

RESEARCH

Open Access



Indicators of economic crises: a data-driven clustering approach

Maximilian Göbel* and Tanya Araújo

*Correspondence:
maximilian.goebel@phd.iseg.ulisboa.pt
UECE/REM- ISEG, Universidade de
Lisboa, R. Miguel Lupi 20, 1249-078,
Lisboa, Portugal

Abstract

The determination of reliable early-warning indicators of economic crises is a hot topic in economic sciences. Pinning down recurring patterns or combinations of macroeconomic indicators is indispensable for adequate policy adjustments to prevent a looming crisis. We investigate the ability of several macroeconomic variables telling crisis countries apart from non-crisis economies. We introduce a self-calibrated clustering-algorithm, which accounts for both similarity and dissimilarity in macroeconomic fundamentals across countries. Furthermore, imposing a desired community structure, we allow the data to decide by itself, which combination of indicators would have most accurately foreseen the exogeneously defined network topology. We quantitatively evaluate the degree of matching between the data-generated clustering and the desired community-structure.

Keywords: Early-Warning Models, Crisis Prediction, Macroeconomic Dynamics, Network Analysis, Community Structure, Great Recession, Clustering Algorithm

Introduction

Methods for community structure detection initially originated in the fields of physics. But nowadays the applications of these tools range from analyzing the rise of the Medici in medieval Florence (Padgett and Ansell 1993) to uncovering community formation in stock indices (Mantegna 1999). Within the field of economic sciences, network theory found entry into risk assessment and financial stability analyses. Demirer et al. (2017) find an overall increase in connectivity in the global banking system during times of crises, easing up the undamped transmission of locally occurring shocks to global threats. In a similar line of thought, Bostanci and Yilmaz (2020) find emerging markets to be the main transmitters of sovereign credit risk, i.e. the risk of a sovereign state to default on its debt, while the originators of these shocks (e.g. Greece, Portugal, Argentina) are rather placed in the periphery of the network. Another stream of studies centered their research on determining the drivers of various types of economic crises. Even if machine learning algorithms such as Artificial Neural Networks (ANN) (e.g. Fioramanti (2008); Sarlin (2014); Ristolainen (2018)) took the lead over community structure detection procedures

in the calibration of early-warning models, the limited interpretability of the output produced by the ANNs Ristolainen (2018), forces researchers to go in search for different methods.

Hence, machine learning algorithms such as regression tree analysis (e.g. Manasse and Roubini (2009)) or random forests (Goulet Coulombe 2020), but also community structure detection algorithms entered the literature on early-warning models of economic crises. Fuertes & Kalotychou (2007) applied the *k-means* algorithm as the basis of their early-warning model for predicting sovereign-debt crises. For determining the drivers of banking-crises, Marghescu et al. (2010) applied a *fuzzy c-means* algorithm, which allows data points – in their case a country-specific vector of indicator values for each of their time-series' months – to not only be assigned to one specific community, but to be partially split up into several clusters. Both approaches do, however, not decisively outperform conventional logit- or probit-models.

In a similar fashion as in Goulet Coulombe (2020), but directly applied to financial crises, Ward (2017) introduces *classification tree ensembles* (CTE) to the literature of early-warning models of financial crises. Using an impressively rich data set, covering the period between 1870 and 2011 on an annual, respectively the time between 1970 and 2011 on a quarterly basis, the CTEs outperform orthodox logit-regressions in the out-of-sample forecasting of banking crises. The best-performing CTE generates a false-positive rate of 30% while calling 90% of the banking crises correctly. Logit models on the other hand produce a 80% false-positive rate in order to reach the same percentage of correctly classified periods.

Rather than nominating a single best-performing model, Fouliard et al. (2019) deployed the framework of *online machine learning* to combine the power of several workhorse early-warning models of financial crises. By imposing time-varying weights, their algorithm selects a group of predictive models among a sample of potential candidates, in order to predict the probability of currently being in a pre-crisis period.

All the aforementioned studies were meant to calibrate a model that can predict a financial crisis, or its pre-crisis period, out-of-sample. Despite these multiple studies, the variables, driving countries onto trajectories leading the economy into a banking- or currency-crisis, are still obscure. Uncovering these, or at least adding another piece to the exploration of the determinants of economic crises, is not only of importance to policy-makers, but also crucial for a proper calibration of any type of early-warning model. Especially the very nature – time-varying or recurring – of these drivers of economic crises is an ongoing debate ((Fioramanti 2008; Reinhart and Rogoff 2008; Ristolainen 2018)). This is why we want to take a step back from a pure forecasting exercise and rather identify particular patterns during the periods leading up to a banking- and/or currency-crisis. Hence, we do not propose a forecasting methodology, but introduce a tool to assess the nature of different financial crises since the early 1990s. This allows us to state our own opinion on Reinhart & Rogoff's (2008) claim that "[w]hile each financial crisis no doubt is distinct, they also share striking similarities".

In a recent study, Göbel and Araújo (2020) explored the existence of patterns in macroeconomic dynamics during pre-crisis periods of several crisis-events. They examined the extent of homophily within a sample of 27 countries based on five macroeconomic variables, which are believed to be valid indicators of currency-crises ((Berg and Pattillo 1999b; Marghescu et al. 2010; Sarlin and Marghescu 2011)). Even

if a clear-cut identification of pure crisis-, respectively non-crisis, clusters did not emerge, the authors detected prominent and startling cross-country similarities. During the run-up period to the Asian crises of 1997/98, the authors detected a disconnected cluster of economies, which had suffered from at least a currency-crisis at the end of the respective pre-crisis period. Investigating the cross-country dynamics during the eight quarters preceding the Great Recession of 2007/08, Göbel and Araújo (2020) found an intriguing community structure. The results revealed similarities in the five underlying macroeconomic variables in the United States, Greece, Spain, Portugal, and Italy – the big players in the upcoming European Sovereign Debt crisis. This raises the questions about the adequacy of subsequent policy responses. While the United States sallied forth onto a path, which resulted in the longest economic expansion in its history, the trajectory of the European sovereigns was rather different.

In this paper, we continue in the spirit of Göbel and Araújo (2020). However, we will extend their approach in two major aspects: firstly, we introduce a clustering algorithm, which is calibrated as to comply with the theoretical underpinnings of this study as outlined in “[The clustering algorithm](#)”. In essence, this study introduces an algorithm that can help to answer the question about the existence of any combination of indicators, foreseeing a given topology of a network of countries.

Secondly, our methodology selects a subset of a given number of macroeconomic variables, in which crisis countries show the largest discrepancies from non-crisis economies throughout a given pre-crisis period. Whereas Göbel and Araújo (2020) used a static model of five indicators, we expand the set of variables and select a sub-sample of indicators according to their ability to tell future crisis countries apart from future non-crisis economies during a given pre-crisis period. In particular, we try to tackle the problem of indicator selection by assessing whether *any* combination of indicators could have separated upcoming crisis from non-crisis countries during the quarters before a particular crisis event.

In doing so, we address two issues, which Göbel and Araújo (2020) encouraged future research to work on: first, we search for early-warning indicators, and second, we propose a clustering algorithm that is designed to account for both *similarity* and *dissimilarity* among countries.

To do so, the remainder of the study is structured as follows: “[Data](#)” gives an overview of the underlying data set. In “[Methodology](#)” we outline the specific methodology applied to answer the questions at hand. That is, we explain our measurement of similarity and introduce our clustering and indicator-selection algorithm. We proceed in “[Run-Up phase & reference dates](#)” with an overview of the different crisis-events, observed in our sample. In “[Pre-Crisis macroeconomic similarities & crisis occurrence](#)” we propose two network topologies, which would match a homogeneous clustering of crisis and non-crisis economies and report that combination of indicators, which most accurately matches the desired structures. “[Conclusion](#)” concludes.

Data

The dataset comprises 10 macroeconomic variables, observed on a quarterly basis between 1990 and 2018. Their abbreviations and the corresponding names are listed in Table 1. The sources range from Thomson-Reuters’ Datastream, the World Bank, the

Table 1 Description of Macroeconomic Variables

Indicator	Description
PubDebt / GDP	Public-Debt in relation to GDP
EX	Exports
IM	Imports
Reserves	International Reserves
GDS / GDP	Gross-Domestic-Savings (% of GDP)
STDebt / Reserves	Outstanding International Debt Coming Due < 1 Year International Reserves
RealFX	Real Exchange Rate [<i>Deviation from Trend</i>]
GFCE / GDP	General-Government-Final-Consumption-Expenditure (in % of GDP)
CA / GDP	Current-Account-Balance in relation to GDP
GFCF / GDP	Gross-Fixed-Capital-Formation (in % of GDP)

IMF, the OECD, the Bank for International Settlements, the European Central Bank, the Federal Reserve Bank of St. Louis' FRED database, via national statistics offices or central banks to Kaminsky's (2007) database. The advantage of sticking to only one single data provider - in order to ensure data consistency - was sacrificed for compiling a richer dataset. The sample covers 27 advanced and developing economies.

Compared to other studies (see Ward (2017)) our sample size may appear rather limited in all its three dimensions: countries, time period, and indicators. Our methodology, however, requires a fully-balanced panel. This is also why we chose a quarterly frequency of observations over monthly data, as many time series are not reported on a monthly frequency. Annual data would have allowed us to increase the sample size in each dimension, but would have at the same time significantly reduced the number of observation for each pre-crisis period – assuming a pre-crisis period to not last longer than two years. Nonetheless, data availability remained a major issue. Including more than the 10 macroeconomic variables, would have forced us to significantly reduce the number of countries. However, we wanted to paint a picture as complete as possible for understanding the similarities and differences among a heterogeneous set of countries. We therefore faced a triple trade-off between time-coverage, number of countries and indicator availability.

Sticking to quarterly observations did not prevent us from having a balanced panel upfront. In the rare case of quarterly observations starting after 1990, the quarterly data was augmented by interpolating the yearly observations via cubic splining, using R's algorithm *'spline'*¹ in case the annualized quarterly data matched the yearly observations. If annualized quarterly data did not match the yearly observations, the interpolated yearly data substituted the quarterly observations. Also a simple linear interpolation would have helped to fill the gaps, but as described and motivated in “[The euclidean distance](#)”, we later on transformed the data into percentiles, which mitigates the effect of using different interpolation methodologies.

Last but not least, we want to emphasize that one of the strengths of data mining/machine learning techniques is their capability to deal with large amounts of data (Athey and Imbens 2019), much larger than this paper's underlying data set. Our methodology is not any different from that. Our initial goal was to further increase the number of countries, variables and quarters than those in Göbel and Araújo (2020), but adequate data availability prevented us from doing so. Nevertheless, our proposed methodology is not limited to a certain sample size.

¹See Becker et al. (1988); Forsythe et al. (1977) for explanations of the algorithm.

Methodology

The present work follows Göbel and Araújo (2020) in assuming that the bilateral similarity among countries – measured on the basis of macroeconomic indicators over a certain period of time – determines a similar or even common subsequent economic development.

Even if the importance of accounting for beliefs and expectations of market participants intruded the literature on early-warning models already in the mid 1990s ((Obstfeld 1994; Eichengreen et al. 1995; Obstfeld 1996)), sentiment does not emerge out of nowhere. We assume expectations to rest on any sort of real-world observation or experience. Whether this perception conforms with rational thinking or not, is irrelevant. The essential rationale, underlying our considerations assumes the expectation-generating process to be rooted in any observed *past* and *present* economic dynamics, which make market participants form their view of the mid-term future.

The upcoming subsections describe the necessary tools for mapping these convictions into our assessment. Whereas “The euclidean distance” presents a measure of similarity, “The clustering algorithm” describes the community structure detection algorithm.

Measuring similarity

The euclidean distance

Following Göbel and Araújo (2020), we use the Euclidean distance as our measure of similarity. We transformed quarterly raw data, described in “Data”, into percentiles of each indicator’s time-series distribution. For determining the country-specific distribution of each indicator, we did not restrict the time-series to range from 1990 Q1 to 2018 Q4 only, but also included earlier observations if available.

Thus, each indicator’s percentiles are calculated based on the distribution covering its *entire* time-series. Following Göbel and Araújo (2020), these percentiles $p_{i,n,t}$ then form the input for the row-vector $\mathbf{v}_{n,t}$, where $p_{i,n,t}$ is the percentile of indicator i for country n at quarter t . $\mathbf{v}_{n,t}$ describes the $t \times I$ row-vector for country n , where I is the total number of indicators i , which the model is composed of. The Euclidean distance, as described in Gan et al. (2007), between country n and country z at a particular quarter t is defined as:

$$d_t(\mathbf{v}_{n,t}, \mathbf{v}_{z,t}) = \left[\sum_{i=1}^I (p_{i,n,t} - p_{i,z,t})^2 \right]^{\frac{1}{2}} = \left[(\mathbf{v}_{n,t} - \mathbf{v}_{z,t}) (\mathbf{v}_{n,t} - \mathbf{v}_{z,t})^T \right]^{\frac{1}{2}}, \quad (1)$$

where $p_{i,n,t}$ and $p_{i,z,t}$ are the percentiles of the i^{th} variable, respectively indicator, of country n , respectively z , at quarter t .

So far, this formula only measures the distance between any two countries n and z at one specific point in time, i.e. the bilateral distance in one specific quarter t . Extending the similarity measure to several periods, $\Delta t = [t, t + I]$, transforms the $t \times I$ row-vector $\mathbf{v}_{n,t}$ into the matrix $\Delta t \times I$. Thus, Eq. (1) is augmented by a time dimension as follows:

$$d_{\Delta t}(\mathbf{v}_{n,\Delta t}, \mathbf{v}_{z,\Delta t}) = \left[\sum_{t=1}^I \sum_{i=1}^I (p_{i,n,t} - p_{i,z,t})^2 \right]^{\frac{1}{2}} = \left[(\mathbf{v}_{n,\Delta t} - \mathbf{v}_{z,\Delta t}) (\mathbf{v}_{n,\Delta t} - \mathbf{v}_{z,\Delta t})^T \right]^{\frac{1}{2}}. \quad (2)$$

The drawback of Eq. (1) is the necessity for the matrix, $\mathbf{v}_{n,\Delta t}$, to display equal dimensions in the cross-section of countries, i.e. a balanced data set is crucial. This imposes the aforementioned trade-off between capturing as many crises as possible and omitting as few countries and indicators as necessary.

The Euclidean distance, as described above, is just one tool to fill the adjacency matrix, i.e. to measure the similarity between countries. For example, Demirer et al. (2017) create a directed network by filling the entry A_{zn} of the adjacency matrix A with the proportion that element n contributes to the forecast error variance of element z . Another metric was introduced by Mantegna (1999) and became a popular tool for assessing the topology of financial markets (Araújo and Louçã 2007; Spelta and Araújo 2012; Araújo and Göbel 2019). Transforming the correlation coefficient in such a way that it complies with all three axioms of a distance metric, Mantegna's (1999) measure is well suited for dealing with high-frequency stock market data. Another often-cited metric is the Mahalanobis distance. Other than the Euclidean distance, the Mahalanobis distance accounts for co-movement among variables. The type of co-movement of a given variable i in countries n and z can provide additional information about the similarity between those two countries. If the two series differ largely in their values as quoted – may the quotation be in levels or growth rates – the Euclidean distance, as measured in Eq. 2, will be quite large. Not accounting for a potential positive co-movement will therefore introduce an upward bias in the measurement of the *distance* between those two countries – and a downward bias in case of a negative correlation. However, also the Mahalanobis distance cannot alleviate this problem: as described in Eq. 2, we take indicator i for any two countries n and z and calculate the Euclidean distance at each time step t . Hence, we have a single time-series i , eight time-steps $\Delta t = 8$, and two countries $N = 2$. As in all our calculations we will face the issue that $\Delta t > N$, the covariance matrix, which is a central component of the Mahalanobis distance, is singular and not invertible (Prekopcsák and Lemire 2012). Thus, the plain-vanilla Mahalanobis distance is not applicable in our setup. Adding an additional weighting scheme, which accounts for cross-country correlation, could be a possible workaround. Prekopcsák and Lemire (2012) describe such extensions in more detail. Last but not least, one might be concerned with Eq. 2 assigning equal weights to the distances at each quarter/time-step t . One could argue that the distances, measured in the proximity of the crisis event, i.e. at $t = 6, 7, 8$, shall matter more than the distances at $t = 1, 2, 3$ – or even the other way around, if the effect of some variables materializes with a lag. Such extensions of time-varying weights are by no means possible. Nevertheless, following Göbel and Araújo (2020), we applied the Euclidean distance as our measurement of similarity, but we would like to stress that our approach is agnostic about the particular methodology used for filling the entries of the adjacency matrix.

Generating sparse networks

After calculating the bilateral Euclidean distances, we get a *weighted* but complete network. The resulting $\frac{N(N-1)}{2}$ edges provide information about those economies, which seem to be most similar to each other, but also reveal those countries with a very different macroeconomic profile. As we only want to group those economies with very similar fundamentals into the same cluster, an adequate filtering is inevitable. Such a filtering could be achieved on the one hand by conditioning on a certain *edge-weight* and on the

other hand by conditioning on preserving only a fraction of the $\frac{N(N-1)}{2}$ edges in the final network representation. Both require the determination of an adequate threshold.

Another popular filtering technique is the Minimal Spanning Tree (MST), initially introduced by Mantegna's study on the topology of a number of stock market indices (Mantegna 1999) and adopted in subsequent papers (e.g. Spelta and Araújo (2012); Araújo and Göbel (2019)). The $N - 1$ links, spanning the MST, create a connected network, which does not allow for disconnected components – *islands* or *cliques* – which is, however, a desired outcome, i.e. a disconnected cluster would at least indicate a completely different macroeconomic profile of countries within that cluster as compared to countries in other communities. Furthermore, the remaining $N - 1$ edges in the MST structure, do not necessarily represent the shortest of the unfiltered $\frac{N(N-1)}{2}$ edges.

Hence, we will abstract from the MST and truncate a certain number of the $\frac{N(N-1)}{2}$ links. The remaining edges form the set of shortest distances, i.e. they connect countries with a very similar macroeconomic profile. This requires, however, the definition of an adequate threshold, TT . Relaxing the threshold too generously may allow too much noise to enter the analysis, whereas a restrictive filtering may cut-off important information. As we assume the shortest distances between countries to hint at their subsequent economic state, we choose a truncation threshold to be based on the ranking of these similarities, i.e. the edge-weight. One possibility is to cut-off a certain percentage of edges (e.g. Piccardi et al. (2011)). This cut-off, however, varies with the number of nodes in the network. Another approach would keep the average degree constant and consider only the $c \times N$ shortest distances of the total $\frac{N(N-1)}{2}$ edges, with c being an arbitrary constant. Tests have, however, favored the percentage-cut-off as in Piccardi et al. (2011) over preserving the average degree in the network. Therefore, the upcoming analysis will set the threshold to $TT = 0.1 \times \frac{N(N-1)}{2}$.

The clustering algorithm

Having measured the cross-country similarities and having set a threshold for selecting the relevant distances, we now introduce the algorithm for uncovering the underlying community-structure. Popular clustering algorithms are Newman's (2006) *modularity* or the *k-means* approach, dating back to MacQueen (1967). Even if such algorithms might have become state of the art in network studies, some problem settings require their own individual specifications.

To recall, this paper assumes countries with similar dynamics over a certain period of time to subsequently experience a similar economic state. Consequently, countries with a rather dissimilar macroeconomic profile are supposed to be located on a different trajectory and not end up in the same economic state. The consequential premise is that large distances reveal larger differences in macroeconomic fundamentals between countries *relative* to small distances. This goes not without admitting that the distribution of edge-weights may be time-varying. The weight of a link, which ranges among the largest distances in one period, may not necessarily occupy the same rank in another period.

The important take-away is that *both* the smallest *and* the largest distances within the adjacency matrix capture important information for network formation processes. Focusing only on one end of the distribution discards valuable information about cross-country dynamics. An appropriate algorithm shall therefore account for distances on both ends of the edge-distribution and group very similar nodes into the same community, but at the

same time prevent vertices from being assigned to the same cluster if their mutual link ranges among the network's largest distances. We call the latter condition the *exclusion restriction*. Thus, we designed a *purely hierarchical* algorithm to operate as follows:

the network structure is built up on an *initial* cluster, which is formed by the network's shortest distance between any two of the N countries. Then the countries, which form the second strongest tie, enter the algorithm. Both nodes are attached to the already existing cluster, if any of the two nodes of this second-shortest distance, is already part of the existing community, and if - at the same time - the counterpart of this second-shortest distance does not form any direct interconnection, that ranges among largest distances, with any of the members of the existing cluster. If the counterpart does form such a long-distance connection with any of the members of the existing community, this node calls for the *exclusion restriction* and creates a cluster of its own. The *exclusion restriction* states that two countries may not be assigned to the same cluster, if their mutual link $d_{\Delta t}(\mathbf{v}_{n,\Delta t}, \mathbf{v}_{z,\Delta t}) \in \left[(1 - TT) \frac{N(N-1)}{2} + 1; \frac{N(N-1)}{2} \right]$ with TT being the truncation threshold (e.g. $TT = 0.1$) as described in "[Generating sparse networks](#)" and with the links being ordered ascendingly from shortest to largest. If none of the nodes, which form the incumbent link, is represented in an already existing cluster, then both the two countries form a new cluster as well. The process continues until all strongest edges have been examined. The assumption about the truncation threshold, TT , being symmetric, is of course alterable.

Run-Up phase & reference dates

The main goal of our analysis is the identification of a specific combination of indicators, which – based on the values of a particular *pre-crisis* period – generates a clustering structure with homogeneous crisis and non-crisis communities. Thus, a proper dating of the time and type of a specific crisis is indispensable.

We follow Göbel and Araújo (2020) and merge the *Exchange Market Pressure Index* – introduced to the literature on early-warning models by Kaminsky et al. (1996) – for the dating of currency-crises, the database of Laeven & Valencia (2018) for banking-crises and the extensive dataset of Reinhart & Rogoff (2009) for both types of crises. This combination yielded the *reference dates* displayed in Table 2 and the corresponding crisis and non-crisis flags as listed in Table 3. *Periods 1-4* mark the crisis events under investigation. As already noted in Göbel and Araújo (2020), the fact that some crisis-events *lead* and some *lag* a particular *reference date* t , requires a *country-specific* calibration of pre-crisis times. Otherwise, a uniform pre-crisis determination would not allow to capture the *distortion-free* pre-crisis macroeconomic dynamics in all of the countries. This results at times in a timely *shift* of pre-crisis periods for certain countries.

Following the seminal paper on early-warning models by Kaminsky et al. (1998), the *pre-crisis* period, respectively the *pre-reference-date* period, comprises eight quarters, ending in the quarter *prior* to the specific *reference date*, t .

Table 3 shows the period-specific economic status of each country. These crisis and non-crisis flags were determined by merging the aforementioned studies: the *Exchange Market Pressure Index* by Kaminsky et al. (1998), the databases of Laeven & Valencia (2018) and Reinhart & Rogoff (2009) and the study by Ristolainen (2018). Looking beyond the pre-crisis time frame, and being aware of previous and forthcoming events, complements the interpretation of the upcoming network structures. Thus, not only those

Table 2 *Reference Dates for Assessment Periods 1–4*

	Period 1	Period 2	Period 3	Period 4
Argentina	Q4 1991	Q1 1995	Q4 1997	Q3 2008
Australia	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Austria	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Belgium	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Brazil	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Canada	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Colombia	Q4 1991	Q4 1994	Q2 1998	Q3 2008
Finland	Q3 1991	Q4 1994	Q4 1997	Q3 2008
France	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Germany	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Greece	Q4 1991	Q4 1994	Q4 1997	Q3 2008
India	Q4 1991	Q4 1993	Q4 1997	Q3 2008
Indonesia	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Italy	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Japan	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Mexico	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Norway	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Netherlands	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Portugal	Q4 1991	Q4 1994	Q4 1997	Q3 2008
South Korea	Q4 1991	Q4 1994	Q3 1997	Q3 2008
Spain	Q4 1992	Q4 1994	Q4 1997	Q3 2008
Singapore	Q4 1991	Q4 1994	Q4 1997	Q3 2008
Sweden	Q3 1991	Q4 1994	Q4 1997	Q3 2008
Thailand	Q4 1991	Q4 1994	Q3 1997	Q3 2008
Turkey	Q4 1991	Q1 1994	Q4 1997	Q3 2008
United Kingdom	Q4 1991	Q4 1994	Q4 1997	Q3 2007
United States	Q4 1991	Q4 1994	Q4 1997	Q4 2007

countries, which suffered from a banking- and/or currency-crisis at the end of the respective period, are flagged with a crisis indication, but also those countries, which had either experienced some banking-sector and/or currency turbulences shortly prior to or at the onset of the respective pre-crisis period ($b;c;t$), were flagged accordingly. For a better interpretation of results, also crisis events happening shortly after the *reference date*, were accounted for ($b1;c1$).

Pre-Crisis macroeconomic similarities & crisis occurrence

In the first part of our analysis we test the validity of our procedure and try to loosen up the rather static model of Göbel and Araújo (2020), as some indicators might turn out to be redundant in explaining a specific clustering structure. As reported in Göbel and Araújo (2020), *Exports* did not seem to carry any discriminatory power in most of the periods.

In a second step, we feed the clustering algorithm a priori with a *targeted* network structure for each of the periods in Table 2. This *targeted* network structure is intended to resemble a community-structure with homogeneous crisis and non-crisis clusters. We also extend the scope of indicators from five to ten macroeconomic variables listed in Table 1. We are interested in the extent to which any combination of indicators could have generated our desired topology.

Table 3 Crisis Flagging: Assessment Periods 1–4

	Period 1	Period 2	Period 3	Period 4
Argentina	t	B; c	b	0
Australia	b*	b*	C	C
Austria	0	0	0	B
Belgium	0	0	0	B
Brazil	t	B; c	b	C
Canada	0	0	0	0
Colombia	c	C	T	0
Finland	B; c1	0	0	0
France	0	B*	0	B
Germany	0	0	0	B
Greece	B*; c	b*	0	B
India	0	B	b*	C
Indonesia	B*	b*	T	C
Italy	b*; C	b*	0	B
Japan	B	b	b	0
Mexico	B	T	b*	C
Norway	B; c1	0	0	C
Netherlands	0	0	0	B
Portugal	0	0	0	B
South Korea	0	0	T	0
Spain	C	0	0	B
Singapore	0	0	C	0
Sweden	B; c1	b	0	T
Thailand	0	0	T	0
Turkey	B*	C	0	C
United Kingdom	B*	B*	b*	B; c1
United States	b	0	0	B

Note: B = banking-crisis at the end of the period; C = currency-crisis at the end of the period; T = banking- & currency-crisis at the end of the period; b = banking-crisis occurrence within four quarters prior to the reference date; c = currency-crisis occurrence within four quarters prior to the reference date; t = twin-crisis occurrence within four quarters prior to the reference date; b1 = banking-crisis occurrence within four quarters after the reference date; c1 = currency-crisis occurrence within four quarters after the reference date

We then evaluate the degree to which the *targeted* clustering structure and the topology generated by the data coincide, according to the *level of satisfaction*, S , as described in “[Targeted network topologies](#)”.

Testing the methodology

In order to test our methodology, we calculate the cross-country similarities over the eight quarters preceding each of the four periods $P = 1, 2, 3, 4$, which are reported in Table 2. For uncovering the community-structure, we apply the clustering-algorithm described in “[The clustering algorithm](#)”. This is why the topology differs from the one in Göbel and Araújo (2020), where the network partition was generated by the modularity algorithm proposed by Newman (2006). We call the resulting topology *NetModel-Ref*, which is pictured in Fig. 1. We then validate our proposed indicator-selection procedure by telling the algorithm a priori the desired community-structure to be *NetModel-Ref*. If calibrated correctly, the algorithm should at least reveal the five-indicator model to exactly generate the topology of *NetModel-Ref*.

As it turns out, the algorithm identified the five indicators used in Göbel and Araújo (2020) to generate the partitioning of *NetModel-Ref*. But in addition to these five variables,

a**Period 1**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	
Austria [0]	Greece [B*; c]	Australia [b*]	Argentina [t]	Indonesia [B*]	
Belgium [0]	Portugal [0]	Canada [0]	Brazil [t]	Thailand [0]	
France [0]	United States [b]	Finland [B; c1]	Turkey [B*]		
Germany [0]		Mexico [B]			
Netherlands [0]		Singapore [0]			
Norway [B; c1]		South Korea [0]			
Sweden [B; c1]					
Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11
Colombia [c]	India [0]	Italy [b*; C]	Japan [B]	Spain [C]	United Kingdom [B*]

b**Period 2**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	
Colombia [C]	Austria [0]	Spain [0]	Australia [b*]	Belgium [0]	Greece [b*]	
Indonesia [b*]	Germany [0]	Netherlands [0]	Brazil [B; c]	Sweden [b]	Portugal [0]	
Singapore [0]			Canada [0]			
South Korea [0]			France [B*]			
Thailand [0]			Norway [0]			
			United Kingdom [B*]			
			United States [0]			
Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12	Cluster 13
Argentina [B; c]	Finland [b]	India [B]	Italy [b*]	Japan [b]	Mexico [T]	Turkey [C]

c**Period 3**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Belgium [0]	Austria [0]	Brazil [b]	Indonesia [T]	Australia [C]	Canada [0]
Spain [0]	Germany [0]	Norway [0]	Thailand [T]	Colombia [T]	Turkey [0]
Finland [0]		Singapore [C]		South Korea [T]	United Kingdom [b*]
Greece [0]		Sweden [0]			
Netherlands [0]		United States [0]			
Portugal [0]					
Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
Argentina [b]	France [0]	India [b*]	Italy [0]	Japan [b]	Mexico [b*]

d**Period 4**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Belgium [B]	Spain [B]	Brazil [C]	Australia [C]	Colombia [0]	Austria [B]
Finland [0]	Greece [B]	Indonesia [C]	Canada [0]	France [B]	Italy [B]
Germany [B]	Portugal [B]	Mexico [C]	Norway [C]	United Kingdom [B; c1]	
Netherlands [B]	United States [B]	Thailand [0]	Singapore [0]		
			Sweden [T]		
Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	
Argentina [0]	India [C]	Japan [0]	South Korea [0]	Turkey [C]	

Fig. 1 Targeted Cluster-Composition: NetModel-Ref

the algorithm also revealed that *Exports* were redundant in explaining the network partition in Periods 2 and 3. Hence, during the mid 1990s the level of *Exports* did not add any additional value to the information already embedded in the other four indicators.

This exercise was intended to show that given a certain network structure – which was built on a known set of indicators – our procedure can successfully detect this known set of indicators and in addition even select those indicators, which carry redundant information. Having proven the functionality of the procedure, we now move on to identify that combination of indicators which would have foreseen a desired network structure most accurately. The intention is to specify a certain topology, which is characterized by homogeneous crisis and non-crisis clusters. We then want to quantitatively evaluate the degree to which any combination of indicators could match our desired community-structure.

Targeted network topologies

We now turn to the actual indicator selection procedure. To do so, we extend the number of macroeconomic variables from five to ten, as reported in Table 1. Furthermore, as it is highly unlikely that any combination of these ten indicators can *exactly* match our desired network structures, we still want to know how close each combination gets to an exact matching. Therefore, we introduce the *level of satisfaction*, $S \in [0, 1]$, as a measure for how accurately any combination of these ten macroeconomic variables would have matched our *targeted* network structure. Hence, we define the *level of satisfaction*, S , as follows:

for any period P , with $P = 1, 2, 3, 4$,

- I. Take a cluster of the *pre-reference-date* period.
- II. Then take a cluster of the *emph-targeted* network and determine the number of countries of the *pre-reference-date* cluster, which are represented in the former.
- III. Divide this number by the larger of the two chosen clusters (*pre-reference-date* cluster and the *targeted* network cluster).
- IV. Apply steps (II) and (III) - for the taken *pre-reference-date* cluster in (I) - to each of the communities in *emph-targeted* network.
- V. The average of all non-zero values of these *degrees of satisfaction* gives the *cluster-specific degree of satisfaction*.
- VI. Return to step (I) and take another cluster of the *pre-reference-date* era and run through steps (II)-(VI) until all clusters of the *pre-reference-date* period have been investigated.
- VII. Take the average of all the *cluster-specific degrees of satisfaction* of Period P and get the *Average Level of Satisfaction*, S .

This procedure allows us to investigate the most promising indicators and even tells us if the most successful combination occurs over and over again or if it is rather an artifact of its time. Here, the economic literature is still highly undecided. Reinhart & Rogoff (2008) detected common patterns in certain macroeconomic variables during several pre-crisis episodes, whereas Fioramanti (2008) emphasizes the time-varying nature of sovereign debt crises. Hence, our procedure allows us to add another piece to this ongoing discussion.

In Figs. 2 and 3 we show two examples for what we would consider a *reasonable* partitioning of crisis and non-crisis clusters. In the remainder we will refer to these topologies

a**Period 2**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Austria [0]	Greece [B*; c]	Australia [b*]	Argentina [t]	Colombia [c]	Spain [C]	Japan [B]
Belgium [0]	Turkey [B*]	United States [b]	Brazil [t]		Italy [b*; C]	Mexico [B]
France [0]	Indonesia [B*]					Finland [B; c1]
Germany [0]	United Kingdom [B*]					Norway [B; c1]
Netherlands [0]						Sweden [B; c1]
Portugal [0]						
Canada [0]						
Thailand [0]						
Singapore [0]						
South Korea [0]						
India [0]						

b**Period 3**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Austria [0]	Australia [b*]	United Kingdom [B*]	Sweden [b]	Mexico [T]	Argentina [B; c]	Turkey [C]
Germany [0]	Indonesia [b*]	France [B*]	Finland [b]		Brazil [B; c]	Colombia [C]
Spain [0]	Greece [b*]		Japan [b]		India [B]	
Netherlands [0]	Italy [b*]					
Norway [0]						
United States [0]						
Canada [0]						
Belgium [0]						
Portugal [0]						
Singapore [0]						
South Korea [0]						
Thailand [0]						

c**Period 4**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Belgium [0]	Singapore [C]	Brazil [b]	Indonesia [T]	Mexico [b*]
Spain [0]	Australia [C]	Argentina [b]	Thailand [T]	India [b*]
Finland [0]		Japan [b]	Colombia [T]	United Kingdom [b*]
Greece [0]			South Korea [T]	
Netherlands [0]				
Portugal [0]				
Austria [0]				
Germany [0]				
Norway [0]				
Sweden [0]				
United States [0]				
Canada [0]				
Turkey [0]				
France [0]				
Italy [0]				

d**Period 5**

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Belgium [B]	Finland [0]	Brazil [C]	Sweden [T]
France [B]	Thailand [0]	Indonesia [C]	
Germany [B]	Singapore [0]	Mexico [C]	
Netherlands [B]	Canada [0]	Australia [C]	
Spain [B]	Colombia [0]	Norway [C]	
Greece [B]	Argentina [0]	India [C]	
Portugal [B]	Japan [0]	Turkey [C]	
United States [B]	South Korea [0]		
Austria [B]			
Italy [B]			
United Kingdom [B; c1]			

Fig. 2 Targeted Cluster-Composition: NetModel1

a**Period 2**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Austria [0]	Greece [B*; c]	Australia [b*]	Argentina [t]	Colombia [c]	Spain [C]
Belgium [0]	Turkey [B*]	United States [b]	Brazil [t]		Italy [b*; C]
France [0]	Indonesia [B*]				
Germany [0]	United Kingdom [B*]				
Netherlands [0]	Japan [B]				
Portugal [0]	Mexico [B]				
Canada [0]	Finland [B; c1]				
Thailand [0]	Norway [B; c1]				
Singapore [0]	Sweden [B; c1]				
South Korea [0]					
India [0]					

b**Period 3**

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Austria [0]	Australia [b*]	United Kingdom [B*]	Mexico [T]	Turkey [C]
Germany [0]	Indonesia [b*]	France [B*]		Colombia [C]
Spain [0]	Greece [b*]	Argentina [B; c]		
Netherlands [0]	Italy [b*]	Brazil [B; c]		
Norway [0]	Sweden [b]	India [B]		
United States [0]	Finland [b]			
Canada [0]	Japan [b]			
Belgium [0]				
Portugal [0]				
Singapore [0]				
South Korea [0]				
Thailand [0]				

c**Period 4**

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Belgium [0]	Singapore [C]	Brazil [b]	Indonesia [T]
Spain [0]	Australia [C]	Argentina [b]	Thailand [T]
Finland [0]		Japan [b]	Colombia [T]
Greece [0]		Mexico [b*]	South Korea [T]
Netherlands [0]		India [b*]	
Portugal [0]		United Kingdom [b*]	
Austria [0]			
Germany [0]			
Norway [0]			
Sweden [0]			
United States [0]			
Canada [0]			
Turkey [0]			
France [0]			
Italy [0]			

d**Period 5**

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Belgium [B]	Finland [0]	Brazil [C]	Sweden [T]
France [B]	Thailand [0]	Indonesia [C]	
Germany [B]	Singapore [0]	Mexico [C]	
Netherlands [B]	Canada [0]	Australia [C]	
Spain [B]	Colombia [0]	Norway [C]	
Greece [B]	Argentina [0]	India [C]	
Portugal [B]	Japan [0]	Turkey [C]	
United States [B]	South Korea [0]		
Austria [B]			
Italy [B]			
United Kingdom [B; c1]			

Fig. 3 Targeted Cluster-Composition: NetModel2

as NetModel1 and NetModel2. In Fig. 2 we group countries according to a similar crisis-flag. However, we do not only discriminate between banking-, currency-, twin- and non-crisis communities, but we further differentiate the crisis-countries: a lower-case letter (b ; c ; t) indicates that shortly prior to or right at the onset of the pre-crisis period – over which we calculate the cross-country similarities – the country has experienced a crisis-event and is in the *aftermath* of a crisis. A capital-case letter (B ; C ; T) indicates that the country is hit by a crisis-event at the *reference-date* and that its pre-crisis period is not distorted by any aftermath-dynamics. Finally, a lower-case letter with an additional 1 ($b1$; $c1$) tells us that a crisis-event is looming shortly after the *reference-date*. Furthermore, we also distinguish between by *systemic* and *non-systemic* banking-crises as indicated by an asterisk (b^* ; B^*). In Fig. 3 we ignore the latter differentiation and only group according to O , B , C , T first, and b , c thereafter. As mentioned previously, we tried to generate *pure* and *distortion-free* pre-crisis periods as far as possible.

Letting the data decide

Before turning to the final results, a small recap of the procedure might facilitate the understanding of the remainder: we have a balanced data set of 27 countries and 10 macroeconomic variables with quarterly observations between 1990 and 2018. For this period, we have flagged each country in every quarter with either a crisis-event or a non-crisis event. This resulted in four distinct *periods* P , i.e. times during which we saw an increased number of crises to occur. Following the assumption of similar macroeconomic dynamics determining a common subsequent economic state, we conjectured that countries with a similar macroeconomic profile during several quarters prior to a specific *reference date*, t , shall also show a common crisis- or non-crisis state precisely at that *reference date*. Picturing this scenario in a network, we should see that cross-country distances, measured on the basis of macroeconomic variables over several quarters prior to certain date t , can generate clusters of countries, which are homogeneously composed of future crisis countries, respectively non-crisis countries.

Telling a self-calibrated clustering algorithm our *desired* clustering structure for each of the four periods (see Figs. 2 and 3), we now want to know, which combination of the 10 macroeconomic variables, can most accurately generate our *targeted* topology.

The most successful combinations for both NetModel1 and NetModel2 are reported in Table 4. As the allocation of *systemic* and *non-systemic* crisis-events is the only difference between NetModel1 and NetModel2 – NetModel2 we grouped them into the same community, whereas NetModel1 further differentiated between these two types of crises – the differences in the model set-up are not overly pronounced. This is reflected in the almost identical composition of the most successful indicator-bundle across models as well as by the very similar *level of satisfaction*, S . The slightly higher value, reported for NetModel1, suggests that a difference between *systemic* (B^* ; b^*) and *non-systemic* (B ; b ; $b1$) banking-crises is nevertheless inherent in the underlying dynamics of the macroeconomic indicators and a distinction between the two is not unreasonable.

However, no combination of our ten indicators could match the desired outcomes of NetModel1 or NetModel2, pictured in Figs. 2 and 3, with more than $S = 43\%$ of accuracy. This *level of satisfaction*, i.e. the overlap of the *data-generated* and the *desired* network structure, seemed to be quite constant for Periods 1 through 3 – for the crises during the 1990s – but sharply deteriorates in Period 4. This suggests cross-country similarities

Table 4 Most Promising Indicator Combinations *Targeted Community-Structure*: NetModel1 & NetModel2

Crisis	NetModel1		NetModel2	
	Indicators	Satisfaction <i>S</i> (%)	Indicators	Satisfaction <i>S</i> (%)
Period 1	PubDebt / GDP	0.43	PubDebt / GDP	0.40
	Reserves		Reserves	
	CA / GDP		CA / GDP	
	GFCF / GDP		GFCF / GDP	
Period 2	GFCF / GDP	0.41	GFCF / GDP	0.37
	RealFX		RealFX	
	GFCE / GDP		GFCE / GDP	
Period 3	GFCE / GDP	0.40	GFCE / GDP	0.37
	CA / GDP		CA / GDP	
	GFCF / GDP		GFCF / GDP	
	RealFX		PubDebt / GDP	
Period 4	EX	0.25	EX	0.25
	Reserves		Reserves	
	CA / GDP		CA / GDP	
	GFCF / GDP		GFCF / GDP	
	STDebt / Reserves		STDebt / Reserves	
	RealFX		RealFX	

in the eight quarters preceding the events of the Great Recession not to be very insightful for inferring the clear-cut distinction between crisis and non-crisis countries that we propose in Period 4 of NetModel1 and NetModel2². Regarding the choice of indicators, the composition is also not robust over time. Even for Periods 1 through 3, which are only separated by a few years, the composition does vary, despite some apparent commonalities.

From a historical perspective, the first period covers the events at the very beginning of the 1990s when the Scandinavian countries suffered from a banking-crisis and the Italian Lira was the target of speculators. Period 3, on the contrary, represents the time of the Asian crises, when several Asian economies saw their banks in trouble and their currencies subject to speculative attacks subsequently. For both these periods, NetModel2 reveals three common indicators – *PubDebt / GDP*, *GFCF / GDP* and *CA / GDP* – which were part of the bundle of variables, matching the imposed partitioning most successfully. This implies that these three variables seemed to play a decisive role in distinguishing crisis from non-crisis countries during the 1990s. The inspection of NetModel1 shows a very similar picture, as already mentioned. However, *RealFX* seem to play a more important role, whereas *PubDebt / GDP* loses its influence. *RealFX* seems to be important in not only distinguishing between crisis- and non-crisis countries, but further separating the *systemic* banking-crises from the *non-systemic* counterparts.

Proceeding to Period 4 – the Great Recession era – we see that *PubDebt / GDP* falls apart and *STDebt / Reserves* and *EX* seemed to enter for the first time. Interestingly, Berg and Pattillo (1999b) emphasize the *STDebt / Reserves* to be a major indicator for the Asian crises (Period 3). A finding, which our analysis cannot confirm. The only recurring variable, which is part of the most promising bundle of indicators in every period, being

²As *non-systemic* banking-crises do not occur in Period 4, the community-structure is identical in NetModel1 and NetModel2.

capable of telling crisis countries apart from non-crisis countries, is the amount of investment relative to GDP ($GFCF / GDP$). While this measure only provides information about the level of investment, it does not tell anything about the quality and productivity of the investment, which may not be overlooked in an appropriate cross-country evaluation.

So far, the analysis has shown that there are indeed differences in macroeconomic dynamics between run-up periods to *systemic* banking-crises on the one hand and *non-systemic* banking-crises on the other hand. In particular, Period 3 shows that the real exchange rate (*RealFX*) helps with distinguishing between *systemic* and *non-systemic* banking-crises, whereas not accounting for the different nature of banking-crises, as presented in Fig. 3, calls for a closer look at the ratio of public debt to GDP ($PubDebt / GDP$). In both scenarios, $GFCF / GDP$ and CA / GDP seem to be important for telling macroeconomic crisis-dynamics apart from macroeconomic non-crisis dynamics. Hence, we will finally have a look at how well the 3-indicator bundle of $GFCF / GDP$, CA / GDP and *RealFX* would have fared in each of the Periods 1 through 4 for NetModel1 and also show the performance of $GFCF / GDP$, CA / GDP and $PubDebt / GDP$ for NetModel2 respectively.

The comparison of Tables 4 and 5 reveals a mixed picture. Whereas for NetModel1 the *level of satisfaction*, S , is remarkably constant over the first three periods and achieves levels of S which are close to those reported in Table 4, the *level of satisfaction*, S , is significantly reduced in NetModel2 in each of the periods. Even if a *level of satisfaction*, $S \sim 0.37$ in NetModel1 is not enormously overwhelming, its persistence over the course of the 1990s is rather unexpected given earlier results and studies (e.g. Göbel and Araújo (2020)).

Whereas we would tend to reject the hypothesis of time-invariant patterns, respectively macro-economic dynamics, during run-up periods to economic crises for NetModel2, the results for NetModel1 suggest otherwise. However, the level of matching between the data-generated clustering and the desired network topology remains rather low.

Nevertheless, Tables 4 and 5 allow us to dig a bit deeper into the dynamics of each period. Starting with the inspection of NetModel1, we can elicit the marginal contribution of $GFCF / GDP$ in Period 3. The numbers reveal only a tiny gain in the level of S following

Table 5 Comparing Recurring Indicator-Bundles over Time *Targeted Community-Structure*: NetModel1 & NetModel2

Crisis	NetModel1		NetModel2	
	Indicators	Satisfaction S (%)	Indicators	Satisfaction S (%)
Period 1	GFCF / GDP	0.36	GFCF / GDP	0.24
	CA / GDP		CA / GDP	
	RealFX		PubDebt / GDP	
Period 2	GFCF / GDP	0.36	GFCF / GDP	0.23
	CA / GDP		CA / GDP	
	RealFX		PubDebt / GDP	
Period 3	GFCF / GDP	0.37	GFCF / GDP	0.28
	CA / GDP		CA / GDP	
	RealFX		PubDebt / GDP	
Period 4	GFCF / GDP	0.17	GFCF / GDP	0.14
	CA / GDP		CA / GDP	
	RealFX		PubDebt / GDP	

the inclusion of $GFCE / GDP$. The bulk of the *level of satisfaction* is achieved by the three indicators listed in column 2 of Table 5. In Period 4, a doubling of the number of variables only increases the accuracy by 50%. While the matching accuracy remains low, the three indicators $GFCF / GDP$, CA / GDP and $RealFX$, seem to achieve similar scores as the period-specific optimal indicator combinations do.

Switching to NetModel2, we see a significantly reduced *level of satisfaction*, S , relative to NetModel1. The most pronounced decrease is reported in Period 1. This is insofar interesting as it reveals the marginal contribution of *Reserves*. This finding suggests that at the beginning of the 1990s the level of international reserves was an important indicator for the likelihood of a looming crisis, which is line with the existing literature (Krugman 1979; Berg and Pattillo 1999a; 1999b)³. We can also infer the marginal contribution of a single variable in Period 3. $GFCE / GDP$ is non-negligible in generating the corresponding community-structure of NetModel2, even though its marginal contribution is not as outstanding as that of *Reserves* in Period 1. The value, which $GFCE / GDP$ seems to add to S , is rather proportional to the average contribution of each indicator: the 4-indicator model achieves $S = 0.37$, whereas the three indicators in Table 5 generate $S = 0.28$.

The results of this section have shown that there seems to be a static bundle of indicators, which can match a network topology with homogeneous crisis and non-crisis clusters, almost as closely as the period-specific best-performing indicator combinations do. Despite the rather low matching-accuracy, such a result hints at least at the existence of common latent dynamics underlying different and timely separated crisis-events.

Conclusion

We were interested in investigating whether the tools of network theory can shed further light on the existence of timely recurring determinants of banking- and currency-crises. We assumed that similarities in the macroeconomic profiles of two countries – measured over several quarters – are indicative of these countries' economic status at the end of this measurement period. Feeding a self-calibrated clustering algorithm with these cross-country similarities, we were interested in how well the emerging community-structure is composed of homogeneous crisis and non-crisis clusters.

We proposed two exogenously defined clustering structures, which we regarded as a successful separation of crisis from non-crisis countries. We then let the data decide by itself, which combination of ten macroeconomic variables could best generate one of our desired topologies. Even if the overall matching was rather low, the procedure revealed the composition of the most-successful indicator bundle to be varying over time. Extracting a single and static bundle out of these optimal indicator composition and evaluating its performance revealed one combination – $GFCF / GDP$, CA / GDP and $RealFX$ – to report a remarkably constant accuracy of matching with scores close to the best-scoring bundle. This supports Reinhart and Rogoff's (2008) claim that even if each crisis is different, there are underlying latent dynamics, which increase the likelihood of a crisis to occur in the near future.

Interestingly, the ten indicators struggled most with matching the community-structure, which targeted the events prior to the Great Recession. In that period, the most successful combination was very different to those bundles, which achieved the highest

³Discarding *Reserves* in Period 4 of Table 4 revealed the marginal contribution of *Reserves* to be 0.08.

scores in matching the desired network-structure of the pre-crisis times in the 1990s. The results also suggest the indicators to better match a *targeted* network-structure, which distinguishes between *systemic* and *non-systemic* banking-crises than a topology, which neglects such a differentiation. Furthermore, the level of international reserves played an important role in generating the crisis and non-crisis clusters at the beginning of the 1990s and seemed to matter again in the period preceding the Great Recession. *GFCF / GDP*, a proxy for investment, was the only indicator which was part of the best indicator bundle in every period.

In general, network theory seems to offer useful tools for finding answers to economic problem settings, even if the only contribution were to challenge orthodox methodologies to be revised and developed further. Nevertheless, open questions for upcoming research still remain. So far, our proposed methodology is rather useful as an observational tool for describing the underlying data set and understanding the inherent dynamics. Making the methodology applicable to forecasting exercises and assessing the statistical significance of the results would be further steps into the direction of an early-warning model of economic crises.

Abbreviations

B: Banking-crisis at *reference date t* b: Banking-crisis within four quarters prior to the reference date *t* b1: Banking-crisis within four quarters after the reference date *t* C: Currency-crisis at *reference date t* c: Currency-crisis within four quarters prior to the reference date *t* c1: Currency-crisis within four quarters after the reference date *t* CA/GDP: Current-Account-Balance in relation to GDP EX: Exports GDS/GDP: Gross-Domestic-Savings (% of GDP) GFCE/GDP: General-Government-Final-Consumption-Expenditure (in % of GDP) GFCF/GDP: Gross-Fixed-Capital-Formation (in % of GDP) IM: Imports PubDebt/GDP: Public-Debt in relation to GDP RealFX: Real Exchange Rate [*Deviation from Trend*] Reserves: International Reserves STDebt/Reserves: $\frac{\text{Outstanding International Debt Coming Due} < 1 \text{ Year}}{\text{International Reserves}}$ T: Banking- & currency-crisis at *reference date t* t: Banking- & currency-crisis within four quarters prior to the reference date *t*

Acknowledgements

The authors acknowledge financial Support from FCT – Fundação para a Ciência e Tecnologia (Portugal), national funding through research grant UIDB/05069/2020.

Authors' contributions

MG compiled the dataset, worked on data manipulation and computations. TA helped in conducting computations and provided expertise knowledge. Both authors conducted their share of computations which fostered understanding the data and shaping the results of the manuscript. Both authors read and approved the final manuscript.

Funding

Funding was provided by FCT – Fundação para a Ciência e Tecnologia (Portugal), national funding through research grant UIDB/05069/2020.

Availability of data and materials

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 6 March 2020 Accepted: 6 July 2020

Published online: 04 August 2020

References

- Araújo T, Louçã F (2007) The geometry of crashes. a measure of the dynamics of stock market crises. *Quant Finance* 7(1):63–74
- Araújo T, Göbel M (2019) Reframing the s&p 500 network of stocks along the 21st century. *Phys A Stat Mech Appl* 526(121062). <https://doi.org/10.1016/j.physa.2019.121062>
- Athey S, Imbens GW (2019) Machine learning methods that economists should know about. *Ann Rev Econ* 11(1):685–725
- Becker RA, Chambers JM, Wilks AR (1988) The new s language: A programming environment for data analysis and graphics. Chapman & Hall
- Berg A, Pattillo C (1999) Predicting currency crises: The indicators approach and an alternative. *J of Int Money Fin* 18:561–586

- Berg, A, Pattillo C (1999) What caused the asian crises: An early warning aystem approach. *Econ Notes Banca Monte dei Paschi di Siena SpA* 28(3):285–334
- Bostanci G, Yilmaz K (2020) How connected is the global sovereign credit risk network?. *J Bank Financ Forthcom*. <http://dx.doi.org/10.2139/ssrn.2647251>
- Demirer M, Diebold FX, Liu L, Yilmaz K (2017) Estimating global bank network connectedness. *J of Appl Econ* 33:1–15
- Eichengreen B, Rose AK, Wyplosz C (1995) Exchange market mayhem: the antecedents and aftermath of speculative attacks. <http://faculty.haas.berkeley.edu/arose/erw3ep.pdf>. Accessed 16 July 2019
- Fioramanti M (2008) Predicting sovereign debt crises using artificial neural networks: A comparative approach. *J of Fin Stab* 4:149–164
- Forsythe GE, Malcolm MA, Moler CB (1977) Computer methods for mathematical computations
- Fouliard J, Howell M, Rey H (2019) Answering the queen: Machine learning and financial crises. *Forthcoming*
- Fuertes A-M, Kalotychou E (2007) Optimal design of early warning systems for sovereign debt crises. *Int J of Forecast* 23:85–100
- Gan G, Ma C, Wu J (2007) Data Clustering: Theory, Algorithms, and Applications. Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9780898718348>. <https://epubs.siam.org/doi/abs/10.1137/1.9780898718348>
- Göbel M, Araújo T (2020) A network structure analysis of economic crises. In: Cherifi H, Gaito S, Mendes J, Moro E, Rocha L (eds). *Complex Networks and Their Applications VIII*. pp 16–27
- Goulet Coulombe P (2020) The macroeconomy as a random forest. *eprint=2006.12724*
- Kaminsky GL, Board of Governors of the Federal Reserve System, Reinhart CM (1996) The twin crises: The causes of banking and balance-of-payments problems. *Int Fin Disc Paper* 1996(544):1–28. <https://doi.org/10.17016/ifdp.1996.544>
- Kaminsky GL, Lizondo S, Reinhart CM (1998) Leading indicators of currency crises. *IMF Staff Papers* 45(1)
- Kaminsky G (2007) Currency crises: Are they all the same?. *J Int Mon Fin* 25:503–527
- Krugman P (1979) A model of balance of payments crises. *J of Money Cred Bank* 11:311–325
- Laeven L, Valencia F (2018) Systemic banking crises revisited. *IMF Work Pap* 18(206):1. <https://doi.org/10.5089/9781484376379.001>
- MacQueen J (1967) Some methods for classification and analysis of multivariate observations. *Proc Fifth Berkeley Symp Math Stat Prob* 1:281–297
- Manasse P, Roubini N (2009) "rules of thumb" for sovereign debt crises. *J of Int Econ* 78:192–205
- Mantegna RN (1999) Hierarchical structure in financial markets. *Europ Phys J B* 11:193–197
- Marghescu D, Sarlin P, Liu S (2010) Early-warning analysis for currency crises in emerging markets: A revisit with fuzzy clustering. *Intell Sys Acc Fin Mgmt* 17:143–165
- Newman MEJ (2006) Modularity and community structure in networks. *Proc Nat Acad Sci* 103(23):8577–8582
- Obstfeld M (1994) The logic of currency crises. *national bureau of economic research*, Vol. 4648
- Obstfeld, M (1996) Models of currency crises with self-fulfilling features. *Europ Econ Rev* 40:1037–1047
- Padgett JF, Ansell CK (1993) Robust action and the rise of the medici, 1400–1434. *Am J of Sociol* 98:1259–1319
- Piccardi C, Calatroni L, Bertoni F (2011) Clustering financial time series by network community analysis. *Int J of Mod Phys C* 22(1):35–50
- Prekopcsák Z, Lemire D (2012) Time series classification by class-specific mahalanobis distance measures. *Adv Data Anal Classif* 6(3):185–200
- Reinhart CM, Rogoff KS (2008) Is the 2007 u.s. sub-prime financial crisis so different? an international historical comparison, Vol. 14587. *NBER Working Paper*, Cambridge
- Reinhart, CM, Rogoff KS (2009) *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton
- Ristolainen K (2018) Predicting banking crises with artificial neural networks: The role of nonlinearity and heterogeneity. *Scand J Econ* 120(1):31–62
- Sarlin P, Marghescu D (2011) Visual predictions of currency crises using self-organizing maps. *Intell Sys Acc Fin Mgmt* 18(1):15–38
- Sarlin P (2014) On biologically inspired predictions of the global financial crisis. *Neural Comp Appl* 24:663–673
- Spelta A, Araújo T (2012) The topology of cross-border exposures: Beyond the minimal spanning tree approach. *Phys A Stat Mech Appl* 391(22):5572–5583
- Ward F (2017) Spotting the danger zone: Forecasting financial crises with classification tree ensembles and many predictors. *J Appl Econ* 32(2):359–378

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.