Deliverable 4.5

MSCA-ITN Training for Big Data in Financial Research and Risk Management
“BigDataFinance”

Grant Agreement: 675044

This is Deliverable 4.5 of the Work Package 4 (WP4) in “Training for Big Data in Financial Research and Risk Management” (BigDataFinance) Innovative Training Network Marie Sklodowska-Curie project 2015-2019.

Name of the deliverable: A report on a tested and validated system for risk management with real data and simulated stressed scenarios.

Description:

The deliverable comprises of one published paper and one working paper as follows:


   Portfolio credit risk models estimate the range of potential losses due to defaults or deteriorations in credit quality. Most of these models perceive default correlation as fully captured by the dependence on a set of common underlying risk factors. In light of empirical evidence, the ability of such a conditional independence framework to accommodate for the occasional default clustering has been questioned repeatedly. Thus, financial institutions have relied on stressed correlations or alternative copulas with more extreme tail dependence. In this paper, we propose a different remedy—augmenting systematic risk factors with a contagious default mechanism which affects the entire universe of credits. We construct credit stress propagation networks and calibrate contagion parameters for infectious defaults. The resulting framework is implemented on synthetic test portfolios wherein the contagion effect is shown to have a significant impact on the tails of the loss distributions.

Robustness of credit portfolio models is of great interest for financial institutions and regulators, since mispecified models translate to insufficient capital buffers and a crisis-prone financial system. In this paper we propose a method to enhance credit portfolio models based on the model of Merton by incorporating contagion effects. While in most models the risks related to financial interconnectedness are neglected, we use Bayesian networks to uncover the direct and indirect relationships between credits, while maintaining the convenient representation of factor models. A range of techniques to learn the structure and parameters of financial networks from real Credit Default Swaps data is studied and evaluated. Our approach is demonstrated in detail in a stylized portfolio and the impact on standard risk metrics is estimated.

Date, place: December 13th, 2018, Amsterdam, Netherlands

Name, position: Ioannis Anagnostou, Marie Sklodowska-Curie Fellow, ING Groep N.V.
Research Article

Incorporating Contagion in Portfolio Credit Risk Models Using Network Theory

Ioannis Anagnostou,1,2 Sumit Sourabh,1,2 and Drona Kandhai1,2

1Computational Science Lab, University of Amsterdam, Science Park 904, 1098XH Amsterdam, Netherlands
2Quantitative Analytics, ING Bank, Foppingadreef 7, 1102BD Amsterdam, Netherlands

Correspondence should be addressed to Ioannis Anagnostou; i.anagnostou@uva.nl

Received 20 September 2017; Accepted 29 November 2017; Published 8 January 2018

Academic Editor: Thiago C. Silva

Copyright © 2018 Ioannis Anagnostou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Portfolio credit risk models estimate the range of potential losses due to defaults or deteriorations in credit quality. Most of these models perceive default correlation as fully captured by the dependence on a set of common underlying risk factors. In light of empirical evidence, the ability of such a conditional independence framework to accommodate for the occasional default clustering has been questioned repeatedly. Thus, financial institutions have relied on stressed correlations or alternative copulas with more extreme tail dependence. In this paper, we propose a different remedy—augmenting systematic risk factors with a contagious default mechanism which affects the entire universe of credits. We construct credit stress propagation networks and calibrate contagion parameters for infectious defaults. The resulting framework is implemented on synthetic test portfolios wherein the contagion effect is shown to have a significant impact on the tails of the loss distributions.

1. Introduction

One of the main challenges in measuring the risk of a bank's portfolio is modelling the dependence between default events. Joint defaults of many issuers over a fixed period of time may lead to extreme losses; therefore, understanding the structure and the impact of default dependence is essential. To address this problem, one has to take into consideration the existence of two distinct sources of default dependence. On the one hand, performance of different issuers depends on certain common underlying factors, such as interest rates or economic growth. These factors drive the evolution of a company's financial success, which is measured in terms of its rating class or the probability of default. On the other hand, default of an issuer may, too, have a direct impact on the probability of default of a second dependent issuer, a phenomenon known as contagion. Through contagion, economic distress initially affecting only one issuer can spread to a significant part of the portfolio or even the entire system. A good example of such a transmission of pressure is the Russian crisis of 1998-1999 which saw the defaults of corporate and subsovereign issuers heavily clustered following the sovereign default [1].

Most portfolio credit risk models used by financial institutions neglect contagion and rely on the conditional independence assumption according to which, conditional on a set of common underlying factors, defaults occur independently. Examples of this approach include the Asymptotic Single Risk Factor (ASRF) model [2], industry extensions of the model presented by Merton [3] such as the KMV [4, 5] and CreditMetrics [6] models, and the two-factor model proposed recently by Basel Committee on Banking Supervision for the calculation of Default Risk Charge (DRC) to capture the default risk of trading book exposures [7]. A considerable amount of literature has been published on the conditional independence framework in standard portfolio models; see, for example, [8, 9].

Although conditional independence is a statistically and computationally convenient property, its empirical validity has been questioned on a number of occasions, where researchers investigated whether dependence on common factors can sufficiently explain the default clustering which
occurs from time to time. Schönbucher and Schubert [10] suggest that the default correlations that can be achieved with this approach are typically too low in comparison with empirical default correlations, although this problem becomes less severe when dealing with large diversified portfolios. Das et al. [11] use data on US corporations from 1979 to 2004 and reject the hypothesis that factor correlations can sufficiently explain the empirically observed default correlations in the presence of contagion. Since a realistic credit risk model is required to put the appropriate weight on scenarios where many joint defaults occur, one may choose to use alternative copulas with tail dependence which have the tendency to generate large losses simultaneously [12]. In that case, however, the probability distribution of large losses is specified a priori by the chosen copula, which seems rather unintuitive [13].

One of the first models to consider contagion in credit portfolios was developed by Davis and Lo [14]. They suggest a way of modelling default dependence through infection in a static framework. The main idea is that any defaulting issuer may infect any other issuer in the portfolio. Giesecke and Weber [15] propose a reduced-form model for contagion phenomena, assuming that they are due to the local interaction of companies in a business partner network. The authors provide an explicit Gaussian approximation of the distribution of portfolio losses and find that, typically, contagion processes have a second-order effect on portfolio losses. Lando and Nielsen [16] use a dynamic model in continuous time based on the notion of mutually exciting point processes. Apart from reduced-form models for contagion, which aim to capture the influence of infectious defaults to the default intensities of other issuers, structural models were developed as well. Jarrow and Yu [17] generalize existing models to include issuer-specific counterparty risks and illustrate their effect on the pricing of defaultable bonds and credit derivatives. Egoiff et al. [18] use network-like connections between issuers that allow for a variety of infections between firms. However, their structural approach requires a detailed microeconomic knowledge of debt structure, making the application of this model in practice more difficult than that of Davis and Lo’s simple model. In general, since the interdependencies between borrowers and lenders are complicated, structural analysis has mostly been applied to a small number of individual risks only.

Network theory can provide us with tools and insights that enable us to make sense of the complex interconnected nature of financial systems. Hence, following the 2008 crisis, network-based models have been frequently used to measure systemic risk in finance. Among the first papers to study contagion using network models was [19], where Allen and Gale show that a fully connected and homogeneous financial network results in an increased system stability. Contagion effects using network models have also been investigated in a number of related articles; see, for example, [20–24]. The issue of too-central-to-fail was shown to be possibly more important than too-big-to-fail by Battiston et al. in [25], where DebtRank, a metric for the systemic impact of financial institutions, was introduced. DebtRank was further extended in a series of articles; see, for example, [26–28].

The need for development of complexity-based tools in order to complement existing financial modelling approaches was emphasized by Battiston et al. [29], who called for a more integrated approach among academics from multiple disciplines, regulators, and practitioners.

Despite substantial literature on portfolio credit risk models and contagion in finance, specifying models, which take into account both common factors and contagion while distinguishing between the two effects clearly, still proves challenging. Moreover, most of the studies on contagion using network models focus on systemic risk and the resilience of the financial system to shocks. The qualitative nature of this line of research can hardly provide quantitative risk metrics that can be applied to models for measuring the risk of individual portfolios. The aforementioned drawback is perceived as an opportunity for expanding the current body of research by contributing a model that would account for common factors and contagion in networks alike. Given the wide use of factor models for calculating regulatory and economic capital, as well as for rating and analyzing structured credit products, an extended model that can also accommodate for infectious default events seems crucial.

Our paper takes up this challenge by introducing a portfolio credit risk model that can account for two channels of default dependence: common underlying factors and financial distress propagated from sovereigns to corporates and subsovereigns. We augment systematic factors with a contagion mechanism affecting the entire universe of credits, where the default probabilities of issuers in the portfolio are immediately affected by the default of the country where they are registered and operating. Our model allows for extreme scenarios with realistic numbers of joint defaults, while ensuring that the portfolio risk characteristics and the average loss remain unchanged. To estimate the contagion effect, we construct a network using credit default swaps (CDS) time series. We then use CountryRank, a network-based metric, introduced in [30] to quantify the impact of a sovereign default event on the credit quality of corporate issuers in the portfolio. In order to investigate the impact of our model on credit losses, we use synthetic test portfolios for which we generate loss distributions and study the effect of contagion on the associated risk measures. Finally, we analyze the sensitivity of the contagion impact to rating levels and CountryRank. Our analysis shows that credit losses increase significantly in the presence of contagion. Our contributions in this paper are thus threefold: First, we introduce a portfolio credit risk model which incorporates both common factors and contagion. Second, we use a credit stress propagation network constructed from real data to quantify the impact of deterioration of credit quality of the sovereigns on corporates. Third, we present the impact of accounting for contagion which can be useful for banks and regulators to quantify credit, model, or concentration risk in their portfolios.

The rest of the paper is organized as follows. Section 2 provides an overview of the general modelling framework. Section 3 presents the portfolio model with default contagion and illustrates the network model for the estimation of contagion effects. In Section 4 we present empirical analysis of two
synthetic portfolios. Finally, in Section 5, we summarize our findings and draw conclusions.

2. Merton-Type Models for Portfolio Credit Risk

Most financial institutions use models that are based on some form of the conditional independence assumption, according to which issuers depend on a set of common underlying factors. Factor models based on the Merton model are particularly popular for portfolio credit risk. Our model extends the multifactor Merton model to allow for credit contagion. In this section, we present the basic portfolio modelling setup, outline the model of Merton, and explain how it can be specified as a factor model. A more detailed presentation of the multivariate Merton model is provided by [9].

2.1. Basic Setup and Notations. This subsection introduces the basic notation and terminology that will be used throughout this paper. In addition, we define the main risk characteristics for portfolio credit risk.

The uncertainty of whether an issuer will fail to meet its financial obligations or not is measured by its probability of default. For comparison reasons, this is usually specified with respect to a fixed time interval, most commonly one year. The probability of default then describes the probability of a default occurring in the particular time interval. The exposure at default is a measure of the extent to which one is exposed to an issuer in the event of, and at the time of, that issuer’s default. The default of an issuer does not necessarily imply that the creditor receives nothing from the issuer. The percentage of loss incurred over the overall exposure in the event of default is given by the loss given default. Typical values lie between 45% and 80%

Consider a portfolio of \( m \) issuers, indexed by \( i = 1, \ldots, m \), and a fixed time horizon of \( T = 1 \) year. Denote by \( e_i \) the exposure at default of issuer \( i \) and by \( p_i \) its probability of default. Let \( q_i \) be the loss given default of issuer \( i \). Denote by \( Y_i \) the default indicator, in the time period \([0, T]\). All issuers are assumed to be in a nondefault state at time \( t = 0 \). The default indicator \( Y_i \) is then a random variable defined by

\[
Y_i = \begin{cases} 
1 & \text{if issuer } i \text{ defaults} \\
0 & \text{otherwise}
\end{cases} \tag{1}
\]

which clearly satisfies \( P(Y_i = 1) = p_i \). The overall portfolio loss is defined as the random variable

\[
L = \sum_{i=1}^{m} q_i e_i Y_i. \tag{2}
\]

With credit risk in mind, it is useful to distinguish potential losses in expected losses, which are relatively predictable and thus can easily be managed, and unexpected losses, which are more complicated to measure. Risk managers are more concerned with unexpected losses and focus on risk measures relating to the tail of the distribution of \( L \).

2.2. The Model of Merton. Credit risk models are typically distinguished in structural and reduced-form models, according to their methodology. Structural models try to explain the mechanism by which default takes place, using variables such as asset and debt values. The model presented by Merton in [3] serves as the foundation for all these models. Consider an issuer whose asset value follows a stochastic process \((V_t)_{t \geq 0}\). The issuer finances itself with equity and debt. No dividends are paid and no new debt can be issued. In Merton’s model the issuer’s debt consists of a single zero-coupon bond with face value \( B \) and maturity \( T \). The values at time \( t \) of equity and debt are denoted by \( S_t \) and \( B_t \) and the issuer’s asset value is simply the sum of these; that is,

\[
V_t = S_t + B_t, \quad t \in [0, T]. \tag{3}
\]

Default occurs if the issuer misses a payment to its debtholders, which can happen only at the bond’s maturity \( T \). At time \( T \), there are only two possible scenarios:

(i) \( V_T > B \): the value of the issuer’s assets is higher than its debt. In this scenario the debtholders receive \( B_T = B \), the shareholders receive the remainder \( S_T = V_T - B \), and there is no default.

(ii) \( V_T \leq B \): the value of the issuer’s assets is less than its debt. Hence, the issuer cannot meet its financial obligations and defaults. In that case, shareholders hand over control to the bondholders, who liquidate the assets and receive the liquidation value in lieu of the debt. Shareholders pay nothing and receive nothing; therefore we obtain \( B_T = V_T, S_T = 0 \).

For these simple observations, we obtain the below relations:

\[
S_T = \max (V_T - B, 0) = (V_T - B)^+, \tag{4}
\]

\[
B_T = \min (V_T, B) = B - (B - V_T)^+. \tag{5}
\]

Equation (4) implies that the issuer’s equity at maturity \( T \) can be determined as the price of a European call option on the asset value \( V_T \) with strike price \( B \) and maturity \( T \), while (5) implies that the value of debt at \( T \) is the sum of a default-free bond that guarantees payment of \( B \) plus a short European put option on the issuer’s assets with strike price \( B \).

It is assumed that under the physical probability measure \( \mathbb{P} \) the process \((V_t)_{t \geq 0}\) follows a geometric Brownian motion of the form

\[
dV_t = \mu_V V_t dt + \sigma_V V_t dW_t, \quad t \in [0, T], \tag{6}
\]

where \( \mu_V \in \mathbb{R} \) is the mean rate of return on the assets, \( \sigma_V > 0 \) is the asset volatility, and \((W_t)_{t \geq 0}\) is a Wiener process. The unique solution at time \( T \) of the stochastic differential equation (6) with initial value \( V_0 \) is given by

\[
V_T = V_0 \exp \left( \left( \mu_V - \frac{\sigma_V^2}{2} \right) T + \sigma_V W_T \right) \tag{7}
\]

which implies that

\[
\ln V_T \sim \mathcal{N} \left( \ln V_0 + \left( \mu_V - \frac{\sigma_V^2}{2} \right) T, \sigma_V^2 T \right). \tag{8}
\]
Hence, the real-world probability of default at time $T$, measured at time $t = 0$, is given by
\[
\mathbb{P}(V_T \leq B) = \mathbb{P}(\ln V_T \leq \ln B) = \Phi\left( \frac{\ln(B/V_0) - \left(\mu_V - \sigma_V^2/2\right) T}{\sigma_V \sqrt{T}} \right).
\]
(9)

A core assumption of Merton’s model is that asset returns are lognormally distributed, as can be seen in (8). It is widely acknowledged, however, that empirical distributions of asset returns tend to have heavier tails; thus, (9) may not be an accurate description of empirically observed default rates.

2.3. The Multivariate Merton Model. The model presented in Section 2.2 is concerned with the default of a single issuer. In order to estimate credit risk at a portfolio level, a multivariate version of the model is necessary. A multivariate geometric Brownian motion with drift vector $\mu_V = (\mu_1, \ldots, \mu_m)^t$, vector of volatilities $\sigma_V = (\sigma_1, \ldots, \sigma_m)$, and correlation matrix $\Sigma$, is assumed for the dynamics of the multivariate asset value process $(V_t)_{t \geq 0}$ with $V_t = (V_{t,1}, \ldots, V_{t,m})^t$, so that for all $i$
\[
V_{T,i} = V_{0,i} \exp \left( \left( \mu_i - \frac{1}{2} \sigma_i^2 \right) T + \sigma_i W_{T,i} \right),
\]
(10)
where the multivariate random vector $W_T$ with $W_T = (W_{T,1}, \ldots, W_{T,m})^t$ is satisfying $W_t \sim N_m(0, \Sigma)$. Default takes place if $V_{T,i} \leq B_i$, where $B_i$ is the debt of company $i$. It is clear that the default probability in the model remains unchanged under simultaneously strictly increasing transformations of $V_{T,i}$ and $B_i$. Thus, one may define
\[
X_i = \frac{\ln V_{T,i} - \ln V_{0,i} - \left(\mu_i - (1/2) \sigma_i^2\right) T}{\sigma_i \sqrt{T}},
\]
\[
d_i = \frac{\ln B_i - \ln V_{0,i} - \left(\mu_i - (1/2) \sigma_i^2\right) T}{\sigma_i \sqrt{T}},
\]
(11)
and then default equivalently occurs if and only if $X_i \leq d_i$. Notice that $X_i$ is the standardized asset value log-return $\ln V_{T,i} - \ln V_{0,i}$. It can be easily shown that the transformed variables satisfy $(X_1, \ldots, X_m)^t \sim N_m(0, \Sigma)$ and their copula is the Gaussian copula. Thus, the probability of default for issuer $i$ is satisfying $p_i = \Phi(d_i)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. A graphical representation of Merton’s model is shown in Figure 1. In most practical implementations of the model, portfolio losses are modelled by directly considering an $m$-dimensional random vector $X = (X_1, \ldots, X_m)^t$ with $X \sim N_m(0, \Sigma)$ containing the standardized asset returns and a deterministic vector $d = (d_1, \ldots, d_m)$ containing the critical thresholds with $d_i = \Phi^{-1}(p_i)$ for given default probabilities $p_i$, $i = 1, \ldots, m$. The default probabilities are usually estimated by historical default experience using external ratings by agencies or model-based approaches.

2.4. Merton Model as a Factor Model. The number of parameters contained in the correlation matrix $\Sigma$ grows polynomially in $m$, and thus, for large portfolios it is essential to have a more parsimonious parameterization which is accomplished using a factor model. Additionally, factor models are particularly attractive due to the fact that they offer an intuitive interpretation of credit risk in relation to the performance of industry, region, global economy, or any other relevant indexes that may affect issuers in a systematic way. In the following we show how Merton’s model can be understood as a factor model. In the factor model approach, asset returns are linearly dependent on a vector $F$ of $p < m$ common underlying factors satisfying $F \sim N_p(0, \Omega)$. Issuer $i$’s standardized asset return is assumed to be driven by an issuer-specific combination $\tilde{F}_i = \alpha'_i F$ of the systematic factors
\[
X_i = \sqrt{\beta_i} \tilde{F}_i + \sqrt{1 - \beta_i} \epsilon_i,
\]
(12)
where $\tilde{F}_i$ and $\epsilon_1, \ldots, \epsilon_m$ are independent standard normal variables and $\epsilon_i$ represents the idiosyncratic risk. Consequently, $\beta_i$ can be seen as a measure of sensitivity of $X_i$ to systematic risk, as it represents the proportion of the $X_i$ variation that is explained by the systematic factors. The correlations between asset returns are given by
\[
\rho(X_i, X_j) = \text{cov}(X_i, X_j) = \sqrt{\beta_i \beta_j \text{cov}(\tilde{F}_i, \tilde{F}_j)} = \sqrt{\beta_i \beta_j \alpha'_i \Omega \alpha'_j}
\]
(13)
since $\tilde{F}_i$ and $\epsilon_1, \ldots, \epsilon_m$ are independent and standard normal and $\text{var}(X_i) = 1$. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{In Merton’s model, default of issuer $i$ occurs if at time $T$ asset value $V_{T,i}$ falls below debt value $B_i$, or equivalently if $X_i := (\ln V_{T,i} - \ln V_{0,i} - (\mu_i - (1/2) \sigma_i^2) T)/\sigma_i \sqrt{T}$ falls below the critical threshold $d_i := (\ln B_i - \ln V_{0,i} - (\mu_i - (1/2) \sigma_i^2) T)/\sigma_i \sqrt{T}$. Since $X_i \sim N(0,1)$, $i$’s default probability, represented by the shaded area in the distribution plot, is satisfying $p_i = \Phi(d_i)$. Note that default can only take place at time $T$ does not depend on the path of the asset value process.}
\end{figure}
3. A Model for Credit Contagion

In the multifactor Merton model specified in Section 2.4, the standardized asset returns $X_i$, $i = 1, \ldots, m$, are assumed to be driven by a set of common underlying systematic factors, and the critical thresholds $d_i$, $i = 1, \ldots, m$, are satisfying $d_i = \Phi^{-1}(p_i)$ for all $i$. The only source of default dependence in such a framework is the dependence on the systematic factors. In the model we propose, we assume that, in the event of a sovereign default, contagion will spread to the corporate issuers in the portfolio that are registered and operating in that country, causing default probability to be equal to their CountryRank. In Section 3.1, we demonstrate how to calibrate the critical thresholds so that each corporate’s probability of default conditional on the default of the corresponding sovereign equals its CountryRank, while its unconditional default probability remains unchanged. In Section 3.2, we show how to construct a credit stress propagation network and estimate the CountryRank parameter.

3.1. Incorporating Contagion in Factor Models. Consider a corporate issuer $C_i$ and its country of operation $S$. Denote by $p_{C_i}$ the probability of default of $C_i$. Under the standard Merton model, default occurs if $C_i$’s standardized asset return $X_{C_i}$ falls below its default threshold $d_{C_i}$. The critical threshold $d_{C_i}$ is assumed to be equal to $\Phi^{-1}(p_{C_i})$ and is independent of the state of the country of operation $S$. In the proposed model, a corporate is subject to shocks from its country of operation; its corresponding state is described by a binary state variable. The state is considered to be stressed in the event of sovereign default. In this case, the issuer’s default threshold increases, causing it more likely to default, as the contagion effect suggests. In case the corresponding sovereign does not default, the corporates liquidity state is considered stable. We replace the default threshold $d_{C_i}$ with $d_{C_i}^*$, where

$$d_{C_i}^* = \begin{cases} d_{C_i}^\text{sd} & \text{if the corresponding sovereign defaults} \\ d_{C_i}^\text{nsd} & \text{otherwise} \end{cases} \quad (14)$$

or equivalently

$$d_{C_i}^* = \mathbb{1}_{[Y_S=1]} d_{C_i}^\text{sd} + \mathbb{1}_{[Y_S=0]} d_{C_i}^\text{nsd}. \quad (15)$$

We denote by $p_{S}$ the probability of default of the country of operation and by $Y_{C_i}$ the CountryRank parameter which indicates the increased probability of default of $C_i$ given the default of $S$. An example of the new default thresholds is shown in Figure 2. Our objective is to calibrate $d_{C_i}^\text{sd}$ and $d_{C_i}^\text{nsd}$ in such way that the overall default rate remains unchanged and $P(Y_{C_i} = 1 | Y_S = 1) = Y_{C_i}$. Denote by

$$\phi_2(x, y; \rho) = \frac{1}{2\sqrt{1 - \rho^2}} \exp \left( -\frac{x^2 + y^2 - 2\rho xy}{2(1 - \rho^2)} \right), \quad (16)$$

the density and distribution function of the bivariate standard normal distribution with correlation parameter $\rho \in (-1, 1)$.
Note that $d_{Ci}^d(\omega) = d_{Ci}^d$ for $\omega \in \{Y_{C_i} = 1, Y_S = 1\} \subset \{Y_S = 1\}$, and $d_{Ci}^d(\omega) = d_{Ci}^d$ for $\omega \in \{Y_{C_i} = 1, Y_S = 0\} \subset \{Y_S = 0\}$. We rewrite $P(Y_{C_i} = 1 | Y_S = 1)$ in the following way:

$$
P(Y_{C_i} = 1 | Y_S = 1) = \frac{1}{P(Y_S = 1)} P(Y_{C_i} = 1, Y_S = 1)
$$

$$
= \frac{1}{P_S} P \left[ X_{C_i} < d_{Ci}^d, X_S < d_S \right]
$$

$$
= \frac{1}{P_S} \Phi \left( d_{Ci}^d, d_S; \rho_{SC_i} \right).
$$

Using the above representation and given $d_S = \Phi^{-1}(\rho_S)$ and $\rho_{SC_i}$, one can solve the equation

$$
P \left( Y_{C_i} = 1 | Y_S = 1 \right) = \gamma_{C_i}
$$

over $d_{Ci}^d$. We proceed to the derivation of $d_{Ci}^d$ in such a way that the overall default probability remains equal to $p_C$. This constraint is important, since contagion is assumed to have no impact on the average loss. Clearly,

$$
p_{C_i} = P \left( Y_{C_i} = 1 \right)
$$

$$
= P \left( Y_{C_i} = 1, Y_S = 1 \right) + P \left( Y_{C_i} = 1, Y_S = 0 \right)
$$

$$
= P \left( Y_{C_i} = 1 | Y_S = 1 \right) P \left( Y_S = 1 \right)
$$

$$
+ P \left( Y_{C_i} = 1, Y_S = 0 \right)
$$

and thus

$$
P \left( Y_{C_i} = 1, Y_S = 0 \right) = p_{C_i} - \gamma_{C_i} \cdot p_S.
$$

The left-hand side of the above equation can be represented as follows:

$$
P \left( \text{corp.def} \cap \text{nsov.def} \right) = P \left[ X_{C_i} < d_{Ci}^d, X_S > d_S \right]
$$

$$
= P \left[ X_{C_i} < d_{Ci}^d \right] - P \left[ X_{C_i} < d_{Ci}^d, X_S < d_S \right]
$$

$$
= \Phi \left( d_{Ci}^d \right) - \Phi \left( d_{Ci}^d, d_S; \rho_{SC_i} \right).
$$

By use of the above and given $d_S = \Phi^{-1}(\rho_S)$ and $\rho_{SC_i}$, one can solve the previous equation over $d_{Ci}^d$.

### 3.2. Estimation of CountryRank

In this section, we elaborate on the estimation of the CountryRank parameter [30], which serves as the probability of default of the corporate conditional on the default of the sovereign. In addition, we provide details on the construction of the credit stress propagation network.

#### 3.2.1. CountryRank

In order to estimate contagion effects in a network of issuers, an algorithm such as DebtRank [31] is necessary. In the DebtRank calculation process, stress propagates even in the absence of defaults and each node can propagate stress only once before becoming inactive. The level of distress for a previously undistressed node is given by the sum of incoming stress from its neighbors with a maximum value of 1. Summing up the incoming stress from neighboring nodes seems reasonable when trying to estimate the impact of one node or a set of nodes to a network of interconnected balance sheets where links represent lending relationships. However, when trying to quantify the probability of default of a corporate given the infectious default of a sovereign node, one has to consider that there is significant overlap in terms of common stress, and thus, by summing we may be accounting for the same effect more than once. This effect is amplified in dense networks constructed from CDS data. Therefore, we introduce CountryRank as an alternative measure which is suited for our contagion model.

We assume that we have a hypothetical credit stress propagation network, where the nodes correspond to the issuers, including the sovereign, and the edges correspond to the impact of credit quality of one issuer on the other. The details of the network construction will be presented in Section 3.2.2. Given such a network, the CountryRank of the nodes can be defined recursively as follows:

(i) First, we stress the sovereign node and as a result its CountryRank is 1.

(ii) Let $\gamma_S$ be the CountryRank of the sovereign and let $e_{(i,j)}$ denote the edge weight between nodes $j$ and $k$. Given a node $C_i$, let $\rho = SC_iC_2 \cdots C_{i-1}C_i$ be a path without cycles from the sovereign node $S$, to the node $C_i$. The weight of the path $\rho$ is defined as

$$
w(\rho) = \gamma_S e_{(1,2)} \cdots e_{(i-1,i)},
$$

where $e_{(i,j)}$ are the respective edge weights between nodes $j$ and $k$ for $j \in \{1, \ldots, i-1\}$ and $k \in \{2, \ldots, i\}$. Let $p_1, \ldots, p_m$ be the set of all acyclic paths from the sovereign node to the corporate node $C_i$ and let $w(p_1), \ldots, w(p_m)$ be the corresponding weights. Then the CountryRank of node $C_i$ is defined as

$$
\gamma_{C_i} = \max_{1 \leq j \leq m} \left( \gamma_S \cdot w \left( p_j \right) \right).
$$

In order to compute the conditional probability of default of a corporate given the sovereign default analytically, we would need the joint distribution of probabilities of default of the nodes, which has an exponential computational complexity, and it is therefore intractable. Thus, we approximate the conditional probability by choosing the path with the maximum weight in the above definition for CountryRank.

The example in Figure 3 illustrates calculation of CountryRank for a hypothetical network. The network consists of a sovereign node $S$ and corporate nodes $C_1, C_2, C_3, C_4$. The edge labels indicate weights in network between two nodes. We initially stress the sovereign node which results in a CountryRank of 1 for node $S$. In the next step, the stress propagates to node $C_1$ and as a result its CountryRank is 0.9.
Then, node $C_2$ gets stressed giving it a CountryRank value 0.8. For node $C_3$, there are two paths from node $S$, so we pick the path through node $C_2$ having a higher weight of 0.48. Finally, there are three paths from node $S$ to node $C_4$, and the path with maximum weight is 0.27.

### 3.2.2. Network Construction.

Credit default swap spreads are market-implied indicators of probability of default of an entity. A credit default swap is a financial contract in which a protection seller $A$ insures a protection buyer $B$ against the default of a third party $C$. More precisely, regular coupon payments with respect to a contractual notional $N$ and a fixed rate $s$, the CDS spread, are swapped with a payment of $N(1 - RR)$ in the case of the default of $C$, where $RR$, the so-called recovery rate, is a contract parameter which represents the fraction of investment which is assumed to be recovered in the case of default of $C$.

**Modified $\epsilon$-Draw-Up.** We would like to measure to what extent changes in CDS spreads of different issuers occur simultaneously. For this, we use the notion of a modified $\epsilon$-draw-up to quantify the impact of deterioration of credit quality of one issuer on the other. Modified $\epsilon$-draw-up is an alteration of the $\epsilon$-draw-ups notion which is introduced in [32]. In that article, the authors use the notion of $\epsilon$-draw-ups to construct a network which models the conditional probabilities of spike-like comovements among pairs of CDS spreads. A modified $\epsilon$-draw-up is defined as an upward movement in the time series in which the amplitude of the movement, that is, the difference between the subsequent local maxima and current local minima, is greater than a threshold $\epsilon$. We record such local minima as the modified $\epsilon$-draw-ups. The $\epsilon$ parameter for a local minima at time $t$ is set to be the standard deviation in the time series between days $t - n$ and $t$, where $n$ is chosen to be 10 days. Figure 5 shows...
the time series of Russian Federation CDS with the calibrated modified ε-draw-ups using a history of 10 days for calibration.

Filtering Market Impact. Since we would like to measure the comovement of the time series $i$ and $j$, we exclude the effect of the external market on these nodes as follows. We calibrate the ε-draw-ups for the CDS time series of an index that does not represent the region in question; for instance, for Russian issuers we choose the iTraxx index which is the composite CDS index of 125 CDS referencing European investment grade credit. Then, we filter out those ε-draw-ups of node $i$ which are the same as the ε-draw-ups of the iTraxx index including a time lag $\tau$. That is, if iTraxx has a modified ε-draw-up on day $t$, then we remove the modified ε-draw-ups of node $i$ on days $t, t+1, \ldots, t+\tau$. We choose a time lag of 3 days for our calibration based on the input data which is consistent with the choice in [32].

Edges. After identifying the ε-draw-ups for all the issuers and filtering out the market impact, the edges in our network are constructed as follows. The weight of an edge in the credit stress propagation network from node $i$ to node $j$ is the conditional probability that if node $i$ has an epsilon draw-up on day $t$, then node $j$ also has an epsilon draw-up on days $t, t+1, \ldots, t+\tau$, where $\tau$ is the time lag. More precisely, let $N_i$ be the number of ε-draw-ups of node $i$ after filtering using iTraxx index and $N_{ij}$ epsilon draw-ups of node $i$ which are also epsilon draw-ups for node $j$ with the time lag $\tau$. Then, the edge weight $\omega_{ij}$ between nodes $i$ and $j$ is defined as $\omega_{ij} = N_i/N_{ij}$. Figure 6 shows the minimum spanning tree of the credit stress propagation network constructed using the CDS spread time series data of Russian issuers.

Uncertainty in CountryRank. We test the robustness of our CountryRank calibration by varying the number of days used for ε-parameter. The figure in Appendix B shows that the ε-parameter for Russian Federation CDS time series remains stable when we vary the number of days. We initially obtain time series of ε-parameters by calculating standard deviation in the last $n = 10, 15, 20$ days on all local minima indices of Russian Federation CDS. Subsequently, we calculate the mean of the absolute differences between the epsilon time series calculated and express this in units of the mean of Russian Federation CDS time series. The percentage difference is 1.38% between the 10-day ε-parameter and 15-day ε-parameter and 2.22% between the 10-day and 20-day ε-parameters.

Further, we quantify the uncertainty in CountryRank parameter as follows. For a corporate node, we calculate the absolute difference in CountryRank calculated using $n = 15$ and 20 days with CountryRank using $n = 10$ days for the ε-parameter. We then calculate this difference as a percentage of the CountryRank calculated using 10 days for ε-parameter for all corporations and then compute their mean. The mean difference between CountryRank calculated using $n = 15$ days and $n = 10$ days is 6.84% and $n = 20$ days and $n = 10$ days is 9.73% for the Russian CDS data set.

### Table 1: Systematic factor: index mapping.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>MSCI EUROPE</td>
</tr>
<tr>
<td>Asia</td>
<td>MSCI AC ASIA</td>
</tr>
<tr>
<td>North America</td>
<td>MSCI NORTH AMERICA</td>
</tr>
<tr>
<td>Latin America</td>
<td>MSCI EM LATIN AMERICA</td>
</tr>
<tr>
<td>Middle East and Africa</td>
<td>MSCI FM AFRICA</td>
</tr>
<tr>
<td>Pacific</td>
<td>MSCI PACIFIC</td>
</tr>
<tr>
<td>Materials</td>
<td>MSCI WRLD/MATERIALS</td>
</tr>
<tr>
<td>Consumer products</td>
<td>MSCI WRLD/CONSUMER DISCR</td>
</tr>
<tr>
<td>Services</td>
<td>MSCI WRLD/CONSUMER SVC</td>
</tr>
<tr>
<td>Financial</td>
<td>MSCI WRLD/FINANCIALS</td>
</tr>
<tr>
<td>Industrial</td>
<td>MSCI WRLD/INDUSTRIALS</td>
</tr>
<tr>
<td>Government</td>
<td>ITRAXX SOVX GLOBAL LIQUID INVESTMENT GRADE</td>
</tr>
</tbody>
</table>

### 4. Numerical Experiments

We implement the framework presented in Section 3 to synthetic test portfolios and discuss the corresponding risk metrics. Further, we perform a set of sensitivity studies and explore the results.

#### 4.1. Factor Model

We first set up a multifactor Merton model, as it was described in Section 2. We define a set of systematic factors that will represent region and sector effects. We choose 6 region and 6 sector factors, for which we select appropriate indexes, as shown in Table 1. We then use 10 years of index time series to derive the region and sector returns $F_{R(j)}$, $j = 1, \ldots, 6$ and $F_{S(k)}$, $k = 1, \ldots, 6$, respectively, and obtain an estimate of the correlation matrix $\Omega$, shown in Figure 7. Subsequently, we map all issuers to one region and one sector factor, $F_{R(i)}$ and $F_{S(i)}$, respectively. For instance, a Dutch bank will be associated with Europe and financial factors. As a proxy of individual asset returns, we use 10 years of equity or CDS time series, depending on the data availability for each issuer. Finally, we standardize the individual returns time series $(X_{ij})$ and perform the following Ordinary Least Squares regression against the systematic factor returns

$$X_{i,t} = \alpha_{R(i)} F_{R(i), t} + \alpha_{S(i)} F_{S(i), t} + \epsilon_{i,t}$$

(24)

to obtain $\hat{\alpha}_{R(i)}, \hat{\alpha}_{S(i)},$ and $\hat{\beta}_i = R^2$, where $R^2$ is the coefficient of determination, and it is higher for issuers whose returns are largely affected by the performance of the systematic factors.

#### 4.2. Synthetic Test Portfolios

To investigate the properties of the contagion model, we set up 2 test portfolios. For these portfolios, the resulting risk measures are compared to those of the standard latent variable model with no contagion. Portfolio A consists of 1 Russian government bond and 17 bonds issued by corporations registered and operating in the Russian Federation. As it is illustrated in Table 2, the issuers are of medium and low credit quality. Portfolio B represents a similar but more diversified setup with 4 sovereign bonds.
Complexity

Table 2: Rating classification for the test portfolios.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Issuers</th>
<th>Portfolio A</th>
<th>%</th>
<th>Issuers</th>
<th>Portfolio B</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>3</td>
<td></td>
<td>3.75%</td>
</tr>
<tr>
<td>AA</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>3</td>
<td></td>
<td>3.75%</td>
</tr>
<tr>
<td>A</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>22</td>
<td></td>
<td>27.50%</td>
</tr>
<tr>
<td>BBB</td>
<td>1</td>
<td></td>
<td>5.56%</td>
<td>39</td>
<td></td>
<td>48.75%</td>
</tr>
<tr>
<td>BB</td>
<td>15</td>
<td></td>
<td>83.33%</td>
<td>9</td>
<td></td>
<td>11.25%</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td></td>
<td>11.11%</td>
<td>3</td>
<td></td>
<td>3.75%</td>
</tr>
<tr>
<td>CCC/C</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>1</td>
<td></td>
<td>1.25%</td>
</tr>
</tbody>
</table>

Table 3: Sector classification for the test portfolios.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Issuers</th>
<th>Portfolio A</th>
<th>%</th>
<th>Issuers</th>
<th>Portfolio B</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>5</td>
<td></td>
<td>27.78%</td>
<td>12</td>
<td></td>
<td>15.00%</td>
</tr>
<tr>
<td>Consumer products</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>12</td>
<td></td>
<td>15.00%</td>
</tr>
<tr>
<td>Services</td>
<td>3</td>
<td></td>
<td>16.67%</td>
<td>19</td>
<td></td>
<td>23.75%</td>
</tr>
<tr>
<td>Financial</td>
<td>7</td>
<td></td>
<td>38.89%</td>
<td>25</td>
<td></td>
<td>31.25%</td>
</tr>
<tr>
<td>Industrial</td>
<td>-</td>
<td></td>
<td>0.00%</td>
<td>6</td>
<td></td>
<td>7.50%</td>
</tr>
<tr>
<td>Government</td>
<td>3</td>
<td></td>
<td>16.67%</td>
<td>6</td>
<td></td>
<td>7.50%</td>
</tr>
</tbody>
</table>

Issued by Germany, Italy, Netherlands, and Spain and 76 corporate bonds by issuers from the aforementioned countries. The sectors represented in Portfolios A and B are shown in Table 3. Both portfolios are assumed to be equally weighted with a total notional of €10 million.

4.3. Credit Stress Propagation Network. We use credit default swap data to construct the stress propagation network. The CDS raw data set consists of daily CDS liquid spreads for different maturities from 1 May 2014 to 31 March 2015 for Portfolio A and 1 July 2014 to 31 December 2015 for Portfolio B. These are averaged quotes from contributors rather than exercisable quotes. In addition, the data set also provides information on the names of the underlying reference entities, recovery rates, number of quote contributors, region, sector, average of the ratings from Standard & Poor’s, Moody’s, and Fitch Group of each entity, and currency of the quote. We use the normalized CDS spreads of entities for the 5-year tenor for our analysis. The CDS spreads time series of Russian issuers are illustrated in Figure 4.

4.4. Simulation Study. In order to generate portfolio loss distributions and derive the associated risk measures, we perform Monte Carlo simulations. This process entails generating joint realizations of the systematic and idiosyncratic risk factors and comparing the resulting critical variables with the corresponding default thresholds. By this comparison, we obtain the default indicator \( Y_i \) for each issuer and this enables us to calculate the overall portfolio loss for this trial. The only difference between the standard and the contagion model is that in the contagion model we first obtain the default indicators for the sovereigns, and their values determine which default thresholds are going to be

<table>
<thead>
<tr>
<th>CDS spreads of Russian entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil Transporting Jt Stk Co Transneft</td>
</tr>
<tr>
<td>Vnesheconombank</td>
</tr>
<tr>
<td>Bk of Moscow</td>
</tr>
<tr>
<td>City Moscow</td>
</tr>
<tr>
<td>JSC GAZPROM</td>
</tr>
<tr>
<td>JSC Gazprom Neft</td>
</tr>
<tr>
<td>Lukoil Co</td>
</tr>
<tr>
<td>Mobile Telesystems</td>
</tr>
<tr>
<td>MDM Bk open Jt Stk Co</td>
</tr>
<tr>
<td>Open Jt Stk Co ALROSA</td>
</tr>
<tr>
<td>OJSC Oil Co Rosneft</td>
</tr>
<tr>
<td>Jt Stk Co Russian Standard Bk</td>
</tr>
<tr>
<td>Russian Agric Bk</td>
</tr>
<tr>
<td>JSC Russian Railways</td>
</tr>
<tr>
<td>Russian Fedn</td>
</tr>
<tr>
<td>SBERBANK</td>
</tr>
<tr>
<td>OPEN Jt Stk Co VIMPEL Comms</td>
</tr>
<tr>
<td>JSC VTB Bk</td>
</tr>
</tbody>
</table>

Figure 4: Time series of CDS spreads of Russian issuers.
Table 4: Portfolio losses for the test portfolios and additional risk due to contagion.

(a) Panel 1: Portfolio A

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Loss standard model</th>
<th>Loss contagion model</th>
<th>Contagion impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>99%</td>
<td>1,115,153</td>
<td>1,162,329</td>
<td>47,176</td>
</tr>
<tr>
<td>99.5%</td>
<td>1,443,579</td>
<td>3,003,949</td>
<td>1,560,370</td>
</tr>
<tr>
<td>99.9%</td>
<td>2,258,857</td>
<td>4,968,393</td>
<td>2,709,536</td>
</tr>
<tr>
<td>99.99%</td>
<td>3,543,441</td>
<td>5,713,486</td>
<td>2,170,045</td>
</tr>
<tr>
<td>Average loss</td>
<td>71,807</td>
<td>71,691</td>
<td></td>
</tr>
</tbody>
</table>

(b) Panel 2: Portfolio B

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Loss standard model</th>
<th>Loss contagion model</th>
<th>Contagion impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>99%</td>
<td>373,013</td>
<td>379,929</td>
<td>6,915</td>
</tr>
<tr>
<td>99.5%</td>
<td>471,497</td>
<td>520,467</td>
<td>48,971</td>
</tr>
<tr>
<td>99.9%</td>
<td>775,773</td>
<td>1,009,426</td>
<td>233,653</td>
</tr>
<tr>
<td>99.99%</td>
<td>1,350,279</td>
<td>1,847,795</td>
<td>497,516</td>
</tr>
<tr>
<td>Average loss</td>
<td>44,850</td>
<td>44,872</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5: Time series for Russian Federation with local minima, local maxima, and modified ε-draw-ups.

4.5. Sensitivity Analysis. In the following, we present a series of sensitivity studies and discuss the results. To achieve a candid comparison, we choose to perform this analysis on the single-sovereign Portfolio A. We vary the ratings of sovereign and corporates, as well as the CountryRank parameter, to draw conclusions about their impact on the loss distribution and verify the model properties.

4.5.1. Sovereign Rating. We start by exploring the impact of the credit quality of the sovereign. Table 5 shows the quantiles of the generated loss distributions under the standard latent variable model and the contagion model when the rating of the Russian Federation is 1 and 2 notches higher than the original rating (BB). It can be seen that the contagion effect appears less strong when the sovereign rating is higher. At the 99.9% quantile, the contagion impact drops from 120% to 62% for an upgraded sovereign rating of BBB. The drop is even higher, when upgrading the sovereign rating to A, with only 11% additional losses due to contagion. Apart from having a less significant impact at the 99.9% quantile, it is clear that, with a sovereign rating of A, the contagion impact is zero at the 99% and 99.5% levels, where the results of the
Figure 6: Minimum spanning tree for Russian issuers.

Figure 7: Estimated systematic factor correlation matrix $\hat{\Omega}$. 
contagion model match those of the standard model. This is to be expected since a rating of A corresponds to a probability of default less than 0.01%, and as explained in Section 4.4, when sovereign default occurs seldom, the contagion effect can hardly be observed.

4.5.2. Corporate Default Probabilities. In the next test, the impact of corporate credit quality is investigated. As Table 6 illustrates, contagion has smaller impact when the corporate default probabilities are increased by 5%, which is in line with intuition since the autonomous (not sovereign induced) default probabilities are quite high, meaning that they are likely to default whether the corresponding sovereign defaults or not. For the same reason, the impact is even less significant when the corporate default probabilities are stressed by 10%.

4.5.3. CountryRank. In the last test, the sensitivity of the contagion impact to changes in the CountryRank is investigated. In Table 7, we test the contagion impact when CountryRank is stressed by 15% and 10%, respectively. The results are in line with intuition, with a milder contagion effect for lower CountryRank values and a stronger effect in case the parameter is increased.

5. Conclusions

In this paper, we present an extended factor model for portfolio credit risk which offers a breadth of possible applications to regulatory and economic capital calculations, as well as to the analysis of structured credit products. In the proposed framework, systematic risk factors are augmented
with an infectious default mechanism which affects the entire portfolio. Unlike models based on copulas with more extreme tail behavior, where the dependence structure of defaults is specified in advance, our model provides an intuitive approach, by first specifying the way sovereign defaults may affect the default probabilities of corporate issuers and then deriving the joint default distribution. The impact of sovereign defaults is quantified using a credit stress propagation network constructed from real data. Under this framework, we generate loss distributions for synthetic test portfolios and show that the contagion effect may have a profound impact on the upper tails.

Our model provides a first step towards incorporating network effects in portfolio credit risk models. The model can be extended in a number of ways such as accounting for stress propagation from a sovereign to corporates even without sovereign default or taking into consideration contagion between sovereigns. Another interesting topic for future research is characterizing the joint default distribution of issuers in credit stress propagation networks using Bayesian network methodologies, which may facilitate an improved approximation of the conditional default probabilities in comparison to the maximum weight path in the current definition of CountryRank. Finally, a conjecture worthy of further investigation is that a more connected structure for the credit stress propagation network leads to increased values for the CountryRank parameter, and, as a result, to higher additional losses due to contagion.
Appendix

A. CDS Spread Data

The data used to calibrate the credit stress propagation network for European issuers is the CDS spread data of Dutch, German, Italian, and Spanish issuers as shown in Figures 9 and 10.

B. Stability of $\epsilon$-Parameter

The plot in Figure 11 shows the time series of the epsilon parameter for different number days used for $\epsilon$-draw-up calibration.
Disclosure

The opinions expressed in this work are solely those of the authors and do not represent in any way those of their current and past employers.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors are thankful to Erik Vynckier, Grigorios Papamanousakis, Sotirios Sabanis, Markus Hofer, and Shashi Jain for their valuable feedback on early results of this work. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement no. 675044 (http://bigdatafinance.eu/), Training for Big Data in Financial Research and Risk Management.

References


Contagious defaults in a credit portfolio: a Bayesian network approach

Ioannis Anagnostou∗1,2, Javier Sánchez Rivero2, Sumit Sourabh1,2, and Drona Kandhai1,2

1Computational Science Lab, University of Amsterdam, Science Park 904, Amsterdam 1098XH, The Netherlands
2Quantitative Analytics, ING Bank, Foppingadreef 7, Amsterdam 1102BD, The Netherlands

November 5, 2018

Abstract
Robustness of credit portfolio models is of great interest for financial institutions and regulators, since misspecified models translate to insufficient capital buffers and a crisis-prone financial system. In this paper, we propose a method to enhance credit portfolio models based on the model of Merton by incorporating contagion effects. While in most models the risks related to financial interconnectedness are neglected, we use Bayesian network methods to uncover the direct and indirect relationships between credits, while maintaining the convenient representation of factor models. A range of techniques to learn the structure and parameters of financial networks from real Credit Default Swaps (CDS) data is studied and evaluated. Our approach is demonstrated in detail in a stylized portfolio and the impact on standard risk metrics is estimated.

Keywords – Portfolio Credit Risk; Bayesian Learning; Credit Default Swaps; Default Contagion; Probabilistic Graphical Models; Network Theory

1 Introduction
In recent years, there has been an increasing interest in modelling credit risk by practitioners as well as academics (see e.g., Gregory, 2015, Green, Kenyon, & Dennis, 2014, Sourabh, Hofer, & Kandhai, 2018, De Graaf, Feng, Kandhai, & Oosterlee, 2014, de Graaf, Kandhai, & Reisinger, 2018, Simaitis, de Graaf, Hari, & Kandhai, 2016). Portfolio credit risk models are concerned with the occurrence of large losses due to defaults or deteriorations in credit quality. In practice, these models have a wide range of applications, such as regulatory and economic capital measurements, portfolio management, and risk-adjusted pricing. The robustness of such models is of great interest both for financial institutions and regulators, since misspecified models could lead to insufficient capital buffers, which in turn would result in a crisis-prone financial system and the need for regular bail-outs.

The key challenge in portfolio credit risk modelling is the incorporation of default dependence. Joint defaults of many issuers to which a portfolio is exposed to may lead to extreme losses. Therefore, understanding the relationship between default events is crucial. In most portfolio credit risk models existing in the literature, defaults of individual issuers depend on a set of common underlying risk factors, describing the state of a sector, region, or the economy as a whole. Notable examples of this approach are the Asymptotic Single Risk Factor (ASRF) model (Gordy, 2003) in the Basel regulations and industrial adaptations of Merton model (Merton, 1974) such as the CreditMetrics (JP Morgan, 1997) and KMV models (Bohn & Kealhofer, 2001; Crosbie & Bohn, 2002).

∗i.anagnostou@uva.nl
Many researchers have challenged the claim that default dependence can be fully explained by
dependence on common underlying factors, on the grounds that such models often fail to capture
default clustering that does occur from time to time. Schönbucher & Schubert, 2001 suggest that
in most cases the default correlations that can be achieved with common factors are not as high as
the ones in empirical data. Das, Duffie, Kapadia, & Saita, 2007 perform statistical tests and reject
the hypothesis that factor correlations can sufficiently explain the empirically observed default
correlations. Thus, it becomes evident that an additional channel of default dependence needs to
be considered.

Besides dependence on common factors, joint defaults might occur as a result of direct links
between issuers, a phenomenon known as contagion. Davis & Lo, 2001 were among the first ones
who introduced contagion in credit risk models, by considering that any default might infect to
any other issuer in the portfolio. Jarrow & Yu, 2001 tried to generalize already existing models,
included particular specifications of the issuers and focused on their effect on bonds and credit
derivatives. Egloff, Leippold, & Vanini, 2007 introduced network theory to allow for a variety of
infections, however the model required detailed information making its application more difficult
than expected.

Following the financial crisis, there has been a significant interest in using network-based meth-
ods for financial stability and systemic risk (see e.g., Battiston, Gatti, Gallegati, Greenwald,
& Stiglitz, 2012, Cont, Moussa, & Santos, 2013, Squartini, Van Lelyveld, & Garlaschelli, 2013,
Battiston et al., 2016, Poledna, Thurner, Farmer, & Geanakoplos, 2014, Musmeci,Nicosia, Aste,
Di Matteo, & Latora, 2017). Nevertheless, the use of these methods for valuation and measurement
of risk charges such as capital is limited. In a recent study, Anagnostou, Sourabh, & Kandhai, 2018
introduced a portfolio credit risk model that can account for both channels of default dependence:
common underlying factors and contagion from sovereigns to corporates and sub-sovereigns. The
authors augment systematic risk factors with a contagious default mechanism where the default
probabilities of issuers in the portfolio are immediately affected by a sovereign default. To estimate
the contagion effect they use a network constructed from CDS time series and introduce Coun-
tryRank, a network based metric that approximates the probability of default of a node conditional
on the infectious default of a sovereign. The article presents a thorough approach of how conta-
gion effects can be introduced to portfolio credit risk models using complex networks. However,
the underlying network in Anagnostou et al., 2018 is based on one-to-one relationships between
issuers. While this can capture the direct relationships effectively, it is well-known that the associ-
ations between entities might be indirect and often mediated through others. In this article we use
Probabilistic Graphical Models (PGM) to learn the network using the data in a holistic manner.
This extends the one-to-one approach and provides a more natural and accurate representation
of the network. Moreover, we can efficiently estimate the joint default probability distribution of the
issuers in a PGM.

PGMs are a powerful framework for representing complex relationships using probability distrib-
utions. Their ability to model associations in complex datasets has proven them particularly useful
for a wide range of machine learning problems, including natural language processing (Galley, McK-
eown, Hirschberg, & Shriberg, 2004), medical diagnosis (Beinlich, Suermondt, Chavez, & Cooper,
1989), and genetic linkage analysis (Fishelson & Geiger, 2004). One of the most important classes
of PGMs is Bayesian networks (BNs). More recently, there have been attempts to utilize BNs for
solving financial problems. In particular, Denev, 2013 presented a method to calculate portfolio
losses in the presence of stress events using BNs. Nevertheless, his approach, relies on the ability
of the risk manager to identify causal links and subjectively assign probabilities. Chong & Klüp-
pelberg, 2018 developed a structural default model for interconnected financial institutions, but
the need for balance sheet data makes its applicability limited. Kitwiwattanachai, 2015 used credit
default swaps (CDS) data to learn the structure of interbank networks, which would enable policy
makers to make decisions on the too-big-to-fail problem. In order to learn the BN, the author uses
the log of CDS spreads under the assumption of normality. However, this strong assumption is
barely supported by empirical evidence.

In this paper, we overcome the need for making assumptions about the distribution of CDS
spreads by introducing a discretization method based on the notion of modified $\epsilon$-drawups (Kaushik
& Battiston, 2013; Anagnostou et al., 2018). This transformation enables us to utilize algorithms
for structure and parameter learning that assume discrete random variables, without having to
sacrifice the interpretability of the resulting models. We use the discretized CDS time series
to learn a BN of interactions between issuers, and to estimate the contagion effects following a
sovereign default. Different techniques to learn the structure and parameters of financial networks are studied and evaluated, with the results confirming that the structures are robust. In order to investigate the impact of these effects on credit losses, we carry out simulations and calculate the quantiles of the loss distribution in the presence of contagion. Finally, we perform a comparative analysis with the results obtained by Anagnostou et al., 2018.

The rest of the article is organized as follows. Section 2 presents BNs and outlines the methods for learning their structure and parameters. Section 3 demonstrates a method to learn BNs for CDS data. Section 4 gives a brief description of factor models for portfolio credit risk, along with a model for credit contagion. In Section 5, we present empirical analysis on a synthetic test portfolio. Finally, in Section 6 we summarize and draw conclusions. Additional information is included in the Appendices.

2 Bayesian networks

In this section we briefly introduce Bayesian networks (BNs). For a detailed background, we refer to Koller & Friedman, 2009. In the rest of the article, we only consider directed graphs \((V, E)\), where \(V\) is a finite vertex set and \(E \subseteq \{(i, j) : i, j \in V, i \neq j\}\) is a set of edges without any self-loops. Moreover, we further assume that the graph has no directed cycles.

Definition 1. For a directed cyclic graph \(G = (V, E)\), a collection \(\{X_i : i \in V\}\) of random variables forms a Bayesian network over \(G\) if

\[
P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | \text{Pa}_{X_i})
\]

where \(\text{Pa}_{X_i}\) is the set of parents of the node \(X_i\) and \(|V| = n\).

BNs have the useful property that they allow us to represent a joint distribution in a tractable manner. For example, even if the random variables \(X_1, \ldots, X_n\) follow a binomial distribution, we would need \(2^n - 1\) probabilities to represent their joint distribution. With a BN, the order of representation of joint distribution is linear in the number of variables.

In order to demonstrate this representation, we provide a commonly used example from (Jensen & Nielsen, 2007), illustrated in Figure 1. We consider that the grass can appear wet in the morning either because the sprinkler was on during the night or because it rained. Note that these events are not mutually exclusive: it is possible that the sprinkler was on and it rained at the same time. Thus we have two binary-valued random variables, \textit{Sprinkler} (S) and \textit{Rain} (R). We also have two more binary-valued variables, \textit{Cloudy} (C) and \textit{Wet} (W), which are correlated to both \textit{Sprinkler} and
Rain. The strength of these relationships is shown in the Conditional Probability Tables (CPTs). For example, we see that \( P(W|S,R) = 0.9 \), and thus, \( P(\neg W|S,R) = 1 - 0.9 = 0.1 \). Since the C node has no parents, its CPT specifies the prior probability that it is cloudy, which in this case is equal to 0.5. Overall, our probability space has \( 2^4 = 16 \) values which correspond to all the possible assignments of these four variables. By the chain rule of probability, the joint probability of all the nodes in the graph is:

\[
P(C, S, R, W) = P(C)P(S|C)P(R|C, S)P(W|C, S, R)
\]  

(2)

By using conditional independence relationships, this can be rewritten as:

\[
P(C, S, R, W) = P(C)P(S|C)P(R|C)P(W|S, R)
\]  

(3)

where it was possible to simplify the third term because \( \text{Rain} \) is independent of \( \text{Sprinkler} \) given its parent \( \text{Cloudy} \), and the last term because \( \text{Wet} \) is independent of \( \text{Cloudy} \) given its parents \( \text{Season} \) and \( \text{Rain} \). Thus, it is clear that the conditional independence relationships allow for a more compact representation of the joint probability distribution.

In this work the nodes in BNs represent random variables characterizing issuers of debt and the edges how these issuers influence each other. More specifically, the random variables of interest are the probabilities of default. To learn the structure and parameters of these BNs we use time series of CDS spreads. The rest of this section describes the process of learning the structure and parameters of a BN from data.

2.1 Learning

In order to estimate the joint probability distribution from a BN, we first need to learn both the structure and the parameters of the network from data. In the first place we will explain the parameter learning and afterwards the structure learning. The reason for this order is the use of some parameter estimation techniques in the latter.

Parameter learning Suppose we have a collection of \( n \) random variables \( X_1, \ldots, X_n \) such that the number of states of the random variable \( X_i \) is \( 1, 2, \ldots, r_i \) and the number of configurations of parents of \( X_i \) is \( 1, 2, \ldots, q_i \). The parameters which have to be estimated in this case are:

\[
\theta_{ijk} = P(X_i = j | \text{Pa}(X_i) = k), \quad i \in \{1, \ldots, n\}, j \in \{1, \ldots, r_i\} \text{ and } k \in \{1, \ldots, q_i\}.
\]

We use \( \Theta = \{\theta_{ijk} | i \in \{1, \ldots, n\}, j \in \{1, \ldots, r_i\}, k \in \{1, \ldots, q_i\}\} \) to denote the parameter vector.

Let us assume that we know the structure of a BN. There are two different methods for learning the parameters: Maximum Likelihood Estimation (MLE) and Bayesian estimation. MLE is based on choosing the parameters which maximize the likelihood of the data. Given a data set \( D \), the MLE method chooses parameters \( \hat{\theta}_D \) such that:

\[
\hat{\theta}_D = \arg \max_{\theta_D \in \Theta_D} \mathcal{L}(\theta_D : D) = \arg \max_{\theta_D \in \Theta_D} P(D : \theta_D) = \arg \max_{\theta_D \in \Theta_D} \prod_m P(\xi[m] : \theta_D)
\]

where \( \Theta_D \) is the space of possible values of \( \theta_D \), \( \xi[m] \) is the \( m \)-th instance of \( D \).

The alternative Bayesian estimation method is based on assuming a prior distribution over the parameters \( P(\theta_D) \), and updating it with each instance of the data to obtain the posterior distribution, \( P(\theta_D | D) \) using the Bayes’ rule as follows:

\[
P(\theta_D | \xi[1], \ldots, \xi[M]) = \frac{P(\xi[1], \ldots, \xi[M] : \theta_D) P(\theta_D)}{P(\xi[1], \ldots, \xi[M])}
\]  

(4)

where the denominator is a normalizing factor and \( P(\xi[1], \ldots, \xi[M] : \theta_D) \) is the likelihood.

The choice of the prior distribution is key for the Bayesian estimation procedure. From Equation 4, we can see that the posterior distribution is proportional to the product of the likelihood and the prior. Therefore, we need to choose the prior in such a way that it can be updated easily after each new sample, while maintaining the form of the posterior distribution. It is well-known that the Dirichlet distribution is the conjugate prior for the multinomial distribution (Koller & Friedman, 2009), which means that if the prior distribution of the multinomial parameters is Dirichlet then the posterior distribution is also a Dirichlet distribution. Since we deal with multinomial variables in our case, we choose Dirichlet distribution as the prior for our experiments.
**Structure learning** The structure learning for a BN is essentially an optimization problem where we minimize a score over the search space of possible configurations of the network. The score measures how likely a particular structure is based on the data, and is divided into two categories: likelihood scores and Bayesian scores.

The likelihood scores rely mainly on the likelihood function, which is the probability of sampling the data given the structure, \( L(\mathcal{G} | \mathcal{D}) = P(\mathcal{D} | \mathcal{G}) \). The notation \( \langle \mathcal{G}, \theta \rangle \) denotes a BN, where \( \mathcal{G} \) represents the structure and \( \theta \) the parameters of the network. The structure of the network is chosen so as to maximize the likelihood score, using the MLE parameters.

\[
\max_{\mathcal{G}, \theta} L(\langle \mathcal{G}, \theta \rangle : \mathcal{D}) = \max_{\mathcal{G}} \left[ \max_{\theta} L(\langle \mathcal{G}, \theta \rangle : \mathcal{D}) \right] = \max_{\mathcal{G}} L(\langle \mathcal{G}, \hat{\theta} \rangle : \mathcal{D})
\]

The Bayesian scores have a similar approach as the Bayesian estimation for the parameters. First, we define a prior distribution over the structure \( P(\mathcal{G}) \) and a conditional prior over the parameters \( P(\theta | \mathcal{G}) \). Then, we obtain the posterior distribution \( P(\mathcal{G} | \mathcal{D}) \) using the Bayes’ rule. Similar to the Bayesian estimation for the parameters, the denominator is a normalizing factor and it remains the same for all the structures. Then, the score can be defined by taking the logarithm of the numerator:

\[
\text{score}_B(\mathcal{G} : \mathcal{D}) = \log P(\mathcal{D} | \mathcal{G}) + \log P(\mathcal{G})
\]

where the second term, the prior, makes no significant difference in the score (see, e.g. Koller & Friedman, 2009). The first term, called *marginal likelihood*, is the average over all the possible choices of \( \theta \) based on the conditional probability provided before:

\[
P(\mathcal{D} | \mathcal{G}) = \int_{\theta} P(\mathcal{D} | \theta, \mathcal{G}) P(\theta | \mathcal{G}) d\theta \tag{5}
\]

The average over all the possible parameters makes the model more conservative, and hence it tries to avoid over-fitting as we take into account the sensitivity to the values of the parameters. Finally, we briefly describe the search space and the optimization procedure for structure learning. The search space is a network itself where each of the nodes is a candidate structure \( \mathcal{G} \). Given a node \( \mathcal{G} \) corresponding to a structure in the search space network, the edges of \( \mathcal{G} \) are structures obtained by either adding an edge, deleting an edge or reversing an edge. The Hill-Climbing algorithm follows the steps below:

1. Set initial BN structure to a network without edges \( \mathcal{G} \)
2. Compute its score
3. Consider all the neighbours of \( \mathcal{G} \), obtained by either adding, deleting or reversing an edge in \( \mathcal{G} \)
4. Choose the neighbour \( \mathcal{G}_{\text{best}} \) which leads to the best improvement in the score
5. Set \( \mathcal{G} \leftarrow \mathcal{G}_{\text{best}} \) and repeat until no further improvement to the score is possible.

Some improvements can be made to this algorithm: (Koller & Friedman, 2009, Glover, 1995).

### 3 Learning Bayesian networks from CDS data

In order to learn the structure and parameters of a BN, a key requirement is that the input data should be either a discrete or a Gaussian distribution (Koller & Friedman, 2009). In Kitwiwattanachai, 2015, the authors made the assumption that the residuals of the regressions on the log returns of the CDS spreads are normally distributed which is not supported by empirical data. Therefore, we transform the continuous CDS time series data into a discrete distribution. For details on CDS contracts, the reader is referred to Section A in the Appendix.
3.1 CDS dataset

The data used for the construction of the network are credit default swaps (CDS) spreads for different maturities obtained from Markit. These consist of daily CDS liquid spreads of Russian issuers from September 14th 2010 until August 15th 2015. Apart from the spreads, the dataset also includes information about the recovery rates, region, sector and average of the ratings from Standard & Poor’s, Moody’s, and Fitch Group of each issuer. We use the CDS spreads of issuers for the 5 year tenor for our analysis, since they are the most liquid quotes. Since recovery rates are not the same for all issuers, we have to normalize CDS spreads to do a consistent analysis. The normalization of CDS spreads for the recovery rate $RR$ is done as follows:

$$\hat{S} = S \frac{0.6}{(1 - RR)},$$

where $\hat{S}$ denotes the normalized CDS spread, which corresponds to a recovery rate of 40%. The choice of normalizing the CDS spreads using a recovery rate of 40% is based on the literature (see e.g., Das & Hanouna, 2009) where the recovery rate is often assumed to be a constant. We use the normalized CDS spread of the five year tenor in our analysis. Figure 2 shows the normalized spreads for the Russian issuers in the portfolio.

3.2 Discretization of CDS data

We use the notion of modified $\epsilon$-drawups to transform the continuous CDS time series data into a discrete distribution. Modified $\epsilon$-drawups build up on the notion of $\epsilon$-drawups (Kaushik & Battiston, 2013, Sornette & Zhou, 2006) and can detect instances in the time series where it shows significant upward jumps. Note that, in the context of CDS spreads, upward jumps translate to rapid deteriorations in credit quality.

**Definition 2.** A modified $\epsilon$-drawup is defined as an upward movement in the time series at a local minima, in which the amplitude of the movement, that is, the difference between the subsequent local maxima and the local minima, is greater than a threshold $\epsilon$. We record such a local minima in the time series as a modified $\epsilon$-drawup.

An illustration of the above definition is shown in Figure 3. The $\epsilon$ parameter at time $t$ is set to be the standard deviation in the time series between days $t - n$ and $t$, where $n$ is chosen to be 10 days consistent with the choice in Kaushik & Battiston, 2013. The CDS data can be transformed into a discrete distribution for learning the BN as follows. Firstly, we compute the modified $\epsilon$-drawups for each of the time series. Thus, each issuer $i$ will either have a modified $\epsilon$-drawup at time $t$ or not. We define a binary random variable $X_i^t$ corresponding to the issuer
A modified $\epsilon$-drawup is defined as an upward movement in the time series at a local minima, in which the amplitude of the movement is greater than a threshold $\epsilon$. The $\epsilon$ parameter at time $t$ is set to be the standard deviation in the time series between days $t - n$ and $t$, where $n$ is chosen to be 10 days. In the above example, the first minimum is recorded as a modified $\epsilon$-drawup, while the second is not.

$i$ such that $X_i^t = 1$, if issuer $i$ has a modified $\epsilon$-drawup on day $t$, and 0, otherwise. Note that a company cannot have two modified $\epsilon$-drawups on consecutive days by definition.

An additional step which is needed to prepare the CDS data for learning the co-dependence of defaults is to introduce the concept of time-lag. This allows us to capture the fact that issuers can impact each other with a slight delay. For time-lag, we introduce a new categorical value, 0.5, in the following way. Let a company $i$ have a modified $\epsilon$-drawup on day $t$, so $X_i^t = 1$. If a different company $j$ has a modified $\epsilon$-drawup on at least one of the following three days, $t + 1$, $t + 2$, or $t + 3$, and not on day $t$, then we set $X_j^t = 0.5$. Hence, the number of modified $\epsilon$-drawups in the time series remains unchanged which ensures that the marginal probability of having a modified $\epsilon$-drawup remains unchanged.

### 3.3 Bayesian network learning

For learning the network we use the hill-climbing algorithm based on two different scores: Bayesian Information Criterion (BIC) (Schwarz, 1978) which is a likelihood score, and Bayesian Dirichlet Sparse (BDs) (Scutari, 2016), which is a Bayesian score. We refer to Section C in Appendix for details on the two scores.

For learning the network, we also applied a bootstrapping technique for ensuring the robustness of results (Friedman, Goldszmidt, & Wyner, 1999). For the structure learning, we obtain the structure (1000) times and then we compute the average structure by including the edges which appear in at least 50% of the networks. The data set $D_k$ used at iteration $k$ is obtained by sampling uniformly $|D|$ instances from the original training data $D$.

Once the BN is learnt, we can evaluate the queries for conditional probabilities $P(Q|E)$, of events $Q$ given evidence $^1E$. We use logic sampling algorithm to perform these queries which uses the following steps. First, it orders the variables in the topological order implied by the structure $G$. This means that the variables with no parents appear first followed by their children. Next, we set the counters $n_E = 0$ and $n_{E,Q} = 0$. Afterwards we generate a sufficiently large number of samples $M$ where each sample consists of a vector of instances of all the random variables in the network. Note that generating the instance for $X_i$ only requires the values of $Pa_{X_i}$. Then, for each sample if it includes $E$, set $n_E = n_E + 1$, and, if it includes both $Q$ and $E$, set $n_{Q,E} = n_{Q,E} + 1$. Finally, we can estimate $P(Q|E)$ by $n_{Q,E}/n_E$. This method is based on a Monte Carlo simulation, therefore a sufficiently large number of simulations is needed to assure a reliable result.

---

1 An event in the BN terminology refers to a (some) random variable(s) taking a particular value(s). An evidence is mathematically the same as an event with the difference that it is known.
4 Portfolio credit risk modelling

The most widely used portfolio credit risk models assume that issuers depend on some common underlying factors. Factor models based on the Merton model are particularly popular for portfolio credit risk. The model presented in Anagnostou et al., 2018 extends the multi-factor Merton model to allow for credit contagion. In this section we provide a brief description of factor models for portfolio credit risk, along with an overview of the contagion model from Anagnostou et al., 2018.

4.1 Factor models

Factor models for portfolio credit risk can be motivated by a multivariate firm-value model based on Merton, 1974. This category includes widely used industry models such as CreditMetrics and KMV. Default occurs for an issuer $i$ if a critical variable $X_i$, representing the standardized asset return, falls below a critical threshold $d_i$. For a portfolio of $m$ issuers, $(X_1, ..., X_m)' \sim N_m(0, \Sigma)$ and thus, the probability of default for issuer $i$ is satisfying $p_i = \Phi(d_i)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The default probabilities are usually estimated by historical default experience using external ratings by agencies or model-based approaches.

In the factor model approach, the critical variables $X_i, i = 1, ..., m$ are linearly dependent on a vector $F$ of $p < m$ common underlying factors satisfying $F \sim N_p(0, \Omega)$. Issuer $i$’s standardized asset return is assumed to be driven by a issuer-specific combination $\tilde{F}_i = \alpha_i' \bar{F}$ of the systematic factors

$$X_i = \sqrt{\beta_i} \tilde{F}_i + \sqrt{1-\beta_i} \epsilon_i,$$

where $\tilde{F}_i$ and $\epsilon_1, ..., \epsilon_m$ are independent standard normal variables, and $\epsilon_i$ represents the idiosyncratic risk. Consequently, $\beta_i$ can be seen as a measure of sensitivity of $X_i$ to systematic risk, as it represents the proportion of the $X_i$ variation that is explained by the systematic factors. The correlations between asset returns are given by

$$\rho(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{\text{var}(X_i) \text{var}(X_j)}} = \beta_i \beta_j \alpha_i' \Omega \alpha_j,$$

where $\tilde{F}_i$ and $\epsilon_1, ..., \epsilon_m$ are independent and standard normal and $\text{var}(X_i) = 1$.

4.2 A model for credit contagion

Consider a corporate issuer $C_i$, and its country of operation $S$. Denote by $p_{C_i}$ the probability of default of $C_i$. Under the standard Merton model, default occurs if $C_i$’s standardized asset return $X_{C_i}$ falls below its default threshold $d_{C_i}$. The critical threshold $d_{C_i}$ is assumed to be equal to $\Phi^{-1}(p_{C_i})$ and is independent of the state of the country of operation $S$. In the proposed model, a corporate is subject to shocks from its country of operation; its corresponding state is described by a binary state variable. The state is considered to be stressed in the event of sovereign default. In this case, the issuer’s default threshold increases, causing it more likely to default, as the contagion effect suggests. In case the corresponding sovereign does not default, the corporates liquidity state is considered stable. We replace the default threshold $d_{C_i}$ with $d_{C_i}^*,$ where

$$d_{C_i}^* = \begin{cases} d_{C_i}^{sd} & \text{if the corresponding sovereign defaults} \\ d_{C_i}^{nsd} & \text{otherwise} \end{cases}$$

or equivalently

$$d_{C_i}^* = \mathbb{I}(Y_S=1) d_{C_i}^{sd} + \mathbb{I}(Y_S=0) d_{C_i}^{nsd}$$

We denote by $p_S$ the probability of default of the country of operation, and by $\gamma_{C_i}$ the parameter which indicates the increased probability of default of $C_i$ given the default of $S$. Our objective is to calibrate $d_{C_i}^{sd}$ and $d_{C_i}^{nsd}$ in such way that the overall default rate remains unchanged and $P(Y_{C_i} = 1 | Y_S = 1) = \gamma_{C_i}$. Denote by

$$\phi_2(x; y; \rho) := \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left( -\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)} \right),$$

$$\Phi_2(h; k; \rho) := \int_{-\infty}^{h} \int_{-\infty}^{k} \phi_2(x; y; \rho) dy dx$$

(12)
the density and distribution function of the bivariate standard normal distribution with correlation
parameter $\rho \in (-1, 1)$. Note that $d^{sd}_{C_i} (\omega) = d^{sd}_{C_i}$ for $\omega \in \{Y_{C_i} = 1, Y_S = 1\} \subset \{Y_S = 1\}$, and $d^{sd}_{C_i} (\omega) = d^{sd}_{C_i}$ for $\omega \in \{Y_{C_i} = 1, Y_S = 0\} \subset \{Y_S = 0\}$. We rewrite $P(Y_{C_i} = 1 | Y_S = 1)$ in the
following way

$$P(Y_{C_i} = 1 | Y_S = 1) = \frac{1}{P(Y_S = 1)} P(Y_{C_i} = 1, Y_S = 1)$$

$$= \frac{1}{p_S} P[X_{C_i} < d^{sd}_{C_i}, X_S < d_S]$$

$$= \frac{1}{p_S} \Phi_2(d^{sd}_{C_i}, d_S; \rho_{SC_i})$$

Using the above representation and given $d_S = \Phi^{-1} (p_S)$ and $\rho_{SC_i}$, one can solve the equation

$$P(Y_{C_i} = 1 | Y_S = 1) = \gamma_{C_i}$$

over $d^{sd}_{C_i}$.

We proceed to the derivation of $d^{sd}_{C_i}$ in such way that the overall default probability remains
equal to $p_{C_i}$. This constraint is important, since contagion is assumed to have no impact on the
average loss. Clearly,

$$p_{C_i} = P(Y_{C_i} = 1)$$

$$= P(Y_{C_i} = 1, Y_S = 1) + P(Y_{C_i} = 1, Y_S = 0)$$

$$= P(Y_{C_i} = 1 | Y_S = 1)P(Y_S = 1) + P(Y_{C_i} = 1, Y_S = 0)$$

and thus

$$P(Y_{C_i} = 1, Y_S = 0) = p_{C_i} - \gamma_{C_i}, p_S$$

The left-hand side of the above equation can be represented as follows

$$P(\text{corp.def} \cap \text{nosov.def}) = P[X_{C_i} < d^{sd}_{C_i}, X_S > d_S]$$

$$= P[X_{C_i} < d^{sd}_{C_i}] - P[X_{C_i} < d^{sd}_{C_i}, X_S < d_S]$$

$$= \Phi(d^{sd}_{C_i}) - \Phi_2(d^{sd}_{C_i}, d_S; \rho_{SC_i})$$

By use of the above and given $d_S = \Phi^{-1} (p_S)$ and $\rho_{SC_i}$, one can solve the previous equation over $d^{sd}_{C_i}$.

5 Numerical study

In this section we describe the setup of the problem and the results obtained in the calibrations
and simulations.

5.1 Bayesian network learning and robustness

We use CDS data as described in Section 3 to learn the structure and the parameters of the BN.
The bnlearn (Scutari, 2010) library in R is used for the Hill-Climbing procedure. The numerical
experiments were performed using two scores: BDs and BIC. Both scores resulted in similar structure
of the network as shown in Figure 5, except for one edge, GAZPRU.Gneft → AKT which is
present in the network learnt with BIC, whereas with BDs it is substituted by CITMOS → AKT,
as we can see in Figure. The table in Appendix A, lists the tickers of the issuers and the complete
names.

After learning the network, the next step is to estimate the conditional probability of default of
an issuer, conditional on the default of the sovereign. In order to estimate these probabilities we run
a Monte Carlo simulation of 100 iterations. For each of these iterations, we calculate the conditional
probability using $4 \times 10^5$ samples. Finally, we take the mean of the Monte Carlo simulations as
an estimate for conditional probabilities. We compare the probabilities for both BIC and BDs
scores including mean, standard deviation and absolute difference in Table 3. Figure 4 shows the
structure with the nodes colored according to their probability of default given sovereign default.
Figure 5: Structure obtained with the different scores. Note that the visualization system is mirroring the plot but close inspection reveals that the structure is not that different.
to have a more visual explanation. The darker the color of the node, the higher is the probability of default of the node conditional on sovereign default. We see that Gazprom and Gazprom Neft, which are the two nodes connected to the sovereign, are the ones more affected and the issuers which are further from the sovereign have relatively lower conditional default probabilities.

Figure 4: Bayesian network learnt with BIC with colored nodes according to its probability of default given sovereign default. Note that having a darker color means that the node has a higher conditional probability of default.

We observe that the standard deviation of the conditional probabilities estimates is quite small for all the issuers, ranging between 0.002 and 0.0035. This implies that the estimates are quite robust, and gives us a strong confidence on the reliability of the results. Moreover, we notice that the absolute difference of the probabilities from the two scores is smaller than the standard deviation except in two cases.

The main difference we note in the table is Transneft, whose probability of default conditional on sovereign default decreases by more than 0.07 when changing the score from BIC to BDs. This is caused by the change in the structure, which directly affects this issuer. With BIC score, Gazprom Neft is a parent of Transneft whereas City Moscow is not, and using BDs score it is the other way around. We see that with both scores Gazprom Neft is more affected by the sovereign than City Moscow. This is also a confirmation of the fact that the stress spreads faster from one issuer to another if they are directly connected. The second difference in conditional probability is for MDM Bank. If we look at the structures, we see that MDM has only one parent, which is Transneft. As the parent (Transneft) is less affected in the structure learnt by BDs, its child being less affected is in line with intuition. This causes a small but still noticeable difference in the conditional probabilities.
5.2 Comparative analysis

We set up a multi-factor Merton model, as it was described in Section 4. We define a set of systematic factors that will represent region and sector effects. We choose 6 region and 6 sector factors, for which we select appropriate indices, as shown in Table 6. We then use 10 years of index time series to derive the region and sector returns \( F_{R(j)}, j = 1, ..., 6 \) and \( F_{S(k)}, k = 1, ..., 6 \) respectively, and obtain an estimate of the correlation matrix \( \Omega \). Subsequently, we map all issuers to one region and one sector factor, \( F_{R(i)} \) and \( F_{S(i)} \) respectively. For instance, a Dutch bank will be associated with Europe and Financial factors. As a proxy of individual asset returns we use 10 years of equity or CDS time series, depending on the data availability for each issuer. Finally, we standardize the individual returns time series \( X_{i,t} \) and perform the following Ordinary Least Squares regression against the systematic factor returns

\[
X_{i,t} = \alpha_{R(i)} F_{R(i),t} + \alpha_{S(k)} F_{S(k),t} + \epsilon_{i,t}
\]

(24)
to obtain \( \hat{\alpha}_{R(i)} \), \( \hat{\alpha}_{S(i)} \), and \( \hat{\beta}_i = R^2 \), where \( R^2 \) is the coefficient of determination, and it is higher for issuers whose returns are largely affected by the performance of the systematic factors.

To investigate the properties of the contagion model, we set up a test portfolio. The resulting risk measures for this portfolio are compared to those of the standard latent variable model with no contagion. The portfolio consists of 1 Russian government bond and 17 bonds issued by corporations registered and operating in the Russian Federation. As it is illustrated in Table 1, the issuers are of medium and low credit quality. The sectors represented are shown in Appendix B, Table 2. The portfolio is assumed to be equally weighted with a total notional of \( €10 \) million.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Issuers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>1</td>
<td>5.56%</td>
</tr>
<tr>
<td>BB</td>
<td>15</td>
<td>83.33%</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>11.11%</td>
</tr>
</tbody>
</table>

Table 1: Rating classification for the test portfolio.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Issuers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>5</td>
<td>27.78%</td>
</tr>
<tr>
<td>Services</td>
<td>3</td>
<td>16.67%</td>
</tr>
<tr>
<td>Financial</td>
<td>7</td>
<td>38.89%</td>
</tr>
<tr>
<td>Government</td>
<td>3</td>
<td>16.67%</td>
</tr>
</tbody>
</table>

Table 2: Sector classification for the test portfolio.

In order to generate portfolio loss distributions and derive the associated risk measures we perform Monte Carlo simulations. This process entails generating joint realizations of the systematic and idiosyncratic risk factors, and comparing the resulting critical variables with the corresponding default thresholds. By this comparison we obtain the default indicator \( Y_i \) for each issuer and this enables us to calculate the overall portfolio loss for this trial. The only difference between the standard and the contagion model is that in the contagion model we first obtain the default indicators for the sovereigns, and their values determine which default thresholds are going to be used for the corporate issuers. A liquidity horizon of 1 year is assumed throughout and the figures are based on a simulation with \( 10^6 \) samples. Moreover, for the results shown we used the probabilities of default computed with the structure learnt with the score BIC. However, this choice does not make any notable difference in the quantiles of the loss distribution because the probabilities were almost the same and such tiny difference would not cause a large disturbance.

We compared the BN model with the CountryRank model of Anagnostou et al., 2018. In Table 4 we can observe the difference between the probabilities obtained with both methods, using the same data, and same parameters, 10 days to compute the standard deviation and 3 days as time lag. Following the sensitivity analysis by Anagnostou et al., 2018 we can expect that the 15% increase of the mean will not have a substantial impact on the quantiles of the loss distribution. This hypothesis can be confirmed by the results shown in Table 5 and the graph depicted in Figure 6.
<table>
<thead>
<tr>
<th>Order</th>
<th>Issuer</th>
<th>BIC</th>
<th>s.d.</th>
<th>BDs</th>
<th>s.d.</th>
<th>Abs diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JSC Gazprom</td>
<td>0.7824</td>
<td>0.0030</td>
<td>0.7809</td>
<td>0.0028</td>
<td>0.0015</td>
</tr>
<tr>
<td>2</td>
<td>JSC Gazprom Neft</td>
<td>0.7024</td>
<td>0.0026</td>
<td>0.7018</td>
<td>0.0028</td>
<td>0.0006</td>
</tr>
<tr>
<td>3</td>
<td>Sberbank</td>
<td>0.6260</td>
<td>0.0034</td>
<td>0.6261</td>
<td>0.0032</td>
<td>0.0001</td>
</tr>
<tr>
<td>4</td>
<td>Russian Agriculture Bank</td>
<td>0.6165</td>
<td>0.0032</td>
<td>0.6167</td>
<td>0.0030</td>
<td>0.0002</td>
</tr>
<tr>
<td>5</td>
<td>Oil Transporting JSC Transneft</td>
<td>0.5754</td>
<td>0.0025</td>
<td>0.4983</td>
<td>0.0026</td>
<td>0.0771</td>
</tr>
<tr>
<td>6</td>
<td>Lukoil Company</td>
<td>0.5417</td>
<td>0.0024</td>
<td>0.5410</td>
<td>0.0023</td>
<td>0.0007</td>
</tr>
<tr>
<td>7</td>
<td>Open JSC Rosneft</td>
<td>0.5394</td>
<td>0.0030</td>
<td>0.5407</td>
<td>0.0025</td>
<td>0.0013</td>
</tr>
<tr>
<td>8</td>
<td>JSC Russian Railways</td>
<td>0.5186</td>
<td>0.0029</td>
<td>0.5189</td>
<td>0.0025</td>
<td>0.0003</td>
</tr>
<tr>
<td>9</td>
<td>JSC VTB Bank</td>
<td>0.4913</td>
<td>0.0027</td>
<td>0.4911</td>
<td>0.0030</td>
<td>0.0002</td>
</tr>
<tr>
<td>10</td>
<td>Vnesheconombank</td>
<td>0.4583</td>
<td>0.0028</td>
<td>0.4583</td>
<td>0.0024</td>
<td>$&lt; 10^{-4}$</td>
</tr>
<tr>
<td>11</td>
<td>Bank of Moscow</td>
<td>0.4576</td>
<td>0.0027</td>
<td>0.4572</td>
<td>0.0025</td>
<td>0.0004</td>
</tr>
<tr>
<td>12</td>
<td>City Moscow</td>
<td>0.4377</td>
<td>0.0031</td>
<td>0.4375</td>
<td>0.0025</td>
<td>0.0002</td>
</tr>
<tr>
<td>13</td>
<td>MDM Bank Open JSC</td>
<td>0.4251</td>
<td>0.0026</td>
<td>0.4028</td>
<td>0.0032</td>
<td>0.0223</td>
</tr>
<tr>
<td>14</td>
<td>Alrosa C.L.</td>
<td>0.3890</td>
<td>0.0025</td>
<td>0.3885</td>
<td>0.0028</td>
<td>0.0005</td>
</tr>
<tr>
<td>15</td>
<td>Mobile Telesystems</td>
<td>0.3542</td>
<td>0.0025</td>
<td>0.3540</td>
<td>0.0023</td>
<td>0.0002</td>
</tr>
<tr>
<td>16</td>
<td>Open JSC VimpelComLimited</td>
<td>0.3523</td>
<td>0.0022</td>
<td>0.3524</td>
<td>0.0023</td>
<td>0.0001</td>
</tr>
<tr>
<td>17</td>
<td>JSC Russian Standard Bank</td>
<td>0.3290</td>
<td>0.0024</td>
<td>0.3296</td>
<td>0.0021</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

Table 3: Probabilities of default given sovereign default with BIC and BDs score.

<table>
<thead>
<tr>
<th>Order</th>
<th>Issuer</th>
<th>BN</th>
<th>CountryRank</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JSC Gazprom</td>
<td>0.7824</td>
<td>0.6220</td>
<td>0.1585</td>
</tr>
<tr>
<td>2</td>
<td>JSC Gazprom Neft</td>
<td>0.7024</td>
<td>0.5610</td>
<td>0.1415</td>
</tr>
<tr>
<td>3</td>
<td>Sberbank</td>
<td>0.6260</td>
<td>0.5854</td>
<td>0.0406</td>
</tr>
<tr>
<td>4</td>
<td>Russian Agriculture Bank</td>
<td>0.6165</td>
<td>0.5854</td>
<td>0.0312</td>
</tr>
<tr>
<td>5</td>
<td>Oil Transporting JSC Transneft</td>
<td>0.5754</td>
<td>0.5732</td>
<td>0.0022</td>
</tr>
<tr>
<td>6</td>
<td>Lukoil Company</td>
<td>0.5417</td>
<td>0.3381</td>
<td>0.2036</td>
</tr>
<tr>
<td>7</td>
<td>Open JSC Rosneft</td>
<td>0.5394</td>
<td>0.5244</td>
<td>0.0150</td>
</tr>
<tr>
<td>8</td>
<td>JSC Russian Railways</td>
<td>0.5186</td>
<td>0.5427</td>
<td>-0.0241</td>
</tr>
<tr>
<td>9</td>
<td>JSC VTB Bk</td>
<td>0.4913</td>
<td>0.6098</td>
<td>-0.1184</td>
</tr>
<tr>
<td>10</td>
<td>Vnesheconombank</td>
<td>0.4583</td>
<td>0.3339</td>
<td>0.1244</td>
</tr>
<tr>
<td>11</td>
<td>Bank of Moscow</td>
<td>0.4576</td>
<td>0.5305</td>
<td>-0.0729</td>
</tr>
<tr>
<td>12</td>
<td>City Moscow</td>
<td>0.4377</td>
<td>0.5122</td>
<td>-0.0745</td>
</tr>
<tr>
<td>13</td>
<td>MDM Bk Open JSC</td>
<td>0.4251</td>
<td>0.3131</td>
<td>0.1120</td>
</tr>
<tr>
<td>14</td>
<td>Alrosa C.L.</td>
<td>0.3890</td>
<td>0.2293</td>
<td>0.1596</td>
</tr>
<tr>
<td>15</td>
<td>Mobile Telesystems</td>
<td>0.3542</td>
<td>0.2446</td>
<td>0.1096</td>
</tr>
<tr>
<td>16</td>
<td>Open JSC VimpelComLimited</td>
<td>0.3523</td>
<td>0.2964</td>
<td>0.0559</td>
</tr>
<tr>
<td>17</td>
<td>JSC Russian Standard Bank</td>
<td>0.3290</td>
<td>0.0869</td>
<td>0.2420</td>
</tr>
</tbody>
</table>

Table 4: Comparison of $\gamma_C$ using BN and CountryRank model.
Figure 6: Quantiles of the Loss distribution without and with contagion with the BN and the CountryRank model.

Table 5: Comparison of the quantiles using Bayesian networks and CountryRank model.

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Loss - Standard model</th>
<th>BN model Contagion impact</th>
<th>CountryRank model Contagion impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>99%</td>
<td>1,115,153</td>
<td>117,341</td>
<td>11%</td>
</tr>
<tr>
<td>99.5%</td>
<td>1,443,579</td>
<td>1,054,844</td>
<td>73%</td>
</tr>
<tr>
<td>99.9%</td>
<td>2,324,088</td>
<td>2,302,667</td>
<td>99%</td>
</tr>
<tr>
<td>99.99%</td>
<td>3,514,968</td>
<td>2,198,517</td>
<td>63%</td>
</tr>
</tbody>
</table>

6 Concluding remarks

In this article, we presented a novel method of estimating contagion effects from CDS data using BNs. Rather than assuming a certain distribution for CDS spreads, we introduced a method for learning BNs using ε-drawups. Different techniques to learn the structure and parameters of financial networks were studied and evaluated. We used CDS spreads of issuers in a stylized portfolio and incorporated the conditional probabilities in the credit portfolio model presented by Anagnostou et al., 2018. Simulations were carried out for a stylized portfolio and the impact on standard risk metrics was estimated. Contagion was shown to have a significant impact in the tails of the credit loss distribution, with the results being in line with results obtained by using the CountryRank metric.

The results presented are a first step in the application of BNs on portfolio credit risk models. However, the BN framework we developed is flexible enough to allow for wider applications. For instance, one can extend the contagion so that stress originates from any issuer and not only at the sovereign. Moreover, one can test scenarios where multiple issuers default. Two examples of such scenarios can be found in Appendix D. These applications can be particularly useful for
risk managers, who are often interested in building scenarios for catastrophic risks and testing the resilience of their portfolios to such scenarios.

In order to extend our analysis, we plan to further investigate applications of the developed probabilistic framework in problems beyond credit portfolio modelling. A promising direction is to use our framework in order to identify systemically important nodes in the financial system and measure systemic risk. This could be done by considering, in a recursive manner, the fact that a node is more systemically important if it impacts many systemically important nodes. Another interesting direction is the modelling of wrong-way risk (WWR) arising in the case of a sovereign default in the pricing of Credit Valuation Adjustment (CVA) and Funding Valuation Adjustment (FVA) for interest-rate and foreign exchange derivatives.

Acknowledgement

The authors are grateful to Marcel Boersma for the useful comments.

Disclosure statement

The opinions expressed in this article are solely those of the authors and do not represent in any way those of their current and past employers. The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement no. 675044 (http://bigdatafinance.eu/), Training for Big Data in Financial Research and Risk Management.
References


A  CDS and Ticker list

In this section we briefly illustrate what a CDS contract is and after that provide the list with the CDS tickers with a mapping to the names of the issuers in the synthetic portfolio.

A credit default swap is a financial contract in which a buyer B gets insurance from a protection seller A against the default of a third party C. More precisely, given a contractual notional $N$, regular coupon payments with respect to $N$ and a fixed rate $s$, the CDS spread, are swapped with a payment of $N(1 - RR)$ in the event of the default of C, where $RR$, so-called recovery rate, is the contractual parameter which represents the part of the investment supposed to be recover in the event of default of C. An extensive description of these contracts including various modifications can be found in O’Kane, 2011.

The following list contains the issuers in the synthetic portfolio and the corresponding ticker which represent them in the networks depicted.

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Issuer name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT</td>
<td>Oil Transporting JSC Transneft</td>
</tr>
<tr>
<td>ALROSA</td>
<td>Alrosa C.L.</td>
</tr>
<tr>
<td>BKECON</td>
<td>Vnesheconombank</td>
</tr>
<tr>
<td>BOM</td>
<td>Bank of Moscow</td>
</tr>
<tr>
<td>CITMOS</td>
<td>City Moscow</td>
</tr>
<tr>
<td>GAZPRU</td>
<td>JSC Gazprom</td>
</tr>
<tr>
<td>GAZPRU.Gneft</td>
<td>JSC Gazprom Neft</td>
</tr>
<tr>
<td>LUKOIL</td>
<td>Lukoil Company</td>
</tr>
<tr>
<td>MBT</td>
<td>Mobile Telesystems</td>
</tr>
<tr>
<td>MDMOJC</td>
<td>MDM Bank Open JSC</td>
</tr>
<tr>
<td>ROSNEF</td>
<td>Open JSC Rosneft</td>
</tr>
<tr>
<td>RSBZAO</td>
<td>JSC Russian Standard Bank</td>
</tr>
<tr>
<td>RUSAGB</td>
<td>Russian Agriculture Bank</td>
</tr>
<tr>
<td>RUSRAI</td>
<td>JSC Russian Railways</td>
</tr>
<tr>
<td>RUSSIA</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>SBERBANK</td>
<td>Sberbank</td>
</tr>
<tr>
<td>VIP</td>
<td>Open JSC VimpelCom Limited</td>
</tr>
<tr>
<td>VTB</td>
<td>JSC VTB Bank</td>
</tr>
</tbody>
</table>

Note that JSC is the acronym for Joint Stock Company.

B  Systematic factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>MSCI EUROPE</td>
</tr>
<tr>
<td>Asia</td>
<td>MSCI AC ASIA</td>
</tr>
<tr>
<td>North America</td>
<td>MSCI NORTH AMERICA</td>
</tr>
<tr>
<td>Latin America</td>
<td>MSCI EM LATIN AMERICA</td>
</tr>
<tr>
<td>Middle East and Africa</td>
<td>MSCI FM AFRICA</td>
</tr>
<tr>
<td>Pacific</td>
<td>MSCI PACIFIC</td>
</tr>
<tr>
<td>Materials</td>
<td>MSCI WRLD/MATERIALS</td>
</tr>
<tr>
<td>Consumer products</td>
<td>MSCI WRLD/CONSUMER DISCR</td>
</tr>
<tr>
<td>Services</td>
<td>MSCI WRLD/CONSUMER SVC</td>
</tr>
<tr>
<td>Financial</td>
<td>MSCI WRLD/FINANCIALS</td>
</tr>
<tr>
<td>Industrial</td>
<td>MSCI WRLD/INDUSTRIALS</td>
</tr>
<tr>
<td>Government</td>
<td>ITRAXX SOVX GLOBAL LIQUID INVESTMENT GRADE</td>
</tr>
</tbody>
</table>

Table 6: Systematic factor - Index mapping.
C Scores

We briefly describe the two scores namely the BIC and BDs used for BN learning. For details, we refer to Koller & Friedman, 2009. BIC is a likelihood score defined as:

$$\text{score}_{BIC}(\mathcal{G} : \mathcal{D}) = \log L(\mathcal{G} : \mathcal{D}) - \frac{1}{2} |\mathcal{G}| \log N$$

where $|\mathcal{G}|$ is the complexity of the network, the number of independent parameters in the network. This score penalizes explicitly the score with the number of independent parameters, hence assigning higher scores to sparser structures.

The following expression is the closed-form derived for the marginal-likelihood of the structure score:

$$P(\mathcal{D} \mid \mathcal{G}) = \prod_{i} \prod_{k=1}^{q_i} \frac{\Gamma(\alpha_{X_i,k})}{\Gamma(\alpha_{X_i,k} + M[k])} \prod_{j=1}^{r_i} \left[ \frac{\Gamma(\alpha_{j,k} + M[j,k])}{\Gamma(\alpha_{j,k})} \right]$$

(25)

where $\alpha_{X_i,k} = \sum_{j=1}^{r_i} \alpha_{j,k}$.

BDs is derived from a different score, BDeu, which is obtained from 25 by assuming a uniform prior distribution over the parameters (Dirichlet distribution with all the hyperparameters taking the same value $\alpha$). Let $\alpha_{j,k} = \alpha/(r_i q_i)$ and $\alpha_i = \alpha$, where $r_i$ is the number of states of $X_i$ and $q_i$ is the number of configurations of parents of $X_i$, the number of parents configuration of $X_i$. Then score BDeu is defined as follows

$$\text{BDeu}(\mathcal{G}, \mathcal{D}; \alpha) = \prod_{i} \prod_{k=1}^{q_i} \frac{\Gamma(r_i \alpha_i)}{\Gamma(r_i \alpha_i + M[k])} \prod_{j=1}^{r_i} \frac{\Gamma(\alpha_i + M[j,k])}{\Gamma(\alpha_i)}$$

(26)

Scutari, 2016 argues that choosing uniform prior distributions over $\theta_{\mathcal{G}X_i \mid \mathcal{P}a_{X_i}}$ and $\mathcal{G}$ can have a negative effect over the quality of the results obtained with the score BDeu. To avoid this he introduces the score BDs.

In the first place we see that if $P(k) = 0$ for some $k \in \{1, \ldots, q_i\}$ and $i \in [n]$, or if the sample size of $\mathcal{D}$ is very small, it may happen that $M[k] = 0$ for some configurations of $\mathcal{P}a_{X_i}$, which do not appear in $\mathcal{D}$, then we can split

$$\text{BDeu}(\mathcal{G}, \mathcal{D}; \alpha) = \prod_{i} \left( \prod_{k=1}^{q_i} \frac{\Gamma(r_i \alpha_i)}{\Gamma(r_i \alpha_i + M[k])} \prod_{j=1}^{r_i} \frac{\Gamma(\alpha_i + M[j,k])}{\Gamma(\alpha_i)} \right)$$

We note that as the number of parent configurations which appear in the data $\mathcal{D}$ decreases, the effective imaginary sample size decreases, as

$$\sum_{k:M[k]>0} \sum_{j} \alpha_i \leq \sum_{k,j} \alpha_i = \alpha$$

(27)

This induces the posterior to converge to the corresponding likelihood estimation and hence leaning towards overfitting and including spurious edges in $\mathcal{G}$. To avoid this problem we define:

$$\tilde{q}_i = |\{k \in \{1, \ldots, q_i\} : M[k] > 0\}| \quad \text{and} \quad \tilde{\alpha}_i = \begin{cases} \alpha/(r_i \tilde{q}_i) & \text{if } \tilde{q}_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

With this new definition the expression 27 becomes an equality, $\sum_{k:M[k]>0} \sum_{j} \alpha_i = \alpha$. Moreover, we see that the uniform prior that we just defined is on the conditional distribution which can be estimated from $\mathcal{D}$, so this is a empirical Bayesian score. Finally, we substitute $\alpha_i$ in 26 by $\tilde{\alpha}_i$ and obtain:

$$\text{BDs}(\mathcal{G}, \mathcal{D}; \alpha) = \prod_{i} \prod_{k=1}^{\tilde{q}_i} \frac{\Gamma(r_i \tilde{\alpha}_i)}{\Gamma(r_i \tilde{\alpha}_i + M[k])} \prod_{j=1}^{r_i} \frac{\Gamma(\tilde{\alpha}_i + M[j,k])}{\Gamma(\tilde{\alpha}_i)}$$

(28)
D Stress scenarios

The first scenario is stressing three banks in the network, BKECON, BOM and SBERBANK. Figure 7 depicts the network with the nodes colored according to the probabilities obtained in that scenario, shown next to it. We observe that the values obtained are similar to the ones of the scenario where the sovereign is stressed. This may be due to the strong connections of the sovereign with big companies such as GAZPRU and GAZPRU.Gneft and because the stressed banks are in the periphery of the network.

<table>
<thead>
<tr>
<th>Issuer</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKECON</td>
<td>1</td>
</tr>
<tr>
<td>BOM</td>
<td>1</td>
</tr>
<tr>
<td>SBERBANK</td>
<td>1</td>
</tr>
<tr>
<td>RUSAGB</td>
<td>0.7757</td>
</tr>
<tr>
<td>VTB</td>
<td>0.7287</td>
</tr>
<tr>
<td>GAZPRU</td>
<td>0.69</td>
</tr>
<tr>
<td>AKT</td>
<td>0.5797</td>
</tr>
<tr>
<td>GAZPRU.Gneft</td>
<td>0.5593</td>
</tr>
<tr>
<td>ROSNEF</td>
<td>0.5514</td>
</tr>
<tr>
<td>RUSSIA</td>
<td>0.5351</td>
</tr>
<tr>
<td>CITMOS</td>
<td>0.5107</td>
</tr>
<tr>
<td>LUKOIL</td>
<td>0.4948</td>
</tr>
<tr>
<td>VIP</td>
<td>0.4742</td>
</tr>
<tr>
<td>RUSRAI</td>
<td>0.4528</td>
</tr>
<tr>
<td>ALROSA</td>
<td>0.4244</td>
</tr>
<tr>
<td>MDMOJC</td>
<td>0.4242</td>
</tr>
<tr>
<td>MBT</td>
<td>0.3379</td>
</tr>
<tr>
<td>RSBZAO</td>
<td>0.3319</td>
</tr>
</tbody>
</table>

Figure 7: Probabilities and network in the scenario of three banks stressed.

As the Russian crisis of 2014 was strongly related to oil industry, in the second scenario we stress four of the largest oil companies: AKT, GAZPRU, LUKOIL, and ROSNEF. Figure 8 shows the network and the probabilities for this case. It is noticeable that the oil companies have a larger impact on the network. Note also that in this case four nodes are stressed. However, one can see that these nodes are more centered and the rest of the nodes are more stressed.

<table>
<thead>
<tr>
<th>Issuer</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT</td>
<td>1</td>
</tr>
<tr>
<td>GAZPRU</td>
<td>1</td>
</tr>
<tr>
<td>LUKOIL</td>
<td>1</td>
</tr>
<tr>
<td>ROSNEF</td>
<td>1</td>
</tr>
<tr>
<td>GAZPRU.Gneft</td>
<td>0.8382</td>
</tr>
<tr>
<td>RUSAGB</td>
<td>0.7867</td>
</tr>
<tr>
<td>SBERBANK</td>
<td>0.7563</td>
</tr>
<tr>
<td>RUSSIA</td>
<td>0.6469</td>
</tr>
<tr>
<td>VTB</td>
<td>0.6028</td>
</tr>
<tr>
<td>RUSRAI</td>
<td>0.5814</td>
</tr>
<tr>
<td>MDMOJC</td>
<td>0.5321</td>
</tr>
<tr>
<td>BKECON</td>
<td>0.5273</td>
</tr>
<tr>
<td>CITMOS</td>
<td>0.5133</td>
</tr>
<tr>
<td>MBT</td>
<td>0.5078</td>
</tr>
<tr>
<td>BOM</td>
<td>0.4972</td>
</tr>
<tr>
<td>ALROSA</td>
<td>0.4708</td>
</tr>
<tr>
<td>RSBZAO</td>
<td>0.4053</td>
</tr>
<tr>
<td>VIP</td>
<td>0.366</td>
</tr>
</tbody>
</table>

Figure 8: Probabilities and network in the scenario of four oil companies stressed.