THE SHERLOC: AN EWS-BASED INDEX OF VULNERABILITY FOR EMERGING ECONOMIES

2019

Irma Alonso and Luis Molina

Documentos de Trabajo
N.º 1946

BANCO DE **ESPAÑA**
Eurosistema

THE SHERLOC: AN EWS-BASED INDEX OF VULNERABILITY FOR EMERGING ECONOMIES

THE SHERLOC: AN EWS-BASED INDEX OF VULNERABILITY FOR EMERGING ECONOMIES (*)

Irma Alonso (**) and Luis Molina (***)

BANCO DE ESPAÑA

(*) The views expressed in this paper are those of the authors and do not necessarily represent the views of the Banco de España or the Eurosystem. We are grateful to Sonia López Senra for her extraordinary and essential work in the earlier stages of the paper. We also thank Daniel Santabárbara, Juan Carlos Berganza, Pedro del Río, Sonsoles Gallego, Enrique Alberola, José Manuel Marqués, Javier Pérez, Luis Orgaz and José Ramón Martínez for their useful comments, as well as the participants in Banco de España seminars and IRC meetings.

(**) irma.alonso@bde.es.
(***) lmolina@bde.es.

Documentos de Trabajo. N.º 1946

2019

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: http://www.bde.es.

ISSN: 1579-8666 (on line)

Abstract

This paper presents a tool to detect the accumulation of risks in emerging market economies based on a synthetic index of "vulnerability" for three different types of crisis (sovereign, currency and banking crises). To build the index we first use a signalling approach (Auroc) to preselect the variables that issue adequate signals before the blown up of a crisis. The short-term interbank rate is a leading indicator for the three different types of crises and short term external debt also plays a prominent role. These variables are then introduced in a logistic estimation to obtain the predicted probability of being in a "vulnerable" state for each type of crisis. These indexes, labelled SHERLOC, which stands for Signalling Heightened Emerging Risks that Lead to the Occurrence of Crises, outperform all best single indicators in terms of in-sample and out-of-sample validation. Additionally, a synthetic index for each type of crisis seems to predict better "vulnerable" states than the use of an aggregate index for all types of crises.

Keywords: emerging economies, crisis, vulnerabilities, early warning models, risks.

JEL classification: E44, F01, F34, F37, G01.

Resumen

Este documento presenta una herramienta para detectar la acumulación de debilidades en las economías emergentes basada en un índice sintético de «vulnerabilidad» frente a tres tipos diferentes de crisis (crisis soberana, cambiaria y bancaria). Para construir el índice, primero nos centramos en un enfoque de señalización (Auroc) que nos permite preseleccionar las variables que emiten señales adecuadas antes de que estalle una crisis. El tipo de interés interbancario a corto plazo parece ser uno de los principales indicadores adelantados para los tres tipos de crisis, mientras que la deuda externa a corto plazo también desempeña un papel destacable. Estas variables preseleccionadas se incorporan en una estimación logística para obtener la probabilidad prevista de estar en un estado «vulnerable» frente a cada tipo de crisis. Estos índices, denominados SHERLOC, que es el acrónimo de *Signalling Heightened Emerging Risks that Lead to the Occurrence of Crises* (en español, señalización de riesgos crecientes en emergentes que conducen a la ocurrencia de crisis), ofrecen mejores rendimientos que los mejores indicadores individuales en términos de validación dentro y fuera de la muestra. Además, un índice sintético para cada tipo de crisis parece predecir mejor estados «vulnerables» que el uso de un índice agregado para todos los tipos de crisis.

Palabras clave: economías emergentes, crisis, vulnerabilidades, modelos de alerta temprana y riesgos.

Códigos JEL: E44, F01, F34, F37, G01.

1 Introduction and motivation

Emerging Market Economies (EMEs) have become increasingly relevant in the world economy, both in terms of trade and financial flows. Given that these economies have traditionally been more prone to suffer crises, early warning signals (EWS) models regain relevance in order to anticipate vulnerable states in EMEs that can affect global financial stability. In that sense, there are two key questions of research. First, could these stress episodes be anticipated sufficiently in advance to tame its global effects? Second, which variables should be prominently monitored to detect the accumulation of vulnerabilities in these economies?

In this paper we try to answer both questions based on the large and growing literature on EWS. The contribution of this paper is twofold. First, it develops a quarterly dataset of sovereign, currency and banking crises, partially based on the seminal work of Laeven and Valencia (2008, 2012 and 2018) and Reinhart and Rogoff (2009). Contrary to the rest of the literature, we define the beginning and the end of crises on a quarterly basis in order to develop a timely Early Warning model. Second, and more relevant, we propose a synthetic index of vulnerability for emerging economies, distinguishing by type of crisis (sovereign, banking and currency), which is labelled SHERLOC (Signalling Heightened Emerging Risks that Lead to the Occurrence of Crises)[1]. We show that a synthetic index for each type of crisis seems to outperform an aggregate index for all types of crises.

Along with their greater presence in global markets both in trade and financial terms[2], emerging economies have experienced a large number of currency, banking and sovereign crises with high costs in terms of economic activity and employment. There are well-known examples of such situations: Mexico in 1995, Asia in 1997, Russia in 1998, or Argentina in 2001[3]. In recent years, crises have been less frequent, partly due to the measures implemented by policymakers to avoid them, adopting floating exchange rate regimes and de-dollarizing public debt. But, in some cases, emerging economies have continued to face periods of heightened financial volatility, related either to their greater integration in trade and financial international flows or to domestic disequilibria. These stress periods resulted in increases in risk premia, declines in stock indexes and significant depreciations of their currencies, which in some cases also fed economic declines, as pointed out by the recent turbulences in Argentina and Turkey.

Although most of these recent episodes reversed faster than in previous periods, it is essential to monitor the build-up of vulnerabilities in emerging economies. Firstly, it enables policy makers and domestic investors to assess the economic situation of these countries and the risks they might face. Secondly, in a more and more globalized world, risks derived from a higher integration in global financial markets are relevant and should be taken into consideration. Financial markets in emerging economies have substantially developed and therefore financial spillovers and spillbacks from emerging economies to advanced economies are becoming increasingly relevant (see IMF's Global financial Stability Report, April 2016). Finally, a financial crisis in one of these countries can affect the financial stability of other economies through the positions of financial institutions and firms in EMEs. That is why, for instance, the European Systemic Risk Board (ESRB) established the possibility of imposing countercyclical capital buffers to European banks that are exposed to third countries with unaddressed risks[4].

In order to build the synthetic indexes to monitor EMEs vulnerabilities, we use a signalling approach (AUROC) to assess, in a univariate setting, the predictive ability of each indicator before the onset of a crisis, be it sovereign, currency or banking crisis and to select the variables that make up these new vulnerability indexes using a standard logit model. Other methodological approaches are also proposed to build the index, such as a factor augmented logistic estimation, a simple average of the risk percentiles or an aggregation of preselected variables based on

[1] The acronym SHERLOC also reminds us the main use of the index, which could provide clues to detect the skeleton in the closet

[2] According to published earnings and profitability reports, some European banks obtained around half of their EBITDA in emerging economies.

[3] For more information on these crises, see Appendix 1.3.

[4] See Decision of the European Systemic Risk Board of 11 December 2015 on the assessment of materiality of third countries for the Union's banking system in relation to the recognition and setting of countercyclical buffer rates (ESRB/2015/3) (2016/C 97/11).

principal component analysis (PCA). Finally, we also assess the in-sample and out-of-sample predictive ability of the respective SHERLOCs, using the methodology of Alessi and Detken (2011). Employing this evaluation method, we show that a synthetic index for each type of crisis outperforms (i) an aggregate index for all types of crises and (ii) the best single indicators for each crisis (mainly sovereign spreads and short-term interbank rate) in the sense that they increase the usefulness of the model. However, it is essential to bear in mind that the aim of these indexes is not to predict crises, but to identify underlying vulnerabilities and imminent tail risks that predispose a country to a crisis.

The rest of the paper is structured as follows. The next section provides a brief literature review and discusses the main pitfalls of EWS methodology. Section 3 presents the main features of the dataset used in the econometric analysis. Section 4 describes the empirical strategy used in the paper, the results of the non-parametric and parametric approaches to detect vulnerabilities, and the construction and validation of the SHERLOCs. Section 5 concludes and highlights future work[5].

2 Literature review and dealing with EWS' pitfalls

There is a widespread and growing literature on early warning models, developed in the aftermath of the seminal paper of Kaminsky, Lizondo and Reinhart (1998). In its early stage, most of the literature on EWS focused on analysing EMEs' risks, such as done by Bussiere and Fratzscher (2006); Lestano and Kuper (2001) for Asian countries; and Kamin et al (2001) for Latin American countries. In the wake of the global financial crisis, most recent papers have focused on developed countries[6] and assessing vulnerabilities that can affect countries' financial stability[7]. Our work is partially based on this large strand of the literature to select the leading indicators of our analysis, considering all types of crises occurred over the period 1993-2018 in emerging markets.

Nevertheless, a large number of theoretical and empirical papers have challenged the EWS methodology. First, on the empirical front, there are too few crisis observations to reach consistent logit estimators. This problem is even more acute when dealing with out-of-sample calibration of EWS, as the number of crisis observations dwarfs. In this paper, as the aim of the synthetic index of vulnerability is not to predict crises, but to identify underlying vulnerabilities, we use six quarters previous to a crisis, which partially mitigates the issue of lack of enough stress periods. Moreover, the use of random effects in the logistic estimation allows us to exploit the information of all the countries in the sample, even those with no crisis observations.[8]

Second, there could be a problem of "real time" EWS. This means that the authorities and economic modellers do not have the entire set of information used by a posteriori estimations[9]. The need to build real time indicators is a time consuming task in macro samples with heterogeneous countries and the use of a common lag of publication for all countries could bias the results[10]. In this analysis, the aggregate indexes are constructed in real time each time the exercise is updated. While in the baseline model we decide to keep the model as simple as possible discarding "the real-time issue", we carry out a validation exercise using a pseudo-real time approach, whose results can be found in the appendix V. Our results are mostly robust to this specification.

[5] This approach can be complemented by the use of vulnerability dashboards (heat maps), which have been useful to detect vulnerabilities in the past, as decribed in Alonso, I. and Molina, L. (forthcoming), "A complementary approach to monitor risk in emerging economies: the vulnerability dashboard", Banco de España Ocassional Paper, forthcoming.

[6] Rose and Spiegel (2012) and Frankel and Saravelos (2012) propose different methodologies to determine the drivers of the global financial crisis using a cross-country approach. Catao and Milesi-Ferreti (2014) highlight the role of foreign liabities to explain external crises in advanced economies.

[7] Oet et al (2013), Gramlich et al (2010) and Castro et al (2016) develop EWS for systemic risk to implement macroprudential tools.

[8] Another way to avoid this problem would be to build a kind of sole country in which crises and variables will be placed consecutively, as explained in Gadea and Pérez Quirós (2012).However, this is appropiate when the degree of homogeneity when a shock occurs is higher than the homogeneity within countries, and this is probably not the case in our sample.

[9] At time t they only have information on GDP in t-1, on NEER on t, on Doing Business indicators at time t-4 and so on.

[10] Another additional problem is what we would labell as "true real time", that is, to rigorously calibrate an EWS we should use the data that were actually published at the time of the calibration, which are usually revised later, especially those referred to activity and the public sector balances. In our case, we do not have the vintages to carry out true real-time EWS.

Third, EWS may present a post-crisis bias of discrete dependent variables as macro variables tend to be very persistent and therefore some indicators show an "erratic behaviour" in the recovery phase. Bussiere and Fratzscher (2006) predict financial crises in EMEs by relying on a multinominal logit model, which allows to distinguish between three states: a normal, pre-crisis, and post-crisis state. In our case, we show that considering the "post-crisis" bias, by eliminating the data of four quarters after a crisis, is relevant to slightly reduce type II errors (false alarms) out-of-sample, but it does not seem to affect the ratio of crisis predicted in-sample.

Fourth, these EWS models might also suffer from a problem of prediction failures outside the sample that could be partially attributed to an in-sample overfitting and variable selection bias. As many combinations of different variables could lead to the same prediction, the final selection of variables can be arbitrary, which leads to data mining or cherry picking issues. A way to circumvent those pitfalls can be related to the use of automated variable selection methods like the Lasso[11] or out-of-sample validation methodologies like the Random Forest[12], which basically rely on the statistical improvement of the results to decide whether a concrete indicator is included or not in the model. In this paper we deal mainly with the overfitting issue in a simpler way than Lasso or random forest models. First, we use a signalling approach (Auroc) to preselect the variables that issue the best signals ahead of a crisis[13], and then use these variables to build the vulnerability indexes SHERLOC, via a panel logit model[14] -estimating the predicted probability of being in a state of vulnerability-, a factor augmented logit model, or simplifying them via principal component analysis techniques[15]. These two methodologies have been broadly used by the literature (Bruggemann et al (2000), Edison (2003), Frankel and Rose (1996) and Lo Duca and Peltonen (2013)), but as far as we know these papers do not provide a clear framework to preselect variables as we do.

The main contribution to the literature of this paper is twofold: 1) we develop a quarterly dataset of events, distinguishing by type of crisis, and 2), we propose a synthetic index of vulnerability for EMEs. First, one of the most relevant issues of early warning models is how to define stress events. There are two main ways to define them. One can rely on a binary indicator to define periods of crisis. This approach has several advantages. First, it is more flexible since you can easily define "vulnerable states" as pre-crisis periods as we do in this paper. Second, we are able to distinguish between different types of crises. While most papers rely on the seminal work of Laeven and Valencia (2008, 2012 and 2018), Reinhart and Rogoff (2009) or Kaminsky and Reinhart (1999), we contribute to the literature by developing a quarterly dataset of sovereign, banking and currency crises, partially based on Laeven and Valencia (2008, 2012) and Reinhart and Rogoff (2009), for 25 emerging economies in order to get a timely early warning system model. This approach faces some limitations since an exogenous definition of crisis can be subject to expert judgment and misclassification, and needs to deal with the "post-crisis" bias, as we do as a robustness check. Moreover, it requires a sufficient number of stress periods to get robust results. That is why some papers rely on the use of continuous variables, such as Financial Stability Indexes (FSI) or Exchange Market Pressure (EMP) (Eichengreen (1995, 1996) and lo Duca and Peltonen (2013) and let the model endogenously choose the periods of crisis through the use of Markov Switching Models (Peria (2002)). However, in our case it is difficult to follow this strategy due to the lack of continuous variables long enough for our sample of EMES countries. For instance, we could use FSI measures, such as the ones developed by Berganza and Molina (forthcoming), but at the cost of fewer observations since most of

---

[11]The Lasso (Least Absolute Shrinkage and Selection Operator) technique was introduced in order to improve the prediction accuracy and interpretability of least square regression models by altering the model fitting process to select only a subset of the provided covariates. It forces the sum of the absolute value of the regression coefficients to be less than a fixed value, which leads to certain coefficients to be set to zero, effectively leading to a simpler model. After repeating the estimation, variables that are retained most often are selected. See Tibshirani, R. (1997).

[12]The random forest methodology operates by constructing a multitude of "decision" trees and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. In the case of EWS, it could consist of selecting the regressor based on their relative importance, that is, on the relative increase of the accuracy of the prediction once the concrete variable is included. See Breiman (2001).

[13]The signalling approach was initially developed by Kaminsky et al (1998) to identify macroeconomic variables that can anticipate currency crises based on a "critical threshold" calculated by minimising the noise-to-signal ratio for each indicator. But, it does not consider the interrelationships between variables.

[14]One of the earliest contributions was from Frankel and Rose (1996) who study the determinants of currency crashes in 100 emerging economies from 1971 to 1992. Lo Duca and Peltonen (2013) also develop a framework for predicting systemic events, which incorporate both domestic and global indicators that improve the forecast performance of the model.

[15]Many variables are introduced in the regression in levels, although some of them (typically stocks variables like credit to GDP or those related to debt) could have a trend that could bias their accuracy as crises' predictors. In some cases we have introduced first time differences to avoid this problem.

them start after 2000 for emerging economies. Moreover, it is debatable which indicator to use in order to capture all the stress events that we are interested in.

Second, we propose a synthetic index of vulnerability for EMEs. Few papers have proposed a synthetic index of vulnerabilities for emerging economies. One of the most recent is the index of vulnerability for emerging economies proposed by Lepers and Sánchez-Serrano (2017). Contrary to our paper, they only focus on financial crises and do not evaluate the performance of their composite index. In this sense, our focus will be on all the crises occurred over the period 1993-2018 in emerging markets in order to exploit the heterogeneity of different types of crises. Additionally, we incorporate a widespread evaluation method, developed by Alessi and Detken (2011), to assess the predictive ability of the SHERLOC both in sample and out-of sample. Alessi and Detken (2011) propose a new measure of Usefulness, which compares the loss of the model with regards to the loss of disregarding the model, taking into account policy maker's preference. While the in-sample performance is quite adequate, in line with the results of Alessi and Detken (2011), the out-of-sample performance is poorer, as in other out of sample validations. Moreover, one needs to be aware of some of the criticisms that still hold. First, noise (excessive issuance of signals) remains high. Second, we are not able to capture non-linearities as pointed out by Eichengreen (2002). Third, even if the use of different types of crises for each country partially mitigates the criticism of "this time is different" and the out-of-sample performance suggests that "not every time is different", a new type of crisis anticipated by other variables can always appear without been detected by the SHERLOC.

3 Data

3.1 Building an Early Warning System: Stress events

The first step of EWS models consists of identifying the relevant periods of crisis in the 25 emerging economies analysed. The definition of crisis is crucial in EWS and needs a thorough analysis to avoid misclassification issues and uncertainty.The identification of stress events are used in order to proxy "vulnerable states" in the econometric analysis, defined as the six quarters previous to the crisis.

One of the contributions of this paper is the construction of a quarterly dataset of stress events, distinguishing between sovereign, banking and currency crisis. To do so, we mainly identify sovereign and banking crises following Laeven and Valencia (2008, 2012), but defining the beginning and the end on a quarterly basis instead of an annual frequency, and updating the dataset until the end of 2018[16]. The beginning of a sovereign crisis is defined as the quarter in which the sovereign defaults or restructures its debt according to S&P while the end of a sovereign crisis is associated with the quarter in which an agreement with debt holders is reached, or alternatively, the date of the debt exchange, which implies a partial reaccess of the sovereign to international markets. For banking crises, we date the quarter of the start of the stress using national authorities' information or IMF's reports. Banking stress are determined when there are significant signs of financial stress in the banking system and banking policy interventions. The end date of the crisis is assigned to the quarter in which the eighth consecutive quarter of growth of both real GDP and real credit is reached, which is the criteria used by Laeven and Valencia (2008, 2012 and 2018)[17]. Finally, for currency crises we rely on a definition similar to that of Reinhart and Rogoff (2009) though with a more restrictive percentage. A crisis is assigned when a quarterly depreciation of the exchange rate against the US dollar is above 30%[18]. As a

---

[16] The update was made before the publication of Laeven and Valencia (2018). That is why there are slight differences. For instance, we consider a banking crisis in Russia in 2015 that is not included in LV (2018).

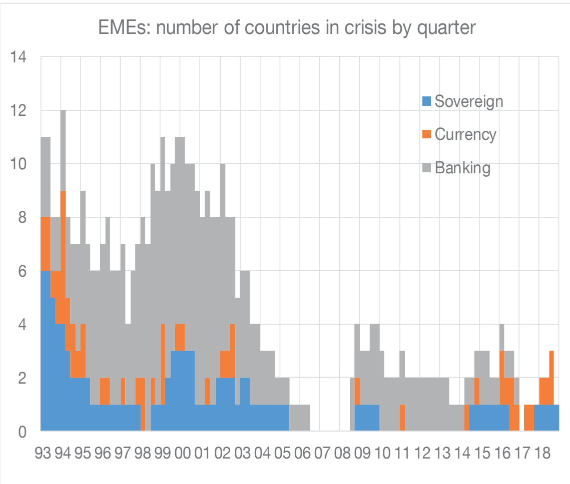
[17] However, in our dataset, we decide not to truncate the end of banking crises. Laeven and Valencia truncate banking crises 5 years after the blown up.

[18] As events are defined in a different way (banking and sovereign are proxied by events while currency by a quantitative threshold), there is a wide heterogeneity in terms of frequency and duration of crises. Indeed, currency crises tend to be more frequent but with a shorter duration. We have discarded the definition of a currency crisis as a change in the exchange rate regime as since the mid 00s almost all countries in our sample have adopted freely floating regimes.

robustness check, we also define a currency crisis as a depreciation of at least 15% on a quarterly basis as long as this depreciation exceeds the average variation of the exchange rate plus a standard deviation (see table A1.3 Annex 1 for a detailed description of the dates of crises).

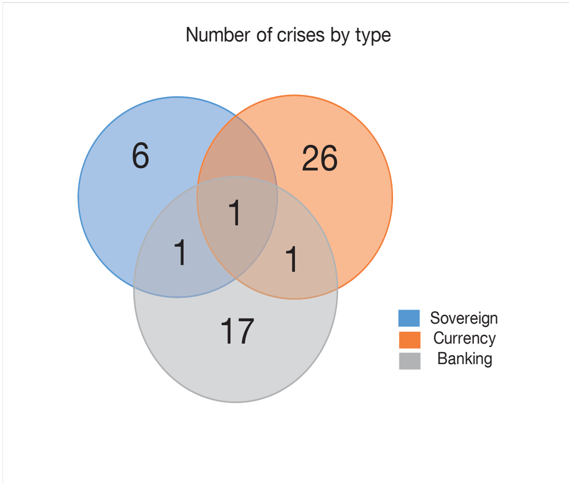
Once the crisis events are identified on a quarterly basis, we construct three dummies, one for each type of crisis, that are used in the econometric approach. Chart 1 plots the number of crisis identified per quarter[19]. As expected, the number of crisis diminished dramatically since 2004. Sovereign and currency crises are less frequent than in the nineties since many countries of the sample that used to rely on a fixed exchange rate to the USD as the main tool to stabilize domestic inflation have now turned to interest rates with floating exchange rates. Finally the percentage of banking crises in terms of total crises has increased after 2008. Chart 2 presents the exact number of crises in our sample, distinguishing between currency, sovereign and banking crises. As shown in Chart 2, in our dataset, there are only a twin sovereign-banking crisis (Argentina 2001 q4), a twin currency-banking crisis (Indonesia 1997 q4) and a triple crisis (Russia 1998 q3)[20].

Chart 1



Number of countries with at least one type of crisis by quarter

Chart 2



Notes: number of crisis only takes into account the quarter in which the crisis begins

As the objective of EWS is to anticipate the appearance of crises with enough time to enable authorities to react, indicators should send a correct signal in advance of the stress periods identified. Considering the type of crisis we are dealing with, we use an evaluation window up to 6 quarters prior to the event (consistent with the results of Kaminsky et al (2000) who stated that warnings of a crisis usually appear 10 to 18 months before the outset). This means that the dummy identifies the one and a half year period before the crisis as the period in which indicators should send correct signals and the Auroc approach assesses the relevance and performance of each variable over this period[21]. Crisis years are not taken into consideration as our objective is to identify leading indicators, and indicators tend to have an erratic behaviour during periods of stress. The rest of years are considered normal or tranquil periods and therefore are equal to zero.

3.2 Data for the econometric approach

Our dataset includes 25 countries, representing around 78% of the GDP of developing economies, and around 45% of world GDP. It comprises 9 Latin American countries (Argentina, Brazil, Mexico, Colombia, Chile, Peru,

[19] Note that these numbers do not correspond exactly with the number of countries in crisis as one country could suffer from two or even three types of crisis at the same time.

[20] This avoid us to conduct a concrete analysis of the determinants of twin crises. Therefore, the possible bias arising from the presence of twin crises does not seem to be too problematic. It is worth noting that the currency crises in Argentina in 2001 takes place a quarter after the sovereign and banking crises.

[21] In pratical terms this implies that the dummies are defined with a 1 six quarters before the onset of the crisis, with missing values during the crisis, and zero in tranquil no pre-crisis times. Four and eight quarters are also considered as robustness checks (see Appendix III).

Venezuela, Ecuador and Uruguay), 5 developing Asian nations (China, South Korea, India, Indonesia and Thailand), 6 Eastern Europe countries (Czech Republic, Hungary, Poland, Romania, Russia and Turkey) and 5 countries from Africa and the Middle East (South Africa, Nigeria, Saudi Arabia, Egypt and Morocco). The selection of countries is determined by the availability of data, which can be easily updated, and also by their stronger commercial and financial ties with Spain[22].

Table 1 reports the 35 vulnerability indicators used in this paper. Their election is mainly based on the most significant variables suggested by the literature on EWS. Notice, however, that we include a large number of variables related to the banking sector, in order to reflect the fact that banking crises have become more relevant for advanced economies in the 21st century. In addition, most indicators are easy to update and allow a frequent monitoring of the risks faced by the countries in the sample. Moreover, we use both national and international sources in order to be able to extend the dataset back to the first quarter of 1993. This enables us to include some of the most relevant idiosyncratic crises (Mexico, Asia or Russia in the 90s).

The 35 variables are divided into four groups, mainly reflecting the frequency of update, and their presumed capacity to react to an increase in risks. Financial markets variables –updated daily- are supposed to react more quickly to situations of high vulnerability, and even an overshooting of them could trigger a crisis. Fundamental variables –montly or quarterly updated- tend to reflect the increase in risks more parsimoniously and constitute the core of the econometric analysis. Finally crises coming from a deterioration of the institutional quality –whose indicators are updated annually- tend to have a longer ripening process, and their transmission channels differ from those of macroeconomic variables. For instance, a worsening of the institutional framework can deter foreign and internal investment and it can lead to sudden stops or capital flights even if fundamental variables are in good shape. The wilingness of governments to pay back their external debt is also a relevant factor, which should be monitored[23]. Finally, some variables have been included in the estimation strategy either to proxy contagion -both from a global turbulence or from EME's interconnectedness and EME's dependence on global commodity prices- and to control for global factors that could bias the results.

For the econometric analysis all the indicators are transformed into a quarterly frequency, using either the last data of the corresponding quarter or the quarterly average (from lower to quarterly frequency) or a linear interpolation (from annual to quarterly frequency). Moreover, the lack of time series large enough leads us to discard many relevant qualitative variables[24].

[22] See Luis Molina, Esther López, Enrique Alberola, 2016. "An index of external positioning for the Spanish economy," Occasional Papers 1602, Banco de España

[23] Most of the data used is comparable since they are extracted from international or official national statistical sources. However, there remain differences, for instance with the definition of reserves which in some cases include reserves that cannot be mobilized by the monetary authority.

[24] Some of them are the BICRA indicators (Standard and Poor's), banking and political risks scores (EIU), or Doing Bussiness and Absence of Violence (World Bank). All of them are introduced in the vulnerability dashborads described in Alonso and Molina (2020 forthcoming)

Table 1

FINANCIAL MARKETS	
* Sovereign spread (bps level) * Sovereign spread (change over 3 months) * Stock Exchange index (change over 3 months) * Exchange rate vis a vis the USD (change over 3 months)	
MACROECONOMIC FUNDAMENTALS	
<u>REAL SECTOR:</u> * GDP (change y-o-y) * Inflation rate * Industrial production (12 month MA, y-o-y) * NEER overappreciation <u>FISCAL SECTOR:</u> * Public sector balance (% GDP) * Public sector gross debt (% GDP) <u>EXTERNAL SECTOR:</u> * Current account balance (% GDP) * Gross external debt (% GDP) * FDI (% GDP) * Short term external debt (% Reserves) * Reserves (% GDP) * External debt service (% exports) * Portfolio gross inflows (% GDP)	<u>BANKING SECTOR:</u> * Real credit to private sector (y-o-y) * Real deposits on domestic banks (y-o-y) * Loan to Deposit ratio * Non performing loans (% total loans) * Net foreign assets of domestic banks (% GDP) * Bank Stock Exchange (change over 3 months) * Spread of bank's external debt (3 months change) * Short term interbank rate (%) * Intermediation margin (loan rate - deposit rate)
WEALTH AND INSTITUTIONAL QUALITY * Per capita GDP (USD PPP and % change) * GPR index * Sovereign rating (average of the 3 agencies)	**CONTAGION RISKS** * VIX *10 year US Treasury bond yield * Short term interest rate * Trade links * EMBI sovereign spread * Oil prices

Variables included in the AUROC exercise, mainly taken from the previous EWS literature

4 Empirical analysis

In order to minimize "data mining" issues, we preselect the best leading indicators by using a signalling approach (the AUROC), and then used these variables in a logistic estimation to get the predicted probability of being in a "vulnerable" state.

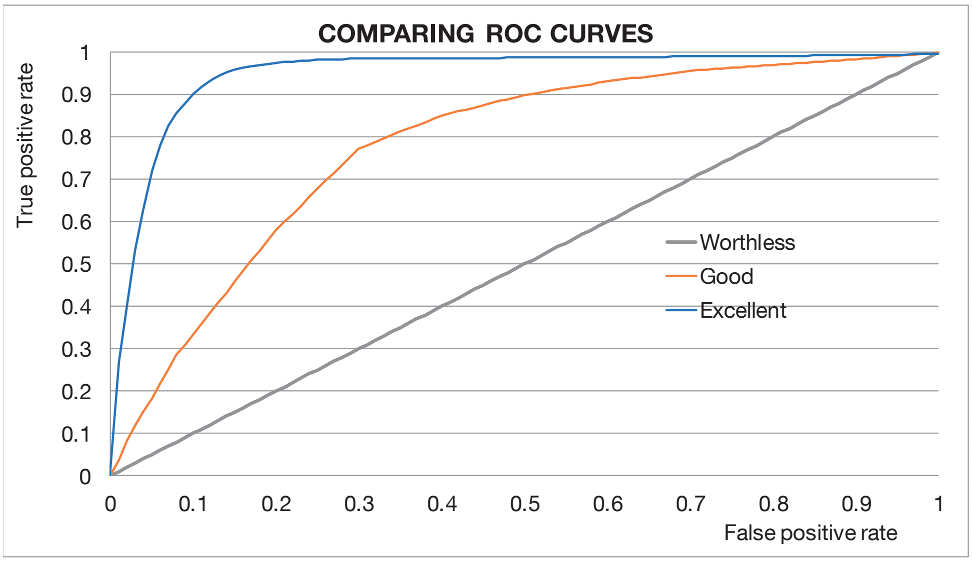
4.1 Signalling approach: AUROC results

The AUROC (Area Under the Receiver Operating Characteristics) approach is a univariate setting that measure the performance of each variable to distinguish between two distributions (in this case normal and stress periods). More specifically, it extracts signals from each indicator when it is above or below a certain threshold calculated with its historical data[25]. Based on these signals, the procedure estimates the performance of each variable depending on the number of correct signals (right guess of a crisis happening or not happening) and false signals (missed crises or predicted crises not happening) issued in the predetermined evaluation window. The Auroc evaluates each indicator performance taking into consideration the true positive rates in terms of the false positive rates. A value of 1 implies that the specific indicator is only issuing correct signals, whereas a value of 0.5 provides no information at all since it would anticipate crises randomly (like flipping a coin).

[25] See, for instance, Castro, Estrada and Martínez (2016) for an application to the case of the financial sector in Spain.

More specifically, the AUROC methodology indicates how well the model is able to distinguish between the probabilities of being in a vulnerable state or in a normal period. Indeed, as the threshold imposed to differentiate between stress and normal time increases, the ratio of true positive rate and false positive rate declines. An AUROC equals to 0.5 implies that there is an overlap between the distributions of crisis and normal times (Chart 3).

Chart 3



The AUROC estimates the performance of each indicator depending on the ratio of true positive rate in terms of false positive rates. An Auroc equal to 0.5 (worthless) cannot distinguish between the distributions of normal times and stress periods. The higher the AUROC, the better.

In technical words the AUROC estimates the area under the curves ROC (Receiver Operating Characteristics), which show the relation between the true positive rate and false positive rate for different thresholds:

$$AUROC = \int_0^1 ROC\ (FP)dFP$$

to evaluate the predictive ability of each variable as an early indicator of a crisis. FP stands for false positive rate. It is worth mentioning that the AUROC model used assumes that policy makers preferences are equally concerned about type I (missing crises) and type II (false positives) errors[26].

The results for the signalling approach are presented in Table 2[27]. The variables that seem to send strong signals to detect an increase in vulnerabilities are marked in red and in orange, indicating a percentage of good signals above 70% and 65%, respectively.

The most outstanding result would be that, in the case of the anticipation of any kind of crisis, only a few variables reach an adequate performance, in terms of the ratio of true positive rates and false positive rates (just 6 out of 48). Moreover, only half of them come from the so-called "fundamental" indicators (the rate of inflation, the short-term interest rate, and the ratio of short term external debt to reserves), together with another variable reflecting market sentiment (the sovereign spread) and contagion from global turbulences (the US 10 year interest rate). As for the "qualitative" variables, just the sovereign rating seems to issue right signals in an adequate proportion. These apparently disappointing results could be due to the double

[26] Nevertheless other AUROC specifications could give more weigth to any of the error type (pseudo AUROC).
[27] Each variable has been transformed so an increase implies greater risks.

Table 2

Variable	By type of crisis				By region				By time	
	All crises	Banking	Sovereign	Currency	Latin America	East. Eur.	Asia	Other EMEs	Before 2007Q1	After 2007q1
Sov.spread	**0.73**	0.59	**0.90**	**0.80**	**0.77**	0.64	0.18	**0.83**	**0.68**	**0.77**
GDP growth	0.60	0.53	**0.83**	**0.67**	**0.67**	0.58	0.35	0.57	0.60	0.63
Inflation rate	**0.74**	0.64	**0.68**	**0.75**	**0.70**	**0.76**	**0.68**	**0.79**	0.64	**0.86**
NEER deviation from mean trend	0.59	**0.68**	0.59	0.58	0.58	0.49	**0.72**	0.62	0.63	0.55
Public sector balance	0.56	0.39	0.57	**0.66**	**0.70**	0.56	0.26	0.57	0.53	0.64
Public debt (first diff)	0.62	0.57	**0.74**	**0.69**	0.62	0.59	0.41	**0.77**	0.64	0.63
Net Foreign Assets dom.banks	0.59	**0.68**	0.56	0.56	0.53	0.51	**0.65**	0.57	**0.66**	0.51
Loan to Deposit ratio	0.54	**0.71**	0.62	0.46	0.45	0.48	**0.93**	0.40	**0.71**	0.31
Short term interbank rate	**0.74**	**0.78**	**0.70**	**0.75**	**0.69**	**0.73**	**0.91**	**0.73**	**0.75**	**0.69**
External debt	0.47	0.52	**0.73**	0.47	0.50	0.30	0.58	0.18	0.50	0.43
Short term ext.debt	**0.67**	**0.66**	**0.87**	**0.71**	**0.72**	0.56	**0.82**	0.25	**0.70**	0.62
Reserves over GDP	0.64	0.54	**0.73**	**0.70**	**0.68**	0.61	**0.76**	**0.64**	0.56	**0.70**
Ext.debt service	0.57	0.58	**0.74**	0.58	**0.65**	0.42	0.52	0.30	0.60	0.52
Portfolio inflows tail2-decline	0.48	0.38	**0.67**	0.51	0.47	0.45	0.43	0.62	0.39	0.57
GDP per capita (change)	0.60	0.50	**0.84**	**0.68**	**0.68**	0.55	0.41	0.58	0.59	0.63
Sov.rating	**0.65**	0.51	**0.88**	**0.73**	**0.74**	0.55	0.34	**0.92**	0.52	**0.79**
US 10Y treasury bond rate	**0.65**	**0.78**	0.56	0.58	0.61	**0.76**	**0.92**	0.30	**0.73**	0.56
US 3M interbank rate	0.62	**0.75**	0.60	0.56	0.57	**0.73**	**0.85**	0.41	0.64	0.53
Int. Trade interconectedness	0.53	0.45	**0.72**	0.57	0.53	0.60	0.17	**0.74**	0.49	0.62

AUROC results: percentage of good signals issued 6 quarters before the onset of each type of crisis. Red bolded cells indicate an AUROC above 70% and orange bolded cells an AUROC above 65%. This table only shows the significant variables.

heterogeneity of our sample, on account of the diverse geographical composition and the different nature of crises. In addition, there seems to be a structural break around 2006 as the number of crises sunk to zero and then a new type of crisis emerge in 2007 (mainly contagion from advanced economies turbulences and banking crises), which can have affected the type of variables that issue anticipatory signals. For all these reasons, the exercise has been carried out distinguishing between banking, currency and sovereign crises. Moreover, we estimate another Auroc splitting the sample into EME regions, and before and after the third quarter of 2006.

Results for banking crises seem to be coherent with the theoretical background. The net foreign position of domestic banks (a proxy for an "excessive" reliance on external funding and for balance sheet currency mismatches) and the loan to deposit ratio (which again proxies a higher reliance on market funding instead of traditional funding) have a ratio of true positive signals close to 70%, as well as the NEER deviation from trend, which could proxy strains in banks' borrowers. Moreover, two variables representing the cost of external funding and external financial conditions -short and long term interest rates in the US-, tend also to anticipate EMEs banking crises, as well as short term domestic rate increases. Finally, non-performing loans have a noise to signal ratio of around 62%[28].

In the case of sovereign crises there are a large number of variables issuing adequate signals (17 out of 48 show a percentage of right signals above 65%, 12 of which have a percentage above 70%). Sovereign ratings and the sovereign spread deteriorates in anticipation to a public debt default. Moreover, activity growth variables deteriorates before a sovereign crisis, which can be a proxy for tax base shrinkages. On the fiscal side, relevant variables seem to be the level and increase of public debt instead of the public sector balance. As a large part of public debt of those countries were placed abroad, variables related to the level of external debt, such as maturity and debt service, are also leading indicators. In this sense, short term domestic rate also presents a ratio of good signals above 70%, probably as it is influenced by monetary authorities to limit capital outflows that could lead to a depreciation of the currency. As reserves constitutes the last line of defence in case of strains derived from its external debt,

[28] For our two tail risk variables we have presented the result of an increase in real credit growth and huge portfolio inflows. Nevertheless, the results for the other tail risk (that could detect the appearance of domestic credit bubbles financed in part with foreign short term capital) are also non-significant (38% and 41% of good signals, respectively).

their level is also relevant to anticipate sovereign crises. Finally, the variable that measures the interconnectedness of EMEs –trade links- also seems to be significant.

Currency crises seem to be anticipated by an increase in sovereign spreads and short term interbank rates, as well as by short term external debt and an acceleration of public debt. A decrease in reserves and a fall in activity growth are also relevant. Surprisingly the overvaluation of NEER does not issue right signals (58%), although the inflation rate could proxy the loss of competitiveness in fixed exchange rate regimes like those of the 90s. Auroc results also point to a high correlation between currency and sovereign crises, as they share some common determinants, though with a slightly higher AUROC ratio for public debt, sovereign rating and spreads and reserves in the latter. As predicted by first generation currency crisis models, an increase in public deficit and public debt can lead to a sudden stop of capital inflows and a strong depreciation of the currency, which ultimately can derive in a sovereign default, as a large part of public debt is denominated in foreign currency and hold by external investors.

Turning to the analysis by region, Latin American crises seem to be well anticipated by a deterioration of public sector balance, an increase of short-term external debt and its service, a decrease of international reserves and a deterioration of the macro environment (higher inflation and lower activity growth). These factors seem to fit well first generation currency crisis models[29]. For Eastern European countries, banking sector variables (a fall in real credit and deposits growth and an increase of short term interest rate) tend to better anticipate the outbreak of crises, along with two global variables and total reserves[30]. Developing Asian crises are correctly signalled by both banking sector variables (those related to currency mismatches as the net foreign position of domestic banks; or those referred to an excessive leverage in financial markets like the loan to deposit ratio)) and external disequilibria variables (current account balance, reserves)[31]. For the five remaining emerging countries (Rest of EMEs), the picture is much less clear, global and financial market variables, an aggregate evaluation of the economies (sovereign rating) or some variables reflecting the macro environment (public sector accounts, reserves and inflation rate) issue good signals[32].

Finally, as mentioned above, from 2006q3 to 2008q2 none of the countries in our sample suffered any kind of crisis. Splitting the sample before and after 2007q1 implies that after 2007 the inflation rate, the sovereign rating and the level of international reserves gain predictive power. The short term interbank rate remains one of the best predictors for any kind of crisis

Summing up, the heterogeneous nature of our sample leads us to divide the dummy by type of crisis instead of just using a dummy for all crises or splitting the sample by region, so that the results could improve substantially[33]. According to the parametric exercise, the short term domestic rate tends to correctly anticipate vulnerable states, especially those related to the banking ones. Public sector disequilibria, both at the domestic and the external side, are relevant to determine the pre-crisis state that could lead to sovereign crises, which seem to be highly correlated with currency crises. Finally, variables measuring the deterioration of debt stocks, rather than flow variables seem to be more useful to anticipate stress periods.

[29] Latin America registered 29 crises from 1993q1 to 2018q4, of which 16 were currency crises (8 of them before 2006q4).

[30] Eastern European countries present by far the large number of banking crises on the sample, around 142 quarters, that is, 24 quarters by each country.

[31] Almost all crises for this group of countries are banking or currency crises, and all are dated back before 2000q1.

[32] The five countries included in this cluster are probably the most heterogeneous of the sample, and for example the standard deviation of GDP growth and of GDP per capita is the highest of the four groups. Moreover the number of missing values is also the highest (almost 10% of the total possible data).

[33] For the sake of robustness, we have build SHERLOCs using the regional results, but out of sample validation is much poorer.

4.2 Building the SHERLOC

Once the variables have been preselected using a signalling approach (AUROC above 65%), we propose four different methodologies to aggregate the relevant variables and build the SHERLOC. First, we estimate the predicted probability of being in a pre-crisis (or vulnerable) state (six quarters previous to the crisis) based on a logistic estimation. Second, the predicted probability is estimated from a factor-augmented logit approach. Third we calculate an index using the two first principal components of the set of (preselected) standardized variables. Fourth, we just calculate the mean of the risk percentiles of the relevant variables. Then, we validate these four approaches both in-sample and out-of-sample to select the most appropriate model to anticipate vulnerable countries.

4.2.1 SHERLOC 1.0: LOGISTIC ESTIMATION

The first methodology consists of extracting the predicted probability of being in a vulnerable state using a logistic estimation. The panel logit approach estimates the probability of being in a "vulnerable state" (i.e, six quarters prior to the crisis), which is assumed to follow a logistic distribution that depends on risk factors, that is:

$$\Pr\left(Y_{it}, = 1 | X_{i,t}, \beta\right) = \frac{e^{\left(\alpha + X_{i,t}{}'\beta + \varepsilon\right)}}{1 + e^{\left(\alpha + X_{i,t}{}'\beta + \varepsilon\right)}}$$

where Y_{it} is the period of vulnerability t (six quarters) prior to a crisis in country i. As already mentioned, in the baseline specification, we use an evaluation window up to six quarters before the crisis and distinguish between a sovereign, currency and a banking crisis. $X_{i,t}$ stands for the factors that signal a vulnerable state according to the AUROC approach. In order to deal with multicollinearity issues, we exclude some of the AUROC preselected variables that are highly correlated (i.e. GDP growth and GDP per capita, or reserves in USD billion and reserves over GDP).

Individual (country) effects are incorporated into the model by using random effects, which assume that the country effects have a distribution[34]. The choice of random effects has several advantages. First, it represents a more efficient combination of "within" and "between information". Second, it enables us to exploit the information of all the countries in the sample, even those which have not suffered any type of crisis. On the contrary, in the case of using a fixed effect model, which is a Conditional Logit model, it eliminates the countries that have never faced a crisis. Understanding why some countries have never suffered a crisis is also relevant in our analysis. However, in order to use a random effect model, one needs to assume that the individual effects are not correlated with the independent variables[35]. As a robustness check, we also estimate the logit using pooled data (See Appendix IX)

In general the results of this multivariate analisys are coherent with those of the univariate (AUROC) estimations. For banking crises, the most relevant factors to explain a pre-crisis state are the proxy for global financial conditions (US long term interest rate) and the loan to deposit ratio, which points to an excessive leverage of domestic banks. Both are closely related with the results for net foreign assets of domestic banks. Finally, the negative sign of

[34] While in the fixed effects, it is assumed to be fixed.

[35] The fixed effect model is also better at minimising the "omitted variable" bias.

Results of the logistic estimation[36] are posted on Table 3:

Table 3

VARIABLES	Banking	Sovereign	Currency	Crises (any kind)
GDP growth		**-0.206*****	0.012	-0.014
		[0.056]	[0.024]	[0.023]
Inflation rate	-0.031***		-0.003	0.000
	[0.008]		[0.004]	[0.006]
NEER deviation from trend	**0.028*****			
	[0.006]			
Public debt (first difference)		0.057	**0.058*****	
		[0.053]	[0.086]	
Public sector balance			**-0.311*****	**-0.143*****
			[0.044]	[0.036]
Net Foreign assets dom banks	**-0.038*****			
	[0.007]			
Loan to Deposit ratio	**0.693*****			
	[0.243]			
Short-term interbank rate	**0.026*****	**0.065*****	**0.059*****	**0.059*****
	[0.010]	[0.022]	[0.009]	[0.010]
External debt		**0.038****		
		[0.018]		
Short-term external debt	-0.001	-0.035	**0.034*****	0.010
	[0.002]	[0.031]	[0.011]	[0.007]
Reserves (mm USD)		**-0.152*****	**-0.006*****	**-0.004*****
		[0.038]	[0.002]	[0.001]
Ext.debt service		-0.002		**0.011****
		[0.009]		[0.004]
Foreign Direct Investment				**-0.212*****
				[0.061]
Rating		-0.217	0.209***	0.194***
		[0.138]	[0.068]	[0.066]
US 10 Y interest rate	**0.689*****			
	[0.166]			
US 3 M interest rate	0.153			
	[0.109]			
Trade links		**0.237*****		
		[0.082]		
Observations	2197	2270	2407	2014
Number of id	25	25	25	25

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Results of a logistic estimation using as regressors the variables with the highest AUROC

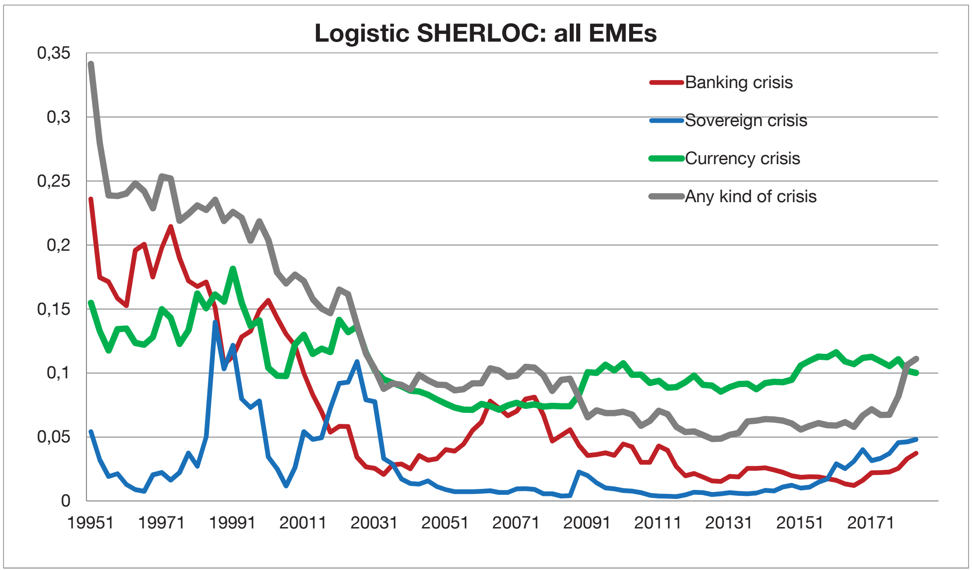
the inflation rate can be explained by the fact that, in high inflation environments, banks have an advantage over their clients in the management of their assets and liabilities, for example lending at variable rates and taking deposits at fixed rates or longer maturities. For sovereign crises, a drop in activity and international reserves, an increase in short term domestic rates or external debt are highly relevant, as well as the trade links between EMEs. As in the AUROC approach, currency crises are closely related with sovereign crises, although in this case flows

[36] All the logistic regressions in this section use a random effects model with standard errors derived from asymptotic theory (OIM) and include a constant term. The use of random effects is validated by a standard Hausmann test.

(public sector balance) also play a role[37,38]. The results for any kind of crisis summarize fundamentally those obtained for sovereign and currency crises, and add the most stable capital inflows, foreign direct investment, with the expected sign.

SHERLOC 1.0 is then calculated as the predicted probability of being in a pre-crisis state (6 quarters before a crisis) derived from the results of the logistic regressions reported in table 3. Chart 4 depicts the Logistic SHERLOC for the whole sample for each type of crisis, and chart 5 the SHERLOCSs by region:

Chart 4



SHERLOC built using the estimated probability of being in a vulnerable state from the logistic regression in Table 4.2.1.1
Simple average of country's SHERLOCs

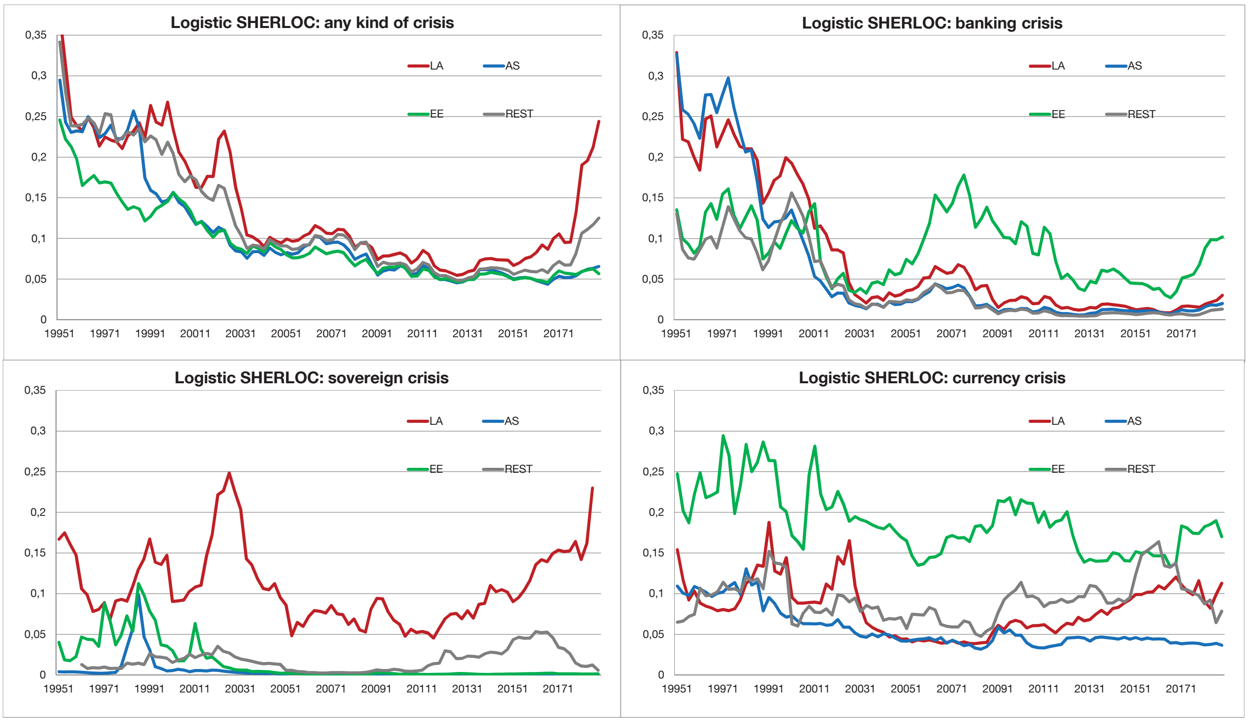
Even in this aggregate form[39], the SHERLOC seems to capture the stress previous to the successive crises of the late 90s, the tranquil times between 2006 and 2009, and the recent increase in vulnerability in some EMEs, especially in Latin America and Eastern Europe. Sovereign stresses are close to historical lows except in Latin America, as well as banking stress, with the exception of Eastern Europe.

[37] The results shown in the table constitute the baseline for building the SHERLOC, although we have tested also the reliability of other specifications (see Appendix II for details).

[38] We have excluded the sovereign spread as it tends to be a simultaneous indicator of crisis instead of being a leading variable. Indeed, the ratio of good signals increases monotonously as we get closer to the crisis date, which is not the case for the rest of variables. In addition, the sovereign differential has the highest pairwise correlation on average with the rest of the regressors, and also the highest among them (-0.62 with the rating). Finally including the sovereign spread in the regression of currency and sovereign crises hardly vary the results.

[39] The aggregate SHERLOC is the simple average of country's SHERLOCs. Using a weighted average (via GDP in PPP terms or the number of crises of each country, to capture those more prone to register turbulences) does not alter the picture.

Chart 5



SHERLOC built using the estimated probability of being in a vulnerable state from the logistic regression in Table 4.2.1.1. LA: Latin America, AS: Asia, EE: Eastern Europe and REST: Rest of countries
Simple average of country's SHERLOCs

4.2.2 SHERLOC 2.0: FACTOR AUGMENTED LOGIT ESTIMATION

Another alternative to estimate the predicted probability of being in a vulnerability state consists of relying on a factor-augmented logit approach. This methodology enables us to exploit all the available information in a dense model by using as explanatory variables the common factors extracted from the largest dataset that we have built. The underlying idea is to capture all the fundamental drivers of the economy that might lead to a vulnerable state. To do so, we first extract the common factors using a principal component analysis and incoporate all the factors with an eigenvalue higher than one in a logistic estimation[40]. The probability of being in a "vulnerable state" (i.e, six quarters prior to the crisis) is, therefore, assumed to follow a logistic distribution that depends on common factors, that is:

$$\Pr(Y_{it}, = 1|X_{i,t}, \beta) = \frac{e^{(\alpha + X_{i,t}{}'\beta + \varepsilon)}}{1 + e^{(\alpha + X_{i,t}{}'\beta + \varepsilon)}}$$

where Y_{it} is the period of vulnerability t (six quarters) prior to a crisis in country i. $X_{i,t}$ stands for the six common factors extracted from the large dataset based on a principal component analysis[41].

In the factor-augmented logit estimation, we follow the methodology proposed by Chen et al (2011) due to two main reasons. First, our dummy variable is equal to one the six quarters previous to a crisis, and therefore we capture all the indicators that might affect crises from one lag to six lags prior to the onset of the crisis. In this sense, we do not consider

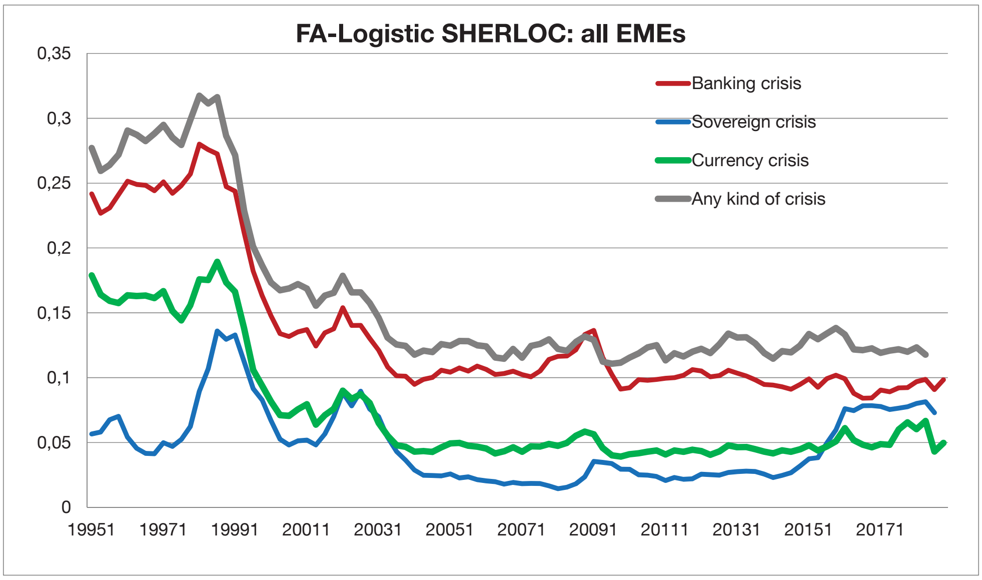
[40] We end up wiht six common factors that account for 60% of the data variation among all the variables included in the principal component analysis.

[41] The logistic regressions use a random effect model with standard errors derived from asymptotic theory (OIM) and include a constant term.

necessary to include dynamics as Bellégo and Ferrara (2012) do[42]. Second, as pointed out by Chen et al (2011), the extracted factors from the PCA "may already incorporate the lags of underlying dynamic factors".

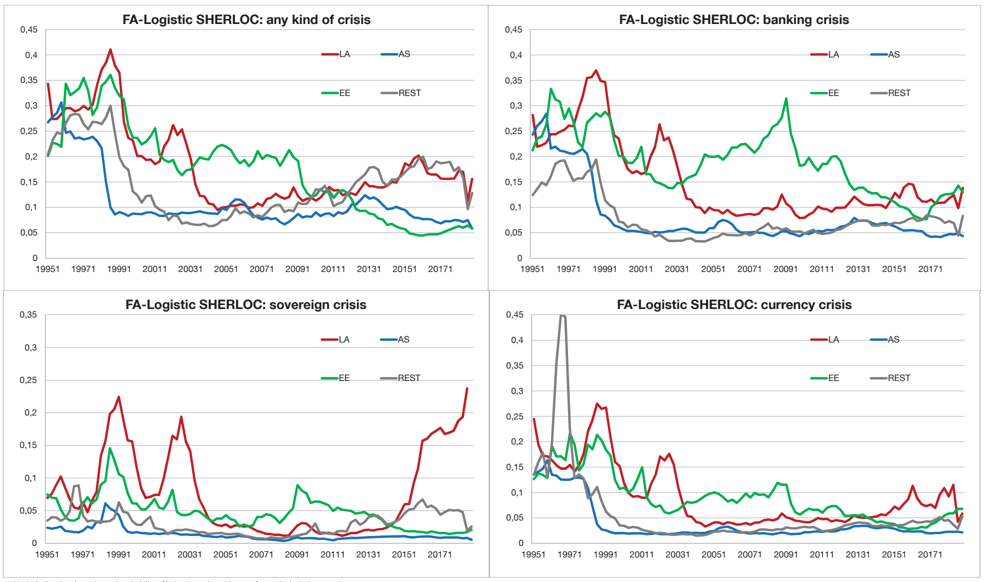
Charts 6 and 7 show the developments of vulnerabilities based on the FA-logistic approach. Results are qualitative very similar to the previous logistic SHERLOC: Latin America is the most vulnerable region nowadays due to an increase in sovereign stress and in currency stress, which has also increased recently in Eastern Europe.

Chart 6



SHERLOC built using the estimated probability of being in a vulnerable state from a FA-logistic regression
Simple average of country's LIARs

Chart 7



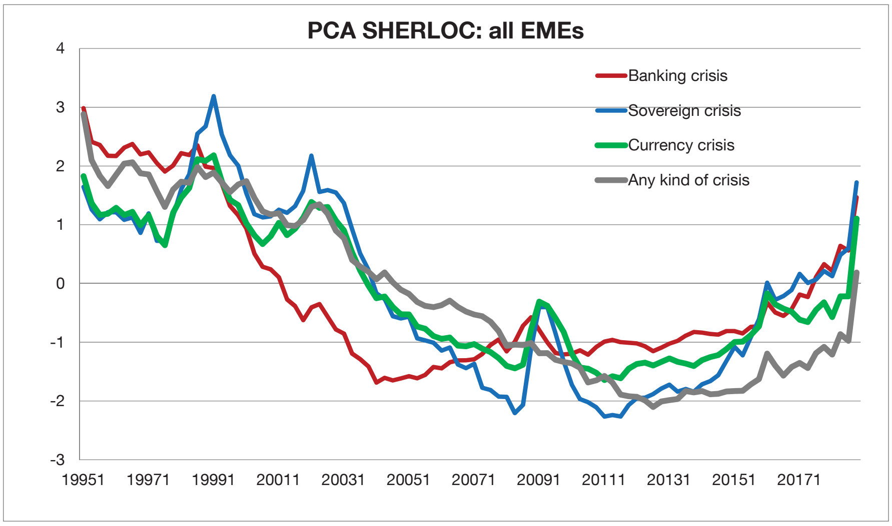
SHERLOC built using the estimated probability of being in a vulnerable state from a FA-logistic regression
Simple average of country's SHERLOCs

[42] In their case, the dummy only takes value 1 when the economy is in recession and, hence, they need to include the dynamic effects in the explanatory variables

4.2.3 SHERLOC 3.0: PRINCIPAL COMPONENTS ANALISYS

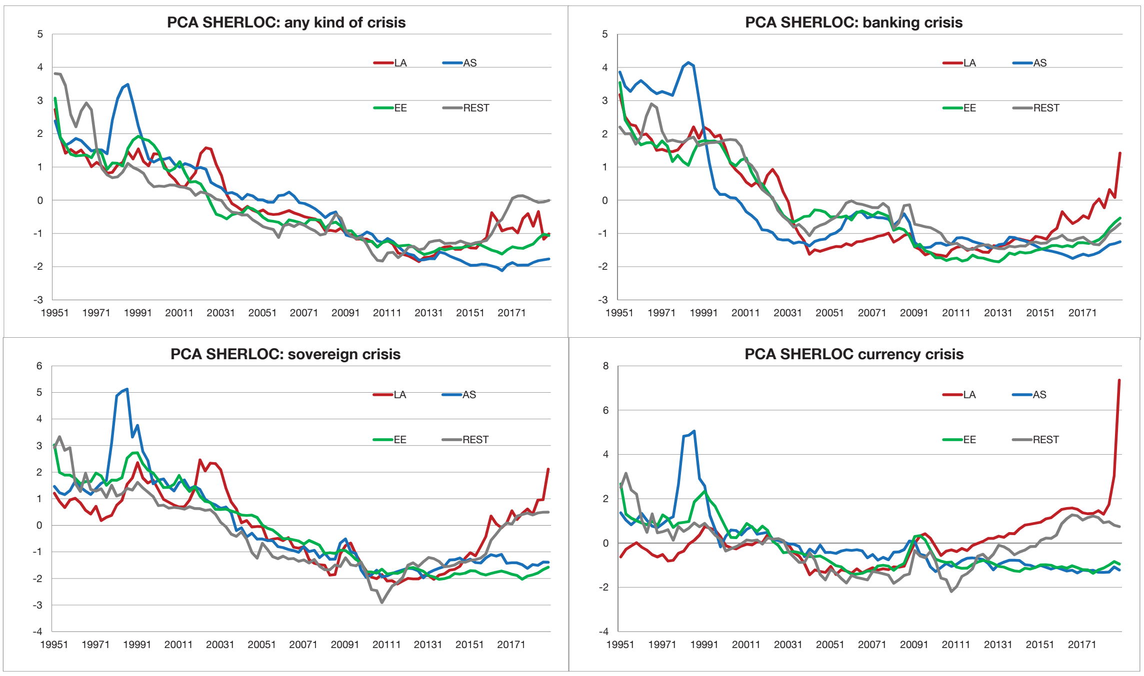
Another way to build the SHERLOC is to synthetize the information contained in the variables that issue the best signals into a single index using principal component analisys techniques, once the indicators are standardized. Unlike the Logistic SHERLOC, the PCA SHERLOC simply indicates whether the vulnerability in a country at a moment of time is greater or lower than its historical average, and it is comparable for a country over time but not between countries. Finally, the possible advantage of the SHERLOC PCA could be that it eliminates the redundant information generated by the correlations between the selected variables, which could bias some results of the logistic estimation. In addition, it enables us to previously assign the expected sign of the relationship between the selected variables and the vulnerability. Again, charts 8 and 9 show the aggregate results for the PCA SHERLOC:

Chart 8



SHERLOC built using 2 principal components of preselected AUROC variables
A value above zero implies a level of risk above historical average

Chart 9



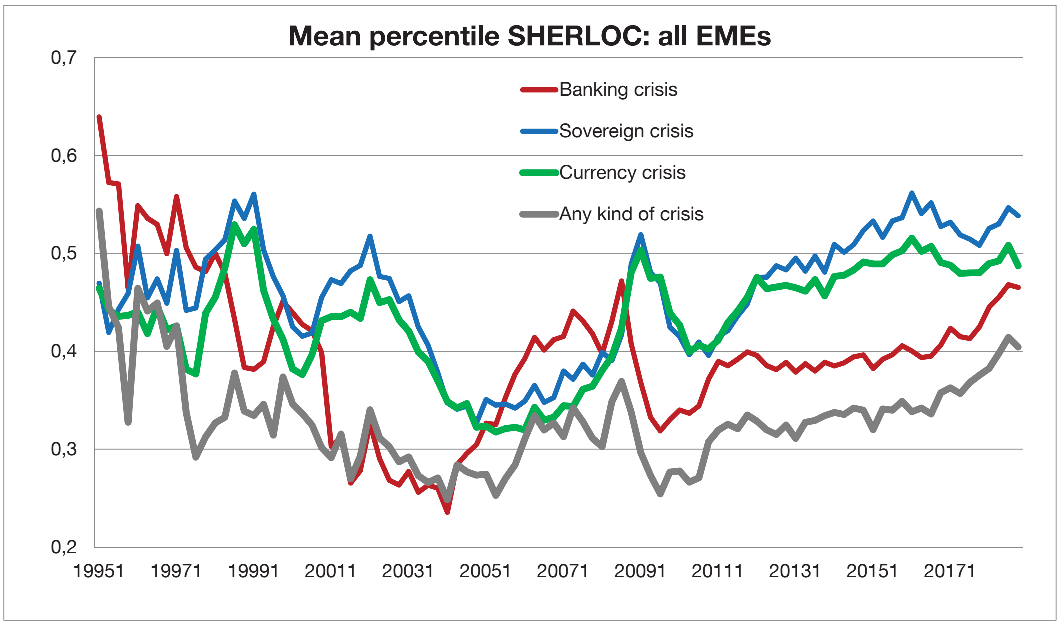
SHERLOC built using 2 principal components of preselected AUROC variables
A value above zero implies a level of risk above historical average

Vulnerability related to banking and sovereign crises has substantially increased in recent quarters. The case of Latin American countries is striking, as the region increased their vulnerability related to currency crisis to reach historical highs. Sovereign and banking vulnerability also rose to levels similar to those of the beginning of the 2000s.

4.2.4 SHERLOC 4.0: MEAN PERCENTILES OF RISKS:

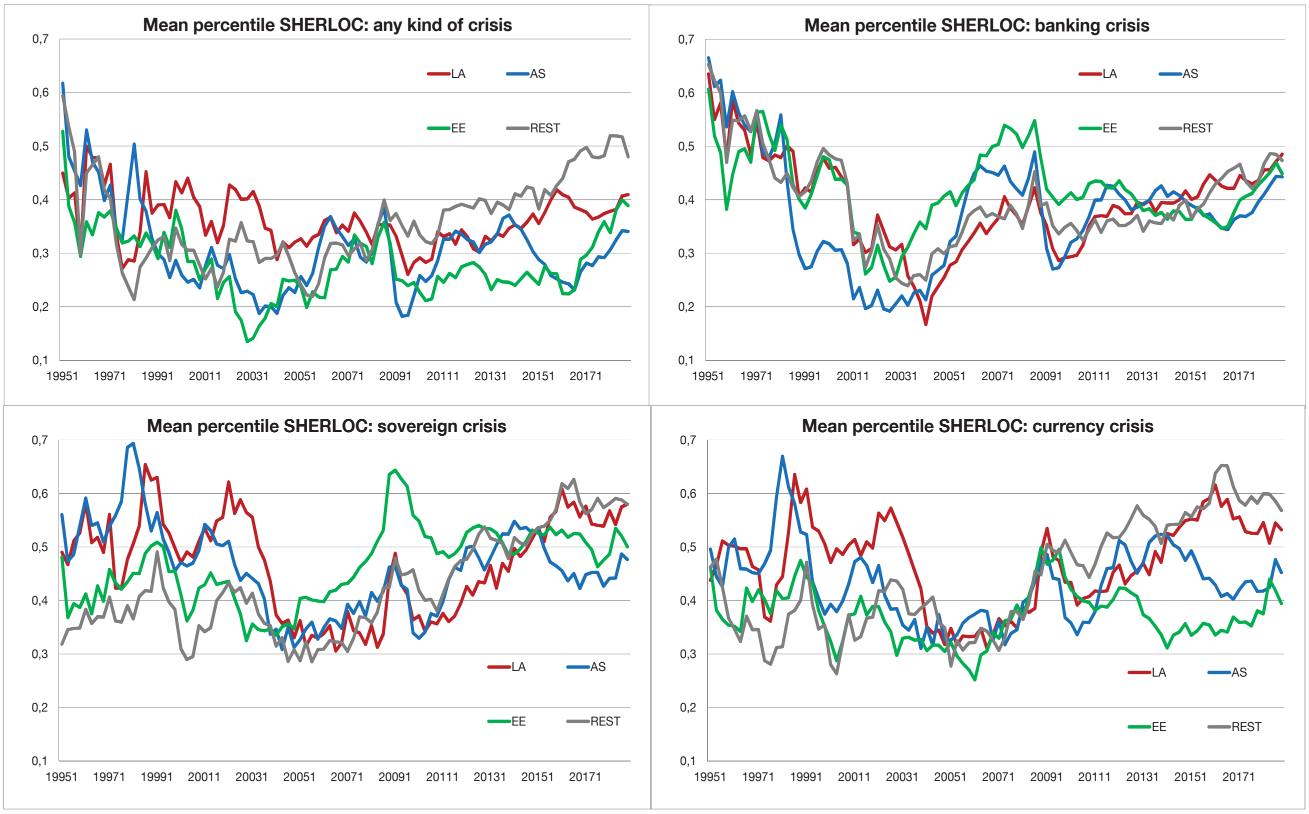
Finally an easier way to compute the SHERLOC would be to calculate the mean of the percentiles of risks for the pre-selected variables. To do this we estimate the frequency distribution of the historical series of the preselected variables, excluding crisis times (to avoid possible biases) and then we calculate the average of the percentile position of each indicator. The advantage of this index lies in the fact that it standardizes all the relevant information using percentiles, bounded between 0 and 1. Therefore, it directly shows the level of risk at each moment of time in comparison with the level of risk over time. Its interpretation is straitghforward: a SHERLOC below 50% represents a level of risk below its historical median. Charts 10 and 11 present the Mean Percentile SHERLOC for the whole sample for each type of crisis and by region, respectively. Currency and sovereign stress vulnerability for all EMEs are slightly above the historical median, mainly on account of the recent development of vulnerability in Latin American countries and the rest of EMEs (Middle East and African countries).

Chart 10



SHERLOC built using percentiles of risks of preselected AUROC variables. Average of individual percentiles

Chart 11



SHERLOC built using percentiles of risks of preselected AUROC variables. Average of individual percentiles

4.3 Validating the SHERLOC

In order to assess the performance of our synthetic indexes in predicting "pre-crisis" events, we follow the methodology proposed by Alessi and Detken (2011), which develop a measure of Absolute Usefulness based on policy maker's preferences. This methodology has also been used by Lo Duca and Peltonen (2013) to validate different models to predict systemic financial crisis and Babecký et al (2012) to assess the performance of early warning indicators of debt crises. An index is considered to be "useful" whenever there is a gain in using this index as compared to ignoring it. In other words, the loss of not using the index is larger than the loss of using the index.

Assume that $L(\theta)$ is the loss function when using the index proposed, which is dependent on policy makers' preferences between type I and type II errors:

$$L(\theta) = \theta\ (Type\ I) + (1-\theta)(Type\ II)$$

θ captures policy maker's preferences between type I and type II errors. While Type I error calculates the share of missed vulnerable periods in terms of total number of stress, Type II errors represents the ratio of false alarms issued in terms of "tranquil periods". A θ larger than 0.5 reveals that a policy maker prefers receiving a false alarm than missing a crisis.

The measure of Absolute Usefulness (U) of Alessi and Detken (2011) is formally defined as:

$$U = \min[\theta; 1-\theta] - L(\theta)$$

Where $\min[\theta; 1-\theta]$ is the loss that the policy maker incurrs when ignoring the index (and therefore she either always assumes there is a signal and has to react if θ >0.5 or there is never a signal and does not react if$\theta < 0.5$); and $L(\theta)$ is the loss function when using the index proposed, previously defined. In our framework, we suppose that policy makers are as concerned about missing a crisis as they are about issuing a false alarm, and therefore we set$\theta = 0.5$.

Optimal thresholds are calculated by maximizing the measure of Absolute Usefulness (U). Additionally, we rank the indexes obtained using different methodological approaches according to its "Usefulness". If this measure is positive, it provides useful signals for policy makers compared to ignoring it. The larger the measures, the higher the benefit.

This methodology has several advantages. First, it takes into consideration policy makers' preferences with regards missed crises versus false alarms. Second, it enables us to calculate the optimal threshold beyond which an early warning signal is issued. Thresholds are optimized for each index for the given preference parameter θ = 0.5 using all the information in the evaluation sample in order to get the percentile of the distribution that maximizes the Absolute Usefulness measure. Finally, it is quite intuitive and easily understood by policymakers. An index or model is "useful" whenever there is a benefit of using the index to detect vulnerable states or a pre-crisis periods.

4.3.1 IN SAMPLE PERFORMANCES OF THE SHERLOCS:

First, we evaluate the performance of the different indexes in predicting a "vulnerable" or a "pre-crisis" state using the sample over the period 1993 to 2018. For each index proposed, we calculate the threshold for the estimated likelihood of a "vulnerable" state that maximizes the usefulness score.

Table 4 reports the evaluation in-sample for the banking, sovereign and currency indexes proposed and an aggregate index for all type of crises, based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crisis (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%)). More specifically, this table presents the performance of the indexes for the four different methodological approaches used to build the indexes: the predicted probability based on a logistic estimation (logistic SHERLOC) and on a factor-augmented logistic approach (FA logistic approach), a principal component analysis from the variables preselected using the AUROC (PCA SHERLOC) and the simple average of the percentiles of the relevant variables according to the AUROC[43] (mean percentile SHERLOC). Finally, for the sake of comparison, we also include the performance of the best single indicators derived from the AUROC: the short-term interbank rate in the case of the banking crises; and the sovereign spread for sovereign and currency indexes. The variation of reserves is also assessed for currency crises as it has been an early warning indicator broadly used in the literature, as well as the stock of banks and deposits in the case of the banking crises[44].

---

[43] We preselect the variables with an AUROC above 65 in the case of banking and currency crises, and above 70 for sovereign crises.
[44] Results can be provided under request.

As expected, all models achieve a positive Usefulness measure, which suggests that the indexes proposed provide gains for policy makers. Indeed, the usefulness ratio found are similar to those reported by Alessi and Detken (2011) in a univariate setting and for the same preference parameter (θ) and by Babecký et al (2012) relying on a composite early warning index. Both papers find similar usefulness values of around 0.15-0.25[45]. Lo Duca and Peltonen (2013) obtain slightly larger Usefulness values than these two papers, with values around 0.19-0.34, in line with the results of the Logistic SHERLOC presented in this paper for each type of crisis (0.20-0.33).

More significantly, all the models proposed outperform their respective best single indicators: the short-term interbank rate for banking crises and the sovereign spread for sovereign and currency crises. Therefore, the use of a composite early warning index (the SHERLOC) seems to be more accurate to anticipate vulnerable states of countries than the use of a single indicator, which accords with the results of Babecký et al (2012) and Lo Duca and Peltonen (2013). However, in this paper we also show that the creation of an index for each type of crisis seems to outperform an aggregate index for all types of stress. Indeed, the usefulness values when using an aggregate index for all crises are much lower than the results of the SHERLOC, regardless of the methodology used and the type of crisis. The underlying reason is the wide heterogeneity of the nature of crises, which therefore should be treated in a different way.

Table 4: In-sample performance of indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
Banking indexes					
Logistic SHERLOC	75	0.33	0.24	86%	17%
FA Logistic SHERLOC	72	0.30	0.29	84%	15%
PCA SHERLOC	66	0.25	0.36	78%	12%
Mean percentile SHERLOC	76	0.20	0.37	63%	11%
Short-term interbank rate	56	0.19	0.50	79%	9%
Sovereign indexes					
Logistic SHERLOC	72	0.27	0.32	79%	6%
FA Logistic SHERLOC	75	0.24	0.32	71%	6%
PCA SHERLOC	90	0.23	0.16	54%	12%
Mean percentile SHERLOC	85	0.19	0.27	52%	7%
Sovereign spread	55	0.17	0.57	77%	3%
Currency indexes					
Logistic SHERLOC	57	0.20	0.50	81%	10%
FA Logistic SHERLOC	74	0.26	0.31	74%	16%
PCA SHERLOC	71	0.20	0.39	67%	13%
Mean percentile SHERLOC	79	0.15	0.39	50%	13%
Sovereign spread	55	0.08	0.74	59%	8%
All crises					
Logistic SHERLOC	74	0.14	0.43	50%	18%
FA Logistic SHERLOC	80	0.22	0.26	60%	27%
PCA SHERLOC	64	0.16	0.47	60%	17%
Mean percentile SHERLOC	55	0.10	0.67	63%	12%

This table presents the in-sample validation exercise for the banking, sovereign, currency indexes proposed and an aggregate index for all types of crises (all crises) and the four different methodological approaches used to build the indexes: the predicted probability from a logistic estimation (Logistic SHERLOC) and from a factor augmented logistic estimation (FA Logistic Sherloc), a principal component analysis from the variables (PCA SHERLOC), and the average of the risk percentiles(mean percentile SHERLOC). The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%). The threshold indicates the percentile beyond which the index issues a signal. Short-term interbank rate and sovereign spread represent the best single indicators according to the AUROC. This exercise is based on a neutral policy maker (θ =0.5) and evaluation window six quarter previous to the crisis.

Additionally, the usefulness measures enables us to rank the different indexes proposed. The higher the usefulness measure, the better. The Logistic SHERLOC seems to outperform the three other methodologies used (the FA logistic SHERLOC; the PCA SHERLOC and the mean percentile SHERLOC) as it achieves higher usefulness scores for the banking and sovereign indexes[46]. In addition, the ratio of predicted crises are very high for the three types of crises considered (around 87% for banking crises, 79% for sovereign crises and 81% for currency crises).

Finally, comparing the Logistic SHERLOC validation exercise with the outcome of previous literature, our index seems to outperform some of the indexes or variables proposed. Indeed, the usefulness measures of our analysis are similar or higher than the ones reported by Alessi and Detken (2011) and Babecky et al (2012). They found values around 0.20 and 0.25, and 0.20 respectively. In the case of our Logistic SHERLOC, we found values of 0.33, 0.27 and 0.20 for banking, sovereign and currency, respectively. Therefore, our index seems to be better to predict banking and sovereign crises at least. However, results are not strictly comparable since their analysis focus on developed economies, instead of emerging economies, over a different period of analysis. In addition, Alessi and Detken (2011) predict asset booms instead of crises. If we compare it with Lo Duca and Peltonen (2013), they find similar values for a mixed sample of advanced and emerging economies. The percentage of predicted crises is also similar. But, they only focus on financial stress instead of taking a broader perspective and considering all types of stress. Finally, Lepers and Sánchez-Serrano do not provide any evaluation methodology, although they point out that their results are better than the usual credit to GDP gap. Nevertheless, the creation of asset bubbles as a vulnerability is captured in our analysis using two tail risks for real credit growth, but the AUROC results lead us to discard it.

Therefore, the SHERLOC proposed will be based on the logistic estimation. However, as the PCA SHERLOC, whose interpretation is more straightforward and its update is simpler, also seems to perform adequately, it can be used as a robustness check. The FA logistic SHERLOC seems quite appropriate for currency crises. The Mean percentile SHERLOC will be discarded as its performance does not seem to be good enough.

4.3.2 OUT OF SAMPLE PERFORMANCES OF THE SHERLOCS:

Second, in order to assess the predictive ability of each index out-of-sample, we split our sample in two subsamples. The period from 1993 to 2007 is used to estimate our models and calibrate the optimal threshold beyond which each index issues an early warning signal. The last years of the sample (2008-2018) are used to estimate the performance out-of-sample[47].

Table 5 presents the results of the out-of-sample validation for the different indexes proposed for the banking, sovereign and currency stress and an aggregate index for all type of crises. As expected, out-of-sample performance is slightly poorer than in-sample performance although usefulness measures remain positive and higher than the best single predictors, with the exception of sovereign crises and the FA logistic SHERLOC for currency and sovereign crises. Indeed, usefulness values hover around 0.06 and 0.15, which is somewhat in line with

[46] The FA logistic SHERLOC achieves a higher usefulness for currency crises.

[47] As we noted before, there seems to be some structural changes around 2006 as the number of crises dwindles (reaching zero in some quarters). Using 2007 as the break point enables us to have more crises out-of-sample to validate the SHERLOCs. Nevertheless we have also tested the performance of SHERLOCs from 2014 onwards, and the results (provided upon request) do not change significantly. Indeed, in some cases, the model improves in terms of performance, for instance for banking crises.

the out-of-sample results reported by Lo Duca and Peltonen (2013). As it was the case in in-sample predictions, the SHERLOC indexes perform much better than the best single indicators and the use of an index for each type of crisis outperforms an aggregate index of all crises in terms of usefulness[48].

Finally, when comparing between the different methologies employed to construct the SHERLOC, the results are mixed. The Logistic SHERLOC performs adequately for currency and banking crises, as suggested by the usefulness measures and the percentage of predicted crises, 49% and 58%, respectively. However, the usefulness value for the sovereign crisis (0.08) is lower than the scores obtained with the PCA and mean percentile (0.12). Still, the percentage of predicted crises (50%) is higher or equal to the percentage predicted by the PCA and mean percentile. In addition, one needs to be aware that there were only three episodes of sovereign crises over this period, including the "hold-out" crisis in Argentina, which is a particular case, difficult to anticipate. In addition, the PCA SHERLOC also provides a positive usefulness measure out-of-sample, which suggests that it can also be an adequate measure to anticipate vulnerable states. The FA logistic SHERLOC only provides substantial gains in the case of currency crises, while its out-of-sample forecast performance is very low in the case of banking and sovereign crises. Therefore, this methodology can be used as a robustness check for currency crises. Finally, the Mean Percentile SHERLOC provides a low gain in the case of the currency index (0.06) and a poor performance in-sample, and, therefore, it is discarded.

Table 5: Out-of-sample performance of indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
Banking indexes					
Logistic SHERLOC	70	0.14	0.53	58%	4%
FA Logistic SHERLOC	67	0.04	0.81	42%	3%
PCA SHERLOC	51	0.13	0.65	75%	3%
Mean percentile SHERLOC	70	0.14	0.54	58%	4%
Short-term interbank rate	53	0.07	0.77	63%	3%
Sovereign indexes					
Logistic SHERLOC	69	0.08	0.68	50%	2%
FA Logistic SHERLOC	70	0.01	0.93	33%	2%
PCA SHERLOC	80	0.12	0.45	44%	3%
Mean percentile SHERLOC	74	0.12	0.52	50%	3%
Sovereign spread	51	0.11	0.69	72%	2%
Currency indexes					
Logistic SHERLOC	76	0.11	0.55	49%	8%
FA Logistic SHERLOC	70	0.15	0.49	60%	9%
PCA SHERLOC	61	0.08	0.70	54%	7%
Mean percentile SHERLOC	79	0.06	0.66	33%	7%
Sovereign spread	51	0.00	1.01	49%	5%
All crises					
Logistic SHERLOC	66	0.07	0.71	49%	10%
FA Logistic SHERLOC	71	0.07	0.65	43%	11%
PCA SHERLOC	51	0.12	0.65	72%	11%
Mean percentile SHERLOC	64	0.04	0.81	46%	9%

This table presents the out-of-sample-sample validation exercise for the banking, sovereign, currency indexes proposed and an aggregate index for all types of crises (all crises) and the four different methodological approaches used to build the indexes: the predicted probability from a logistic estimation (Logistic SHERLOC) and from a factor augmented logistic estimation (FA Logistic Sherloc), a principal component analysis from the variables (PCA SHERLOC), and the average of the risk percentiles(mean percentile SHERLOC). The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%). The threshold indicates the percentile beyond which the index issues a signal. Short-term interbank rate and sovereign spread represent the best single indicators according to the AUROC. This exercise is based on a neutral policy maker (θ =0.5) and evaluation window of six quarters previous to the crisis.

[48] Although the PCA SHERLOC for all crises also seems to perform adequately and the performance of the FA logistic approach for sovereign and currency crises is quite poor (0.04 and 0.01 respectively).

4.3.3 SUMMING UP THE VALIDATION EXERCISE

In light of the in-sample and out-of-sample exercises, the (preferred) vulnerability index proposed will be based on a logistic estimation as it outperforms other indexes in-sample and performs adequately out-of-sample. As a robustness check, the PCA SHERLOC can also be used since its out-of-sample performance remains good and the FA logistic SHERLOC can be monitored but only for currency crises. The Mean percentile SHERLOC will be discarded as its predictions do not seem to be good enough, especially in-sample.

The high type II errors are a source of concern as suggested by the high noise to signal ratios and the low percentage of conditional probability both in-sample and out-of-sample. Indeed, early signals tend to be noisier, but they are more valuable for policymakers. A possible way to reduce the elevated noise is to take into consideration the "post-crisis bias". Indeed, macroeconomic variables tend to show a strong persistence and therefore they have an "erratic" behavior in the recovery phase. This is one of the issues of binary choice models since the model cannot distinguish between a period where the economy is back to normal times and a period in which the country is still adjusting. As a robustness check, we remove post-crisis years (4 quarters), as done with crisis years, at a cost of fewer observations. We show that considering the "post-crisis" bias is relevant to slightly reduce type II errors out-of-sample, but it does not seem to affect the ratio of crises predicted in-sample and it has only a marginal impact out-of-sample (See Appendix IV).

Finally, our model is robust to different specifications and, broadly speaking, it tends to outperform these specifications. As robustness checks, we validate the performance of our SHERLOC using different logistic estimations: including and excluding some variables (Appendix II), using a different definition of currency crisis (Appendix III), taking into consideration the "pseudo-real time approach" (Appendix V), relying on different evaluation windows (Appendix VI), assuming different preferences of policy makers (Appendix VII), calculating the index by region (Appendix VIII), or using pooled data (Appendix IX).

5 Conclusions and work ahead

Is it possible to detect vulnerabilities in EMEs with enough time to implement measures to tame its effects from a financial stability point of view? In this paper we propose a user friendly tool to monitor vulnerabilities in 25 EMEs based on an index of vulnerabilities (labelled SHERLOC) for each type of crisis. These indexes capture the developments of the leading indicators to anticipate pre-crisis periods, selected using a signalling approach (Auroc) to mitigate the usual data mining issue. Statistical validation techniques suggest that the use of different SHERLOCs for each type of crisis implies a gain of utility with respect to the use of an aggregate index for all crises and with respect to the individual indicators with the best performance according to the Auroc. Results seem to be robust to different specifications of the SHERLOC and different subsamples, so we consider that this tool can be very useful to monitor risks in EMEs.

Nevertheless one needs to be aware of the main pitfalls of the methodology used. First, noise remains elevated, partly due to the "post-crisis bias". A possible solution is to remove post-crisis episodes at the cost of fewer observations. Second, the criticism of "this time is different" also applies. However, as "not every time is different" is also possible (even more probable in EMEs than in other countries), the use of variables that anticipate crises in the past is useful. Third, binary choice models require a sufficient number of stress events to get robust results. To overcome this shortcoming, we can use continuous variables, such as Financial Stress Indicators (FSI) and apply Markov Switching models that endogenously determine the beginning and exit of the crisis, at the cost of not evaluating the "vulnerable" state but the "crisis period". Additionally, the main drawback is to define continuous variables long enough to capture all the events and cover a large period of time. Fourth, the results depend on policymakers' preferences for Type I and Type II errors, and also on the definition of crises. Finally, these models are not able to capture non-linearities. All of them would be interesting topics for future research.

REFERENCES

Alessi, L., & 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., Detken, C., & Oprică, S. (2016). On the Evaluation of Early Warning Models for Financial Crises. In *Handbook of Research on Financial and Banking Crisis Prediction through Early Warning Systems* (pp. 80-96). IGI Global.

Babecký, J., Havranek, T., Mateju, J., Rusnák, M., Smidkova, K., & Vasicek, B. (2012). Banking, debt and currency crises: early warning indicators for developed countries.

Bellégo, C., & Ferrara, L. (2012). Macro-financial linkages and business cycles: A factor-augmented probit approach. *Economic Modelling*, *29*(5), 1793-1797.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Bruggemann, A., Linne, T., & Pankki, S. (2000). Are the central and Eastern European transition countries still vulnerable to a financial crisis? *Results from the Signals Approach. Bank of Finland Istitute for Economies in Transition Discussion Paper*.

Bussiere, M., & Fratzscher, M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, *25*(6), 953-973.

Castro, C. E., Estrada, Á., & Martínez-Pagés, J. (2016). The countercyclical capital buffer in Spain: an analysis of key guiding indicators.

Catão, L. A., & Milesi-Ferretti, G. M. (2014). External liabilities and crises. *Journal of International Economics*, *94*(1), 18-32.

Chen, Z., Iqbal, A., & Lai, H. (2011). Forecasting the probability of US recessions: a Probit and dynamic factor modelling approach. *Canadian Journal of Economics/Revue canadienne d'économique*, *44*(2), 651-672.

Drehmann, M., & Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting*, *30*(3), 759-780.

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

Eichengreen, B., Rose, A. K., & Wyplosz, C. (1995). Exchange market mayhem: the antecedents and aftermath of speculative attacks. Economic policy, 10(21), 249-312.

Eichengreen, B., Rose, A. K., & Wyplosz, C. (1996). Contagious currency crises (No. w5681). National Bureau of Economic Research.

Eichengreen, B. (2002). International financial crises: is the problem growing? *Jahrbuch für Wirtschaftsgeschichte/Economic History Yearbook*, *43*(1), 89-104.

Edison, H. J. (2003). Do indicators of financial crises work? An evaluation of an early warning system. *International Journal of Finance & Economics*, *8*(1), 11-53.

Frankel, J. A., & Rose, A. K. (1996). Currency crashes in emerging markets: An empirical treatment. *Journal of international Economics*, *41*(3-4), 351-366.

Frankel, J., & Saravelos, G. (2012). Can leading indicators assess country vulnerability? Evidence from the 2008–09 global financial crisis. *Journal of International Economics*, *87*(2), 216-231.

Frankel, J. A., & Wei, S. J. (2004). *Managing macroeconomic crises* (No. w10907). National Bureau of Economic Research.

Gadea Rivas, M. D., & Pérez-Quirós, G. (2012). The Failure to Predict the Great Recession-The Failure of Academic Economics? A View Focusing on the Role of Credit.

Gramlich, D., Miller, G., Oet, M. V., & Ong, S. J. (2010). Early warning systems for systemic banking risk: critical review and modeling implications. *Banks and Bank Systems*, *5*, 199-211.

Kamin, S. B., Schindler, J. W., & Samuel, S. L. (2001). The contribution of domestic and external factors to emerging market devaluation crises: An early warning systems approach. *FRB International Finance Discussion Paper*, (711).

Kaminsky, G., Lizondo, S., & Reinhart, C. M. (1998). Leading indicators of currency crises. *Staff Papers*, *45*(1), 1-48.

Kaminsky, G. L., & Reinhart, C. M. (1999). The twin crises: the causes of banking and balance-of-payments problems. *American economic review*, *89*(3), 473-500.

Laeven, L., & Valencia, F. (2008). Sytemic Bank Crises: A new database.

Laeven, Luc, and Fabian Valencia. "Systemic banking crises database: An update." (2012).

Laeven, M. L., & Valencia, M. F. (2018). *Systemic banking crises revisited*. International Monetary Fund.

Lepers, E., & Sánchez Serrano, A. (2017). *Decomposing financial (in) stability in emerging economies* (No. 39). ESRB Working Paper Series.

Lestano, J. J., & Kuper, G. H. (2003). Indicators of financial crises do work! An early-warning system for six Asian countries crises, Univerty of Groningen Working Paper

Molina, L., López, E. and Alberola, E. (2016). An index of external positioning for the Spanish economy, Occasional Papers 1602, Banco de España

Oet, M. V., Bianco, T., Gramlich, D., & Ong, S. J. (2013). SAFE: An early warning system for systemic banking risk. *Journal of Banking & Finance*, *37*(11), 4510-4533.

Peria, M. S. M. (2002). A regime-switching approach to the study of speculative attacks: a focus on EMS crises. In *Advances in Markov-Switching Models* (pp. 159-194). Physica, Heidelberg.

Reinhart, C. M., & Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly*. Princeton university press.

Rose, A. K., & Spiegel, M. M. (2012). Cross-country causes and consequences of the 2008 crisis: early warning. *Japan and the World Economy*, *24*(1), 1-16.

Sarlin, P. (2013). On policymakers' loss functions and the evaluation of early warning systems. *Economics Letters*, *119*(1), 1-7.

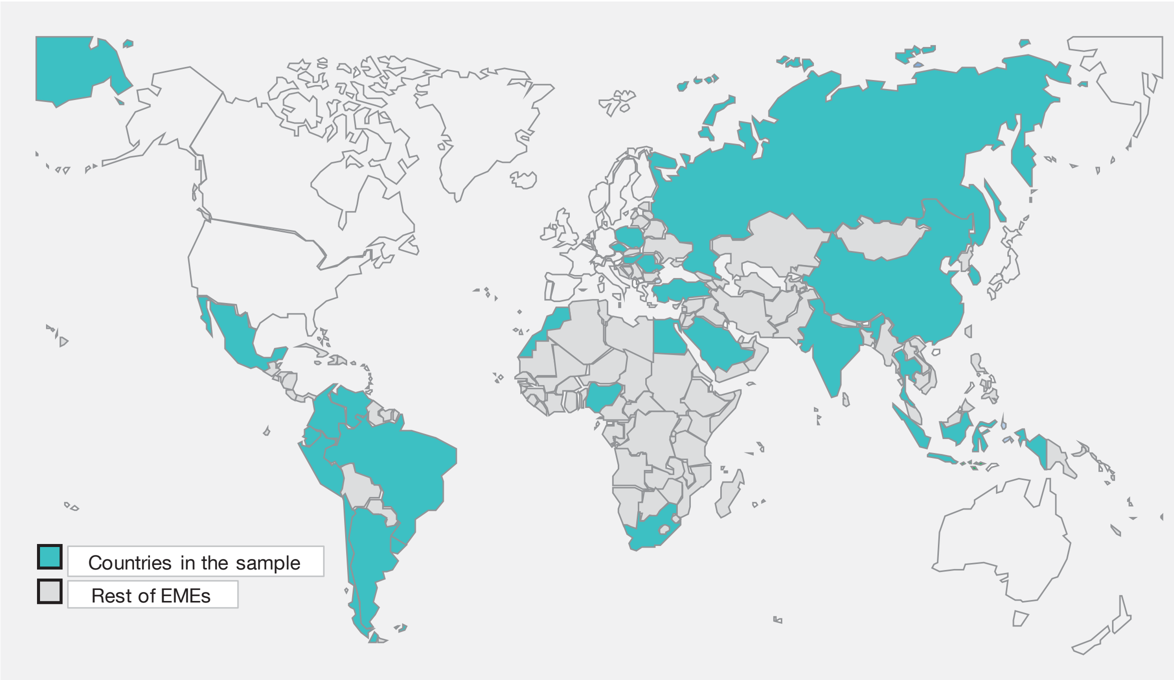
Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. Statistics in medicine, 16(4), 385-395.

APPENDIX I: MAIN FEATURES OF THE DATABASE

1. Countries

Countries included in the sample:

Chart A1.1



2. Variables

Table A1.1

VARIABLE	SOURCE
FINANCIAL MARKETS	
* Sovereign spread (bps level)	JP Morgan EMBI or equivalent
* Sovereign spread (change over 3 months)	JP Morgan EMBI or equivalent
* Stock Exchange index (change over 3 months)	Thomsom Reuters
* Exchange rate vis a vis the USD (change over 3 months)	Thomsom Reuters
MACROECONOMIC FUNDAMENTALS	
<u>REAL SECTOR:</u>	
* GDP (change y-o-y)	National Statistics, Oxford
* Inflation rate	National Statistics
* Industrial production (12 month MA, y-o-y)	National Statistics
* NEER overappreciation	JP Morgan and own calculations
<u>FISCAL SECTOR:</u>	
* Public sector balance (% GDP)	National Statistics, Oxford, IMF IFS
* Public sector gross debt (% GDP)	National Statistics, Oxford, IMF IFS
<u>EXTERNAL SECTOR:</u>	
* Current account balance (% GDP)	National Statistics, Oxford, IMF IFS
* Gross external debt (% GDP)	National Statistics, Oxford, IMF IFS
* FDI (% GDP)	National Statistics, Oxford, IMF IFS
* Short term external debt (% Reserves)	National Statistics, Oxford, IMF IFS
* Reserves (% GDP)	National Statistics, Oxford, IMF IFS
* External debt service (% exports)	National Statistics, Oxford, IMF IFS
* Portfolio gross inflows (% GDP)	National Statistics, Oxford, IMF IFS
<u>BANKING SECTOR:</u>	
* Real credit to private sector (y-o-y)	IMF IFS
* Real deposits on domestic banks (y-o-y)	IMF IFS
* Loan to Deposit ratio	IMF IFS
* Non performing loans (% total loans)	National Central Banks, World Bank
* Net foreign assets of domestic banks (% GDP)	IMF IFS
* Bank Stock Exchange (change over 3 months)	Thomsom Reuters
* Spread of bank's external debt (3 months change)	JP Morgan
* Short term interbank rate (%)	National Central Banks
* Intermediation margin (loan rate - deposit rate)	National Central Banks
WEALTH AND INSTITUTIONAL QUALITY	
* Per capita GDP (USD PPP and % change)	World Bank
* GPR index	Dario Caldara and Matteo Iacoviello
* Sovereign rating (average of the three main agencies)	Standard and Poor's, Fitch, Moody's
CONTAGION RISKS	
* VIX	Thomsom Reuters
* 10 year US Treasury bond yield	Thomsom Reuters
* Short term interest rate	Thomsom Reuters
* Trade links	Thomsom Reuters / own calculations
* EMBI sovereign spread	Thomsom Reuters
* Oil prices	Thomsom Reuters

Variables included in the AUROC exercise, mainly taken from the previous EWS literature

What does our data look like? Table A1.2 presents the mean for the variables included in the estimations of both parametric and non-parametric models, for the whole sample and for tranquil and stress times. Some of the results of the previous literature could be traced here. In a sovereign crisis, sovereign spreads soar and sovereign rates sunk in comparison with tranquil times. Moreover, a high external debt[49] and short term external debt are also representative features of sovereign crises, which tend to be hand in hand with a sudden stop in capital inflows (portfolio) and are more frequent in poorer and more unstable countries (lower GDP per capita and higher political risk). Additionally, oil prices during sovereign crises tend to be lower, pointing to a lost of fiscal revenues in oil or commodity exporting countries that in the end could lead to a sovereign default. In the case of a currency crisis, nominal exchange rate deviation from the trend (currency overvaluation), activity data and public sector imbalances seem to be relevant. Moreover the authorities tend to sustain the exchange rate reducing their level of international reserves and increasing the short term interest rate. In some cases, the external position of the economy became unsustainable due to the net foreing asset position of domestic banks. Finally, banking crises are characterized by huge non-performing loans, a strong decrease in real credit and a negative net external position of domestic banks.

Table A1.2

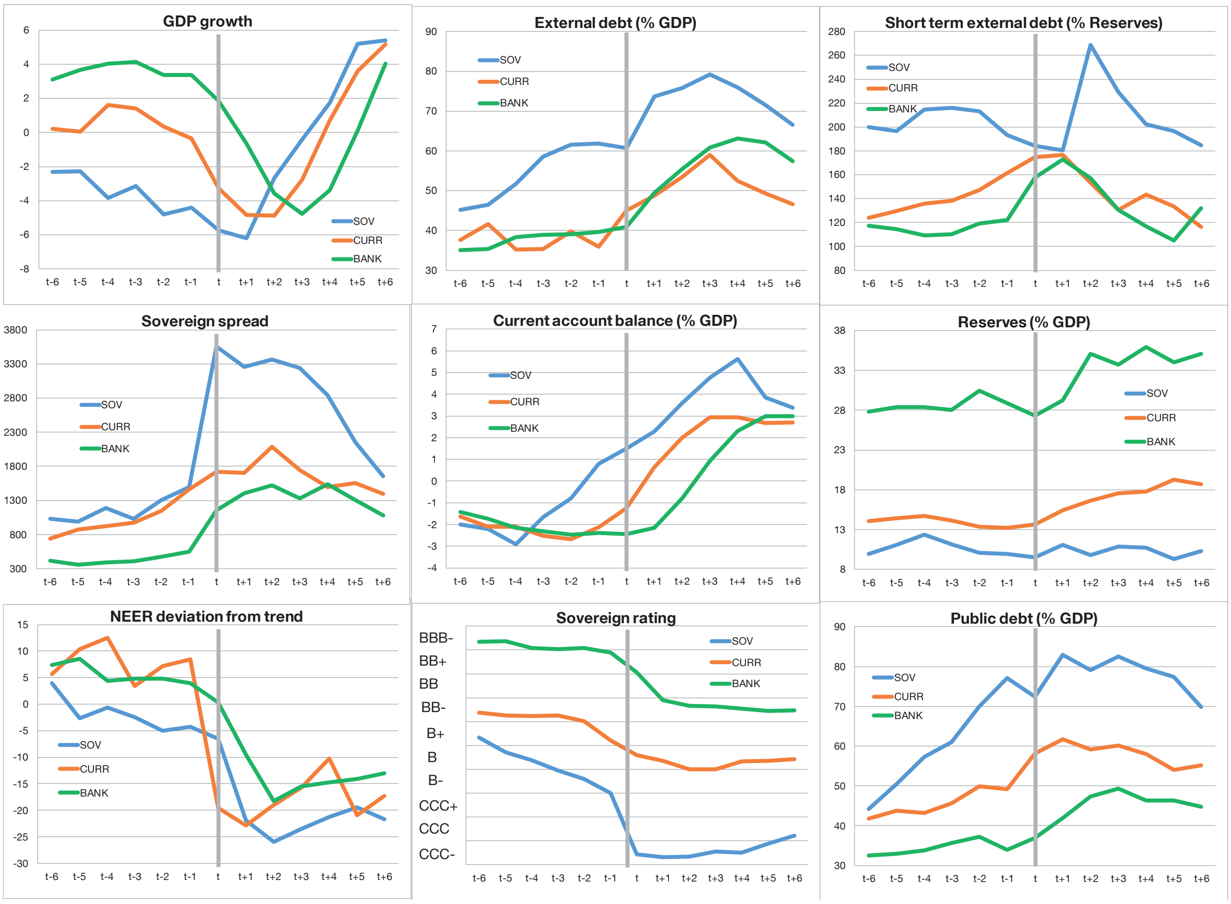
Variable	Observations Whole sample	Mean Whole sample	No crisis	Sovereign	Currency	Banking
Sov.spread	2083	450.9	337	2947	1981	993
3M change sov.spread	2042	2.6	1.9	12.6	22.3	6.2
3M change stock exchange	2316	5.4	3.6	40.4	82.4	5.1
3M change ER vs USD	2456	-1.6	-0.9	-10.4	-33.4	-3.1
GDP growth	2595	3.7	4.2	2.6	-4.4	1.5
Inflation rate	2591	67.1	10.7	1341.0	952.7	24.2
12M change Ind. production	2352	2.9	3.2	2.0	-5.2	0.5
NEER deviation from trend	2580	-5.4	-6.2	16.2	39.3	-14.5
Public sector balance	2560	-2.1	-2.0	-1.8	-5.2	-2.4
Public debt	2578	42.9	38.5	125.5	107.3	50.8
Real credit (y-o-y)	2577	7.1	9.0	1.6	-4.3	-7.8
Real deposits (y-o-y)	2578	7.0	8.0	6.1	-6.9	-0.9
Net Foreign Assets dom.banks	2594	-1.8	2.0	-6.1	-14.9	-8.6
Non performing loans (% total)	1716	6.5	5.6	7.3	7.2	15.2
Loan to Deposit ratio	2594	1.0	1.0	1.1	1.2	1.3
3M change stock banks	2290	5.7	4.6	28.5	48.8	4.6
Short term interbank rate	2533	11.0	9.4	20.7	47.5	17.5
Lending minus deposit rates	2514	7.5	6.9	10.0	21.0	10.9
Current account balance	2592	-0.5	-0.5	0.6	0.4	0.1
FDI	2589	2.6	2.5	2.0	1.6	3.5
External debt	2577	38.2	33.8	79.3	66.1	59.8
Short term ext.debt	2574	77.7	68.2	221.8	187.7	105.8
Reserves	2597	29.1	27.4	11.0	11.4	47.4
Ext.debt service	2562	29.5	26.5	48.8	45.7	44.8
Portfolio inflows	2529	0.9	1.0	-0.8	0.4	-0.1
GDP per capita	2600	13220.7	13716	7955	9467	11141
GDP per capita (change)	2600	4.8	5.2	3.7	-3.0	3.1
GPR index	1456	96.7	96.3	114.2	99.1	96.5
Sov.rating	2440	BBB-	BBB-	CCC	B-	BB
Int. Trade interconectedness	2574	25.7	25.9	32.2	28.2	23.0
VIX	2600	19.3	19.1	19.2	18.3	21.4
10Y Treasury yield	2600	4.2	4.1	5.2	4.9	5.0
3M US interbank rate	2600	2.9	2.7	3.6	3.4	3.9
Oil Price	2600	50.3	53.3	29.4	31.4	34.5
EMBI Global	2600	523.0	493.8	700.1	670.2	688.8

Chart A1.2 depicts the developments of some of the relevant variables six quarters before and six quarters after the onset of a crisis, labelled t, by type of crisis. This preliminary analysis suggests that in the case of sovereign crises the level and increase of external debt and the rise of public debt over GDP play a relevant role, while for currency crises the currency

[49] Which tend to be correlated with high public debt in EMEs -especially at the beginning of the sample- as most of the external debt was public or publicly guaranteed.

overvaluation and the low level of reserves behave abnormally before a crisis. This points to the need of using the variables of the sample both in levels and, in some cases, in differences when estimating their accuracy as leading indicators of a crisis.

Chart A1.2



3.- Crises

This section provides a global overview of the trigger of the most relevevant crises in EMES; it briefly revises the literature on "generations" of currency crises; and it presents the dates of crises on which the Auroc and the estimations are based on.

- **Mexico (1995):** The so-called Tequila crisis blew up in December 1994 in the midst of increasing political strains -the assassination of the PRI candidate for President-, a deteriorated banking sector, and a huge amount of dollar denominated short-term debt (Tesobonos). Its spillovers hit Argentina throughout 1995.

- **Asia (1997):** The trigger of this crisis was a series of consecutive devaluations of dollar-linked currencies of Asian countries, starting with the Thai bath in July 1997, on account of the strong devaluation of the Chinese Renminbi in 1994, which reduced the competitiveness of the economies that competed with China in similar value-added goods in third markets. The currency and maturity mismatches -due to the accumulation of dollar denominated short-term liabilities and domestic currency denominated long-term assets- in both firms and banks of the region fuelled the crisis.

- **Russia (1998)** issued short-term dollar denominated debt (GKOs) which were put under default in August 1998 after a huge increase of their interest rates (up to 150%) and the strong decline in commodity prices that followed the Asian crisis. The Russian default led to the bankruptcy of the US fund LTCM and the intervention of the US Federal Reserve.

- **Argentina (2001)** abandoned its currency board with the US dollar –established in 1991- and declared the default on its external debt at the end of 2001 in the aftermath of unsuccessful fiscal consolidation with the support of the IMF and in the midst of a deep political crisis.

From a theoretical point of view, there are three "generations" of currency crisis models. The first generation (Krugman, 1979), explains speculative attacks on fixed exchange rates as a rational behaviour of markets that can anticipate the unsustainability of maintaining the parity and the increase of fiscal imbalances, reflected in a widening of the current account deficit as the economy is less and less competitive and cannot generate export revenues at the needed pace. When the attack occurs, authorities tend to sell international reserves and increase official rates until they finally devaluate the currency. For those models, variables like the public and current account balance, short term rates, the overvaluation of the currency, the inflation rate (which represents an effective overvaluation as the exchange rate used to be fixed and it is also related to the monetization of the fiscal deficit) and the level of reserves are essential. Second Generation Models (Obstfeld 1986) introduced an objective function of authorities that could decide to abandon the fixed rate whenever the loss is too high. This objective function depends on the deviations of GDP from its natural level and inflation. When markets' expectations align on the prospect of a devaluation (due, for example, to a widening current account deficit) but the government does not devalue, inflation will be too low and GDP will be below the natural rate (unemployment will be too high). As authorities could face a high loss, they opt to devalue the currency (multiple equilibria and self-fulfilling prophecies). In this context, activity growth rates, inflation, the current account, and financial markets variables has to be taken into account. Finally third generation models (Eichengreen and Hausmann, 1999) highlight the interaction between banks and currency crises through balance sheets mismatches: banks could have borrowed short term foreign currency to finance long term projects, whose revenues are denominated in local currency. When an external event increases the cost of external funding, banks could go bankrupt, increasing the loss of currency crises. On the contrary, a domestic event (like a recession) that strains banks borrowers could lead to banks defaulting on its external loans and to a devaluation of the currency. Finally, if the government takes over banks, the risk of a fiscal crisis increases which feedbacks the currency crisis. This model adds to the previous set of variables net external positions of banks, loan to deposit ratio, and the change of credit and deposits. On sovereign crises, in the case of EMEs, fiscal imbalances tend also to be reflected in an increase of external debt, especially short term, and as usually domestic savings rates are low all variables related to external financing, such as FDI (more stable) and portfolio (less stable) foreign inflows are relevant.

Table A1.3

Country	Event		
	Banking crises	**Sovereign crises**	**Currency crises**
Argentina	1995q1-1996q2 2001q4-2004q3	1993q1-1993q2 2001q4-2005q2 2014q3-2016q2	2002q1-2002q2 2016q1 2018q3
Brazil	1994q4-1997q3	1993q1-1994q1	1993q1-1994q3 2008q4
China	1998q1-1998q2		1994q1
Chile			
Colombia	1998q2-2002q1		
Czech Republic	1996q2-2003q2		
Ecuador	1998q3-2001q3	1993q1-1995q2 1999q3-2000q3 2008q4-2009q4	1999q1 1999q4-2000q1
Egypt			2016q4
Hungary	1993q1-2015q2		
India	1993q1-1993q2		
Indonesia	1997q4-2000q4	1999q2-2002q3	1997q4-1998q1
Korea	1997q3-1998q4		1998q1

Date of crises, by type

Table A1.3 (contd.)

Country	Event Banking crises	Sovereign crises	Currency crises
Mexico	1994q4-1997q1		1995q1
Nigeria	2009q3-2013q2		1999q1 2016q3
Peru	1993q1-1997q4		
Poland	1993q1-1994q2		
Romania	1993q1-1994q3		1993q4-1994q1 1997q1
Russia	1998q3-2001q3 2008q3-2010q1 2015q1-2016q4	1998q3-2000q3	1993q1-1995q1 1998q3 2014q4
South Africa		1993q1-1993q3	
Thailand	1997q3-2002q2		
Turkey	2000q4-2003q4		1994q1-1994q2 2001q2
Uruguay	2002q1-2006q2	2003q1-2003q2	2002q3
Venezuela	1994q1-1997q1	2017q4-2018q4	1996q1-1996q2 2002q3 2011q1 2014q2 2016q1-2016q3 2017q2-2017q3 2018q1-2018q3

Date of crises, by type

APPENDIX II: AUROC RESULTS AND DIFFERENT SPECIFICATIONS FOR THE SHERLOCS

Table A2.1 presents the AUROC results for all the variables included in the sample.

Table A2.1

Variable	By type of crisis				By region				By time	
	All crises	Banking	Sovereign	Currency	Latin America	East. Eur.	Asia	Other EMEs	Before 2007Q1	After 2007q1
Sov.spread	0.73	0.59	0.90	0.80	0.77	0.64	0.18	0.83	0.68	0.77
3M change sov.spread	0.58	0.56	0.59	0.61	0.57	0.64	0.43	0.63	0.61	0.56
3M change stock exchange	0.51	0.54	0.54	0.51	0.45	0.55	0.60	0.67	0.59	0.42
3M change ER vs USD	0.60	0.59	0.64	0.63	0.59	0.57	0.62	0.59	0.64	0.55
GDP growth	0.60	0.53	0.83	0.67	0.67	0.58	0.35	0.57	0.60	0.63
Inflation rate	0.74	0.64	0.68	0.75	0.70	0.76	0.68	0.79	0.64	0.86
12M change Ind. production	0.59	0.58	0.68	0.64	0.61	0.59	0.52	0.65	0.61	0.59
NEER deviation from mean trend	0.59	0.68	0.59	0.58	0.58	0.49	0.72	0.62	0.63	0.55
NEER deviation from HP trend	0.46	0.44	0.39	0.48	0.39	0.43	0.64	0.55	0.43	0.49
Public sector balance	0.56	0.39	0.57	0.66	0.70	0.56	0.26	0.57	0.53	0.64
Public sector balance (first diff)	0.57	0.54	0.49	0.57	0.50	0.56	0.65	0.67	0.63	0.50
Public debt	0.45	0.38	0.70	0.53	0.56	0.39	0.12	0.52	0.36	0.57
Public debt (first diff)	0.62	0.57	0.74	0.69	0.62	0.59	0.41	0.77	0.64	0.63
Real credit (y-o-y)	0.57	0.59	0.45	0.50	0.43	0.74	0.79	0.55	0.54	0.61
Real credit (y-o-y) tail2	0.43	0.41	0.56	0.50	0.57	0.26	0.21	0.45	0.46	0.39
Real deposits (y-o-y)	0.53	0.58	0.45	0.49	0.42	0.65	0.69	0.58	0.51	0.55
Net Foreign Assets dom.banks	0.59	0.68	0.56	0.56	0.53	0.51	0.65	0.57	0.66	0.51
Non performing loans (% total)	0.48	0.62	0.46	0.47	0.47	0.32		0.48	0.54	0.44
Loan to Deposit ratio	0.54	0.71	0.62	0.46	0.45	0.48	0.93	0.40	0.71	0.31
3M change stock banks	0.51	0.57	0.55	0.52	0.47	0.54	0.54	0.59	0.57	0.43
Short term interbank rate	0.74	0.78	0.70	0.75	0.69	0.73	0.91	0.73	0.75	0.69
Lending minus deposit rates	0.55	0.51	0.59	0.60	0.50	0.52	0.31	0.75	0.48	0.42
Current account balance	0.57	0.60	0.52	0.58	0.58	0.40	0.79	0.54	0.68	0.45
Curr.acc.balance (first diff)	0.61	0.55	0.46	0.61	0.60	0.64	0.47	0.74	0.62	0.61
FDI	0.56	0.52	0.56	0.61	0.72	0.54	0.42	0.54	0.47	0.63
FDI (first diff)	0.47	0.45	0.51	0.47	0.47	0.50	0.39	0.54	0.48	0.47
External debt	0.47	0.52	0.73	0.47	0.50	0.30	0.58	0.18	0.50	0.43
External debt (first diff)	0.58	0.60	0.66	0.60	0.56	0.62	0.59	0.54	0.61	0.57
Short term ext.debt	0.67	0.66	0.87	0.71	0.72	0.56	0.82	0.25	0.70	0.62
Reserves over GDP	0.64	0.54	0.73	0.70	0.68	0.61	0.76	0.64	0.56	0.70
Reserves / GDP (first diff)	0.52	0.50	0.52	0.55	0.49	0.53	0.46	0.57	0.52	0.53
Reserves (m.USD)	0.61	0.56	0.76	0.62	0.56	0.65	0.82	0.45	0.54	0.64
Ext.debt service	0.57	0.58	0.74	0.58	0.65	0.42	0.52	0.30	0.60	0.52
Portfolio inflows	0.52	0.62	0.33	0.49	0.53	0.55	0.57	0.34	0.61	0.43
Portfolio inflows (first diff)	0.49	0.50	0.47	0.48	0.49	0.51	0.50	0.44	0.51	0.47
Portfolio inflows tail2-decline	0.48	0.38	0.67	0.51	0.47	0.45	0.43	0.62	0.39	0.57
Portfolio inflows (first diff) tail 2	0.51	0.50	0.53	0.52	0.51	0.49	0.50	0.53	0.48	0.53
GDP per capita	0.57	0.59	0.60	0.55	0.52	0.69	0.67	0.68	0.52	0.53
GDP per capita (change)	0.60	0.50	0.84	0.68	0.68	0.55	0.41	0.58	0.59	0.63
Sov.rating	0.65	0.51	0.88	0.73	0.74	0.55	0.34	0.92	0.52	0.79
Political risk	0.45	0.41	0.50	0.44	0.45	0.47	0.36		0.40	0.54
Political risk (change)	0.53	0.53	0.57	0.50	0.54	0.53	0.49		0.54	0.51
VIX	0.52	0.53	0.63	0.50	0.51	0.46	0.53	0.65	0.52	0.49
US 10Y treasury bond rate	0.65	0.78	0.56	0.58	0.61	0.76	0.92	0.30	0.73	0.56
US 3M interbank rate	0.62	0.75	0.60	0.56	0.57	0.73	0.85	0.41	0.64	0.53
Oil Price	0.38	0.30	0.39	0.36	0.41	0.37	0.16	0.47	0.29	0.53
Int. Trade interconectedness	0.53	0.45	0.72	0.57	0.53	0.60	0.17	0.74	0.49	0.62

AUROC results: percentage of good signals issued 6 quarters before the onset of each type of crisis. Red bolded cells indicate an AUROC above 70% and orange bolded cells an AUROC above 65%.

As a robustness check, we validate the performance of several indexes based on different specifications. Indeed, one of the striking results is that some variables that are commonly found to be leading indicators by the literature have been excluded in our index since its performance was very poor in terms of its AUROC. This is the case of non-performing loans for banking crises, or the deviations (overvaluation) of effective exchange rates for currency stress. That is why, we decide to carry out a robustness check exercise including some of these variables. Moreover, we also wonder if our index would have performed better if we exclude external factors, such as short and long term interest rates in the US for banking crises and trade links in the case of sovereign stress. Finally, we also check if the inclusion of sovereign spread (in level or in changes) improves the sovereign index in terms of usefulness.

Table A2.2, A2.3 and A2.4 present the results for the banking, sovereign and currency stress, respectively. Broadly speaking, the indexes based on the AUROC results seem to outperform the other specifications both in-sample and out-of-sample. This is true for the banking and sovereign indexes. However, in the case of currency crises, the inclusion of the effective exchange rate and exclusion of rating seem to improve in terms of usefulness. We decide to maintain the results of the AUROC in order to avoid data mining and cherry-picking issues, though we are aware that policymakers also need to monitor the development of the effective exchange rate.

Table A2.2: Robustness check. Banking indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	75	0.33	0.24	86%	17%
Banking RC1	67	0.28	0.35	86%	13%
Banking RC2	81	0.24	0.25	63%	12%
Banking RC3	83	0.25	0.21	63%	14%
Out-of-sample performance					
Banking	70	0.14	0.53	58%	4%
Banking RC1	74	0.03	0.83	33%	3%
Banking RC2	71	-0.11	3.62	8%	1%
Banking RC3	67	-0.07	1.65	21%	1%

This table presents the in-sample and out-of-sample exercises for different specifications for the banking index. Banking is the model proposed, based on the AUROC results. RC1 includes the same variables except for external factors. RC2 also excludes external factors, but includes NPL. Finally, RC3 includes NPL and external factors. The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%). The threshold indicates the percentile beyond which the index issues a signal. This exercise is based on a neutral policy maker (θ =0.5) and an evaluation window of six quarters previous to the crisis.

Table A2.3: Robustness check. Sovereign indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Sovereign	72	0.27	0.32	79%	6%
Sovereign RC1	82	0.27	0.22	69%	9%
Sovereign RC2	78	0.22	0.31	64%	6%
Sovereign RC3	80	0.28	0.24	74%	8%
Out of sample performance					
Sovereign	69	0.08	0.68	50%	2%
Sovereign RC1	78	0.04	0.78	33%	2%
Sovereign RC2	70	0.08	0.68	50%	2%
Sovereign RC3	69	0.13	0.56	61%	3%

This table presents the in-sample and out-of-sample exercises for different specifications for the sovereign index. Sovereign is the model proposed, based on the AUROC results. RC1 includes the same variables except for external factors. RC 2 also excludes external factors, but includes sovereign spread. Finally, RC3 includes the sovereign spread and external factors. The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%). The threshold indicates the percentile beyond which the index issues a signal. This exercise is based on a neutral policy maker (θ =0.5) and an evaluation window of six quarters previous to the crisis.

Table A2.4: Robustness check. Currency indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Currency	57	0.20	0.50	81%	10%
Currency RC1	71	0.21	0.38	69%	13%
Currency RC2	63	0.24	0.41	83%	12%
Currency RC3	65	0.22	0.42	77%	13%
Out of sample performance					
Currency	76	0.11	0.55	49%	8%
Currency RC1	59	0.11	0.66	65%	7%
Currency RC2	76	0.14	0.48	54%	10%
Currency RC3	78	0.11	0.50	44%	10%

This table presents the in-sample and out-of-sample exercises for different specifications for the currency index. Currency is the model proposed, based on the AUROC results. RC1 includes the same variables except for rating. RC 2 also excludes rating, but includes nominal effective exchange rate. RC3 includes the same variables than in "Currency" and sovereign spread. The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob (%). Threshold indicates the percentile beyond which the index issues a signal. This exercise is based on a neutral policy maker (θ =0.5) and an evaluation window of six quarters previous to the crisis.

APPENDIX III: DIFFERENT DEFINITION OF CURRENCY CRISIS

As currency crises are defined using a quantitative threshold instead of using events, such as done for the banking and sovereign crises, we carry out a robustness exercise using a broader definition of currency crisis. Currency crisis is defined, in this case, as a depreciation of at least 15% on a quarterly basis as long as this depreciation exceeds the average variation of the exchange rate plus a standard deviation.

The AUROC and logistic results are very similar to the outcome of the baseline specification. Almost the same variables are preselected using the AUROC, though rating and reserves are not included. The variables included in the logistic regression have the expected sign and they are significant. Table A3.1 reports the in-sample and out-of-sample validation results for the currency index using a broader definition. While in-sample results remain similar to those reported in the baseline specification in terms of usefulness, out-of-sample results are very poor. There is no gain in using this model out-of-sample, and therefore we conclude that it is more appropiate to stick to a narrow definition of currency crisis.

Table A3.1: Robustness check. Currency crisis

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Currency	72	0.21	0.38	68%	13%
Out of sample performance					
Currency	76	0.00	1.03	25%	5%

This table presents the in-sample and out-of-sample exercises for the currency index, based on the predicted probility from a logistic estimation, using a different definition of event. Currency crisis is defined as a depreciation of at least 15% on a quarterly basis as long as this depreciation exceeds the average variation of the exchange rate plus a standard deviation.

APPENDIX IV: DEALING WITH THE POST-CRISIS BIAS

The use of a binary indicator to define periods of stress suffers from a "post-crisis" bias, that is, it is not able to distinguish between normal periods and the recovery phase after a crisis. As a robustness check, we remove post-crisis years (4 quarters), as done with crisis years, at a cost of fewer observations. The AUROC and the logistic approach are re-estimated using this new definition of stress periods. The results of our baseline model are robust to this new specification since estimates hardly vary.

Table A4.1 presents the in-sample and out-of-sample exercise when dealing with the post-crisis bias. In our case, we show that considering the "post-crisis" bias, by eliminating the data of four quarters after a crisis, is relevant to slightly reduce type II errors (false alarms) out-of-sample, but it does not seem to affect the ratio of crises predicted in-sample and it has only a marginal impact out-of-sample.

Table A4.1: Robustness check. Dealing with the post-crisis bias

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	76	0.33	0.23	85%	18%
Sovereign	74	0.26	0.31	75%	7%
Currency	69	0.20	0.41	69%	12%
Out of sample performance					
Banking	70	0.14	0.53	58%	4%
Sovereign	70	0.11	0.61	56%	3%
Currency	68	0.14	0.55	63%	9%

This table presents the in-sample and out-of-sample exercises for three different indexes minimising the post-crisis bias, based on the predicted probability from a logistic estimation. To assess the post-crisis bias, we remove the four quarters after the crisis in both the AUROC and the logistic estimations.

APPENDIX V: PSEUDO REAL TIME SHERLOCS

A relevant issue in EWS is to carry out real-time exercises of the models proposed, which implies the development and estimation of its performance using the real data that policy makers would have at the moment of making decisions. In the tool proposed, the composite indexes are updated with the available information up to this moment, and therefore are real-time exercises. However, the validation exercise has been done without considering the "real-time" approach for two main reasons. First, we do not have the vintages and therefore, we can only assess its performance using a "pseudo-real time" approach. Second, due to the wide heterogeneity of countries and indicators used in our model, publications lags are difficult to assess and the use of a common lags to all the countries could bias the results. That is why, we decide to keep the exercise as simple as possible. However, in order to confirm that our results remain robust, we carry out a "pseudo-real exercise" for validating the performance of the model proposed, the Logistic SHERLOC.

Table A5.1 presents the results for in-sample and out-of-sample exercises[50], using a "pseudo-real time" approach. To do so, we include common lags to the variables preselected

[50] For the out-of-sample exercise, we divide our sample into two subsamples. 1993-2006 and 2007-2018. The first subsample is used to calibrate the model and the second subsample to validate its performance.

using the AUROC: one lag (one quarter) for all the macroeconomic variables except for the loan-to-deposit ratio, in which we include two lags. Financial variables are included contemporaneously.

In-sample and out-of-sample results are very similar to those reported in Tables 4 and 5 of the main text. Usefulness measures are positive for the three indexes, and close to the values obtained when discarding the real-time approach. Indeed, while in the in-sample exercise, the usefulness gains are slightly lower, they are considerably higher for the out-of-sample validation for the sovereign indexes, confirming our results.

Table A5.1: Robustness check. "Pseudo-real time" indexes

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	73	0.32	0.26	85%	16%
Sovereign	84	0.25	0.21	65%	9%
Currency	79	0.19	0.33	57%	15%
Out of sample performance					
Banking	63	0.11	0.63	58%	3%
Sovereign	64	0.14	0.59	67%	3%
Currency	66	0.10	0.63	54%	7%

This table presents the in-sample and out-of-sample validation exercise for the banking, sovereign, currency indexes proposed for the predicted probability using a logistic estimation (Logistic SHERLOC). In order to calculate the pseudo-real time idex, common lags have been included for most variables.The validation exercise is based on the measure of Usefulness (U), the noise to signal ratio (NtSr), the percentage of predicted crises (% Predicted) and the number of correct signals in terms of total signals issued (Cond Prob) (%). The threshold indicates the percentile beyond which the index issues a signal. This exercise is based on a neutral policy maker (θ =0.5) and an evaluation window of six quarters previous to the crisis.

APPENDIX VI: EVALUATION WINDOW: FOUR AND EIGHT QUARTERS

The SHERLOC has been constructed using an evaluation window of six quarters prior to the crisis in the AUROC and calculating the probability of being in a “vulnerable state”, defined as six quarters before the onset of the crisis. The election of this period is based on the seminal work of Kaminsky et al (2000), which show that early indicators tend to appear 10 to 18 months before the beginning of the crisis. Additionally, policymakers need to be warned in advance to allow for an earlier implementation of policy tools.

However, as a robustness check, we estimate the AUROC, the Logistic SHERLOC and the validation exercise using an evaluation window of four and eight quarters. The results of the latter are reported in Table A6.1 and A6.2, respectively. Results are robust to this new specification. In-sample validation resuts are quite similar in both cases (four and eight quarters). Out-of-sample results improve in the case of four quarters for sovereign and banking crises in terms of the usefulness measure, but the gain is lower in the case of currency stresses. The opposite results are obtained when using an evaluation window of two years. In any case, the range of usefulness measures remain adequate for an out-of-sample exercise.

Table A6.1: Robustness check. Evaluation window 4 quarters prior to the crisis.

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	74	0.32	0.25	86%	16%
Sovereign	74	0.26	0.31	75%	6%
Currency	73	0.17	0.42	59%	11%
Out-of-sample performance					
Banking	74	0.14	0.48	56%	3%
Sovereign	62	0.12	0.65	67%	2%
Currency	72	0.06	0.72	44%	7%

This table presents the in-sample and out-of-sample exercises for the banking, sovereign and currency indexes using an evaluation window of four quarters instead of six. The "vulnerable state" in the logistic estimation is defined as the four quarters previous to the onset of the stress period.

Table A6.2: Robustness check. Evaluation window 8 quarters prior to the crisis.

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	77	0.32	0.23	83%	18%
Sovereign	84	0.29	0.19	71%	10%
Currency	57	0.20	0.50	81%	10%
Out-of-sample performance					
Banking	62	0.05	0.79	50%	3%
Sovereign	79	0.06	0.66	33%	2%
Currency	78	0.12	0.48	46%	10%

This table presents the in-sample and out-of-sample exercises for the banking, sovereign and currency indexes using an evaluation window of eight quarters instead of six. The "vulnerable state" in the logistic estimation is defined as the eight quarters previous to the onset of the stress period.

APPENDIX VII: POLICY MAKERS' PREFERENCES

In our baseline model, we assume that policy makers are as concerned about missing a crisis as they are about issuing a false alarm. As a result, the validation exercise in sub-section 5.3 is carried out by imposing that θ=0.5. However, it is relevant from a policy maker point of view to know how sensitive these results are to the choice of θ. That is why, in this appendix, we validate the logistic SHERLOC assuming different preferences of policy makers.

Table A7.1 presents the results for the banking, sovereign and currency indexes proposed for the Logistic SHERLOC assuming different values for the parameter θ. Regardless of their preferences, policy makers would benefit from using the indexes. This is particularly true for the case of assuming preferences close to a neutral policy maker (i.e, θ=0.4 or θ=0.6). All models have a positive usefulness except for the case of currency crises assuming a strong preference with regards to not missing crises (θ=0.8).

Table A7.1: Robustness check. Performance of the model depending on the choice of θ

Model	Threshold (percentile)	θ	U	NtSr	% Predicted	Cond Prob (%)
Banking	75	0.5	0.33	0.24	86%	17%
Banking	73	0.6	0.23	0.26	88%	16%
Banking	80	0.4	0.23	0.20	81%	20%
Banking	66	0.7	0.15	0.33	91%	13%
Banking	80	0.3	0.13	0.20	81%	20%
Banking	66	0.8	0.07	0.33	91%	13%
Banking	81	0.2	0.03	0.20	79%	20%
Sovereign	72	0.5	0.27	0.32	79%	6%
Sovereign	84	0.4	0.18	0.20	67%	10%
Sovereign	72	0.6	0.17	0.32	79%	6%
Sovereign	86	0.3	0.10	0.19	63%	10%
Sovereign	63	0.7	0.08	0.42	83%	5%
Sovereign	91	0.2	0.03	0.16	48%	12%
Sovereign	63	0.8	0.00	0.42	83%	5%
Currency	57	0.5	0.20	0.50	81%	10%
Currency	57	0.6	0.12	0.50	81%	10%
Currency	86	0.4	0.12	0.25	48%	19%
Currency	86	0.3	0.06	0.25	48%	19%
Currency	52	0.7	0.05	0.55	83%	9%
Currency	93	0.2	0.01	0.19	29%	23%
Currency	50	0.8	-0.02	0.57	84%	9%

This table presents the in-sample exercise for the banking, sovereign and currency logistic SHERLOC indexes using different values for the parameter θ using an evaluation of six quarters.

APPENDIX VIII: INDEXES BY REGION

In the baseline specification, we decide to build the SHERLOCs distinguishing by type of stress event, instead of focusing on a dummy for all crises or constructing the SHERLOC by region due to the different nature of crises. However, as a robustness check, we also build the SHERLOC using a regional approach. This means that the AUROC and the logistic SHERLOC are estimated using only the information of each region.

The validation exercise confirms our approach (See Table A8.1). In-sample performance is extremely good, with usefulness measures around 0.45 in the case of Asia, but out-of-sample results are very disappointing with a negative value in the case of Africa and not predicting any crise in the case of Asia. This suggests that the regional approach suffers from the criticism of “this time is different”. That means that it is able predict past crises, as regional results are biased with their own past stress episodes, but not future crises. By contrast, an approach distinguishing by type of stress seems to partially mitigate this issue since we are considering different kind of stress that can take place in other countries.

Table A8.1: Robustness check. Indexes by region.

Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Latin America	66	0.21	0.38	68%	27%
Eastern Europe	61	0.32	0.31	91%	27%
Asia	88	0.46	0.09	100%	39%
Africa	86	0.25	0.20	61%	25%
Out-of-sample performance					
Latin America	73	0.08	0.61	42%	17%
Eastern Europe	59	0.03	0.86	50%	8%
Asia	77	-	-	-	0%
Africa	99	-0.19	3.25	17%	4%

This table presents the in-sample and out-of-sample exercises for indexes based only on the regional information. The AUROC and the logit have been estimated considering only the information of each region. In the case of Africa, the out-of-sample validation has been carried out splitting the sample before and after 2014 due to the lack of enough crisis events before 2007.

APPENDIX IX: INDEXES USING POOLED DATA

Finally, in the baseline model, individual country effects are incorporated into the model by using random effects, as suggested by the Hausman test. However, some papers opt to pool observations across countries whenever the group is homogeneous and there are not enough crises to separate countries. In our case, the sample of countries is quite heterogeneous, as already discussed, and therefore we opt to include individual country effects.

As a robustness check, we calculate indexes using pooled data. In the logistic estimation, we get the expected sign and significance in the results. In-sample and out-of-sample validation results are quite similar to the results of the baseline model, though slightly worse in terms of usefulness values. However, we still believe that in our case the random effect approach is the correct way of dealing with this exercise due to the high heterogeneity of countries analysed.

Table A9.1: Robustness check. Indexes using pooled data.

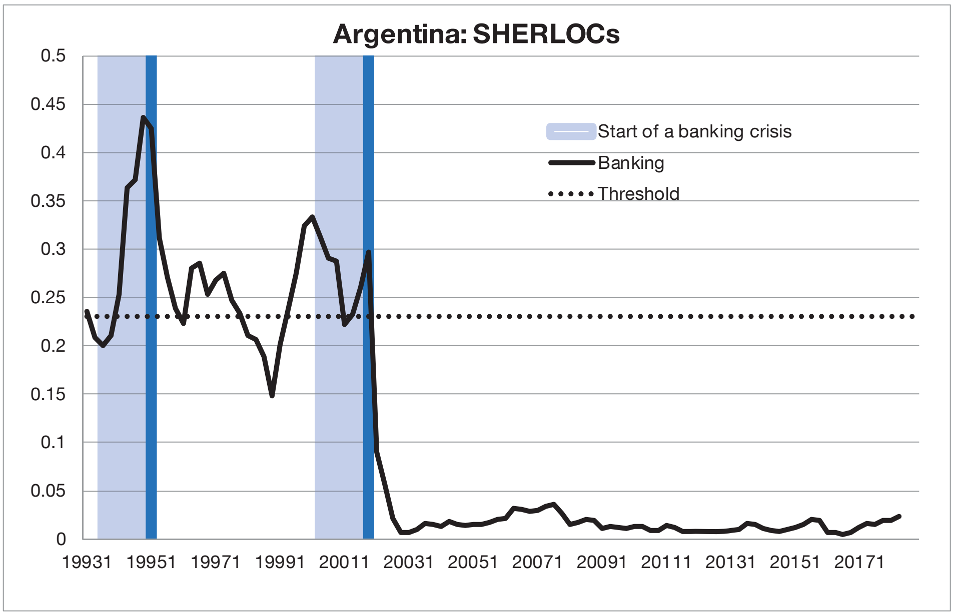
Model	Threshold (percentile)	U	NtSr	% Predicted	Cond Prob (%)
In-sample performance					
Banking	76	0.30	0.25	80%	17%
Sovereign	74	0.25	0.32	73%	6%
Currency	58	0.15	0.57	71%	9%
Out-of-sample performance					
Banking	73	0.18	0.44	63%	5%
Sovereign	69	0.09	0.62	50%	2%
Currency	59	0.04	0.83	49%	6%

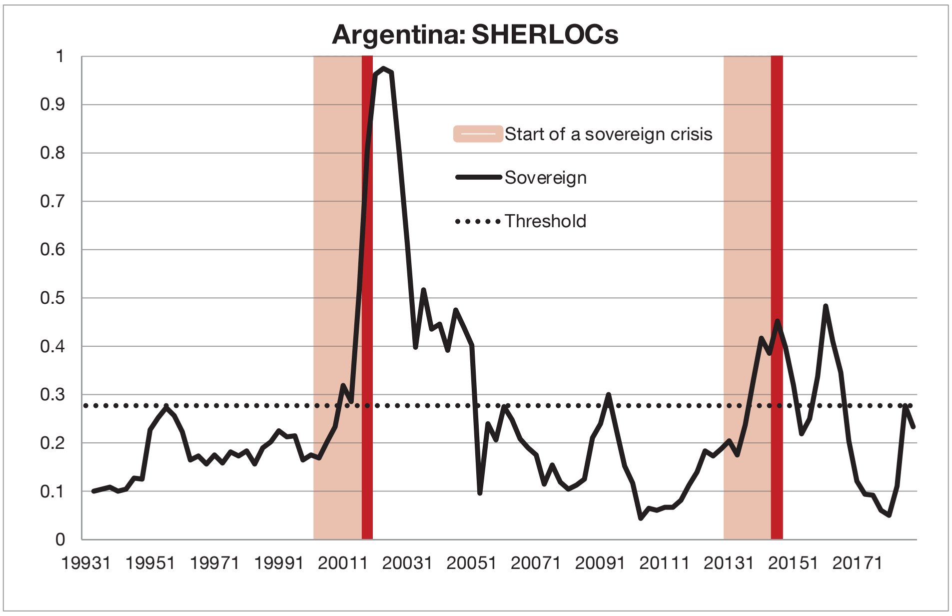
This table presents the in-sample and out-of-sample exercises for different specifications for the banking, sovereign and currency indexes using pooled data in the logistic estimation instead of random effects.

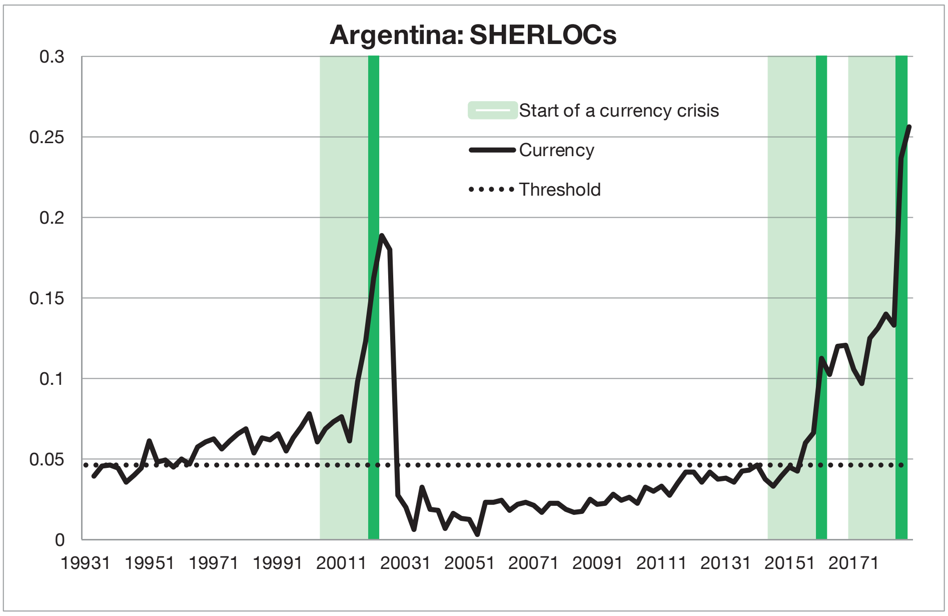
APPENDIX X: SHERLOCS FOR MOST RELEVANT EMES

In this section, we present the Logistic SHERLOCS for some countries of our sample. Thresholds are calculated using the percentiles given by the validation model (75% for banking crises, 72% for sovereign crises and 57% for currency crises). The SHERLOCs of the rest of countries can be provided under request.

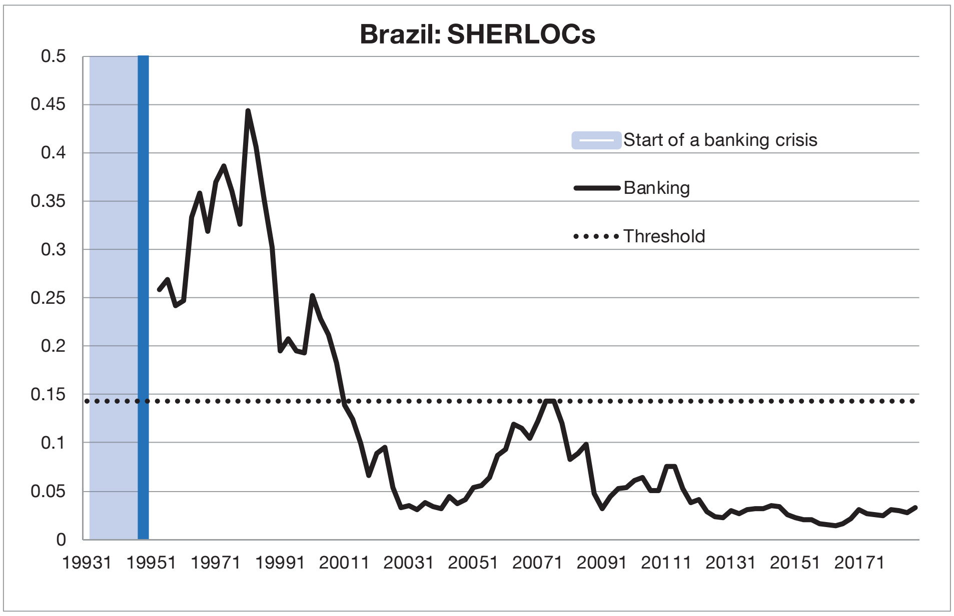
Appendix IX (1)

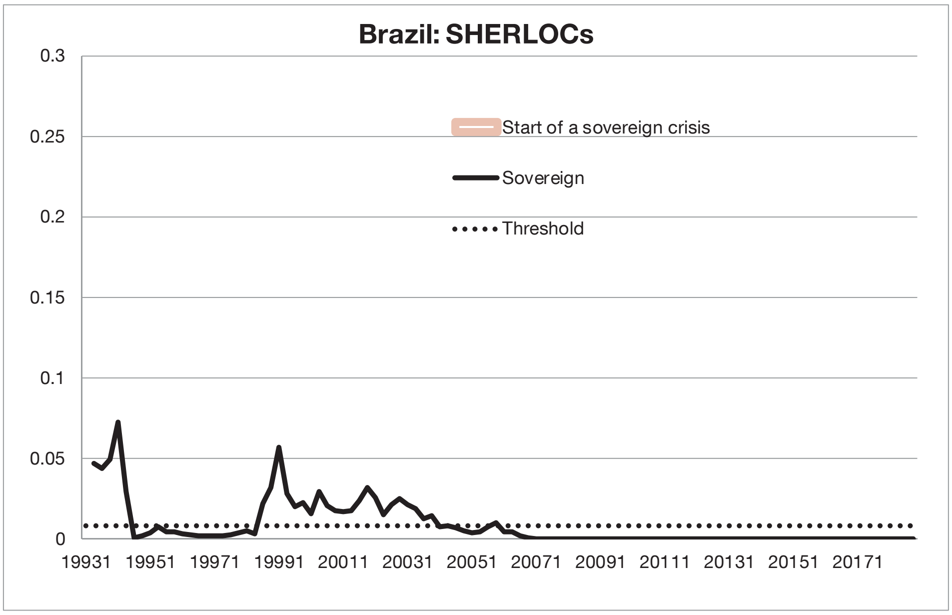


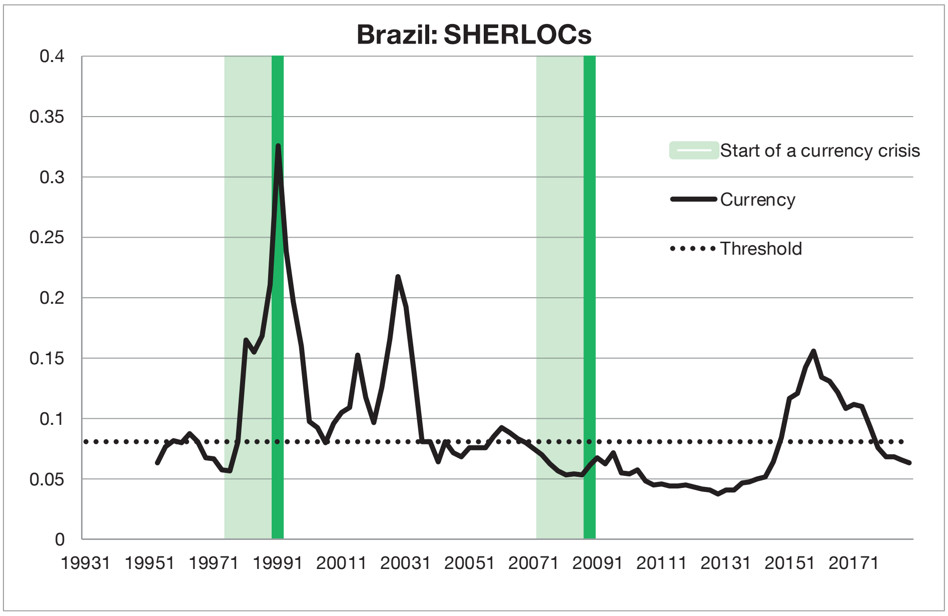


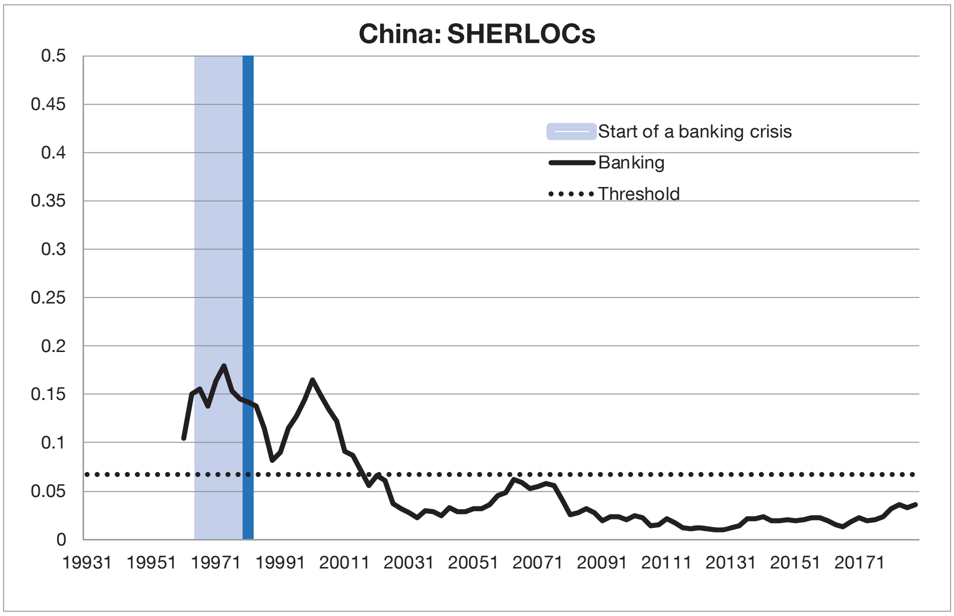


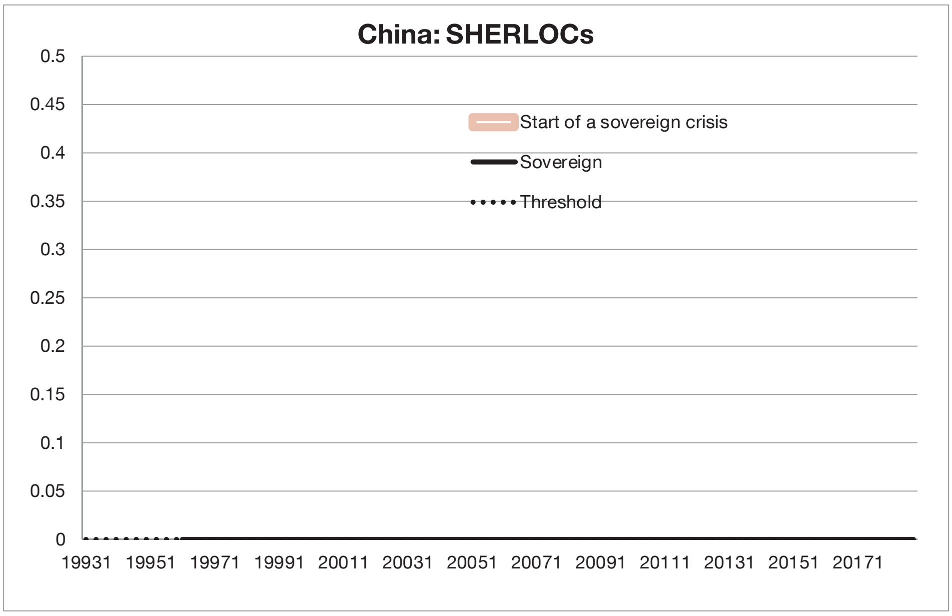
Appendix IX (2)

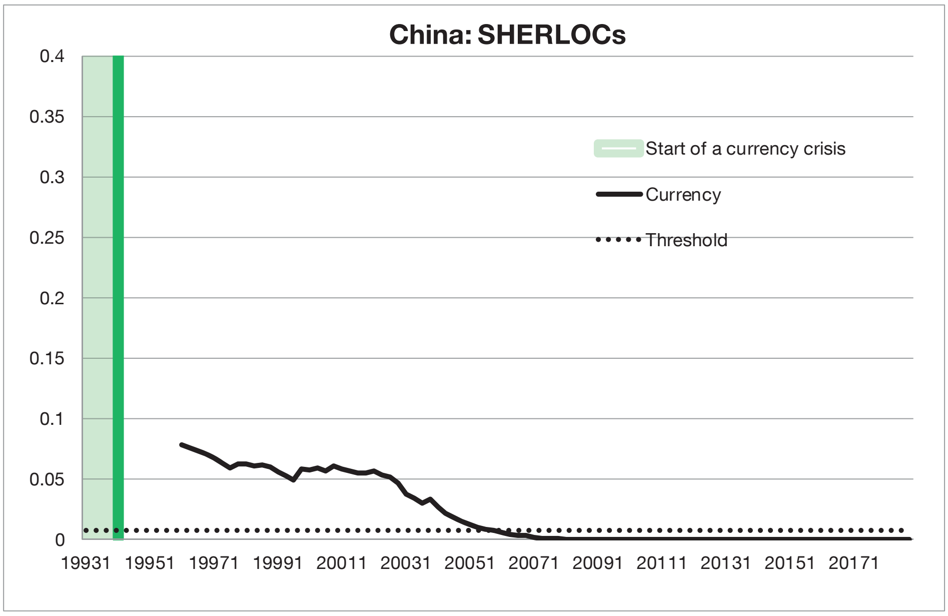




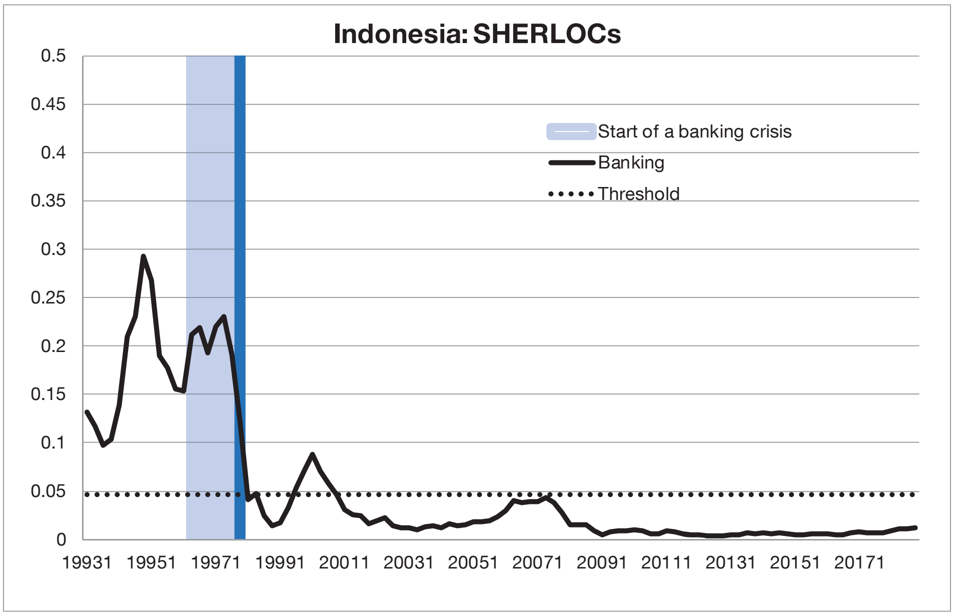


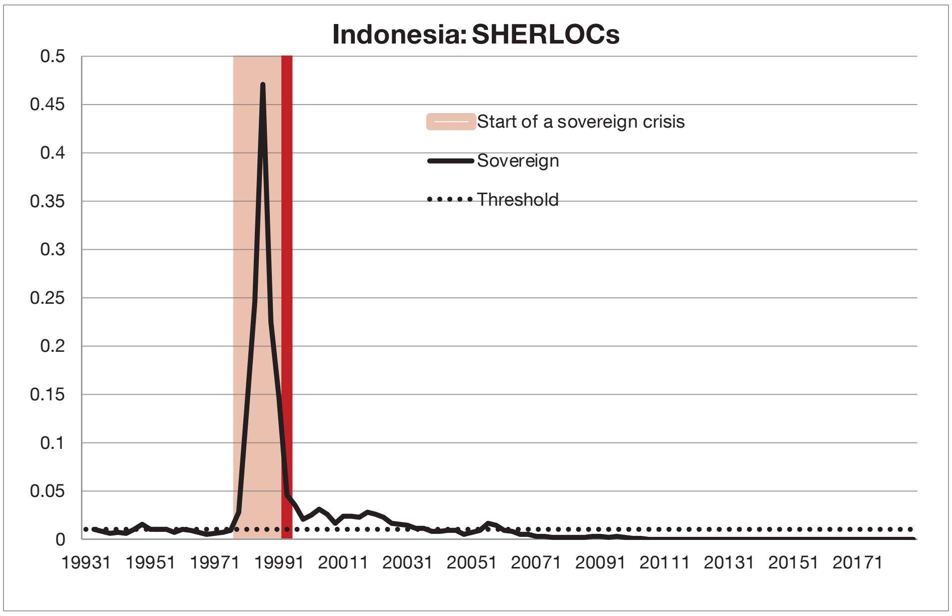

China: SHERLOCs
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Start of a banking crisis
Banking
Threshold
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

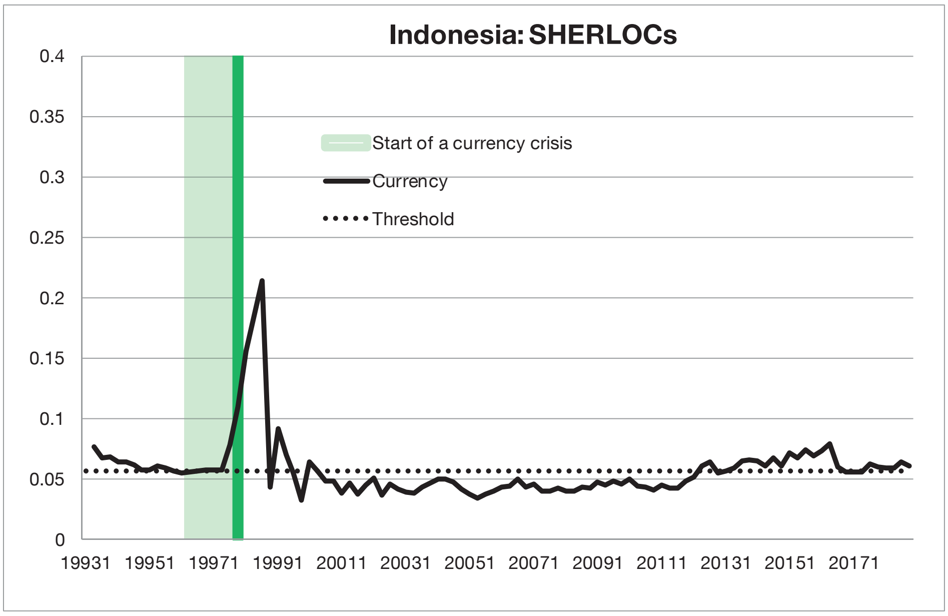

China: SHERLOCs
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Start of a sovereign crisis
Sovereign
Threshold
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171


China: SHERLOCs
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Start of a currency crisis
Currency
Threshold
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

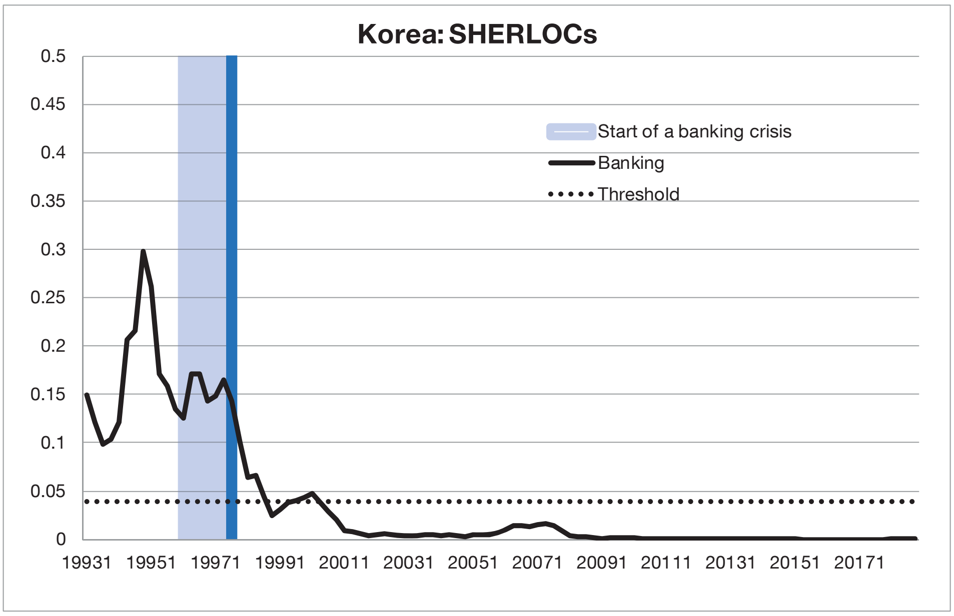
Appendix IX (4)

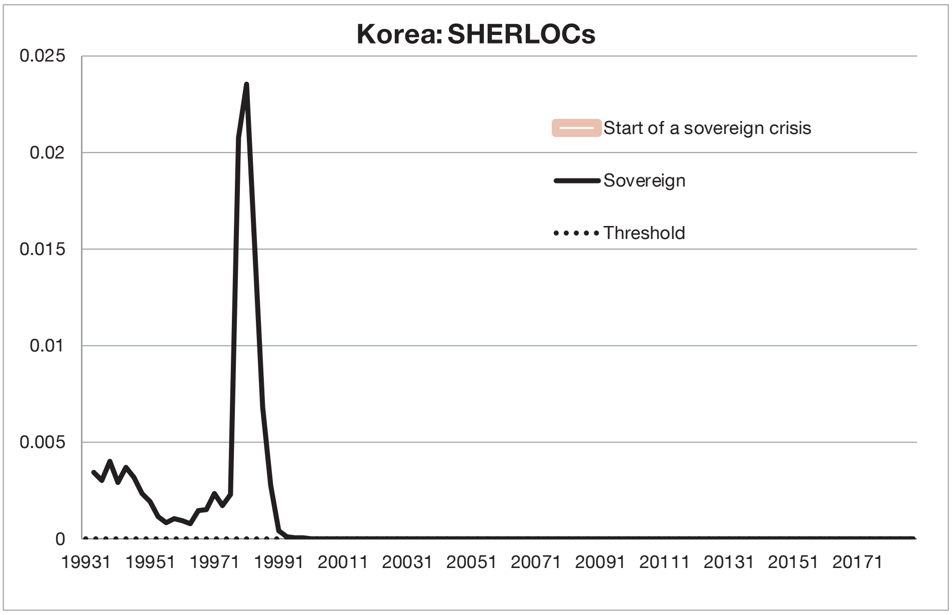


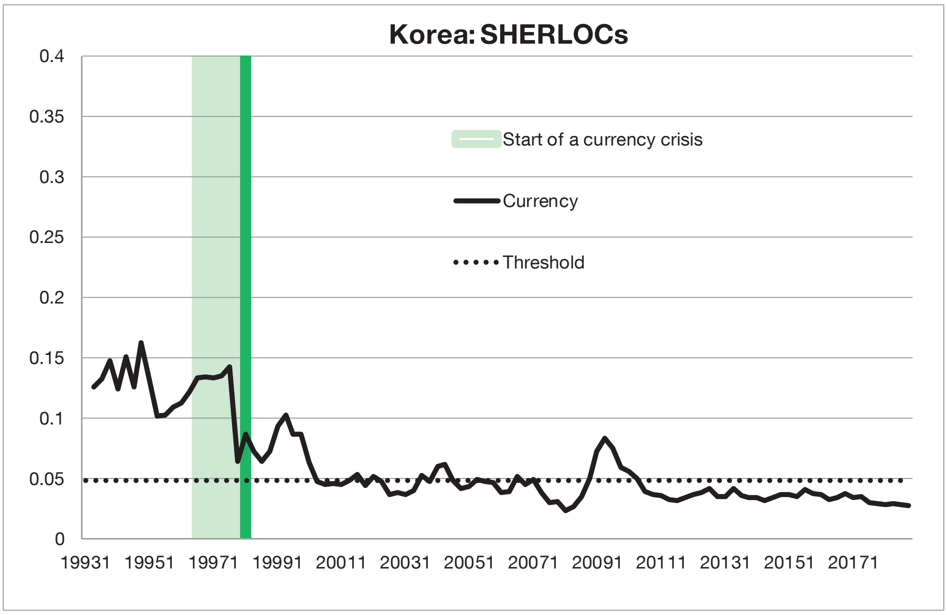


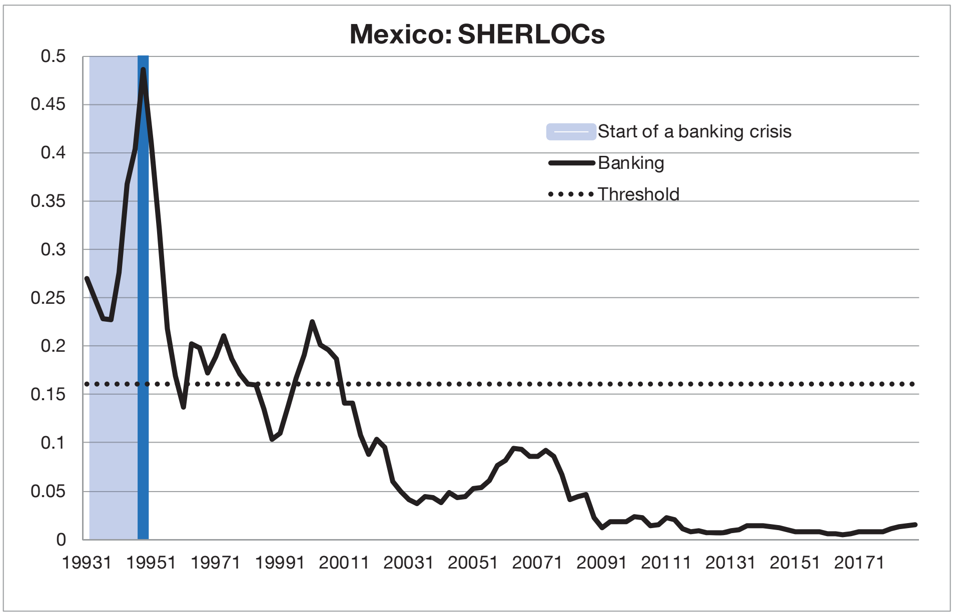


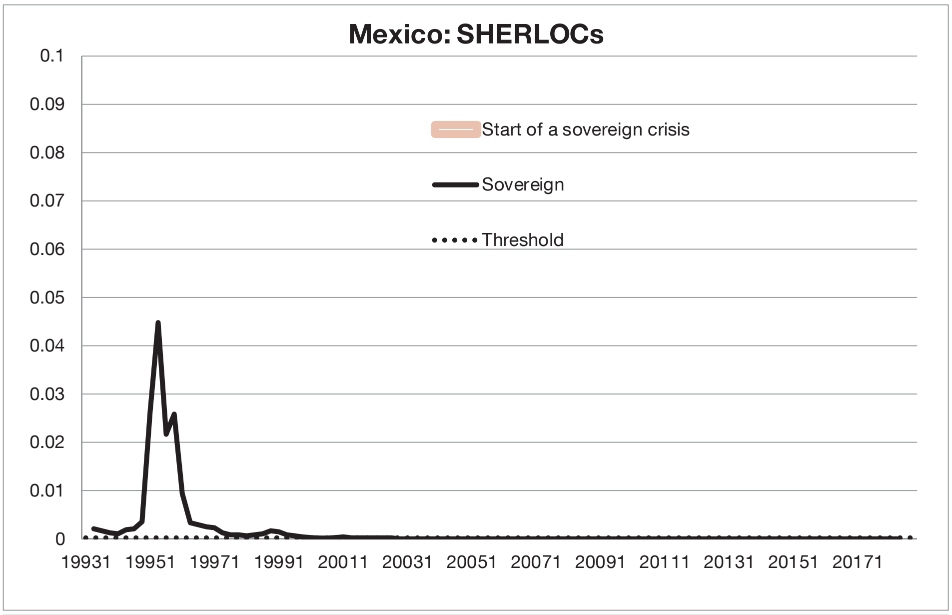
Appendix IX (5)

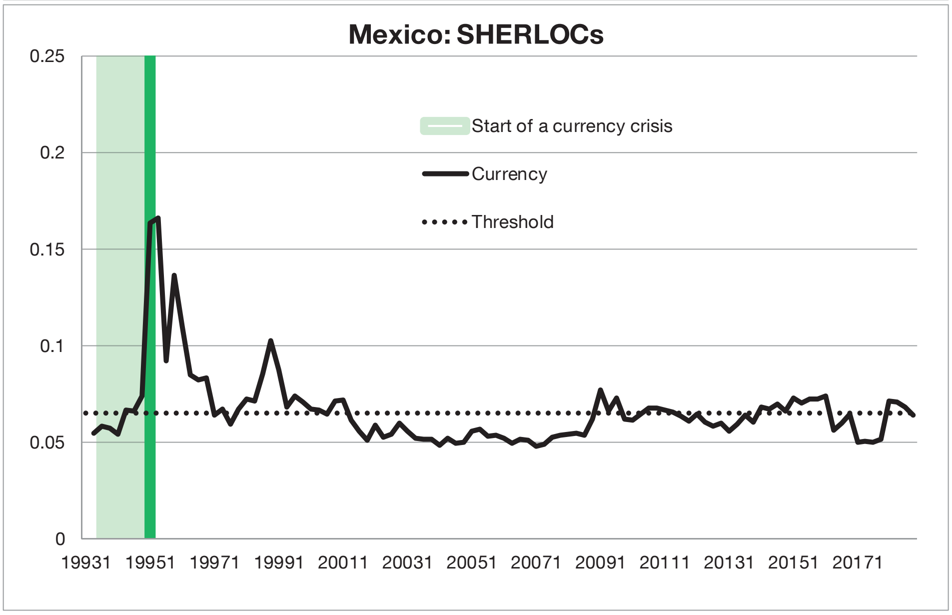




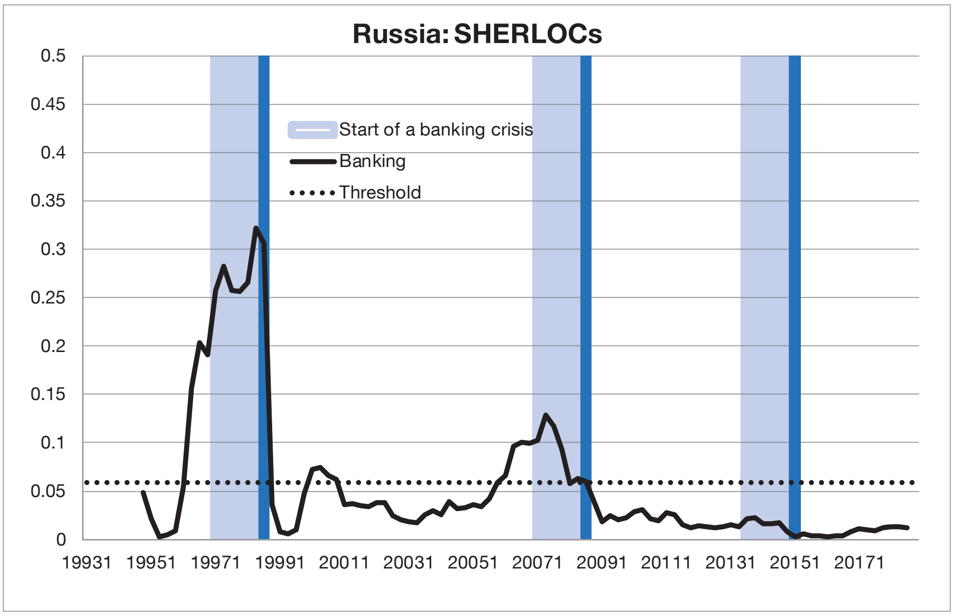


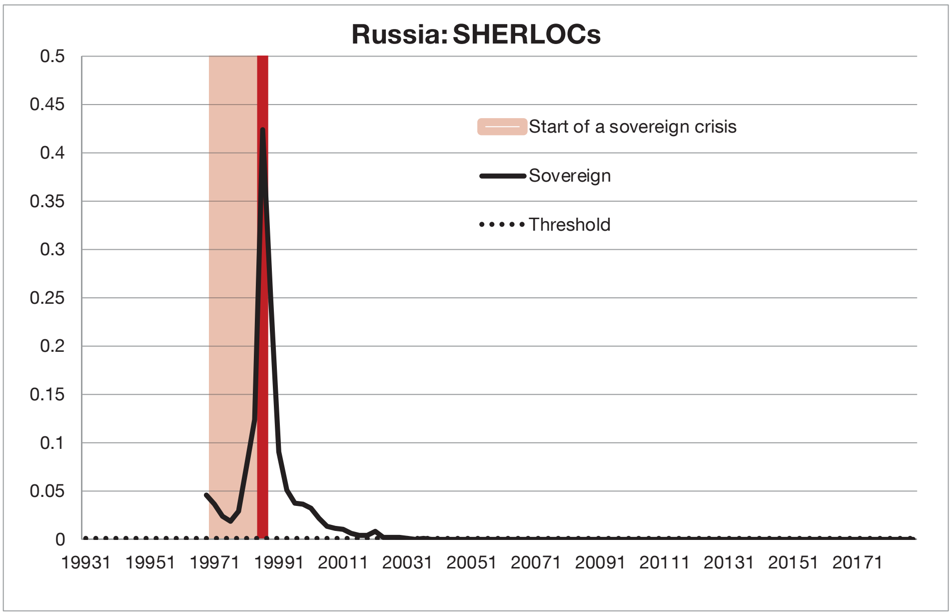

Mexico: SHERLOCs
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a banking crisis
Banking
Threshold

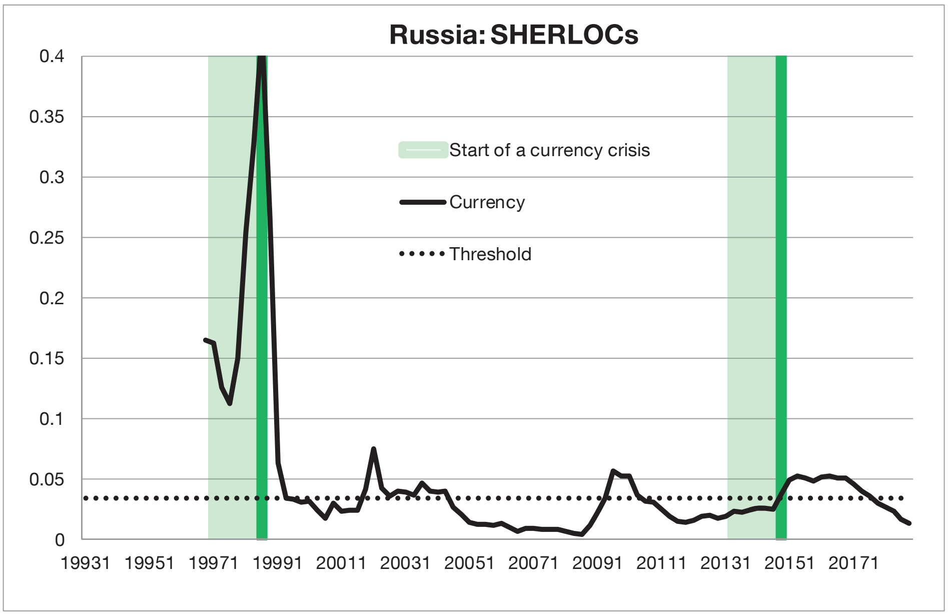

Mexico: SHERLOCs
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a sovereign crisis
Sovereign
Threshold

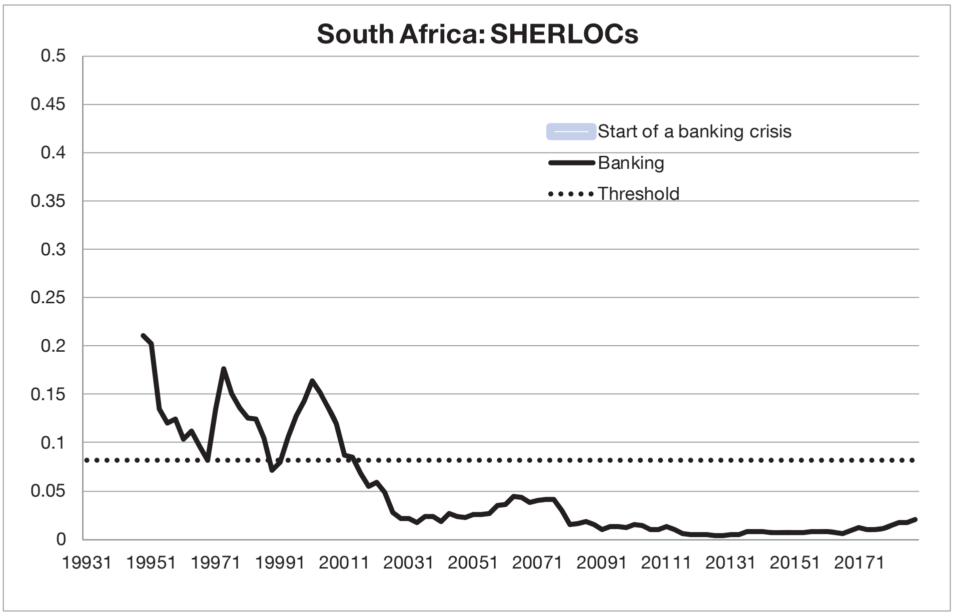

Mexico: SHERLOCs
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a currency crisis
Currency
Threshold

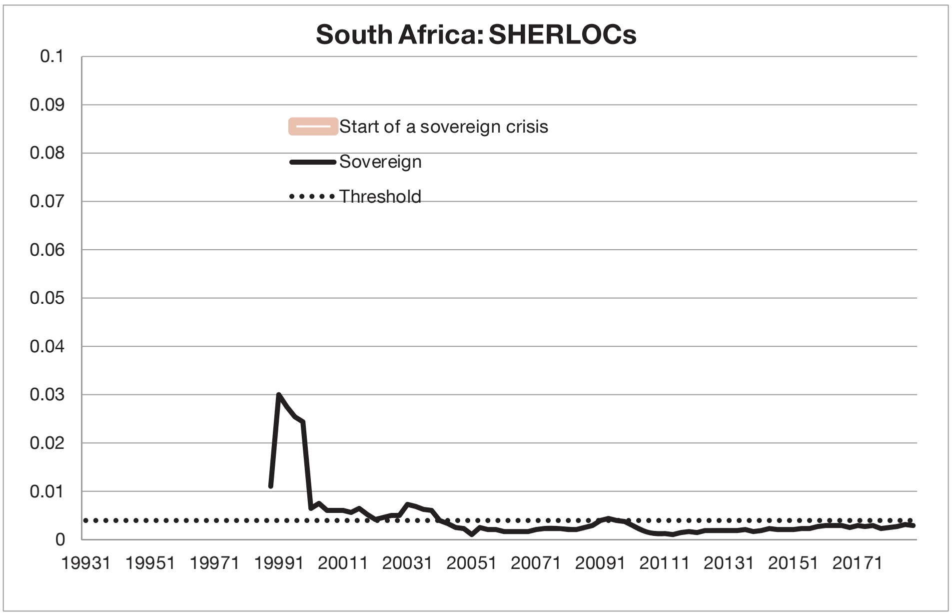
Appendix IX (7)

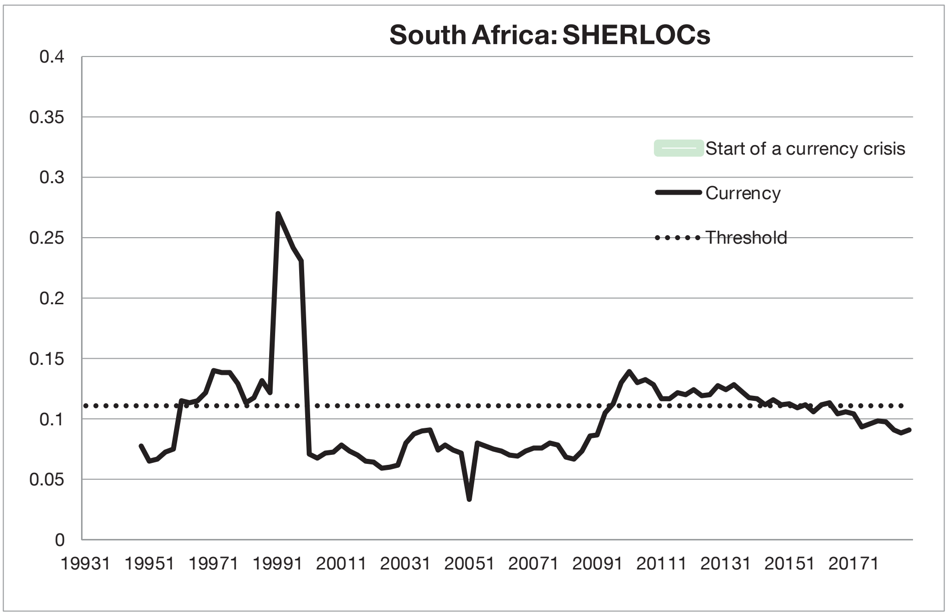


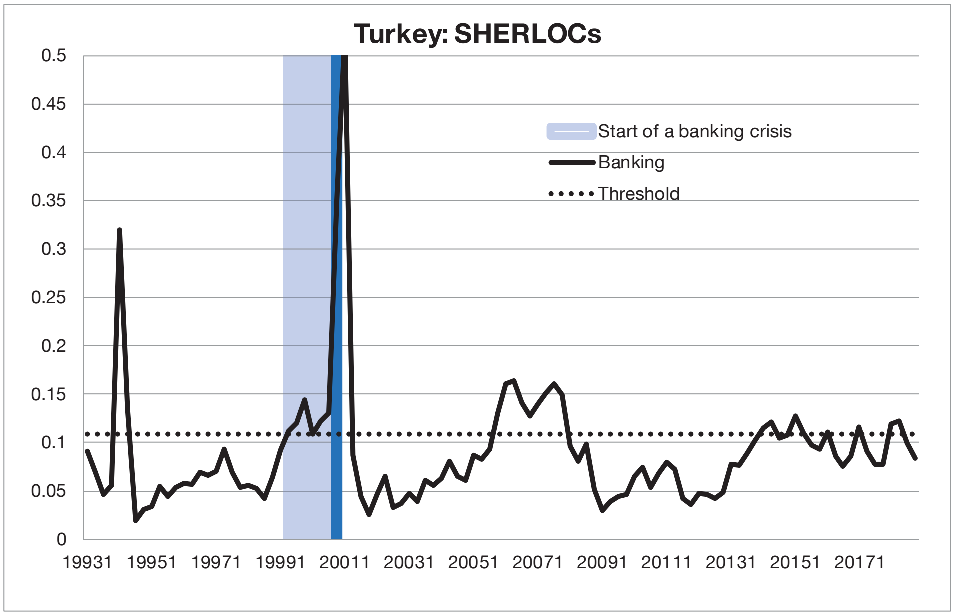


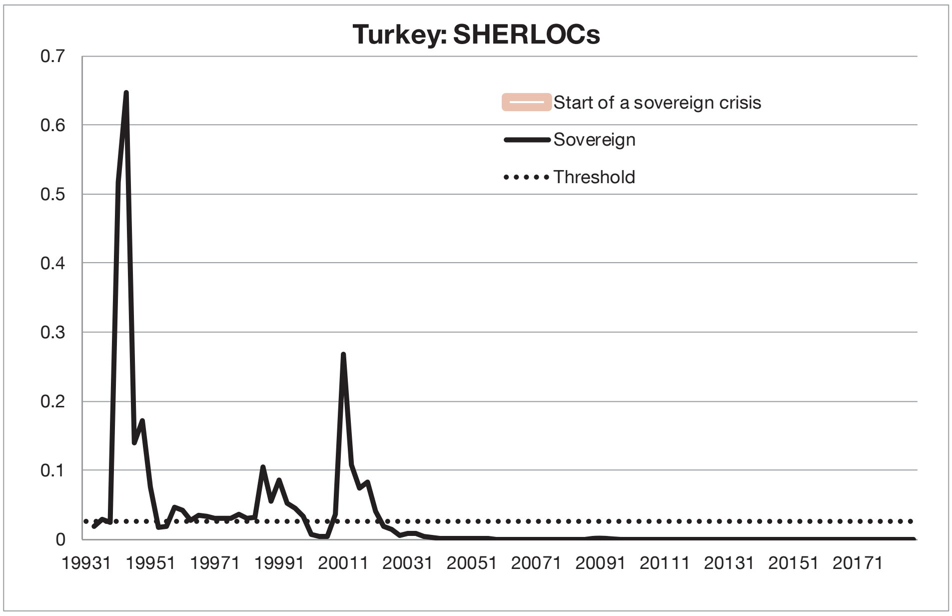


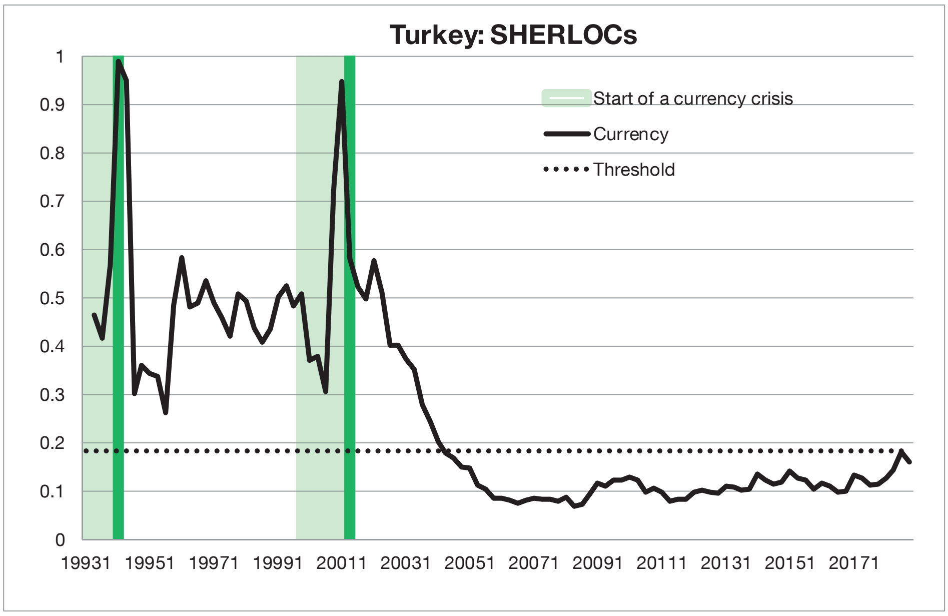

South Africa: SHERLOCs
Start of a banking crisis
Banking
Threshold
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

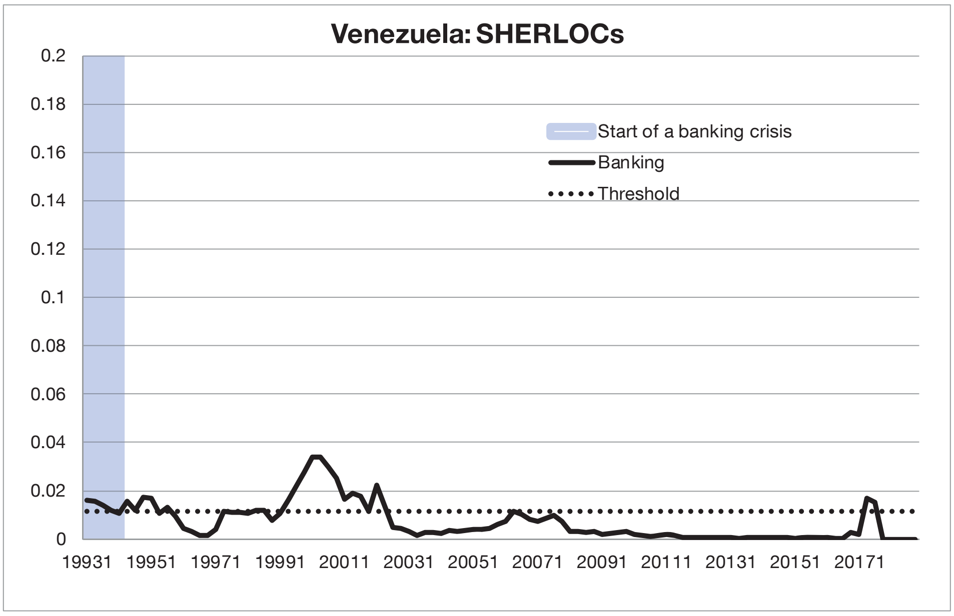

South Africa: SHERLOCs
Start of a sovereign crisis
Sovereign
Threshold
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

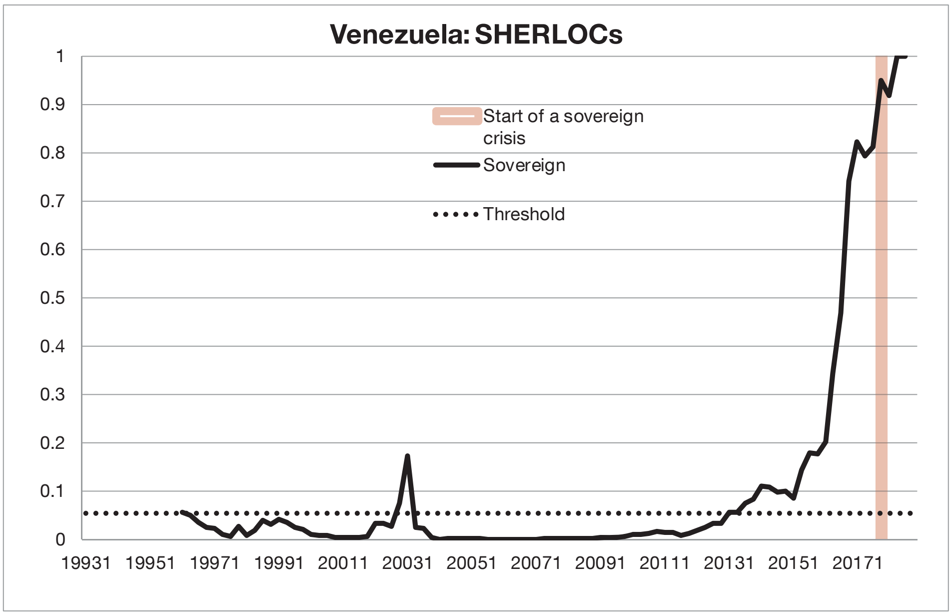

South Africa: SHERLOCs
Start of a currency crisis
Currency
Threshold
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

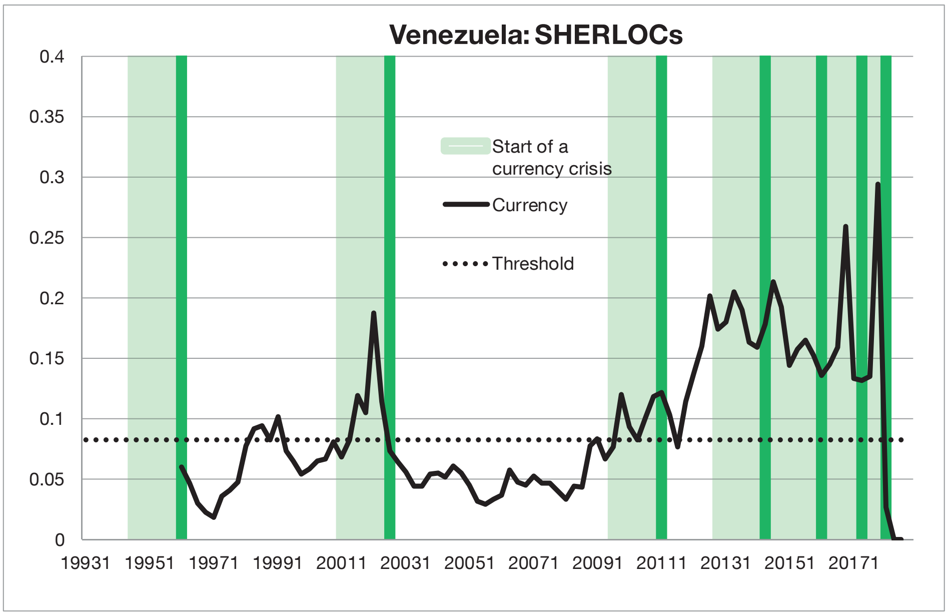

Turkey: SHERLOCs
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a banking crisis
Banking
Threshold


Turkey: SHERLOCs
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a sovereign crisis
Sovereign
Threshold


Turkey: SHERLOCs
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171
Start of a currency crisis
Currency
Threshold


Venezuela: SHERLOCs
Start of a banking crisis
Banking
Threshold
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171


Venezuela: SHERLOCs
Start of a sovereign crisis
Sovereign
Threshold
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171


Venezuela: SHERLOCs
Start of a currency crisis
Currency
Threshold
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
19931 19951 19971 19991 20011 20031 20051 20071 20091 20111 20131 20151 20171

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

1840 ALESSIO MORO and OMAR RACHEDI: The changing structure of government consumption spending.
1841 GERGELY GANICS, ATSUSHI INOUE and BARBARA ROSSI: Confidence intervals for bias and size distortion in IV and local projections – IV models.
1842 MARÍA GIL, JAVIER J. PÉREZ, A. JESÚS SÁNCHEZ and ALBERTO URTASUN: Nowcasting private consumption: traditional indicators, uncertainty measures, credit cards and some internet data.
1843 MATÍAS LAMAS and JAVIER MENCÍA: What drives sovereign debt portfolios of banks in a crisis context?
1844 MIGUEL ALMUNIA, POL ANTRÀS, DAVID LÓPEZ-RODRÍGUEZ and EDUARDO MORALES: Venting out: exports during a domestic slump.
1845 LUCA FORNARO and FEDERICA ROMEI: The paradox of global thrift.
1846 JUAN S. MORA-SANGUINETTI and MARTA MARTÍNEZ-MATUTE: An economic analysis of court fees: evidence from the Spanish civil jurisdiction.
1847 MIKEL BEDAYO, ÁNGEL ESTRADA and JESÚS SAURINA: Bank capital, lending booms, and busts. Evidence from Spain in the last 150 years.
1848 DANIEL DEJUÁN and CORINNA GHIRELLI: Policy uncertainty and investment in Spain.
1849 CRISTINA BARCELÓ and ERNESTO VILLANUEVA: The risk of job loss, household formation and housing demand: evidence from differences in severance payments.
1850 FEDERICO TAGLIATI: Welfare effects of an in-kind transfer program: evidence from Mexico.
1851 ÓSCAR ARCE, GALO NUÑO, DOMINIK THALER and CARLOS THOMAS: A large central bank balance sheet? Floor vs corridor systems in a New Keynesian environment.
1901 EDUARDO GUTIÉRREZ and ENRIQUE MORAL-BENITO: Trade and credit: revisiting the evidence.
1902 LAURENT CAVENAILE and PAU ROLDAN: Advertising, innovation and economic growth.
1903 DESISLAVA C. ANDREEVA and MIGUEL GARCÍA-POSADA: The impact of the ECB's targeted long-term refinancing operations on banks' lending policies: the role of competition.
1904 ANDREA ALBANESE, CORINNA GHIRELLI and MATTEO PICCHIO: Timed to say goodbye: does unemployment benefit eligibility affect worker layoffs?
1905 CORINNA GHIRELLI, MARÍA GIL, JAVIER J. PÉREZ and ALBERTO URTASUN: Measuring economic and economic policy uncertainty, and their macroeconomic effects: the case of Spain.
1906 CORINNA GHIRELLI, JAVIER J. PÉREZ and ALBERTO URTASUN: A new economic policy uncertainty index for Spain.
1907 ESTEBAN GARCÍA-MIRALLES, NEZIH GUNER and ROBERTO RAMOS: The Spanish personal income tax: facts and parametric estimates.
1908 SERGIO MAYORDOMO and OMAR RACHEDI: The China syndrome affects banks: the credit supply channel of foreign import competition.
1909 MÓNICA CORREA-LÓPEZ, MATÍAS PACCE and KATHI SCHLEPPER: Exploring trend inflation dynamics in Euro Area countries.
1910 JAMES COSTAIN, ANTON NAKOV and BORJA PETIT: Monetary policy implications of state-dependent prices and wages.
1911 JAMES CLOYNE, CLODOMIRO FERREIRA, MAREN FROEMEL and PAOLO SURICO: Monetary policy, corporate finance and investment.
1912 CHRISTIAN CASTRO and JORGE E. GALÁN: Drivers of productivity in the Spanish banking sector: recent evidence.
1913 SUSANA PÁRRAGA RODRÍGUEZ: The effects of pension-related policies on household spending.
1914 MÁXIMO CAMACHO, MARÍA DOLORES GADEA and ANA GÓMEZ LOSCOS: A new approach to dating the reference cycle.
1915 LAURA HOSPIDO, LUC LAEVEN and ANA LAMO: The gender promotion gap: evidence from Central Banking.
1916 PABLO AGUILAR, STEPHAN FAHR, EDDIE GERBA and SAMUEL HURTADO: Quest for robust optimal macroprudential policy.
1917 CARMEN BROTO and MATÍAS LAMAS: Is market liquidity less resilient after the financial crisis? Evidence for US treasuries.
1918 LAURA HOSPIDO and CARLOS SANZ: Gender Gaps in the Evaluation of Research: Evidence from Submissions to Economics Conferences.
1919 SAKI BIGIO, GALO NUÑO and JUAN PASSADORE: A framework for debt-maturity management.
1920 LUIS J. ÁLVAREZ, MARÍA DOLORES GADEA and ANA GÓMEZ-LOSCOS: Inflation interdependence in advanced economies.

1921 DIEGO BODAS, JUAN R. GARCÍA LÓPEZ, JUAN MURILLO ARIAS, MATÍAS J. PACCE, TOMASA RODRIGO LÓPEZ, JUAN DE DIOS ROMERO PALOP, PEP RUIZ DE AGUIRRE, CAMILO A. ULLOA and HERIBERT VALERO LAPAZ: Measuring retail trade using card transactional data.

1922 MARIO ALLOZA and CARLOS SANZ: Jobs multipliers: evidence from a large fiscal stimulus in Spain.

1923 KATARZYNA BUDNIK, MASSIMILIANO AFFINITO, GAIA BARBIC, SAIFFEDINE BEN HADJ, ÉDOUARD CHRÉTIEN, HANS DEWACHTER, CLARA ISABEL GONZÁLEZ, JENNY HU, LAURI JANTUNEN, RAMONA JIMBOREAN, OTSO MANNINEN, RICARDO MARTINHO, JAVIER MENCÍA, ELENA MOUSARRI, LAURYNAS NARUŠEVIČIUS, GIULIO NICOLETTI, MICHAEL O'GRADY, SELCUK OZSAHIN, ANA REGINA PEREIRA, JAIRO RIVERA-ROZO, CONSTANTINOS TRIKOUPIS, FABRIZIO VENDITTI and SOFÍA VELASCO: The benefits and costs of adjusting bank capitalisation: evidence from Euro Area countries.

1924 MIGUEL ALMUNIA and DAVID LÓPEZ-RODRÍGUEZ: The elasticity of taxable income in Spain: 1999-2014.

1925 DANILO LEIVA-LEON and LORENZO DUCTOR: Fluctuations in global macro volatility.

1926 JEF BOECKX, MAARTEN DOSSCHE, ALESSANDRO GALESI, BORIS HOFMANN and GERT PEERSMAN: Do SVARs with sign restrictions not identify unconventional monetary policy shocks?

1927 DANIEL DEJUÁN and JUAN S. MORA-SANGUINETTI: Quality of enforcement and investment decisions. Firm-level evidence from Spain.

1928 MARIO IZQUIERDO, ENRIQUE MORAL-BENITO and ELVIRA PRADES: Propagation of sector-specific shocks within Spain and other countries.

1929 MIGUEL CASARES, LUCA DEIDDA and JOSÉ E. GALDÓN-SÁNCHEZ: On financial frictions and firm market power.

1930 MICHAEL FUNKE, DANILO LEIVA-LEON and ANDREW TSANG: Mapping China's time-varying house price landscape.

1931 JORGE E. GALÁN and MATÍAS LAMAS: Beyond the LTV ratio: new macroprudential lessons from Spain.

1932 JACOPO TIMINI: Staying dry on Spanish wine: the rejection of the 1905 Spanish-Italian trade agreement.

1933 TERESA SASTRE and LAURA HERAS RECUERO: Domestic and foreign investment in advanced economies. The role of industry integration.

1934 DANILO LEIVA-LEON, JAIME MARTÍNEZ-MARTÍN and EVA ORTEGA: Exchange rate shocks and inflation comovement in the euro area.

1935 FEDERICO TAGLIATI: Child labor under cash and in-kind transfers: evidence from rural Mexico.

1936 ALBERTO FUERTES: External adjustment with a common currency: the case of the euro area.

1937 LAURA HERAS RECUERO and ROBERTO PASCUAL GONZÁLEZ: Economic growth, institutional quality and financial development in middle-income countries.

1938 SILVIA ALBRIZIO, SANGYUP CHOI, DAVIDE FURCERI and CHANSIK YOON: International Bank Lending Channel of Monetary Policy.

1939 MAR DELGADO-TÉLLEZ, ENRIQUE MORAL-BENITO and JAVIER J. PÉREZ: Outsourcing and public expenditure: an aggregate perspective with regional data.

1940 MYROSLAV PIDKUYKO: Heterogeneous spillovers of housing credit policy.

1941 LAURA ÁLVAREZ ROMÁN and MIGUEL GARCÍA-POSADA GÓMEZ: Modelling regional housing prices in Spain.

1942 STÉPHANE DÉES and ALESSANDRO GALESI: The Global Financial Cycle and US monetary policy in an interconnected world.

1943 ANDRÉS EROSA and BEATRIZ GONZÁLEZ: Taxation and the life cycle of firms.

1944 MARIO ALLOZA, JESÚS GONZALO and CARLOS SANZ: Dynamic effects of persistent shocks.

1945 PABLO DE ANDRÉS, RICARDO GIMENO and RUTH MATEOS DE CABO: The gender gap in bank credit access.

1946 IRMA ALONSO and LUIS MOLINA: The SHERLOC: an EWS-based index of vulnerability for emerging economies.

BANCO DE ESPAÑA Eurosistema	Unidad de Servicios Auxiliares Alcalá, 48 - 28014 Madrid E-mail: publicaciones@bde.es www.bde.es