Ca' Foscari
University
of Venice

Master's Degree Programme

in Economics and Finance

Final Thesis

Bayesian Networks

and Financial Stress Testing

Assessing the Probability of Default for a Credit Institution

Supervisor
Debora Slanzi

Co-Supervisor
Monica Billio

Graduand
Andrea Carraro

Matriculation Number
848303

Academic Year
2017 / 2018

Abstract

The thesis is focused on the practice of *Financial Stress Testing* analysis as a specific risk management instrument that may be employed by credit institutions to monitor the levels of selected key performance indicators. These statistics may support management in decision making when performing assessments to evaluate banks' current and expected economic healthiness conditions, which could be influenced by factors either internal to the firm or belonging to the external macroeconomic environment. Then, with regard to the financial context, such analyses may be useful to serve regulatory compliance purposes, other than being of guidance for the implementation of crisis-prevention actions, as well as for the execution of resource planning and control activities. With the aim to adequately respond to the needs of credit institutions in performing risk management tasks, the present work proposes a methodology which is grounded on the use of *Bayesian Networks* as specific data analysis tools that may be utilized in the conduction of Financial Stress Testing practices. Indeed, Bayesian Networks are instruments deemed to be capable of providing accurate indications on dependencies between and among business-specific factors. The assessment of such relations, via simulation procedures, may allow the identification of criticalities relative to the single banking institution, which could consequently be able to decide where to focus efforts and, in case necessary, evaluate the execution of corrective actions. Therefore, in this sense, Bayesian Networks may be considered useful and adequate tools in supporting Financial Stress Testing practices. To this end, the present work provides a case study analysis, based on real-world data, on the application of the previously-mentioned risk monitoring activities.

Contents

List of Tables

List of Figures

List of Abbreviations

ADF	-	Augmented Dickey Fuller (Test)
BGe	-	Bayesian Gaussian equivalent uniform
BIC	-	Bayesian Information Criterion
BN	-	Bayesian Network
CA	-	Capital Adequacy
$CET1$	-	Common Equity Tier 1 Ratio
CI	-	Confidence Interval
CoR	-	Cost of Risk Ratio
CPT	-	Conditional Probability Table
DAG	-	Directed Acyclic Graph
EAD	-	Exposure At Default
EBA	-	European Banking Authority
ECB	-	European Central Bank
ECM	-	Error Correction Model
$Euribor$	-	Euribor 3-months Rate
GBN	-	Gaussian Bayesian Network
GC	-	Granger Causality
GDP	-	Gross Domestic Product
HC	-	Hill Climbing (Algorithm)
$Infl$	-	Inflation Rate
IR	-	Interest Rate Risk
LGD	-	Loss Given Default
LR	-	Leverage Ratio
MC	-	Monte Carlo (Simulation)
$Mkt.Sh$	-	Market Share of Company Assets
MLE	-	Maximum Likelihood Estimation
NPL	-	Non-Performing Loans Ratio
$NSFR$	-	Net Stable Funding Ratio
Oil	-	Oil Price
OLS	-	Ordinary Least Squares
PD	-	Probability of Default
$RORAC$	-	Return On Risk-Adjusted Capital
RWA	-	Risk-Weighted Assets
SSM	-	Single Supervisory Mechanism
UN	-	Unemployment Rate
VAR	-	Vector Auto-Regressive (process)
VaR	-	Value at Risk

Acknowledgements

I gratefully acknowledge my supervisors for their guidance and support during the development of this thesis.

I acknowledge those who let me utilize actual data as a basis for the realization of a real-world case study analysis, with the surely interesting purpose of verifying the truthfulness of theoretical assertions on a practical application.

I am also thankful to all those people I had the chance to interact with and who induced me into thinking critically through cues which helped me elaborate ideas for the realization of this work.

Special thanks and blessings go to my loved ones who constantly support me in life, being them always the true reasons of my joy.

Introduction

The aim of this work is to answer the needs of executing a *Financial Stress Testing* analysis by applying a statistical approach grounded on the use of *Bayesian Networks*. In particular, an introduction to the Stress Testing concepts and theory of Bayesian Networks would be provided, as well as a practical example relative to the application of such statistical tool in the bank-related environment.

Financial Stress Testing analyses may be identified as part of the risk management practices and deemed to be especially useful for credit institutions, provided that such analyses are often focused on a wide array of financial indicators, hence they may provide insights relative to key aspects of a bank's business. Moreover, other than assessing and providing forecasts on the level of several statistics under different market conditions, a rigorous Stress Testing analysis would demand further investigations on the factors that contribute to the definition of key performance indicators and critical thresholds. Indeed, one may wish to identify their root and common dependencies, so that to gain more informative analysis' outputs to be exploited when making business-related decisions. For this particular instance, among the selection of statistical instruments that a credit institution may employ, it would be appropriate to choose a tool which allows to account for interactions between and among the variables selected for the analysis.

A Bayesian Network is a statistical model which combines the properties of graphical and probabilistic theories to derive joint probability distributions on the variables of a system under study. Moreover, such tool would be of use in assessing the strength of dependencies that may persist throughout a specific network of elements. This is a fundamental feature which allows to perform analyses on *as is* conditions and to provide estimates for potential future scenarios. With this in mind, concerning the financial and economic environments, entities operating in the industry would need to adequately address their needs of managing risks, as well as operating monitoring activities to ensure business solidity, even in case of crisis-like events. Therefore, given the significant amount and variety of market and firm-related data tracked by financial institutions, Bayesian Networks may support the aforementioned analyses and controlling activities, provided that they are flexible enough tools which could handle the complexity of data at disposal and may allow to grasp those critical relations that persist between and among the elements object of the analysis. Thus, Bayesian Networks are deemed to be adequate statistical models and represent a valid solution to employ when performing activities related to Financial Stress Testing.

The present work is structured in a *broad-to-specific* fashion. To be clear, the thesis starts from providing notions for a general understanding of the financial context of application relative to Stress Testing. Then, the focus is narrowed toward the statistical instruments utilized to carry out business-specific analyses, hence providing a dedicated theoretical description on Bayesian Networks. Lastly, elements converge to display the results of a specific analysis performed utilizing real-world data for a credit institution.

In particular, the thesis is structured as follows.

Chapter 1 deals with presenting an overview relative to the specific financial risk management context. In particular, since financial institutions necessitate to thoroughly analyze their own creditworthiness, either for self-control needs or to comply with regulations into force, the practice of *Financial Stress Testing* may be employed to serve these purposes. Indeed, the aim would be to test via simulation the solidity of credit institutions, both under normal and downturn market conditions. To this end, Bayesian Networks would allow to take into consideration those dependence relations persisting throughout and between the macroeconomic and firm-specific contexts. Hence, such statistical tool may convey information on the possible short-term trends relative to selected key performance indicators; in turn, by assessing both the *status quo* and forecasts for such measures, financial entities may be able to take the necessary actions in face of potential future adverse economic conditions.

Chapter 2 provides an overview of the main concepts and functionalities relative to Bayesian Networks as statistical and data analysis tools. In particular, they are characterized by graphical and probabilistic properties, as mentioned already, and one of their key features is that they are backed up by selected learning algorithms, which allow the definition of networks' structure and parameters based on data at hand. Notice in particular that, for the purposes of the present work, the focus will be on the use of a specific type of instrument, namely the Gaussian Bayesian Network.

In chapter 3 a specific Financial Stress Testing simulation is carried out with the aid of Bayesian Networks. To be more specific, a case study analysis would be executed relative to actual data concerning a credit institution operating in the European context. The main purpose of the application would be to assess the probability of default for such financial institution, as well as providing an assessment on the level of selected firm-specific key performance indicators. To this end, the analysis would consist in defining historical trends and simulating potential future ones, either in a context of normal or critical market-wide economic conditions.

At last, some conclusions and closing remarks are provided.

1 Financial Stress Testing

The present chapter will be dedicated to analyzing and reviewing the core elements and principles which relate to the conduction of a *Financial Stress Testing* analysis, with particular focus toward those aspects that concern the case study presented in the third chapter. This will help clarify and put stress on those key concepts which are a prerequisite for the understanding and application of Bayesian Networks to the empirical analysis context.

1.1 Definition and Macroeconomic Regulatory Context

In the first place, Financial Stress Testing is often defined as being a supervisory and monitoring instrument utilized to verify the resistance of credit institutions to potential future economic downturns. In other words, from a regulatory standpoint, the test mainly consists in a simulation study assessing the bank's ability to bear losses and evaluating its resilience capabilities, on the basis of predefined critical thresholds.

In the aftermath of the 2008 Crisis, regulators and practitioners casted doubt on the appropriateness of previously-utilized *Quantitative Risk Management* instruments and felt the need to define tools to provide a more realistic picture on the solvency state of credit institutions. One of the pre-Crisis tools mostly in use for the assessment of risks related to the banking sector was Value at Risk (VaR). However, as it turned out, VaR revealed itself as being not an adequate instrument to evaluate tail scenarios (Gao, Mishra and Ramazzotti, 2017). Therefore, as already-mentioned, the need for a more adequate risk management tool led to the creation and rise of Stress Testing analyses on financial institutions.

Notice that, such risk assessment practice may be characterized by aspects that are tailored and relevant to the bank object of the analysis, hence defining a *micro-level* dimension focus. Still, it is also important not to lose sight of the broader picture in which a credit institution may be operating, that is to say the relative macroeconomic context. In fact, shocks and imbalances at the *macro-level* dimension could generate significant downside effects (like for instance liquidity or credit-related troubles) on each single banking institution (Dees, Henry and Martin, 2017).

Other than the interaction with the real economy and the interconnections between and across banks, Stress Testing analyses should take into account the dynamicity in bank's economic and financial cycles, in a way that reflects the most closely its solvency situation and accounts for its actions in response to external systemic events (Dees, Henry and Martin, 2017). For instance, a credit institution may decide to adopt corrective/strategic actions like deleveraging or capital raises, hence changing its balance sheet composition to counter the effects of macroeconomic stresses. Then, in order to take adequate decisions, a Stress Test may be structured in a way that ensures the appropriate granularity in data and time instants, so that to output proper information on the bank's financial situation.

The principles intrinsic to the practice of Stress Testing and the obtained financial indicators are often demanded by authorities (like the European Central Bank (ECB), the European Banking Authority (EBA) or the Basel Committee on Banking Supervision) in order to oversight and monitor over the broader macroeconomic and financial context. Still, the simulated measures help the financial institution in assessing its own economic state and possibly may indicate the need for corrective actions. Indeed the simulation output may be utilized, say, for resource planning/control, compliance and risk monitoring purposes. Therefore, provided that the quality and quantity of data would be sufficient and adequate to carry out analyses, the Stress Testing analysis could also be exploited to effectively implement continuous improvement activities on the credit institutions (Basel Committee, 2017).

With reference to the macro-prudential regulatory frameworks topic, the ECB puts into place the so called Single Supervisory Mechanism (SSM): the ECB is in charge of a supervisory role in monitoring the financial stability relative to the countries within the Eurozone. Notice, in particular, that Financial Stress Testing is actually part of the SSM. In this regard, it is worth providing a quick overview on a specific model utilized by the ECB to conduct analysis for the assessment of credit institutions' solvency. To provide a summary of such model, it is enough to report its four core pillars (Dees, Henry and Martin, 2017):

- Identify the elements and scenarios belonging to the macroeconomic environment which are relevant to the credit institution object of the analysis;
- Generate variables describing the previously-identified elements to evaluate the bank's projected loss absorption capacity;
- Starting from the estimated future losses obtained in the previous step, calculate the overall impacts on the financial institution's solvency position;
- Evaluate the effects generated by the bank's own initial solvency state toward the broader financial and economic environment.

Hence, in general, the Stress Testing analysis would be focused on forecasting future values of a chosen set of financial ratios, under baseline and adverse scenarios, in order to assess the financial healthiness of the credit institution, conditional on some critical thresholds set for those same ratios (Dees, Henry and Martin, 2017).

1.2 Approaches to Stress Testing

Keeping in mind that the main objective of the simulation would be to assess the frequency and impacts with which adverse events would affect financial institutions, specific and different approaches may be applied to the Stress Testing analysis, based on the analyst's choice. Hence, it is worth reviewing some of the most relevant methods to be implemented, reported and briefly described next (Rebonato, 2017):

- The *Extreme-Tail Approach* essentially gives more importance and weight to tail events, selecting those data which reflect the most adverse conditions mapped in the past.
  The reliability of such method may be questioned in that it is heavily reliant on past data and actually excludes some possible scenarios from the overall picture, somehow outputting less informative conclusions;

- The *Vulnerability-Driven Approach*, in the first place, focuses on identifying vulnerabilities that are present in the credit institutions' activities. In order to accomplish that, for instance, one may perform a comparison between the actual values of chosen financial performance indicators and their relative institutional/regulatory threshold values. Secondly, on the basis of the former step, the perimeter for potential adverse scenarios should be defined.
  Notice that one difficulty with such approach rests in the fact that quantitative measures of events' probabilities are often difficult to obtain;

- The *Coherent-Stress-Testing Approach* (a) is designed to identify the so called *root causes* of effects influencing other variables of the analysis; (b) aims at defining connections between and among variables, often expressed through (conditional) probabilities. Then, the ultimate aim of the analysis would be to evaluate the variables' effects jointly, usually in order to draw conclusions on a specific financial performance indicator, depending on the specific objective for which the Stress Testing was designed and implemented.

The main challenge relative to this method would be to accurately assign events' and variables' probabilities, which often are derived from other models or assigned subjectively.

In practice, a number of quantitative methodologies may be employed to support the conduction of analyses within the boundaries of the afore-mentioned methods. Then, for instance, financial institutions may rely on the following statistical instruments for Financial Stress Testing purposes: *VAR* and *ARIMA* models are rather common tools to utilize when assessing *as is* and *to be* conditions of selected variables, while *Markov Switching*[1] models are seldom employed. However, as better exposed later, the choice for the present work is to focus on the use of Bayesian Networks, aiming to investigate for the presence of dependencies among the variables under study, aspect which may not be grasped by other types of modeling methodologies.

Notice that, with reference to all the possible approaches which could be adopted in the context of Stress Testing and, more in general, to Risk Management, a certain degree of subjectivity must be accepted and adequately utilized in these kind of analyses. Such qualitative component would serve to complement the quantitative side of outputs generated by statistical models, in order to obtain a more informative picture on the state of the credit institution under study.

The above-mentioned Financial Stress Testing methodologies may be implemented by regulators and authorities to monitor the financial sector at large. Still, as explained in the following section, such approaches may be applied by single banking institutions either for regulatory or risk monitoring purposes.

[1] For more information on *Markov Switching* models applied to Financial Stress Testing see Jacobs Jr. and Sensenbrenner (2018).

1.3 Key Indicators and Specific Regulatory Context

Other than being useful to regulators for macro-prudential purposes, the Stress Testing analysis is also demanded by supervisory authorities in order to oversight the single credit institutions, primarily on liquidity and capital adequacy matters. Furthermore, in order to perform the necessary assessments, analyses are often focused on the evaluation of the bank's risks relating to credit, market, interest rates and liquidity indicators. In this respect, it would be key to dispose of sufficiently accurate data, with particular attention to the proper mapping of exposure-related information (Basel Committee, 2017).

A more thorough explanation of the elements that may be part of a Stress Testing analysis will be provided in the current section, together with some considerations relative to regulatory topics on single banking entities. Moreover, what comes next will also be of use with regard to the understanding of the empirical case study presented in the third chapter.

Starting from the regulatory context, it must be stressed that some indicators utilized in the analysis may also be interpreted as measures to evaluate a credit institution on its level of compliance with the rules and regulations into force. Therefore, it is of interest to report some key concepts from the *Basel III* framework, which directly apply to credit institutions at international level and recall some features of the Stress Testing analysis practice.

The core idea behind the *Basel III* framework is that exposures should be *"backed by a high quality capital base"* (Basel Committee, 2011). Banks should possess enough capital reserves to serve as a buffer against potential losses associated to the creditworthiness deterioration of other financial players, as well as against possible future adverse market-wide cycles. Furthermore, other than enhancing the quality and quantity of capital reserves, the framework demands that banks hold a so called *capital buffer* accumulated during good times and exploited during recessions, hence countering the effects of economic/financial cycles (Bologna and Segura, 2016).

For these reasons, banks should carry out analyses keeping into consideration the simulation of adverse scenarios to better define the amount of capital reserves to allocate. The key measures imposed by the *Basel III* framework for capital requirements follow (Basel Committee, 2011):

- The *Common Equity Tier 1* (CET1) *Ratio* must always be at least at a level of 4.5%. The next formula illustrates how this indicator is calculated:

$$CET1\ Ratio = \frac{CET1}{RWAs} \tag{1.1}$$

  Where *RWAs* stands for Risk-Weighted Assets[2].

- The *Capital Adequacy Ratio* must always be at least equal to 8.0%. The following formula shows its composition:

$$Capital\ Adequacy\ Ratio = \frac{Tier\ 1\ Capital + Tier\ 2\ Capital}{RWAs} \tag{1.2}$$

Notice that, although being a key requirement, the creation of adequate capital buffers is not enough for an appropriate management of credit-related risks. Indeed, this must be supported by internationally harmonized rules concerning liquidity and leverage requirements. To this end, a financial institution should dispose of sufficient short-term liquid assets to survive sudden severe adverse events, along with more stable resources to ensure resilience over the long term (Basel Committee,

[2] Usually defined as bank's assets or exposures weighted according to the relative risk, which may be assessed either through the *Standardized* or the *Internal Ratings-Based* (IRB) approaches. For more on this topic, see Basel Committee (2011).

2011). In this regard, the *Basel III* framework proposes the use of measures like:

- The *Net Stable Funding Ratio*, which in principle should be grater or equal to 100%, calculated utilizing the following formula:

$$Net\ Stable\ Funding\ Ratio = \frac{Available\ funding\ for\ the\ medium\ term}{Total\ amout\ of\ stable\ funding} \quad (1.3)$$

- The *Leverage Ratio*, which should be grater or equal than 3%, computed in the following way:

$$Leverage\ Ratio = \frac{Tier\ 1\ Capital}{Total\ Exposures} \quad (1.4)$$

One may assert that the indicators and thresholds reported above are the main principles a banking institution must comply with in order to be in line with the *Basel III* framework key requirements. Still, performing an effective control on potential financial risks would also imply monitoring and assessing counterparty credit risk (by keeping track of all exposures) and market-wide risk factors (for instance interest rates, exchange rates and equities).

Other than the *Basel III* framework's requirements, an adequate supervision on risk would be supplemented by the use of the aforementioned Stress Testing analysis, which may be performed via simulation to obtain estimates (through forecasting models) of the potential impacts and vulnerabilities affecting the bank. Such estimated conditions would be interpreted through different elements (financial indicators), each providing a specific evaluation on the state in which the bank is expected to be in future times[3]. Therefore, the overall assessment is performed by considering different indicators' values across the

[3] It is usually the case that indicators are estimated to obtain forecasts up until one year ahead with respect to present time in which the Stress Testing analysis is performed.

financial institution, which may be affected by the realization of exogenous and/or endogenous plausible scenarios. To put it another way, macroeconomic and bank-specific variables contribute to the definition of the aforementioned financial indicators and the relative forecasts, also through interactions among them. Notice that such interconnections are mostly based on the relative time series values, hence being characterized by a specific time trend as well as random components to allow for uncertainty in future times.

It is usually the case that the key indicators chosen for the analysis fall in the following four broad categories:

- *Risk-Adjusted Framework* (RAF), including indicators which values are adjusted for the relative risk. For instance, one may consider all those indicators calculated with the contribution of Risk-Weighted Assets or exposures adjusted for risk. In general, financial indicators of this type may be *CET1 Ratio*, *Return on Equity*, *Non-Performing Loans Ratio* and *Cost of Risk Ratio*;
- *Liquidity*, which measures the quantity of liquid assets at disposal of the bank. This category includes indicators like *Self-Funding Ratio* and *Net Stable Funding Ratio*;
- *Credit*, measuring the level of indebtedness of the credit institution. Indicators relative to this category may be *Leverage Ratio*, *Non-Performing Loans Ratio* and *Probability of Default*;
- *Rates*, that mainly include those measures related to *Interest* and *Exchange Rates Risks*.

Notice that, baseline values for the financial indicators are directly estimated through modeling and forecasting procedures. The reference (threshold) values are either set by regulatory bodies or self-imposed by the credit institution itself. In turn, depending on each indicator’s value (and whether it crosses a predefined threshold) the bank may decide to implement corrective actions, in case necessary.

More in depth explanations and definitions relative to financial indicators will be provided in the chapter dedicated to the case study analysis, in particular, focusing on those that will effectively be utilized to carry out the empirical project.

Nonetheless, it is necessary to expand more on a specific element that reveals itself being a central factor in the third chapter: the *Probability of Default*. The following sub-section provides further concepts relative to such measure and its components.

1.3.1 Probability of Default

One of the key elements which would be estimated in a Stress Testing analysis is the *Probability of Default* (PD) connected to the analyzed credit institution. Intuitively, this measure may be interpreted as being a summary indicator for the overall riskiness connected to such banking institution. Indeed, many factors, which may be either bank-specific and/or macroeconomic, contribute to the definition of the PD value relative to that credit institution.

Depending on whether the bank is a public or private company, computations to obtain the relative PD may differ, indeed, data at disposal also would differ. In general, one may exploit historical data (if available) obtained from market-related sources or from previously performed Stress Tests in order to assess the present state of PD. In particular, concerning a credit institution which shares are traded on the market, it would be possible to obtain proxies for PD values estimation mainly from (Dees, Henry and Martin, 2017):

a) *Bank default rates realized in past times*. In such case the historical information utilized may not be totally adequate, given these kind of rates are obtained from data reflecting company-specific characteristics. Consequently, there may be no correspondence between such features and the ones possessed by the bank which PD has to be estimated;

b) *Default rates obtained from Non-Performing Loans* (NPL) *statistics*. This approach is often preferred to the former, provided that NPL data do not necessarily imply that a company would be insolvent. Moreover, NPL data better reflect the composition of a credit institution's portfolio, hence characterizing its own PD;

c) *Default rates derived from market data of financial companies*. PD would be estimated via inputs like leverage (derived from the long-term and short-term liabilities of the bank) and asset volatilities (obtained from the bank equities' volatility).

Notice that a distinction has to be made between two types of estimated PDs: *actual PDs* and *risk-neutral PDs* (Dees, Henry and Martin, 2017). In the former case PD would be obtained from financial expected default rates based on the relative time series of financial institutions. In the latter case PDs may be implied by market instruments like Credit Default Swaps (CDS) which also reflect market-related risks; hence, PD values would be partly biased and not uniquely reflecting a bank's core features and risks.

There exist cases in which the bank object of the analysis may not be a publicly traded company; as a consequence, it would not be possible to exploit market-related data on such financial institution. The way to estimate the relative PD would be to utilize a model which exploits the bank's balance sheet data among its inputs. One proposed solution to the derivation of PD values would be to utilize the method proposed by the Basel Committee for Banking Supervision, that is the *Internal-Ratings Based* (IRB) approach for PD estimation, which main passages will be reviewed next.

Key elements to consider when assessing the probability of a credit institution to default are the relative future projected losses and the regulatory capital buffer against unexpected adverse trends. These factors are also reflected on the *IRB* approach which, in particular, is based on the following assumptions for PD calculation (Gómez, Parrado and Partal, 2018):

a) *Basel III* assertions on credit risks are correct;

b) The single credit institutions' risks are in line with the *Basel III* framework measures;

c) All risks relative to a bank may be interpreted in terms of credit risks.

In particular, the inputs to calculate *PD* estimate relative to a credit institution are reported next (Hurliny, Leymariez and Patinx, 2017):

- The amount of regulatory capital as defined by the *Basel III* framework (that would be the *Capital Adequacy* measure for capital, *CA*, in this case), which corresponds to the sum of all capital requirements for each exposure;

- The *Exposure At Default* (*EAD*), that is the total value of outstanding debt (exposures) in case the bank defaults on its loans;

- The *Loss Given Default* (*LGD*), which is a portion of EAD that is effectively lost in case the relative debtors default on their loans;

- The *Maturity Adjustment* term, expressed by

$$\gamma(M) = \frac{(1 + (M - 2.5)\, b(PD))}{1 - 1.5\, b(PD)} \tag{1.5}$$

with a smoothing factor $b(PD) = (0.11856 - 0.05478 \ln(PD))^2$ and M as the maturity measure.

- A term $\delta(PD)$ relating to probability defined as

$$\delta(PD) = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\rho(PD)}\Phi^{-1}(99.9\%)}{\sqrt{1-\rho(PD)}}\right) - PD \tag{1.6}$$

where $\Phi(.)$ denotes the cumulative distribution function of a standard normal distribution and $\rho(PD)$ would be identified as the *default correlation* of probability which utilized formula would be

$$\rho(PD) = 0.12\left(\frac{1-e^{-50\,PD}}{1-e^{-50}}\right) + 0.24\left(1-\left(\frac{1-e^{-50\,PD}}{1-e^{-50}}\right)\right) \tag{1.7}$$

Given all the components and bank's balance sheet data, it is possible to derive a value for the Probability of Default in a given moment in time ($t = 0, 1, 2, \dots, T$) by solving formula (1.8) for PD, which would be contained within both $\gamma(M)$ and $\delta(PD)$, as shown in the formulas above:

$$(EAD_t \cdot LGD_t \cdot \delta(PD)_t \cdot \gamma(M)_t) - CA_t = 0. \tag{1.8}$$

Then, by computing the PD in different time instants, for which data are updated, it may be possible to construct a time series in order to observe its evolution and possibly perform forecasts on such indicator. Moreover, changes in PD may be useful in that they could stimulate the credit institution's effort to execute further and deeper analyses on other key financial indicators, with the aim to identify potential issues and possibly take corrective actions where necessary.

1.4 Assessing Causation in Stress Testing

In general, the output of a Stress Testing analysis should be suggesting whether corrective actions may be taken by the bank as to either comply with regulations or improve stability and solidity concerning specific indicators. Hence, to properly act on the basis of the results provided by the analysis, the utilized tools (models) should be informative on which causal relations would link the considered variables. Then, one may be able to understand on which indicators to intervene and which corrective measures would be most adequate to implement.

In this sense, with reference to the PD estimation case, it may be possible to grasp which relations persist among macroeconomic or firm-specific factors contributing to the definition of the PD value. In particular, notice that links dictated by correlations may not be really informative for corrective action implementation. Indeed, what would actually be useful is the knowledge of links between and among variables generated by *causal* relations. Hence, an important aspect of a Stress Testing analysis would be to avoid evaluating the effects of the considered variables separately and, instead, study each relative contribution in conjunction with all the others jointly.

To this end, in order to better represent and visually display causal relations among variables, it may be possible to utilize statistical and probabilistic models. In particular, the present work focuses on the use of *Bayesian Networks*, grounded on the concept that there exist reasons for which such modeling technique may be exploited to perform Stress Testing on banking institutions.

Bayesian Networks are deemed to be especially useful since, through the use of specific learning algorithms, they may be able to describe which variables are causing the value of, say, a bank's PD. Moreover, to possibly increase the precision of such an instrument, one may exploit expert knowledge (if available) to obtain more accurate results.

Therefore, by providing a straightforward and intuitive output, Bayesian Networks may be adequate in supporting the Stress Testing practice for the following reasons:

- They allow the description of causal relations in terms of probability and temporal ordering (Gao, Mishra and Ramazzotti, 2017), hence permitting the estimation of causal structures at different time instants for specific financial indicators;

- It may be possible to forecast via simulation the most probable values, expressed through confidence intervals, that a specific indicator may take in the foreseeable future;
- Bayesian Networks mirror in several aspects the aforementioned *Coherent-Stress-Testing Approach*, by allowing the identification of root and intermediate causes for a specific indicator;
- They include some features of previously-mentioned ECB method utilized for assessing the healthiness of a credit institution, in that Bayesian Networks allow the selection of the most relevant macroeconomic and firm-specific variables affecting a specific financial indicator, as well as evaluating the effects of changes in such model's variables.

Although Bayesian Networks seem to be useful and accurate tools, it is always important to keep in mind that they may only provide an approximate description of reality, given they concentrate uniquely on a rather specific selection of risk factors. Indeed, by considering the realization of exceptional events like crises, market and financial conditions would likely be altered and chances are that, depending on the severity and nature of the downturn, predictive models may fail to account for certain aspects and events (Rebonato, 2017). Therefore, one may be wary and careful in interpreting the model's output by itself. It would be more adequate, instead, to contextualize those results and complement them with insightful expert knowledge on the phenomena surrounding a specific financial indicator.

In general, despite a few limitations that Bayesian Networks possess (like any other model would), they probably are a rather useful and somewhat flexible tool to be exploited in the description of various phenomena: their most important feature is the possibility to identify key links connecting variables, possibly on different layers of causal relations, which ultimately affect a specific target financial indicator.

Notice that Bayesian Networks may not be exploited uniquely as a tool to support Stress Testing analyses. Indeed, it is worth mentioning that they possess a much wider application across different fields of knowledge, among which Biology, Computer Science and Medicine, for instance (Neapolitan, 2004).

Anyway, for the specific purposes of the present thesis, the next chapter will display those features of Bayesian Networks that are relevant to the implementation of a financial Stress Testing analysis relative to a credit institution, as demonstrated also in the empirical project application presented in the third chapter.

2 Bayesian Networks

This section will be dedicated to a brief review of the fundamental elements relating to Bayesian Networks (simply BN). Exposing the key notions on BNs is deemed to be a necessary prerequisite to the understanding of the analysis carried out in the following chapter.

A BN is composed of two core elements, one of qualitative and the other of quantitative nature. The former is the structure of the network (the so called *Directed Acyclic Graph*, as explained below), while the latter represents the probabilistic side of the net, relative to the nodes of the BN: in case of discrete data, this element translates into the use of *Conditional Probability Tables* (CPTs), meanwhile for continuous data, *Gaussian probability distributions* are assumed for each node in the network. Hence, a BN is deemed to be included within the class of *Probabilistic Graphical Models*.

It is also important to stress that the nodes (or vertices) of the network actually correspond to the variables that are meant to be part of the model. Moreover, the lines (*edges*) connecting the nodes may indicate the presence of specific dependence relations between variables, which, in some cases, may also be interpreted in causal terms.

Essentially, a BN can be exploited to define probability distributions in a compact manner, distributions which may otherwise be potentially complex to display. It is possible to achieve such representation since the characteristics of BNs allow for the specification of relations between and among variables via the definition of the relative probabilistic conditional dependencies and independencies (Darwiche, 2010). Moreover, the main purpose of a BN would generally be to model beliefs expressed in probabilistic terms and, by analyzing its output, take appropriate decisions on a specific matter.

After providing the general idea and key elements of a BN, it is necessary to analyze in further detail and with more thorough specifications the core matter. In order to accomplish that, the next subsections will be dedicated to the following topics:

- Analysis on the graphical features of BNs;
- Bayes Theorem and some Probability notions;
- The link between graphical and probabilistic properties of BNs and further details on the related key concepts;
- Bayesian Networks and their characteristics;
- Continuous Bayesian Networks.

2.1 Graphical Properties

As already mentioned, BNs belong to a specific class of Graphical Models, more specifically *Probabilistic Graphical Models*, which allow to clearly display conditional and joint probabilities of a dataset composed by multiple variables, which often are identified as random variables $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$. To be clear, the overall distribution defined by $\boldsymbol{X}$ is the *global distribution* (also known as the joint probability distribution of the variables), while each single $X_i \in \boldsymbol{X}$ (with $i = 1, 2, \ldots, n$) may be defined as being the distribution of each random variable linked to the data, in other words, a *local distribution* (Scutari and Denis, 2015). Notice that, in the construction of the BN, such local distributions relative to a variable correspond to the so called *Conditional Probability Table* of that node, for discrete data, and *Gaussian probability distributions*, in the continuous data instance.

Before further exploring the properties of BNs, a brief recap on Graph Theory is needed.

A fundamental concept which must be explained is the one of a *Directed Acyclic Graph* (DAG), which may be denoted by $G = (\boldsymbol{V}, A)$ notation.

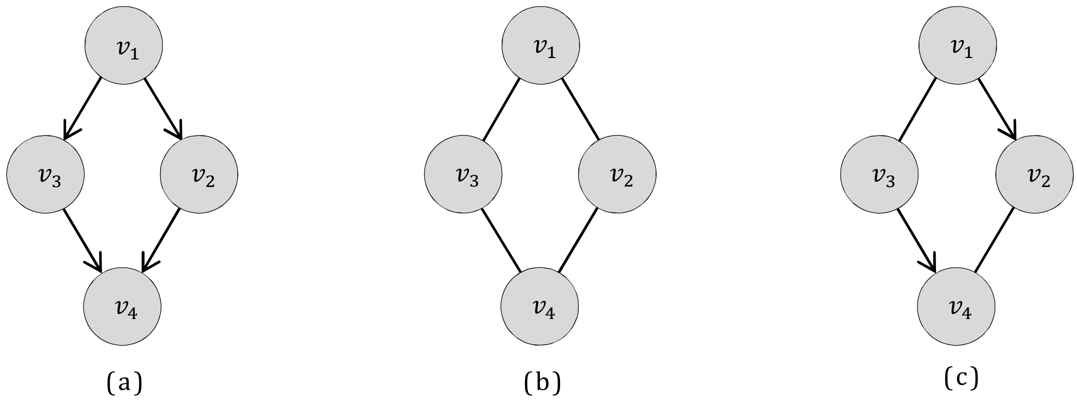


Figure 1: (a) The graph on the left represents a *directed* DAG, (b) the one in the center an *undirected* one and (c) the DAG on the right is of a *mixed* type.

In this context the nodes of the graph are connected to the single variables X_is of the model; therefore, they are identified as $v_i \in \boldsymbol{V}$ (with $i = 1, 2, \dots, n$). To be more specific, a DAG is a peculiar graphical structure concept which may be understood by simply expanding on the words *Directed* and *Acyclic*: the first, is linked to the fact that the arcs (or edges) may be given a direction pointing to one of the nodes which they connect, in order to identify potential causal relations; the second one, concerns the fact that there is no cyclicality in the graphical structure, meaning that edges "departing" from a node are not allowed to create a directed path through other vertices that leads back to that same original node. As a consequence, there exist nodes which are the origin of the graph having no incoming arcs (known as *root nodes*) and others that terminate the graph with no outgoing arcs (known as *leaf nodes*). For instance, in Figure 1, the root node is v_1, while the leaf node is v_4 and this same concept can be applied to more complex networks.

In a more formal way, by considering an arc $a \in A$ which connects two vertices v_i and v_j (with $i = 1, 2, \dots, n-1$ and $j = 2, 3, \dots, n$) denoted by $a = (v_i, v_j)$, in the event that v_i is the *tail*, and v_j is the *head*, it is possible to assert that the arc is *directed* $(v_i \to v_j)$. Furthermore, if the interpretation of causality is deemed to be valid, v_i would be seen as the cause, while v_j would be the effect, otherwise, only a dependence relationship would persist. In case there is no directionality, the edge

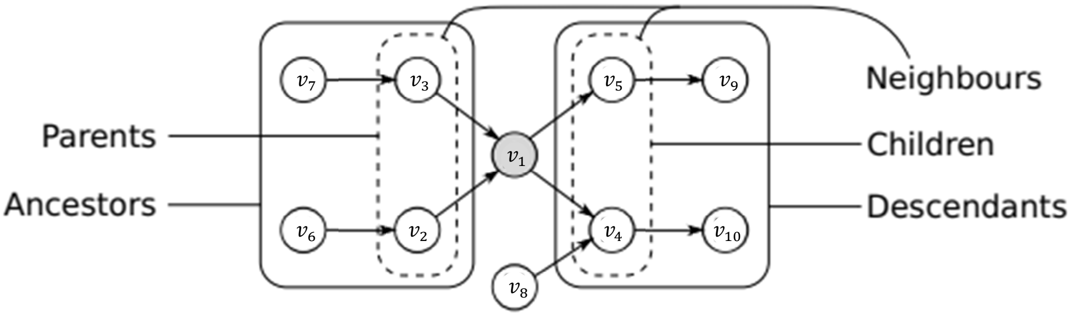


Figure 2: Graphical display of the relations forming the sets of Ancestors, Parents, Neighbors, Children and Descendants [source: Nagarajan, Scutari and Lebre (2013)].

would be defined as *undirected* $(v_i - v_j)$. In the former instance, the edges of the graphical model which connect the various nodes, denoted by $a \in A$, serve as a way to identify the dependence or independence relationships between the vertices and, hence, the variables. In the latter case, instead, edges would be denoted by $e \in E$ and the corresponding DAG notation would be $G = (\boldsymbol{V}, E)$ and corresponding to an *undirected graph*. For the sake of completeness, it is worth mentioning the existence of another type of graph, which would be the *mixed* one: $G = (\boldsymbol{V}, A, E)$.

Moreover, direct dependence relations between variables are put in evidence by the directionality of the arcs. Notice that indirect dependencies are not depicted in graphs, however, they may be grasped by looking at the sequences of arcs (paths) that connect non-adjacent nodes. Therefore, an indirect dependence exists in the event that between a couple of variables (nodes) there exist intermediate variables (nodes).

To better understand the above-mentioned concepts, Figure 1 displays graphical examples of the different DAGs.

The way arcs are defined on a graph, determines the so called *structure* of the graph, which depends on the presence of arcs and their relative directionality (if any). In turn, the structure of the graph defines its own properties (which later will be analyzed in further detail). Recalling the concepts of *root* and *leaf* nodes, it is key to define other notions relative to each single node in the network. For instance, considering what is depicted in Figure 1(a), on the one hand, vertex v_1 can be defined as being the *ancestor* of v_2, v_3 and v_4, in other words, the node that precedes the others in the ordered sequence of arcs. On the other hand, v_2, v_3 and v_4 are defined as *descendants* of v_1. Furthermore, v_2 and v_3 are the *parents* of node v_4, like v_1 is the parent of both v_2 and v_3. Moreover, v_2 and v_3 are the *children* of v_1, like v_4 is the child of both v_2 and v_3. Therefore, the concepts of *parents* and *children* are related to nodes which are located in the graph close to each other (*neighbors*). At this point it is clear that the parents are a subset of the ancestors and the children are a subset of the descendants. These concepts are made clearer by the graph reported in Figure 2.

To expand a little more on the way a BN may be constructed, other than its elements, it must be said that there exist approaches in the building of the network. In fact, it is usually the case that the BN reflects the knowledge or beliefs of some experts on the matter object of the analysis. In such case, the aim would be to try capturing the relations between variables, relations which may be established by causality, and the network would be a so called *expert network*.

Another way to determine a BN would be to define it according to stricter rules with the purpose, for instance, of determining a network to monitor the functioning of an information system (Darwiche, 2010). Another popular method worth mentioning is to build up the network by *learning* its structure from the data (as explained in section 2.5.1).

In particular, an expert network and structures *learnt* from data will be the instruments utilized to carry out the case study analysis in the following chapter.

2.2 Probability Concepts

One of the core elements relating to BNs surely is Bayesian Probability, therefore, a brief review is deemed necessary in order to recap the key concepts.

It is important to stress the difference that stands between *frequentist* and *subjective* approaches in the way probability is defined. Relative to the former case, in essence, the frequency (meaning the number of times out of the number of trials) of an event happening determines the probability to be assigned to that specific event. In the latter approach instead, with particular reference to the use of conditional probabilities, a set of subjective assumptions and knowledge apply to the way the desired probability is obtained.

Focusing on the subjective approach, which is relevant to BNs, it may be inferred that initial assignments of probability to a specific event U actually imply the definition of a conditional probability statements like $P(U|W)$ (or more simply $P(U)$), where W identifies the knowledge on the context of that specific event (Fenton and Neil, 2013). Consequently, this kind of probability is determined by way of personal or expert *beliefs*, which may be informative in a setting where no events are yet to be observed. Such principle is reflected on the notation $P(U)$ which may be corresponding to the notion of *prior belief* about an event U. Moreover, by utilizing evidence K on the actual event U it may be possible to update the prior belief and obtain a *posterior belief* on U denoted by $P(U|K)$. In order to compute this probability in terms of $P(U)$ it would be necessary to utilize the joint probability $P(U \cap K)$, which often is unknown. What is usually known, instead, is the *likelihood* of K, that is $P(K|U)$. The aim is to update beliefs in order to compute the following conditional probability:

$$P(U|K) = \frac{P(U \cap K)}{P(K)}. \qquad (2.1)$$

In order to solve the issue concerning the joint probability in formula (2.1), the Bayes' Rule comes to the rescue by allowing the computation of $P(U|K)$ utilizing the *likelihood*:

$$P(U|K) = \frac{P(K|U)\,P(U)}{P(K)}. \qquad (2.2)$$

The principle built in the Bayes' Rule is not directly visible in BNs, however, it is applied for the quantitative side of networks and may be used to calculate conditional probabilities relative to nodes where necessary. This is connected to the afore-mentioned concept of CPTs and Gaussian probability distributions, each relating to a specific vertex and carrying the probability distribution of a specific node which is conditioned upon the node's parents' probabilities, made exception for the case in which the node in question would be a root node. In this sense, the CPT or the Gaussian probability distributions for a specific vertex aims at capturing the strength of the relationship between the vertex and its parents (Fenton and Neil, 2013).

Concerning the probability distribution linked to a specific node, it is useful to define the so called *marginalization* of variables (which in turn will be linked to the *chain rule* topic as reported in section 2.4), in case one wishes to extract marginal probability values from a joint probability distribution. In brief, by considering for instance two variables X_1 and X_2, marginalization on X_1 may be performed by applying the following principle:

$$P(X_1) = \sum_{X_2} P(X_1|X_2)\, P(X_2) \tag{2.3}$$

Furthermore, notice that

$$\sum_{A} P(X_1|X_2) = 1 \tag{2.4}$$

Hence, given formulas (2.3) and (2.4), the marginalization principle may be similarly applied on joint distributions with more than two variables.

2.3 Graphical Structure and Probability

In order to investigate whether dependence relations are present between nodes, it is important to distinguish between the notions of *graphical* ($\perp_G$) and *probabilistic* ($\perp_P$) independence. In the former case, it is possible to observe independence, for instance, between two nodes that are separated by no arc (also defined as *conditional independence* in probabilistic terms). In the latter case, instead, independence is rooted in the probabilistic relation between variables. As it turns out, these two types of independencies are actually linked in a way that the structure of the BN may influence their determination (Scutari and Denis, 2015).

In a more formal way, suppose M to be the structure defining dependencies in the probability distribution P of the variables in $\boldsymbol{X}$. Define also $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ to be disjoint subsets of $\boldsymbol{X}$. It is possible to infer that a graph G is a *dependency map* of M in case there is perfect correspondence between $\boldsymbol{X}$ and the vertices in $\boldsymbol{V}$, so that

$$\boldsymbol{A} \perp_P \boldsymbol{B}|\boldsymbol{C} \Rightarrow \mathbf{A} \perp_G \boldsymbol{B}|\boldsymbol{C} \tag{2.5}$$

In case G would be and *independency map* of M, then

$$\boldsymbol{A} \perp_P \boldsymbol{B}|\boldsymbol{C} \Leftarrow \mathbf{A} \perp_G \boldsymbol{B}|\boldsymbol{C}. \tag{2.6}$$

The graph G may be defined as a perfect map of M in case it is both a *dependency* and an *independency map*, meaning that there is independence under both a structural and a probabilistic point of view. Then

$$\boldsymbol{A} \perp_{\boldsymbol{P}} \boldsymbol{B}|\boldsymbol{C} \Leftrightarrow \boldsymbol{A} \perp_{\boldsymbol{G}} \boldsymbol{B}|\boldsymbol{C} \tag{2.7}$$

and, for this specific instance, G would be *faithful* to M (defining the *faithfulness* condition).

Notice that in the first two cases the mapping of independence is not necessarily the same between the structure of the graph and the probability distribution underlying the variables. However, it is important

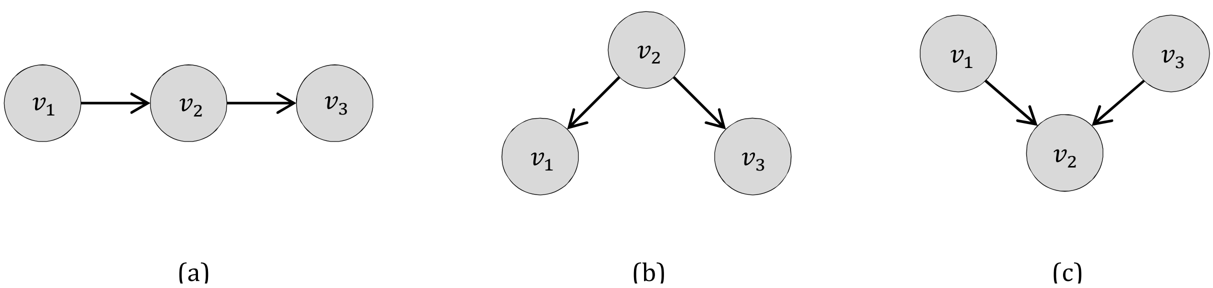


Figure 3: (a) Serial connection; (b) divergent connection; (c) convergent connection (or $\boldsymbol{v}$-structure in such case).

to stress that in the case in which G is faithful to M, separation in the graphical network and conditional independence relative to the variables perfectly correspond (Scutari and Denis, 2015).

2.3.1 *d*-separation

With reference to the concept of graphical separation, the key principle that stands behind it is *d-separation* (directed separation).

As before, by considering $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ to be disjoint subsets of $\boldsymbol{X}$ for a DAG G, then, the subset $\boldsymbol{C}$ d-separates $\boldsymbol{A}$ and $\boldsymbol{B}$. As previously seen above, the relative notation is $\mathbf{A} \perp_G \boldsymbol{B}|\boldsymbol{C}$. The condition of *d-separation* is satisfied only if in between any path that leads from the vertices in $\boldsymbol{A}$ to the ones in $\boldsymbol{B}$ there exist nodes like v, having the following properties:

- Arcs converge to v and such node and its relative descendants do not belong to $\boldsymbol{C}$;
- v belongs to $\boldsymbol{C}$ but converging arcs pointing at it do not exist.

In support of the definition relating to *d-separation*, it is necessary to describe the possible configurations which may apply to three vertices and two edges, known as *fundamental connections*, which are utilized when constructing a BN.

Consider for example the existence of the three vertices v_1, v_2 and v_3. Then:

- A *serial connection* is a structure so that $v_1 \rightarrow v_2 \rightarrow v_3$;
- A *divergent connection* would be $v_1 \leftarrow v_2 \rightarrow v_3$;
- A *convergent connection* is identified by $v_1 \rightarrow v_2 \leftarrow v_3$. Moreover, in case the parents (v_1 and v_3) are not liked by an edge, the connection would be a *v-structure*.

To better appreciate the various definitions relating to the different types of connections, Figure 3 displays the three instances described above.

2.4 Defining Bayesian Networks

By knowing how to define and classify connections between vertices, keeping into consideration the probability distributions of the underlying variables, it may be possible to provide a definition for a BN:

> Provided there is a probability distribution P relative to a set of variables $\boldsymbol{X}$, we may define a DAG $G = (\boldsymbol{X}, A)$, where A identifies the set of arcs connecting the nodes of the network, to be a BN denoted by $B = (G, \boldsymbol{X})$ if and only if the graph G is a minimal independency map, in the sense that, in case any arc is removed, the network would not be an independency map anymore (Scutari and Denis, 2015). Therefore, it is clear that graphical separation entails probabilistic independence; still, not all conditional independencies from a probabilistic standpoint are depicted on the network graphically.

From the definition of BN it is necessary to expand and provide further details on the properties of these types of networks.

Firstly, by knowing that a BN is also an independency map, it follows that the decomposition of the joint probability distribution $Pr(\boldsymbol{X})$ is applicable and defined as follows:

$$Pr(\boldsymbol{X}) = \prod_{i=1}^{p} Pr(X_i | \prod X_i) \tag{2.8}$$

where $\prod X_i$ is identified as being the set of parents of X_i. Another way to express the joint probability would be through the *chain rule*:

$$Pr(\boldsymbol{X}) = \prod_{i=1}^{p} Pr(X_i | X_{i+1}, \dots, X_p) \tag{2.9}$$

which allows to simplify the decomposition of the variables in the event that the graph would be acyclic. In general, however, formula (2.8) is to be preferred as first choice.

Secondly, BNs do possess the so called *Markov Property* (an application of the chain rule in formula (2.9)) which allows to represent a joint probability distribution of the variables as a product of conditional probability distributions (Nagarajan, Scutari and Lebre, 2013).

Recalling the previously defined *fundamental connections* and blending those concepts with the ones of the *Markov Property*, it is possible to provide an exemplification of the chain rule formula. In order to accomplish that, knowing there is a direct correspondence between nodes $v_i \in V$ and the underlying variables $X_i \in \boldsymbol{X}$ ($v_1 \leftrightarrow X_1$; $v_2 \leftrightarrow X_2$; $v_3 \leftrightarrow X_3$; …):

- For the *serial connection* in Figure 3(a) and the *divergent connection* in Figure 3(b), the DAGs may be defined as *independency maps* and, hence, the following formulas apply respectively:

$$P(X_1, X_2, X_3) = P(X_3|X_2)P(X_2|X_1)P(X_1)$$

$$P(X_1, X_2, X_3) = P(X_1|X_2)P(X_3|X_2)P(X_2)$$

- Relating to the *convergent connection* (*v-structure*) in Figure 3(c), node v_2 violates the conditions for *d-separation*, therefore, vertices v_1 and v_3 are not independent from v_2:

$$P(X_1, X_2, X_3) = P(X_2|X_1, X_3)P(X_1)P(X_3)$$

From the example just presented, follows that serial and divergent connections do have the same factorizations in terms of conditional probabilities, then, they are deemed to be Markov Equivalent Structures, which together form a so called equivalence class: DAGs with the same underlying structure and v-structures are said to be equivalent. From what asserted so far, it is evident that v-structures differ from the other types of connections, in particular, under graphical and probabilistic independence standpoints (Scutari and Denis, 2015).

2.4.1 Markov Blankets

In connection to the *d-separation* topic, a *Markov Blanket* is defined to be a set of vertices d-separating a given vertex from the remaining part of the network (Nagarajan, Scutari and Lebre, 2013). Therefore, given a specific node, the *Markov Blanket* is composed by the node's parents, children and the nodes with which there are children in common.

In a formal way: the *Markov Blanket* of a node $v_i \in \boldsymbol{V}$ (with underlying variables $X_i \in \boldsymbol{X}$) is that minimal subset $\boldsymbol{S}$ in $\boldsymbol{V}$ so that

$$X_i \perp_G \boldsymbol{V} - \boldsymbol{S} - X_i | \boldsymbol{S} \tag{2.10}$$

and in case the faithfulness assumption is valid, then

$$X_i \perp_P \boldsymbol{V} - \boldsymbol{S} - X_i | \boldsymbol{S}. \tag{2.11}$$

It is also possible to assert that a *Markov Blanket* includes all those nodes ($\boldsymbol{S}$) that make all the other ones not necessary when inference is performed on a given vertex, say v_i. In this regard, it is usually the case that parent nodes values are the ones majorly affecting v_i's value, especially when causal relations are valid.

2.5 Bayesian Networks with Continuous Data

On the basis of the empirical case study in the next chapter, it is useful to focus on defining the elements to adopt when building a *Gaussian Bayesian Network* (GBN). A GBN is a BN built to model continuous data assuming the relative underlying joint probability distribution would be a multivariate Gaussian one. Hence, it follows naturally that, instead of *CPTs*, each node's values would be associated to a Normal distribution conditioned on the parent's values. In turn, this translates into defining a series of multivariate regressions, one for each node present in the network. In fact, every vertex would have an associated value (the response variable of the regression, which explanatory variables are the parent's values) and variance (the error term of the regression), that together are respectively the first and the second moment of the relative Gaussian distribution for that specific node.

More formally, GBNs may be defined by the following key features (Scutari and Denis, 2015):

- Each vertex follows a Gaussian distribution;
- Root nodes are defined by the relative marginal distributions;
- The intercept term and explanatory variables in the regression for each node define the conditioning effect on that particular node and, moreover, these elements do not affect the variance that is indeed specific of the node considered;
- Nodes' local distributions are defined by Gaussian linear regressions, with no interaction terms for the explanatory variables;
- The global distribution of all the nodes in the network corresponds to a multivariate Gaussian and, as already-mentioned in previous sections, such joint distribution is the product of all the single local distributions.

Furthermore, by considering data D and the BN $B = (G, \boldsymbol{X})$ and denoting the parameters associated with $\boldsymbol{X}$ by $\boldsymbol{\Theta}$, it is possible to define the parametric distributions used to describe data D and the relative BN as $B = (G, \boldsymbol{\Theta})$ (Scutari and Denis, 2015). Then, the entire learning process for structure and parameters would be summarized by $P(B|D) = P(G, \boldsymbol{\Theta}|D)$, which also corresponds to:

$$P(G, \boldsymbol{\Theta}|D) = P(G|D)\, P(\boldsymbol{\Theta}|G, D) \tag{2.12}$$

where the right-hand side first element represents the process of structure learning from data, while the second one is the subsequent step dedicated to parameter estimation.

In the following, the key steps of *Structure Learning* (inferring the structure for a DAG, that is, setting the arcs and their relative direction) and *Parameter Learning* (estimate local distributions on the basis of the previously-learned structure) will be described, as they represent important steps in the definition of a Network. The focus will be on the notions connected to GBNs; however the following underlying principles will be also valid for all other types of BNs.

2.5.1 Structure Learning

This is the first and probably most crucial step in the definition of a BN, given it defines whether relations between nodes should really exist and which direction the estimated arcs should have, not to mention that it will set the basis for the parameter estimation step.

In particular, structure learning may be performed by utilizing the Bayes' theorem to define the components of the posterior probability relative to the graph, that is

$$P(G|D) \propto P(G)P(D|G) \tag{2.13}$$

where $P(G)$ represents the prior distribution carrying information on dependence between nodes, while $P(D|G)$ is the probability of the data given the graph (Scutari and Denis, 2015).

It is usually the case that $P(G)$ is chosen to allocate the same probability to all the DAGs it represents (*non-informative prior*), including information on the conditional dependencies of the network's variables $\boldsymbol{X}$.

Concerning $P(D|G)$, instead, by decomposing the function as

$$P(D|G) = \int P(D|G, \boldsymbol{\Theta})P(\boldsymbol{\Theta}|G)d\boldsymbol{\Theta} \tag{2.14}$$

it may be proven that a further equivalent representation of equation (2.14) is the following:

$$P(D|G) = \prod_{i=1}^{p} E_{\Theta X_i}[P(X_i|\prod_{X_i})] \tag{2.15}$$

where p represents the number of nodes, and therefore variables, present in the network. For a more detailed derivation of equation (2.15), see Scutari and Denis (2015).

As it turns out, in case all the expected values in equation (2.15) may be computable in closed form, it would be possible to get a measure of $P(D|G)$ in relatively short computational time. This principle is valid for the already-mentioned multivariate Normal distribution characterizing GBNs and $P(D|G)$ may be computed through the relative conjugate[4] *Inverse Wishart Distribution*[5]. Moreover, the relative posterior distribution may be defined as the so called *Bayesian Gaussian equivalent uniform* (BGeu or BGe), which is also to be seen as a type of Network Score[6]. Notice that the assumptions on BGe include a non-informative prior on both DAGs $[P(G) \propto 1]$ and parameters of each vertex $[P(\Theta_{X_i}|\prod_{X_i}) = \alpha_{ij} = \alpha/|\Theta_{X_i}|]$. Moreover, the parameter belonging to the BGe is the imaginary sample size α which is connected to the prior and the higher its value, the more weight is assigned to such prior (Scutari and Denis, 2015).

Alternative ways of computing $P(D|G)$ reside in the use of Conditional Independence Tests[7] or other Network Scores, to approximate its value.

For the sake of the analysis carried out in the following chapter, where BGe score will be utilized, it would be more appropriate to spend a few more words on the latter method, just to review the definition of the *Bayesian Information Criterion* (BIC), as an alternative approximate Network Score methodology (Scutari and Denis, 2015):

$$BIC(G,D) \rightarrow \log[BDe(G,D)] \qquad \text{as sample size } n \rightarrow \infty. \tag{2.16}$$

---

[4] The concept linked to the term *conjugate distribution* comes from Bayesian Statistics and, in general, applies in case the *posterior distribution* belongs to the same probability distribution family of the *prior distribution*.

[5] An *Inverse Wishart Distribution* is the multivariate extension to the *Inverse Gamma Distribution* and has the following general form $W^{-1}(\boldsymbol{\Psi}, v)$:

- $\boldsymbol{\Psi} > 0$ and represents a $p\,x\,p$ scale matrix, that is, a sum of squares matrix belonging to a multivariate Gaussian distribution;
- the v parameter corresponds to $v > p - 1$ degrees of freedom.

For more on this topic, see Schuurman, Grasman and Hamaker (2016).

[6] Score which may be compared to the ones of other estimated structures to set a ranking of the different potential DAGs that may characterize data.

[7] Just to mention, in case of continuous data, an appropriate Conditional Independence Test would be to utilize *partial correlations* to verify whether the null hypothesis of conditional independence between two variables, given the value of the relative parents, is to be accepted or rejected.

Moreover, it is possible to perform a decomposition of the BIC as

$$BIC(G, D) = \sum_{i=1}^{p} \{\log[P(X_i | \prod_{X_i})] - (|\Theta_{X_i}|/2)\}. \tag{2.17}$$

In addition, if *causal* BNs are the tool utilized for the analysis, it is important to provide and review some further specific notions.

Firstly, in case expert knowledge is available on a determinate phenomenon, there would be no issues in setting causal connections between variables (nodes).

Secondly, if instead the *causal* BN structure is to be learnt from data, additional assumptions are needed (Nagarajan, Scutari and Lebre, 2013):

- *Causal Markov Assumption*, that is, each variable $X_i \in \boldsymbol{X}$ is conditionally independent of its direct and indirect non-effects, given the relative direct causes. Also, notice that, the relations expressed by the directionality of edges in the BN indicate causal connections between variables, where, in general, parents (or in some cases root nodes) are identified as the causes of their neighboring descendants.
- A network must exist so that it is *faithful* to the probability distribution of $\boldsymbol{X}$, in order to obtain probabilistic causal dependencies which are generated by *d-separations* in the DAG;
- Unobserved variables that may influence those variables that are actually present in the BN must not exist, otherwise the *faithfulness* condition and *Causal Markov Assumption* may be violated.

Still, in real-world settings there may be a considerable number of connections and correlations between and among variables, relations which if not captured by the model would correspond to unobserved (*latent*) variables. Hence, rather than assuming that causal relationships described by the network correspond to the real world connections, it is more appropriate to keep in mind that, given a context of application, the variables considered in the BN would identify causal relations specific to that particular instance and only approximate reality (Neapolitan, 2004).

To perform the process of structure learning in the context of BNs, a number of a*lgorithms* are available, each tailored to the problem that needs to be addressed and specific to the type of data at disposal.

Broadly speaking, it is possible to distinguish among three learning algorithms types: *Constraint-based*, *Score-based* and *Hybrid*. Notice that, all of these are subject to some assumptions (Scutari and Denis, 2015):

- A one-to-one correspondence between vertices of the network and variables in $\boldsymbol{X}$ must exist;
- Conditional independencies must exist between variables in $\boldsymbol{X}$;
- The DAG structure must reflect actual dependencies between all possible combinations of variables in order to ensure the uniqueness of Markov Blankets and, in turn, of the BN structure;
- The observed values are regarded as being independent realizations for variables coded by network vertices.

For the purposes of the analysis carried out in chapter 3, among all possibilities, the focus will be on the *Hill Climbing* (HC) *Algorithm*.

In general, HC belongs to the class of *greedy search algorithms*, which perform structure search by starting from a set of nodes and gradually adding, erasing or reversing arcs in the network. This process continues until it is no more possible to improve the network's score; notice that such element allows HC to be classified also as a Score-based algorithm, which performs structure optimization via heuristics like BGe or BIC for instance. The pseudo-code relative to the HC algorithm follows (Nagarajan, Scutari and Lebre, 2013):

Hill-Climbing Algorithm

1. Choose an empty DAG G (with no arcs) over the set of variables $\boldsymbol{X}$
2. Calculate the score of G according to a defined scoring criteria and denoted by $Score_G$
3. Set the maximum score as $maxscore = Score_G$
4. Loop the following as $maxscore$
   a) For every possible addition, erasure or reversal of arcs that do not make the DAG a cyclic network:
      i. compute the score of the modified DAG G^*, that is $Score_{G^*}$
      ii. if $Score_{G^*} > Score_G$, set $G = G^*$ and $Score_G = Score_{G^*}$
   b) update $maxscore$ with the obtained value for $Score_G$
5. Return the DAG structure G

The way in which the HC algorithm operates, dependent on the specific problem and the relative data at hand, may result in a DAG with excessive and non-coherent directed arcs. Moreover, in case no information is present on the underlying data structure, the estimated DAG may not be able to reveal clear relationships between and among variables, hence, being uninformative.

A solution to the above-mentioned problem may be to utilize other types of learning algorithms (like for instance *Grow-Shrink* or *PC*) and/or different scoring criteria to estimate the DAGs. Subsequently, it may be possible to compare all of them to identify the arcs that are most recurring across the structures, so that one may infer the presence of relationships between specific nodes. Indeed, this approach may be identified as a kind of *model averaging* procedure.

A similar but alternative approach may be to perform multiple simulations by utilizing a single type of algorithm and eventually "average" the entirety of the obtained structures.

In case the underlying data structure and relationships are not totally obscure to the analyst, it may be possible to exploit expert information to compel the algorithm to either set determinate arcs (*white list*) or prohibit the presence of specific relations (*black list*). In such a way the algorithm may output more informative DAGs, relative to data considered in the analysis. By utilizing this approach the algorithm may be classified under the *Constraint-based* class of algorithms.

As it turns out, depending on the specific features (scores and constraints) that one may utilize to characterize an algorithm, the output may significantly vary. The choice is of course based on the purposes of the analysis and the type/quantity of data available.

Furthermore, notice that in some instances, as performed in the case study analysis, it could be appropriate to exploit *Hybrid* algorithms, which are characterized by both constraints and scoring criteria features.

It is worth only mentioning that in presence of data observed across time, one approach may be to estimate a *Dynamic Bayesian Network* structure in order to account not only for relations among variables in a precise temporal moment, but also for those relationships that exist between variables over time. Notice that such method has not been implemented in the case study analysis. For more on *Dynamic Bayesian Networks*, see Nagarajan, Scutari and Lebre (2013).

A different approach may be to set up a learning procedure aiming at discovering the underlying (causal) structure linking data and persisting through time. To accomplish this, a possible technique could be similar to the one described below:

Step (1): Perform repeated simulations on each moment in time $(t_0, t_1, t_2, \ldots, T)$ and obtain a representative network structure for each of those moments.

Step (2): Average all the obtained structures $(DAG_{t_1}, DAG_{t_2}, \ldots, DAG_T)$ to estimate a guiding underlying structure (DAG^*) that may be valid and invariant across time.

Once the structure of the underlying network has been either learnt or defined through expert knowledge, the following step would be to perform inference on the parameters of the network.

2.5.2 Parameter Estimation

The last step in the learning process consists in implementing inferential procedures to estimate the joint probabilistic distribution of the BN. In order to do that, it would be appropriate to estimate the distributions associated to each node of the network obtained in the previous step. Then, by obtaining the statistics relative to each node, it would be possible to assess the strength of the dependence relationships (possibly causal) between that node and its parents (Scutari and Denis, 2015). Hence, by simply looking at the estimated parameters of the BN, one may be able to infer which (causal) links connecting nodes are the most relevant.

Given GBNs are the designated models for the analysis, nodes' local distributions would be defined by Gaussian linear regression models: the response variable would be identified by the node itself and the regressors would correspond to its parents (Scutari and Denis, 2015). Notice that, in practice, one approach to parameter inference is Maximum Likelihood Estimation (MLE)[8]. In general, Bayesian estimation of parameters would only be employed in case new evidence (data) becomes available. For this reason, the initial estimation of parameters is usually performed via MLE.

It would also be useful to define the Confidence Intervals (CIs) relative to each estimated coefficient, in order to define a range over which the value of the actual parameter may fall in. Furthermore, in predictive terms, it may be appropriate to obtain prediction CIs through the use of standard error measures like the Mean Squared Error or the Mean Absolute Error. Alternatively, one may perform multiple simulations on the prediction range to obtain a different indication of the area over which series' values are most likely to be observed in the future.

Lastly, when estimating parameters, it is important to remember that the results of inferential procedures are highly dependent on the data at hand and especially on the relative sample size (Scutari and Denis, 2015).

It is worth mentioning that other topics relative to parameter estimation do exist and apply to various different settings[9]. However, with regard to the empirical analysis that will follow in the next chapter, the

---

[8] In general, MLE is a technique which aims at finding parameter values θ so that the likelihood function $\boldsymbol{L}(\theta; x)$ is maximized. Notice that oftentimes the natural logarithm of the maximum likelihood function $ln[\boldsymbol{L}(\theta; x)]$ is utilized for convenience. Then, the resulting MLE parameter estimate would be $\hat{\theta} \in arg\max\boldsymbol{L}(\theta; x)$.

[9] For instance, in presence of evidence, one may be able to implement inferential queries to investigate differences in the obtained parameter estimations. Such procedure is also known as *Belief Updating* (which may be either *exact* or *approximate*). Notice that evidence may be of two types: *hard*, in case it is possible to dispose of new observations for variables instantiation, or *soft*, in case new distributions for the network's variables are available.

For more details on this topic see Scutari and Denis (2015).

notions provided in the current section are sufficient for an early overview and understanding of the concepts that will be applied on the data in the case study analysis.

This concludes the overview on the properties, learning procedures and inference that relate to BNs, with a particular focus on approaches used when continuous data are analyzed. Such notions serve as a preliminary step necessary to the understanding of computational methods applied in the following chapter concerning the empirical analysis.

For more in depth notions and further topics on the use of BNs, see Neapolitan (2004), Scutari and Denis (2015) and Nagarajan, Scutari and Lebre (2013).

3 Estimating a Bank's Probability of Default

The present chapter is focused on explaining and reporting the methodology and the results obtained from the use of a Bayesian Network to address the needs of a Stress Testing analysis application, based on actual data relative to a specific credit institution (identified by the expression *The Bank* in the following paragraphs and sections).

Aim of the analysis would be to provide an indication on *The Bank*'s Probability of Default (*PD*) up to 1 year ahead from the last observations available in the database. Furthermore, a 1-year-ahead assessment will also be provided for the variables deemed to be closely related and directly affecting *PD*. Therefore, differently from a wide-scope Financial Stress Testing (including the assessment of various elements as described in the first chapter), the following analysis will be rather focused on the *PD* component and the elements related to it.

All the computations were performed by utilizing specific packages of the *R* Software Environment (www.r-project.org). *APPENDIX B* provides more information relative to *R* and the packages utilized in the analysis.

3.1 Data Description

In the first place, data utilized to develop the Financial Stress Testing analysis and further descriptions concerning the variables employed are presented. Notice that raw data for the whole database have been directly extracted/defined by *The Bank*; furthermore, variables employed in the analysis have been determined by elaborating data, in line with the aim to perform an assessment on *The Bank*'s *PD*. To this end, among the indicators and data provided by *The Bank*, Table 1 and Table 2 show an overview of the specific components deemed to be useful for the purposes of the analysis.

Notice that, concerning the *Firm-specific Variables* for this particular instance, RWAs would include measures of risk-weighted elements owned

Variable Name	Alias	Description
Unemployment	**UN**	Average unemployment percentage level for the EU countries
Inflation Rate	**Infl**	The measure for inflation is expressed via the Consumer Price Index for the EU countries
Gross Domestic Product	**GDP**	Change in average GDP for the EU countries
Market Share of Company Assets	**Mkt.Sh**	This measure indicates the percentage proportion of assets belonging to *The Bank* concerning the relative market
Oil Price	**Oil**	Change in mean price of oil across the EU countries
Euribor 3-months	**Euribor**	The Euribor 3-months interest rate

Table 1: The *Macroeconomic Variables* utilized in the assessment of *The Bank*'s *PD* together with the relative alias (code of the element utilized in the analysis on R Software) and description.

and calculated by *The Bank* concerning *non-corporate retail assets*[10] and *corporate retail assets*[11]. More specifically, the two aforementioned categories form the RWAs relative to credit activities, which combined with the RWAs generated through operations executed by *The Bank* form the overall utilized measure for RWAs.

Concerning the time dimension of data, the available observations span between April the 1st, 2015 and January the 1st, 2018. Notice however that, some of these measures were observed monthly (*NPL*, *CoR*, *NSFR*, *Infl*, *Mkt.Sh*, *Oil* and *Euribor*) while others were observed quarterly (*CET*1, *RORAC*, *LR*, *IR*, *UN* and *GDP*).

[10] Such elements would relate to central governments/banks, regional governments and local authorities, public sector entities, credit institutions and connected financial instruments (either high or low risk).

[11] Such elements would relate to firms and corporations. On this measure the Credit Valuation Adjustment had been applied: it represents the discrepancy between risk-free and risky assets' value, taking into account the possibility for counterparties' default, hence defining the value for *counterparty credit risk*.

Variable Name	Alias	Description
Common Equity Tier 1 Ratio	**CET1**	As defined in formula (1.1) it is the ratio of CET1 capital amount over RWAs
Return On Risk-Adjusted Capital	**RORAC**	Change in Operating Income Net Taxes which has been adjusted for risk by exploiting the contribution of RWAs measure
Non-Performing Loans Ratio	**NPL**	The ratio of doubtful exposures value (in other words Non-Performing Loans) over the total value of all exposures
Cost of Risk Ratio	**CoR**	The ratio of losses and costs (associated with assets exposed to a certain degree of risk) over the average value of exposures
Net Stable Funding Ratio	**NSFR**	As defined in formula (1.3) it is the ratio of medium-term resources (medium-term financing and *The Bank*'s value of equity) over the total value of stable funding resources (net value of exposures and reserves for bad debts)
Leverage Ratio	**LR**	As defined in formula (1.4) it is the ratio of total exposures' value over the amount of Tier 1 capital
Interest Rate Risk	**IR**	It is the ratio of net exposures (in this case computed by assessing the discrepancies between changes in interest rates for assets and liabilities of different maturities) over total capital amount

Table 2: The *Firm-specific Variables* utilized in the assessment of *The Bank*'s *PD* together with the relative alias (code of the element utilized in the analysis on *R Software*) and description.

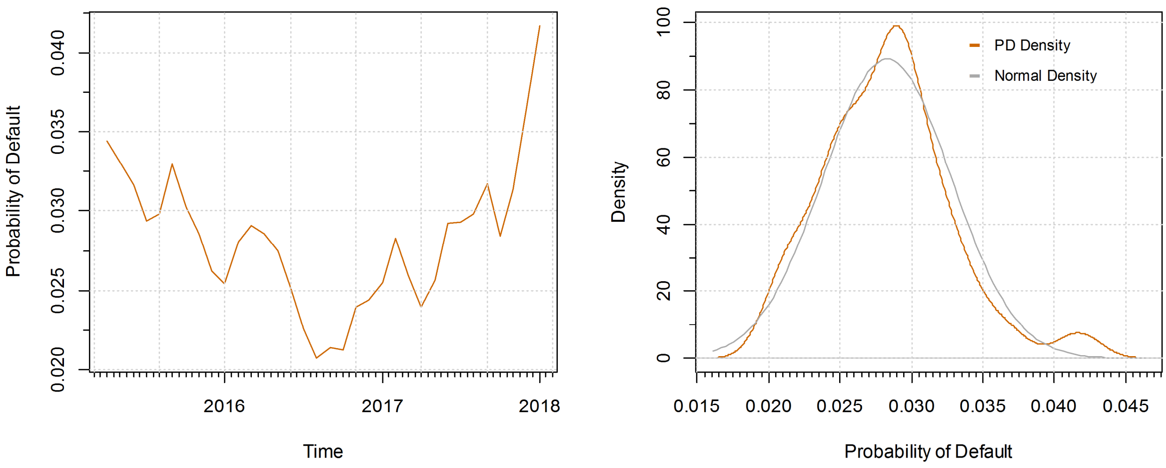
In a preliminary phase, the analysis had been performed utilizing quarterly data relative to all the variables. Nevertheless, results actually turned out to be uninformative and this fact is probably due to the scarcity of data at disposal (12 observations per variable) over a time period of almost 3 years. Therefore, the decision was to apply interpolation (more specifically the *spline* method) in order to extract *synthetic* observations from those time series possessing quarterly data. In such a way it had been possible to obtain monthly observations on all the considered variables, hence getting series of 34 time points length.

Variables	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	Maximum
FIRM-SPECIFIC VARIABLES						
CET1	0.1156	0.0041	-0.5348	3.0899	0.1062	0.1242
RORAC	0.4754	0.5952	-0.4696	2.3649	-0.7191	1.3031
NPL	0.0160	0.0025	0.3242	1.8869	0.0122	0.0213
CoR	0.0016	0.0011	0.5978	3.1796	-0.0001	0.0044
NSFR	1.1130	0.0233	-0.9009	3.8568	1.0510	1.1540
LR	0.0950	0.0033	0.2126	2.6262	0.0887	0.1017
IR	0.0007	0.0005	0.5973	2.8635	-0.0001	0.0019
MACROECONOMIC VARIABLES						
UN	0.0853	0.0081	0.0223	2.1934	0.0728	0.1020
Infl	0.0065	0.0072	0.4504	1.5858	-0.0022	0.0198
GDP	0.0081	0.0252	-0.6397	2.6323	-0.0546	0.0429
Mkt.Sh	0.0747	0.0073	0.2772	2.6056	0.0606	0.0916
Oil	0.0003	0.0264	-0.5736	3.3793	-0.0766	0.0416
Euribor	-0.0022	0.0013	0.7367	1.8580	-0.0033	0.0003
RESPONSE VARIABLE						
PD	0.0283	0.0045	0.6595	3.9142	0.0207	0.0417

Table 3: Summary statistics for the variables deemed to be influencing, either directly or indirectly, the value of *PD* and summary statistics for the baseline value of *PD*.

Moreover, recalling that the aim of the Stress Testing analysis would be to provide an indication of future values on *PD* and firm-specific variables affecting it, monthly predictions would relate to the time span between February the 1st, 2018 and January the 1st, 2019 (*prediction period*).

Notice that no raw time series of *PD* is at disposal of *The Bank* to be utilized for the starting observations of the analysis' target variable. Hence, in order to generate a monthly time series relative to the baseline implied unconditional *PD*, as explained in section 1.3.1, the IRB approach



(a) (b)

Figure 4: (a) Time series for the baseline *PD*, obtained through the *IRB Approach*, which shows an initial decrease followed by a relatively significant increase in *PD*'s value; (b) comparison of the *PD* density with the one of a Normal distribution.

had been exploited. Other than the previously-mentioned assumptions and conditions imposed by such methodology, data on time series of firm-specific variables (spanning between April the 1st, 2015 and January the 1st, 2018) relative to *Capital Adequacy* and *Total Exposures* measures had been used: the former identifies the amount of *Regulatory Capital* needed, while the latter represents the *Exposure at Default* value. Furthermore, in *The Bank*'s specific instance, assumptions of the IRB methodology allow the use of constant values over time for the *Loss Given Default* and the *Maturity* measures, which would respectively be fixed at 0.45 and 2.50.

Then, by solving formula (1.8) for the unknown variable, baseline time series for *PD* had been obtained, from April the 1st, 2015 until January the 1st, 2018.

To summarize, the *PD* assessment analysis had been based on the use of 6 macroeconomic and 9 firm-specific time series, of which 2 where exploited for calculating the baseline values for *PD*, as previously-explained. Each time series has 34 monthly observations and predictions would be performed up until 12 months ahead. Summary statistics for the time series relating to *PD*, and the variables which are deemed to be causing it, are shown in Table 3. Notice in particular that the mean baseline value for *PD* is around 2.83% over the time period considered; the minimum value observed is 2.07% while the maximum value is 4.17%; moreover, the distribution for *PD* is slightly heavy-tailed and positively skewed. Lastly, Figure 4 provides a visual representation of the time series and density relating to the baseline *PD* computed via the *IRB Approach*.

3.2 Assessing Causal Relations

As a first step into exploring causal relations between variables, it may be possible to infer either the presence or the absence of such connections by applying the principles of *Granger Causality* (GC): this approach allows to assess causal relations between couples of variables by utilizing a specific type of test performed on a *Vector Auto-Regressive* (VAR) *process* (see *APPENDIX A* for a brief review on this topic).

On the basis of the assumptions for the application of GC principles, one would start by assessing the stationarity of the time series considered. Then, one may obtain it via differencing or adjusting for seasonality: transformations which are applied on the basis of additional graphical analyses and formal testing procedures, like the Augmented Dickey Fuller (ADF) test, aimed at either supporting or rejecting the application of these transformations. Once completed such preliminary phase, the order of integration associated to each series may be defined.

Notice that a GC test[12], performed on a specific VAR process composed by elements belonging to two distinct stationary time series, may be applied by exploiting previously-obtained transformed stationary time

[12] The *Granger Causality test* is essentially based on the use of an F-type test aimed at verifying the significance of the causal relation between the considered variables. Notice that the null hypothesis (H_0) would translate into the absence of GC, while the alternative hypothesis (H_1) would correspond to the inferred presence of GC between the variables included in the relative VAR model.

FIRM-SPECIFIC VARIABLES							
Variable	CET1	RORAC	NPL	CoR	NSFR	LR	IR
Order of Integration	$I(1)$	$I(0)$	$I(1)$	$I(2)$	$I(2)$	$I(1)$	$I(1)$
MACROECONOMIC VARIABLES							
Variable	UN	Infl	GDP	Mkt.Sh	Oil	Euribor	
Order of Integration	$I(0)$	$I(2)$	$I(1)$	$I(2)$	$I(1)$	$I(2)$	
RESPONSE VARIABLE							
Variable	PD						
Order of Integration	$I(1)$						

Table 4: Orders of integration relative to the time series of firm-specific and macroeconomic variables, as well as the one concerning the response variable.

series. However, the use of such modified series would imply the inevitable loss of their original features. As a consequence, test results may provide an unfeasible interpretation of the causal relations linking variables.

Therefore, a possible solution to the problem would be to utilize the original time series to build the relative VAR process, on condition that the cointegration property is verified (see *APPENDIX A*). This methodology has the advantage of preserving the information carried by each time series in its original form; hence, test results may potentially be more accurate on the assessment of causality between variables. However, such approach has a drawback: the cointegration principle is only applicable to those couple of series which possess the same order of integration; then, some interactions between variables will likely not be tested a priori. Nevertheless, at the expense of potentially missing out on part of the actual causal relations, it is probably still a better choice to perform the analysis by utilizing untransformed time series, to hopefully grasp those that are deemed to be the most relevant causal relations.

To clarify what said so far and contextualize it in relation to *The Bank*'s Stress Testing case, the first step would be to assess the order of integration associated to each time series for the variables considered, either firm-specific or macroeconomic. Graphical analyses and ADF tests had been utilized to assess the relative orders of integration; the key results of the analysis are shown in Table 4.

In accordance with the cointegration principle, it would make sense to set up VAR models each utilizing, on the explanatory side, one among the series possessing the same order of integration as PD (the response variable). Similarly, the same would apply for relations between macroeconomic and firm-specific variables (response variables for these instances). In the Stress Testing analysis, the selection process of the necessary lagged values to be included in such VAR models had been based on a function which automatically selects the lags to consider, on the basis of the relative BIC scores for the series' past values. Then, it had been possible to create VAR models to perform a GC test on each of them.

The core assumption utilized to define the high-level causal relations applied throughout the whole Stress Testing analysis is the

> ***Structural Assumption****: The PD variable may directly be affected by either firm-specific or macroeconomic variables. Firm-specific variables may be affected by macroeconomic variables.*

On the basis of such assertion, in consideration of the GC approach and provided that PD is $I(1)$, the following would hold:

- The firm-specific variables which may directly affect PD would be $CET1$, NPL, LR and IR;
- The macroeconomic variables which may directly affect PD would be GDP and Oil;
- The macroeconomic variables which may directly affect firm-specific variables (those which in turn may Granger cause PD) would be GDP and Oil.

Notice that, by adopting this approach, one would directly exclude to perform any GC test on $RORAC$, CoR, $NSFR$, UN, $Infl$, $Mkt.Sh$ and $Euribor$. Such classification would define the perimeter over which to perform GC tests. To this end, Table 5 clarifies the concept and displays the p-values

Variable	GC Test p-value	Presence of GC?
Firm-specific variables which may Granger cause PD		
CET1	0.0275	Yes
NPL	0.7201	No
LR	0.0479	Yes
IR	0.0937	No
Macroeconomic variables which may Granger cause PD		
GDP	0.0073	Yes
Oil	0.5659	No
Macroeconomic variables which may Granger cause $CET1$		
GDP	0.0011	Yes
Oil	0.0134	Yes
Macroeconomic variables which may Granger cause LR		
GDP	0.0058	Yes
Oil	0.0083	Yes

Table 5: Display of p-values for the executed GC Tests and related interpretation of Granger causality between variables.

together with their related practical interpretation, obtained from performing the GC tests on the variables included in the perimeter. As the analysis' output demonstrates, p-values below 0.05 are associated to the rejection of the null hypothesis H_0 at a 5% significance level. Then:

- $CET1$ and LR are the firm-specific variables deemed to be Granger causing PD;
- GDP is the macroeconomic variable deemed to be Granger causing PD;
- GDP and Oil are the macroeconomic variables deemed to be Granger causing both $CET1$ and LR.

The aforementioned conclusions are also reflected in a straightforward way in Figure 5, which graph puts into evidence those that are believed to be the relations between variables for the Stress Testing analysis, according to the GC methodology.

As the graphical representation in Figure 5 shows, *The Bank*'s probability of default is believed to be directly caused by the level of leverage relative to *The Bank*'s exposures, the amount of regulatory CET1 capital set aside and the average European GDP level. Indirect effects on the probability of default are deemed to be provided by both the *GDP* variable and the mean level for oil prices across European countries.

Nevertheless, as previously mentioned, the application of GC principles could exclude a priori variables which may instead turn out to be relevant for explaining causality on *PD*. As a consequence, chances are that GC methodology will not convey a thorough coverage for analyzing the variables at hand.

Therefore, it is advisable to perform further analyses with the use of other more adequate and flexible tools.

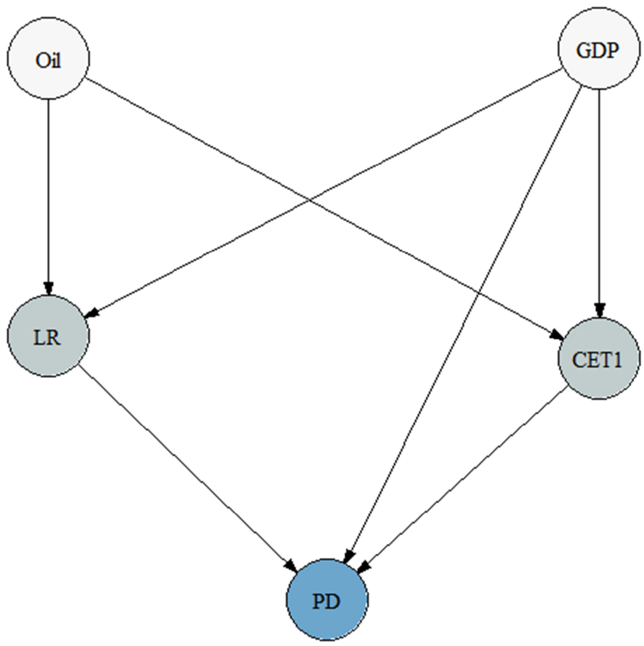


Figure 5: Graphical representation of Granger Causal relations linking variables and flowing towards *PD*.

3.3 Bayesian Networks Structures

An alternative approach to identify causal relations linking variables and explaining the effects on *PD*, is the estimation of BNs relative to the Stress Testing case. In particular, given data's characteristics, the structure and the parameters of a GBN model would be inferred. Moreover, to allow for results comparisons, an *Expert Network* will be developed.

3.3.1 Expert Network

A first way of defining the GBN structure may be to exploit expert knowledge in order to describe causal relations between and among variables, keeping the focus on assessing *PD*'s values. Figure 6 displays the expert network deemed to be best describing such causal links, as for the data and variables available in *The Bank*'s Stress Testing analysis.

Explanations are given next on the way the *Expert Network* had been constructed; furthermore, reasons for the inclusion of the variables deemed to be causing *PD* are also provided.

In the first place, concerning the layout of the expert network and the relative categories of variables (*Firm-specific*, *Macroeconomic* and *Response*) involved, it suffices to say that the guiding principle defining the overall structure of the system actually is the one previously exposed in the *Structural Assumption*.

On the macroeconomic side, in general, a number of phenomena at the European level are believed to potentially generate impacts for *The Bank*'s business activities. To this end, the following variables are deemed to indirectly cause *PD*:

- ***Oil***: in consideration of *The Bank*'s business-specific characteristics, oil price movements (which are often associated with high volatility) may determine significant causal effects on *NPL*, *CoR* and *CET*1;

- ***UN***: depending on its level, it may have important effects on both *NPL* and *CoR* for credit-related reasons. Relative to the former, an increase in unemployment would probably correspond to an increase in NPLs. As for the latter, an increase in *UN* may lead to the creation of riskier exposures for *The Bank*, hence raising the level of *CoR*;
- ***Mkt.Sh***: performances of *The Bank*'s stock on the market may reflect changes in funding and regulatory capital levels. In this sense, there may me an effect on the level of *CET*1;
- ***Infl***: its fluctuations may have weak effects on *NPL* and *CoR*, given that abnormal levels of inflation may indirectly cause increases in the number of Non-Performing Loans as well as more costs to firms for potentially increased risks.

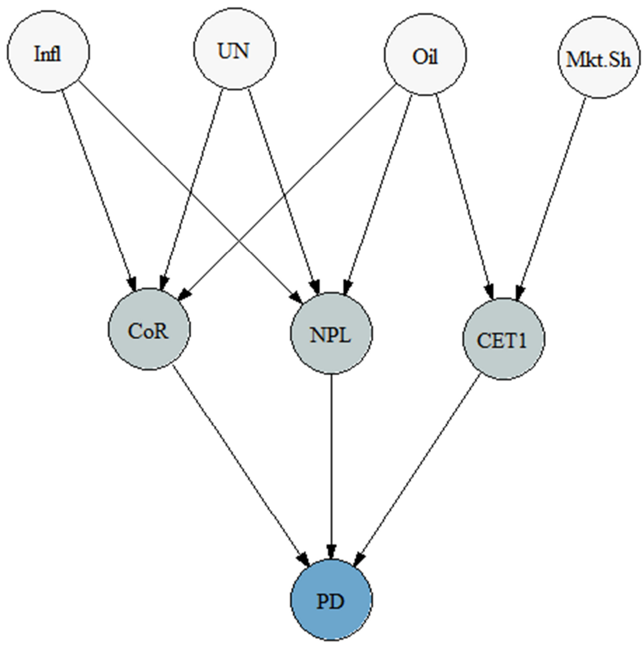


Figure 6: Structure defined by expert knowledge to ultimately identify the direct and indirect causal links influencing *PD*.

On the firm-specific side, instead, the following variables are deemed to directly affect the value of *PD*:

- ***NPL***: given the features of *The Bank*'s business, it probably is the variable which influences the most *PD*'s level. In other words, the Non-Performing Loans ratio is believed to be critical in terms of causal effects toward the target variable of the analysis. Indeed, a rise in NPLs would directly affect *The Bank*'s exposures, which really are a central factor for the business' performances;
- ***CoR***: it also possesses a strong causal link towards *PD*. Indeed, a raising Cost of Risk may be associated with downward effects on the business income, which in turn will likely have consequences on *The Bank*'s *PD*;
- ***CET1***: compared to the other firm-specific variables, it has a weaker causal effect on *PD*. Still, scarce capital buffers may correspond to the incapacity of the business to put aside capital amounts to counter future potential crises. This would consequently raise the probability that, in correspondence of a downturn, the credit institution may actually default.

For the purposes of the present Stress Testing analysis, the expert network would represent the reference structure for this specific instance.

Notice that, according to the results obtained in Table 5, the previously utilized GC approach does have shortcomings in comparison with the expert network: the only correspondence between the two methodologies is provided by the path $Oil \rightarrow CET1 \rightarrow PD$ and an important deficiency of the GC method implies that it fails to recognize *NPL* among the variables that directly affect *PD*.

The following section will illustrate a GBN methodology that may be adopted for those cases in which expert knowledge is not available.

3.3.2 GBN Structure Learning

Suppose expert knowledge on specific variables' features was absent and the only information on the network layout would be provided by the intuitive concept exposed in the *Structural Assumption*. Then, it would be necessary to define the structure of causal relations linking variables by elaborating on the raw data available. To this end, a GBN may be estimated.

The first step in the definition of a GBN would be to implement a procedure in order to learn its structure and causal relations from data. To accomplish that, it would be necessary to dispose of probability distributions for each variable in every time instant. In the Stress Testing case a Monte Carlo (MC) simulation had been executed and performed on all variables, from baseline *PD* to macroeconomic and firm-specific variables. More specifically, the procedure consisted in the following two simple phases:

1. Fitting ARIMA models on each time series[13];
2. Perform a simulation of 1000 steps[14] on each time series, exploiting the data available and the ARIMA models defined in the previous phase.

Hence, the use of MC procedure allowed to obtain enough simulated observations to define the needed probability distributions associated with each variable for all the 34 time instants.

---

[13] Each series had been estimated via the use of an automatic procedure which selected the best parameters according to certain criteria. For more details on the function applying such principles refer to the *R* documentation on *forecast* package, which is also mentioned in *APPENDIX B*.

[14] The number of steps chosen for the MC simulation was arbitrary. However, notice that such choice was dictated by previously performed trial simulations with different number of steps employed. What turned out from this preliminary analysis was that 1000 steps would probably be an adequate compromise, provided that it allows a balance in computational speed and accuracy of simulated resulting probability distributions.

To infer the structure of the GBN the HC algorithm is exploited; its core features are described in section 2.5.1. It is important to mention some aspects on the way such instrument was utilized. In particular, for each moment in time t_k, with $k = 1, 2, \dots, 34$, the following holds:

- The structure optimization procedure is based on *boostrap resampling*, which the HC algorithm applies in this case;
- Input data for the HC algorithm correspond to the previously-simulated variables' probability distributions;
- The HC algorithm utilizes the BGe heuristic as the scoring criteria for estimating the network's structure.

Hence, by applying this procedure, the output would correspond to 34 estimated GBNs (graphs are not shown in this paper), one for each t_k.

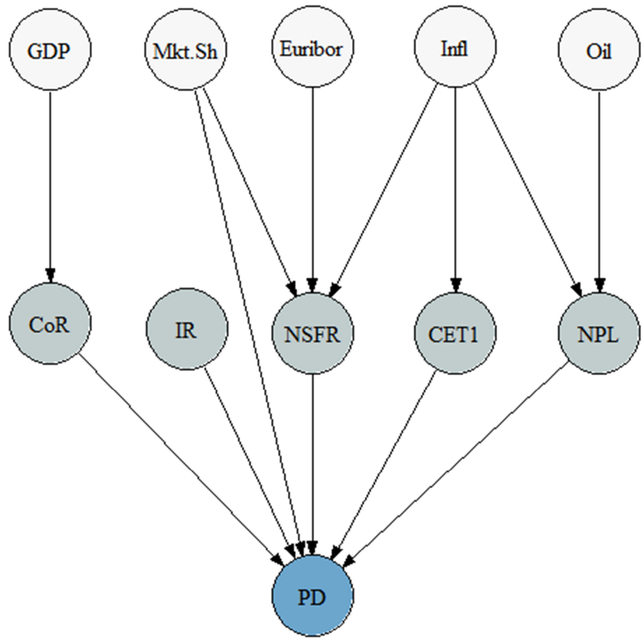
The following step would be to keep trace of the arcs set by the HC algorithm across the 34 estimated GBNs. In this way, it would be possible to evaluate the strength of the causal relations between and among variables, by observing the most recurring causal links in correspondence of different time instants. Table 6 displays the estimated links in all 34 GBNs and provides an indication of which relations may be the strongest.

By assessing the frequency with which arcs are estimated via HC algorithms, for each time step, it would be possible to define a guiding structure for the learned GBN. Such approach may be interpreted as a type of *model averaging* procedure: the aim would be to obtain a network valid throughout all time steps, reflecting those causal relations that are expected to be persisting over time. In particular, this may be accomplished by setting thresholds on the arcs' percentage of appearance over all t_ks. Notice that choosing high threshold levels would results into the definition of more robust GBNs, at the expense of losing trace of some weaker relations that may be present. In order to represent the GBN layout for this specific instance, 50% and 85% threshold values on the arcs' percentage of appearance over all t_k had been chosen.

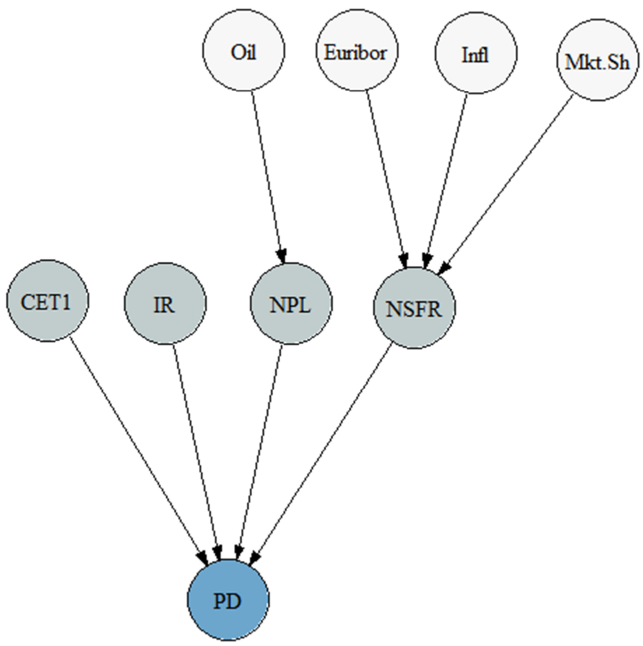
Arcs linking Variables			Arcs' Frequency of Appearance over all t_k	Arcs' Percentage of Appearance over all t_k
From		**To**		
CET1	→	PD	34	100.00%
CoR	→	PD	23	67.65%
Euribor	→	CET1	14	41.18%
Euribor	→	CoR	4	11.76%
Euribor	→	IR	7	20.59%
Euribor	→	LR	33	97.06%
Euribor	→	NSFR	31	91.18%
Euribor	→	RORAC	6	17.65%
GDP	→	CET1	12	35.29%
GDP	→	CoR	33	97.06%
GDP	→	IR	10	29.41%
GDP	→	NSFR	5	14.71%
GDP	→	PD	9	26.47%
GDP	→	RORAC	34	100.00%
Infl	→	CET1	25	73.53%
Infl	→	CoR	2	5.88%
Infl	→	IR	2	5.88%
Infl	→	LR	34	100.00%
Infl	→	NPL	24	70.59%
Infl	→	NSFR	34	100.00%
Infl	→	RORAC	1	2.94%
IR	→	PD	34	100.00%
Mkt.Sh	→	CoR	1	2.94%
Mkt.Sh	→	IR	2	5.88%
Mkt.Sh	→	NPL	2	5.88%
Mkt.Sh	→	NSFR	34	100.00%
Mkt.Sh	→	PD	26	76.47%
Mkt.Sh	→	RORAC	10	29.41%
NPL	→	PD	34	100.00%
NSFR	→	PD	34	100.00%
Oil	→	IR	1	2.94%
Oil	→	NPL	33	97.06%
Oil	→	PD	1	2.94%
Oil	→	RORAC	2	5.88%
RORAC	→	PD	4	11.76%
UN	→	CET1	3	8.82%
UN	→	IR	3	8.82%
UN	→	LR	22	64.71%
UN	→	NSFR	4	11.76%
UN	→	PD	14	41.18%
UN	→	RORAC	6	17.65%

Table 6: Arcs estimated via HC algorithms' *bootstrap resampling* procedure and relative Frequency and Percentage of appearance for such arcs over all t_ks.

Then, on the basis of results shown in Table 6, the GBN defined via the 50% threshold would include all those arcs that are present more than half the times over all t_ks, in accordance with the output of the HC



(a)



(b)

Figure 7: Networks obtained from *model averaging* procedures based on (a) the 50% threshold principle and (b) the 85% threshold one.

algorithm estimation procedure. A similar principle would hold when applying the 85% threshold, keeping into consideration all those arcs that are estimated with a percentage frequency above such level. The output from the application of the two thresholds is reported in the GBNs depicted in Figure 7.

It is useful to compare the estimated networks with the one previously-defined via expert knowledge.

Firstly, considering Figure 7(a), the arcs and paths present in both the 50% threshold network and the expert network are the following:

$Infl \rightarrow NPL \rightarrow PD$;

$Oil \rightarrow NPL \rightarrow PD$;

$CoR \rightarrow PD$;

$CET1 \rightarrow PD$.

Notice that the variables deemed to be not relevant, according to expert knowledge, but still estimated in the 50% threshold network are GDP, $Euribor$, IR and $NSFR$.

Secondly, with reference to Figure 7(b), the arcs and paths present in both the 85% threshold network and the expert network are:

- $Oil \rightarrow NPL \rightarrow PD$;
- $CET1 \rightarrow PD$.

The estimated variables irrelevant if compared with the expert network are $Euribor$, IR and $NSFR$.

Both the estimated GBNs include most of the variables defined through the expert network. Comparatively, the 50% threshold network grasps all the relations between firm-specific variables and PD, while the 85% threshold network fails to get the $CoR \rightarrow PD$ arc. Notice that, concerning the macroeconomic variables, UN is never present in the estimated networks. Moreover, paths and links between variables often do not correspond to the ones identified via expert knowledge. Nevertheless, it is important to notice that both the 50% and 85% threshold networks are able to grasp the most relevant causal path linking macroeconomic, firm-specific and response variable elements, that is $Oil \rightarrow NPL \rightarrow PD$. Indeed, the causal effects generated by *Oil* and *NPL* on *PD* are deemed to be the most crucial ones.

Therefore, one may assert that the implementation of HC algorithms and model averaging principles for estimating GBNs provides rather decent approximations of those causal links which are expected to persist on *PD*.

3.4 Inference on Bayesian Networks' Parameters

On the basis of the previously-defined causal structures (estimated and expert networks) the following phase would imply to conduct inference on the relative GBN parameters. This would translate into assessing the value of explanatory variables' coefficients associated with

Response Variable	**Explanatory Side**	**Coefficients' Estimates**	**Confidence Interval**	
			Lower Bound (2.5%)	**Upper Bound** (97.5%)
NPL	Intercept	-0.0038	-0.0095	0.0019
	Infl	-0.0771	-0.1475	-0.0067
	Oil	-0.0136	-0.0260	-0.0012
	UN	0.2383	0.1757	0.3010
CoR	Intercept	0.0118	0.0060	0.0175
	Infl	-0.1354	-0.2059	-0.0649
	Oil	0.0010	-0.0114	0.0134
	UN	-0.1085	-0.1712	-0.0457
CET1	Intercept	0.1265	0.1111	0.1419
	Mkt.Sh	-0.1462	-0.3515	0.0591
	Oil	0.0161	-0.0403	0.0726
Mkt.Sh	Intercept	0.0747	0.0722	0.0772
Infl	Intercept	0.0065	0.0040	0.0090
Oil	Intercept	0.0003	-0.0089	0.0095
UN	Intercept	0.0853	0.0825	0.0882
PD	Intercept	-0.0836	-0.1056	-0.0615
	NPL	-0.0677	-0.3678	0.2324
	CoR	0.0590	-0.6680	0.7859
	CET1	0.9763	0.7849	1.1678

Table 7: Maximum Likelihood coefficient estimates and relative confidence intervals concerning the expert network.

each node of the network. For these specific instances, MLE (in accordance with the assertions provided in section 2.5.2) had been employed for inferring key networks' parameters, based on data available for all the time instants. Table 7, Table 8 and Table 9 report the output on significant coefficients' estimates and confidence intervals relative to the expert network and the estimated networks.

With reference to results shown in Table 7, in relation to the explanatory variables and the relative confidence intervals, notice that the most precise estimates are provided by *Oil* coefficients, relative either to $CET1$,

Response Variable	**Explanatory Side**	**Coefficients' Estimates**	**Confidence Interval**	
			Lower Bound (2.5%)	**Upper Bound** (97.5%)
CET1	Intercept	0.1151	0.1131	0.1170
	Infl	0.0788	-0.1256	0.2833
CoR	Intercept	0.0015	0.0011	0.0018
	GDP	0.0195	0.0060	0.0331
Euribor	Intercept	-0.0022	-0.0027	-0.0017
GDP	Intercept	0.0080	-0.0007	0.0168
Infl	Intercept	0.0065	0.0040	0.0090
IR	Intercept	0.0008	0.0006	0.0010
Mkt.Sh	Intercept	0.0747	0.0722	0.0772
NPL	Intercept	0.0179	0.0172	0.0186
	Infl	-0.2847	-0.3606	-0.2089
	Oil	-0.0051	-0.0259	0.0156
NSFR	Intercept	1.1856	1.1240	1.2473
	Euribor	-7.6628	-14.1443	-1.1814
	Infl	-2.8031	-3.9416	-1.6645
	Mkt.Sh	-0.9605	-1.8204	-0.1006
Oil	Intercept	0.0003	-0.0089	0.0095
PD	Intercept	-0.0221	-0.0976	0.0534
	CET1	0.7995	0.5602	1.0389
	CoR	-0.2461	-1.0246	0.5324
	IR	-2.6676	-4.6998	-0.6354
	Mkt.Sh	-0.0664	-0.1980	0.0652
	NPL	-0.3117	-0.6647	0.0412
	NSFR	-0.0266	-0.0728	0.0196

Table 8: Maximum Likelihood coefficient estimates and relative confidence intervals concerning the learnt 50% threshold network.

CoR or *NPL* as response variables; while the least accurate ones mostly concern *NPL*, *CoR* and *CET*1 relative to *PD* as response variable.
In a similar way, the output displayed in Table 8 and Table 9, shows that, in general, the least accuracy in estimates is associated with coefficients regarding both *PD* and *NSFR* as response variables. In particular, for instance, one may interpret with caution *Euribor*'s coefficient (relative to *NSFR*) provided that its confidence interval suggests there may be imprecision in the estimate.

Then, one may assert that coefficient estimates associated with the relative intermediate and leaf nodes (each possessing layers of parent nodes) of the expert or learned networks are increasingly less accurate. This observation may be supported by the values associated to the respective coefficients' confidence intervals; indeed, this may be plausible given the following: the deeper a node is located in the network, the more

Response Variable	**Explanatory Side**	**Coefficients' Estimates**	**Confidence Interval**	
			Lower Bound (2.5%)	**Upper Bound** (97.5%)
CET1	Intercept	0.1156	-0.0089	0.0095
Euribor	Intercept	-0.0022	0.1141	0.1170
Infl	Intercept	0.0065	-0.0027	-0.0017
IR	Intercept	0.0008	0.0040	0.0090
Mkt.Sh	Intercept	0.0747	0.0006	0.0010
NPL	Intercept	0.0160	0.0151	0.0169
	Oil	-0.0113	-0.0459	0.0233
NSFR	Intercept	1.1856	1.1240	1.2473
	Euribor	-7.6628	-14.1443	-1.1814
	Infl	-2.8031	-3.9416	-1.6645
	Mkt.Sh	-0.9605	-1.8204	-0.1006
Oil	Intercept	0.0003	0.0722	0.0772
PD	Intercept	-0.0380	-0.1031	0.0272
	CET1	0.8472	0.6307	1.0638
	IR	-2.9457	-4.8552	-1.0363
	NPL	-0.3272	-0.6618	0.0074
	NSFR	-0.0217	-0.0631	0.0197

Table 9: Maximum Likelihood coefficient estimates and relative confidence intervals concerning the learnt 85% threshold network.

it could be influenced by other nodes' interactions, hence generating increased uncertainty relative to its coefficients' estimates. Such phenomenon may also be explained by the fact that, starting from root nodes, the causal flows propagate down the network and, consequently, may generate slight gradual increases in noise associated to the coefficients' estimates.

Anyway, it is worth stressing that, for these specific instances, estimated errors associated to networks' nodes are often small in magnitude.

As the next section will explain, the obtained coefficients' estimates actually become useful when applied to forecasting future variables' values relative to the considered GBNs.

3.5 Variables Prediction and Analysis Output

The previously-performed procedures aimed at the definition of structures and parameters of the expert and learnt GBNs, allow to dispose of instruments which may be utilized for predictive purposes in accordance with *The Bank*'s specific objectives. More specifically, this section deals with the use of the expert network and the 85% threshold GBN as tools to assess the level that *PD* and the relevant firm-specific variables may take relative to the period ranging from February 1st, 2018 until January 1st, 2019, namely the prediction period (a 12 months period ahead with respect to the last observation available in the baseline dataset, as previously defined).

Notice that given the 85% threshold GBN is a more conservative structure compared to the 50% threshold one, it is possible to assert that their predictive performances would be rather similar, moreover, the former is composed only by the more robust variables and stronger arcs learnt. For these reasons, the predictive analysis will be carried out comparing the results obtained via the expert network and the 85% threshold GBN only. The 50% threshold system will not be included since it would not add significant predictive information other than the ones already provided by the other estimated network.

The advantage of utilizing systems of variables for predictive purposes is that by providing forecasted measures on the root nodes of the GBNs one may observe the effects of such input data on all other descendants. Indeed, thanks to causal edges linking variables, it is possible to *propagate* the effects of input data at the root-nodes level through the firm-specific variables and, in turn, down towards the end node PD. To this end, the first step into forecasting had been to define predictive estimates based on the relative historical values of each time series, for all the variables involved in the analysis. As already mentioned, such preliminary step is necessary to define networks' input data, which actually find correspondence with the point predictions relative to each macroeconomic variable. In particular, forecasted values for $Infl$, UN, Oil, $Mkt.Sh$ and $Euribor$ would be utilized. Remember that each descendant node is associated with a specific Gaussian linear regression model, hence, forecasts on PD and firm-specific variables (NPL, CoR, $CET1$, $NSFR$, IR) had been used for the definition of the relative random shocks

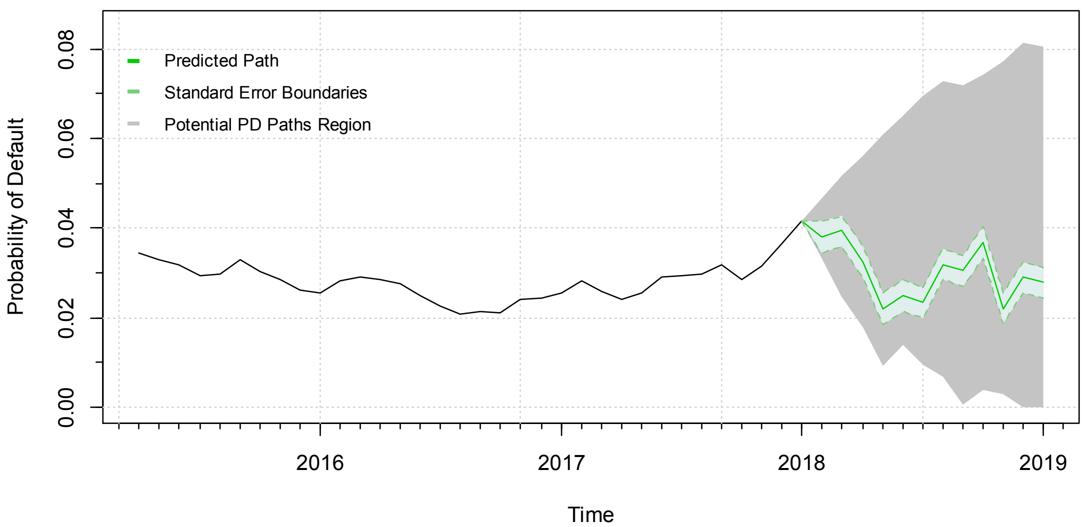


Figure 8: Predicted values for PD, standard error boundaries and prediction interval for the possible range of paths/values PD may take based on the 85% threshold GBN.

associated to each regression's error term. Then, it should be clear that by inputting macroeconomic forecasted values into children nodes, one would obtain predicted measures for each relative linear model of the firm-specific variables. In turn, these latter forecasted estimates would become input to the linear model associated with the end node *PD*.

By following the above-mentioned forecasting procedure it had been possible to infer which values *The Bank*'s *PD* is likely to take in the prediction period. Figure 8 and Figure 9, respectively referring to the estimated and the expert networks, show the predicted path, the relative standard errors and the prediction intervals concerning *PD*.

As one may notice, there is a marked similarity between the two prediction outputs: both forecasted paths suggest that *The Bank*'s *PD* is likely to oscillate around 2% and 4% relative to the prediction period; the resemblance is also supported by the relative standard error and prediction intervals ranges of values. Moreover, according to data available, notice that generated forecasts do not suggest the presence of any marked future trend; overall, *PD* level is expected to slightly decrease over the prediction period.

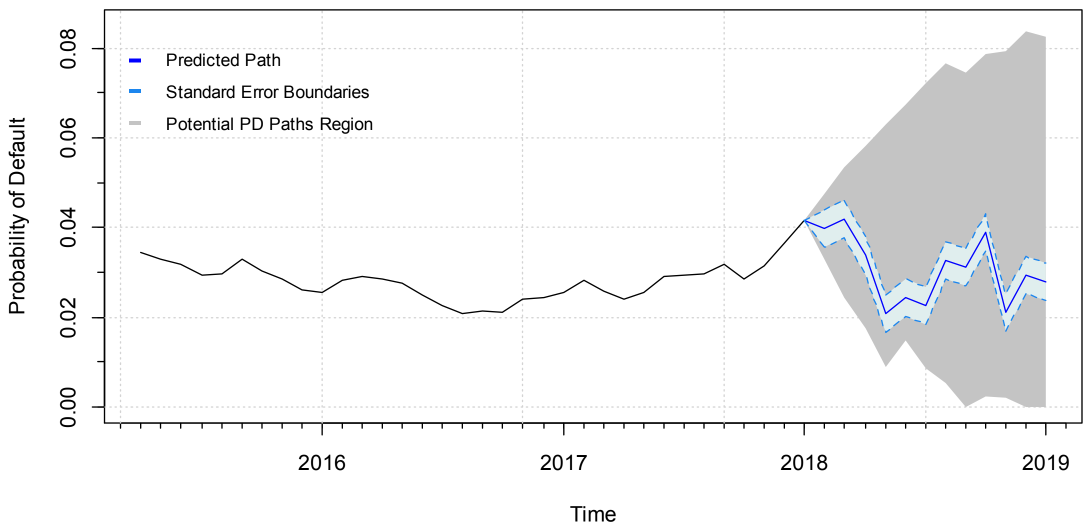


Figure 9: Predicted values for *PD*, standard error boundaries and prediction interval for the possible range of paths/values *PD* may take based on the expert GBN.

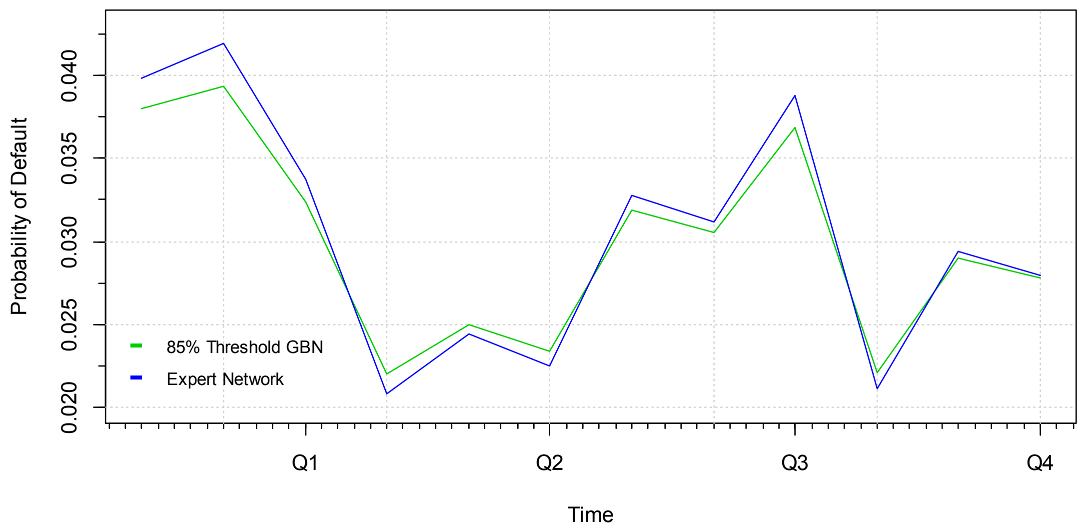


Figure 10: Predicted values for *PD* obtained via the 85% threshold GBN and the expert network relative to the prediction period.

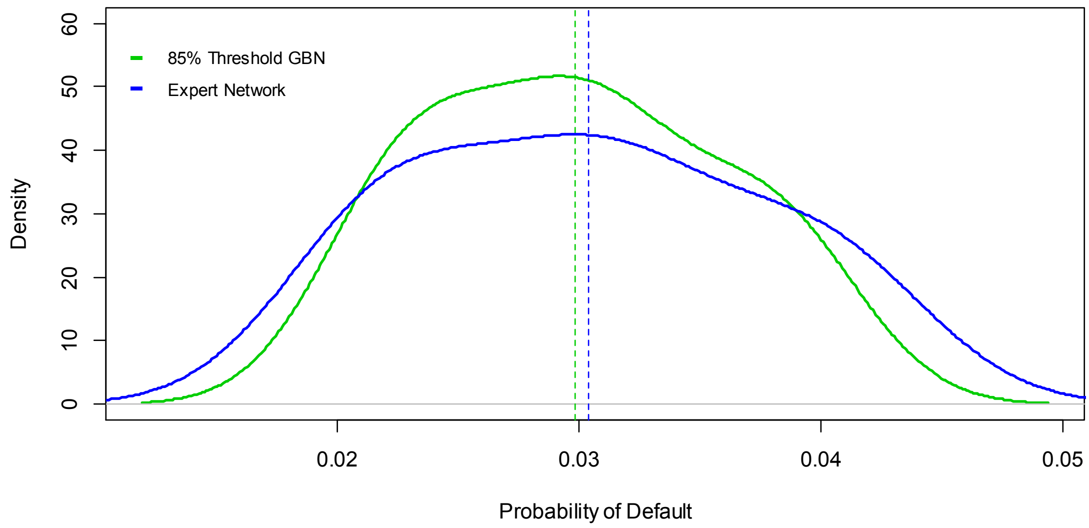


Figure 11: Comparison of the distributions relative to *PD* forecasts obtained via the 85% threshold GBN and the expert network. Dashed lines are depicted in correspondence of the relative mean values.

To better evaluate whether any significant discrepancies between such forecasted *PD* values exist, Figure 10 displays the two predicted paths in the same graph, while Figure 11 provides a view on the shape of the two relative distributions in a way that facilitates comparisons.

Following, some statistics relative to the two predicted distributions:

85% threshold GBN:
- Mean: 0.0296
- Standard Deviation: 0.0065
- Skewness: 0.1433
- Kurtosis: 1.7639

Expert Network:
- Mean: 0.0304
- Standard Deviation: 0.0073
- Skewness: 0.1628
- Kurtosis: 1.7733

Then, from the analysis' output it appears evident that the expert and learnt GBNs provide similar estimates of *The Bank*'s *PD*, despite presenting some discrepancies with regard to the variables and arcs composing such networks.

For the purposes of the Stress Testing analysis, other than assessing *PD*'s future path, it is of interest to provide an indication of potential trends concerning some relevant firm-specific variables. In particular, the focus will be on those that are believed to directly affect *PD*, as defined by the previously estimated 85% threshold GBN and the expert network. In the first place, prediction outputs relative to firm-specific variables in common between the two networks (*NPL* and *CET*1) will be exposed and compared. Following that, the remaining firm-specific variables' outputs of the GBNs would be presented.

By looking at Figure 12 one may notice a difference in the forecasted output obtained through the learnt network and the expert one. More specifically, in relation to the 85% threshold GBN, the predicted output and relative prediction intervals are somewhat close to the mean value, as shown in Figure 12(c), and do not suggest the presence of any upward or downward trends in *NPL* values, as depicted in Figure 12(a). Concerning the expert network prediction output, *NPL* values appear comparatively more spread out; this may be due to the fact that the expert GBN estimated the presence of a likely future downward trend, as may be deduced from Figure 12(b). In turn, one may assert the presence of more variability in *NPL* estimated values.

Figure 13, instead, shows the results obtained relative to *CET*1 variable. By simply looking at the graphs one may conclude that there is no significant difference in the predicted values for *CET*1, either they be obtained through the 85% threshold network or via the expert network. Notice that, in both cases, forecasted values appear to be concentrated around the mean of the distribution and variability in values seems to be rather limited.

Notice that in Figure 12(c) and (d), as well as in Figure 13(c) and (d), some of the dash-dotted vertical lines stand to indicate particular

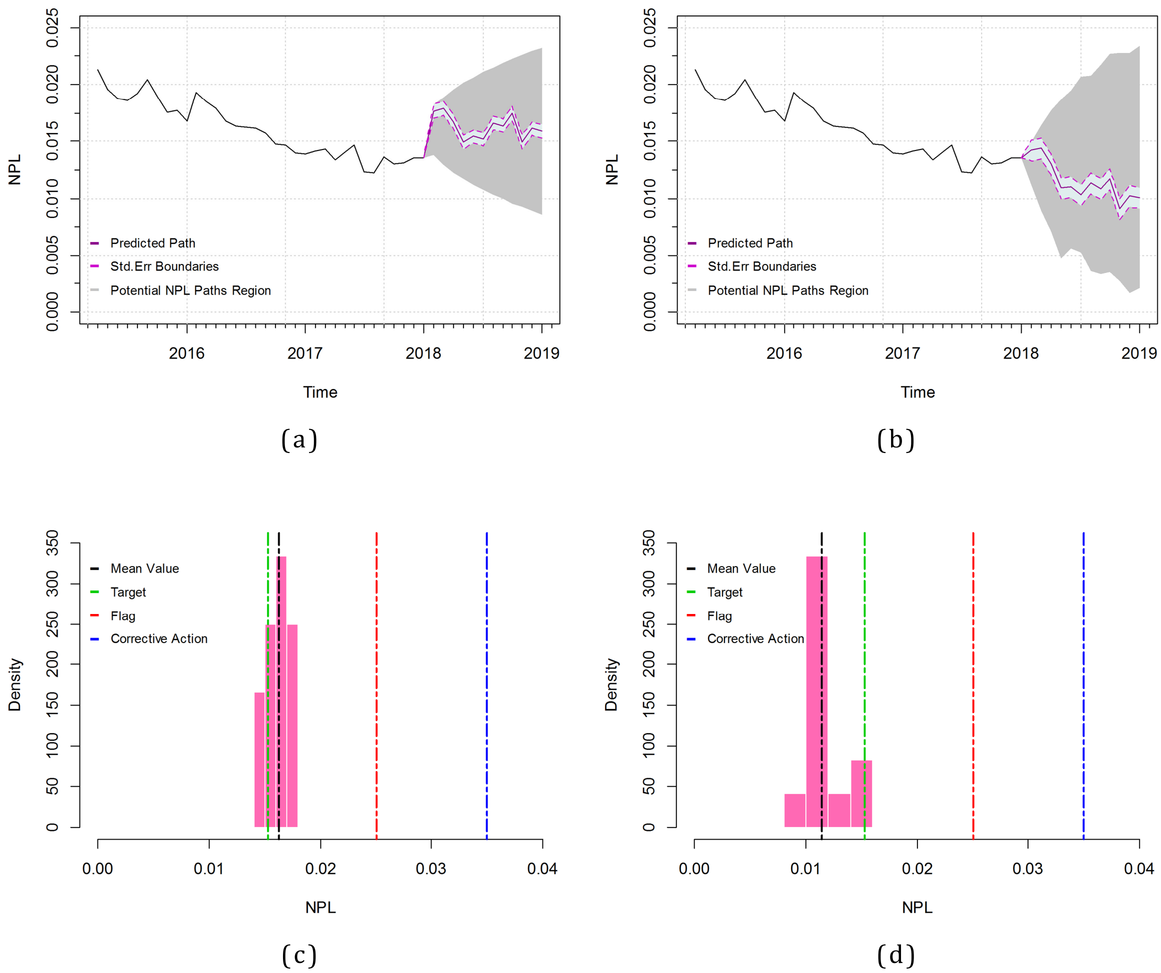


Figure 12: Predicted values for *NPL*, relative standard error boundaries and prediction intervals based on the use of (a) the 85% threshold GBN and (b) the expert network. Simulated forecasted distributions for *NPL* generated by utilizing (c) the 85% threshold GBN and (d) the expert network; dash-dotted vertical lines mark critical thresholds and values relative to *The Bank*.

thresholds set by *The Bank*. More specifically such limits may stand for:

- *Target*: level which *The Bank* aims to either reach or maintain;
- *Flag*: level which alerts *The Bank* and advised for a close monitoring of the specific measure;
- *Corrective Action*: level which requires *The Bank* to implement actions in order to adjust the specific measure;
- *Regulatory Requirement*: level imposed by regulatory authorities.

Both the learnt GBN and the expert network outputs, relative to *NPL* and *CET*1, suggest these measures are expected to remain close to *The Bank*'s

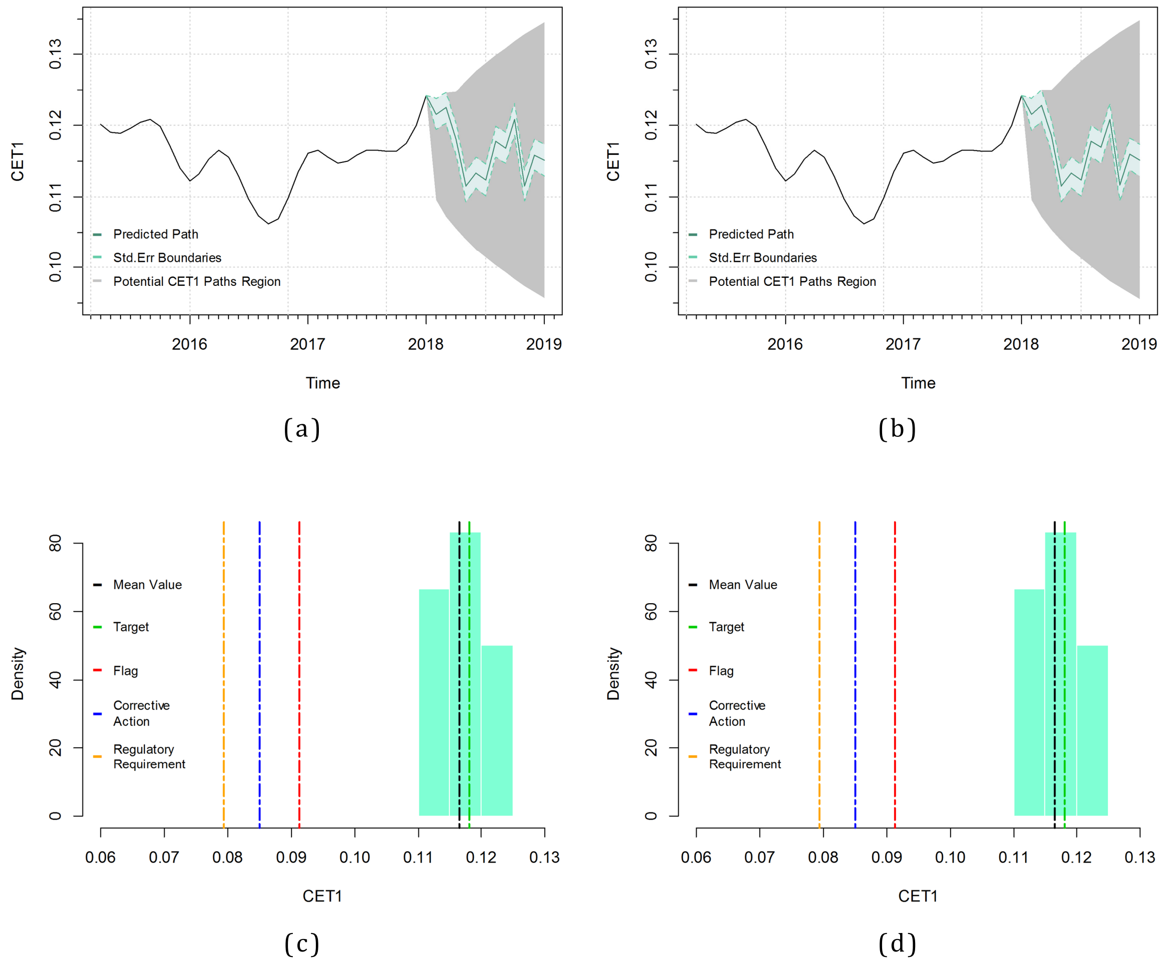


Figure 13: Predicted values for *CET1*, relative standard error boundaries and prediction intervals based on the use of (a) the 85% threshold GBN and (b) the expert network. Simulated forecasted distributions for *CET1* generated by utilizing (c) the 85% threshold GBN and (d) the expert network; dash-dotted vertical lines mark critical thresholds and values relative to *The Bank*.

target level. Therefore, on its part, *The Bank* may only continue monitoring the level of these two measures; no corrective actions are deemed to be required concerning the forthcoming period.

GBN	Variables	Mean	Standard Deviation	Skewness	Kurtosis
Expert	NPL	0.0114	0.0017	0.6994	2.4218
	CoR	0.0026	0.0004	-0.1201	1.2824
	CET1	0.1165	0.0039	0.1783	1.7818
85% threshold	NPL	0.0163	0.0010	0.1803	1.7804
	CET1	0.1164	0.0039	0.1785	1.7796
	NSFR	1.1030	0.0048	0.2651	1.8732
	IR	0.0009	0.0001	0.5864	1.7640

Table 10: Summary statistics for firm-specific variables' forecasted values relative to the 85% threshold GBN and the expert network.

To better appreciate what said so far concerning *NPL* and *CET*1 predicted values, in a more formal way, Table 10 displays summary statistics for all firm-specific variables included in the learnt and expert GBNs.

The previously performed learning procedure generated the 85% threshold GBN which, as already stressed, differs from the expert one in some aspects. In any case, it is still deemed valuable for *The Bank* to account for the most relevant results relative to those firm-specific variables which only find correspondence in one of the two networks[15]. Indeed, such type of discrepancies probably are of secondary importance, still, they should not be seen as totally wrongful estimates/forecasts because they may actually be revealing of non-obvious connections between variables.

[15] Notice that *IR* is deemed not to be a really significant variable for the purposes of the Stress Testing analysis. Indeed, no specific thresholds or targets had been set by *The Bank* in relation to such measure. Hence, the graphical display of its forecasted trend and distributional features are not shown in the present thesis since they would not add any significant insight to the analysis, for this specific instance. The only key information on *IR*'s predicted distribution are reported in Table 10. Notice, moreover, that the magnitude of *IR*'s mean and standard deviation is quite modest.

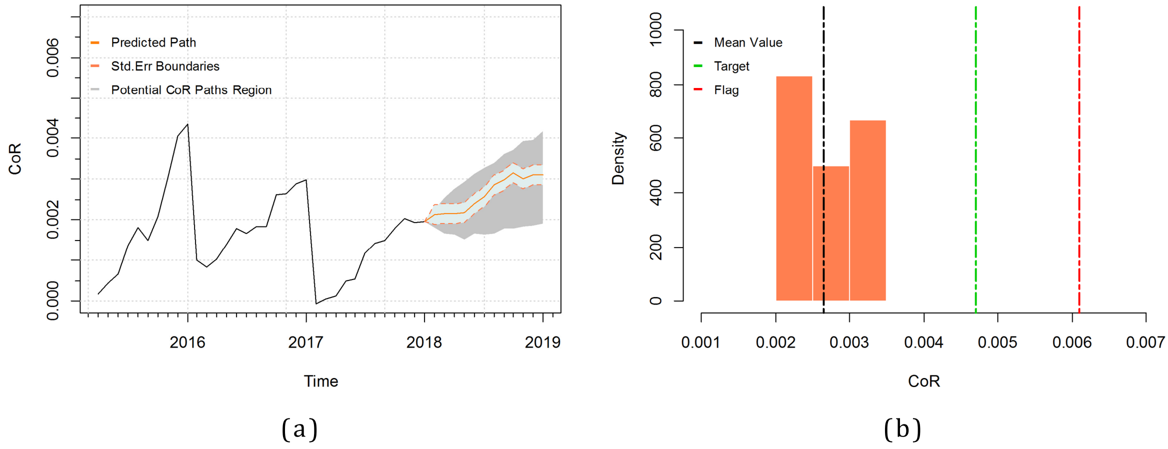
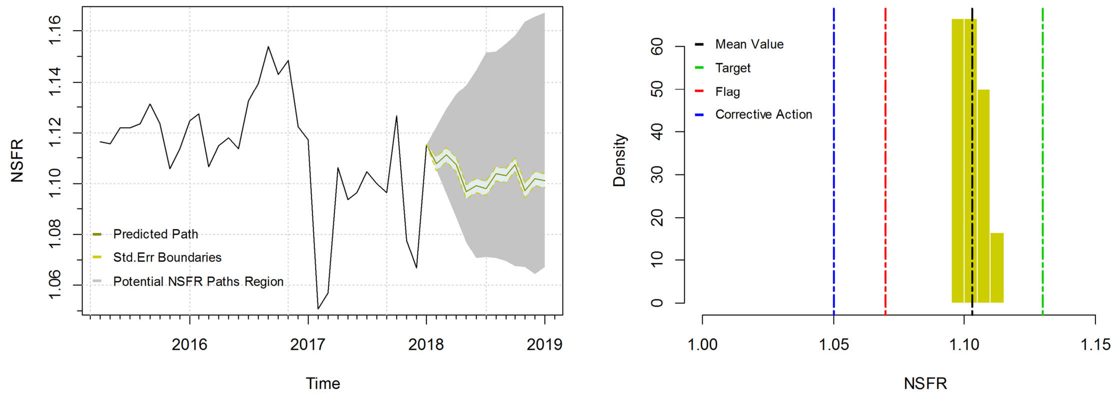


Figure 14: (a) Predicted values for *CoR,* relative standard error boundaries and prediction intervals based on the use of the expert network. (b) Simulated forecasted distributions for *CoR* generated by utilizing the expert network; dash-dotted vertical lines mark critical thresholds and values relative to *The Bank*.

In relation to the variable *CoR* of the expert network, it is necessary to stress a couple of aspects: first, by analyzing Figure 14(a), one may assert that the predictive model misses out on grasping the periodicity of the measure. Indeed, at the beginning of the prediction period one would probably expect a drop in *CoR* instead of an increase in its value[16]. Nevertheless, despite predicted values are likely to suffer from a slight upward bias, the forecasts seem to be capturing *CoR*'s trend correctly. Furthermore, in relation to Figure 14(b), *CoR*'s predicted values are, in any case, likely to be off-target with regard to *The Bank*'s objectives; still, this would probably be not a matter of particular concern provided that the *flag* threshold is even farther than *The Bank*'s target for such measure. Lastly, notice that deviations in values for *CoR* appear to be, in general, of small magnitude. In turn, one may infer that *CoR*'s causal effects on *PD* are likely to be small.

[16] This phenomenon may be explained by the fact that utilized input series' values remained untouched before being subject to data analysis procedures; this is in line with the monitoring process implemented by *The Bank* so far. Therefore, original values have been preserved and directly fed into learning algorithms/procedures to derive GBNs structures and parameters for the present Stress Testing analysis. Of course, with this in mind, *CoR* measure may be interpreted carefully.



(a) (b)

Figure 15: (a) Predicted values for *NSFR*, relative standard error boundaries and prediction intervals based on the use of the 85% threshold GBN. (b) Simulated forecasted distributions for *NSFR* generated by utilizing the 85% threshold GBN; dash-dotted vertical lines mark critical thresholds and values relative to *The Bank*.

Then, on the one hand, expert knowledge is suggesting *CoR* could be one of the relevant variables in the determination of *PD*'s value. On the other hand, it is interesting to notice that, in the 85% threshold version, the learnt GBN excludes *CoR* from being among the most important firm-specific variables causing *PD*. Therefore, one may conclude that *CoR* probably is a measure to be considered when carrying out the Stress Testing analysis; however, chances are that it would not be particularly impacting on *PD*'s value, given the identified magnitude in its values and deviations.

Concerning the *NSFR* variable included in the 85% threshold GBN, Figure 15(a) shows that predicted values are likely to be characterized by a slight downward trend for the prediction period. Moreover, Figure 15 evidences that variability in *NSFR* may be somewhat significant. Nevertheless, the major concentration of values for the distribution of forecasts appears to be closer to *The Banks* target level than to the *flag* one; therefore, chances are that corrective actions may not be immediately necessary for the foreseeable future.

Overall, structural discrepancies existing between the 85% threshold GBN and the expert network do not seem to significantly impact the forecasted value of *The Bank*'s *PD*. Firm-specific variables' should be analyzed carefully, given they are believed to be causing *PD*. Moreover, differences between the learnt and the expert GBN could result into opportunities to gain insight and get a more thorough picture on variables' interactions and importance.

Notice that, the performed predictions were obtained assuming normal market conditions, as reflected by historical data available. The following subsection, instead, will show an application of the GBNs to a different economic scenario that could take place in the future.

3.5.1 Scenario Analysis

The present subsection focuses on performing a *Scenario Analysis* which, in particular, is aimed at feeding into the estimated and expert GBNs crisis-like data assumed for the prediction period. Such simulation study is useful to analyze the outputs provided by such networks and to assess *The Bank*'s *PD* value, together with the firm-specific variables, in a possible future downturn scenario. Moreover, this type of analysis would also provide insight concerning *The Bank*'s own sensitivity and solidity in face of significant downward market pressures, hence effectively stress testing the credit institution.

A few considerations are necessary on the way data utilized for the scenario analysis had been created:

- In order to simulate a crisis-like scenario for the prediction period, changes in monthly data relative to the *crisis period*[17], ranging from April 2008 until March 2009, had been utilized as reference values. These were useful to generate simulated observations for a 12 month period starting from January 2018. Where real data were

[17] Period in which the 2008 Crisis' consequences and turmoil were probably felt the most by the economic environment surrounding *The Bank*.

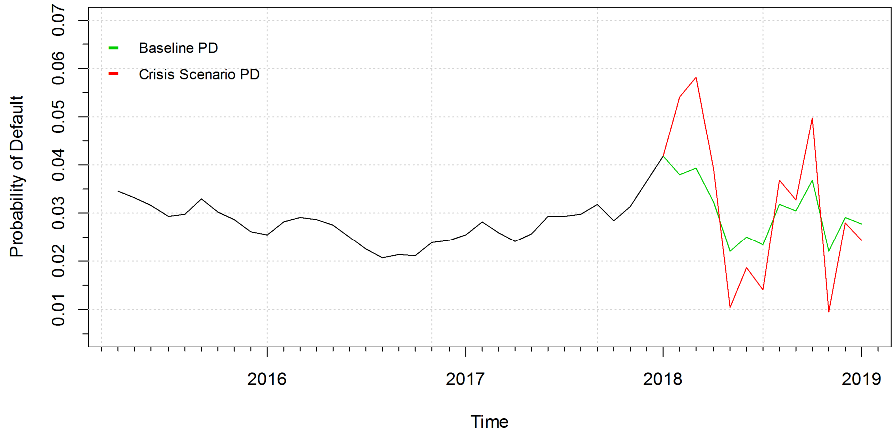
absent, as to simulate adverse conditions for *The Bank*, arbitrarily set changes to data had been employed;

- All simulated macroeconomic variables' data are based on actual changes registered during the crisis period, made exception for *Mkt.Sh*'s which had been generated assuming a decreasing trend over the prediction period;
- All simulated firm-specific data, relevant to the 85% threshold and the expert GBNs, had been created arbitrarily in a way that reflects worsening economic conditions for *The Bank*. The only variable which data are available for the crisis period is *NPL*; hence, its relative simulated values are based on actual changes registered during past times.

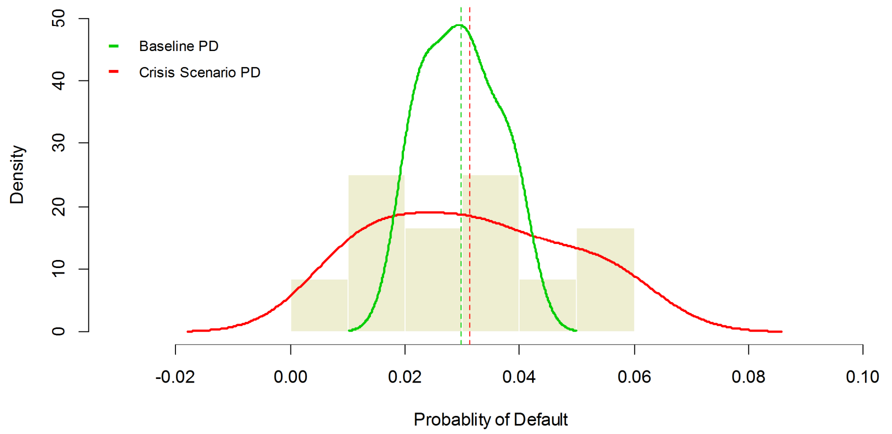
Notice that synthetic input data, relating to firm-specific variables, would be utilized to simulate the error terms of the linear models underlying firm-specific nodes. Macroeconomic simulated input data, instead, would be critical in defining descendants' values, thanks to the propagation of causal effects throughout the considered networks.

It is of interest to observe the output of the crisis simulation analysis relative to *The Bank*'s *PD* and subsequently backtrack to the firm-specific elements which influenced its relative value. As before, both the 85% threshold network and the expert network had been exploited to obtain the relative predictions.

Concerning the 85% threshold GBN, Figure 16 displays a comparison of *PD*'s prediction obtained in the previous section with the one generated via the application of the simulated crisis scenario; Figure 17 shows similar graphs relative to the use of the expert network. In general, for either case, one may conclude that crisis-like macroeconomic data caused *PD* predictions to take more extreme values and, in turn, to possess an increased variability in forecasts. This is actually in line with expectations, provided that market turmoil often generates uncertainty in the macroeconomic environment and such modified external conditions may also be reflected onto the *The Bank* itself.



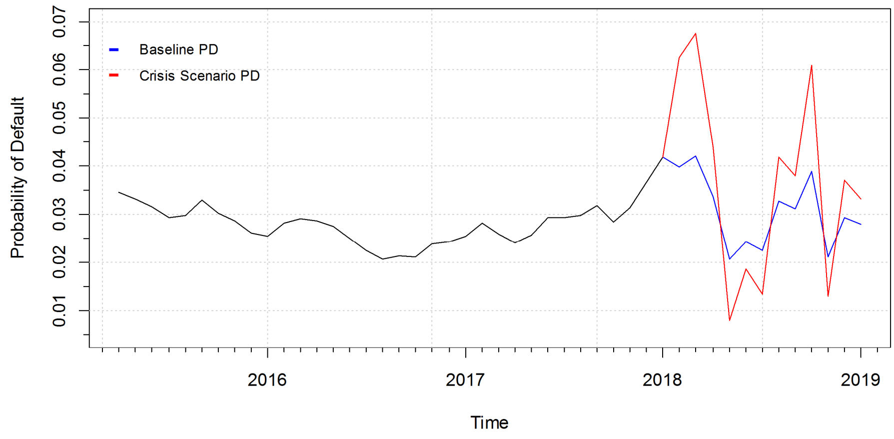
(a)



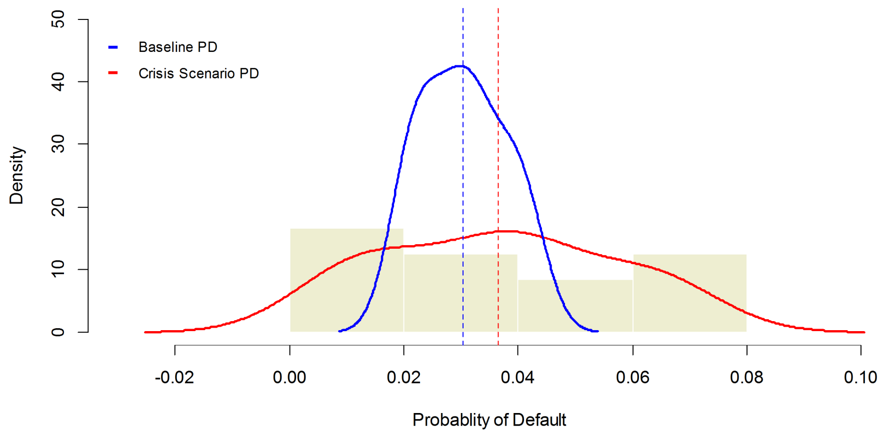
(b)

Figure 16: Comparison of *PD* forecasted values obtained through the 85% threshold GBN for both *normal* and *crisis* market conditions: (a) time series plot and forecasted paths; (b) histogram bars correspond to simulated crisis predictions, while kernels approximate the shape of the relative distributions.

To better appreciate differences between *PD*'s forecasts, Table 11 reports the summary statistics of such predictions for either the baseline and the crisis scenario data: standard deviation measures for crisis-related predictions are of higher magnitude than their baseline counterparts; moreover, *PD*'s level is believed to be, on average, higher during an economic downturn. Indeed, in such a scenario, *PD*'s predicted values are estimated to potentially reach values up to almost 7%.



(a)



(b)

Figure 17: Comparison of *PD* forecasted values obtained through the expert network for both *normal* and *crisis* market conditions: (a) time series plot and forecasted paths; (b) histogram bars correspond to simulated crisis predictions, while kernels approximate the shape of the relative distributions.

GBN	Scenario	Mean	Standard Deviation	Skewness	Kurtosis
Expert	Baseline	0.0304	0.0073	0.1628	1.7733
	Crisis	0.0364	0.0203	0.0941	1.7939
85% threshold	Baseline	0.0298	0.0061	0.1787	1.7796
	Crisis	0.0313	0.0168	0.2293	1.8026

Table 11: Summary statistics for learnt and expert GBNs relative to baseline's and crisis scenario's *PD* estimates.

To better compare the crisis scenario forecasts obtained through the 85% threshold and the expert GBNs, the relative predicted paths had been plotted in the same graph, as depicted in Figure 18.

In addition to what said so far, it may be interesting to notice that the expert network tends to return higher *PD* prediction values in comparison with the ones forecasted by the learnt network. In this sense, despite the modest discrepancies between the two, it may be possible to assert that the former provides more conservative estimates/forecasts than the latter.

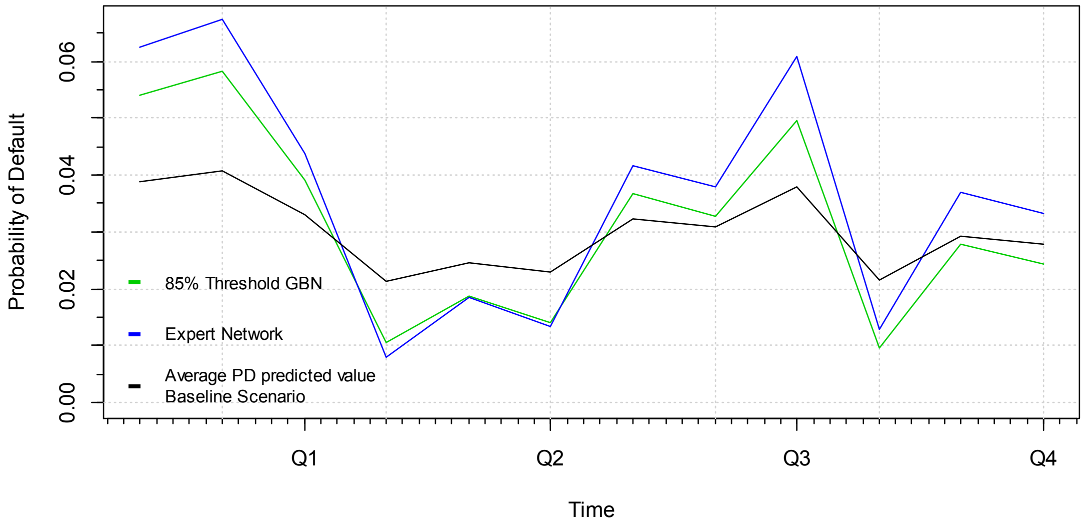


Figure 18: Comparison of predicted values, relative to the crisis scenario analysis, obtained via the 85% threshold GBN and the expert network. The mean *PD* predicted baseline path was created by averaging the baseline forecasts of the two networks.

As conducted before relative to the baseline scenario forecasts, it would be interesting to provide an assessment concerning firm-specific variables for the crisis scenario at hand, in order to identify the most critical firm-specific indicators that may influence *The Bank*'s *PD* level.

In particular, results would be provided only relative to those firm-specific elements that are causally related to macroeconomic variables, as defined by the 85% threshold and the expert GBNs. Indeed, such elements would be the ones affected by the changed market conditions set by the simulated crisis scenario.

Consider first the *NPL* variable, which graphical output is displayed in Figure 19. The crisis-related trend predicted by the 85% threshold GBN resembles the case in which market conditions are normal, with the only difference being that variability of forecasts is comparatively increased and the mean is slightly shifted to the right, with reference to Figure 19(c). Similar considerations are valid for the crisis-related predictions obtained via the expert network, however, notice: in comparison with the learnt network, variability in forecasts is somewhat higher and the mean (Figure 19(d)) is farther apart from the average value, referred to the

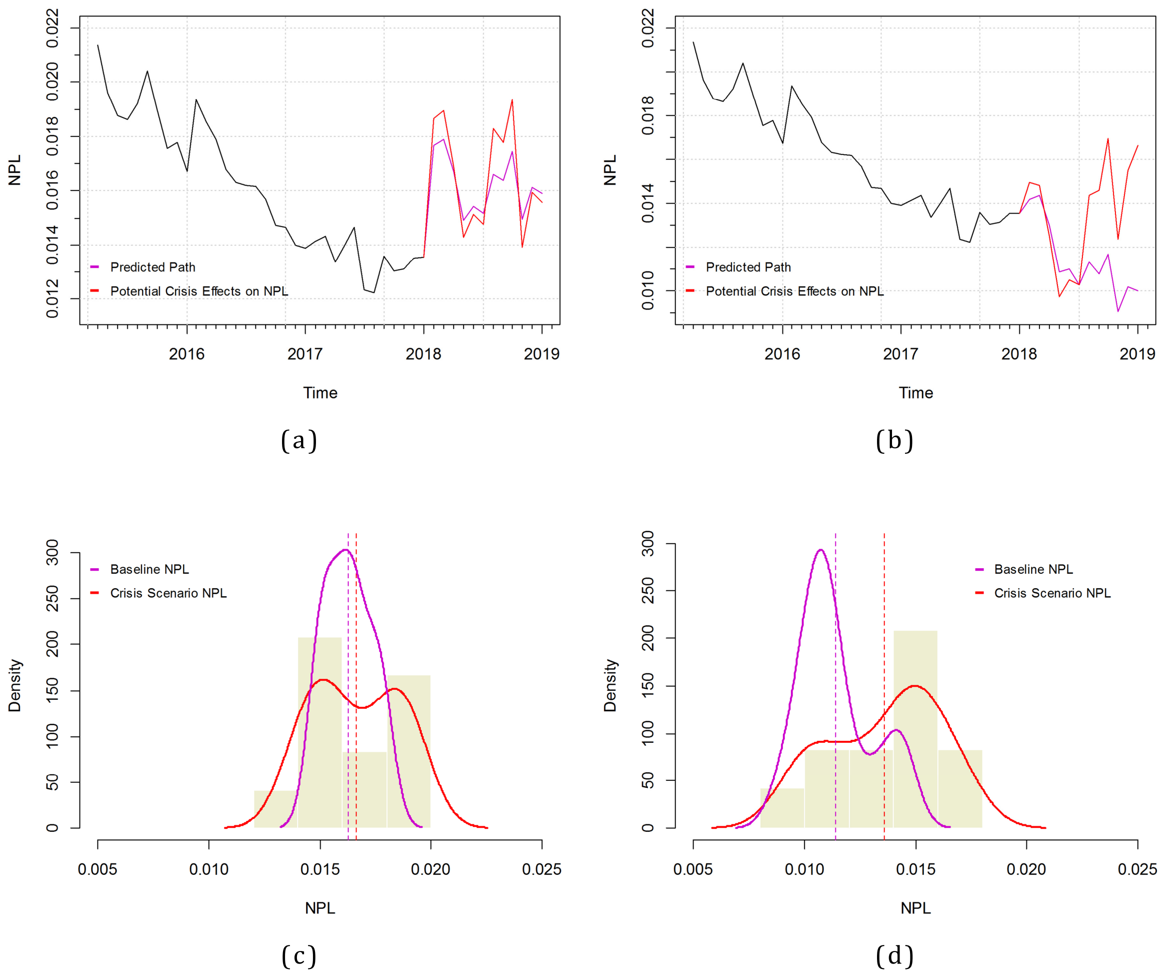


Figure 19: Crisis scenario predicted values for *NPL*, relative standard error boundaries and prediction intervals based on the use of (a) the 85% threshold GBN and (b) the expert network. Crisis scenario simulated forecasted distributions for *NPL* generated by utilizing (c) the 85% threshold GBN and (d) the expert network; dash-dotted vertical lines mark critical thresholds and values relative to *The Bank*.

case in which normal market conditions are assumed. Moreover, Figure 19(b) clearly shows that the expert network identifies an increasing trend in presence of crisis-like conditions, which is opposite to what it suggests for the normal market scenario. This is interesting because such information could be critical in assessing *The Bank*'s *PD* value, given *NPL* is deemed to be the most influencing variable on the credit institution's probability of default. Furthermore, one may infer that the previously forecasted upward shift in *PD*'s average value (Figure 18) could be partly

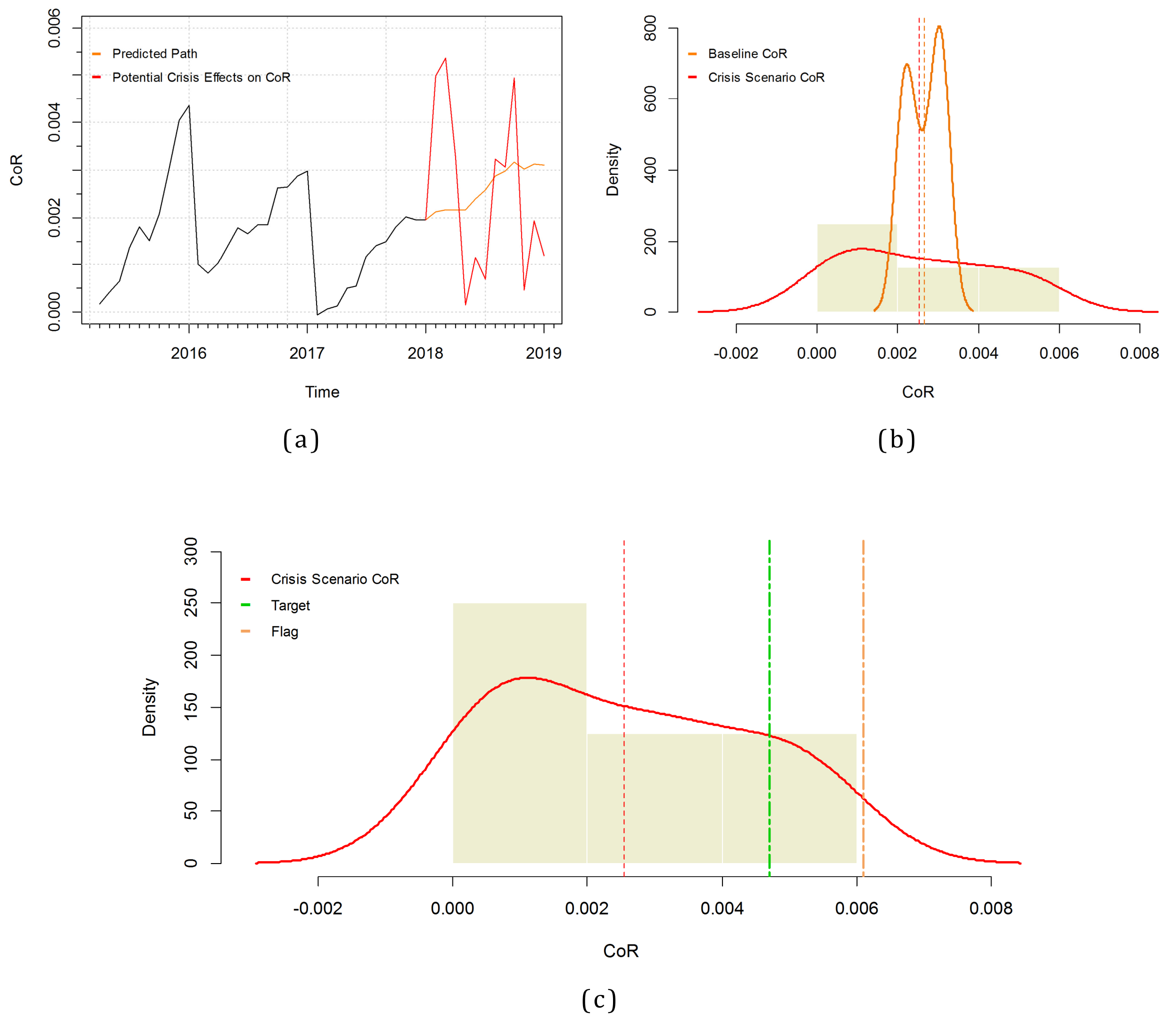


Figure 20: Simulated values for *CoR* and generated via the expert network; dashed vertical lines indicate the relative mean values. (a) Crisis scenario predicted values, relative standard error boundaries and prediction intervals. (b) Comparison of the crisis and normal scenarios simulated forecasted distributions. (c) Crisis scenario simulated forecasted distribution and *The Bank*'s critical thresholds marked by dash-dotted vertical lines.

due to the increase in *NPL* level predicted by the expert network. Lastly, despite the shifts *NPL* would be subject to in a downturn period, notice that the *flag* level (set at around 0.025 by *The Bank*) is unlikely to be reached even during and economic crisis. This may be explained by the fact that the last value observed for *NPL* was rather low in magnitude, hence, setting a good margin which may serve as a buffer against potential adverse conditions relative to the prediction period.

Turning now to analyzing the firm-specific variables which are present in only either of the learnt or the expert network, consider first the variables *CoR* and *CET*1 which are contemplated in the latter network.

Figure 20 relates to *CoR* and, as one may notice, its distribution's average values are rather close to each other; instead, variability in the crisis scenario forecasts is quite evident compared to the other instance, in which predicted values are rather concentrated around the mean value. Also, Figure 20(a) suggests that no particular trend is expected to be present in a market downturn situation. Lastly, as Figure 20(c) displays, *CoR* could potentially reach the flag level set by *The Bank*. Then, the credit institution should monitor *CoR*'s values closely in correspondence of events similar to the simulated ones.

Concerning the *CET*1 variable and its forecasted values for the crisis-like scenario, there appears to be no predicted trend in forecasts (Figure 21(a)) and mean values of the distributions seem to be somewhat similar (Figure 21(b)). Most importantly, though, as Figure 21(c) depicts, the simulated downturn conditions may cause the variability in *CET*1 to increase significantly and this could erode *The Bank*'s capital buffers up to the point in which the credit institution may need to take corrective actions: it may necessitate to put in place measures in order to increase the capital base, for instance by searching further possible funding solutions. Then, such simulated scenario is in line with the expectation that a crisis is likely to generate difficulties for many businesses and credit institutions. Consequently, even the *CET*1 variable may be regarded as a contributor to the forecasted increased level in *The Bank*'s *PD*.

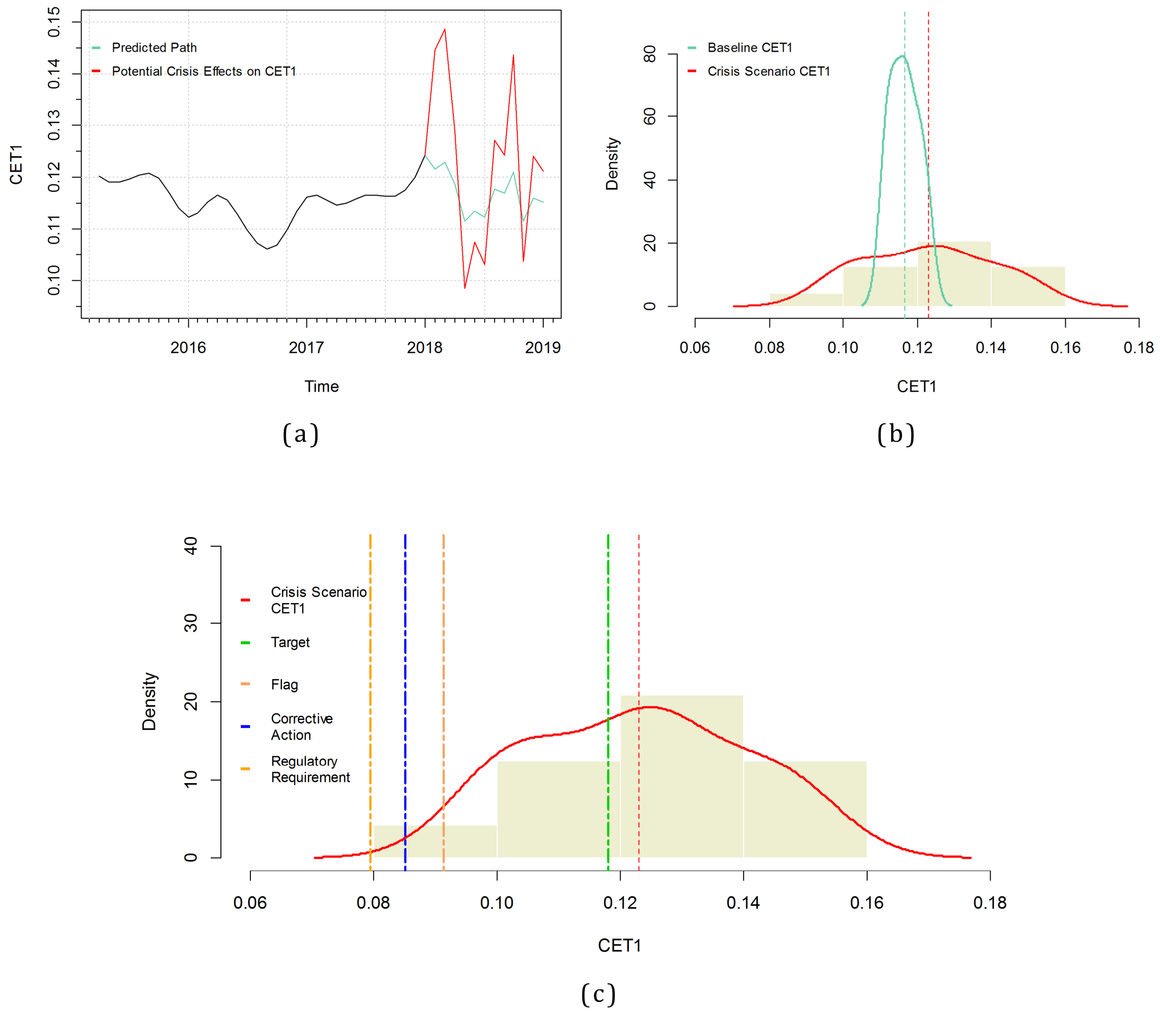
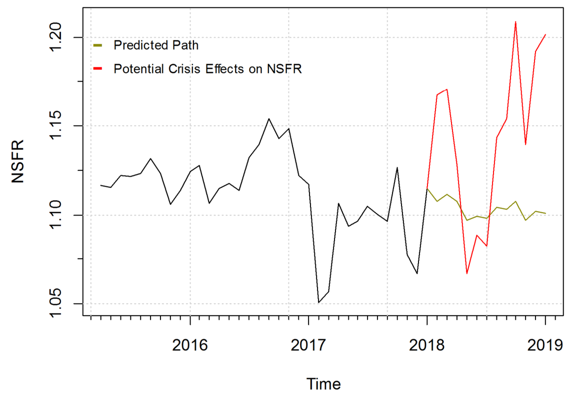


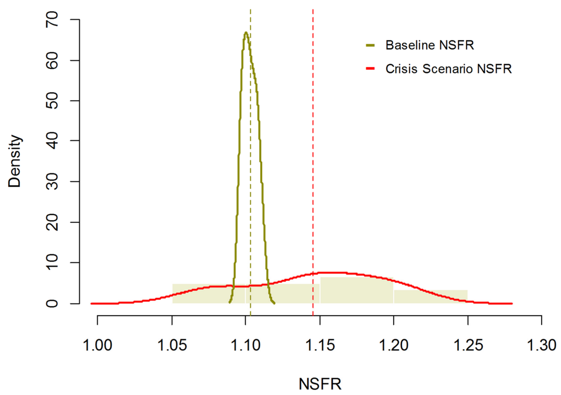
Figure 21: Simulated values for *CET1* and generated via the expert network; dashed vertical lines indicate the relative mean values. (a) Crisis scenario predicted values, relative standard error boundaries and prediction intervals. (b) Comparison of the crisis and normal scenarios simulated forecasted distributions. (c) Crisis scenario simulated forecasted distribution and *The Bank*'s critical thresholds marked by dash-dotted vertical lines.

Concerning the 85% threshold GBN, the focus would be on the *NSFR* variable, which graphical analysis is reported in Figure 22. As also noted in other cases, *NSFR* suffers from an increased variability in forecasts when severe adverse market changes do take place. For this particular instance, predictions suggest the presence of an upward trend in *NSFR* forecasted values. Such behavior may be plausible in the event that, after a reduction in sources of funding due to immediate crisis' adverse effects, *The Bank* succeeds in finding alternative funding solutions. However, this really is an optimistic view of countering the effects of adverse economic conditions. Indeed, due to the fact that market turmoil often negatively affects several businesses and other credit institutions, what would probably happen in reality is that sources of funding for *The Bank* would be very limited in the short term. This interpretation is in line with the output shown in Figure 22(c), which puts in evidence that the increased uncertainty of a crisis scenario would likely require *The Bank* to take corrective actions to ensure the survival of the credit institution.

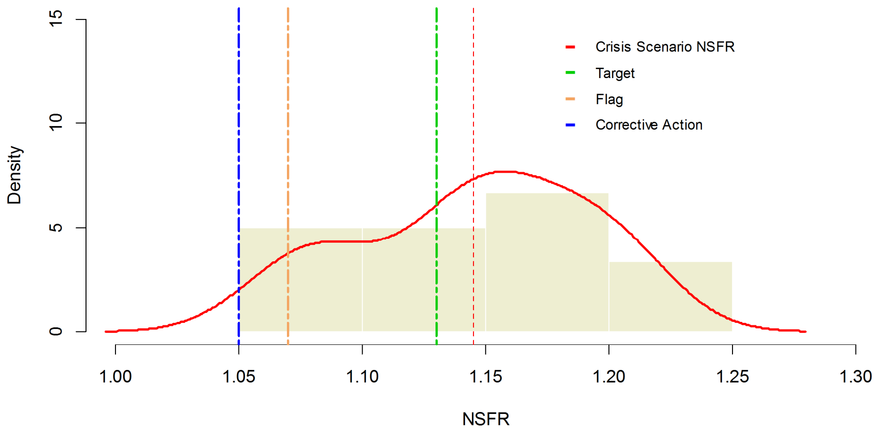
Notice that the interpretation provided on *NSFR* may actually be in line



(a)



(b)



(c)

Figure 22: Simulated values for *NSFR* and generated via the 85% threshold GBN; dashed vertical lines indicate the relative mean values. (a) Crisis scenario predicted values, relative standard error boundaries and prediction intervals. (b) Comparison of the crisis and normal scenarios simulated forecasted distributions. (c) Crisis scenario simulated forecasted distribution and *The Bank*'s critical thresholds marked by dash-dotted vertical lines.

with some aspects of the *CET1* variable previously analyzed: difficulties in maintaining *The Bank*'s capital buffers, in a crisis-like situation, may be worsened by contemporaneous reduced funding opportunities.

In general, one may infer that, despite elements/variables be included in different GBNs (either learnt or expert ones), sometimes it could be helpful to combine different instruments/tools and blend their outputs in order to obtain a synergic effect, to possibly end up with a more informed opinion relative to the specific matter object of the analysis.

GBN	Variable	Scenario	Mean	Standard Deviation	Skewness	Kurtosis
Expert	NPL	Baseline	0.0114	0.0017	0.6994	2.4218
		Crisis	0.0136	0.0024	-0.3107	1.7846
	CoR	Baseline	0.0026	0.0004	-0.1201	1.2824
		Crisis	0.0025	0.0019	0.2602	1.6450
	CET1	Baseline	0.1165	0.0039	0.1783	1.7818
		Crisis	0.1229	0.0171	0.0619	1.7963
85% threshold	NPL	Baseline	0.0163	0.0010	0.1803	1.7804
		Crisis	0.0166	0.0019	0.0375	1.5044
	NSFR	Baseline	1.1030	0.0048	0.2651	1.8732
		Crisis	1.1450	0.0469	-0.3247	1.9552

Table 12: Summary statistics for learnt and expert GBNs relative to baseline's and crisis scenario's firm-specific estimates.

Lastly, firm-specific variables' summary statistics are reported in Table 12, which allows for comparisons between the baseline and the economic downturn scenarios. As a general consideration for such Stress Testing analysis, in line with expectations, it is evident that adverse market movements generate uncertainty, which in turn affects variability in performance indicators and, ultimately, tends to raise either the magnitude and the variability relative to *The Bank*'s *PD*.

Therefore, what emerges from the performed *Scenario Analysis* is that a rise in the level of *PD* in correspondence of crisis-like scenarios is to be expected. In any case, the important concept to remember is that, under any market scenario, it is crucial to try to predict and monitor the movements in key performance indicators. This may allow an appropriate assessment on the credit institution’s economic situation, which in turn may provide to management the opportunity to anticipate adverse trends and implement adequate timely corrective actions, when necessary.

Conclusions and Closing Remarks

Before any closing remark and consideration, some key takeaways are to be mentioned relative to the case study analysis' results.

First, the generated output suggests that *The Bank* is expected to be relatively stable and healthy, as well as compliant with regulatory requirements (relative to capital buffers), either under baseline or crisis-like scenarios for the prediction period. Hence, on the basis of the results and scope of the Stress Testing analysis, the credit institution may choose to focus its efforts on the conduction of risk monitoring activities, resource allocation planning and business as usual tasks. Indeed, no corrective actions are expected to be needed in the short-run.

Second, concerning the specific outcomes of the learnt and expert networks, the most noticeable discrepancies between the two have been observed relative to the *NPL* variable, for any of the two scenarios considered. As already-mentioned, such factor is deemed to be the most relevant in terms of dependence relations toward *The Bank*'s *PD* and may actually be the major reason for the difference in *PD*'s estimates, however modest.

Third, notice that the specific results obtained via the performed Stress Testing analysis are heavily dependent on the type and availability of input data. In fact, such inputs contribute to determining both the structure and parameters of networks, on condition that learning procedures are applied. In this sense, ensuring to dispose of sufficient and reliable data is always fundamental when performing these types of analyses.

Overall, the focus of this work has been to build a statistical Bayesian Network approach to support the conduction of Financial Stress Testing tasks for credit institutions. In particular, compared to other modeling methodologies, BNs turn out to be rather flexible instruments, which may be utilized on an ongoing basis. Indeed, once set up in a sufficiently adaptable way, a credit institution would only need to input data as they become available to subsequently dispose of updated stats and prediction outputs on the relative indicators of interest. Moreover, thanks to the learning features of BNs, if utilized continuously over time, their estimated structures may adapt to reflect changes concerning the underlying features and dependence relations jointly characterizing data. Therefore, such tool may provide opportunities to grasp non-trivial information hidden in between variables.

Most often, on the practical side, BNs could be utilized to identify potential weaknesses in business and, in this sense, they may be included among the array of risk management tools. In fact, practices like Financial Stress Testing analysis serve as ways to control and monitor risks, with the aim of implementing actions to ensure business solidity as well as continuous improvement for the firm.

Anyway, aside from the specific context of financial institutions, remember that BNs do have a wider application relative to other fields of knowledge, as noted in an earlier section. In most instances, they are deemed to perform well in case one disposes of large datasets of numerous variables, which could potentially be intertwined by several different relations.

In general, it is usually the case that utilizing BNs implies possessing knowledge of theoretical notions relative to the model and its properties, as well as an understanding of the specific context under study. Used in isolation, they would likely only provide unstructured information or outputs of difficult interpretation. Therefore, one would always accept a certain degree of subjectivity or interpretation when carrying out data analysis activities, especially when trying to produce predictive estimates. The purpose is not to limit BNs data exploration capabilities, on the contrary, uncommon or non-trivial insights extracted by such instruments from data would be welcomed, and posed to more thorough analysis if needed. However, it is important to always define the perimeter over which the underlying algorithms may work, in order to avoid nonsense outputs a priori. Moreover, a critical aspect for a correct interpretation and understanding of the output produced by BNs would be to always count on critical thinking when analyzing results.

In light of what asserted so far, as the Financial Stress Testing output demonstrated, both expert networks and estimated networks are deemed to be useful and should be utilized jointly to support analytical activities, provided that they may actually work in synergy by way of complementarities, if applicable.

In conclusion, Bayesian Networks may be considered as a valid tool among the array of risk management instruments that credit institutions could employ when performing Financial Stress Testing analyses. Moreover, to summarize the general idea behind the use of Bayesian Networks, one may wish to keep in mind a rather simple, still valuable concept:

> *What is truly interesting is not how well the output produced by analytical tools matches expectations, but rather how discrepancies between expert knowledge and the estimated results may allow for opportunities to gain further insights and knowledge.*

Bibliography

Basel Committee (2011), "Basel III: A global regulatory framework for more resilient banks and banking systems", Bank for International Settlements, 2-46.

Bologna Pierluigi and Anatoli Segura (2016), "Integrating stress tests within the Basel III capital framework: a macroprudentially coherent approach", Bank of Italy, Questioni di Economia e Finanza, Occasional Paper No. 360, 5.

Cobb Barry R., Rafael Rumí and Antonio Salmerón (2007), "Bayesian Network Models with Discrete and Continuous Variables", Studies in Fuzziness and Soft Computing.

Darwiche Adnan (2010), "What are Bayesian networks and why are their applications growing across all felds?", Communications of the ACM; Volume 53, Issue 12, 80-82.

Dees Stéphane, Jérôme Henry and Reiner Martin (2017), "Stress-Test Analytics for Macroprudential Purposes in the Euro Area", European Central Bank, EU catalogue No. QB-02-17-041-EN-N, 5-47.

European Banking Authority (2018), "2018 EU-Wide Stress Test - Methodological Note".

Fenton Norman and Martin Neil (2012), "Risk Assessment and Decision Analysis with Bayesian Networks", 113-119, CRC Press.

Gao Gelin, Bud Mishra and Daniele Ramazzotti (2017), "Efficient Simulation of Financial Stress Testing Scenarios with Suppes-Bayes Causal Networks", Procedia Computer Science, Volume 108, 272-275.

Gao Gelin, Bud Mishra and Daniele Ramazzotti (2018), "Causal Data Science for Financial Stress Testing", Journal of Computational Science, Volume 26.

Gómez Pilar, Purificación Parrado and Antonio Partal (2018), "Risk Profile Indicators and Spanish Banks' Probability of Default from a Regulatory Approach", Sustainability, Volume 10, Issue 4, 3-6.

Hurlin Christophe, Jérémy Leymarie and Antoine Patin (2018), "Loss functions for LGD model comparison", Le Centre pour la Communication Scientifique Directe - HAL, 6-36.

Jacobs Jr. Michael and Sensenbrenner Frank J. (2018), "A comparison of methodologies in the stress testing of credit risk – alternative scenario and dependency constructs", Quantitative Finance and Economics, 2 (2), 294-324.

Kipkoech Rotich Titus, George Orwa and Joseph Mung'atu (2013), "Granger Causality and Error Correction Models in Economics: A Case study of Kenyan Market", Mathematical Theory and Modeling, Volume 3, No. 14, 101-102.

Konečný Tomáš, Seidler Jakub, Belyaeva Aelita, Belyaev Konstantin (2017), "The time dimension of the links between loss given default and the macroeconomy", ECB Working Paper 2037.

Mihajlovic Vojkan and Petkovic Milan (2001), "Dynamic Bayesian Networks: A State of the Art", University of Twente, CTIT Technical Report Series, vol. 01-34.

Nagarajan Radhakrishnan, Marco Scutari and Sophie Lebre (2013), "Bayesian Networks in R with Applications in Systems Biology", 2-47, Springer.

Neapolitan Richard E. (2004), "Learning Bayesian Networks", Prentice Hall series in Artificial Intelligence, Pearson Prentice Hall.

Schuurman N. K., Grasman R. P. P. P. and Hamaker E. L. (2016), "A Comparison of Inverse-Wishart Prior Specifications for Covariance Matrices in Multilevel Autoregressive Models", Multivariate Behavioral Research, 51:2-3, 185-206.

Rebonato Riccardo (2017), "Stress Testing with Bayesian Nets and Related Techniques: Meeting the Engineering Challenges", EDHEC Risk Institute, 1-8.

Scutari Marco and Jean-Baptiste Denis (2015), "Bayesian Networks with Examples in R", 2-112, CRC Press.

Stephenson Todd Andrew (2000), "An Introduction to Bayesian Network Theory and Usage", IDIAP Publications.

Venables William N. and Smith David M. (2018), "An Introduction to R - Notes on R: A Programming Environment for Data Analysis and Graphics", R Core Team, Version 3.5.2.

Verbeek Marno (2005), "A Guide to Modern Econometrics", 313-319, 2nd Edition, John Wiley & Sons.

Zou Cunlu and Jianfeng Feng (2009), "Granger causality vs. dynamic Bayesian network inference: a comparative study", BioMed Central Ltd., BMC Bioinformatics.

APPENDIX A: Overview on Granger Causality

This section will be dedicated to an overview of the most important notions relating to *Granger Causality*, which aim is to identify the causal inter-dependences between the observed variables. The application of such methodology, in its original form, is based on the following simple assertion:

By letting X_t and Y_t be two time series, X_t is said to Granger-cause Y_t if, by exploiting the lagged elements of both X_t and Y_t, the prediction on Y_t would be improved compared to the case in which one may utilize Y_t past elements only.

More formally, applying the Granger Causality principle on two time series, firstly involves the definition of Y_t as a *Vector Auto-Regressive* (VAR) *process*:

$$\alpha(L)y_t = \varepsilon_t \tag{A.1}$$

where $\alpha(L) = (1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_p L^p)$, L^p indicates the lag order p of Y_t and ε_t is a White Noise process $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

Secondly, the inclusion of X_t series would imply a *VAR(p,q)* model with the addition of exogenous lagged variables such that:

$$\alpha(L)\, y_t = \beta(L)\, x_t + \varepsilon_t \tag{A.2}$$

where $\beta(L) = (\beta_0 + \beta_1 L + \beta_2 L^2 + \cdots + \beta_q L^q)$, L^q indicates the lag order q of X_t and ε_t is a White Noise, as before.

Notice that, in accordance with the definition provided above, the lagged variables of X_t would be kept if and only if they provide meaningful explanatory power to the prediction of the response variable[18] y_t in

[18] To identify which lagged variables are to keep in the model, it may be possible to perform *t-tests* for single parameter significance and *F-tests* for joint significance (Kipkoech, Orwa and Mung'atu, 2013).

equation (A.2). Therefore, in case X_t variables are to be included in the VAR process, X_t is said to Granger-cause Y_t.

When assessing the presence (or absence) of Granger Causality relationships, as a prerequisite, it is important to ensure that the considered time series X_t and Y_t actually are stationary (Kipkoech, Orwa and Mung'atu, 2013). To this end, notice that an important feature of VAR processes is stability, expressed via the use of stationary time series.

In case processes turn out being non-stationary, procedures like differencing or correcting for seasonality may be applied in order to recover stationarity. However, such measures may actually imply the loss of some information that would otherwise be conveyed via the original form of the process. Hence, an alternative that may allow the avoidance of time series transformation is the application of *Cointegration* (where appropriate): in case two non-stationary processes have the same trend in common it would still be possible to get OLS estimates of coefficients by regressing Y_t on X_t (Verbeek, 2005).

To be clear, for instance, consider the case in which both Y_t and X_t are non-stationary processes first-order integrated ($I(1)$) and either process is extended up until the first lag, reflected in a $VAR(1,1)$ model as follows:

$$\alpha(L)\, y_t + \beta(L)\, x_t = \varepsilon_t \tag{A.3}$$

where $\alpha(L) = (1 - \alpha_1 L)$, $\beta(L) = (\beta_0 - \beta_1 L)$ and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$; notice that formula (A.3) is a simple rearrangement of equation (A.2). Then, in case ε_t results being a stationary process, it is evident from equation (A.3) that also Y_t and X_t together would be stationary. Hence, regardless of Y_t and X_t being $I(1)$ processes, the cointegration condition would be verified and procedures to assess Granger Causality may be performed.

In general, one may follow the so called *Engle-Granger two-step procedure* in order to verify the presence of cointegration between time series:

- In the first place, one may wish to assess the degree of integration relative to each time series (which in the majority of circumstances may either be $I(0)$, $I(1)$ or $I(2)$) via statistical methodologies like the Augmented Dickey Fuller (ADF) test. This, in turn, would define the model specification on which the OLS estimation procedure would be performed;
- Secondly, given the model specification in the previous step, it may be possible to obtain the relative OLS estimates. Thereafter, it would be necessary to verify that model's residuals ($\hat{\varepsilon}_t$) actually are stationary; this phase may be performed once again via an ADF test.

As previously stated, cointegration would be verified when Y_t and X_t share a common trend: in order to reach this "equilibrium", the variables of each time series would converge to such condition at a certain rate, which behavior is modeled by the so called *Error Correction Model* (ECM) (Kipkoech, Orwa and Mung'atu, 2013).

With an illustrative purpose, to grasp the meaning of the ECM, it is sufficient to briefly display its formula relative to a $VAR(1{,}1)$ model (equivalent to the one in formula (A.3)) in its extended version. The VAR process would be

$$y_t - \alpha_1 y_{t-1} = \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t. \tag{A.4}$$

While its relative ECM form would correspond to

$$\begin{aligned} y_t - y_{t-1} &= \beta_0(x_t + x_{t-1}) + (\alpha - 1)[y_{t-1} - kx_{t-1}] + \varepsilon_t \\ &= \beta_0(x_t + x_{t-1}) + (\alpha - 1)u_{t-1} + \varepsilon_t \end{aligned} \tag{A.5}$$

where ε_t would still be a White Noise process, $(\alpha - 1)$ is the correction coefficient of the error u_{t-1} and $k = (\beta_0 + \beta_1)/(1 - \alpha)$ is the so called longrun coefficient, which *de facto* expresses the effect on y_t of a one-unit variation in x_t.

To conclude this brief review on the core concepts related to Granger Causality (tailored to the topics of the present paper), notice that such methodology is usually deemed adequate when a significant number of observations is available for the time series considered in the analysis. Moreover, in case data length at disposal is deemed to be short, one may assert that Bayesian Networks would be a more adequate tool to be utilized in such circumstance (Zou and Feng, 2009). Lastly, keep in mind that the notions presented above only relate to pairwise Granger-causal relations, hence applying uniquely to bivariate time series[19].

[19] Notice that extensions to the *classic* Granger Causality methodology presented in this section do exist. For instance, (Zou and Feng, 2009) presented an approach to address the issue of assessing causality in multivariate contexts via the *Conditional Granger Causality* method. Still, tools like Bayesian Networks (and others) could be more adequate in assessing causal and dependence relationships in multivariate environments.

APPENDIX B: *R* Software Environment

The present section's purposes are, first, to provide a brief definition of the *R* Software Environment and, second, to acknowledge the authors of packages utilized for the development of this work.

R is to be intended as an integrated and structured software environment grounded on the underlying *S* programming language, first created and developed at *Bell Laboratories* by Rick Becker, John Chambers and Allan Wilks. Such software suite is primarily meant for data handling and data analysis activities. Indeed, R may perform data manipulation, computational tasks and display graphical outputs, features which are often of use for statistical applications and analyses purposes.

For a more thorough description on *R* and relative introductory notions, see Venables and Smith (2018).

It must be mentioned that *R* software is extended by several *packages*, which development is to be acknowledge either to the *R* Core Team or to users and contributors.

Following, a list of the packages utilized to elaborate and produce the outputs of the present work, together with relative brief descriptions and references:

bnlearn Provides functions and algorithms to perform Bayesian network structure and parameter learning, as well as parameter inference.

Reference: Scutari Marco (2018), "Bayesian network structure learning, parameter learning and inference", Version: 4.4

fakeR Allows the creation of simulated data based on random draws from a multivariate normal distribution. It also applies on the datasets of time series' observations.

Reference: Zhang Lily and Tingley Dustin (2016), "Simulates Data from a Data Frame of Different Variable Types", Version: 1.0

forecast Provides instruments to perform analyses on univariate time series predictions, including the use of automatic functions relative to ARIMA models.

Reference: Hyndman Rob, O'Hara-Wild Mitchell, Bergmeir Christoph, Razbash Slava and Wang Earo (2017), "Forecasting Functions for Time Series and Linear Models", Version: 8.2

gmodels Includes instruments related to model fitting tasks.

Reference: Warnes Gregory R., Bolker Ben, Lumley Thomas and Johnson Randall C. (2018), "Various R Programming Tools for Model Fitting", SAIC-Frederick Inc., Version: 2.18.1

graph Dedicated to the handling of data structures for the provision of a graphical output.

Reference: Gentleman Robert, Whalen Elizabeth, Huber Wolfgang and Falcon Seth (2016), "graph: A package to handle graph data structures", Version: 1.52.0

Hmisc Mainly contains several functions relative to data manipulation and handling, data analysis and plotting features. Notice that it depends on the package *lattice*.

Reference: Harrell Jr. Frank E. (2018), "Harrell Miscellaneous", Version: 4.1-1

igraph Dedicated to the creation of graphical outputs, in particular, relative to network analysis topics.

Reference: Csardi Gabor, Nepusz Tamas and *Authors** (2018), "Network Analysis and Visualization", Version: 1.2.1
*https://cran.r-project.org/web/packages/igraph/AUTHORS

lattice It allows the creation of graphical outputs, with particular focus on multivariate data.

Referenece: Sarkar Deepayan (2017), "Trellis Graphics for R", Version: 0.20-35

lmtest Includes a series of statistical test functions for diagnostics and inference on linear regression models.

Reference: Hothorn Torsten, Zeileis Achim, Farebrother Richard W. and Cummins Clint (2018), "Testing Linear Regression Models", Version: 0.9-36

moments Includes functionalities to compute and execute tests concerning distributions' moments.

Reference: Komsta Lukasz and Novomestky Frederick (2015), "Moments, cumulants, skewness, kurtosis and related tests", Version: 0.14

pracma Provides specific functions for numerical analysis, especially focused on optimization and time series matters.

Reference: Borchers Hans Werner (2017), "Practical Numerical Math Routines", Version: 2.1.4

Rgraphviz Connects *R* with *AT* and *T graphviz* packages in order to produce graphical outputs via the specific *graph* library.

Reference: Hansen Kasper Daniel, Gentry Jeff, Long Li, Gentleman Robert, Falcon Seth, Hahne Florian and Sarkar Deepayan (2016), "Provides plotting capabilities for R graph", Version: 2.18.0

stringr Provides functions to manipulate strings. It is usually exploited for data cleaning and preparation purposes.

Reference: Wickham Hadley (2018), "Simple, Consistent Wrappers for Common String Operations", Version: 1.3.1

tseries Includes statistical tests and functionalities mainly tailored to computational finance topics.

Reference: Trapletti Adrian, Hornik Kurt and LeBaron Blake (2018), "Time Series Analysis and Computational Finance", Version: 0.10-44

vars Dedicated to diagnostics tests, predictive and causal analyses for VAR type models.

Reference: Pfaff Bernhard (2018), "VAR Modelling", Version: 1.5-3

xlsx Among its functionalities, it allows to either read or write *Microsoft Excel* files, as well as manage other types of interactions between the Excel suite and *R*.

Reference: Dragulescu Adrian A. (2014), "Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files", Apache POI project for Microsoft Excel format, Version: 0.5.7

xts Function which allows to generate *extensible time-series* objects from an input dataset. Notice that such package is linked to the *zoo* library.

Reference: Ryan Jeffrey A. and Ulrich Joshua M. (2018), "Create Or Test For An xts Time-Series Object", Version: 0.11-2