S Y S T E M I C R I S K

Understanding the sources of systemic risk and vulnerabilities that may lead to a systemic fi nancial crisis is of central importance to policy-makers as it allows policy actions to be taken to prevent the further build-up of vulnerabilities or to enhance the shockabsorption capacity of the fi nancial system. Much of the empirical literature deals with early warning systems that rely on conventional statistical modelling methods, such as the univariate “signals” approach or multivariate logit/probit models. For example, Berg and Pattillo apply a discrete choice model to predicting currency crises; Fuertes and Kalotychou to predicting debt crises; and Lo Duca and Peltonen to predicting systemic crises. Berg et al. provide a comprehensive review of the literature. Given the changing nature of fi nancial crises, stand-alone numerical predictions are unlikely to convey a comprehensive picture. This motivates the development of tools with clear visual capabilities to complement numerical predictions. The dimensionality of the problem complicates visualisation, since a large number of indicators are often required to accurately assess vulnerabilities that could lead to a fi nancial crisis. In addition to the limitation of standard two and three-dimensional visualisations for describing higher dimensions, there is also the challenge of visualising temporal or cross-sectional information relevant for predicting financial stress. While composite indices of leading indicators and predicted crisis probabilities provided by early warning systems enable comparisons across countries and over time, these indices fall short in disentangling the individual sources of vulnerability. Methods for exploratory data analysis can, to some extent, overcome thes types of shortcomings. Exploratory data analysis attempts to describe the relevant phenomena in easily understandable forms. The Self-Organising Map(SOM) is a method that combines two groups of methods for exploratory data analysis: data and dimensionality reduction techniques.

The SOM provides a non-linear description of the multi-dimensional data distribution on a two-dimensional plane without losing sight ofindividual indicators. Thus, the two-dimensional output of the SOM makes it particularly useful for visualisations, or summarisations, of large amounts of information. We presents a methodology for mapping the state of financial stability and the sources of systemic risks using the SOM. The Self-Organising Financial

Stability Map (SOFSM) enables a two-dimensional representation of a multidimensional financial stability space that makes it possible to disentangle the individual sources of vulnerabilities which have an impact onsystemic risks. The SOFSM can be used to monitor macro-financial vulnerabilities by locating a particular country in the fi nancial stability cycle: either in the pre-crisis, crisis, post-crisis or tranquil state. In addition, the
SOFSM model performs at least as well as a standard logit model in classifying in-sample data and in making out-of-sample predictionsregarding the ongoing global f nancial crisis.

THE SELF-ORGANISING FINANCIAL STABILITY MAP

This section introduces the elements that are necessary to construct the SOFSM, namely the methodology based on the SOM; the dataset defi ning systemic financial crises and macro-f nancial vulnerabilities; the evaluation framework for assessing the signals of the SOFSM and their suitability for policy use; and the training of the SOFSM.
The SOM is a data and dimensionality reduction method that uses an unsupervised learning method developed by Kohonen. It maps input data onto a two-dimensional array of output nodes and attempts to preserve the neighbourhood relations in the data rather than the absolute distances.
On a two-dimensional grid, the numbers on the *x* and *y*-axes do not carry a numeric meaning in a parametric sense: they represent positions in the data space of the map, where each of these positions (*x*,*y*) is a mean profi le (cluster).
In addition, a second-level clustering can be applied on the nodes of the SOM. That is, data can be separated into nodes and nodes into clusters. The intuition of the basic SOM algorithm is presented below (see also Chart D.1).
The algorithm used in the analysis consists of constructing an SOM grid-based on a userspecified number of nodes, which represent the same dimensions (number of variables) as the actual dataset. Generally, the SOM algorithm operates according to the following step (see Chart D.1).

1. Compare all data points with all nodes to find the nearest node for each data point (i.e., the best-matching unit).
2. Update each node to averages of the attracte data, including data located in a specified neighbourhood.
3. Repeat steps 1 and 2 a specified number of times.
4. Group nodes into a reduced number of clusters using Ward's hierarchical clustering.

The parameters relevant for the SOM ar the radius of the neighbourhood, the number of nodes, the map format (ratio of *X* and *Y* dimensions), and the number of training iterations. Large radii result in stiff maps that stress topology preservation at the cost of quantisation accuracy, while small radii lead to a standard *k*-means clustering with no topologypreserving mapping. The number of nodes determines the granularity of the results.

For the purpose of this analysis, the output of the SOM algorithm is visualised on atwo-dimensional plane. The rationale for not using a one-dimensional map is that there are differences within clusters. A three-dimensional map, while adding a further dimension, impairs the interpretability of data visualisations. For each individual indicator, a feature plane represents the distribution of its values on the two-dimensional map. As the feature planes are different views of the same map, one unique point represents the same node on all planes. The feature planes are produced in colour. Cold colours represent low values of the indicator and warm colours represent high values, as defi ned by the colour scale below each feature plane. Shading on the two-dimensional map indicates the distance between each node and its corresponding second-level cluster centre, i.e. those close to the centre have a lighter shade and those farther away have a darker shade. Chart D.2 presents an example of a feature plane for a vulnerability indicator (real equity growth). Feature planes representing the distribution of individual variables on the SOFSM are created for the vulnerability indicators. SOM quality measures, such as quantisation error, distortion measure and topographic error, are usually used to determine the quality of a SOM. As the class information (financial stability cycle states) is available, classifi cation performance measures are used for evaluating the quality of the SOM.

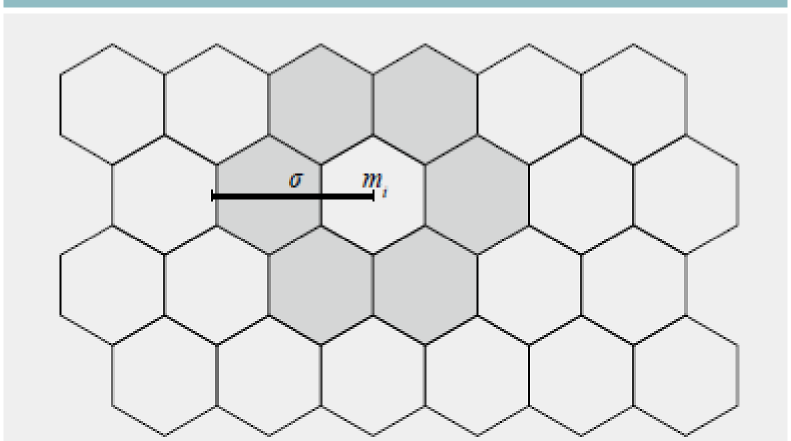
Systemic events and vulnerabilities

Systemic fi nancial crises are defi ned using a fi nancial stress index (FSI). This approach provides an objective criterion for the definition of the starting date of a systemic fi nancial crisis. The rationale behind the FSI is that the larger and broader the shock is (i.e. the more systemic the shock), the higher the co-movement among variables refl ecting tensions in different market segments (money, equity and foreignexchange market).
To define systemic financial crises, the FSI is first transformed into a binary variable.

In order to capture the systemic nature of the financial stress episodes, the focus is on episodes of extreme fi nancial stress that have, in the past, led (on average) to negative consequences for the real economy. In practice, a binary “crisis” variable is created that takes a value of 1 in the quarter when the FSI moves above the predefi ned threshold of the 90th percentile of its country-specifi c distribution, and otherwise. This approach identifi es a set of 94 systemic events between 1990 and 2010 for 28 economies (The countries in the sample are: Argentina, Australia, Brazil, China, the Czech Republic, Denmark, the euro area, Hong Kong, Hungary, India, Indonesia, Japan, Malaysia, Mexico, New Zealand, Norway, Poland, Russia, Singapore, South Africa, Sweden, Switzerland, Taiwan, Thailand, the Philippines, the United Kingdom, the United States and Turkey)

Chart D.3 illustrates the FSI and identifi ed systemic financial crises for Hong Kong SAR. To describe the fi nancial stability cycle, other class variables besides the crisis variable are created. First, a “pre-crisis” class variable is created by setting the binary variable to 1 for the 18 months preceding the systemic financial crisis, and to 0 for all other periods. The pre-crisis variable mimics an ideal leading indicator that perfectly signals a systemic fi nancial crisis in the 18 months before the event. Similarly, a “postcrisis” class variable is set to 1 for the 18 months after the systemic event. Finally, all other time periods are classifi ed as “tranquil” periods

Chart D.1 The Self-Organising Map

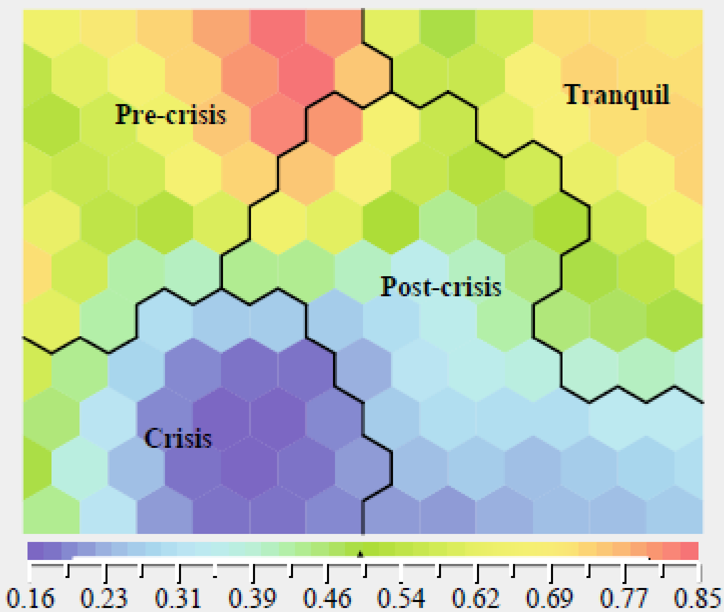


Notes: The chart displays the self-organisation of the SOM. First, each data point is assigned to its best-matching unit m_i. Then, each node m_i is updated to averages of attracted data, including data located in a specified neighbourhood σ.

To analyse the sources of systemic risk and vulnerability, a set of indicators consisting of commonly used metrics in the macroprudential literature for capturing the build-up of vulnerabilities and imbalances in the domestic and global economy are used.12 The key variables are asset price developments and valuations, and variables proxying for credit developments and leverage. In addition, other common variables (e.g. government budget defi cit and current account defi cit) are used to control for vulnerabilities stemming from macroeconomic imbalances.

Chart D.2 An example of a feature plane for a vulnerability indicator

(Q1 1990 – Q1 2005)

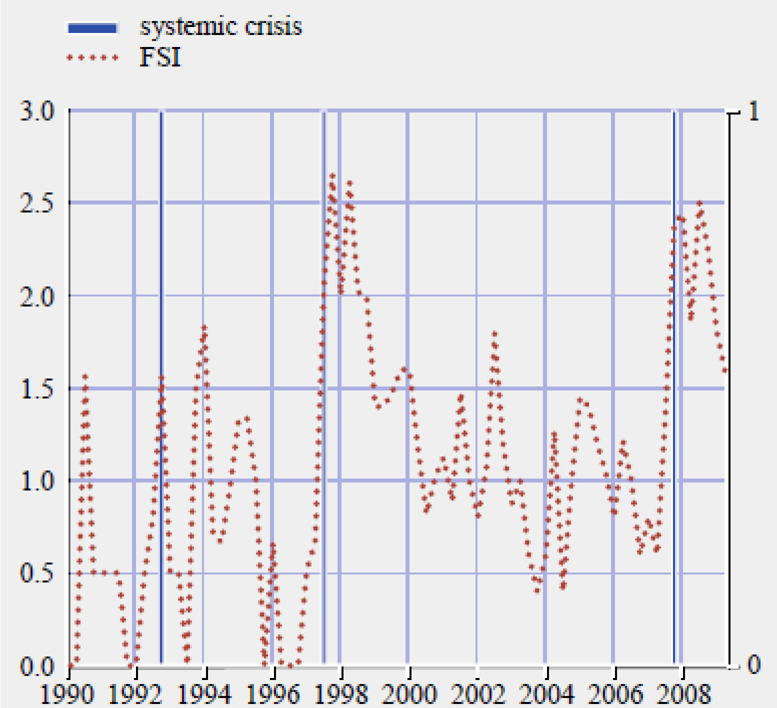


Sources: P. Sarlin and T.A. Peltonen, "Mapping the state of financial stability", *ECB Working Paper Series*, No 1382, 2011.
Notes: Feature planes are layers of the SOFSM in Chart D.4. This feature plane shows the distribution of a vulnerability indicator (real equity growth) on the grid. Cold colours represent low values of the indicator and warm colours represent high values, as defined by the colour scale.

In line with the literature, several transformations of the indicators (e.g. annual changes and deviations from moving averages or trends) are constructed to proxy for misalignments and a build-up of vulnerabilities. To proxy for global macro-fi nancial imbalances and vulnerabilities, a set of global indicators are calculated by averaging the transformed variables for the United States, the euro area, Japan and the United Kingdom. The fi nal set of indicators is chosen based on their univariate performance in predicting systemic events.

Chart D.3 Financial stress index and identified systemic financial crises for Hong Kong SAR

(Q1 1990 – Q1 2009; the FSI is scaled to [0.3] (left-hand scale), while systemic crises are measured in terms of a binary variable [0.1] (right-hand scale))



Sources: Bloomberg and Haver Analytics.
Note: Systemic crises are defined as periods where the FSI exceeds the 90th percentile of its country-specific distribution. Over the sample period of 1990 to 2010, these periods of extremely high financial stress have (on average) led to negative real economic consequences. For more details, see M. Lo Duca and T.A. Peltonen, op. cit.

Evaluation framework

To evaluate the performance of models in terms of predicting systemic fi nancial crises and to calibrate an optimal model and threshold for policy action, the approach pioneered in Demirgüç-Kunt and Detragiache 14 is adapted with the technical implementation as in Alessi and Detken, which also accounts for differencesin class size. The loss function of the policymaker is thus defi ned as:

$$L(\mu)=\mu(FN/(FN+TP))+(1-\mu)(FP/(FP+TN)) \quad (1)$$

where the parameter μ represents the relative preference of the policy-maker between false negatives and false positives. When $\mu = 0.5$, the policy-maker is equally concerned about missing crises and issuing false signals. The policy-maker is less concerned about issuing false alarms when $\mu > 0.5$ and more concerned when $\mu < 0.5$. To fi nd out the usefulness of a prediction, the loss is subtracted from the expected value of a guess with the given preferences, i.e. $Min(\mu,1-\mu)$. From this, the usefulness of the model is obtained:

$$U = Min(\mu, 1-\mu) - L(\mu). \qquad (2)$$

When using the above framework with a predefi ned preference parameter value, crisis and tranquil events are classifi ed by setting the threshold on the probability of a crisis so as to maximise the usefulness of the model for policy action. The extent to which policy-makers might be more or less concerned about failing to identify an impending crisis than issuing a false alarm is not explicitly assessed. The benchmark preference parameter of 0.5 belongs to a policymaker who is equally concerned about missing crises as issuing false alarms.

Training and evaluating the Self-Organising Financial Stability Map

In the analysis, a semi-supervised SOM is employed by making use of the information about the class variables in the model training. To partition the map according to the stages in the fi nancial stability cycle, the secondlevel clustering is performed using Ward clustering with respect to the class variables (see Chart D.4). The crisp clustering, which is given by the lines that separate the map into four parts, should only be interpreted as an aid forfinding the four stages of the fi nancial stability cycle, not as completely distinct clusters.
The predictive feature of the model is obtained by assigning to each data point the pre-crisis value of its best-matching unit.18 The performance of a model is evaluated using the framework introduced earlier based on the usefulness criterion for a policy-maker. The performance is computed using static and pooled models, i.e. the coefficients or maps are not re-estimated recursively over time or across countries. To test the predictability of the ongoing global fi nancial crisis, the sample is split into two sub-samples: the training sample from the fourth quarter of 1990 to the first quarter of 2005, and the test sample from the second quarter of 2005 to the second quarter of 2009. The training framework and choice of the SOM is implemented so that: (1) the model does not overfi t the in-sample data (parsimonious); (2) the framework does not include out-of-sample performance (objective); and (3) visualisation is taken into account (interpretability). The chosen SOM has 137 nodes on an 11x13 grid.
A standard logit model is estimated using the same in-sample data as was used for the SOFSM and later used for classifying in-sample data and predicting out-of-

sample data. For the benchmark models, the SOFSM and the logit model perform similarly overall (see Table D.1 and Chart D.5). With regard to the training set, the SOFSM performs slightly better than the logit model in terms of usefulness, recall positives, precision negatives and the area under the curve measure, while the logit model outperforms on the other measures. The classifi cation of the models are of opposite natures, as the SOFSM issues more false alarms (FP rate = 31%) than it misses crises (FN rate = 19%), whereas the logit model misses more crises (31%) than it issues false alarms (19%). That also explains the difference in the overall accuracy, since the class sizes are unbalanced (around 20% crisis periods and 80% tranquil periods). The difference in performance of the models on the test set is similar to the training set, except for the SOFSM having slightly higher overall accuracy. This is, in general, due to the higher share of crisis episodes in the out-of-sample dataset.

MAPPING THE STATE OF FINANCIAL STABILITY

Detecting signs of vulnerabilities and potential for contagion

In contrast to early warning systems which use binary classifi cation methods, such as discrete choice techniques, the SOFSM enables simultaneous assessment of the correlations with all four stages of the fi nancial stability cycle. Thus, new models need not be derived for different forecast horizons or defi nitions of the dependent variable. By assessing the feature planes of the SOFSM, the following strong correlations are found. First, one can differentiate between early and late signs of systemic crises by assessing differences within the pre-crisis cluster. The strongest early signs of a crisis (upper right part of the cluster) are increases in high domestic and global real equity growth and equity valuation, while most important late signs of a crisis (lower left part of the cluster) are increases in domestic and global real GDP growth as well as domestic real credit growth, leverage, budget surplus and current account deficits. Second, the highest values of global leverage and real credit growth in the crisis cluster exemplify the fact that increases in some indicators may refl ect a rise in fi nancial stress only up to a specifi c threshold. Increases beyond that level are, in this case, morem concurrent than preceding signals of a crisis. Similarly, budget deficits characterise the late post-crisis and early tranquil periods. The topological ordering of the SOFSM enables the assessment, in terms of macro financial conditions, of neighbouring financial states of a particular position on the map. While transmission of fi nancial contagion is often defi ned by neighbourhood measures like fi nancial or trade linkages, proxies of financial shock propagation, equity market co-movement or geographical relations, the SOFSM neighbourhood is based upon macro financial vulnerabilities. When assessing the SOFSM, the concept of neighbourhood of a country represents the similarity of the current macro-fi nancial conditions. Thus, a crisis in one position on the map indicates propagation of financial instabilities to adjacent locations. This type of representation may help to identify the changing nature of crises that surpasses historical experience.

Temporal analysis of the euro area and the United States

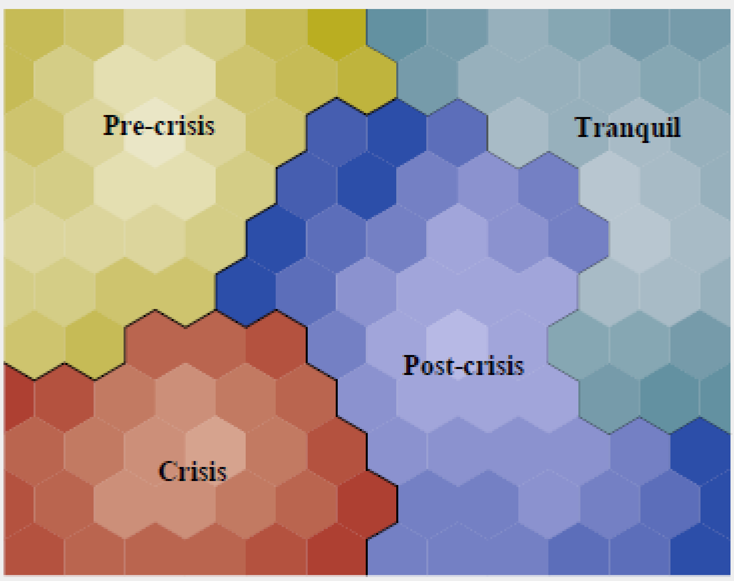
Employing the methodology for the United States and the euro area for the fi rst quarters of 2002 to 2011 and the second quarter of 2011, which is the latest data point in the sample, the SOFSM clearly recognises the pre-crisis, crisis and post-crisis stages of the fi nancial stability cycle for both economies (see ChartD.6). Regarding the capability of the SOFSM to make out-of-sample predictions of the onset of the current fi nancial crisis, the following observations can be made. First, the SOFSM maps the United States and the euro area in the pre-crisis state as early as the fi rst quarter of 2006 (see Chart D.6). While in the first quarter of 2007 the United States remains in the pre-crisis state, the SOFSM maps the euro area in the crisis state. Then, when moving to the first quarter of 2008, both the United States and the euro area are classifi ed as being in the crisis state. Looking at more recent periods, the euro area is located in the tranquil cluster in the fi rst quarter of 2010. This indicates that the aggregated macro-prudential metrics for the euro area as a whole did not refl ect the crisis in certain euro area countries. It also coincides with a relatively low financial stress index for the aggregate euro area at that point in time. This can be explained by the vulnerabilities and financial stress in smaller economies being averaged out by better conditions in the larger euro area countries (e.g. Germany), highlighting the importance of country-level analysis. It also stresses the importance of including a broad set of vulnerability indicators in the SOFSM and of cross-checking with other models. The macro-fi nancial vulnerabilities currently used in the SOFSM are best suited for capturing the build-up of vulnerabilities in the form of boom-bust cycles. However, they are less useful in identifying situations, where, for example, bank funding constraints or counterparty risks in a post-crisis recovery phase cause elevated fi nancial stress that feeds back to the real economy, increasing the probability of a financial crisis. Furthermore, by using the traditional macro-fi nancial vulnerabilities, it is rather difficult to capture situations where, as in the current crisis, self-fulfilling expectations drive the equilibrium outcomes. Nevertheless, according to the SOFSM, in the second quarter of 2011 the euro area moved to the border of the pre-crisis cluster. With a policy-maker's preference parameter μ =0.4, this particular location in the SOFSM is an early warning unit. At the same time, the United States was located in the post-crisis cluster in the first quarter of 2010 and in the tranquil cluster in the second quarter of 2011.

CONCLUDING REMARKS

This special feature puts forward a Self-Organising Financial Stability Map (SOFSM) based upon data and dimensionality reduction methods for mapping the state of financial stability and visualising the sources of systemic risks. Moreover, the SOFSM can be used as an early warning system, and to analyse contagion on the basis of similarities in macro-fi nancial vulnerabilities across countries. According to the results, the SOFSM makes an out-of-sample prediction identifying the onset of the global financial crisis in the United States as early as the first quarter of 2006

Chart D.4 The two-dimensional grid of the Self-Organising Financial Stability Map

(Q1 1990 – Q1 2005)



Sources: P. Sarlin and T.A. Peltonen, op. cit.
Notes: The chart displays the two-dimensional grid of the SOFSM, which represents a multi-dimensional financial stability space. The lines separate the SOFSM into four parts which represent the financial stability states, but should only be interpreted as an aid for finding the states, not as completely distinct clusters. Within each cluster, the shading on the SOFSM shows the distance of each node to the centre of the financial stability state.

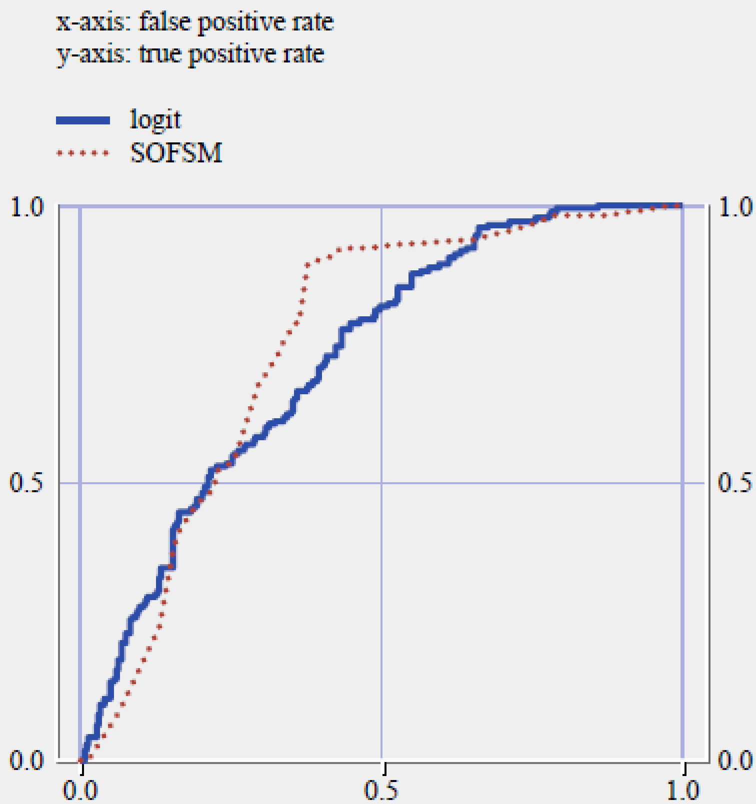
Table D.1 In-sample and out-of-sample results for the SOFSM and the logit model

							Positives		Negatives				
Mode	Data set	Threshold	TP	FT	TN	FN	Precision	Recall	Precision	Recall	Accuracy	Usefulness	AU
Logit	Training	0.72	162	190	830	73	0.46	0.69	0.92	0.81	0.79	0.25	0.
SOFSM	Training	0.60	190	314	706	45	0.38	0.81	0.94	0.69	0.71	0.25	0.
Logit	Test	0.72	77	57	249	93	0.57	0.45	0.73	0.81	0.68	0.13	0.
SOFSM	Test	0.60	112	89	217	58	0.56	0.66	0.79	0.71	0.69	0.18	0.

Sources: P. Sarlin and T.A. Peltonen, op. cit.
Notes: The table reports results for the logit and SOFSM for the training and test datasets and the optimal threshold. Thresholds ar calculated for $\mu=0.5$ and a forecast horizon of six quarters. The table also reports the following measures to assess the performance of th models: TP = true positive, FP = false positive, TN= true negative, FN = false negative, precision positive = TP/(TP+FP), recall positive TP/(TP+FN), precision negative = TN/(TN+FN), recall negative = TN/(TN+FP), accuracy = (TP+TN)/(TP+TN+FP+FN), usefulnes (see formulae 1 and 2), and AUC = area under the ROC curve (TP rate to FP rate, see Chart D.5).

Chart D.5 Out-of-sample receiver operating characteristic (ROC) curves for the Self-Organising Financial Stability Map and the logit model

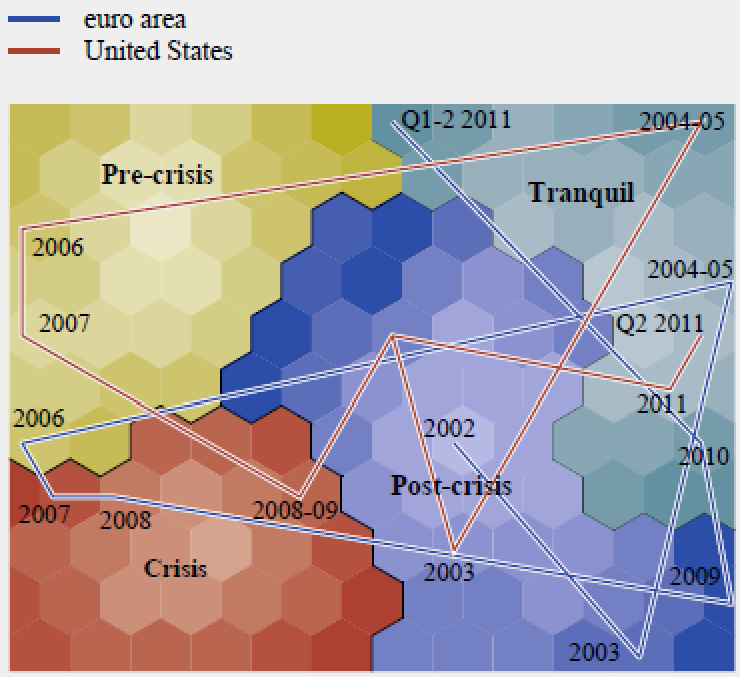
(Q2 2005 – Q2 2009)



Sources: P. Sarlin and T.A. Peltonen, op. cit.
Notes: The vertical and horizontal axes represent the true positive rate (TP/(TP+FN)) and the false positive rate (FP/(FP+TN)). The assumption used is that μ=0.5. The forecast horizon is 18 months.

Chart D.6 A mapping of the financial stability states of the United States and the euro area in the period 2002–11

(Q1 2002 – Q2 2011)



Sources: P. Sarlin and T.A. Peltonen, op. cit.
Notes: The chart displays the two-dimensional grid of the SOFSM. The lines that separate the map into four parts are based on the distribution of the four underlying financial stability states. Within each cluster, the shading on the SOFSM shows the distance of each node to the centre of the financial stability state. The data represent the first quarters of 2002 to 2011 and the second quarter of 2011. Data are mapped onto the grid by projecting them to their best-matching units. Consecutive time series data are linked by red and blue lines.