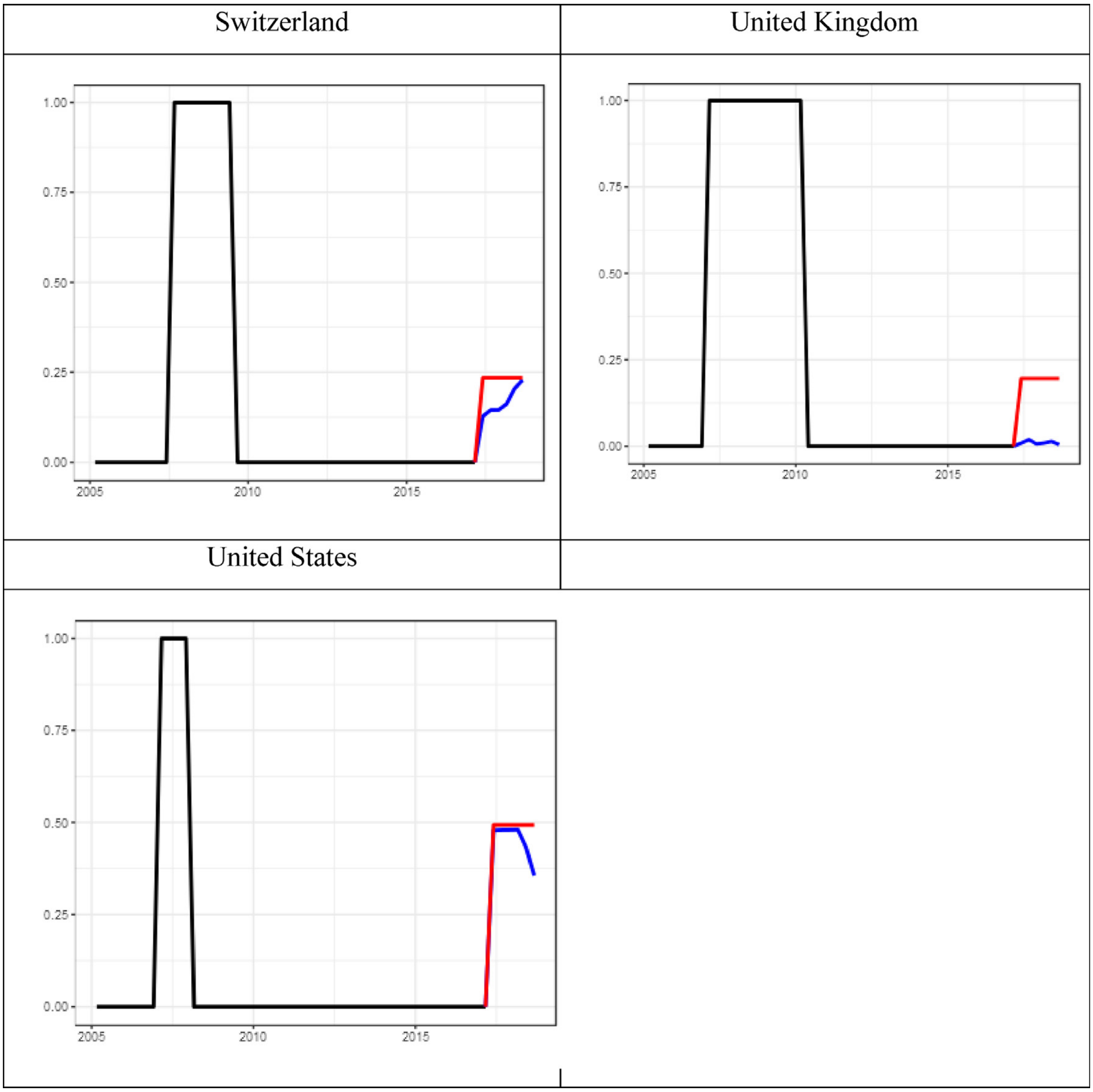


Fig. 14. (*continued*).

filters are included in the best models only once, while neural network models are never the best specification. Using desynchronized ratios and country-specific models often improve the quality of prediction. In addition, the variability of the AUCs among the best models by country and the different alarm thresholds suggest that a one-size-fits-all approach is not suited for estimating banking crises probabilities. As a result, we can recommend our methodology for operational purposes keeping in mind that additional variables can guarantee its robustness across prediction horizons.

4.3. A normative view of the benchmark methodology

For each possible model, our framework returns an estimated threshold to detect crisis. We name it the "alert threshold". Thus, an alert threshold can be built even for the method defined by the Basel gap standards. These alert thresholds are represented in Fig. 12 (in red), together with the predicted probability of crisis (in blue). We can notice that the benchmark model cannot be used as a one-size-fits-all method as the alert threshold significantly differs across countries. It is also remarkable that in some cases, we are unable to deliver a reliable threshold (too close from 0 or 1). This highlights that using the benchmark methodology could be misleading because, whatever the threshold chosen, the power of detection cannot be accurate.

4.4. Some case studies

Results obtained on a set of countries allow for a comparison of different in-sample and out-of-sample features. Our results are based on the last available observation for each series in 2017Q4. The prediction for crisis detection is hence made up to 6 quarters ahead that is to say until 2019Q2.

For most countries, we get a probability of banking crisis quite different and often inferior to the one stemming from the standard credit-to-GDP model. This underlines the importance of using alternative and better designed models, as well as auxiliary variables to evaluate the risk of a crisis at a proper level. In the following Fig. 13, we present a comparison of the estimated probability of crisis stemming from the benchmark model and from our described methodology in France and Germany.

In Fig. 14, we present the expected probability of crisis from our framework (in blue) with the estimated alert threshold (in red): when the blue line goes above the red line, there is a risk of banking crisis in the corresponding country at the time of crossing.

From a more economic and financial stability point of view, we find that some countries experience a state of the economy that should be considered as perilous: Denmark and Ireland on the short-term, and Switzerland on the long-term.

5. Conclusion

Predicting banking crises is challenging. We have presented the standard credit-to-GDP methodology and its limitations. In fact, its ability to predict incoming crises precisely is questionable. In order to assess the predictive power of this method, we have proposed alternatives in terms of filtering and modelling.

We have proposed a local level model as a modern way to extract the trend from a time series. It has many advantages, including being easily useable in real-time. We have demonstrated that including several lag versions of the credit and the GDP series in the models helps to detect crises. In fact, dynamic evolutions are better suited than a single punctual value when it comes to describing a phenomenon that evolves over time. Concerning models, we have proven that it is better to model the crisis with the data than just using arbitrary thresholds. Besides, using additional variables such as the inflation rate, the short-term interest rate, share prices, house prices in addition to lagged versions of the credit and the GDP, improves the detection.

According to our study, a framework well suited to detect banking crisis has to implement in priority the HP filter, the logit univariate models and the essential feature of leads and lags adjustments between credit and GDP dynamics. The other features such as country-specific or cross-country modeling and the use of additional variables are important and often improve the results. However, the use of fancy models such as random forest and multivariate logit and filters such as Kalman filtering proved to be useful only in very specific cases.

In future works, some improvements could be considered such as challenging the smoothing parameter λ of the HP filter, optimizing the specification of neural networks and fine-tuning thresholds in order to prioritize the ability to detect crisis more than limiting the false alarm rate.

Conflicts of interest

All authors have none to declare.

Appendix 1. Systemic banking crisis data sources combined in the database

1	Caprio and Klingebiel (2003)[8]	The annual dataset (1970–2002) includes information on 117 episodes of systemic banking crises in 93 countries and on 51 episodes of borderline and non-systemic banking crises in 45 countries. A systemic crisis is defined as "much or all of bank capital was exhausted." Expert judgment was also employed "for countries lacking data on the size of the capital losses, but also for countries where official estimates understate the problem."
2	Kaminsky and Reinhart (1999)[20]	The monthly dataset (1970–1995) includes 26 episodes of banking crisis in 20 countries. Banking crises are defined by two types of events: "(1) bank runs that lead to the closure, merging, or takeover by the public sector of one or more financial institutions; and (2) if there are no runs, the closure, merging, takeover, or large-scale government assistance of an important financial institution (or group of institutions) that marks the start of a string of similar outcomes for other financial institutions." The dataset of banking crises was compiled using existing studies of banking crises and the financial press.
3	Laeven and Valencia (2008, 2010, 2012)[21–23]	The annual dataset (1970–2011) covers systemically important banking crises (147 episodes) in over 100 countries all over the world and provides information on crisis management strategies. A banking crisis is considered to be systemic if the following two conditions are met: "(1) Significant signs of financial distress in the banking system (as indicated by significant bank runs, losses in the banking system, and/or bank liquidations); and (2) Significant banking policy intervention measures in response to significant losses in the banking system." The first year that both criteria are met is considered to be the starting year of the banking crisis, and policy interventions in the banking sector are considered significant if at least three out of the following six measures were used: "(1) extensive liquidity support; (2) bank restructuring costs; (3) significant bank nationalizations; (4) significant guarantees put in place; (5) significant asset purchases; and (6) deposit freezes and bank holidays." The dataset is compiled using the authors' calculations combined with some elements of judgment for borderline cases.
4	Reinhart and Rogoff (2013, 2011)[28,29]	The annual dataset (1800–2010, from the year of independence) covers banking crises in 70 countries. The definition of banking crisis is the same as in [20] (see above). The dataset of banking crises was compiled using existing studies of banking crises and the financial press.

References

1. Aldasoro I, Borio C, Drehmann M. *Early Warning Indicators of Banking Crises: Expanding the Family, BIS Quarterly Review*. 2018. March 2018.
2. Babecký J, Havránek T, Matějů J, Rusnák M, Šmídková K, Vašíček B. Banking, debt, and currency crises in developed countries: stylized facts and early warning indicators. *J Financ Stabil*. 2014;15(December):1–17, 2014.
3. Baron M, Verner E, Xiong W. Identifying banking crises. Princeton working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3116148; 2018.
4. Barrell R, Davis P, Karim D, Liadze I. Bank regulation, property prices and early warning systems for banking crises. *J Bank Finance*. 2010;34(9):2255–2264. September 2010.
5. Basel Committee on Banking Supervision. Guidance for national authorities operating the countercyclical capital buffer. Available at: https://www.bis.org/publ/bcbs187.htm; 2010.
6. Borgy V, Clerc L, Renne J-P. Measuring aggregate risk: can we robustly identify asset-price boom–bust cycles? *J Bank Finance*. 2014;46:132–150.
7. Borio C, Lowe P. *Asset prices, financial and monetary stability: exploring the nexus*. In: *BIS Working Papers*. vol. 114. Bank for International Settlements; 2002.
8. Caprio G, Klingebiel D. Episodes of systemic and borderline financial crises. In: Klingebiel D, ed. Washington DC: The World Bank; 2003.
9. Chirita M. *Usefulness of Artificial Neural Networks for Predicting Financial and Economic Crisis*. Annals of "Dunarea de Jos" University of Galati, Fascicle I. Economics and Applied Informatics; 2012. no2/2012.
10. Coffinet J, Montornès J. Le financement de l'habitat en France et la crise financière. *Rev Econ Financ*. 2014;(3):209–228. Association d'économie financière.

11. Coudert V, Idier J. An early warning system for macro-prudential policy in France. In: *Banque de France Working Paper Series no. 609, November*. 2016.
12. Drehmann M, Borio C, Gambacorta L. *Countercyclical Capital, Exploring Options, BIS Working Papers*. vol. 317. Bank for International Settlements; 2010.
13. Drehmann M, Borio C, Tsatsaronis K. Anchoring countercyclical capital buffers: the role of credit aggregates. *International Journal of Central Banking*. 2011;7(4):189−240. December.
14. Drehmann M. *Total Credit as an Early Warning Indicator for Systemic Banking Crises, BIS Quarterly Review*. June: Bank for International Settlements; 2013.
15. Drehmann M, Tsatsaronis K. *The credit-to-GDP gap and countercyclical capital buffers: questions and answers. BIS Quarterly Review*. March: Bank for International Settlements; 2014.
16. Ecer F. Comparing the bank failure prediction performance of neural networks and support vector machines: the Turkish case. *Ekonomska istraživanja-economic research*. 2013;26(3):81−98.
17. Fielding D, Rewilak J. Credit booms, financial fragility and banking crises. *Econ Lett*. 2015;136:233−236, 2015.
18. Hamilton JD. Why You Should Never Use the Hodrick-Prescott. *Rev Econom Statistics*. 2017. https://doi.org/10.1162/REST_a_00706.
19. Hodrick RJ, Prescott EC. Postwar U.S. Business cycles: an empirical investigation. *J Money Credit Bank*. 1997;29(1):1−16. February.
20. Kaminsky GL, Reinhart CM. The twin crises: the causes of banking and balance-of-payments problems. American Economic Association *Am Econ Rev*. 1999;89(3):473−500 [June].
21. Laeven L, Valencia F. Systemic banking crises: a new database. In: *International Monetary Fund Working Papers 08/224*. 2008.
22. Laeven L, Valencia F. Resolution of banking crises: the good, the bad, and the ugly. *International Monetary Fund Working Papers*. 2010;10:146.
23. Laeven L, Valencia F. Systemic banking crises database: an update. *International Monetary Fund Working Papers*. 2012;12:163.
24. Lo Duca M, Koban A, Basten M, et al. In: Peltonen T/, ed. *A New Database for Financial Crises in European Countries*. ECB/ESRB EU crises database, *European Central Bank Occasional Paper Series*; 2017. No 194/July 2017.
25. O'Brien E, O'Brien M, Velasco S. *Measuring and Mitigating Systemic Risk in Ireland: The Application of the Countercyclical Capital Buffer, Central Bank of Ireland Financial Stability Notes N°4*. 2018.
26. Petris G, Petrone S, Campagnoli P. *Dynamic Linear Models with R*. New York: Springer-Verlag; 2009.
27. Ravn MO, Uhlig H. On Adjusting the Hodrick-Prescott Filter for the Frequency of observations. *Rev Econ Stat*. 2002;vol. 84:371−375, 2, MIT Press.
28. Reinhart CM, Rogoff KS. Banking crises: an equal opportunity menace. *J Bank Finance*. 2013;37(11):4557−4573. November 2013.
29. Reinhart CM, Rogoff KS. From financial crash to debt crisis. *Am Econ Rev*. 2011;101(5):1676−1706. August, American Economic Association.
30. Ristolainen K. Predicting banking crises with artificial neural networks: the role of nonlinearity and heterogeneity. *Scand J Econ*. 2018;120(1):pp31−pp62.
31. Spanish Foundation for Science and Technology A neural network model predicts whether a bank can go bust available at: http://www.sciencenewsline.com/news/2015050610170014.html, 6 May 2015
32. Zidong A, Jalles JT, Loungani P. Ow well do economists forecast recessions? *International Monetary Fund Working Papers*. 2018:18−39.