Journal of Financial Stability 35 (2018) 226–241

Contents lists available at ScienceDirect

Journal of Financial Stability

journal homepage: www.elsevier.com/locate/jfstabil

Network linkages to predict bank distress☆

Andreea Constantin[a], Tuomas A. Peltonen[b], Peter Sarlin[c,d,*]

[a] Université de Lausanne, Dept. of Finance, Quartier de Dorigny, 1015 Lausanne, Switzerland
[b] European Systemic Risk Board, Sonnemannstrasse 22, 60314 Frankfurt am Main, Germany
[c] Hanken School of Economics, Arkadiankatu 22, 00100 Helsinki, Finland
[d] RiskLab Finland, Arcada University of Applied Sciences, Jan-Magnus Janssons plats 1, 00560 Helsinki, Finland

ARTICLE INFO

Article history:
Received 20 January 2016
Received in revised form 6 July 2016
Accepted 26 October 2016
Available online 22 November 2016

JEL classification:
G21
G33
C54
D85

Keywords:
Bank distress
Bank networks
Systemic risk

ABSTRACT

Building on the literature on systemic risk and financial contagion, the paper introduces estimated network linkages into an early-warning model to predict bank distress among European banks. We use multivariate extreme value theory to estimate equity-based tail-dependence networks, whose links proxy for the markets' view of bank interconnectedness in case of elevated financial stress. The paper finds that early warning models including estimated tail dependencies consistently outperform bank-specific benchmark models without networks. The results are robust to variation in model specification and also hold in relation to simpler benchmarks of contagion. Generally, this paper gives direct support for measures of interconnectedness in early-warning models, and moves toward a unified representation of cyclical and cross-sectional dimensions of systemic risk.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The global financial crisis has stimulated research on deriving tools for monitoring systemic risk and contagion risk. This is usually approached from two perspectives: the structural and cyclical dimensions of systemic risk. While early-warning models tackle the cyclical dimension and build-up of systemic risk, various network approaches address structural or cross-sectional aspects related to an interconnected financial system. This paper contributes to the field by combining a model of bank distress with bank networks of interconnectedness, in order to account for the propensity of distress to spread in early-warning exercises.

We address systemic risk surveillance by introducing bank networks into an early-warning model to predict bank distress. With a two-step estimation, we rely on the assumption that the vulnerability of one bank is also defined by the vulnerability of its neighbors. This paper provides a general-purpose framework that enables combining any type of networks with any type of distress models. While previous literature on bank-level early-warning models have ignored potential network effects by focusing solely on individual bank distress, the key contribution of the paper is that it explicitly combines potential contagion effects through tail dependencies in a bank failure model. The network perspective is modeled with the multivariate extreme value theory approach of Poon et al. (2004) to estimate tail-dependence networks based on equity prices, which proxy markets' view of bank interconnectedness via direct bilateral or common exposures. Despite being estimated networks, they are not necessarily inferior to real

☆ The paper includes a supplementary network visualization: http://vis.risklab.fi/#/tailnet (for a further discussion of the VisRisk platform see Sarlin (2014)). The authors thank two anonymous reviewers, Philip Davis, Marco Geraci, Christophe J. Godlewski, Christophe Hurlin, Dilruba Karim, Jan-Hannes Lang, Michael Rockinger, Bernd Schwaab, Christoph Siebenbrunner and René Stulz for discussions and support. Thanks also to participants at the ECB Financial Stability Conference (2012, Frankfurt), MAFIN (2012, Genoa), CEQURA (2012, Munich), MaRs Conference (2012, Frankfurt), Cleveland Fed/OFR Financial Stability Conference (2013, Washington), INFINITI (2013, Sciences Po), CREDIT (2013, Venice), International Symposium on Forecasting (2014, Rotterdam), ECB Financial Stability seminar (2014, Frankfurt), SYRTO Conference (2015, Amsterdam), INFINITI (2015, Ljubljana), FEBS (2015, Nantes), EconomiX Conference in International Macroeconomics and Financial Econometrics (2015, Paris), AFFI (2015, Paris), SFI Research Days (2015, Gerzensee), and Banco de Mexico, Journal of Financial Stability and University of Zurich Conference on Network Models and Stress Testing (2015, Mexico City). The views in this paper are those of the authors and do not reflect those of the ESRB, the ECB or the Eurosystem.

* Corresponding author at: Hanken School of Economics, Arkadiankatu 22, 00100 Helsinki, Finland.
E-mail addresses: andreea.piloiu@gmail.com (A. Constantin), tuomas.peltonen@ecb.int (T.A. Peltonen), peter@risklab.fi (P. Sarlin).

https://doi.org/10.1016/j.jfs.2016.10.011
1572-3089/© 2016 Elsevier B.V. All rights reserved.

exposure data, as the market's view also accounts for more indirect sources of interdependence, such as common and correlated exposures and behavioral aspects. These networks are combined with the early-warning model of Betz et al. (2014) using bank-specific and country-level indicators to provide information on the potential spread of distress through interconnectedness in the banking system. We apply our approach in a European setting with 171 listed banks over 1999Q1–2012Q3.

The paper is related to several strands of literature. First, the sole assumption of an interconnected financial system relates to the theoretical literature on (indirect) contagion (e.g., Freixas et al., 2000; Cifuentes et al., 2005; Brunnermeier, 2008; Brunnermeier and Pedersen, 2009; Tirole, 2011). More concretely, our approach to estimating tail-dependence networks relates mainly to the literature on multivariate extreme value theory (e.g., Poon et al., 2004), as well as more generally to the literature on financial contagion through extreme value theory (e.g., Bae et al., 2003; Hartmann et al., 2004, 2005; Gropp and Moerman, 2004 April; Longin and Solnik, 2001; Gropp et al., 2009). Beyond this, the literature has obviously also proposed a number of other approaches to estimating tail-dependence networks, such as Diebold and Yılmaz (2014), Hautsch et al. (2014), Hautsch et al. (2014), and Betz et al. (2014). On a more general note, the literature has its basis in network structures and contagion as described in a seminal paper by Allen and Gale (2000), as well as in Battiston et al. (2012), Gai et al. (2011) and Battiston et al. (2012), and also surveyed in Nier et al. (2007) and Allen and Babus (2009). Moreover, at the bank level, a directly related study is Hale et al. (2014), in which they show the impact of crises through direct and indirect exposures on bank profitability.

A more related strand of literature has focused on network effects in early-warning models. While being few in number, previous works have accounted for the interconnectedness in assessing and predicting systemic risks. In particular, Mikhail et al. (2013) used indicators of the cross-sectional dimension of systemic risk through connectivity indicators, such as CoVaR, in order to signal banking crises, Minoiu et al. (2013) assess the link between overall cross-country financial connectedness and vulnerability to banking crises, and Peltonen et al. (2014) analyze the impact of both cross-country and domestic interconnectedness in terms of four different financial instruments as vulnerability to banking crises. Yet, in relation to the present paper, these are all at the country level and compute only overall interconnectedness as a vulnerability rather than allowing for distress pass through in networks. In contrast, this paper builds upon and extends the bank failure model by Betz et al. (2014), by complementing it with the estimated tail-dependence network. Beyond country vulnerability indicators, it also includes a country-specific fixed effect to proxy for cross-country heterogeneity like supervisory standards.

The paper finds that models including estimated tail dependencies consistently outperform the benchmark model, which is based solely on bank-specific and country-specific data and does not account for any type of vulnerability transmission. For country-specific data, the variables cover both sector level and macro-financial variables. Our results are robust to a wide range of variation in model specification, such as different network estimations, policymaker's preferences, forecast horizons and selections of explanatory factors. For assessing the out-of-sample performance of different early-warning models we use signal evaluation concepts for classification problems, which are wide-spread in machine learning and statistics. The methodology can be summarized as follows: we start by splitting the data sample into an in-sample period, used to estimate the early-warning model, and an out-of-sample period, used to make predictions and assess the model's performance. The out-of-sample predictions are made iteratively, one quarter at a time, while the in-sample period increases by one quarter after each iteration. After the full iteration, we compute the performance of the model based on the out-of-sample signals, by comparing the predictions delivered by the model to the historically observed bank distress events.

For comparison purposes, we construct contagion variables that are either based on estimated network linkages or location of banks' incorporation (country-level contagion). The results show that for the in-sample estimations, all country and network contagion coefficients are statistically significant and have the expected sign: the probability of banks being vulnerable increases if the bank is exposed to contagion from already vulnerable neighbors. The network contagion coefficients also have the highest magnitude when compared to the country contagion ones. In out-of-sample evaluations, the results of the network-based contagion outperform those of simpler contagion benchmarks, such as geographically neighboring banks. Even though the magnitude of the improvement in out-of-sample performance for the two models with network contagion variables is modest, it is statistically significant. This improvement comes from better performance both in terms of missing less crises (reduced false-negative rate) and giving fewer false signals (reduced false-positive rate). When the contagion variables are built using the location of banks' incorporation, there is almost no change in the results compared to the benchmark case, where no contagion is assumed.

These results give a direct support for including measures of interconnectedness and proxies for contagion when building early-warning models. From a policy perspective, they emphasize the need for macro-prudential perspective to complement micro-prudential analysis of individual bank's risk drivers to monitor systemic risk and analyze contagion risk. It is not only enough to either identify vulnerabilities due to linkages among entities or individual distress probabilities, but clearly useful to combine this information. In particular, this provides early steps toward a unified representation of cyclical and cross-sectional dimensions of systemic risk.

The remainder of the paper is organized as follows: Section 2 describes the modeling framework and Section 3 the data used in the analysis. Section 4 presents the results and also discusses their robustness, while Section 5 concludes. Appendix A includes summary statistics and additional robustness tests.

2. Modeling framework

This section presents a modeling framework for combining early-warning models with bank networks. We estimate individual probabilities of bank vulnerability and complement them with network linkages that account for possible transmission of vulnerabilities between banks. The rationale behind the simultaneous use of early-warning models and networks is that this allows for capturing the vulnerabilities that descend directly from each entity itself as well as indirectly from other interlinked entities. Although this paper uses market data to estimate how the realization of negative shocks for any bank's returns may depend on the realization of negative shocks of other banks' returns, it is worth noting that this is a general-purpose framework that is independent of the techniques used for deriving the probabilities and the network linkages. In addition to tail-dependence networks, this section presents the approaches used for deriving and evaluating early-warning models, as well as their combination with network linkages.

2.1. Tail-dependence networks

Given that data on interbank lending and exposures in Europe is not publicly available, we use market data to estimate how the realization of negative shocks for any bank i's returns may depend on the realization of negative shocks of other banks' returns. We thus use an extrema dependence measure that is based on the probability of having extreme negative values that occur simultaneously for any two banks' return series. This effectively amounts to augmenting the individual bank-vulnerability probabilities with conditional probabilities that account for possible simultaneous vulnerabilities among banks. Despite being estimated networks, they are not necessarily inferior to real exposure data. Even if the data on real exposures would be available, it would only offer a partial view on contagion channels, while the market's view also accounts for more indirect sources of interdependence, such as common and correlated exposures and behavioral aspects. Moreover, markets are forward-looking, which is important for early-warning models.

The multivariate extreme-value approach used to estimate the dependence structure among banks' returns is based on modeling the joint-tail distribution of pairs of banks' returns, using the methodology introduced by Poon et al. (2004). Multivariate extreme value theory has proven to be an efficient way to study dependence structures that emerge rarely, such as systemic events. In practice, various fields like portfolio management, allocation decisions, risk management, and hedging largely use statistical approaches that account for more complex dependence structures than simple correlations. Poon et al. (2004) show that dependence measures based on extreme value theory lead to better portfolio risk assessment compared to traditional dependence measures.

The use of multivariate extreme value theory to estimate interbank exposures and contagion risk is further justified by the need of capturing bank interdependencies beyond what can be expected in normal times. Moreover, by filtering stock returns of exogenous factors and time-varying volatility, the estimated interdependencies are in excess of what can be explained by economic fundamentals. Hartmann et al. (2005) also identify contagion risk among banks with extreme negative co-movements between individual bank stocks, such that the extreme events correspond to crisis situations and are severe enough to always be of importance for policymakers.

In order to account for possible common factors that drive banks' returns, we start by regressing each bank's demeaned return series, denoted $r_{i,t}$, on its own lag and the European banking sector and country-specific demeaned return indices. The dependence structure between banks will be based on the innovations of banks' returns, beyond what we could expect from aggregate factors and after filtering for possible heteroscedasticity in the error terms.

The following specification allows us to account for external and common factors' effects on banks' i demeaned returns:

$$r_{i,t} = \beta^{i}_{i,t} * r_{i,t-1} + \beta^{C^i}_{i,t} * r_{C^i,t} + \beta^{E}_{i,t} * r_{E,t} + \epsilon^{i}_{t},$$

where $r_{E,t}$ represents the demeaned European banking sector return index, $r_{C^i,t}$ represents country's i demeaned return index and ϵ^i_t is the error term. The time-varying parameters of the model are estimated using the Dynamic Conditional Beta approach proposed by Engle (2012), which uses time-varying variance-covariance matrices in the ordinary least squares estimation of the coefficients.

Next, our interest shifts to the error terms ϵ^i_t, which can hide non-linear dependencies in the cross-section. Given that heteroscedasticity is a source of tail dependence, we apply univariate asymmetric GARCH(1,1) models (introduced by Glosten et al., 1993) to the error terms. The asymptotic dependence measure should only account for return co-crashes that are not triggered by common factors or increased volatility. The asymmetric GARCH model also accounts for the empirical observation that negative returns tend to be followed by larger increases in volatility than equally large positive returns:

$$\begin{aligned}
&\epsilon_{i,t} = \sigma_{i,t} z_{i,t},\\
&\sigma^2_{i,t} = \omega_i + \alpha_i \epsilon^2_{i,t-1} + \beta_i \sigma^2_{i,t-1} + \gamma_i \epsilon^2_{i,t-1} 1_{\epsilon_{i,t} \le 0},\\
&z_t \sim f(z_{i,t}; \nu_i, \lambda_i).
\end{aligned}$$

The innovation process $z_{i,t}$ is assumed to have a univariate skewed t-distribution, with ν_i degrees of freedom and λ_i the asymmetry parameter, which is known to suit well the fat-tailed conditional distribution of stock returns. In order to focus purely on the dependence between the innovation series, it is suitable to transform the data such that all series have a common marginal distribution. This allows us to remove the possible effects of marginal aspects and assure that the differences in joint-tail probabilities are solely due to differences in the dependence structure. Given the extensively researched fat-tailed distribution of asset returns, we transform the innovation series (z_i,z_j) to Fréchet marginals (S,T) as follows:

$$S = \frac{-1}{\log F_{z_i}(z_i)} \quad \text{and} \quad T = \frac{-1}{\log F_{z_j}(z_j)},$$

where F_{z_i} and F_{z_j} are the respective marginal distributions for z_i and z_j.

For determining whether any two banks i and j are linked by some dependence structure, we use the transformed innovations series to compute the extremal dependence measure $\bar{\chi}$ described by Poon et al. (2004). This measure was developed as a complement to the asymptotic dependence measure χ, which indicates the degree of extremal dependence for any two variables S and T and is defined as:

$$\chi = \lim_{s\to\infty} Pr(T > s | S > s) = \lim_{s\to\infty} \frac{Pr(T > s, S > s)}{Pr(S > s)}.$$

with $0 \le \chi \ge 1$. For $\chi > 0$, the two variables S and T are asymptotically dependent, for $\chi = 1$ they are perfectly dependent and for $\chi = 0$ they are asymptotically independent. The complementary measure $\bar{\chi}$ [1] indicates the rate at which $Pr(T > s|S > s)$ approaches zero. Coles et al. (1999) defined the measure $\bar{\chi}$ as:

$$\bar{\chi} = \lim_{s\to\infty} \frac{2 \log Pr(S > s)}{\log Pr(S > s, T > s)} - 1,$$

where $-1 < \bar{\chi} \le 1$. For perfect dependence between two variables S and T, $\bar{\chi}$=1 and for perfect independence $\bar{\chi}$=0.

Therefore, in order to conclude on the dependence structure between two innovation series, we need to test if $\bar{\chi} = 1$. We will interpret that there is a link between two banks if the measure $\bar{\chi}$ applied to their innovation series in not statistically different from one, meaning that we cannot reject the null hypothesis that the two innovation series may be perfectly dependent. The nonparametric estimation for $\bar{\chi}$ was developed by Ledford and Tawn (1996) and is based on the tail index η, also called the shape parameter, of any heavy-tail variable $Z = \min(S, T)$.[2] It was established that under weak conditions

$$\bar{\chi} = 2\eta - 1.$$

[1] The asymptotic independence measure $\bar{\chi}$ was developed by Ledford and Tawn (1996).

[2] Ledford and Tawn (1996) argue that the bivariate dependence structure is a regular varying function under fairly general conditions

Thus, for estimating the dependence structure between two variables S and T, we need to start by estimating the tail index of the variable $Z=\min(S, T)$. This follows from

$$Pr(Z > z) = Pr(\min(S, T) > z) = Pr(S > z, T > z).$$

The tail index parameter η is estimated using a modified version of Hill's estimator (Hill, 1975) developed by Huisman et al. (2001) for small samples. The modified estimator is based on a weighted average of Hill's estimators for different tail-threshold values. We choose to use the modified tail-index because it provides accurate estimates when the exact tail threshold of the innovations' distribution is not known. For the tail-threshold k of any innovation series of length n, we generally apply an empirically driven rule proposed by Loretan and Phillips (1994) of $k=n^{2/3}/\log(\log(n))$. Moreover, we will compute the variance of the $\hat{\bar{\chi}}$ estimator as:

$$var(\hat{\bar{\chi}}) = \frac{(\hat{\bar{\chi}}+1)^2}{k}.$$

We interpret that there is a link between two banks if the measure $\bar{\chi}$ applied to their innovation series is less than two standard deviations away from one, which also corresponds to $\eta=1$ and the innovation series being asymptotically dependent. To this end, we use the $\hat{\bar{\chi}}$ and $var(\hat{\bar{\chi}})$ estimators to compute the standard z-scores of $\bar{\chi}$, which effectively indicate the distance in units of standard deviation between the raw dependence measure and one. The network used in the early-warning model is thus based on a matrix with binary links that are constructed by comparing the standard z-score of the dependence measure between any two banks to one.

The data used for constructing the European financial network is composed of market data covering daily stock prices for 171 listed European banks, the European banking sector's equity price index and country-level equity price indices. Banks' stock prices were taken from Bloomberg while the equity price indices were taken from Thomson Reuters Datastream; all series cover the period from January 1999 to April 2014. We then extract log-returns from the price series and use their demeaned values to perform the above estimation. The banking network is estimated in an expanding-window fashion for each quarter in the period 2007Q1–2013Q2. The network for the first quarter of 2007 is estimated using the data from January 1999 to December 2006, the network for the second quarter of 2007 is based on the period January 1999 to March 2007, and so on.

Finally, Fig. 1 shows an example of the estimated tail-dependence network for all banks in the sample at 2008Q4, i.e., right after the collapse of the Lehman Brothers and at the start of the Global Financial Crisis. In the figure the nodes represent banks and links are based on tail dependencies. As can be seen from the figure, the estimated tail-dependence networks are dense and one can observe a core-periphery structure commonly found in the literature analyzing interbank contagion. The tail-dependence network can provide the policymaker with additional information about the interconnectedness of a bank either through its direct bilateral exposures or through common or similar exposures. Thus, it

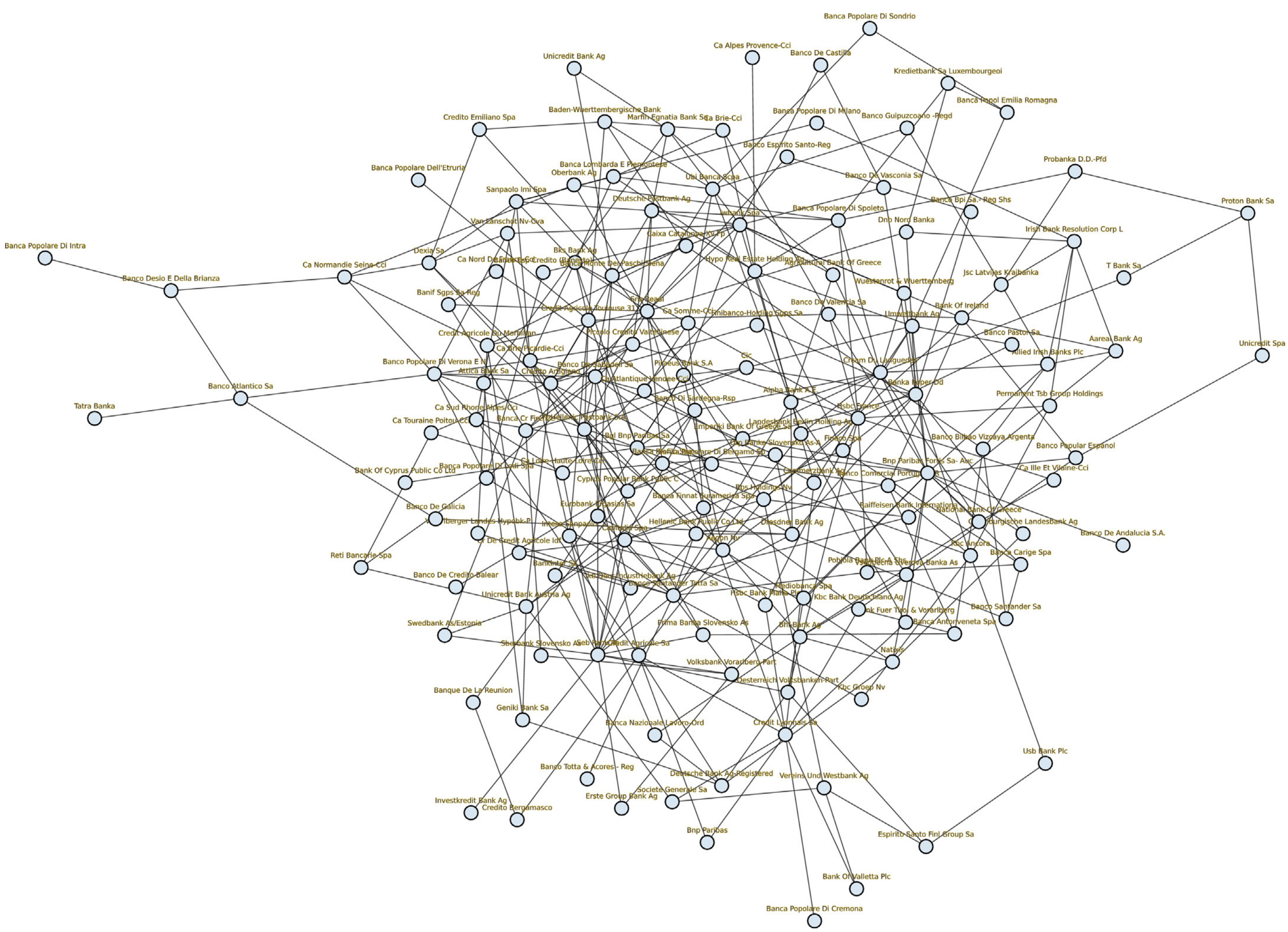

Fig. 1. Estimated tail-dependence network for banks in the estimation sample at 2008Q4, 171 banks.

Table 1
Contingency matrix.

		Actual class C_j	
		1	0
Predicted class P_j	1	*True positive (TP)*	*False positive (FP)*
	0	*False negative (FN)*	*True negative (TN)*

The table shows the relationship between model prediction and actual outcomes. Observations are classified into those where the indicator issues a warning that is indeed followed by a banking crises 12–7 quarters ahead (TP), those where the indicator issues a warning that is not followed by a crisis (FP), those where the indicator issues no warning and there is no crises 7–12 quarters ahead (TN), and those where the indicator issues no warning although there is a crisis coming (FN).

can provide information of potentially vulnerable banks following a bank's failure.

2.2. Signal evaluation framework

Prior to linking the above described networks to early-warning models, we need to understand the objective function of these types of models. The problem of signaling vulnerable banks can be summarized to classifying the banks into two categories. This paper uses a loss function that accounts for type I and type II errors that arise in standard classification problems, and a policymaker's preferences for the errors. The loss function and so-called Usefulness measures make use of the specifications described in Sarlin (2013). In this framework, a model issues a warning signal whenever its estimated probability for a specific bank in a certain period exceeds a threshold $\lambda \in [0, 1]$. In this way, model predictions for each observation j are transformed into binary predictions P_j that are equal to 1 if the respective thresholds are exceeded for this observation and 0 otherwise. Predictive abilities of the variables and the models can then be evaluated by comparing the signals issued by the model to the actual outcome C_j for each observation.[3] Each observation is assigned to the contingency matrix depicted in Table 1.

In order to obtain the optimal threshold λ one needs to define a policymaker's preferences between type I errors (missing a crisis, $T_1(\lambda) = FN/(TP + FN) \in [0, 1]$) and type II errors (issuing a false alarm, $T_2(\lambda) = FP/(FP + TN) \in [0, 1]$). This is done with a loss function that depends on the two types of errors as well as the policymaker's relative preference for either type. The optimal threshold is then chosen by minimizing the loss function. Taking into account the relative frequencies of crises $P_1 = P(C_j = 1)$ and tranquil periods $P_2 = P(C_j = 0)$, the loss function is defined as follows[4]:

$$L(\mu, \lambda) = \mu P_1 T_1(\lambda) + (1 - \mu) P_2 T_2(\lambda),$$

where $\mu \in [0, 1]$ denotes the policymakers' relative preference between type I and type II errors. A μ larger than 0.5 indicates that the policymaker is more averse against missing a crisis than against issuing a false alarm, which is a realistic assumption if one assumes that a signal leads to an internal investigation. With the loss function $L(\mu, \lambda)$, we can assess the usefulness of a model in two ways. First, following Sarlin (2013), the absolute usefulness is also defined to account for unconditional probabilities as follows:

$$U_a = \min(\mu P_1, (1 - \mu) P_2) - L(\mu, \lambda).$$

Note that U_a computes the extent to which having the model is better than having no model. This is because a policymaker can always achieve a loss of $\min(\mu P_1, (1 - \mu)P_2)$ by either always issuing a signal (i.e., $T_1(\lambda) = 0$) or never issuing a signal (i.e., $T_2(\lambda) = 0$). The fact that P_1 is significantly smaller than P_2 in our sample (i.e., few vulnerable states) implies that a policymaker needs to be more concerned about vulnerable states than the avoidance of false alarms. Otherwise, with a suboptimal performing model, it would easily pay off for the policymaker to never issue a signal.

A second measure, the relative usefulness U_r, is computed as follows (see Sarlin, 2013):

$$U_r = \frac{U_a}{\min(\mu P_1, (1 - \mu) P_2)}.$$

The relative usefulness U_r reports U_a as a percentage of the usefulness that a policymaker would gain from a perfectly performing model.[5] The relative usefulness is our preferred performance indicator as it allows the comparison of models for policymakers with different values for the preference parameter μ.

In addition to assessing the relative and absolute usefulness of a model, we also employ receiver operating characteristics (ROC) curves and the area under the ROC curve (AUC). These are also viable measures for comparing performance of early-warning models, particularly when the policymakers' preferences for type I and II errors are not known. While the Usefulness measure shows performance at one specific point of the ROC curve, the AUC measures the performance of models for all preferences $\mu \in [0, 1]$. More specifically, the ROC curve shows the trade-off between the benefits and costs of a certain threshold λ. When two models are compared, the better model has a higher benefit (TP rate (*TPR*) on the vertical axis) at the same cost (FP rate (*FPR*) on the horizontal axis).[6] Thus, as each FP rate is associated with a threshold, the measure shows performance over all thresholds, and hence all preference parameters μ. In this paper, the size of the AUC is computed using trapezoidal approximations. The AUC measures the probability that a randomly chosen vulnerable state is ranked higher than a tranquil period. A perfect ranking has an AUROC equal to 1, whereas a coin toss has an expected AUC of 0.5.

2.3. Logit analysis and recursive modeling

The next task is to connect the above described networks and loss functions to bank distress through approaches for deriving early-warning models. In order to predict bank distress events and to assess the predictive abilities of tail dependence and other contagion measures, we estimate a logit model using pooled data of the following form:

$$Pr(y_{it} = 1) = \frac{e^{X'_{it}\beta}}{1 + e^{X'_{it}\beta}}.$$

where $Pr(y_{it} = 1)$ denotes the probability that bank i is in a vulnerable state (i.e., preceding distress events in quarter t). Thus, rather than using lagged explanatory variables, the dependent variable (denoting the vulnerable state) is defined as 1 eight quarters prior to distress events and 0 otherwise. Following the estimation strategy in Betz et al. (2014), the early-warning model augmented with

[3] C_j is equal to 1 if a bank failure occurs up to eight quarters ahead of the respective period and 0 otherwise.

[4] As pointed out by Sarlin (2013), policymakers should be concerned about the absolute number of misclassification rather than the share of misclassifications in relation to class size (i.e., unweighted type I and type II errors). Therefore, a failure to account for the relative frequency of crisis episodes and tranquil periods results in a bias on the weighting of type I and type II errors in the loss function.

[5] A perfectly performing indicator would achieve $T_1 = T_2 = 0$, implying $L = 0$. Consequently, U_a would reduce to $\min(\mu P_1, (1 - \mu)P_2)$.

[6] The TPR (also called sensitivity) gives the ratio of periods where the model correctly issues a warning to all periods where a warning should have been issued, formally $TPR = TP/(TP + FN)$. The FPR (also called specificity) gives the ratio of periods where the model wrongly issues a signal to all periods where no signal should have been issued, formally $FPR = FP/(FP + TN)$. An ideal model would achieve a TPR of one (no missed crises) and a FPR of zero (no false alarms).

country fixed effects is a recursive logit model that makes a prediction at each quarter $t = 1, 2, \ldots, T$ with an estimation sample that grows in an expanding-window fashion. As independent variables, the vector X_{it} includes country fixed effects and measures of tail dependence and other contagion proxies beyond bank-specific, banking sector and macro-financial variables (see Sections 2.3 and 3 for a precise definition of the variables).

The iterative estimations of the early-warning model imply that bank vulnerabilities are predicted by re-estimating the model at each quarter. This way, all available information is used before making any prediction. The data sample is split into an in-sample period, used to estimate the early-warning model, and into an out-of-sample period, used to make predictions and assess the model's performance. The in-sample period initially covers 1999Q1–2006Q4 and increases by one quarter at each iteration. After each iteration we extract bank vulnerability predictions that will amount in the end to the out-of-sample period 2007Q1–2012Q3.

In order to test the usefulness of the estimated tail dependence to predict bank failures and to introduce contagion-related information, the early-warning model is based on two successive estimations that use the in-sample data for each iteration. The first estimation uses the benchmark model specification to obtain vulnerability signals for each individual bank, which are further used to construct contagion variables. These contagion variables are either based on estimated network linkages or location of banks' incorporation. For the second estimation we add the contagion variables to the benchmark model and re-estimate the model to obtain final vulnerability probabilities, set an optimal threshold, and make the final out-of-sample predictions. In this way, we account for the possible transmission of vulnerabilities between banks that either are linked in the network or because they are incorporated in the same country.

We construct the following variables to introduce network effects and contagion-related information to the early-warning model[7]:

- "Network Dummy": a dummy variable that indicates for each bank whether there are any vulnerable banks to which it is estimated to be connected ("neighbors") through tail dependence.
- "Network Sum": a variable that counts how many vulnerable neighboring banks the bank has in its estimated tail-dependence network.
- "Country Dummy": a dummy variable that indicates for each bank whether there are other banks being signaled as vulnerable in the same country.
- "Country Share": the share of vulnerable banks of total banks in the respective country.

The algorithm is described below. Starting from quarter 2007Q1, we conduct the following iterative exercise for each quarter q in the out-of sample window:

1. Estimate the benchmark early-warning model on the in-sample period, using all available information up to quarter q:

   $$p_i = Pr(y_{it} = 1) = \Lambda(\beta X_{it}),$$

   where $\Lambda(\beta X_{it})$ is the benchmark logit model that uses only bank-specific and country-level indicators X_{it}, p_i represent distress probabilities for bank i, and y_{it} is the distress signal.

2. Collect the distress probabilities $\hat{p}_i$ for the in-sample period, compute the Usefulness measure for all probability thresholds λ and choose the threshold that maximizes in-sample Usefulness. The initial distress signals are given by:

   $$y_{it} = \begin{cases} 1 & \text{if } \hat{p}_i > \lambda, \\ 0 & \text{otherwise.} \end{cases}$$

3. Collect signals from the previous estimation and use a network to identify the neighbors of vulnerable banks. Create contagion variables that indicate for each bank whether it has any or the number of vulnerable neighbors in a specific quarter. Introduce a contagion variable in the benchmark model and re-estimate it using the same in-sample period:

   $$p_i^* = Pr(y_{it} = 1) = \Lambda(\beta X_{it} + \gamma NC_{it}),$$

   where $\Lambda(\beta X_{it} + \gamma NC_{it})$ is the logit model augmented with the contagion variable NC_{it} and p_i^* are the updated distress probabilities.
4. Collect the new distress probabilities $\hat{p}_i^*$ from the model and choose the new optimal threshold λ^* with respect to Usefulness.
5. With the model from Step 4, estimate distress probabilities for quarter q and use the threshold λ^* to signal vulnerable banks. The final distress signals are given by:

   $$y_{it}^* = \begin{cases} 1 & \text{if } \hat{p}_i^* > \lambda^*, \\ 0 & \text{otherwise.} \end{cases}$$

After the full iteration, we compute the performance of the model based on the out-of-sample signals by comparing the predictions delivered by the model to the historically observed bank distress events. The performance is assessed by the relative usefulness U_r measure as well as the receiver operating characteristic (ROC) and the area under the curve (AUC) as explained in Section 2.2.

3. Distress events and indicators

This section describes the data sources and variable definitions and is divided into two sub-sections. We first describe the definition of bank distress events, and then cover the used bank-specific risk drivers and country-level vulnerability indicators. The dataset covers 171 listed European banks over the period 1999Q1–2012Q3. To support comparability, we follow the data collection procedure in Betz et al. (2014), but update coverage and limit our focus on only listed banks.

3.1. Bank distress events

The distress events used in this paper descend from the unique database collected in Betz et al. (2014). As European banks have experienced only few direct bank failures, the events also include state interventions and forced mergers to represent bank distress.

The first type of events include bankruptcies, liquidations and defaults, with the aim of capturing direct bank failures. Bankruptcies occur if the net worth of a bank falls below the country-specific guidelines, whereas liquidations occur if a bank is sold according to the guidelines of the liquidator and the shareholders do not receive full payment for their ownership. Defaults occur if a bank has failed to pay interest or principal on at least one financial obligation beyond any grace period specified by the terms or if a bank completes a distressed exchange. The data on bankruptcies and liquidations are collected from Bankscope, and defaults from Moody's

[7] The labels refer to the variable names in the estimation tables.

and Fitch. The distress events are defined to start when a failure is announced and end at the time of the de facto failure.

The second type of events comprise the use of state support to identify banks in distress. A bank is in distress if it receives a capital injection by the state or participates in asset relief programmes (i.e., asset protection or asset guarantees). It is worth to note that this includes only assistance on the asset side, whereas liquidity support or guarantees on banks' liabilities are not used for defining distressed banks. The state interventions are sourced from the European Commission, as well as data collected by the authors from market sources (Reuters and Bloomberg) for cross-checking. The start dates of the events refer to the announcement of the state aid and the end date to the execution of the state support programme.

The third type of events are forced mergers, which capture private sector solutions to bank distress. Distressed mergers are defined to occur if (i) a parent receives state aid within 12 months after a merger or (ii) if a merged entity exhibits a negative coverage ratio within 12 months before the merger. The coverage ratio is computed using data from Bloomberg (ratio of capital equity and loan reserves minus non-performing loans to total assets, as in e.g., González-Hermosillo, 1999), whereas data on mergers are obtained from Bankscope. The events identified using these definitions of distressed mergers have also been cross-checked using market sources (Reuters and Bloomberg). The dates for these two types of distress events are defined as follows, respectively: (i) the start date is when the merger occurs and the end date when the parent bank receives state aid, and (ii) the start date is when the coverage ratio falls below 0 (within 12 months before the merger) and the end date when the merger occurs.

Table 2 summarizes the frequency of distress events by type. From the table, we can observe that only 13 of the distress events are direct failures, while there are 130 state interventions and 40 distressed mergers. In total, there are 172 distress events at the bank-quarter level. This figure is smaller than the sum of events across categories as they are not mutually exclusive. Table 3 shows the number of banks and distress events by country. As this paper focuses on vulnerable states, or pre-distress events, it is worth noting that we transform in the benchmark case the distress events into a binary pre-distress variable, which is defined to take the value one during the eight quarters prior to the distress events, and otherwise zero.

3.2. Bank-specific and country-level indicators

We use indicators from three categories to capture various aspects of a bank vulnerability: bank-specific, banking sector and macro-financial indicators. To measure bank-specific vulnerabilities, the first category includes indicators based upon banks' income statements and balance sheets. As is common in the literature (e.g., Flannery, 1998; González-Hermosillo, 1999; Poghosyan and Čihak, 2011), we use indicators covering all dimensions in the CAMELS rating system, which have been constructed using Bloomberg data. In contrast to studies like Agarwal and Taffler (2008), we do not consider market-based indicators because we aim at predicting underlying vulnerabilities two or even three years prior to distress, whereas market-based signals tend to have a shorter forecast horizon (e.g., Bongini et al., 2002; Milne, 2014). In the following list, we describe the indicators and their assumed relation to distress.

- Capital adequacy (C): The equity-to-assets ratio (capital ratio) is used to proxy the level of bank capitalization. Higher level of capital acts as a buffer against financial losses protecting a bank's solvency and is expected to reduce the probability of a bank failure.
- Asset quality (A): This dimension is measured with return on assets (ROA), reserves for non-performing loans as a share of non-performing assets, and the share of loan loss provisions to total average loans. Overall, weaker asset quality is expected to be positively associated with bank distress, whereas exaggerated returns may also proxy 'excessive' risk taking. A large share of provisions for loan losses to total average loans is expected to increase the probability of failure. However, the effect of reserves for loan losses as a share of non-performing assets is potentially ambiguous, as higher reserves should correspond to a higher cover for expected losses, but could also proxy for higher expected losses.
- Management soundness (M): Even though being a challenging dimension to measure, we use the cost-to-income ratio to proxy for the efficiency of firms in minimizing costs while increasing profits, which is expected to reduce the probability of a bank failure.
- Earnings (E): To measure bank's profitability, we use the return on equity (ROE), which is expected to be negatively associated with bank distress.
- Liquidity (L): This dimension is measured with the share of interest expenses to total liabilities, the deposits-to-funding ratio and the ratio of net short-term borrowing to total liabilities. A higher deposits-to-funding ratio is expected to be negatively associated with bank distress, as deposits are usually considered as a more stable funding source than the interbank market or securities funding. Yet, large interest expenses to total liabilities and net short-term borrowing to total liabilities are both expected to be positively related with a bank failure.
- Sensitivity to market risk (S): This dimension is proxied for with the share of trading income. The relation of this variable to bank distress could be positive through a riskier business model (i.e., trading income is a volatile source of earnings), whereas investment securities are more liquid than for instance loans (i.e., allows banks to minimize fire sale losses in case of exogenous shocks).

The next two categories of indicators are at the country level. Banking sector indicators proxy for banking system imbalances. Though they are commonly cited as key early-warning indicators for banking crises (e.g., Demirgüç-Kunt and Detragiache, 1998; Demirgüç-Kunt and Detragiache, 2000; Kaminsky and Reinhart, 1999 June; Borio and Lowe, 2002; Lainà et al., 2015), they can also be assumed to impact the vulnerability of individual banks. We proxy the size and rapid increases in banks' balance sheets with total assets to GDP and growth in non-core liabilities; banking-system leverage with debt-to-equity and loans-to-deposits ratios; securitization with debt securities to liabilities; and property booms with the ratio of mortgages to loans. We construct all indicators with the ECB's statistics on Balance Sheet Items (BSI) of the Monetary, Financial Institutions and Markets (MFI). Finally, the third category of indicators consists of macro-financial measures to identify macro-economic imbalances and control for conjunctural variation in asset prices and business cycles. Macro-economic imbalances are controlled for with selected internal and external indicators from the EU Macroeconomic Imbalance Procedure (MIP), such as private sector credit flow, government debt, and international investment position (European Commission, 2012). In addition, we

Table 2
Number of distress events per category.

Total distress events	172
Defaults, liquidations and bankruptcies	13
State interventions:	130
Asset protection	15
Capital injection	96
Guaranteed loans	23
Distressed-merger:	40
State aid	9
Coverage ratio	34

Table 3
Number of banks and distress events by country.

Country	No. of banks	Direct failures	State interventions			Distressed mergers
			Asset prot.	Capital inj.	Guaranteed loans	
Austria	10	0	0	6	0	0
Belgium	4	0	2	5	1	0
Cyprus	4	1	0	1	4	0
Estonia	2	0	0	0	0	0
Finland	3	0	0	0	0	0
France	27	0	0	6	0	3
Germany	16	0	3	11	1	7
Greece	12	4	0	27	16	22
Ireland	5	6	10	11	0	0
Italy	37	1	0	10	0	0
Latvia	2	0	0	0	0	0
Lithuania	3	0	0	0	0	0
Malta	2	0	0	0	0	0
Netherlands	5	0	0	2	0	0
Portugal	9	0	0	3	0	0
Slovakia	5	0	0	0	0	0
Slovenia	4	1	0	1	1	0
Spain	21	0	0	13	0	8
Total	171	13	15	96	23	40

use asset prices (stock and house prices) and business cycle indicators (real GDP growth and CPI inflation) to capture conjunctural variation. Except house price indicators that are retrieved from the ECB, all macro-financial indicators are sourced from Eurostat and Bloomberg. Table 4 presents a list of the risk drivers used in the benchmark estimation, their definitions and the data sources. Table 8 in Appendix A presents descriptive statistics of the risk drivers used in the subsequent analysis. Statistical tests applied show that the data are non-normally distributed and exhibit most often a positive skew with a leptokurtic distribution.

4. A European bank network with a distress model

For a large sample of European banks, this section presents the estimation results, discusses provides illustrative case studies, as well as discusses the robustness analysis.

4.1. Results

The tail-dependence network, estimated as discussed in Section 2.1, offers information about the extremal-dependence structure of listed European banks. Fig. 2 shows the density of the network over the period 2007Q1–2014Q1, indicating what proportion of all possible connections between the banks are actually realized. The higher the density of the network the more links there are among the nodes (i.e., the banks). This measure brings information about such phenomena as the speed at which information diffuses among the nodes. Given that the estimated connections represent high probabilities of simultaneous extreme events between banks, high-density networks can amplify shocks. The blue line represents the density for an expanded sample of European listed banks (243 banks) while the green line is for the sample of banks used in the early warning model (171 banks). The banks' sample for the early warning model is a subset of the whole sample of listed banks because of the availability of their

Table 4
List of explanatory variables, description and data sources.

Variable type		Name	Source
Bank-specific balance sheet variables	C	Total leverage ratio	Bloomberg
	A	Reserves for NPLs to non-performing assets	Bloomberg
		ROA	Bloomberg
		Loan loss provisions to total loans	Bloomberg
	M	Cost to income	Bloomberg
	E	ROE	Bloomberg
	L	Interest expenses to liabilities	Bloomberg
		Deposits to liabilities	Bloomberg
		Net short-term borrowing to liabilities	Bloomberg
	S	Share of trading income to revenue	Bloomberg
Country-specific banking sector variables		Total assets to GDP	ECB MFI Statistics
		Non-core liability growth	ECB MFI Statistics
		Debt to equity	ECB MFI Statistics
		Loans to deposits	ECB MFI Statistics
		Debt securities to liabilities	ECB MFI Statistics
		Mortgages to loans	ECB MFI Statistics
Country-specific marco-financial variables		Real GDP growth	Eurostat
		Inflation	Eurostat
		Stock price growth	Bloomberg
		House price growth	ECB MFI Statistics
		Long-term government bond yield	Bloomberg
		International investment position to GDP	Eurostat/AlertMechanismReport
		Government debt to GDP	Eurostat/AlertMechanismReport
		Private sector credit flow to GDP	Eurostat/AlertMechanismReport

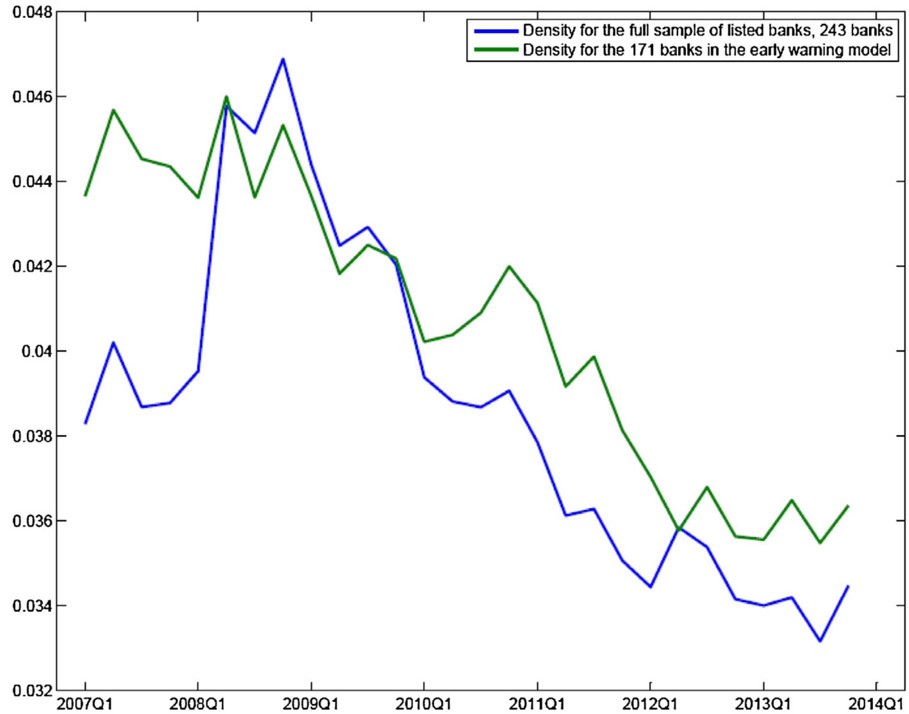

Fig. 2. Network density for the expanded (full) and the estimation sample of banks.

balance-sheet data and their size, as we only use with a minimum of EUR 1bn in total assets during the period under consideration.

It is interesting to observe that, while the density is fairly small at about 4%, it is varying over time. We can observe a strong increase in the number of extremal dependencies during the 2 quarters preceding the failure of Lehman Brothers. The density for the expanded sample of listed banks hits its maximum in 2008Q3. If we reduce the banks' sample in the network to the 171 ones that have the balance sheet data needed for the early-warning model, we can observe in Figure 2 that the links' density was close to its highest levels also during the 6 quarters before 2008Q3. This may be due to the fact that the reduced sample of banks that does not cover smaller banks.

Table 5 presents summary statistics of the nodes' degrees, which represent the number of neighbors each bank has. We can see that the banks have on average 5.8 neighbors among the 170 banks we focus on. However, looking at the 50% quantile we can see that half of the banks have up to 5 neighbors. This indicates that there are few nodes with very high number of neighbors, which is confirmed by a maximum degree of 31 and the 90% quantile at a level of 11. The summary statistics for the eigenvector-centrality measure, which measures the broader influence of a node in the network, are in line with those for the degree and confirm that the network structure is far from being random.

The estimation results of the early-warning models are presented in Table 6. Given that the in-sample estimation of the prediction model is performed each quarter with an expanding time window, we cannot report the coefficient values for each estimation. Table 6 contains indicative results based on the full data sample for the benchmark estimation and the second-step estimations that include contagion variables. Note that the coefficients are only indicative for sign and magnitude, as the real coefficients are re-estimated with every additional quarter data.

In addition to the benchmark model,[8] Table 6 shows four model specifications that account for possible contagion among banks. Two of the specifications include variables to proxy contagion at the country-level ("Country Dummy" and "Country Share"), while two specifications include variables containing information from the estimated tail-dependence network ("Network Dummy" and "Network Sum").[9] Given that the benchmark model is similar to that in Betz et al. (2014), and as the purpose of the paper is to analyze whether accounting for potential contagion could improve the forecast ability of the early-warning model, we focus on presenting the results concerning the country and network contagion variables.

Compared to the benchmark case, where only one estimation of the Betz et al. (2014) early-warning model with country fixed effects is performed for signaling vulnerable banks, the two-step estimation framework shows that contagion effects exist and should be accounted for. All country and network contagion coefficients are statistically significant and have the expected sign: exposure to contagion from vulnerable neighbors increases the vulnerability probabilities of banks. The two network contagion coefficients also have the highest magnitude when compared to the country contagion ones. We also report results for estimated R-squared[10]; the models accounting for the number of vulnerable neighbors ("Network Sum") or the proportion of vulnerable banks

Table 5
Summary statistics of network characteristics for the EWS sample, 171 banks.

	Degree	Eigen-centrality
Mean	5.807	0.060
Std. dev	4.342	0.048
Minimum	0	0.000
Maximum	31	0.321
0.10 quantile	1	0.010
0.50 quantile	5	0.050
0.90 quantile	11	0.118

[8] The benchmark model is based on Betz et al. (2014) augmented with country-fixed effects and an updated dataset.

[9] See Section 2.3 for more information how these variables are constructed.

[10] There are many different ways to calculate R-squared for logistic regressions and no consensus on which is the best. In this paper we use the method of Tjur (2009) for a new coefficient of discrimination for logistic regressions.

Table 6
Early-warning models with bank-specific, banking sector and macro-financial variables, 1999Q1–2012Q3, 171 banks, $\mu = 0.85$.

Variables	Benchmark	Country Dummy	Country Share	Network Dummy	Network Sum
Intercept	−6.0746***	−5.9026***	−5.5761***	−6.109***	−6.6528***
Total leverage ratio	−4.5459***	−4.4711***	−4.4025***	−4.3756***	−3.9523***
Reserves for NPLs to non−performing assets	0.0195	0.0159	0.002	0.0146	−0.0044
ROA	0.6969***	0.6938***	0.4076	0.6571**	0.5369*
Loan loss provisions to total loans	−0.4914	−0.4764	−0.4808	−0.4932	−0.453
Cost to income	−4.0265***	−3.8689***	−3.3875***	−3.8942***	−3.5063***
ROE	−0.0968	−0.0638	−0.6931	−0.1256	−0.2401
Interest expenses to liabilities	0.1089	0.1074	0.1292	0.1032	0.0892
Deposits to liabilities	0.1401	0.1391	0.146	0.1341	0.1034
Net short-term borrowing to liabilities	0.5097***	0.5048***	0.4948***	0.4807***	0.4062***
Share of trading income to revenue	−2.5662***	−2.4883***	−2.2333***	−2.4421***	−2.0866***
Total assets to GDP	13.7318***	12.4496***	9.4916***	13.1504***	10.6329***
Non-core liability growth	0.0147	0.0171	−0.0723	0.0179	−0.0084
Debt to equity	−1.0687***	−1.0565***	−1.0947***	−1.0505***	−0.8629**
Loans to deposits	0.8174*	0.7539	0.838*	0.7909*	0.817*
Debt securities to liabilities	1.0263**	0.8243	0.3785	0.9914*	1.1618**
Mortgages to loans	0.6155	0.5638	0.684	0.611	0.4966
Real GDP growth	0.203*	0.1898	0.1372	0.1842	0.1083
Inflation	−0.0104	−0.0062	−0.0498	−0.0106	−0.0136
Stock price growth	−0.0286	−0.0242	−0.0014	−0.0247	−0.0059
House price growth	0.2263	0.2062	0.0725	0.2132	0.126
Long-term government bond yield	0.5121***	0.4869***	0.2256*	0.4869***	0.3703**
International investment position to GDP	−0.5786	−0.5426	−0.4081	−0.5464	−0.4212
Government debt to GDP	−1.8555***	−1.6642***	−1.8184***	−1.8172***	−1.5327***
Private sector credit flow to GDP	0.3306**	0.3045*	0.1171	0.3055*	0.1984
Country contagion dummy	NA	8.5107***	NA	NA	NA
Country contagion share	NA	NA	5.9262**	NA	NA
Network contagion dummy	NA	NA	NA	9.2621***	NA
Network contagion sum	NA	NA	NA	NA	8.7984***
Observations	3150	3150	3150	3150	3150
R-squared	0.4453	0.4447	0.4662	0.4451	0.4673

The table reports estimation results for the benchmark early-warning model, which does not account for any type of contagion among banks, and four model specifications that introduce contagion based either on estimated network linkages ("Network Dummy" and "Network Sum") or location of banks' incorporation ("Country Dummy" and "Country Sum"). All early-warning models are pooled Logit regressions where the dependent variable is a dummy variable indicating whether a specific bank was in a pre-distress period or not. The benchmark model consists solely of bank-specific risk drivers and country-level vulnerability indicators. The "Dummy" contagion variables indicate for each bank whether there are any vulnerable banks in its vicinity, while the "Sum" contagion variables account for the number of vulnerable banks in each bank's vicinity. Policymakers' preference parameter μ is 0.85 and country fixed-effects are introduced for all five model specifications.
* Significance codes: 0.1.
** Significance codes: 0.05.
*** Significance codes: 0.01.

in a country ("Country Share") have the highest *R*-squared (about 46.7% compared to 44.5% for the benchmark model). However, note that the *R*-squared is not a suitable measure for the goodness-of-fit of the logistic signaling models like the one we discuss here. While the typical R-squared only looks at in-sample fit for individual estimations, the model performance in our case is concerned with the precision of the iterative out-of-sample predictions.[11]

The rest of the explanatory variables used in the benchmark model have coefficients with the expected sign that generally keep their significance levels after adding the contagion variables. Among the CAMELS variables, banks' leverage ratio (C), return on assets (A), cost-to-income ratio (M), the ratio of net short-term borrowing to total liabilities (L), and the share of trading income (S) have highly statistically significant coefficients. The share of trading income has a negative sign, meaning that the liquidity of investment securities reduces bank vulnerability risk by allowing it to minimize fire sale losses in case of a changing macro-financial environment. Among the country-specific indicators, the ratio of the size of the banking system's balance sheet to GDP, its leverage ratio and the securitization ratio are the most important factors, having statistically significant coefficients with high magnitudes.

In order to evaluate the improvement in prediction performance, we consider two benchmark models: the first is identical to the one developed in Betz et al. (2014) and uses only one estimation with no contagion effects ("1est Benchmark") while the second benchmark uses the two step estimation but still with no contagion effects ("2est Benchmark"). For the "2est Benchmark", the variable added to the second part of the estimation simply uses the signals delivered by the early-warning model in the first part, without any type of network-based information. While not being an economically or statistically meaningful estimation strategy, this enables us to assess what would be the improvement in performance from including an additional variable to the model and show that the performance of the network-based models is indeed related to contagion effects. Given that this variable is binary, depending on some external probability threshold, it is unlikely it introduces multicollinearity problems. Indeed, most results show almost no change compared to the main benchmark. The statistical significance of the improvement in prediction performance is assessed by using bootstrapped standard errors. For all the significance tests performed, the results were similar when using "1est Benchmark" or "2est Benchmark" as reference; in the rest of the paper we will use the "1est Benchmark" model when we refer to the benchmark model.

Regarding the out-of-sample performance, Table 7 summarizes the AUC and relative usefulness (U_r) measures across different model specifications. As a reminder, the U_r represents the proportion of usefulness that a policymaker would obtain compared to a perfectly performing model. It looks at the predictions made by the model for a given threshold, and compares the rates of missed signals and false alarms. If we plot the relative usefulness values for all

[11] The out-of-sample performance, over all iterative estimations, is further analyzed with AUC and relative usefulness measures.

Table 7
Out-of-sample performance of the early warning model for different specifications, 2007Q1–2012Q3, $\mu = 0.85$.

Model	AUC	U_r	FN rate	FP rate	TN rate	TP rate
1est Benchmark	0.8941	0.5800	0.1799	0.2095	0.7905	0.8201
2est Benchmark	0.8944	0.5770	0.1799	0.2125	0.7875	0.8201
Country Dummy	0.8933	0.5807	0.1691	0.2214	0.7786	0.8309
Country Share	0.8959	0.5904	0.1799	0.1991	0.8009	0.8201
Network Dummy	0.8992**	0.6060	0.1367	0.2340	0.7660	0.8633
Network Sum	0.8986*	0.6444	0.1655	0.1620	0.8380	0.8345
Network Dummy w/Lag	0.8966	0.6030	0.1367	0.2370	0.7630	0.8633
Network Sum w/Lag	0.8974	0.6478	0.1835	0.1374	0.8626	0.8165

This table summarizes and compares the area under the ROC curve (AUC) and relative usefulness (U_r) goodness of fit measures across different model specifications, for policymakers' preference parameter μ=0.85. Model performance is evaluated out-of-sample, for the period 2007Q1 - 2012Q3. Detailed values for the false negative, false positive rates and their counterparts are also provided. Significance codes for differences of AUC measures between contagion models and the single estimation benchmark model are: 0.01 '***'; 0.05 '**'; 0.1 '*'.

policymaker's preference values, we obtain the ROC curve. In order to quantify the overall performance of a classifier, we compute the area under this curve (AUC), where more than 50% means there is an improvement. Given the distribution of pre-distress events and tranquil periods, we focus on the high-end range of policymaker's preference parameter μ. The relative usefulness measure in Table 7 is computed using a policymaker preference of 0.85,[12] as building an early-warning model with imbalanced data necessitates a policymaker to be more concerned about the rare distress cases, which translates into having a preference to predict distress. We can see that the model with "Network Sum" has the best out-of-sample results with the highest relative usefulness (0.64), while the highest AUC is with the specification "Network Dummy" (0.90). Even though the magnitude of the improvement in AUC for the two models with network contagion variables is modest, it is statistically significant when using bootstrapped standard errors.

In the case where the number of vulnerable neighbors for each bank is introduced in the model, represented by "Network Sum", the relative usefulness measure jumps from 0.58 in the benchmark model to 0.64. This improvement comes from better performance both in terms of missing less crises (reduced false-negative rate) and giving fewer false signals (reduced false-positive rate). When the contagion variables are built using the location of banks' incorporation (i.e., "Country Dummy" and "Country Share"), there is almost no change in the results compared to the benchmark case. In the case of country-based contagion, the improvement in the out-of-sample performance is rather small and not significant. This suggests that intra-country contagion is already accounted for in the early-warning model through country-level indicators and country-fixed effects, as well as highlights the need to account for cross-border contagion.

Finally, we also consider the second benchmark model, where the variable added to the second part of the estimation simply uses the signals delivered by the early-warning model in the first part, without any type of contagion. We can see that simply including an additional variable of model signals to the second step of the estimation does not improve model performance. The AUC for the two types of benchmark models are not statistically different from each other. This confirms that contagion effects are indeed the factor that drives our results. The last two lines in Table 7 show the results for two robustness models that use one-quarter lagged networks for constructing the "Network dummy" and "Network Sum" contagion variables. The numbers indicate that using the previous quarter network when accounting for contagion does not bring any important change in the results.

[12] Other values for the policymaker's preference parameter have also been considered and are discussed in the robustness analysis in Section 4.3.

To sum up, the specifications with country and network contagion variables outperform the benchmark model both in their in-sample fit and their out-of-sample forecast performance. Overall, the models with tail dependencies have the highest in-sample fit and out-of-sample forecast performance given their consistently highest AUC measures that in most cases are statistically significantly larger than those of the benchmark model. These results give a direct support for proxying for interconnections or contagion beyond focusing on individual banks when building an early-warning model for bank failures.

4.2. Case studies

In order to illustrate the basic idea of the early-warning models incorporating tail-dependence measures as well as to show its performance, selected case studies of the model predictions are displayed in Figs. 3–6. Figs. 3 (Dexia) and 5 (National Bank of Greece) show the predicted distress probabilities for the selected banks when the full benchmark model (using bank-specific and country-level banking sector and macro-financial vulnerabilities) is estimated without the tail-dependence network or contagion dummies. The figures show the out-of-sample predicted (absolute) probabilities for a bank to be in a pre-distress period (vulnerable state), during which a distress event could occur in the next eight quarters. Moreover, the figures also show the percentile distress probabilities, which illustrate how high the predicted probabilities are compared to the predicted probabilities in the sample. Finally, in the figures the distress probability lines are shaded with grey when the predicted probabilities are above the time-varying optimal threshold, i.e., when the model issues an early-warning signal. As can be seen from the figures, the benchmark model seems to perform rather well for selected cases in predicting individual bank failures out-of-sample.

Figs. 4 (Dexia) and 6 (National Bank of Greece) show the model predictions for the individual banks as well as for their estimated neighbors in the tail-dependence network. In this case, the benchmark model is augmented with "Network Sum". In other words, the model takes into account potential signals from its estimated neighborhood when assessing the distress probability of a bank i and whether the probability is considered as an early-warning signal or not (the optimal threshold is calibrated accordingly). As can be seen from the figures, in both example cases, there are bank distress events (marked with filled dots) in the estimated neighborhood of the banks. Moreover, the model is in most cases able to predict the distress events in advance. This points to the regularity that neighboring banks to a distressed bank have also experienced distress. Further, the figures also clearly illustrate the high interconnectedness and complexity of the European banking sector that goes beyond national borders and thus motivates the use of

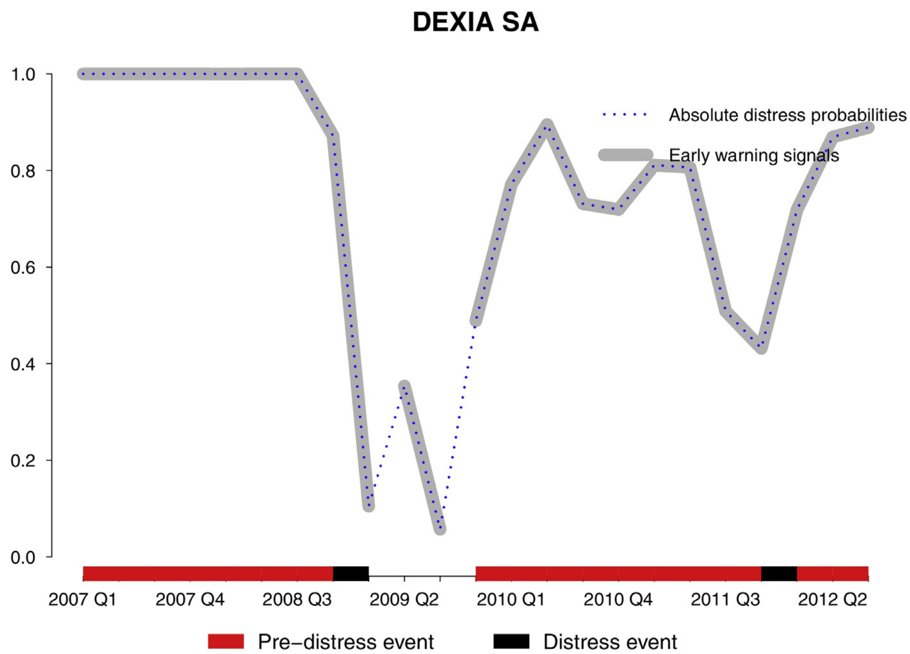

Fig. 3. Predicted probability for Dexia, 2007Q1–2012Q3.

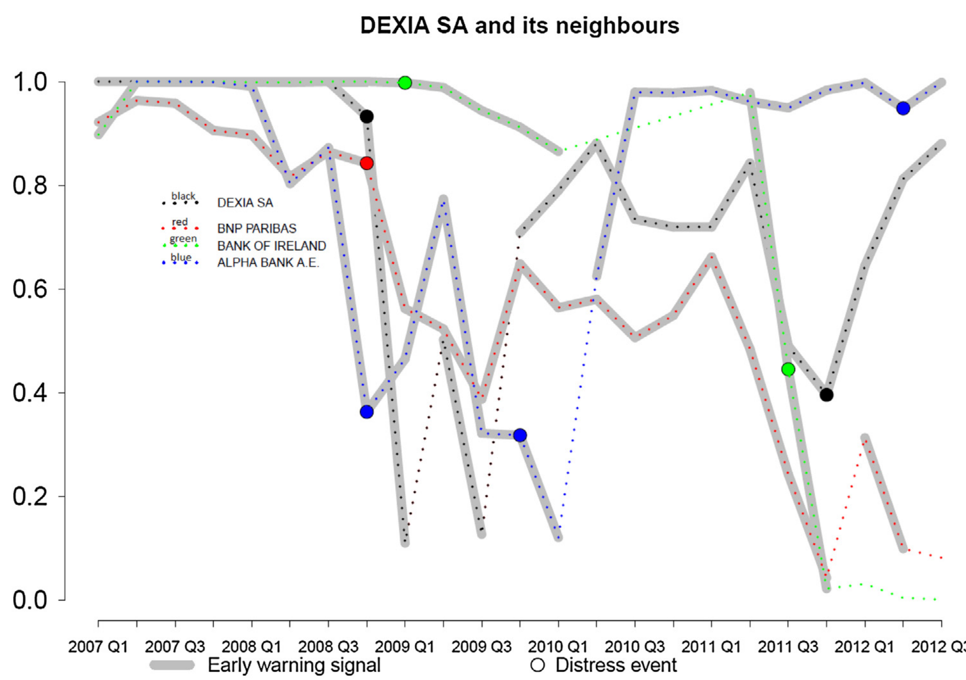

Fig. 4. Predicted probability for Dexia and its neighbors, 2007Q1–2012Q3.

proxies for interconnections and contagion, such as the tail-dependence network presented in this paper.

4.3. Robustness analysis

The robustness of the results has been assessed with a wide range of different tests. The results of the robustness evaluations can be found in Table 9 in Appendix A.[13] In all cases, the importance of proxying for interconnections or contagion beyond focusing on the bank-specific factors is still valid.

First, we test a number of different early-warning specifications. To start with, the analysis is performed for three different policymaker's preference parameters $\mu = 0.80$, $\mu = 0.85$ and $\mu = 0.90$. Given the distribution of pre-distress events and tranquil periods, we focus on the high-end range of policymaker's preference μ. From Table 9, it appears that the model performs very well for all three parameter values. In terms of the relative usefulness, the best performing model is for $\mu = 0.80$. The results are qualitatively the same for all three parameter values (i.e., the importance of proxying for interconnections is strongly supported). The magnitude of the

[13] Due to space constraints the estimated coefficients of the robustness analysis are not shown in Appendix A. However, the results are available upon request.

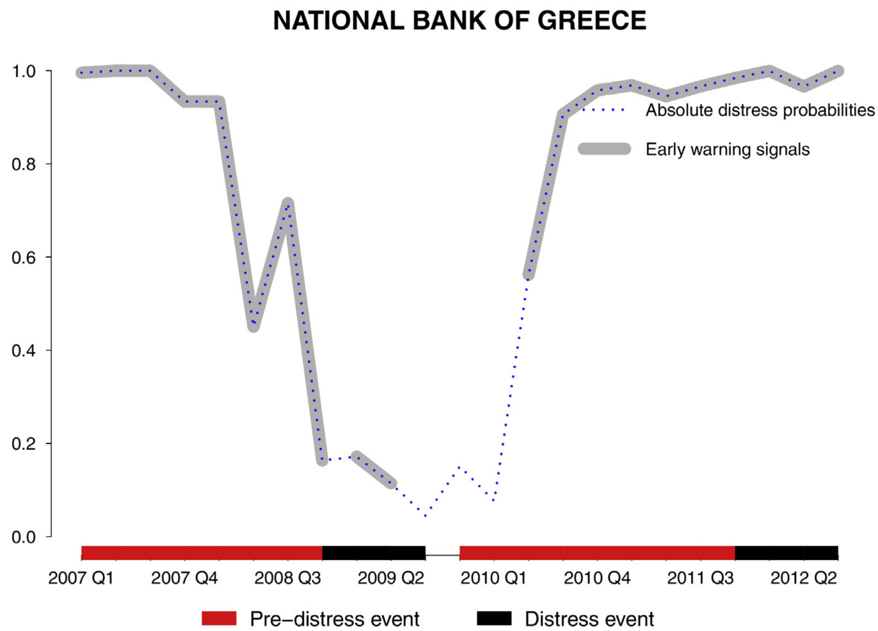

Fig. 5. Predicted probability for National Bank of Greece, 2007Q1–2012Q3.

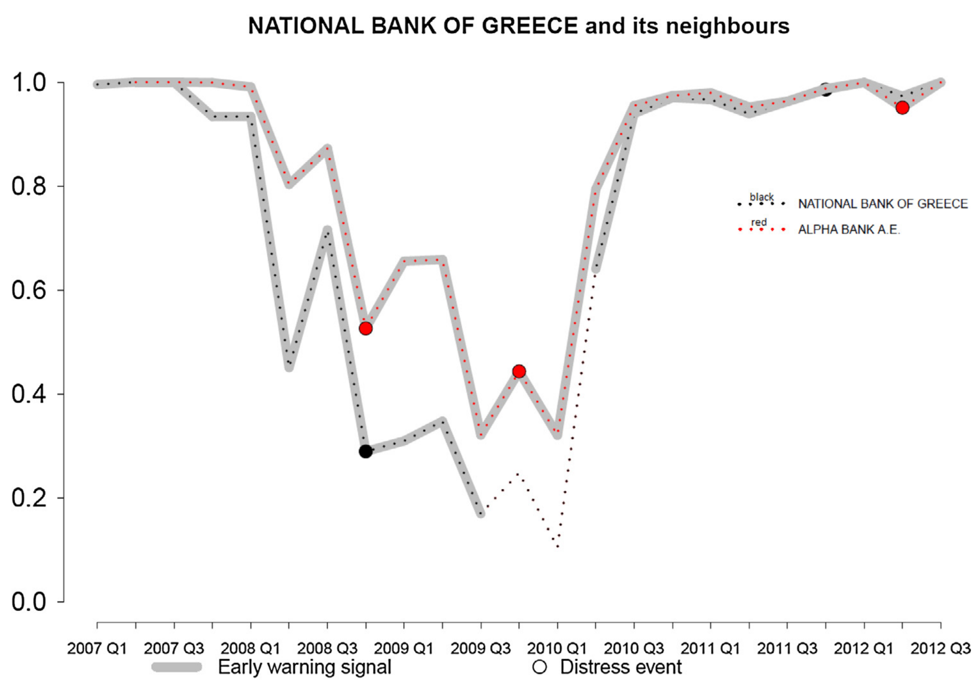

Fig. 6. Predicted probability for National Bank of Greece and its neighbors, 2007Q1–2012Q3.

relative Usefulness changes as expected: lower μ values greatly reduce the false positive rate, and implicitly the costs related to false alarms, which is for some cases done at the expense of missing crises. Moreover, these results are also robust to different model specifications and changes in forecast horizons.

Second, we consider a possible second benchmark model, where we compute the performance of the signaling model when the extra variable added to the second step of the estimation simply uses the signals of the initial estimation. That is, the second step is estimated but with no information of contagion. In this way, we are able to assess whether the performance improvement descends from simply including an additional variable or whether it is actually related to contagion effects. This model is called "2est Benchmark" in Table 9. We can see for all robustness estimations that simply including one additional variable with initial distress signals to the second estimation step does not improve model performance. This further confirms that contagion effects among banks are indeed the driver behind our results.

Moreover, we tested whether the results are robust for alterations of the estimated tail-dependence network. As can be seen in the last two rows of the panels in Table 9, we tested whether lagging network relations by one quarter improves model performance, but could not find an improvement compared to the no-lag case. Further, in the third part of Table 9 we also test whether using actual bank distress events in addition to predicted distress probabilities would improve model performance. The information of actual distress events slightly improved the model predictive performance but this difference is not statistically significant.

We also test the convergence of the predicted distress probabilities for the out-of-sample quarters by repeating the third and fourth steps of the empirical strategy presented in Section 2.3. This amounts to re-computing the network-based contagion variables and re-estimating the augmented logit model until the change in the predicted distress probabilities is less than one percentage point.[14] We perform this convergence test in order to avoid possible inconsistencies between the predicted distress probabilities of the different estimation steps. The results concerning the convergence for network-based contagion models, with policymaker's preference parameter μ of 0.85, are presented in the last part of Table 9. We can see that the results are very similar to the simpler two estimation procedure, with a slightly smaller magnitude. This confirms that only one estimation of the contagion augmented early-warning model should be enough to obtain stable signals and supports the robustness of our previous results.

Finally, among other tests[15], we assessed whether the results would change if a persistence condition would be applied to the estimated dependence network (i.e., estimated network relation has persisted for 2 quarters instead of a concurrent relation). The results are qualitatively similar, but the model performance is slightly poorer. Likewise, we have also tested various z-scores for deriving binary network links without significant changes in results.

5. Conclusion

Building on the literature on systemic risk and financial contagion, the paper introduced estimated network linkages into an early-warning model to predict bank distress. The approach applied in this paper estimates tail-dependence networks (Poon et al., 2004) from equity returns of 171 European banks in 1999Q1–2012Q3 and combines that with a bank-level early-warning model (Betz et al., 2014). Beyond standard bank-level risk drivers and macro-financial indicators, a tail-dependence network provides additional information about market's view on bank interconnectedness in situations of elevated financial stress. Thus, it can provide information of potentially vulnerable banks following an early-warning signal or a bank failure, and the potential for financial contagion and a systemic banking crisis.

The paper finds that the early warning models including estimated tail dependencies consistently outperform the benchmark models, which cover solely vulnerabilities coming from bank-specific, sector-level and macro-financial imbalances in order to predict bank distress events. These results highlight the importance of including measures of interconnectedness and proxies for contagion when building early-warning models. From a policy perspective, they underline the necessity for including information on the structure of the banking system to complement macro and micro-prudential analysis of individual bank's risk. Overall, the paper provides early steps toward a unified representation of cyclical and cross-sectional dimensions of systemic risk and contagion risk.

For future work, it is worth noting the general nature of the framework for combining networks and distress models. While this study is at the bank level, the same procedures would equally well apply at the country level or for another industry, such as insurers and other financial institutions. Another line of future work could focus on the use of the provided framework with a range of interconnectedness measures. In contrast to the classification task performed by early-warning models, for which "ground truth" exists and ex-post evaluations are possible, bank networks and interconnectedness measures do not have similar outcome data to steer modeling. Hence, the framework presented in this paper provides an ideal setting for testing the extent to which interconnectedness measures are proxying for the true channel of distress contagion.

[14] The country-based contagion model using the Country Dummy does not require the re-estimation of the augmented early-warning model.

[15] We do not report herein all the results of the robustness tests but they are available upon request.

Appendix A.

Table 8
Descriptive statistics for the full sample from 2000Q1 to 2012Q3 and 171 banks.

	N	Min	Max	Mean	St. Dev.	Median	Skewness	Kurtosis
Total leverage ratio	6974	0.00	0.48	0.07	0.04	0.06	2.66	16.08
Reserves for NPLs to non-performing assets	4146	0.00	96.50	1.40	3.23	0.91	24.92	723.75
ROA	6942	−0.05	0.04	0.01	0.01	0.01	−1.63	11.09
Loan loss provisions to total loans	6096	−0.01	0.06	0.01	0.01	0.00	2.75	12.04
Cost to income	6761	0.00	1.71	0.60	0.17	0.60	1.01	6.21
ROE	6944	−1.45	0.47	0.07	0.14	0.08	−4.72	34.78
Interest expenses to liabilities	6782	0.00	0.12	0.03	0.01	0.03	1.67	5.65
Deposits to liabilities	6762	0.00	0.99	0.50	0.22	0.50	0.02	−0.65
Net short-term borrowing to liabilities	5743	−0.35	0.95	0.21	0.17	0.2	0.68	1.21
Share of trading income to revenue	6698	−0.76	0.82	0.06	0.13	0.04	−0.31	9.66
Total assets to GDP	8558	0.83	35.24	3.25	4.12	2.60	6.20	40.33
Non-core liability growth	8045	−0.12	0.12	0.00	0.02	0.00	−0.08	3.27
Debt to equity	8729	3.87	40.86	14.32	4.16	13.00	0.99	1.31
Loans to deposits	8729	1.00	7.42	2.34	0.79	2.31	2.20	9.17
Debt securities to liabilities	8729	0.00	0.28	0.15	0.08	0.15	−0.40	−0.73
Mortgages to loans	8729	0.01	0.34	0.16	0.07	0.15	0.41	−0.71
Real GDP growth	9088	−0.10	0.08	0.00	0.01	0.01	−1.01	11.56
Inflation	9098	−0.03	0.07	0.01	0.01	0.01	0.73	2.78
Stock price growth	9140	−0.44	1.45	0.00	0.13	0.01	0.62	7.62
House prices growth	8739	−0.40	0.41	0.01	0.03	0.01	0.80	51.38
Long-term government bond yield	8905	0.02	0.35	0.05	0.02	0.04	7.58	81.65
International investment position to GDP	8921	−1.89	1.87	−0.22	0.37	−0.17	0.05	2.13
Government debt to GDP	9098	0.03	1.7	0.73	0.3	0.66	0.12	−0.41
Private sector credit flow to GDP	8912	−0.22	0.71	0.10	0.08	0.09	1.40	5.21

Table 9
Robustness evaluation of the early-warning models, out-of-sample 2007Q1–2012Q3: sensitivity to μ and use of historical distresses.

Model	AUC	U_r	FN rate	FP rate	TN rate	TP rate
μ = 0.80						
1est Benchmark	0.8941	0.6295	0.2230	0.1218	0.8782	0.7770
2est Benchmark	0.8948	0.6286	0.2230	0.1226	0.8774	0.7770
Country Dummy	0.8951	0.6277	0.2158	0.1293	0.8707	0.7842
Country Share	0.8990***	0.6250	0.2194	0.1285	0.8715	0.7806
Network Dummy	0.8985	0.6214	0.1799	0.1642	0.8358	0.8201
Network Sum	0.9009**	0.6610	0.1906	0.1226	0.8774	0.8094
Network Dummy w/Lag	0.8974	0.6259	0.1799	0.1605	0.8395	0.8201
Network Sum w/Lag	0.9009**	0.6655	0.1978	0.1129	0.8871	0.8022
μ = 0.90						
1est Benchmark	0.8941	0.4978	0.1079	0.3016	0.6984	0.8921
2est Benchmark	0.8930	0.4933	0.1223	0.2793	0.7207	0.8777
Country Dummy	0.8936	0.4733	0.1259	0.2927	0.7073	0.8741
Country Share	0.8974**	0.4970	0.1187	0.2823	0.7177	0.8813
Network Dummy	0.8972**	0.5022	0.1043	0.3039	0.6961	0.8957
Network Sum	0.8978	0.5208	0.1079	0.2786	0.7214	0.8921
Network Dummy w/Lag	0.8961*	0.5022	0.1079	0.2972	0.7028	0.8921
Network Sum w/Lag	0.8969	0.5260	0.1151	0.2600	0.7400	0.8849
μ = 0.85, with historical and estimated distresses						
1est Benchmark	0.8941	0.5800	0.1799	0.2095	0.7905	0.8201
2est Benchmark	0.8944	0.5770	0.1799	0.2125	0.7875	0.8201
Country Dummy	0.8942	0.5879	0.1655	0.2184	0.7816	0.8345
Country Share	0.8966*	0.5877	0.1619	0.2229	0.7771	0.8381
Network Dummy	0.8973	0.6320	0.1475	0.1954	0.8046	0.8525
Network Sum	0.8974	0.6454	0.1691	0.1568	0.8432	0.8309
Network Dummy w/Lag	0.8973	0.6169	0.1547	0.2021	0.7979	0.8453
Network Sum w/Lag	0.8970	0.6399	0.1763	0.1538	0.8462	0.8237
μ = 0.85, with convergence of out-of-sample distress signals						
1est Benchmark	0.8941	0.5800	0.1799	0.2095	0.7905	0.8201
2est Benchmark	0.8944	0.5770	0.1799	0.2125	0.7875	0.8201
Network Dummy	0.8980*	0.5998	0.1331	0.2444	0.7556	0.8669
Network Sum	0.8985*	0.6308	0.1835	0.1545	0.8455	0.8165
Network Dummy w/Lag	0.8969	0.5830	0.1475	0.2444	0.7556	0.8525
Network Sum w/Lag	0.8970	0.6230	0.1793	0.1838	0.8207	0.8162

This table presents the performance of different early-warning model specifications. The analysis is performed for three different policymaker's preference parameters μ = 0.80, μ = 0.85 and μ = 0.90. Results are presented for the area under the ROC curve (AUC) and relative usefulness (U_r) goodness of fit measures. Model performance is evaluated out-of-sample, for the period 2007Q1 - 2012Q3. Detailed values for the false negative (FN), false positive (FP) rates and their counterparts are also provided. Significance codes for differences of AUC measures between contagion models and the single estimation benchmark model are:

* 0.1.
** 0.05.
*** 0.01.

References

Agarwal, V., Taffler, R., 2008. Comparing the performance of market-based and accounting-based bankruptcy prediction models. J. Banking Finance 32 (8), 1541–1551.

Allen, F., Babus, A., 2009. Networks in finance. In: The Network Challenge: Strategy, Profit, and Risk in An Interlinked World, p. 367.

Allen, F., Gale, D., 2000. Financial contagion. J. Polit. Econ. 108 (1), 1–33.

Bae, K.-H., Karolyi, G.A., Stulz, R.M., 2003. A new approach to measuring financial contagion. Rev. Financial Stud. 16 (3), 717–763.

Battiston, S., Gatti, D.D., Gallegati, M., Greenwald, B., Stiglitz, J.E., 2012. Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. J. Econ. Dyn. Control 36 (8), 1121–1141.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., Caldarelli, G., 2012. Debtrank: too central to fail? Financial networks, the fed and systemic risk. Sci. Rep. 2.

Betz, F., Oprica, S., Peltonen, T.A., Sarlin, P., 2014. Predicting distress in European banks. J. Banking Finance 45, 225–241.

Bongini, P., Laeven, L., Majnoni, G., 2002. How good is the market at assessing bank fragility? A horse race between different indicators. J. Banking Finance 26 (5), 1011–1028.

Borio, C., Lowe, P., 2002. Asset Prices, Financial and Monetary Stability: Exploring the Nexus. BIS Working Papers 114. Bank for International Settlements.

Brunnermeier, M.K., 2008. Deciphering the Liquidity and Credit Crunch 2007–08. Technical Report. National Bureau of Economic Research.

Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. Rev. Financial Stud. 22 (6), 2201–2238.

Cifuentes, R., Ferrucci, G., Shin, H.S., 2005. Liquidity risk and contagion. J. Eur. Econ. Assoc. 3 (2–3), 556–566.

Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extreme value analyses. Extremes 2 (4), 339–365.

Demirgüç-Kunt, A., Detragiache, E., 1998 March. The determinants of banking crises in developing and developed countries. IMF Staff Papers 45 (1), 81–109.

Demirgüç-Kunt, A., Detragiache, E., 2000. Monitoring banking sector fragility. A multivariate logit approach. World Bank Econ. Rev. 14 (2), 287–307.

Diebold, Francis X., Yılmaz, Kamil, 2014. On the network topology of variance decompositions: measuring the connectedness of financial firms. J. Econ.

Engle, R., 2012. Dynamic Conditional Beta, Working Paper, Available at SSRN: http://ssrn.com/abstract=2084872.

Flannery, M.J., 1998. Using market information in prudential bank supervision: a review of the US empirical evidence. J. Money Credit Banking, 273–305.

Freixas, X., Parigi, B.M., Rochet, J.-C., 2000. Systemic risk, interbank relations, and liquidity provision by the central bank. J. Money Credit Banking, 611–638.

Gai, P., Haldane, A., Kapadia, S., 2011. Complexity, concentration and contagion. J. Monetary Econ. 58 (5), 453–470.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48, 1779–1801.

González-Hermosillo, B., 1999. Developing indicators to provide early warnings of banking crises. Finance Dev. 36, 36–39.

Gropp, R., Moerman, G., 2004 April. Measurement of contagion in banks' equity prices. J. Int. Money Finance 23 (3), 405–459.

Gropp, R., Duca, M.L., Vesala, J., 2009. Cross-Border Bank Contagion in Europe. Technical Report 1.

Hale, G., Kapan, T., Minoiu, C., 2014. Crisis Transmission in the Global Banking Network.

Hartmann, P., Straetmans, S., De Vries, C.G., 2004. Asset market linkages in crisis periods. Rev. Econ. Stat. 86 (1), 313–326.

Hartmann, P., Straetmans, S., De Vries, C.G., 2005. Banking System Stability: A Cross-Atlantic Perspective. Technical Report. National Bureau of Economic Research.

Hautsch, N., Schaumburg, J., Schienle, M., 2014. Financial network systemic risk contributions. Rev. Finance.

Hautsch, N., Schaumburg, J., Schienle, M., 2014. Forecasting systemic impact in financial networks. Int. J. Forecast.

Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. Ann. Stat. 3 (5), 1163–1174.

Huisman, R., Koedijk, K.G., Kool, C.J.M., Palm, F., 2001. Tail-index estimates in small samples. J. Business Econ. Stud. 19 (1), 208–216.

Kaminsky, G.L., Reinhart, C.M., 1999 June. The twin crises: the causes of banking and balance-of-payments problems. Am. Econ. Rev. 89 (3), 473–500.

Lainà, P., Nyholm, J., Sarlin, P., 2015. Leading indicators of systemic banking crises: Finland in a panel of EU countries. Rev. Financial Econ.

Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83 (1), 169–187.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. J. Finance 56 (2), 649–676.

Loretan, M., Phillips, P.C.B., 1994. Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets. J. Emp. finance 1 (2), 211–248.

Milne, A., 2014. Distance to default and the financial crisis. J. Financial Stab. 12, 26–36.

Minoiu, C., Kang, C., Subrahmanian, V.S., Berea, A., 2013. Does Financial Connectedness Predict Crises? IMF Working Papers 13/267. International Monetary Fund.

Nier, E., Yang, J., Yorulmazer, T., Alentorn, A., 2007. Network models and financial stability. J. Econ. Dyn. Control 31 (6), 2033–2060.

Oet, M., Bianco, T., Gramlich, D., Ong, S.J., 2013. SAFE: an early warning system for systemic banking risk. J. Banking Finance 37 (11), 4510–4533.

Peltonen, T., Rancan, M., Sarlin, P., 2014. Interconnectedness of the banking sector as a vulnerability to crises. Mimeo.

Poghosyan, T., Čihak, M., 2011. Determinants of bank distress in Europe: evidence from a new data set. J. Financial Serv. Res. 40 (3), 163–184.

Poon, S.-H., Rockinger, M., Tawn, J., 2004. Extreme value dependence in financial markets: diagnostics, models, and financial implications. Rev. Financial Stud. 17 (2), 581–610.

Sarlin, P., 2014. Macroprudential Oversight, Risk Communication and Visualization. J. Financial Stab. (forthcoming).

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

Tirole, J., 2011. Illiquidity and all its friends. J. Econ. Lit. 49 (2), 287–325.

Tjur, Tue, 2009. Coefficients of determination in logistic regression models – a new proposal: the coefficient of discrimination. Am. Stat. 63 (4), 366–372.